Mathematische Zeitschrift

Band 187 · 1984

Springer-Verlag
Berlin · Heidelberg · New York · Tokyo

"**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. He was succeeded in 1973 by *H. Salzmann* and 1977 by *H. Heyer*. As of 1984 *W.P. Barth* and *W. v. Wahl* are the managing editors.

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. 1973 ging die Federführung an *H. Salzmann* und 1977 an *H. Heyer* über. Seit 1984 sind *W.P. Barth* und *W. v. Wahl* geschäftsführende Herausgeber.

Manuscripts may be sent to / Manuskripte nehmen entgegen:

Springer-Verlag/Berlin Heidelberg New York
Printers: Universitätsdruckerei H. Stürtz AG, Würzburg

Inhalt/Contents

Covered by Zentralblatt für Mathematik and
Current Mathematical Publications

Mathematische Zeitschrift

Band 187 Heft 1 1984

Springer-Verlag
Berlin Heidelberg New York Tokyo

209 Math. Z. ISSN 0025-5874 MAZEAX 187 (1) 1-144 (1984)

August 1984

Mathematische Zeitschrift

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. 1973 ging die Federführung an *H. Salzmann* und 1977 an *H. Heyer* über. Seit 1984 sind *W.P. Barth* und *W. v. Wahl* geschäftsführende Herausgeber.

Die **„Mathematische Zeitschrift"** dient der Pflege der reinen und angewandten 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 100 Sonderdrucke kostenlos zur Verfügung gestellt, weitere können zum Selbstkostenpreis bezogen werden.

Hinweis für Autoren

Die Autoren werden gebeten, ihre Manuskripte gemäß den „Instructions to Authors" für den Satz vorzubereiten. Ungenügend vorbereitete Manuskripte müßten von der Redaktion zurückgesandt werden, was die Publikation verzögern würde.

"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. He was succeeded in 1973 by *H. Salzmann* and 1977 by *H. Heyer*. As of 1984 *W. P. Barth* and *W. v. Wahl* are the managing editors.

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 will not be published.

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

Notice to Authors

Authors are asked to prepare their manuscripts according to the "Instructions to Authors". Otherwise the manuscripts may have to be returned to the authors, which will cause a delay in publication.

Math. Z. 187, 1–11 (1984)

Mathematische Zeitschrift

On Differential Polynomials and Results of Hayman and Doeringer

J.K. Langley

Department of Mathematics, University of Illinois, 1409 West Green St., Urbana, IL 61801, USA

1. Introduction

Let $f(z)$ be non-constant and meromorphic in the plane, with Nevanlinna characteristic $T(r,f)$, and let $a_0(z), \ldots, a_k(z)$ be meromorphic in the plane and satisfy

$$T(r, a_i) = S(r, f) \tag{1.1}$$

for each $i=0, \ldots, k$, where, using standard notation from [3], we denote by $S(r,f)$ any quantity such that $S(r,f) = o(T(r,f))$ nearly everywhere (n.e.) as $r \to \infty$, i.e. as $r \to \infty$ outside a set of finite linear measure. Then if $k \geqq 1$ and

$$\psi(z) = \sum_{i=0}^{k} a_i(z) f^{(i)}(z) \tag{1.2}$$

is non-constant, the following result ([3], p. 57) gives an upper bound for the growth of $T(r,f)$ in terms of the number of zeros and poles of $f(z)$, and 1-points of $\psi(z)$:

Theorem A. *Suppose that $f(z)$ is non-constant and meromorphic in the plane, and that $\psi(z)$ is given by* (1.2) *and* (1.1) *and is non-constant. Then*

$$T(r,f) < \bar{N}(r,f) + N\left(r, \frac{1}{f}\right) + \bar{N}\left(r, \frac{1}{\psi - 1}\right) - N_0\left(r, \frac{1}{\psi'}\right) + S(r,f) \tag{1.3}$$

where $N_0\left(r, \frac{1}{\psi'}\right)$ counts only zeros of ψ' which are not zeros of $\psi - 1$, and the $\bar{N}$ counting functions count points without regard to multiplicity.

Hayman [2] showed that, if $f(z)$ is transcendental, and $\psi(z) = f^{(k)}(z)$ for some $k \geqq 1$, then the term $\bar{N}(r,f)$ may be eliminated from (1.3):

Theorem B. *If $f(z)$ is transcendental and meromorphic in the plane, then*

$$T(r,f) < \left(2 + \frac{1}{k}\right) N\left(r, \frac{1}{f}\right) + \left(2 + \frac{2}{k}\right) \bar{N}\left(r, \frac{1}{f^{(k)} - 1}\right) + S(r,f).$$

We use a modified version of Hayman's proof to extend this result by showing that $\bar{N}(r,f)$ may be eliminated in (1.3) unless $\psi(z)$ satisfies a certain differential equation; whether or not this exceptional case is possible depends on the nature of the poles, if any, of $a_{k-1}(z)/a_k(z)$. In the case where $a_0,\dots,a_k$ are constants we are able to completely classify the possibilities. We have, in fact:

Theorem 1. *Suppose that $f(z)$ is transcendental and meromorphic in the plane, and that $\psi(z)$ is non-constant and is given by* (1.2) *and* (1.1), *with $a_k(z)$ not identically zero. Then for any non-zero complex number b:*

(i) *either*

$$T(r,f) < 3N\left(r,\frac{1}{f}\right) + 4\bar{N}\left(r,\frac{1}{\psi-b}\right) + S(r,f) \tag{1.4}$$

or $\psi(z)$ satisfies the identity

$$(k+1)\frac{\psi''(z)}{\psi'(z)} - (k+2)\frac{\psi'(z)}{\psi(z)-b} = \frac{2}{k}\frac{a_{k-1}(z)}{a_k(z)} - 2\frac{a_k'(z)}{a_k(z)};$$

(ii) *if $a_k(z) \equiv 1$ and $a_{k-1}(z)$ is entire, then either* (1.4) *holds or $\psi(z)$ has no b-points at all, and in particular, if $a_k(z) \equiv 1$ and $a_{k-1}(z) \equiv 0$,* (1.4) *always holds;*

(iii) *if $a_0(z),\dots,a_k(z)$ are constants with $a_k = 1$, then either* (1.4) *holds or*

$$T(r,f) = O\left(\bar{N}\left(r,\frac{1}{f}\right)\right),$$

or

$$f(z) = e^{\alpha z}(A + Be^{\alpha z})^{-1}$$

and

$$\psi(z) = (D+\alpha)\dots(D+k\alpha)f(z)$$

where A and B are constants, $\alpha = \dfrac{2}{k(k+1)}a_{k-1}$, *and* $D \equiv \dfrac{d}{dz}$.

The second part of the paper is concerned with non-linear differential polynomials (see [3], p. 68). Doeringer [1] has proved:

Theorem C. *If $f(z)$ is a transcendental entire function, and $Q[f]$, $P[f]$ are differential polynomials in $f(z)$, both not identically zero, then if*

$$\phi(z) = f^n Q[f] + P[f]$$

with $n \geqq 2+\gamma$, where γ is the degree of $P[f]$ as a polynomial in $f, f', \dots,$ we have

$$\limsup_{r\to\infty} \frac{\bar{N}\left(r,\dfrac{1}{\phi}\right)}{T(r,\phi)} > 0.$$

This is the best possible general result in the sense that, if w is a suitably chosen transcendental solution of $w^{(j)} = -2\rho\sigma(w-\sigma)$ with ρ,σ non-zero constants, then $w^{(j)}+\rho w^2$ may omit altogether the value $\rho\sigma^2$. However it is remarked by Hayman in [2] that Clunie showed (in a lecture) that these are the only entire functions of finite order for which this is possible. We enlarge upon this observation with the following result and corollary:

Theorem 2. *Suppose that $w(z)$ is a transcendental entire function, and that*

$$\sum_{i=0}^{n} \alpha_i w^{(i)}(z)+w^2(z)-A=\phi(z) \tag{1.6}$$

where $\alpha_0,\ldots,\alpha_n$, and A are constants, and $\phi(z)$ satisfies $\bar{N}\left(r,\frac{1}{\phi}\right)=S(r,\phi)$. Then

$$w(z)=a+p(z)e^{\alpha z}$$

where a and α are constants, with $\alpha_0 a+a^2-A=0$, and $p(z)$ is a polynomial.

Corollary. (i) *With the hypotheses of Theorem 2, if $A=\alpha_0=0$, then $\sum_{i=1}^{n}\alpha_i w^{(i)}\equiv 0$.*

(ii) *If $w(z)$ is a transcendental entire function, then $\bar{N}\left(r,\frac{1}{w^{(k)}+w^2}\right)\neq S(r,w)$ for any $k\geqq 1$.*

Remark. Part (ii) of the corollary was proved in a stronger form by Hayman [2] in the case $k=1$.

Acknowledgement. I wish to thank Professor Steven Bank for his invaluable advice and suggestions, particularly with reference to Theorem 2, which arose from joint discussions. I would also like to acknowledge valuable conversations with G. Gundersen.

2. Preliminary Lemmas

Lemma 1. *Suppose that A, B, α, γ and d are complex numbers with $AB\alpha\neq 0$ and $A+Be^{\alpha d}\neq 0$, and that k is an integer not less than 1. Then:*

(i) $I_{k,\gamma}(z)=\int_d^z \frac{e^{\gamma t}}{(A+Be^{\alpha t})^{k+1}}\,dt$

is meromorphic in the plane if and only if $\gamma/\alpha\in\{1,\ldots,k\}$;

(ii) *if $r\in\{1,\ldots,k\}$,*

$$G_{k,r}(z)=\int_d^z\int_d^\xi \frac{e^{r\alpha t}}{(A+Be^{\alpha t})^{k+1}}\,dt\,d\xi$$

is not meromorphic in the plane.

Proof. We assume without loss of generality that $\alpha=1$, and prove both propositions by induction. To prove (i), we note that integration by parts yields:

$$\begin{aligned} I_{k,\gamma}(z)&=\int_d^z e^{(\gamma-1)t}\frac{e^t}{(A+Be^t)^{k+1}}\,dt\\ &=e^{(\gamma-1)z}(-(Bk)^{-1})(A+Be^z)^{-k}+C_{k,\alpha}+\frac{\gamma-1}{Bk}\int_d^z\frac{e^{(\gamma-1)t}}{(A+Be^t)^k}\,dt \end{aligned} \tag{2.1}$$

where $C_{k,\alpha}$ is a constant. Now the first term on the right-hand side of (2.1) is clearly meromorphic in the plane; however, if $k=1$, the integral on the right-hand side of (2.1) is not meromorphic in the plane, since the integrand has simple poles. Thus $I_{1,\gamma}(z)$ is meromorphic in the plane if and only if $\gamma=1$; for $k>1$, we see that $I_{k,\gamma}(z)$ is meromorphic in the plane if and only if $\gamma=1$ or $I_{k-1,\gamma-1}(z)$ is meromorphic in the plane, and the proposition follows by induction.

To prove (ii), we note that, if $r=1$,

$$G_{k,1}(z)=\int_d^z((-Bk)^{-1}(A+Be^{\xi})^{-k}+D_k)\,d\xi$$

for some constant D_k, and hence is not meromorphic in the plane, using part (i) in the case $k>1$. Also, if $k\geqq 2$ and $r\in\{2,\dots,k\}$, integration by part yields

$$G_{k,r}(z)=\int_d^z(-Bk)^{-1}e^{(r-1)\xi}(A+Be^{\xi})^{-k}+E_{k,r}\,d\xi+\left(\frac{r-1}{Bk}\right)\int_d^z\int_d^{\xi}\frac{e^{(r-1)t}}{(A+Be^t)^k}\,dt\,d\xi \quad (2.2)$$

for some constant $E_{k,r}$. Now, the first integral on the right-hand side is meromorphic in the plane, by part (i); since $r\neq 1$ we see that if $G_{k-1,r-1}(z)$ is not meromorphic in the plane, then neither is $G_{k,r}(z)$ and proposition (ii) follows by induction.

The following lemma, due to Steven Bank, seems not without interest.

Lemma 2. *Suppose that $\lambda_1,\dots,\lambda_n$ are distinct complex numbers, and that $Q_1,\dots,Q_n$ are functions meromorphic in the plane such that*

$$T(r,Q_i)=o(r) \quad \text{as } r\to\infty$$

for each $i=1,\dots,n$. Then if

$$\Pi(z)=\sum_{i=1}^{n}Q_i(z)e^{\lambda_i z} \quad (2.3)$$

satisfies

$$\liminf_{r\to\infty}\frac{1}{r}\bar{N}\left(r,\frac{1}{\Pi}\right)=0 \quad (2.4)$$

we have $Q_i(z)\equiv 0$ for all but at most one $i\in\{1,\dots,n\}$.

Proof. We prove first the following claim:

Claim. If $\mu_1,\dots,\mu_n$ are distinct complex numbers, and if $E_1,\dots,E_m$ are meromorphic functions of finite order in the plane, each satisfying

$$T(r_k,E_i)=o(r_k)$$

through some common sequence (r_k) tending to infinity, then if

$$\sum_{i=1}^{m}E_i(z)e^{\mu_i z}\equiv 0 \quad (2.5)$$

we have $E_i(z)\equiv 0$ for each i.

Since the claim is trivial for $m=1$, we may suppose that $m\geqq 2$ is the least positive integer for which the claim fails, and that (2.5) holds with, say, $E_1\not\equiv 0$. Also we may assume that $\mu_1=0$. But then

$$\sum_{i=2}^{m} R_i(z)e^{\mu_i z}\equiv -1 \tag{2.6}$$

where each $R_i(z)=E_i(z)(E_1(z))^{-1}$ satisfies $T(r_k, R_i)=o(r_k)$ as $r_k\to\infty$. Differentiating (2.6) we have

$$\sum_{i=2}^{m} (R_i'(z)+\mu_i R_i(z))e^{\mu_i z}\equiv 0. \tag{2.7}$$

Since

$$T(r, R_i')<2T(r, R_i)+O(\log r)$$

([3] pp. 36, 55), we have, from (2.7),

$$R_i'(z)+\mu_i R_i(z)\equiv 0$$

which leads to, for some constant c_i,

$$R_i(z)=c_i e^{-\mu_i z}$$

if $R_i\not\equiv 0$. But then $T(r_k, R_i)\neq o(r_k)$ as $r_k\to\infty$, and we conclude that $E_i\equiv 0$ for $i=2,\dots,m$, and hence $E_1\equiv 0$, establishing the claim.

To prove the lemma, we assume that $Q_1\not\equiv 0$ and set $h=\Pi'/\Pi$. Since Π has finite order (by 2.3) we have

$$T(r,h)<\bar{N}\left(r,\frac{1}{\Pi}\right)+O(\log r),$$

and thus

$$T(r_k, h)=o(r_k)$$

for some sequence (r_k) tending to infinity.

Now, $\Pi'=h\Pi$, and so

$$\sum_{i=1}^{n} H_i(z)e^{\lambda_i z}\equiv 0$$

where each

$$H_i(z)=Q_i'(z)+(\lambda_i-h(z))Q_i(z)$$

has finite order and satisfies

$$T(r_k, H_i)=o(r_k)$$

as $r_k\to\infty$. By the claim above, $H_i\equiv 0$ for each i, and so

$$h(z)=\lambda_1+\frac{Q_1'(z)}{Q_1(z)}$$

which gives

$$\Pi(z)=cQ_1(z)e^{\lambda_1 z}$$

for some non-zero constant c. From the claim above, we conclude that $c=1$ and $Q_i(z)\equiv 0$ for $i=2,\dots,n$.

3. Proof of Theorem 1

Suppose that $\psi(z)$ is non-constant and is given by (1.1) and (1.2), where $f(z)$ is transcendental and meromorphic in the plane, and $a_k(z)$ is not identically zero. Suppose further that z_0 is a simple pole of $f(z)$ which is not a pole of any of $a_0(z), \ldots, a_k(z)$, nor a zero of $a_k(z)$. Then we may write

$$f(z)=\frac{d_1}{z-z_0}+O(1) \tag{3.1}$$

near $z=z_0$, for some $d_1 \neq 0$. Now from (1.2) we see that $\psi(z)$ has a pole of order $(k+1)$ at z_0, and that the coefficient of $(z-z_0)^{-k-1}$ in the Laurent series expansion of $\psi(z)-b$ about z_0 is

$$(-1)^k k! d_1 a_k(z_0) \tag{3.2}$$

while that of $(z-z_0)^{-k}$ is

$$(-1)^{k-1}(k-1)! d_1 a_{k-1}(z_0)+(-1)^k k! d_1 a_k'(z_0). \tag{3.3}$$

Combining (3.2) and (3.3) we obtain, near $z=z_0$,

$$\frac{\psi'(z)}{\psi(z)-b}=\frac{-(k+1)}{(z-z_0)}+\left(\frac{a_k'(z_0)}{a_k(z_0)}-\frac{a_{k-1}(z_0)}{k a_k(z_0)}\right)+O(z-z_0). \tag{3.4}$$

From (1.2) we have

$$\psi'(z)=a_k(z)f^{(k+1)}(z)+(a_k'(z)+a_{k-1}(z))f^{(k)}(z)+\ldots+a_0'(z)f(z)$$

and by the same reasoning as above, noting that $\psi'(z)$ has a pole of order $(k+2)$ at z_0, we obtain, near z_0,

$$\frac{\psi''(z)}{\psi'(z)}=\frac{-(k+2)}{(z-z_0)}+\left(\frac{a_k'(z_0)}{a_k(z_0)}-\frac{(a_k'(z_0)+a_{k-1}(z_0))}{(k+1)a_k(z_0)}\right)+O(z-z_0). \tag{3.5}$$

It follows that, near z_0,

$$(k+1)\frac{\psi''(z)}{\psi'(z)}-(k+2)\frac{\psi'(z)}{(\psi(z)-b)}=\frac{2}{k}\frac{a_{k-r}(z_0)}{a_k(z_0)}-2\frac{a_k'(z_0)}{a_k(z_0)}+O(z-z_0). \tag{3.6}$$

Setting

$$h(z)=(k+1)\frac{\psi''(z)}{\psi'(z)}-(k+2)\frac{\psi'(z)}{\psi(z)-b}+2\frac{a_k'(z)}{a_k(z)}-\frac{2}{k}\frac{a_{k-1}(z)}{a_k(z)}, \tag{3.7}$$

we distinguish 2 possible cases:

Case I. Suppose that $h(z)$ is not identically zero. From (3.6) and (3.7) we see that $h(z_0)=0$, and thus, if $N_1(r,f)$ counts the simple poles of $f(z)$, we have, using (1.1),

$$N_1(r,f)<N\left(r,\frac{1}{h}\right)+S(r,f). \tag{3.8}$$

We may now follow Hayman's proof ([3], p. 60). Applying Jensen's formula to $h(z)$, (3.8) yields

$$\begin{aligned} N_1(r,f) &< m(r,h)+N(r,h)+S(r,f) \\ &< N(r,h)+S(r,f) \end{aligned} \tag{3.9}$$

noting that

$$T(r,\psi)=O(T(r,f)) \text{ n.e.}$$

(Note that (3.9) is trivial if $h(z)$ is a non-zero constant.) Now $h(z)$ can only have simple poles at zeros and poles of $\psi'(z)$ and $\psi(z)-b$ which do not satisfy the condition above on z_0, and possibly other poles at zeros and poles of $a_k(z)$ and $a_{k-1}(z)$. If $\bar{N}_2(r,f)$ counts the points at which f has a multiple pole, each counted only once irrespective of multiplicity, then from (3.9) we obtain

$$N_1(r,f)<\bar{N}_2(r,f)+\bar{N}\left(r,\frac{1}{\psi-b}\right)+N_0\left(r,\frac{1}{\psi'}\right)+S(r,f), \tag{3.10}$$

where $N_0\left(r,\frac{1}{\psi'}\right)$ is as defined in Theorem A.

Now, using Theorem A with ψ replaced by $b\psi$, we have

$$T(r,f)<\bar{N}(r,f)+N\left(r,\frac{1}{f}\right)+\bar{N}\left(r,\frac{1}{\psi-b}\right)-N_0\left(r,\frac{1}{\psi'}\right)+S(r,f) \tag{3.11}$$

and noting that

$$N_1(r,f)+2\bar{N}_2(r,f)\leqq N(r,f)\leqq T(r,f)$$

and

$$\bar{N}(r,f)=N_1(r,f)+\bar{N}_2(r,f)$$

we obtain, from (3.11),

$$\bar{N}_2(r,f)<N\left(r,\frac{1}{f}\right)+\bar{N}\left(r,\frac{1}{\psi-b}\right)-N_0\left(r,\frac{1}{\psi'}\right)+S(r,f). \tag{3.12}$$

Finally, on combining (3.10), (3.11) and (3.12) we have

$$T(r,f)<3N\left(r,\frac{1}{f}\right)+4\bar{N}\left(r,\frac{1}{\psi-b}\right)+S(r,f)$$

provided $h(z)$ does not vanish identically. This proves part (i).

We consider now:

Case II. Suppose that $h(z)\equiv 0$. Then we have the identity

$$(k+1)\frac{\psi''(z)}{\psi'(z)}-(k+2)\frac{\psi'(z)}{\psi(z)-b}=\frac{2}{k}\frac{a_{k-1}(z)}{a_k(z)}-2\frac{a_k'(z)}{a_k(z)}, \tag{3.13}$$

which of course is impossible if $a_{k-1}(z)/a_k(z)$ has a multiple pole or a simple pole whose residue is not an integer multiple of $k/2$. Now if $a_k(z)\equiv 1$ and $a_{k-1}(z)$ is entire then (3.13) yields, for some $d_2 \neq 0$,

$$(\psi'(z))^{k+1}(\psi(z)-b)^{-k-2}=d_2\exp\left(\frac{2}{k}\int_0^z a_{k-1}(t)\,dt\right)$$

which we may solve, using the substitution $\psi - b = e^v$, to obtain

$$\psi(z) - b = \left[A_1 + A_2 \int_0^z \exp\left[\frac{2}{k(k+1)} \int_0^s a_{k-1}(t)\,dt\right] ds\right]^{-k-1} \tag{3.14}$$

for suitable constants A_1, A_2. We note that (3.14) implies that $\psi(z)$ has no b-points at all. Also, if $a_{k-1}(z) \equiv 0$, then from (3.14) we see that $\psi(z)$ must be rational, and hence from (1.2) $f(z)$ has only finitely many poles which are not poles of any of the functions $a_0(z), \ldots, a_{k-2}(z)$, and hence $\bar{N}(r,f) = S(r,f)$ in (3.11). This proves part (ii).

If seems worth observing that the exceptional case (3.13) is quite possible. For any non-constant entire function $g(z)$, set

$$f(z) = ((1+e^g)g')^{-1}. \tag{3.15}$$

Then $f(z)$ is non-vanishing. Also

$$f' + \left(\frac{g''}{g'} - g'\right) f = -1 + \frac{e^{2g}}{(1+e^g)^2}$$

and from (3.15) we see that (see [3], p. 54)

$$T\left(r, \frac{g''}{g'} - g'\right) = S(r, f).$$

To prove part (iii), we observe that if $a_0, \ldots, a_{k-1}$ are constants, then either (1.4) holds, or we may write (3.14) in the form

$$\psi(z) = b + (A + Be^{\alpha z})^{-(k+1)} \tag{3.16}$$

where $a_{k-1} = \dfrac{k(k+1)}{2}\alpha$ and A and B may be assumed non-zero, since otherwise we may solve (3.16) for $f(z)$, and conclude that $f(z)$ is entire. We may also write

$$\psi(z) = (D+\gamma_1)(D+\gamma_2)\ldots(D+\gamma_k) f(z) \tag{3.17}$$

where $D \equiv \dfrac{d}{dz}$, and $\gamma_1, \ldots, \gamma_k$ are constants satisfying $\sum_{i=1}^{k} \gamma_i = a_{k-1}$. Since the operators $(D+\gamma_i)$ commute, we may combine (3.16) and (3.17) to obtain, for any i,

$$(D+\gamma_i)u(z) = \psi(z)$$

and hence

$$D(e^{\gamma_i z} u(z)) = e^{\gamma_i z}(b + (A + Be^{\alpha z})^{-(k+1)}),$$

where $u(z)$ is meromorphic. It follows from Lemma 1 that $\gamma_i \in \{\alpha, 2\alpha, \ldots, k\alpha\}$. We now show that the γ_i are distinct (in the case $k \geqq 2$). For suppose that, say, $\gamma_1 = \gamma_2 = \gamma$. Then we have

$$(D+\gamma)^2 w(z) = F(z) = b + (A + Be^{\alpha z})^{-(k+1)} \tag{3.18}$$

where $w(z)$ is meromorphic. But it is easily checked that (3.18) has solution

$$w(z)=e^{-\gamma z}\left(q(z)+\int_a^z\int_a^\xi e^{\gamma t}F(t)\,dt\,d\xi\right) \tag{3.19}$$

where $q(z)$ is linear with arbitrary coefficients and a is some point with $F(a)\neq\infty$. However, since $\gamma\in\{\alpha,\dots,k\alpha\}$, it follows that the right-hand side of (3.19) is not meromorphic, by Lemma 1, and we have a contradiction.

Since we now have $\{\gamma_1,\dots,\gamma_k\}=\{\alpha,\dots,k\alpha\}$ we may write

$$(D+\alpha)\dots(D+k\alpha)\,f(z)=b+(A+Be^{\alpha z})^{-(k+1)}$$

and solve explicitly for $f(z)$, to obtain

$$f(z)=\frac{b}{k!\,\alpha^k}+\frac{(-1)^k}{k!\,B^k\alpha^k}\,\frac{e^{-k\alpha z}}{(A+Be^{\alpha z})}+\mu_1 e^{-\alpha z}+\dots+\mu_k e^{-k\alpha z}$$

where $\mu_1,\dots,\mu_k$ are constants arising from integration. Thus

$$f(z)=\frac{1}{(A+Be^{\alpha z})}(\lambda_1 e^{\alpha z}+\lambda_2+\lambda_3 e^{-\alpha z}+\dots+\lambda_{k+2}e^{-k\alpha z})$$

where $\lambda_1=bB(k!\,\alpha^k)^{-1}\neq 0$. It follows from Lemma 2 that $T(r,f)=O\left(\bar{N}\left(r,\frac{1}{f}\right)\right)$ unless $\lambda_2=\lambda_3=\dots=\lambda_{k+2}=0$ (noting that $T(r,f)=O(r)$). This proves part (iii). As an example of the exceptional case, set $f_1(z)=(2+e^{-2z})^{-1}$. We have $f_1(z)\neq 0$; also $f_1'(z)+2f_1(z)$ is non-constant but omits the value 1.

4. Proof of Theorem 2

Suppose that

$$\sum_{i=0}^{n}\alpha_i w^{(i)}+w^2=A+\phi \tag{4.1}$$

where $w(z)$ is a transcendental entire function, $\alpha_0,\dots,\alpha_n$ and A are complex numbers, and ϕ satisfies

$$\bar{N}\left(r,\frac{1}{\phi}\right)=S(r,\phi).$$

From (4.1) we see that $T(r,\phi)=O(T(r,w))$ n.e., and thus, by the Tumura-Clunie theory ([3], Theorem 3.9, p. 69) we have

$$\phi(z)=(w(z)-a(z))^2$$

where $a(z)$ is entire and satisfies $T(r,a)=S(r,w)$. We set $\phi_1(z)=w(z)-a(z)$, and note that, from (4.1),

$$\begin{aligned}2m(r,w)&\leqq m\left(r,\sum_{i=0}^{n}\alpha_i w^{(i)}\right)+m(r,\phi)+O(1)\\&\leqq m\left(r,\sum_{i=0}^{n}\alpha_i\frac{w^{(i)}}{w}\right)+m(r,w)+m(r,\phi)+O(1)\end{aligned}$$

so that

$$m(r,w)=O(m(r,\phi_1)) \text{ n.e.} \tag{4.2}$$

and

$$T(r,a)=S(r,\phi_1). \tag{4.3}$$

Substituting $w=\phi_1+a$ in (4.1) we obtain

$$\sum_{i=1}^{n} \alpha_i \phi_1^{(i)}+(\alpha_0+2a)\phi_1=A-\sum_{i=0}^{n} \alpha_i a^{(i)}-a^2. \tag{4.4}$$

We set

$$D(z)=\sum_{i=1}^{n} \alpha_i \frac{\phi_1^{(i)}(z)}{\phi_1(z)}+\alpha_0+2a$$

and note that, since $\bar{N}\left(r,\frac{1}{\phi_1}\right)=S(r,\phi_1)$, we have

$$T(r,D)=S(r,\phi_1).$$

But then, unless $D(z)\equiv 0$, we have

$$\phi_1(z)=(D(z))^{-1}B(z) \tag{4.5}$$

where

$$B(z)=A-\sum_{i=0}^{n} \alpha_i a^{(i)}-a^2 \tag{4.6}$$

satisfies $T(r,B)=S(r,\phi_1)$, by (4.3). But (4.5) leads to

$$T(r,\phi_1)=o(T(r,\phi_1) \text{ n.e.}$$

and we conclude that $D(z)\equiv 0$, and hence $B(z)\equiv 0$. But then

$$-\sum_{i=0}^{n} \alpha_i a^{(i)}-a^2+A\equiv 0 \tag{4.7}$$

and the Wiman-Valiron theory ([4], p. 341) shows that the entire function a cannot be transcendental; by a degree argument, a cannot be a non-constant polynomial, and we conclude that a is a constant satisfying $\alpha_0 a+a^2-A\equiv 0$. Returning to (4.4) we see that ϕ_1 satisfies a linear differential equation with constant coefficients, and thus we may write

$$\phi_1(z)=\sum_{i=1}^{s} q_i(z)e^{\lambda_i z}$$

where $s\leqq n$, $\lambda_1,\dots,\lambda_s$ are distinct complex numbers, and $q_1,\dots,q_s$ are polynomials. Since $\bar{N}\left(r,\frac{1}{\phi_1}\right)=S(r,\phi_1)=o(r)$ n.e. it follows from Lemma 2 that $q_i(z)\equiv 0$ for all but at most one i, and Theorem 2 is proved.

To prove the corollary, we note that if $A=0$ and $\alpha_0=0$ in (4.1), and ϕ has few zeros, then $a=0$ and $w(z)=p(z)e^{\alpha z}$ where α is a constant and $p(z)$ is a polynomial. But then, for some polynomial $q(z)$,

$$\sum_{i=1}^{n} \alpha_i w^{(i)} + w^2 = q(z) e^{\alpha z} + p^2(z) e^{2\alpha z}$$

and, by Lemma 2, we must have $q(z) \equiv 0$, and hence $\sum_{i=1}^{n} \alpha_i w^{(i)} \equiv 0$.

Remarks. Since writing the first draft of this paper I have learned (through Professor W.K. Hayman and correspondence) of progress made by Yang Lo ("A General Criterion for Normality" - not yet published) on the problem of Theorem 1 and wish to thank both for their comments. Professor Hayman has also pointed out to me that Lemma 2 (due to Steven Bank) could be shortened by reference to a result of R. Nevanlinna, but it seems worth including in its present, self-contained form.

References

1. Doeringer, W.: Exceptional Values of Differential Polynomials. Pacific Journal of Math. **98**, 55–62 (1982)
2. Hayman, W.K.: Picard Values of Meromorphic Functions and Their Derivatives. Ann. of Math., II. Ser. **70**, 9–42 (1959)
3. Hayman, W.K.: Meromorphic Functions. Oxford: Clarendon Press 1964
4. Hayman, W.K.: The Local Growth of Power Series: a Survey of the Wiman-Valiron Method. Canad. Math. Bull. **17**, 317—358 (1974)

Received November 14, 1983

Math. Z. 187, 13-21 (1984)

Mathematische Zeitschrift

Measurable Refinement Monoids and Applications to Distributive Semilattices, Heyting Algebras, and Stone Spaces

Hans Dobbertin

Institut für Mathematik, Universität Hannover, Welfengarten 1, D-3000 Hannover 1, Federal Republic of Germany

Main results of this paper are that every Heyting algebra can be embedded into the Heyting algebra of all ideals of some atomless Boolean algebra and that the second countable compact Stone spaces can be characterized as the T_0-images of the Cantor set under continuous open mappings. The first section is devoted to the study of certain measures on Boolean algebras with values in refinement monoids (see also [3]). A criterion is given which guarantees that a refinement monoid admits, loosely speaking, many such measures. An application implies that every distributive lattice with zero is the image of some atomless generalized Boolean lattice under a V-homomorphism. In Sects. 2 and 3 we investigate mappings between algebraic Heyting algebras and between Stone spaces, respectively, which are induced by V-homomorphisms between their associated distributive sup-semilattice with zero. These categorical considerations together with results of the first section form the background which allows to derive the two results mentioned in the beginning.

1. Refinement Monoids and Measures

First we shall recall some definitions from [3]. Let $M=(M;+,0)$ be a commutative monoid. M possesses the *refinement property* if whenever $\sum x_i=\sum y_j$ (x_i, $y_j \in M$, $i<n$, $j<m$, $n,m \geqq 2$) then there are $z_{ij} \in M$ with $x_i=\sum_j z_{ij}$ and $y_j=\sum_i z_{ij}$. We call M a *refinement monoid* if the refinement property is valid for M and the sum of non-zero elements of M is always non-zero. A homomorphism $h\colon M \to N$ between commutative monoids is a *V-homomorphism* if $h(x)=y_1+y_2$ ($x \in M$, $y_1, y_2 \in N$) implies $x=x_1+x_2$ and $h(x_i)=y_i$ for some elements $x_i \in M$ and if $h(x)=0$ only for $x=0$ ($x \in M$).

Let B be a Boolean algebra. Then a mapping $\mu\colon B \to M$ is said to be a *V-measure* provided that

(i) $b_1 b_2 = 0$ ($b_1, b_2 \in B$) implies $\mu(b_1 + b_2) = \mu(b_1) + \mu(b_2)$,

(ii) $\mu(b) = 0$ if and only if $b = 0$ ($b \in B$),

(iii) if $\mu(b) = x_1 + x_2$ ($b \in B$, $x_1, x_2 \in M$) then there are $b_i \in B$ with $b = b_1 \dotplus b_2$ (i.e., $b = b_1 + b_2$ and $b_1 b_2 = 0$) and $\mu(b_i) = x_i$.

Moreover, μ is a *measure* (resp., *semimeasure*) if (i) and (ii) (resp., (i) and $\mu(0) = 0$) are valid.

Lemma 1.1. *Let μ: $B \to M$ be a semimeasure, and suppose that $x + y = 0$ implies $x = y = 0$ for all $x, y \in M$. Then $I = \mu^{-1}(0)$ is an ideal of B, and there is a unique measure $\tilde{\mu}$: $B/I \to M$ such that $\mu = \tilde{\mu}h$, where h: $B \to B/I$ denotes the canonical homomorphism:*

$$\begin{tikzcd} B \arrow[r, "\mu"] \arrow[d, "h"'] & M \\ B/I \arrow[ur, "\tilde{\mu}"'] & \end{tikzcd}$$

Let μ: $B \to M$ and ν: $C \to M$ be measures. If e: $B \to C$ is a (Boolean algebra) embedding satisfying $\mu = \nu e$ then we say that e: $(B, \mu) \to (C, \nu)$ is an *embedding*. Conversely, when B is a subalgebra of C and $\mu = \nu | B$ then (C, ν) is an *extension* of (B, μ).

M is *measurable* if for all $x \in M$ there exists some Boolean algebra B and a V-measure μ: $B \to M$ with $\mu(1) = x$. Note that in this case M is necessarily a refinement monoid. We say that M has the *measure extension property* (*MEP*) provided that for each measure μ: $B \to M$ there is an extension (C, ν) such that ν is a V-measure. Observe that if M has the MEP then M is measurable.

Throughout the remainder of this section let M be a refinement monoid.

Theorem 1.2. *M has the MEP if and only if the following condition is satisfied:*

Whenever μ: $B \to M$ is a measure on some Boolean algebra B and $\mu(1) = x_1 + x_2$ ($x_i \in M$) then there are semimeasures μ_i: $B \to M$ ($i = 1, 2$) such that $\mu_i(1) = x_i$ and $\mu = \mu_1 + \mu_2$.

Proof. "$\Rightarrow$". Let (C, ν) be an extension of (B, μ), where ν is a V-measure. Then we find $c_i \in C$ with $\nu(c_i) = x_i$ and $1 = c_1 \dotplus c_2$, because $\nu(1) = x_1 + x_2$. Set $\mu_i(b) = \nu(bc_i)$ ($b \in B$). "$\Leftarrow$". For the sake of brevity, we need a technical expression: if μ: $B \to M$ is a measure and $H \subseteq B \times M \times M$ then (B, μ) is called *H-splitting* provided that for each $(b, x_1, x_2) \in H$ there are $b_i \in B$ with $b = b_1 \dotplus b_2$ and $\mu(b_i) = x_i$. First we show that the following condition is satisfied:

($*$) Whenever μ: $B \to M$ is a measure and $\mu(1) = x_1 + x_2$ then there is a $(1, x_1, x_2)$-splitting extension (C, ν) of (B, μ).

In fact, choose semimeasures μ_i: $B \to M$ according to our assumption. Let d: $B \to B \oplus B$ denote the diagonal embedding. Clearly,

$$\mu_1 \oplus \mu_2\colon\ B \oplus B \to M, \qquad (b_1, b_2) \mapsto \mu_1(b_1) + \mu_2(b_2)$$

is a semimeasure, and $\mu = (\mu_1 \oplus \mu_2)d$. Define $C = B \oplus B / (\mu_1 \oplus \mu_2)^{-1}(0)$, and let h: $B \oplus B \to C$ be the canonical homomorphism. By Lemma 1.1, there exists a measure ν: $C \to M$ such that $\mu_1 \oplus \mu_2 = \nu h$. Obviously, $e = hd$ embeds (B, μ) into (C, ν), and we have $\nu(h(1,0) = x_1$, $\nu(h(0,1)) = x_2$, and $e(1) = h(1,0) \dotplus h(0,1)$. Thus (C, ν) has the desired property, up to isomorphism, and $(*)$ is verified.

Next, by an inductive construction (cf. the first part of the proof of [3; Lemma 3.3]) one can sharpen $(*)$ as follows:

$(**)$ Whenever μ: $B \to M$ is a measure then there exists a H-splitting extension of (B, μ), where

$$H = \{(b, x_1, x_2) \in B \times M \times M\colon \mu(b) = x_1 + x_2\}.$$

Now, if μ: $B \to M$ is a measure then, applying $(**)$, we obtain a sequence (B_n, μ_n) $(n < \omega)$ such that:

(i) $(B_0, \mu_0) = (B, \mu)$,

(ii) (B_n, μ_n) is a H_n-splitting extension of (B_{n-1}, μ_{n-1}) with

$$H_n = \{(b, x_1, x_2) \in B_{n-1} \times M \times M\colon \mu_{n-1}(b) = x_1 + x_2\}.$$

Set $C = \bigcup_{n<\omega} B_n$ and $\nu = \bigcup_{n<\omega} \mu_n$, then ν is a V-measure. This completes the proof.

We shall use Theorem 1.2 to show the MEP for some classes of refinement monoids.

Corollary 1.3. *Every distributive lattice with zero is a refinement monoid (under supremum) which possesses the MEP.*

Proof. Let L be a distributive lattice with zero. Under the assumption of the criterion in Theorem 1.2 set $\mu_i(b) = x_i \mu(b)$ $(b \in B,\ x_i \in L)$.

Corollary 1.4. *The positivity domain of an ordered field is a refinement monoid (under addition) which possesses the MEP.*

Proof. Let K be an ordered field and $P = \{x \in K\colon x \geqq 0\}$. Using Theorem 1.2, put $\mu_i(b) = x_i \mu(1)^{-1} \mu(b)$ $(b \in B,\ x_i \in P)$.

Given a commutative group G, we obtain the associated *extended commutative group* by adding a new zero element to G (see [4]).

Corollary 1.5. *Every extended commutative group is a refinement monoid which possesses the MEP.*

Proof. Let M be an extended commutative group. In order to verify the condition of Theorem 1.2 suppose that μ: $B \to M$ is a measure and $\mu(1) = x_1 + x_2$ $(x_i \in G = M - \{0\})$. Choose an arbitrary prime ideal P of B. For $x \in G$ define the "point" semimeasure δ_x: $B \to M$ by setting

$$\delta_x(b) = \begin{cases} 0 & \text{for } b \in P, \\ x & \text{for } b \in B - P. \end{cases}$$

Now put $\mu_1 = \delta_{x_1}$ and $\mu_2 = \delta_{-x_1} + \mu$.

Corollary 1.6. *Distributive lattices with zero, positivity domains of ordered fields, and extended commutative groups are measurable.*

Corollary 1.7. (i) *Every bounded distributive lattice is the image of some atomless Boolean lattice under a V-homomorphism.* (ii) *Every distributive lattice with zero is the image of some atomless generalized Boolean lattice under a V-homomorphism.*

Proof. First suppose that L is a bounded distributive lattice. By Corollary 1.6, there exists a V-measure $\mu\colon B\to L$ on some Boolean algebra B such that $\mu(1)=1$. Note that B is necessarily atomless. We claim that μ is in fact an onto V-homomorphism (where B and L are viewed as sup-semilattices). It remains to show that μ preserves finite suprema:

$$\begin{aligned}\mu(b_1+b_2)&=\mu(b_1\,b_2)+\mu(b_1\,b_2')+\mu(b_1'\,b_2)\\&=\mu(b_1\,b_2)+\mu(b_1\,b_2')+\mu(b_1\,b_2)+\mu(b_1'\,b_2)\\&=\mu(b_1)+\mu(b_2).\end{aligned}$$

In case L has no greatest element let, for each $x\in L$, h_x be a V-homomorphism from some atomless Boolean lattice B_x into L with $h_x(1)=x$. Define

$$B=\{b\in\prod_{x\in L}B_x\colon b_x=0 \text{ for almost all } x\}.$$

Then B is an atomless generalized Boolean lattice (under pointwise operations), and

$$h\colon B\to L,\, b\mapsto\sum_{x\in L}h_x(b_x)$$

is an onto V-homomorphism.

The preceding proof shows that every measurable distributive sup-semilattice L with zero is the V-homomorphic image of some atomless generalized Boolean lattice. According to [3; Theorem 3.4], L is measurable if each principal ideal of L has at most $\aleph_1$ many elements. The general case remains open; we even do not know whether there is any non-measurable refinement monoid.

The measurability of positivity domains of ordered fields (Corollary 1.6) is a special case of the following theorem:

Proposition 1.8. *Positivity domains of linearly ordered commutative groups are measurable refinement monoids.*

Proof. Suppose that $G=(G;+,0,\leqq)$ is a linearly ordered commutative group, $P=\{x\in G\colon x\geqq 0\}$, and $x_0\in P-\{0\}$. Let B be the Boolean algebra generated by the chain $C=\{x\in P\colon 0<x\leqq x_0\}$, that is, B consists of all finite unions of half-open intervals $(y,z]$ $(y,z\in C\cup\{0\}, y\leqq z)$. Obviously there is a unique measure $\mu\colon B\to P$ such that $\mu((y,z])=z-y$. It is easy to see that μ is a V-measure. Further, we have $\mu(1)=\mu((0,x_0])=x_0$ showing that P is measurable.

2. Distributive Semilattices and Heyting Algebras

For undefined terminology and notations used in this and the next section, the reader is referred to the Compendium of Continuous Lattices [5]. We use the abbreviation sup (resp., Sup) for *finite supremum* (resp., *infinite supremum*). Similarly, inf (resp., Inf) stands for *finite infimum* (resp., *infinite infimum*).

Let $\mathscr{DistrSemilattice}$ be the category of distributive sup-semilattices with zero and V-homomorphisms. A mapping $h: L \to K$ between complete lattices is said to be a *strong V-homomorphism* provided that

(i) $h(x)=0$ if and only if $x=0$,

(ii) h preserves Sups, i.e., $h(\sum_{i\in I} x_i)=\sum_{i\in I} h(x_i)$,

(iii) if $h(x)=\sum_{i\in I} y_i$ then there are x_i $(i\in I)$ such that $x=\sum_{i\in I} x_i$ and $h(x_i)=y_i$.

A Heyting algebra is *algebraic* if it is complete and each element is a Sup of compact elements. Let $\mathscr{AlgHeyt}$ (resp., $\mathscr{AlgHeyt}_0$) be the category of algebraic Heyting algebras and strong V-homomorphisms (resp., maps preserving Sups, Infs, and relative pseudocomplements).

Theorem 2.1. *The categories $\mathscr{DistrSemilattice}$ and $\mathscr{AlgHeyt}$ are equivalent.*

Proof. As usual, we associate to $L\in \mathrm{obj}(\mathscr{DistrSemilattice})$ the ideal lattice of L and to $H\in \mathrm{obj}(\mathscr{AlgHeyt})$ the sup-semilattice of compact elements of H. This gives an equivalence, since every V-homomorphism between distributive sup-semilattices with zero extends uniquely to a strong V-homomorphism between the respective ideal lattices and on the other hand every strong V-homomorphism between algebraic Heyting algebras induces a V-homomorphism between the respective semilattices of compact elements.

Lemma 2.2. *Let H_1 and H_2 be complete Heyting algebras.*

(i) *If $h: H_1 \to H_2$ is a strong V-homomorphism then the upper adjoint $h^+: H_2 \to H_1$ of h exists and h^+ preserves Sups, Infs, and relative pseudocomplements.*

(ii) *Conversely, if $g: H_2 \to H_1$ preserves Sups, Infs, and relative pseudocomplements then the lower adjoint $g^-: H_1 \to H_2$ of g exists and g^- is a strong V-homomorphism.*

Proof. (i). Since h preserves Sups, the upper adjoint $g=h^+$ exists and preserves Infs (see [5; Corollary 0-3.5]). By definition g and h are connected by the equivalence:

$$x \leqq g(y) \quad \text{if and only if} \quad h(x) \leqq y. \tag{1}$$

In particular it follows that

$$x \leqq g\,h(x) \quad \text{and} \quad h\,g(y) \leqq y. \tag{2}$$

In order to see that g preserves Sups, we first note that $h(g(\sum y_i)) \leqq \sum y_i$. Hence there are x_i's with $g(\sum y_i)=\sum x_i$ and $h(x_i)=h(g(\sum y_j))\,y_i \leqq y_i$, i.e. $x_i \leqq g(y_i)$. Consequently, $g(\sum y_i) \leqq \sum g(y_i)$ and therefore $g(\sum y_i)=\sum g(y_i)$, since g is monotone. Next we claim that

$$h(x)\,y = h(x\,g(y)). \tag{3}$$

In fact, "$\geqq$" is obvious. Conversely, since $h(x)y \leqq h(x)$, there exists $x_0 \leqq x$ with $h(x_0) = h(x)y$. Since $h(x_0) \leqq y$, it follows that $x_0 \leqq x\,g(y)$. Thus $h(x_0) \leqq h(x\,g(y))$. Now using (1) and (3), we conclude that the following are equivalent: $x \leqq g(y_1 \Rightarrow y_2)$, $h(x) \leqq y_1 \Rightarrow y_2$, $h(x)\,y_1 = h(x\,g(y_1)) \leqq y_2$, $x\,g(y_1) \leqq g(y_2)$, and $x \leqq g(y_1) \Rightarrow g(y_2)$; where "$\Rightarrow$" denotes the relative pseudocomplement operation. This proves $g(y_1 \Rightarrow y_2) = g(y_1) \Rightarrow g(y_2)$.

(ii) The lower adjoint $h = g^-$ exists and preserves Sups, since g preserves Infs. Again (1) and (2) from above are valid. Assume $h(x) = \sum y_i$. We have to find x_i's with $x = \sum x_i$ and $h(x_i) = y_i$. Indeed, we claim that the elements $x_i = x\,g(y_i)$ have the required property. First we show that

$$v \leqq h(u) \quad \text{implies} \quad h(u\,g(v)) = v. \tag{4}$$

Abbreviate $w = h(u\,g(v))$. By (2), we have $w \leqq h\,g(v) \leqq v$. Therefore it remains to verify $v \leqq w$: in view of $u\,g(v) \leqq g(w)$ it follows that $u \leqq g(v) \Rightarrow g(w) = g(v \Rightarrow w)$. Hence $v \leqq h(u) \leqq v \Rightarrow w$, i.e. $v \leqq w$.

As $y_i \leqq h(x)$, we infer from (4) that $h(x_i) = y_i$. Further $x \leqq g\,h(x) = g(\sum y_i) = \sum g(y_i)$, thus $x = \sum x\,g(y_i) = \sum x_i$.

Theorem 2.3. *The categories $\mathcal{AlgHeyt}$ and $\mathcal{AlgHeyt}_0$ are dually equivalent.*

Proof. The assertion immediately follows from Lemma 2.2.

Theorem 2.4. *Let H be a compact coherent frame (see e.g. [1]), i.e., H is an algebraic Heyting algebra in which the unit is compact and infs of compact elements are always compact. Then H can be embedded in the ideal lattice of some atomless Boolean algebra under a map preserving Sups, Infs, and relative pseudocomplements.*

Proof. According to the duality of $\mathcal{DistrSemilattice}$ and $\mathcal{AlgHeyt}_0$, the assertion is just the translated version of Corollary 1.7.(i). (Explicitly, if $h\colon B \to L$ is an onto V-homomorphism then the associated embedding $\varphi\colon \mathrm{Id}(L) \to \mathrm{Id}(B)$ is given by $\varphi(I) = h^{-1}(I)$.)

Corollary 2.5. *Every bounded distributive sup-semilattice can be embedded in the ideal lattice of some atomless Boolean algebra under a mapping which preserves sups, all existing Infs, and all existing relative pseudocomplements.*

Proof. Obviously, the canonical embedding of a bounded distributive sup-semilattice L into its ideal lattice $\mathrm{Id}(L)$ preserves sups and all existing Infs and relative pseudocomplements. Now apply Theorem 2.4 for $H = \mathrm{Id}(L)$.

Corollary 2.6. *Every Heyting algebra can be embedded in the Heyting algebra of all ideals of some atomless Boolean algebra under a map preserving infs, sups, and relative pseudocomplements.*

Lakser has shown that every distributive p-algebra (i.e., bounded distributive lattice with pseudocomplementation) can be embedded in the ideal lattice of some *atomic* Boolean algebra. (Note that Corollary 2.5 yields the same for some atomless Boolean algebra.) The proof given in [6; p. 180] is easy and very elegant. The reason is that one can reduce the assertion to subdirectly irreducible distributive p-algebras, and there are very few of them. Nevertheless such an argument cannot be applied to prove Corollary 2.6, since subdirectly irreducible Heyting algebras are as complicated as Heyting algebras themselves [2; Exercise 8.9].

3. Heyting Algebras, Sober Spaces, and Stone Spaces

A Heyting algebra is called *spatial* [1] if it is complete and each element is an Inf of prime elements. It is well-known [5; V-4.7.(ii)] that the category of spatial Heyting algebras with maps preserving infs and Sups is dual to the category of sober spaces (i.e., T_0-spaces in which the non-empty $\cup$-irreducible closed sets are precisely the point closures) with continuous functions. In the following we shall consider other morphisms which are important from our present point of view. Let $\mathcal{SpaHeyt}$ (resp., $\mathcal{SpaHeyt}_0$) be the category of spatial Heyting algebras and strong V-homomorphisms (resp., maps preserving Sups, Infs, and relative pseudocomplements).

By $\mathsf{O}(X)$ we denote the lattice of open subsets of a space X. Of course, $\mathsf{O}(X)$ is a spatial Heyting algebra. The *specialization order* on X is defined by setting $x \leqq y$ if and only if $\{x\}^- \subseteq \{y\}^-$. Let us call a subset U of X *almost open* (cf. [7]) if $\uparrow U$ is open and U is a *strict* subset of $\uparrow U$ (that is, for all open $V, W \subseteq \uparrow U$, $U \cap V = U \cap W$ implies $V = W$). A continuous function $f: X \to Y$ is *almost open* if f maps almost open subsets (or equivalently, open subsets) of X onto almost open subsets of Y. Note that, for example, *sobrification maps* (see [5; p. 255]) are almost open. In particular, each T_0-space lies almost open in its sobrification (if it is considered as a subspace). Let $\mathcal{Sober}$ be the category of sober spaces with continuous and almost open mappings.

Lemma 3.1. *Suppose that $f: X \to Y$ is continuous. Then the mapping*

$$\mathsf{O}(f): \mathsf{O}(Y) \to \mathsf{O}(X), \qquad V \mapsto f^{-1}(V)$$

preserves Infs and relative pseudocomplements if and only if f is almost open.

Proof. If $g = \mathsf{O}(f)$ has a lower adjoint $h: \mathsf{O}(X) \to \mathsf{O}(Y)$ (i.e., $h(U) \subseteq V$ if and only if $f[U] \subseteq V$) then, as implicitly shown in the proof of Lemma 2.2,

(*) g preserves Infs and relative pseudocomplements iff

$$h(U) \cap V = h(U \cap g(V)) \qquad \text{for all } U \in \mathsf{O}(X), \quad V \in \mathsf{O}(Y).$$

"$\Rightarrow$" First, g has a lower adjoint h, since g preserves Infs. It is easy to see that $h(U) = \uparrow f[U]$. To show that $f[U]$ is a strict subset of $h(U)$, assume $f[U] \cap V = f[U] \cap W$ for open $V, W \subseteq h(U)$. Then $f[U \cap g(V)] = f[U] \cap V = f[U] \cap W = f[U \cap g(W)]$, and therefore $h(U \cap g(V)) = h(U \cap g(W))$. Thus by (*), we have $V = W$. "$\Leftarrow$" Because f is almost open, we can define the map

$$h: \mathsf{O}(X) \to \mathsf{O}(Y), \qquad U \mapsto \uparrow f[U].$$

Obviously, h is the lower adjoint of g, and h preserves Sups. Hence g is Inf-preserving. Moreover, since $f[U] \cap V = f[U \cap g(V)]$, it follows that $h(U) \cap V = h(U \cap g(V))$. In view of (*), this proves the assertion.

Theorem 3.2. (cf. [7]). *The categories $\mathcal{Sober}$, $\mathcal{SpaHeyt}$, and $\mathcal{SpaHeyt}_0^{\mathrm{op}}$ are equivalent.*

Proof. From Lemma 3.1 and [5; V-4.7.(ii)] it follows that the equivalence of $\mathcal{Sober}$ and $\mathcal{SpaHeyt}_0^{\mathrm{op}}$ is established by the functor

$$\mathsf{O}: \begin{cases} \mathcal{Sober} \to \mathcal{SpaHeyt}_0^{\mathrm{op}} \\ X \mapsto \mathsf{O}(X), \quad f \mapsto \mathsf{O}(f). \end{cases}$$

Further, $\mathcal{SpaHeyt}$ and $\mathcal{SpaHeyt}_0$ are dual by Lemma 2.2.

A sober space is called a *Stone space* if its system of compact open subsets forms an open base. Let $\mathcal{Stone}$ be the full subcategory of $\mathcal{Sober}$ whose objects are the Stone spaces. It is well-known that the functor O gives a dual equivalence between the category of Stone spaces with continuous mappings and the category of algebraic Heyting algebras with maps preserving infs and Sups. Thus we conclude from Lemma 3.1 that $\mathcal{Stone}$ and $\mathcal{AlgHeyt}_0$ are dual categories. Together with the Theorems 2.1 and 2.3 we can summarize:

Theorem 3.3. *The categories $\mathcal{DistrSemilattice}$, $\mathcal{AlgHeyt}$. $\mathcal{AlgHeyt}_0^{\mathrm{op}}$, and $\mathcal{Stone}$ are equivalent.*

Let us take a closer look at the equivalence of $\mathcal{Stone}$ and $\mathcal{DistrSemilattice}$. We shall define mutually inverse functors

$$\mathsf{Comp}: \mathcal{Stone} \to \mathcal{DistrSemilattice} \quad \text{and} \quad \mathsf{Prim}: \mathcal{DistrSemilattice} \to \mathcal{Stone}$$

which implement this equivalence directly. If X is a Stone space then let $\mathsf{Comp}(X)$ be the semilattice, under union, of all compact open subsets of X. For $f \in \mathrm{mor}(X, Y)$ $(X, Y \in \mathrm{obj}(\mathcal{Stone}))$ define

$$\mathsf{Comp}(f): \mathsf{Comp}(X) \to \mathsf{Comp}(Y), \quad C \mapsto \uparrow f[C].$$

Conversely, if L is a distributive sup-semilattice with zero then let $\mathsf{Prim}(L)$ be the prime filter space of L. Recall that the sets $\hat{c} = \{P \in \mathsf{Prim}(L): c \in P\}$ $(c \in L)$ form a base. For $h \in \mathrm{mor}(L, K)$ $(L, K \in \mathrm{obj}(\mathcal{DistrSemilattice}))$ set

$$\mathsf{Prim}(h): \mathsf{Prim}(L) \to \mathsf{Prim}(K), \quad P \mapsto \uparrow h[P].$$

All continuous open functions between sober spaces are morphisms in the category $\mathcal{Sober}$. Probably the converse is not true. Let us, however, mention a positive result in this direction:

Proposition 3.4. *Suppose that X and Y are Stone spaces, Y is first countable, and $f: X \to Y$ is continuous and almost open. Then f is open.*

Proof. We may assume that $X = \mathsf{Prim}(L)$, $Y = \mathsf{Prim}(K)$, and $f = \mathsf{Prim}(h)$ for some $\vee$-homomorphism $h: L \to K$ between distributive sup-semilattices with zero (that is, we have $f(P) = \uparrow h[P]$ for all $P \in \mathsf{Prim}(L)$). We claim that $f[\hat{a}] = \widehat{h(a)}$ $(a \in L)$. Since $f[\hat{a}] \subseteq \widehat{h(a)}$ is obvious, suppose $Q \in \widehat{h(a)}$ in order to prove the converse inclusion. Q is generated (as an upper set) by a descending chaim $h(a) = b_0 \geqq b_1 \geqq \ldots$, because Y is first countable. Define the ideal $I = h^{-1}(K - Q)$. Using the fact that h is a $\vee$-homomorphism, we find a descending chain $a = a_0 \geqq a_1 \geqq \ldots$ in $L - I$ such that $h(a_i) = b_i$. Choose $P \in X$ with the property that $a_i \in P$ $(i < \omega)$ and $P \cap I = \varnothing$, then $f(P) = Q$.

In topological terms, the question of whether every distributive sup-semilattice with zero is the V-homomorphic image of some generalized Boolean lattice (see Sect. 2) can be sharpened as follows: *Is every Stone space the continuous open image of some locally compact zero-dimensional Hausdorff space?* A partial answer is given by the following theorem:

Theorem 3.5. *Second countable compact Stone spaces are precisely the T_0-images of the Cantor set under continuous open functions.*

Proof. Let Y be a second countable compact Stone space. Without loss of generality, $Y = \mathrm{Prim}(L)$ for some countable bounded distributive semilattice L. Using the measurability of countable refinement monoids [3; Theorem 3.4, Lemma 3.6] and the argument of the proof of Corollary 1.7.(i), we find an atomless countable Boolean algebra B and an onto V-homomorphism $h: B \to L$. By Proposition 3.4 and Theorem 3.3, $f = \mathrm{Prim}(h)$ is a continuous open function from $X = \mathrm{Prim}(B)$ in Y. Note that X is a second countable perfect Boolean space, and therefore X is homeomorphic to the Cantor set. Further, we have $\widehat{h(a)} = f[\hat{a}]$ for all $a \in L$ (see proof of Proposition 3.4), in particular $f[X] = f[\hat{1}] = \widehat{h(1)} = \hat{1} = Y$.

The converse is a consequence of the following fact: if X is a second countable Stone Space and Y is a T_0-space which is the image of X under a continuous open mapping f, then Y is also a second countable Stone space. Indeed, let $s: Y \to Y^s$ be the sobrification map of Y. It is easy to see that Y has countably many compact open subsets and that these form a base; thus the same is true for Y^s. Moreover, s is continuous and almost open. Hence $sf: X \to Y^s$ is continuous and almost open too, and Proposition 3.4 shows that sf is in fact open (since Y^s is a Stone space). In particular, we infer $s[Y] = sf[X] = \uparrow sf[X] = \uparrow s[Y] = Y^s$, i.e., s is onto, which means that Y is sober. This completes the proof.

Acknowledgement. The author would like to thank Karl Heinrich Hofmann and Peter T. Johnstone for their valuable comments and suggestions concerning this paper.

References

1. Banaschewski, B.: Coherent frames. In: Continuous lattices, Proceedings of the Conference on Topological and Categorical Aspects of Continuous Lattices (Workshop IV), Bremen 1979 (B. Banaschewski, R.-E. Hoffmann eds.). Lecture Notes in Math. **871**, pp. 1–11. Berlin-Heidelberg-New York: Springer 1981
2. Burris, S., Sankappanavar, H.P.: A course in universal algebra. Berlin-Heidelberg-New York: Springer 1981
3. Dobbertin, H.: Refinement monoids, Vaught monoids, and Boolean algebras. Math. Ann. **265**, 473–487 (1983)
4. Dobbertin, H.: Primely generated regular refinement monoids. J. Algebra (to appear)
5. Gierz, G., Hofmann, K.H., Keimel, K., Lawson, J.D., Mislove, M., Scott, D.S.: A compendium of continuous lattices. Berlin-Heidelberg-New York: Springer 1980
6. Grätzer, G.: Lattice theory: First concepts and distributive lattices. San Francisco: Freeman 1971
7. Johnstone, P.T.: Open maps of toposes. Manuscripta Math. **31**, 217–247 (1980)
8. Johnstone, P.T.: The point of pointless topology. Bull. Amer. Math. Soc. **8**, 41–53 (1983)

Received August 25, 1983; received in final form January 7, 1984

Math. Z. 187, 23–24 (1984)

Mathematische Zeitschrift

Multipliers of AW^*-Algebras

Gert K. Pedersen

Mathematics Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

Given a unital C^*-algebra B, when is it true that for each (non-unital, non-degenerate) C^*-subalgebra A of B, the idealizer of A in B coincides with the set of double centralizers of A? To formulate the problem more precisely, recall from [3] (or [5, 3.12]) that a double centralizer is a pair (ρ_r, ρ_l) of (necessarily) bounded linear operators on A such that

$$\rho_r(x\,y) = x\,\rho_r(y), \qquad \rho_l(x\,y) = \rho_l(x)\,y, \qquad \rho_r(x)\,y = x\,\rho_l(y)$$

for all x, y in A. Each element in the idealizer

$$M = \{b \in B \mid bA + Ab \subset A\}$$

determines a double centralizer (ρ_r, ρ_l), viz. $\rho_r(x) = x\,b$, $\rho_l(x) = b\,x$, and we ask whether this map $b \to (\rho_r, \rho_l)$ is surjective. If B is a von Neumann algebra the affirmative answer is well-known, see e.g. [5, 3.12.3], and it can also easily be established for C^*-algebras that are monotone complete (in the sense that each bounded, monotone increasing net of self-adjoint elements has a least upper bound in the algebra). In [6] Reid noticed that an affirmative answer (for every $A \subset B$) implies that B is necessarily an AW^*-algebra (which by definition means that every maximal abelian C^*-subalgebra of B is monotone complete, see [1, § 7]), and Johnson then proved in [4] that the AW^*-condition suffices. We present below a simple proof of Johnson's result. Since his theorem immediately gives the main result in [2], a revival of this circle of ideas may not be amiss. We shall use the monographs [1] and [5] as references for the theories of AW^*-algebras and C^*-algebras.

Theorem. *Let A be a C^*-subalgebra of an AW^*-algebra B. Assume further that the annihilator of A in B is zero. Then the set M of (two-sided) multipliers of A in B is isometrically *-isomorphic to the set of double centralizers of A (i.e. the set $M(A)$ of multipliers of A in its enveloping von Neumann algebra A'') via an isomorphism that extends the identity map on A.*

Proof. Identifying multipliers of A with double centralizers, cf. [5, 3.12.3], we obtain an isometric *-isomorphism π of M into $M(A)$ that extends the identity map on A. It only remains to show that π is surjective.

Assume first that p is a projection in $M(A)$. Thus $L=pA$ and $K=(1-p)A$ are closed right ideals with $L^*K=0$ and $L+K=A$. Let (x_λ) and (y_μ) be approximate units for $L\cap L^*$ and $K\cap K^*$, respectively, [5, 1.4.2]. Thus $x_\lambda y_\mu=0$ for all λ and μ. Working now inside the AW^*-algebra B we note that the range projections $[x_\lambda]$ and $[y_\mu]$ exist in B and are pairwise orthogonal for all λ and μ. Since the projections in B form a complete lattice, we have $q=\bigvee[x_\lambda]$ and $r=\bigvee[y_\mu]$ in B, with $qr=0$. As $x_\lambda\leqq q$ for every λ it follows that q is a left unit for L. Similarly r is a left unit for K. Since $A=L+K$, we see that $1-(q+r)$ annihilates A, whence $q+r=1$ by assumption. Furthermore, the decomposition $A=L+K$ shows that q is a left multiplier of A. Since $q=q^*$ it follows that $q\in M$. Now $qA=L$ and $(1-q)A=K$ and we conclude that $\pi(q)=p$.

To handle the general case we consider the 2×2-matrix algebras

$$\tilde{A}=A\otimes\mathbb{M}_2,\qquad \tilde{M}=M\otimes\mathbb{M}_2,\qquad \tilde{B}=B\otimes\mathbb{M}_2.$$

By a non-trivial result of Berberian, cf. [1, §62], $\tilde{B}$ is an AW^*-algebra. Note that $M(\tilde{A})=M(A)\otimes\mathbb{M}_2$ and that $\tilde{M}$ is the set of multipliers of $\tilde{A}$ in $\tilde{B}$. Furthermore, the map $\tilde{\pi}\colon \tilde{M}\to M(\tilde{A})$ given by $(\tilde{\pi}(x))_{ij}=\pi(x_{ij})$, for $1\leqq i,j\leqq 2$, is the canonical isometry from $\tilde{M}$ into $M(\tilde{A})$ mentioned before. For each x in $M(A)$ with $0\leqq x\leqq 1$ we define the projection p in $M(A)$ by

$$p=\begin{pmatrix} x & (x-x^2)^{\frac{1}{2}} \\ (x-x^2)^{\frac{1}{2}} & 1-x \end{pmatrix}.$$

As the first part of the proof showed, there is then a projection q in $\tilde{M}$ such that $\tilde{\pi}(q)=p$. Thus with $y=q_{11}$ we have an element in M with $0\leqq y\leqq 1$, such that $\pi(y)=x$. Consequently $\pi(M_+)=M(A)_+$, whence π is surjective as claimed.

Corollary 1. *Let A be an essential, closed ideal in an AW^*-algebra M. Then $M=M(A)$.*

Corollary 2 [2]. *Each isomorphism $\pi\colon A_1\to A_2$ between closed essential ideals of AW^*-algebras M_1 and M_2, respectively, extends to an isomorphism $\pi\colon M_1\to M_2$.*

References

1. Berberian, S.K.: Baer *-Rings. Grundlehren Math. Wiss. **195**. Berlin-Heidelberg-New York: Springer 1972
2. Berberian, S.K.: Isomorphisms of large ideals of AW^*-algebras. Math. Z. **182**, 81–86 (1983)
3. Johnson, B.E.: An introduction to the theory of centralizers. Proc. London Math. Soc. **14**, 299–320 (1964)
4. Johnson, B.E.: AW^*-algebras are QW^*-algebras. Pacific J. Math. **23**, 97–99 (1967)
5. Pedersen, G.K.: C^*-Algebras and their Automorphism Groups. London Math. Soc. Monographs **14**. London-New York: Academic Press 1979
6. Reid, G.A.: A generalisation of W^*-algebras. Pacific J. Math. **15**, 1019–1026 (1965)

Received October 4, 1983; received in final form December 23, 1983

Math. Z. 187, 25–27 (1984)

Mathematische Zeitschrift

The p-adic Topology on a Free Group: A Counterexample

James Howie*

University of Glasgow, Glasgow G12 8QW, Great Britain

In his paper [1], Stallings formulated several inter-related families of conjectures about properties of free groups and about equations over groups. The purpose of this note is to describe a counterexample to some of these conjectures, specifically concerning the p-adic topology on a free group.

I am grateful to John Stallings for removing some errors from my arguments.

1. Notation

We will use (a subset of) the notation of [1], which we nevertheless recall here for the convenience of the reader. Let p be either a prime number or 0. The *p-central series* of a group G is defined inductively by ${}_pG_1=G$ and, for $i\geqq 1$, by ${}_pG_{i+1}=({}_pG_i)^p\cdot[{}_pG_i, G]$ (the subgroup generated by elements of the form $x^p[y,z]$, $x, y\in{}_pG_i$, $z\in G$). In the case $p=0$ this is just the lower central series of G. The *p-adic topology* on G is that defined by taking the sets ${}_pG_i$ to be a basis of neighborhoods of 1. We are interested in the case of a free group F of finite rank. Then $\bigcap_i {}_pF_i=\{1\}$, so the p-adic topology on F is metrizable in a natural way. The *p-adic completion* of F is its completion with respect to this metric, namely the profinite group ${}_p\hat{F}=\lim\operatorname{inv}(F/{}_pF_i)$.

If g is an element of a group G, then $[g]^G$ denotes the conjugacy class of g in G. If $g_1,\dots,g_n\in G$ then $[g_1]^G\dots[g_n]^G$ denote the subset of G consisting of elements of the form $a_1\dots a_n$ with $a_i\in[g_i]^G$. Such a set will be called a *product of n conjugacy classes*. Clearly any product of n conjugacy classes can be regarded as a product of $(n+1)$ conjugacy classes by setting $g_{n+1}=1$. The normal closure of a subset X in a group G is denoted $\langle X\rangle_G$. A subgroup H of G is *normal-convex* in G if $\langle N\rangle_G\cap H=N$ for every normal subgroup N of H.

* The author is an SERC Advanced Fellow

2. The Conjectures

We recall two of the classes of conjectures from [1].

Conjecture $A_n(p)$. *In a finitely generated free group F, any product of n conjugacy classes is a closed set in the p-adic topology.*

Conjecture $C(p)$. *If F is a finitely generated free group, then F is normal-convex in its p-adic completion ${}_p\hat{F}$.*

Among other implication relations proved in [1] is the relation (for all n $A_n(p)) \Rightarrow C(p)$. It is also clear that $A_n(p) \Rightarrow A_m(p)$ whenever $m \leqq n$. Conjecture $A_1(p)$ is true ([1], Theorem 3.6), but we will see below that Conjectures $C(p)$ and $A_2(p)$ (and hence also $A_n(p)$ for all $n \geqq 2$) are false.

3. The Counterexample

Our counterexample arises from the following result.

Proposition. *Let $a, b_1, \ldots, b_n$ be elements of a free group F such that $b_1, \ldots, b_n$ generate F $\operatorname{mod}[F, F]$, and such that*

$$a \equiv b_1 \ldots b_n \operatorname{mod}[F, F].$$

Then there exist $g_1, \ldots, g_n \in {}_0\hat{F}$ such that

$$a = (g_1^{-1} b_1 g_1) \ldots (g_n^{-1} b_n g_n) \quad \text{in } {}_0\hat{F}.$$

Corollary. *Let F, a, $b_1, \ldots, b_n$ be as in the proposition, and let p be a prime number. Then there exist $g_1, \ldots, g_n \in {}_p\hat{F}$ such that*

$$a = (g_1^{-1} b_1 g_1) \ldots (g_n^{-1} b_n g_n) \quad \text{in } {}_p\hat{F}.$$

Proof. One checks that ${}_0\hat{F}$ is naturally embedded as a subgroup of ${}_p\hat{F}$ and applies the proposition.

To get the counterexample, apply the proposition (if $p=0$) or the corollary (if $p>0$) to the case $F=\langle x, y\rangle$, $n=2$, $a=xy^2$, $b_1=x^{-2}y^{-3}$, $b_2=x^{-2}(xy)^5$. Then $a \in \langle N\rangle_{{}_p\hat{F}} \cap F$, where $N=\langle\{b_1, b_2\}\rangle_F$. However, the permutation representation $x \mapsto (12)(34)$, $y \mapsto (135)$ sends b_1 and b_2 to 1, but sends a to (12534), so $a \notin N$. This shows that F is not normal-convex in ${}_p\hat{F}$, contrary to Conjecture $C(p)$.

Furthermore, it follows from the proof of (5.4) of [1] that this example also contradicts Conjecture $A_2(p)$. Explicitely, it shows that a must lie in the p-adic closure of the set $X=[b_1]^G[b_2]^G$. But clearly $a \notin X$, since $X \subset N$.

Remark. In the above example F/N is the binary icosahedral group of order 120. Similar counterexamples to Conjecture $C(p)$ can be constructed, beginning from any finite presentation of any nontrivial perfect group.

4. Proof of Proposition

It is sufficient to find, for each $i=1,\dots,n$, a sequence $(g_{ir})_{(r\geqq 1)}$ of elements of F such that, for each i and r:

$$g_{ir}\equiv g_{i(r+1)} \bmod {}_0F_r; \quad \text{and}$$
$$a\equiv(g_{1r}^{-1}b_1 g_{1r})\dots(g_{nr}^{-1}b_n g_{nr}) \bmod {}_0F_{r+1}.$$

For then each sequence (g_{ir}) is Cauchy, and so has a unique limit g_i in ${}_0\hat{F}$, and

$$a=(g_1^{-1}b_1 g_1)\dots(g_n^{-1}b_n g_n) \quad \text{in } {}_0\hat{F},$$

as required.

We define such sequences simultaneously, by induction on r, beginning with $g_{11}=\dots=g_{n1}=1$. Suppose then that $r\geqq 2$ and that g_{ij} have been suitably defined for $1\leqq i\leqq n$ and for $1\leqq j\leqq r-1$. In particular

$$a=(g_{1(r-1)}^{-1}b_1 g_{1(r-1)})\dots(g_{n(r-1)}^{-1}b_n g_{n(r-1)})\,c$$

for some $c\in {}_0F_r$. Since F is generated mod$[F,F]$ by the b_i, it follows that ${}_0F_r$ is generated mod ${}_0F_{r+1}$ by elements of the form $[b_i,w]=b_i^{-1}w^{-1}b_i w$ with $w\in {}_0F_{r-1}$. Using standard commutator identities, we can express c in the form

$$c=[b_1,w_1]\dots[b_n,w_n]\,d$$

for some $w_1,\dots,w_n\in {}_0F_{r-1}$ and some $d\in {}_0F_{r+1}$. Using the fact that $[b_i,w_i]$ is central mod ${}_0F_{r+1}$, and computing mod ${}_0F_{r+1}$, we see that

$$\begin{aligned}a&\equiv(g_{1(r-1)}^{-1}b_1 g_{1(r-1)})\dots(g_{n(r-1)}^{-1}b_n g_{n(r-1)})\,[b_1,w_1]\dots[b_n,w_n]\\&\equiv(g_{1(r-1)}^{-1}b_1[b_1,w_1]\,g_{1(r-1)})\dots(g_{n(r-1)}^{-1}b_n[b_n,w_n]\,g_{n(r-1)})\\&=(g_{1(r-1)}^{-1}w_1^{-1}b_1 w_1 g_{1(r-1)})\dots(g_{n(r-1)}^{-1}w_n^{-1}b_n w_n g_{n(r-1)}).\end{aligned}$$

Since also $w_i\equiv 1 \bmod {}_0F_{r-1}$, we may now define

$$g_{ir}=w_i\,g_{i(r-1)}.$$

Reference

1. Stallings, J.R.: Surfaces in three-manifolds and non-singular equations in groups. Math. Z. **184**, 1–17 (1983)

Received October 7, 1983

Math. Z. 187, 29–47 (1984)

Mathematische Zeitschrift

Group Actions with Inequivalent Representations at Fixed Points

Krzysztof Pawałowski

Instytut Matematyki UAM, Matejki 48/49, 60-769 Poznań, Poland

Let G be a compact Lie group and let M be a smooth manifold. Then any smooth action of G on M induces a linear representation of G on the tangent space $T_x(M)$ at any point $x \in M$ left fixed by G. If x and y lie in the same connected component of the fixed point set M^G, then the representations of G on $T_x(M)$ and $T_y(M)$ are equivalent.

Let $\mathscr{P}$ be the class of compact Lie groups G whose identity connected component G_0 is abelian and the quotient group G/G_0 has prime power order. It follows from Smith Theory that for each $G \in \mathscr{P}$ acting on an acyclic manifold M, M^G is non-empty and connected. Thus, if the action is smooth, the representations of G at any two fixed points are equivalent.

For smooth actions of a compact Lie group G on spheres, disks or euclidean spaces, the question of the equivalence of the representations of G at any two fixed points and the weaker question of the equality of the dimensions of any two fixed point set connected components go back to Smith ([24], the footnote on p. 406), Hsiang and Hsiang ([9], Problem 16), and Bredon ([3], the second remark on p. 58).

Theorem A of this paper shows for which compact Lie groups G these questions have the affirmative answers for actions on disks or euclidean spaces.

First, denote by $\mathscr{R}$ the class of compact Lie groups G such that each element of G/G_0 has prime power order (here G_0 need not be abelian). Clearly, $\mathscr{R} \supset \mathscr{P}$ and $\mathscr{R} - \mathscr{P}$ includes such groups as the symmetric groups S_3 and S_4, the alternating groups A_5 and A_6, the dihedral groups D_{p^a} of order $2p^a$, p odd prime, $a \geqq 1$, and the classical groups $SO(n)$ and $O(n)$, $n \geqq 3$.

Now, the announced theorem reads as follows.

Theorem A. *Let G be a compact Lie group. Then the following three conditions are equivalent.*

(1) *For any smooth action of G on a disk (resp., euclidean space) at any two fixed points the representations of G are equivalent.*

(2) *For any smooth action of G on a disk (resp., euclidean space) each fixed point set connected component has the same dimension.*

(3) $G \in \mathscr{R}$.

In Theorem A, (1) implies (2) by the Slice Theorem. In order to show that (2) implies (3), for any $G \notin \mathscr{R}$, we construct smooth actions of G on disks and euclidean spaces (as well as on spheres) with fixed point set connected components of different dimensions (see Example 6.1). According to the author's knowledge these are the first examples of such smooth actions. Finally, (3) implies (1) because for any smooth action of G on an acyclic manifold, at any two fixed points the representations of G are equivalent whenever $G \in \mathscr{R}$ (see Proposition 7.1). Since the last result holds also for actions on homology spheres with at least three fixed points (see Proposition 7.2), Theorem A remains true when we replace disks by spheres or homology spheres with at least three fixed points.

As just mentioned, a compact Lie group G has a smooth action on a disk (resp., euclidean space, sphere) with fixed point set connected components of different dimensions and thus with inequivalent representations of G at fixed points in different connected components whenever $G \notin \mathscr{R}$, i.e., G/G_0 has a cyclic subgroup of order pq for some two distinct primes p and q. For example, this holds for $G = \mathbb{Z}_r$, the cyclic group of order r, provided r is not a prime power. However, when constructing smooth group actions on acyclic manifolds or homology spheres with inequivalent representations at fixed points, the cases of isolated fixed points are especially interesting. Cases already known include smooth G-manifolds M with exactly two fixed points at which the representations of G are inequivalent under the following conditions.

1. $G = \mathbb{Z}_r$, r not a prime power, and M is some euclidean space (see Edmonds and Lee [8]).

2. G is a finite odd order abelian group with at least four noncyclic Sylow subgroups and M is some homotopy sphere (see Petrie [18] and Dovermann and Petrie [7]).

3. $G = \mathbb{Z}_{2^a q}$ with $a \geqq 2$ and odd $q \geqq 3$, and M is the n-dimensional sphere with odd $n \geqq 9$. Also, $G = \mathbb{Z}_{4q}$ with $q \geqq 2$, and M is some 9-dimensional homotopy sphere (see Cappell and Shaneson [4, 5]).

The results under the conditions 1 and 2 both can be generalized by replacing two fixed points by any finite number of fixed points (see Edmonds and Lee [8], Remark in § 6 and Petrie [19], Theorem A).

However, if $\mathbb{Z}_r$ acts smoothly on a disk with finite fixed point set F, then F is just one point (see Corollary 2.3). Thus, for smooth $\mathbb{Z}_r$-actions on spheres, more than two isolated fixed points cannot occur.

More generally, if a compact Lie group G acts smoothly on a disk with finite fixed point set, then there is the following restriction on the number k of fixed points. For $G \in \mathscr{P}$, $k = 1$ by Smith Theory. For $G \notin \mathscr{P}$, k is congruent to 1 $\operatorname{mod} n_G$, where n_G is the number defined by Oliver (see Theorem 2.2 and Corollary 2.3). Beside this restriction on k, there are two obvious restrictions on the representations of G at fixed points. The first is that each representation has no trivial summand by the Slice Theorem, and the second is that any two representations are equivalent when restricted to any $P \in \mathscr{P}(G)$, the family of all subgroups of G which are in $\mathscr{P}$, by Smith Theory.

We show, for some finite groups G, that these restrictions are stably necessary and sufficient. Namely, the following theorem holds.

Theorem B. *Assume G is an extension of the form*

$$0 \to \mathbb{Z}_r \to G \to \mathbb{Z}_s \to 0$$

where r and s are relatively prime and s is a product of distinct primes. Let $V_1, \ldots, V_k$ be complex representations of G. Then, there is a smooth action of G on a disk with exactly k fixed points at which the representations of G are equivalent to $V_1 \oplus W, \ldots, V_k \oplus W$ for some complex representation W of G, if and only if the following three conditions hold.

(a) $k \equiv 1 \pmod{n_G}$.

(b) $V_i^G = 0$ *for* $i = 1, \ldots, k$.

(c) $\operatorname{Res}_P(V_i) \cong \operatorname{Res}_P(V_j)$ *for each* $P \in \mathscr{P}(G)$, $1 \leq i, j \leq k$.

The necessity of these conditions has just been discussed. The sufficiency is proved in Theorem 7.4 and in Example 7.5, for any integer $k > 1$, pairwise inequivalent representations $V_1, \ldots, V_k$ of G fulfilling the conditions (b) and (c) are constructed whenever $n_G \neq 0$. Thus, the following corollary holds.

Corollary C. *Let G be as in Theorem B and assume further $n_G \neq 0$. Then, for each integer $k > 1$ with $k \equiv 1 \pmod{n_G}$, there is a smooth action of G on a disk with exactly k fixed points and inequivalent representations of G at different fixed points.*

For example, take $G = D_r$ with r odd and not a prime power. Then G is an extension of the required form and $n_G = 2$ (see Proposition 2.4).

This paper is divided into seven sections. In § 1 we give some needed notations and results from equivariant topology. In § 2 we briefly recall Oliver Construction producing finite contractible G-CW-complexes with a given fixed point set. In § 3 we describe an equivariant thickening procedure used for converting G-CW-complexes into smooth G-manifolds. This procedure requires the existence of a suitable G-vector bundle over a given G-CW-complex. In § 4 we study G-spaces useful for constructing G-vector bundles. In § 5 we obtain a fixed point set realization theorem which enables us to give examples of smooth "exotic" actions. In § 6 we construct such examples and we answer some problems in compact transformation groups. In § 7 we deal with representations at fixed points of smooth actions on acyclic manifolds and homology spheres.

Acknowledgments. Most of the results of this paper are contained in my Ph.D. thesis [16] and it is a pleasure for me to thank to Professor Andrzej Białynicki-Birula for his encouragement and the help that he has given me during my work on the thesis.

I would like to express my sincere thanks to Bob Oliver for his invaluable advice, stimulating conversations, and continuing support in my research. Without his understanding of the subject and clarifying it to me, I would have not been able to write this paper.

I am also indebted to Stefan Jackowski and Ryszard Rubinsztein for helpful conversations and all of the help I have obtained from them.

§ 1. Preliminaries

Let $\rho\colon G \to GL(V)$ be a representation of a compact Lie group G on an n-dimensional complex (resp., real) vector space V. Then ρ is equivalent to a unitary (resp., orthogonal) representation of G on $\mathbb{C}^n$ (resp., $\mathbb{R}^n$). Moreover, ρ decomposes uniquely into the direct sum of irreducible representations. In particular, the vector space V with the linear action determined by ρ, called also a representation of G, decomposes uniquely into the direct sum of the trivial summand V^G and the nontrivial summand $(V^G)^\perp$. For any unitary representation $\rho\colon G\to U(n)$, $\mathbb{C}^n(\rho)$ denotes the complex vector space $\mathbb{C}^n$ with the linear G-action $G\times\mathbb{C}^n\to\mathbb{C}^n$, $(g,v)\mapsto\rho(g)\cdot v$.

Consider $\mathbb{Z}_r$ as the subgroup of S^1, the group of complex numbers of absolute value 1, generated by the primitive r-th root of unity. For any integer k, t^k denotes the representation $\mathbb{Z}_r\to U(1)$, $z\mapsto z^k$. We refer to $t=t^1$ as to the *standard* 1-dimensional complex representation of $\mathbb{Z}_r$ and we put $t^0=1$ provided no confusion is possible. The representations $1,t,\dots,t^{r-1}$ form a complete set of irreducible complex representations of $\mathbb{Z}_r$.

Let X be a G-space. Then, the isotropy subgroups G_x occuring at points $x\in X$ form a family of subgroups of G closed under conjugation. For a closed subgroup H of G, we use the following standard notations:

$$F(H,X)=X^H=\{x\in X\,|\,G_x\supseteq H\}$$
$$X_H=\{x\in X\,|\,G_x=H\}$$
$$X_{(H)}=\bigcup_{g\in G}X_{gHg^{-1}}.$$

Since the orbit type of a given orbit $G(x)$ is determined by the conjugacy class (G_x), thus $X_{(H)}$ is the sum of all orbits in X of type (H).

If a G-vector bundle ξ over X is given an invariant inner product, i.e., ξ is an orthogonal G-vector bundle over X, then for any $x\in X^G$, the fiber ξ_x of ξ over x is an orthogonal representation of G and the bundle $\xi|X^G$ decomposes uniquely into the Whitney sum of the fixed point bundle ξ^G over X^G (the trivial summand) and the nontrivial summand $(\xi^G)^\perp$.

For any representation V of G, there is the product G-vector bundle ξ over X with the total space $E(\xi)=X\times V$ and the bundle projection $X\times V\to X$, $(x,v)\mapsto x$. We denote this bundle by ε^V and we use the standard notation ε^n when V is just $\mathbb{R}^n$ with the trivial action of G.

Consider a smooth G-manifold M and choose an invariant riemannian metric on M. Then, the tangent bundle τ_M is a smooth orthogonal G-vector bundle over M. In particular, for any $x\in M^G$, the tangent space $T_x(M)$ is an orthogonal representation of G and the Slice Theorem says that the exponential map is an equivariant diffeomorphism of a sufficiently small open disk around 0 in $T_x(M)$ onto an invariant open neighbourhood of x in M.

We shall construct smooth G-manifolds from G-CW-complexes; i.e., G-spaces built up by attaching equivariant cells of the form $G/H\times D^n$ (where G acts trivially on D^n) to lower dimensional skeleta via equivariant maps α defined on $G/H\times S^{n-1}$.

Theorem 1.1 (Illman [11]). *Let G be a compact Lie group. Then, any smooth G-manifold (with or without boundary) has the structure of a G-CW-complex. Moreover, the manifold is compact if and only if the complex is finite.*

We shall use the following definition.

Definition 1.2. Let G be a finite group. Then, a G-CW-complex X is called $\mathscr{P}$-*essential* if for each proper subgroup H of G with $H \notin \mathscr{P}(G)$, each equivariant cell in X of the form $G/H \times D^n$ has the attaching map α such that α is constant on $\{gH\} \times S^{n-1}$, $g \in G$.

§ 2. Oliver Construction

Following Oliver's notation (used for finite groups), $\mathscr{G}^1$ denotes the class of compact Lie groups G such that G_0 is abelian and G/G_0 is "cyclic mod p" for some prime p, i.e., G/G_0 has a normal subgroup P of prime power order, such that $(G/G_0)/P$ is cyclic. Clearly, $\mathscr{P} \subset \mathscr{G}^1$. For $G \in \mathscr{G}^1$, the following Euler characteristic condition on the fixed point sets of finite contractible G-CW-complexes follows from Oliver [13], Proposition 1.

Proposition 2.1. *Assume $G \in \mathscr{G}^1$ and X is a finite contractible G-CW-complex. Then $\chi(X^G)=1$.*

Oliver has shown that for a given compact Lie group $G \notin \mathscr{P}$, the Euler characteristic of a finite CW-complex F is the only obstruction for F to occur as the fixed point set of a finite contractible G-CW-complex. Namely, the following theorem holds.

Theorem 2.2 (Oliver [13, 14]). *Let G be a compact Lie group. If $G \notin \mathscr{P}$ there is an integer n_G such that a finite CW-complex F is the fixed point set of a finite contractible G-CW-complex if and only if $\chi(F) \equiv 1 \pmod{n_G}$.*

Now, Theorems 1.1 and 2.2 and Proposition 2.1 give the following corollary.

Corollary 2.3. *Let G be a compact Lie group acting smoothly on a compact contractible manifold M. Then the following two statements are true.*

(1) *$G \in \mathscr{G}^1$ implies $\chi(M^G)=1$.*
(2) *$G \notin \mathscr{G}^1$ implies $\chi(M^G) \equiv 1 \pmod{n_G}$.*

In Oliver [13], Theorem 2.2 is proved for finite groups and in Oliver [14], the result is shown for nonfinite compact Lie groups, so that $n_G = n_{G/G_0}$ when G_0 is abelian and $n_G=1$ when G_0 is nonabelian. In Oliver [15], the integer n_G is calculated for finite groups G. In particular, $n_G=0$ if and only if $G \in \mathscr{G}^1$. Moreover, the following proposition follows from Oliver [15], Sect. 4.

Proposition 2.4. *Assume G is an extension of the form*

$$0 \to \mathbb{Z}_r \to G \to \mathbb{Z}_s \to 0$$

where r and s are relatively prime. Let $p_1, \ldots, p_m$ be all primes dividing r and let H be the subgroup of $\mathbb{Z}_r$ of order $r/p_1 \ldots p_m$. Then $n_G = n_{G/H}$. Moreover, if $G \notin \mathscr{G}^1$ and s is a prime, then $n_G = s$.

Now, let G be a finite group not of prime power order and let F be a finite CW-complex with $\chi(F) \equiv 1 \pmod{n_G}$. Then, by Theorem 2.2, there is a finite contractible G-CW-complex X with $X^G = F$. According to the construction given by Oliver [13], such X is build up one orbit type at a time, adding cells $G/H \times D^n$ (for various n) until the fixed point set of H has the desired homology. More precisely, the construction is done as follows.

First, cells $G/H \times D^n$ are added, so that to obtain a finite G-CW-complex with fixed point set F, such that the fixed point set of H has the desired Euler characteristic for each $H \neq 0$, and it is $\mathbb{Z}_p$-acyclic for each p-subgroup $H \neq 0$.

Then, free orbits of cells $G \times D^n$ are added, so that to obtain a finite 1-connected G-CW-complex with all of the reduced homology concentrated in one dimension.

Finally, again free orbits of cells are added, so that to obtain a finite contractible G-CW-complex X with $X^G = F$.

We refer to this construction of X as to Oliver Construction (see Oliver [13] for the details).

Remark 2.5. It follows from Oliver Construction that each time the fixed point set of a p-subgroup is made $\mathbb{Z}_p$-acyclic, the dimension need be raised by at most 1. Thus, for any $P \in \mathscr{P}(G)$, we may assume that

$$\dim X^P = \max\{1, \dim F\} + \operatorname{leng}(P),$$

where $\operatorname{leng}(P)$ means the length of the longest chain $P = P_1 \subsetneqq P_2 \subsetneqq \ldots \subsetneqq P_n$ of p-subgroups P_i of G.

Remark 2.6. It also follows from Oliver Construction that for $H \notin \mathscr{P}(G)$, X^H has only the desired Euler characteristic. Thus, we may assume that X is $\mathscr{P}$-essential (see Definition 1.2).

§ 3. Equivariant Thickening

Equivariant thickenings produce open smooth G-manifolds or compact smooth G-manifolds with boundary and thus boundaries and interiors of such G-manifolds. Usually, the idea is to convert a given G-CW-complex into a smooth G-manifold by using a G-vector bundle over the complex (see, e.g., Assadi [1] and Edmonds and Lee [8]).

Let G be a finite group. We describe an equivariant thickening procedure which also enables us to produce smooth G-manifolds from G-CW-complexes and G-vector bundles. More precisely, for a compact smooth G-manifold B and a smooth G-vector bundle ν over B, we are interested in the problem of constructing a compact smooth G-manifold M of a given nonequivariant homotopy type, containing B as a smooth invariant submanifold with equivariant normal bundle ν. Moreover, we are interested in the construction of M such that $M - B$ and $N - B$ have the same isotropy subgroups, where N is the invariant closed tubular neighbourhood of B in M. Note that $N - B$ always has the same isotropy subgroups as $S(\nu)$, the total space of the invariant sphere bundle of ν, does. Moreover, M has the structure of a finite G-CW-complex containing B as an invariant subset, and the tangent bundle τ_M has the structure of a G-vector bundle over M, such that $\tau_M|B$ and $\tau_B \oplus \nu$ are equivalent as G-vector bundles.

Thus, in order to construct such M, it is necessary to have a finite G-CW-complex X of the given nonequivariant homotopy type and to have a G-vector bundle ξ over X such that the following two conditions hold.

(NB) *Normal Bundle Condition.* X contains B as an invariant subset and the bundles $\xi|B$ and $\tau_B \oplus \nu \oplus \varepsilon^k$ (where $k \geqq 0$) are equivalent as G-vector bundles.

(IS) *Isotropy Subgroup Condition.* Each isotropy subgroup in $X - B$ occurs also in $S(\nu)$.

In turns out that the existence of such X and ξ is also sufficient whenever ν fulfills the following general position condition.

(GP) *General Position Condition.* For each isotropy subgroup H in $X - B$ and each connected component $E(\nu)_i^H$ of $E(\nu)^H$,

$$\dim E(\nu)_i^H > \dim X^H + \max_C \{\dim X^H, \dim C\},$$

where C runs through the set of all connected components of $E(\nu)_i^H - E(\nu)_H$.

In fact, the following theorem holds (cf. Edmonds and Lee [8], Theorem (4.1) and Addendum (4.4) and Assadi [1], Theorem III.2.3).

Theorem 3.1. *Let B be a compact smooth G-manifold, let ν be a smooth G-vector bundle over B, let X be a finite G-CW-complex, and let ξ be an $(n+k)$-dimensional G-vector bundle over X such that the conditions* (NB), (IS), (GP) *stated above all hold. Then, there is an n-dimensional compact smooth G-manifold M and there is an equivariant map $f\colon M \to X$, such that M contains B as a smooth invariant submanifold with equivariant normal bundle ν, $M - B$ and $S(\nu)$ have the same isotropy subgroups, f is a G-homotopy equivalence coinciding with the identity on B, and the induced bundle $f^*\xi$ and $\tau_M \oplus \varepsilon^k$ are equivalent as G-vector bundles.*

The idea of the proof is to take $N = D(\nu)$, the total space of the invariant disk bundle of ν, and then to build M and f from N and the bundle projection $N \to B$ inductively by thickening each equivariant cell in $X - B$ to an equivariant handle in a way prescribed by ξ (see Pawałowski [16] for the details). This also works for compact Lie group actions (see Pawałowski [17]).

Remark 3.2. With the hypotheses of Theorem 3.1, assume further X is connected. Then, the n-dimensional manifold M is constructed so that M and ∂M both are connected manifolds with the same fundamental group (cf. Assadi [1], Remark III.2.4). Thus, if X is contractible and $n \geqq 6$, it follows from the h-Cobordism Theorem that M and D^n are diffeomorphic.

Corollary 3.3. *Let F be a compact smooth manifold, let X be a finite contractible G-CW-complex with $X^G = F$, and let ξ be a G-vector bundle over X whose fixed point bundle ξ^G over F is equivalent to $\tau_F \oplus \varepsilon^k$, where $k \geqq 0$. Then, there is a representation W of G with $W^G = 0$, and there is a smooth action of G on the disk D^n, $n = \dim \xi + \dim W - k$, such that D^n contains F as the fixed point set with equivariant normal bundle $(\xi^G)^\perp \oplus \varepsilon^W$. In particular, the nontrivial summand of the* *representation of G at any $x \in F$ is equivalent to $(\xi_x^G)^\perp \oplus W$. Moreover, there is a G-homotopy equivalence $f: D^n \to X$ coinciding with the identity on F, such that $f^* \xi \oplus \varepsilon^W$ and $\tau_{D^n} \oplus \varepsilon^k$ are equivalent as G-vector bundles.*

Proof. Put $\nu = (\xi^G)^\perp$. Thus ν is a smooth G-vector bundle over F and as G-vector bundles:

$$\xi | F \cong \xi^G \oplus \nu \cong \tau_F \oplus \nu \oplus \varepsilon^k.$$

Hence, the condition (NB) stated before Theorem 3.1 holds for ν with $B = F$. However, ν may not fulfill the conditions (IS) and (GP). But we may choose a representation W of G with $W^G = 0$, so that $\nu \oplus \varepsilon^W$ fulfills these conditions with ξ replaced by $\xi \oplus \varepsilon^W$. Then, of course, $S(\nu \oplus \varepsilon^W)^G = \varnothing$. For example, as W take a sufficiently large multiple of the nontrivial summand of the real regular representation of G (cf. Edmonds and Lee [8], the proof of (4.1)). Since we may assume that $\dim \xi + \dim W - k \geqq 6$, the result follows from Theorem 3.1 and Remark 3.2. □

Remark 3.4. It is clear that if in Corollary 3.3, ξ is a *complex* G-vector bundle and W is a sufficiently large multiple of the nontrivial summand of the *complex* regular representation of G, then the result remains true with $n = \dim_{\mathbb{R}} \xi + \dim_{\mathbb{R}} W - k$.

§ 4. Constructing of G-vector Bundles

We will construct complex G-vector bundles over a contractible G-CW-complex X. In particular, the underlying vector bundles will have the form $X \times \mathbb{C}^n$ over X. When constructing these G-vector bundles we will make use of Proposition 4.1 which contains a result analogous to that dealing with real G-vector bundles given by Bredon [3], Chap. VI, Proposition 11.1.

Let G be a compact Lie group and let S be either $SU(n)$ or $U(n)$. Consider the space $\mathrm{Map}_\cdot(G, S)$ of maps $\theta: G \to S$ with $\theta(e) = I$, the identity matrix, with the compact-open topology. Let G act on $\mathrm{Map}_\cdot(G, S)$ by

$$(h\theta)(g) = \theta(gh)\,\theta(h)^{-1} \qquad \text{for } g, h \in G.$$

Proposition 4.1. *For any G-space X, there is a natural one-one correspondence between unitary (resp., special unitary) G-vector bundle structures on $X \times \mathbb{C}^n$ over X and equivariant maps $\xi: X \to \mathrm{Map}_\cdot(G, S)$ with $S = U(n)$ (resp., $S = SU(n)$). For a given $\xi: x \mapsto \xi_x$, the corresponding action of G on $X \times \mathbb{C}^n$ is defined by*

$$g(x, v) = (gx, \xi_x(g) \cdot v).$$

Clearly, the fixed point set $\mathrm{Map}_\cdot(G, S)^G$ is just the space $\mathrm{Hom}(G, S)$ of homomorphisms $\rho: G \to S$. If G is a finite group, we show that for a subgroup H of G, $\mathrm{Hom}(H, S)$ occurs as a factor in a product decomposition of the fixed point set $\mathrm{Map}_\cdot(G, S)^H$ which, of course, consists of all θ with $\theta(gh) = \theta(g)\,\theta(h)$ for $g \in G$ and $h \in H$. First, for each coset gH, choose a representative a_{gH} with $a_H = e$, consider G/H as a space with base point H, and write $\mathrm{Map}_\cdot(G/H, S)$ for the space of maps $\sigma: G/H \to S$ with $\sigma(H) = I$. Then, the following proposition holds provided G is a *finite* group.

Proposition 4.2. *The map* f: $\mathrm{Map.}(G,S)^H \to \mathrm{Hom}(H,S) \times \mathrm{Map.}(G/H,S)$, $\theta \mapsto (\theta|H, \sigma)$ *with* $\sigma(gH) = \theta(a_{gH})$ *is a homeomorphism.*

Proof. For the inverse of f, take the map f': $(\rho,\sigma) \mapsto \theta$ with $\theta(g) = \sigma(gH)\,\rho(a_{gH}^{-1}g)$. Note that θ does lie in $\mathrm{Map.}(G,S)^H$ because for $g \in G$ and $h \in H$,

$$\begin{aligned}\theta(gh) &= \sigma(gH)\,\rho(a_{gH}^{-1}gh)\\ &= \sigma(gH)\,\rho(a_{gH}^{-1}g)\,\rho(h)\\ &= \theta(g)\,\theta(h).\end{aligned}$$

A straightforward verification shows that the compositions $f' \circ f$ and $f \circ f'$ both are the identity maps. □

Now, we assume that G is a *finite* group and for a subgroup H of G, we describe the connected components of $\mathrm{Map.}(G,S)^H$. According to Proposition 4.2, it suffices to deal with $\mathrm{Hom}(H,S)$ because

$$\mathrm{Map.}(G/H,S) \cong S \times \ldots \times S,\ (G:H)-1 \text{ times}.$$

To begin, consider the action of $U(n)$ on $\mathrm{Hom}(H,S)$ given by

$$(A\rho)(h) = A\,\rho(h)\,A^{-1}.$$

Then, for any $\rho \in \mathrm{Hom}(H,S)$, the isotropy subgroup $U(n)_\rho$ consists of all $A \in U(n)$ with $A\rho(h)A^{-1} = \rho(h)$ for $h \in H$, i.e., $U(n)_\rho$ is equal to the centralizer $Z_{\rho(H)}$ of $\rho(H)$ in $U(n)$. Note that the connected component of $\mathrm{Hom}(H,S)$ containing ρ is just the equivalence class of ρ, i.e., the orbit of ρ under the action of $U(n)$. Thus, this connected component is homeomorphic to the homogeneous space $U(n)/Z_{\rho(H)}$.

Now, the following corollary follows from Proposition 4.2.

Corollary 4.3. *Let* $\rho_0, \rho_1 : G \to S$ *be two representations. Then,* $\rho_0|H$ *and* $\rho_1|H$ *are equivalent if and only if* ρ_0 *and* ρ_1 *lie in the same connected component of* $\mathrm{Map.}(G,S)^H$.

Proposition 4.4. *Each connected component of* $\mathrm{Hom}(H,S)$ *is* 1-*connected.*[1]

Proof. Consider any $\rho \in \mathrm{Hom}(H,S)$. As noticed above, the connected component of $\mathrm{Hom}(H,S)$ containing ρ is homeomorphic to $U(n)/Z_{\rho(H)}$. A straightforward verification shows that $Z_{\rho(H)}$ has the form (up to conjugation)

$$U(n_1,\ldots,n_k) = U(n_1) \times \ldots \times U(n_k) \subset U(n)$$

with $n_1 + \ldots + n_k = n$ and $n_i \geqq 1$ for all i. Since the embedding of $U(n_1,\ldots,n_k)$ into $U(n)$ induces an epimorphism on the fundamental groups, it follows from the exact sequence of the fibration

$$U(n_1,\ldots,n_k) \to U(n) \to U(n)/U(n_1,\ldots,n_k)$$

that $U(n)/U(n_1,\ldots,n_k)$ is 1-connected. □

Since $SU(n)$ is 2-connected, so is $\mathrm{Map.}(G, SU(n))$. Moreover, the following corollary follows from Propositions 4.2 and 4.4.

[1] We assume further H is abelian

Corollary 4.5. *Each connected component of* $\mathrm{Map.}(G, SU(n))^H$ *is* 1-*connected.*[2]

When constructing smooth actions we will make use of the following proposition (see Theorems 5.1 and 7.4).

Proposition 4.6. *Let G be a finite group and let X be a finite $\mathscr{P}$-essential G-CW-complex such that for any $0 \neq P \in \mathscr{P}(G)$, each equivariant cell of type (P) has dimension less or equal 2 and each free cell has dimension less or equal 3. Let $F_1, \ldots, F_k$ be all connected components of F, the fixed point set of X, and let $V_1, \ldots, V_k$ be complex representations of G, such that $V_i = \mathbb{C}^n(\rho_i)$ with $\rho_i \in \mathrm{Hom}(G, SU(n))$ and $\mathrm{Res}_P(V_i) \cong \mathrm{Res}_P(V_j)$ for $P \in \mathscr{P}(G)$, $1 \leqq i, j \leqq k$. Then, there is a G-vector bundle ξ over X, such that $\xi | F_i \cong \varepsilon^{V_i}$ for $i = 1, \ldots, k$.*[3]

Proof. Let X_0 be the sum of F and all equivariant 0-cells of X. Let

$$\xi_0 \colon X_0 \to \mathrm{Hom}(G, SU(n)) \subset \mathrm{Map.}(G, SU(n))$$

be an equivariant map which maps all of F_i into ρ_i and which maps $X_0 - F$ into the set $\{\rho_1, \ldots, \rho_k\}$. We extend ξ_0 to an equivariant map ξ: $X \to \mathrm{Map.}(G, SU(n))$. Once such an extension has been obtained the proof is completed by using Proposition 4.1.

Now, X is build up from X_0 by attaching a finite sequence of equivariant cells of the form $G/H \times D^m$ via equivariant maps α defined on $G/H \times S^{m-1}$. Thus, in order to obtain the extension on a cell $G/H \times D^m$, it suffices to show that α composed with the previous extension restricted to $\{H\} \times S^{m-1}$ is null-homotopic in $\mathrm{Map.}(G, SU(n))^H$.

For $H \notin \mathscr{P}(G)$, α is a constant map on $\{H\} \times S^{m-1}$ because X is $\mathscr{P}$-essential and we choose the extension so that it is the constant map on $\{H\} \times D^m$.

For $H = P \in \mathscr{P}(G)$, the extension on cells $G/P \times D^1$ and $G/P \times D^2$ exists. This follows, respectively, from Corollary 4.3 (since $\rho_i | P \cong \rho_j | P$) and Corollary 4.5.

Finally, the extension on free cells $G \times D^m$, $m \leqq 3$, exists because $\mathrm{Map.}(G, SU(n))$ is 2-connected. $\square$

§ 5. Fixed Point Set Realization Theorem

The following fixed point set realization theorem allows us to construct smooth actions on disks, spheres, and euclidean spaces with fixed point set connected components of different dimensions.

Theorem 5.1. *Let F be a compact smooth parallelizable manifold whose all connected components are either even or odd dimensional. Assume further F has the structure of a CW-complex containing as a deformation retract a subcomplex L with $\dim L = 1$ and $\chi(L) = 1$. Then, for any two distinct primes p and q, F is the fixed point set of a smooth action of $\mathbb{Z}_{pq}$ on a disk.*

Proof. Let $F_1, \ldots, F_k$ and $L_1, \ldots, L_k$ be all connected components of F and L, respectively. We may assume that F_i contains L_i as a deformation retract. Put

[2] We assume further H is abelian

[3] We assume further G has abelian Sylow subgroups

$$n_i = [(\dim F_i + 1)/2],$$

the greatest integer in $(\dim F_i + 1)/2$. Again, we may assume that $n_1 \leqq \ldots \leqq n_k$. Now, for $n = 2n_k + 1$, consider the representations $\rho_i\colon G \to SU(n)$ given by

$$\rho_i = n_i(t^{p+q} \oplus 1) \oplus (n_k - n_i)(t^p \oplus t^q) \oplus t^{-n_k(p+q)}.$$

Then, $\rho_i|\mathbb{Z}_p \cong \rho_j|\mathbb{Z}_p$ and $\rho_i|\mathbb{Z}_q \cong \rho_j|\mathbb{Z}_q$ for $1 \leqq i, j \leqq k$.

Since $\chi(L) = 1$, thus for $G = \mathbb{Z}_{pq}$, Oliver Construction gives a finite contractible G-CW-complex Y such that $Y^G = L$, $\dim Y^{\mathbb{Z}_p} = \dim Y^{\mathbb{Z}_q} = 2$, and $\dim Y = 3$ (see Remark 2.5).

Consider the finite G-CW-complex $X = Y \cup_L F$, the sum of Y and F along L. Since F contains L as a deformation retract, X is contractible. Clearly, $X^G = F$.

Now, it follows from Proposition 4.6 that there is a G-vector bundle ξ over X such that $\xi|F_i \cong \varepsilon^{V_i}$ with $V_i = \mathbb{C}^n(\rho_i)$ for $i = 1, \ldots, k$. Hence, the fixed point bundle ξ^G over F is equivalent to τ_F when all F_i are even dimensional, and ξ^G is equivalent to $\tau_F \oplus \varepsilon^1$ when all F_i are odd dimensional. Thus, by Corollary 3.3, the result follows. □

Let G be a compact Lie group and let H be a closed normal subgroup of G. If the quotient group G/H acts on a space X with fixed point set F, then the epimorphism $G \to G/H$ allows us to consider X as a G-space with $X^G = F$. Clearly, the action of G on X may not be effective, but it is easy to build up X to a space with an effective action of G and with the same fixed point set. For example, consider an orthogonal action of G on a disk D^n given via a faithful representation $\rho\colon G \to O(n)$, such that the fixed point set is just the origin. Then, $X \times D^n$ (with the diagonal action of G) has the required properties. Hence, the following proposition holds.

Proposition 5.2. *Let G be a compact Lie group and let H be a closed normal subgroup of G. If F is the fixed point set of a smooth action of G/H on a disk, then F is the fixed point set of a smooth effective action of G on a disk.*

Now, let G be a finite group and let H be a subgroup of G with index n (here H need not be normal). Assume H acts on a space X with fixed point set F. In the proof of Lemma 6 in [13], Oliver defines an action of G on X^n, the n-fold cartesian product of X, with the fixed point set equal to the image of F under the diagonal map $X \to X^n$. It follows easily from the construction of the action of G that if H acts smoothly on a smooth manifold X, then the action of G on X^n is also smooth. Thus, the following proposition holds.

Proposition 5.3. *Let G be a finite group and let H be a subgroup of G. If F is the fixed point set of a smooth action of H on a disk, then F is the fixed point set of a smooth action of G on a disk.*

The following corollary follows immediately from Propositions 5.2 and 5.3.

Corollary 5.4. *Let G be a compact Lie group. If F is the fixed point set of a smooth action of some subgroup of G/G_0 on a disk, then F is the fixed point set of a smooth effective action of G on a disk.*

Finally, Theorem 5.1 and Corollary 5.4 give the following corollary.

Corollary 5.5. *Let G be a compact Lie group and let F be as in Theorem 5.1. If $G \notin \mathscr{R}$, then F is the fixed point set of a smooth action of G on a disk.*

§ 6. Examples of Smooth Group Actions

In this section we give examples of smooth group actions on disks, spheres, and euclidean spaces with fixed point set connected components of different dimensions. We also construct smooth group actions which enable us to solve (in some cases) the following problems in compact transformation groups.

Let G be a compact Lie group. The problems are of interest also for the case of finite cyclic G but not of prime power order.

Problem 1. If G acts on D^n with the origin as a stationary point, then for $H \neq G$ with $D^n_{(H)} \neq \varnothing$, is $S^{n-1}_{(H)} \neq \varnothing$? Is this the case when the action on the boundary S^{n-1} is orthogonal?

Problem 2. If G acts on D^n and if $F(G, D^n) \subset \operatorname{int} D^n$, does $F(G, D^n)$ contain *at most* one point?

Problem 3. If G acts on $S^n \times I$ (I the unit interval), then can the set F of stationary points touch $S^n \times \{0\}$ without also touching $S^n \times \{1\}$? Can this happen when the action on both ends is orthogonal?

Problem 4. Let G act on $\mathbb{R}^n$ with a stationary point x. If $\mathbb{R}^n_{(H)} \neq \varnothing$, must x be in the closure of $\mathbb{R}^n_{(H)}$?

Problem 5. Let G act on $\mathbb{R}^n$. If G is finite and contains $\mathbb{Z}_{pq}$ as a proper subgroup, where p and q are two distinct primes, then can $F(G, \mathbb{R}^n)$ and $F(\mathbb{Z}_{pq}, \mathbb{R}^n)$ be two different compact sets? Can the dimensions of these sets be different?

Problems 1, 2, 3, and 4 are posed in Bredon [3], p. 205. For $G = S^1$, Problem 4 is due to Raymond. Stein [25] has obtained the first counterexample by constructing a smooth action of S^1 on $\mathbb{R}^5$ with $F(S^1, \mathbb{R}^5) = \mathbb{R}^1$ and $\mathbb{R}^5_{(H)} = S^1/H$ for $H = \mathbb{Z}_6$.

Problem 5 goes back to Smith [24], Question on p. 412. For $G = \mathbb{Z}_{pq}$, Smith asked: can $F(G, \mathbb{R}^n)$ be compact? Edmonds and Lee [8] have answered affirmatively that question by showing that each closed smooth stably parallelizable manifold is the fixed point set of a smooth action of $\mathbb{Z}_{pq}$ on some $\mathbb{R}^n$. For $G = \mathbb{Z}_{pqr}$, where p, q, and r are three relatively prime integers, Assadi [1] has constructed a smooth action of G on $\mathbb{R}^n$ such that $F(G, \mathbb{R}^n) = \{x_0\}$ and $F(\mathbb{Z}_{pq}, \mathbb{R}^n) = \{x_0, x_1, \ldots, x_r\}$ providing the affirmative answer when $G \neq \mathbb{Z}_{pq}$.

Assadi [1] has also obtained counterexamples to Problems 1, 2, and 4 for $G = \mathbb{Z}_{pqr} \oplus \mathbb{Z}_{pqr} \oplus \mathbb{Z}_s$, where p, q, r, and s are four relatively prime integers, and for any finite group G not of prime power order, he has defined a smooth action of G on a disk with exactly k fixed points, where $k \equiv 1 \pmod{n_G}$, providing counterexamples to Problem 2 whenever $n_G \neq 0$ (see Assadi [1], Chap. V).

Example 6.1. Let $n_0, n_1, \ldots, n_k$ be integers such that $n_0 \geqq 0$, $n_i \geqq 1$, and $n_0 \equiv n_i \pmod 2$ for $i=1,\ldots,k$. Consider the disks D^{n_0+1} with origin x_0 and D^{n_i} with origin x_i for $i=1,\ldots,k$. Then, take the disjoint unions

$$F = D^{n_0+1} \amalg \left(\coprod_{i=1}^{k} D^{n_i} \times S^1 \right)$$

and

$$L = \{x_0\} \amalg \left(\coprod_{i=1}^{k} \{x_i\} \times S^1 \right)$$

and put

$$Q = \coprod_{i=1}^{k} S^{n_i - 1} \times S^1.$$

Clearly, F contains L as a deformation retract, $\dim L = 1$, and $\chi(L) = 1$.

Let G be a compact Lie group, and assume that $G \notin \mathscr{R}$. Then, by Corollary 5.5, there is a smooth action of G on a disk D^{n+1} with fixed point set F.

(6.1.1) This action restricted to the boundary ∂D^{n+1} is a smooth action of G on S^n with

$$F(G, S^n) = S^{n_0} \amalg Q.$$

(6.1.2) Choose a fixed point $x \in S^{n_0}$ and (using the Slice Theorem) take a sufficiently small invariant disk N in S^n around x, so that G acts orthogonally on N. Since $S^n - \operatorname{int} N \cong D^n$, we get a smooth action of G on D^n with

$$F(G, D^n) = D^{n_0} \amalg Q,$$

such that the action on the boundary ∂D^n is orthogonal.

(6.1.3) Since $\operatorname{int} D^n \cong \mathbb{R}^n$, we get now a smooth action of G on $\mathbb{R}^n$ with

$$F(G, \mathbb{R}^n) = \mathbb{R}^{n_0} \amalg Q.$$

Put $n_0 = 0$ and assume $k \geqq 1$. Then $n_1, \ldots, n_k$ are even, $D^{n_0} = \{x_0\}$, and $F(G, D^n)$ is a closed manifold. Thus,

$$F(G, D^n) \subset \operatorname{int} D^n$$

and we get counterexamples to Problem 2. Recall that the action on the boundary ∂D^n is orthogonal.

Now, choose a fixed point $x \in Q$ and take a sufficiently small invariant disk N in D^n around x, so that G acts orthogonally on N. Since $D^n - \operatorname{int} N \cong S^{n-1} \times I$, we get a smooth action of G on $S^{n-1} \times I$, such that the action on both ends $S^{n-1} \times \{0\}$ and $S^{n-1} \times \{1\}$ is orthogonal, and the fixed point set touch only one end. It follows from Theorem A that this cannot happen for $G \in \mathscr{R}$, because if this can happen, then there would exist a smooth action of G on D^n with fixed point set connected components of different dimensions. Thus, in the case of smooth actions, Problem 3 has the affirmative answer if and only if $G \notin \mathscr{R}$.

Example 6.2. Let p, q, and r be three distinct primes. Consider the cyclic group G of order pqr and its subgroup H of order pq. Then, by Example 6.1, there is a smooth action of H on a disk M with $M^H = \{x\} \amalg (D^1 \times S^1)$. Since M has the structure of a finite pointed H-CW-complex with x as the base point, thus the r-fold wedge of M,

$$X = M \vee \overset{r}{\cdots} \vee M,$$

has the structure of a finite H-CW-complex with $X^H = B$, where

$$B = \{x\} \sqcup (D^1 \times S^1) \sqcup \overset{r}{\cdots} \sqcup (D^1 \times S^1).$$

Now, the action of H on X is extended to the action of G on X by permuting the copies of $M - \{x\}$ in X and leaving x fixed. Thus X is a finite contractible G-CW-complex containing B as an invariant subset. Clearly, $X^G = \{x\}$ and $X_H = B - \{x\}$. Moreover, the family of isotropy subgroups in $X - B$ consists of 0, $\mathbb{Z}_p$, and $\mathbb{Z}_q$.

For each copy of M in X, consider the H-equivariant tangent bundle τ_M and the H-equivariant normal bundle $\nu(M^H, M)$ of M^H in M. Then, consider the H-vector bundles ξ over X and ν over B obtained, respectively, from all copies of τ_M and $\nu(M^H, M)$ by piecing together all copies of the tangent space $T_x(M)$. Finally, extend the action of H on $E(\xi)$ to the action of G on $E(\xi)$ by permuting the copies of $T(M) - T_x(M)$ in $E(\xi)$ and leaving $T_x(M)$ fixed. Similarly, extend the action of H on $E(\nu)$ to the action of G on $E(\nu)$. Hence, ξ is a G-vector bundle over X and ν is a smooth G-vector bundle over B, such that $\xi|B$ and $\tau_B \oplus \nu$ are equivalent as G-vector bundles. Moreover, the family of isotropy subgroups in $S(\nu)$ consists of 0, $\mathbb{Z}_p$, $\mathbb{Z}_q$, and $\mathbb{Z}_r$.

Thus, for X, ξ, B, and ν, the conditions (NB) and (IS) stated before Theorem 3.1 both hold. Moreover, we may choose a representation W of G, so that $\nu \oplus \varepsilon^W$ fulfills the condition (GP). For example, let W be a sufficiently large multiple of $\mathbb{C}^3(t \oplus t^p \oplus t^q)$, where t is the standard 1-dimensional complex representation of G. Then, $S(\nu \oplus \varepsilon^W)$ and $S(\nu)$ have the same isotropy subgroups, and the conditions (NB), (IS), and (GP) all hold with ν replaced by $\nu \oplus \varepsilon^W$ and ξ replaced by $\xi \oplus \varepsilon^W$. Since $\dim \xi \geqq 6$, it follows from Theorem 3.1 and Remark 3.2 that there is a smooth action of G on the disk D^n, $n = \dim M + \dim_{\mathbb{R}} W$, containing B as an invariant submanifold, such that the family of isotropy subgroups in $D^n - B$ consists of 0, $\mathbb{Z}_p$, $\mathbb{Z}_q$, and $\mathbb{Z}_r$. Hence, $F(G, D^n) = \{x\}$ and $D^n_H = B - \{x\}$.

(6.2.1) Take the equivariant double of D^n, i.e., first consider $D^n \times D^1$ with the diagonal action of G (where G acts trivially on D^1) and then restrict the action to the boundary $\partial(D^n \times D^1)$. Thus, we get a smooth action of G on S^n with $F(G, S^n) = \{x, y\}$ and

$$S^n_H = T^2 \sqcup \overset{r}{\cdots} \sqcup T^2, \qquad \text{where } T^2 = S^1 \times S^1.$$

(6.2.2) Take a sufficiently small invariant disk N in S^n around y, so that G acts orthogonally on N. Since $S^n - \operatorname{int} N \cong D^n$, we get a smooth action of G on D^n, such that the action on the boundary ∂D^n is orthogonal, with $F(G, D^n) = \{x\}$ and

$$D^n_H = T^2 \sqcup \overset{r}{\cdots} \sqcup T^2 \subset \operatorname{int} D^n.$$

This provides counterexamples to Problem 1.

(6.2.3) Since $\operatorname{int} D^n \cong \mathbb{R}^n$, we get a smooth action of G on $\mathbb{R}^n$ with $F(G, \mathbb{R}^n) = \{x\}$, $F(H, \mathbb{R}^n) = \{x\} \amalg \mathbb{R}^n_H$, and

$$\mathbb{R}^n_H = T^2 \amalg \overset{r}{\ldots} \amalg T^2.$$

This, in turn, provides counterexamples to Problems 4 and 5.

Example 6.3. Let G be a finite group not of prime power order and let F be a compact smooth manifold with $\chi(F) \equiv 1 \pmod{n_G}$. Then, by Theorem 2.2, there is a finite contractible G-CW-complex X with $X^G = F$. Assume further F is stably parallelizable and all connected components of F have the same dimension. Choose a representation V of G with $\dim V^G = \dim F + 1$. Then, for the G-vector bundle $\xi = \varepsilon^V$ over X, the fixed point bundle ξ^G over F is equivalent to $\tau_F \oplus \varepsilon^1$. Hence, by Corollary 3.3, there is a smooth action of G on a disk D^n with fixed point set F.

Assume that $\partial F = \varnothing$. Then $F \subset \operatorname{int} D^n$ and thus for $F \neq pt$, we get again counterexamples to Problem 2. For example, let $G = A_5$, the alternating group of order 60. Then $n_G = 1$, and hence F can be any closed smooth stably parallelizable manifold with all connected components of the same dimension. Note that it follows from Theorem A that for any smooth A_5-action on a disk, the fixed point set connected components *must have* the same dimension.

§ 7. Representations at Fixed Points

In this section we deal with representations at fixed points of smooth G-actions on acyclic manifolds and homology spheres.

Proposition 7.1. *Let G be a compact Lie group acting smoothly on an acyclic manifold M, and assume $G \in \mathcal{R}$. Then, at any two fixed points x and y, the representations of G are equivalent.*

Proof. Let $V = T_x(M)$ and $W = T_y(M)$. Let χ_V and χ_W be the characters of V and W, respectively.

Case 1. G is finite. Since $G \in \mathcal{R}$, it follows that for any cyclic subgroup C of G, $C \in \mathcal{P}$ and hence M^C is nonempty and connected by Smith Theory. Thus, χ_V and χ_W coincide on any cyclic subgroup of G, hence $\chi_V = \chi_W$.

Case 2. $G \in \mathcal{P}$. Again, by Smith Theory, the set $M^G = (M^{G_0})^{G/G_0}$ is connected, and hence $\chi_V = \chi_W$.

Case 3. G is connected. For any $g \in G$, consider a maximal torus T of G, such that $g \in T$. Since $T \in \mathcal{P}$, thus by Case 2, $\chi_V(g) = \chi_W(g)$.

Case 4. G is neither finite nor connected. For any $g \in G$, there is a Cartan subgroup S of G, such that $g \in S$ (see Segal [22], Definition 1.1 and Proposition 1.2). The projection $G \to G/G_0$ maps S onto a cyclic subgroup C of G/G_0, such that $(C:1) | (S:S_0)$ and $(S:S_0) | (C:1)^2$ (see Segal [22], Remarks on p. 117). Since $G \in \mathcal{R}$, $(C:1)$ is a prime power and hence so is $(S:S_0)$. Thus, $S \in \mathcal{P}$ and by Case 2, $\chi_V(g) = \chi_W(g)$. $\square$

Using Smith Theory one can prove the following proposition similarly as Proposition 7.1.

Proposition 7.2. *Let G be a compact Lie group acting smoothly on a homology sphere with at least three fixed points. If $G \in \mathscr{R}$, then at any two fixed points the representations of G are equivalent.*

Now, let G be a compact Lie group acting smoothly on a homology sphere with exactly two fixed points and let V and W be the representations of G at the fixed points. We want to mention examples of those G for which V and W are equivalent.

First, recall the theorem of Atiyah, Bott, and Milnor which says that if the action is semifree, i.e., free outside the fixed points, then V and W are equivalent (see Atiyah and Bott [2], Theorem 7.27 and Milnor [12], Theorem 12.11). It follows easily from this theorem that without the assumption that the action is semifree, V and W are equivalent in each of the following three cases.

(1) $G = \mathbb{Z}_p \times \ldots \times \mathbb{Z}_p$, where p is a prime.

(2) G is a torus.

(3) G is connected.

Then, recall the result of Sanchez which says that for $G = \mathbb{Z}_{p^n}$, where p is an odd prime, V and W are equivalent again without the assumption that the action is semifree (see Sanchez [21], Corollary 1.11). Thus, V and W are also equivalent in each of the following three cases (cf. the proof of Proposition 7.1).

(1)′ G is a finite p-group, where p is an odd prime.

(2)′ G is a Cartan group such that $(G:G_0)$ is an odd prime power.

(3)′ $G \in \mathscr{R}^{\mathrm{odd}}$, the class of compact Lie groups G such that each element of G/G_0 has odd prime power order.

According to the author's knowledge, the complete list of groups G for which V and W are equivalent, is still unknown. We note here that beside the linear equivalence between V and W, one studies other kind of equivalence between them, such as the topological equivalence and Smith equivalence (see, e.g., Cappell and Shaneson [5, 6], Illman [10], and Petrie [20]).

It turns out that when dealing with complex representations at fixed points of smooth finite group actions an acyclic manifolds, we may assume that these representations stably are special unitary representations. This follows from Smith Theory and the following lemma.

Lemma 7.3. *Let G be a finite group and let $\rho_i\colon G \to U(n)$, $i = 1, \ldots, k$, be representations such that $\rho_i | P$ and $\rho_j | P$ are equivalent for each $P \in \mathscr{P}(G)$, $1 \leq i, j \leq k$. Then, there is a representation $\rho\colon G \to U(1)$ such that $\rho_i \oplus \rho\colon G \to SU(n+1)$ for $i = 1, \ldots, k$.*

Proof. First, we claim that $\det \rho_i(g) = \det \rho_j(g)$ for each $g \in G$. To see this, for a given element $g \in G$, consider the cyclic subgroup C of G generated by g. Let t be the standard 1-dimensional complex representation of C and let

$$\rho_i | C \cong t^{a_{i1}} \oplus \ldots \oplus t^{a_{in}}$$

be the decomposition of $\rho_i | C$ into the sum of irreducible representations of C.

Since $\rho_i|P \cong \rho_j|P$ for $P \in \mathscr{P}(C)$, it follows that

$$a_{i1} + \ldots + a_{in} \equiv a_{j1} + \ldots + a_{jn} \pmod{(P:1)}.$$

Thus,

$$a_{i1} + \ldots + a_{in} \equiv a_{j1} + \ldots + a_{jn} \pmod{(C:1)}$$

and hence

$$\det \rho_i(g) = \zeta^{a_{i1} + \ldots + a_{in}} = \zeta^{a_{j1} + \ldots + a_{jn}} = \det \rho_j(g),$$

where ζ is the primitive $(C:1)$-rooth of unity, proving the claim.

Now, the required representation $\rho\colon G \to U(1)$ is given by $\rho(g) = (\det \rho_i(g))^{-1}$ for $g \in G$. □

Theorem 7.4. *Assume G is an extension of the form*

$$0 \to \mathbb{Z}_r \to G \to \mathbb{Z}_s \to 0$$

where r and s are relatively prime and s is a product of distinct primes. Let $V_1, \ldots, V_k$ be complex representations of G such that the following three conditions hold.

(a) $k \equiv 1 \pmod{n_G}$.

(b) $V_i^G = 0$ *for* $i = 1, \ldots, k$.

(c) $\mathrm{Res}_P(V_i) \cong \mathrm{Res}_P(V_j)$ *for each* $P \in \mathscr{P}(G)$, $1 \leqq i, j \leqq k$.

Then there is a smooth action of G on a disk with exactly k fixed points at which the tangential representations of G are equivalent to $V_1 \oplus W, \ldots, V_k \oplus W$ for some complex representation W of G.

Proof. Let $p_1, \ldots, p_m$ be all primes which divide r. Consider the normal subgroup $H = \mathbb{Z}_{r/p_1 \cdots p_m}$ of G. Then, by Proposition 2.4, $n_G = n_{G/H}$. Since the order of G/H is a product of distinct primes and (a) holds, Oliver Construction gives a finite contractible $\mathscr{P}$-essential (G/H)-CW-complex X with fixed point set $F = \{x_1, \ldots, x_k\}$, $\dim X = 3$ and $\dim X^P = 2$ for each $0 \neq P \in \mathscr{P}(G/H)$ (see Theorem 2.2 and Remarks 2.5 and 2.6). Now, the epimorphism $G \to G/H$ allows us to consider X as a finite contractible $\mathscr{P}$-essential G-CW-complex with $X^G = F$ and $\dim X^P = 2$ for each $0 \neq P \in \mathscr{P}(G)$.

According to Lemma 7.3, we may assume that $V_i = \mathbb{C}^n(\rho_i)$ with $\rho_i \in \mathrm{Hom}(G, SU(n))$ for $i = 1, \ldots, k$. Since (c) holds, it follows from Proposition 4.6 that there is a G-vector bundle ξ over X such that $\xi_{x_i} = V_i$ for $i = 1, \ldots, k$. Since (b) holds, the fixed point bundle ξ^G over F is equivalent to τ_F. Thus, Corollary 3.3 and Remark 3.4 complete the proof. □

If in Theorem 7.4, G were in $\mathscr{G}^1$, then Corollary 2.3 would say that $k = 1$. However, for $G \notin \mathscr{G}^1$, $n_G \neq 0$ and hence it is possible to choose $k > 1$ with $k \equiv 1 \pmod{n_G}$.

Example 7.5. Let G be as in Theorem 7.4, and assume further $G \notin \mathscr{G}^1$. For any integer $k > 1$, we construct inequivalent complex representations $V_1, \ldots, V_k$ of G, such that $V_i^G = 0$ and $\mathrm{Res}_P(V_i) \cong \mathrm{Res}_P(V_j)$ for $P \in \mathscr{P}(G)$, $1 \leqq i, j \leqq k$.

To begin, note that $r = pq$ for some two relatively prime integers p and q both greater than 1 (otherwise $G \in \mathscr{G}^1$) and consider the following two inequivalent representations of $\mathbb{Z}_r$:

$$W_1 = \mathbb{C}^2(t^{p+q+2} \oplus t^2) \quad \text{and} \quad W_2 = \mathbb{C}^2(t^{p+2} \oplus t^{q+2}),$$

where t is the standard 1-dimensional complex representation of $\mathbb{Z}_r$. Note that W_1 and W_2 are equivalent when restricted to $\mathbb{Z}_p$ or $\mathbb{Z}_q$.

Now, take the induced representations of G,

$$U_1 = \mathrm{Ind}_{\mathbb{Z}_r}^G(W_1) \quad \text{and} \quad U_2 = \mathrm{Ind}_{\mathbb{Z}_r}^G(W_2).$$

Thus, it follows from Serre [23], 3.3 and 7.3, that $U_1^G = U_2^G = 0$, $U_1 \not\cong U_2$, and $\mathrm{Res}_P(U_1) \cong \mathrm{Res}_P(U_2)$ for $P \in \mathscr{P}(G)$. Finally, put

$$V_i = (k-i)\, U_1 \oplus (i-1)\, U_2$$

for $i = 1, \ldots, k$. Then $V_1, \ldots, V_k$ are the required representations of G.

Remark 7.6. Such examples of representations one can also obtain using the Adams operations in the representation ring of G (see, e.g., Petrie [19] and Pawałowski [16]).

References

1. Assadi, A.: Finite group actions on simply connected manifolds and *CW* complexes. Mem. Amer. Math. Soc. **257** (1982)
2. Atiyah, M.F., Bott, R.: A Lefschetz fixed point formula for elliptic complexes II; Applications. Ann. of Math. **88**, 451-491 (1968)
3. Bredon, G.E.: Introduction to compact transformation groups. New York: Academic Press 1972
4. Cappell, S.E., Shaneson, J.L.: Fixed points of periodic maps. Proc. Nat. Acad. Sci. USA **77**, 5052-5054 (1980)
5. Cappell, S.E., Shaneson, J.L.: Fixed points of periodic differentiable maps. Invent. Math. **68**, 1-20 (1982)
6. Cappell, S.E., Shaneson, J.L.: Class numbers and periodic smooth maps. Comment. Math. Helv. **58**, 167-185 (1983)
7. Dovermann, K.H., Petrie, T.: *G*-surgery II. Mem. Amer. Math. Soc. **260** (1982)
8. Edmonds, A.L., Lee, R.: Fixed point sets of group actions on euclidean space. Topology **14**, 339-345 (1975)
9. Hsiang, W.-C., Hsiang, W.-Y.: Some problems in differentiable transformation groups. In: Proceedings of the Conference on Transformation Groups (New Orleans 1967), pp. 223-234. Berlin-Heidelberg-New York: Springer 1968
10. Illman, S.: Representations at fixed points of actions of finite groups on spheres. In: Current Trends in Algebraic Topology. Canadian Math. Soc. Conference Proceedings (London, Ontario 1981), Vol. 2, Part 2, pp. 135-155. Providence, Rhode Island: Amer. Math. Soc. 1982
11. Illman, S.: The equivariant triangulation theorem for actions of compact Lie groups. Math. Ann. **262**, 487-501 (1983)
12. Milnor, J.W.: Whitehead torsion. Bull. Amer. Math. Soc. **72**, 358-426 (1966)
13. Oliver, R.: Fixed point sets of group actions on finite acyclic complexes. Comment. Math. Helv. **50**, 155-177 (1975)
14. Oliver, R.: Smooth compact Lie group actions on disks. Math. Z. **149**, 79-96 (1976)
15. Oliver, R.: *G*-actions on disks and permutation representations II. Math. Z. **157**, 237-263 (1977)
16. Pawałowski, K.: Smooth actions of finite groups on disks. Ph.D. Thesis, University of Warsaw 1981
17. Pawałowski, K.: Equivariant thickening for compact Lie group actions. Preprint.

18. Petrie, T.: G-surgery $I-A$ survey. In: Algebraic and Geometric Topology. Proceedings of a Symposium (Santa Barbara 1977), pp. 197–233. Lecture Notes in Math. **664**. Berlin-Heidelberg-New York: Springer 1978
19. Petrie, T.: Isotropy representations of group actions on spheres. Preprint
20. Petrie, T.: Smith equivalence of representations. Math. Proc. Cambridge Philos. Soc. **94**, 61–99 (1983)
21. Sanchez, C.U.: Actions of groups of odd order on compact orientable manifolds. Proc. Amer. Math. Soc. **54**, 445–448 (1976)
22. Segal, G.: The representation ring of a compact Lie group. Inst. Hautes Études Sci. Publ. Math. **34**, 113–128 (1968)
23. Serre, J.-P.: Linear representations of finite groups. Graduate Texts in Math. **42**. Berlin-Heidelberg-New York: Springer 1977
24. Smith, P.A.: New results and old problems in finite transformation groups. Bull. Amer. Math. Soc. **66**, 401–415 (1960)
25. Stein, E.: On the orbit types in a circle action. Proc. Amer. Math. Soc. **66**, 143–147 (1977)

Received December 19, 1983

Math. Z. 187, 49–60 (1984)

Mathematische Zeitschrift

Chern-Klassen von ganzzahligen und rationalen Darstellungen diskreter Gruppen

Dominique Arlettaz

Mathematisches Seminar, ETH-Zentrum, CH-8092 Zürich, Schweiz

Einleitung

G bezeichne eine diskrete Gruppe und BG ihren klassifizierenden Raum; BG ist ein Eilenberg-MacLane-Raum $K(G,1)$ und für alle $n \geqq 0$ ist die singuläre Cohomologiegruppe $H^n(BG;\mathbb{Z})$ natürlich isomorph zur algebraisch definierten Cohomologiegruppe $H^n(G;\mathbb{Z})$ der Gruppe G. Es sei $\rho\colon G \to GL\mathbb{C} = \bigcup_{m=1}^{\infty} GL_m\mathbb{C}$ eine Darstellung von G; $B\rho\colon BG \to BGL\mathbb{C}$ induziert den Homomorphismus

$$B\rho^*\colon H^*(BGL\mathbb{C};\mathbb{Z}) = \mathbb{Z}[c_1, c_2, c_3, \ldots] \to H^*(BG;\mathbb{Z}),$$

wobei die Klassen c_j $(j \geqq 1)$ die universellen Chern-Klassen sind. Für $j \geqq 1$ ist die j-te *Chern-Klasse der Darstellung* ρ folgenderweise definiert: $c_j(\rho) := B\rho^*(c_j) \in H^{2j}(BG;\mathbb{Z})$. Die Chern-Klassen von ρ sind die Chern-Klassen des zu ρ assoziierten flachen komplexen Vektorbündels $\xi(\rho)$ über dem klassifizierenden Raum BG.

Sei nun ρ eine ganzzahlige Darstellung der diskreten Gruppe G:

$$\rho\colon G \xrightarrow{\phi} GL\mathbb{Z} \overset{i}{\hookrightarrow} GL\mathbb{C},$$

wobei i die übliche Inklusion bezeichnet. Diese Darstellung induziert den Homomorphismus

$$B\rho^*\colon H^*(BGL\mathbb{C};\mathbb{Z}) \xrightarrow{Bi^*} H^*(BGL\mathbb{Z};\mathbb{Z}) \xrightarrow{B\phi^*} H^*(BG;\mathbb{Z}).$$

Wir definieren $c_j(GL\mathbb{Z}) := c_j(i) = Bi^*(c_j) \in H^{2j}(BGL\mathbb{Z};\mathbb{Z})$, $j \geqq 1$. Diese Chern-Klassen $c_j(GL\mathbb{Z})$ sind Torsionsklassen für alle $j \geqq 1$ (vgl. [8] oder [10]). Wegen $c_j(\rho) = B\phi^*(c_j(GL\mathbb{Z}))$ ist die Ordnung von $c_j(GL\mathbb{Z})$ eine obere Schranke für die Ordnung von $c_j(\rho)$. Da die Inklusion $GL\mathbb{Z} \overset{i}{\hookrightarrow} GL\mathbb{C}$ selbst auch eine ganzzahlige Darstellung ist, ergibt die Ordnung von $c_j(GL\mathbb{Z})$ die beste obere Schranke für die Ordnung von $c_j(\rho)$ für alle ganzzahligen Darstellungen ρ von diskreten Gruppen.

Deswegen interessieren wir uns in dieser Arbeit für die Ordnung von $c_j(GL\mathbb{Z})$ in $H^{2j}(BGL\mathbb{Z};\mathbb{Z}), j \geqq 1$. Es ist schon bekannt, daß $c_j(GL\mathbb{Z})$ die Ordnung 2 besitzt, wenn j eine ungerade Zahl ist; das kommt aus der Untersuchung der Chern-Klassen der Darstellungen der zyklischen Gruppe der Ordnung 2 und aus der Tatsache, daß $c_j(GL\mathbb{Z}) = (-1)^j c_j(GL\mathbb{Z})$ für alle $j \geqq 1$, d.h. $2c_j(GL\mathbb{Z}) = 0$ für ungerade j ist.

Es genügt also dieses Problem für gerade j zu behandeln. In diesem Fall haben wir schon einige Informationen. Eckmann und Mislin haben die Ordnung der Chern-Klassen der Darstellungen endlicher Gruppen untersucht (vgl. [7]). Daraus folgt, daß die Ordnung von $c_j(GL\mathbb{Z})$ ein positives Vielfaches von E_j für gerade j ist, wobei E_j den Nenner von $\frac{B_j}{j}$ bezeichnet (B_j ist die j-te Bernoullische Zahl: $B_2 = \frac{1}{6}$, $B_4 = \frac{1}{30}, \ldots$; $E_2 = 12$, $E_4 = 120, \ldots$). Andererseits ist $2E_j$ eine obere Schranke für die Ordnung von $c_j(GL\mathbb{Z})$ nach [8] oder [10]. Für gerade j stellt sich also die Frage, ob $c_j(GL\mathbb{Z})$ die Ordnung E_j oder $2E_j$ in $H^{2j}(BGL\mathbb{Z};\mathbb{Z})$ besitzt.

Der erste Teil der Arbeit befaßt sich mit dem Fall $j=2$ ($E_2 = 12$): wir bestimmen die Cohomologiegruppe $H^4(BGL\mathbb{Z};\mathbb{Z})$ und zeigen, daß $c_2(GL\mathbb{Z})$ ein Element der Ordnung 24 in $H^4(BGL\mathbb{Z};\mathbb{Z})$ ist. In dem zweiten Teil beantworten wir die obige Frage für $j \equiv 2 \bmod 4$: $c_j(GL\mathbb{Z})$ hat die Ordnung $2E_j$ in $H^{2j}(BGL\mathbb{Z};\mathbb{Z})$. Als Anwendung betrachten wir in dem dritten Teil die ganzzahligen Darstellungen der Kongruenzuntergruppen Γ_m ($m \geqq 2$) und beweisen gewisse Resultate über die Ordnung ihrer Chern-Klassen, besonders der zweiten. Schließlich behandeln wir im Anhang das analoge Problem der Ordnung der Chern-Klassen von rationalen Darstellungen diskreter Gruppen.

Die vorliegende Arbeit ist eine Zusammenfassung meiner Dissertation [2], die ich unter der Leitung von Herrn Professor Guido Mislin ausgeführt habe. An dieser Stelle möchte ich ihm für seine Anregungen und wertvollen Ratschläge meinen herzlichen Dank aussprechen.

1. Die Ordnung von $c_2(GL\mathbb{Z})$ in $H^4(BGL\mathbb{Z};\mathbb{Z})$

Wir wollen zuerst bestimmen, ob die Ordnung von $c_2(GL\mathbb{Z})$ in $H^4(BGL\mathbb{Z};\mathbb{Z})$ gleich 12 oder 24 ist. Dafür benötigen wir einige Vorbereitungen.

Lemma 1.1. $H^4(BGL\mathbb{Z};\mathbb{Z}) \cong H_3(BGL\mathbb{Z};\mathbb{Z})$.

Beweis. Aus einem Ergebnis von Borel [3] folgt, daß $H^4(BGL\mathbb{Z};\mathbb{Z})$ und $H_3(BGL\mathbb{Z};\mathbb{Z})$ Torsionsgruppen sind; da $BGL\mathbb{Z}$ ein CW-Komplex mit endlichen Skeletten ist, sind diese Gruppen endlich. Die Behauptung folgt dann aus dem universellen Koeffizienten-Theorem.

Um die ersten Homologiegruppen von $BGL\mathbb{Z}$ und $BSl\mathbb{Z}$ zu berechnen, betrachten wir die $+$Konstruktion von Quillen, welche die Räume $BGL\mathbb{Z}^+$ und $BSL\mathbb{Z}^+$ liefert. Aus den Eigenschaften der $+$Konstruktion folgt:

$$\Pi_1 BGL\mathbb{Z}^+ \cong GL\mathbb{Z}/SL\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \quad \text{und} \quad \Pi_1 BSL\mathbb{Z}^+ \cong SL\mathbb{Z}/SL\mathbb{Z} = 0.$$

An dieser Stelle brauchen wir folgendes Lemma.

Lemma 1.2. $BGL\mathbb{Z}^+ \simeq BSL\mathbb{Z}^+ \times B\mathbb{Z}/2\mathbb{Z}$.

Beweis. Aus [2], S. 7 oder [14], S. 351, (d) folgt die Homotopie-Äquivalenz $BSL\mathbb{Z}^+ \simeq \widetilde{BGL\mathbb{Z}}^+$. Wir haben also die Faserung $BSL\mathbb{Z}^+ \to BGL\mathbb{Z}^+ \to B\mathbb{Z}/2\mathbb{Z}$. Weil $BGL\mathbb{Z}^+$ ein H-Raum ist und weil es einen Schnitt $B\mathbb{Z}/2\mathbb{Z} \to BGL\mathbb{Z}^+$ gibt, gilt dann $BGL\mathbb{Z}^+ \simeq BSL\mathbb{Z}^+ \times B\mathbb{Z}/2\mathbb{Z}$.

Wegen Lemma 1.2 gilt $\Pi_n BSL\mathbb{Z}^+ \cong K_n\mathbb{Z}$ für $n \geqq 2$, insbesondere $\Pi_2 BSL\mathbb{Z}^+ \cong \mathbb{Z}/2\mathbb{Z}$ und $\Pi_3 BSL\mathbb{Z}^+ \cong \mathbb{Z}/48\mathbb{Z}$ (vgl. [12]).

Lemma 1.3.

$$H_1(BSL\mathbb{Z};\mathbb{Z}) \cong H_1(BSL\mathbb{Z}^+;\mathbb{Z}) = 0$$
$$H_2(BSL\mathbb{Z};\mathbb{Z}) \cong H_2(BSL\mathbb{Z}^+;\mathbb{Z}) \cong \mathbb{Z}/2\mathbb{Z}.$$

Beweis. Da $BSL\mathbb{Z}^+$ einfach zusammenhängend ist, gilt $H_1(BSL\mathbb{Z}^+;\mathbb{Z})=0$ und der Hurewicz-Homomorphismus liefert einen Isomorphismus

$$H_2(BSL\mathbb{Z}^+;\mathbb{Z}) \cong \Pi_2 BSL\mathbb{Z}^+ \cong \mathbb{Z}/2\mathbb{Z}.$$

Lemma 1.4. $H_3(BGL\mathbb{Z};\mathbb{Z}) \cong H_3(BSL\mathbb{Z};\mathbb{Z}) \oplus 2\mathbb{Z}/2\mathbb{Z}$.

Beweis. Nach Lemma 1.2 und dem Künneth-Theorem gilt:

$$\begin{aligned} &H_3(BGL\mathbb{Z};\mathbb{Z}) \\ &\quad \cong H_3(BSL\mathbb{Z}^+;\mathbb{Z}) \oplus H_2(BSL\mathbb{Z}^+;\mathbb{Z}) \otimes \mathbb{Z}/2\mathbb{Z} \oplus H_0(BSL\mathbb{Z}^+;\mathbb{Z}) \otimes \mathbb{Z}/2\mathbb{Z} \\ &\quad \cong H_3(BSL\mathbb{Z};\mathbb{Z}) \oplus 2\mathbb{Z}/2\mathbb{Z}. \end{aligned}$$

Bemerkungen. a) Es genügt also $H_3(BSL\mathbb{Z};\mathbb{Z})$ zu bestimmen, um $H^4(BGL\mathbb{Z};\mathbb{Z})$ zu kennen. Analog wie oben können wir zeigen, daß $H_3(BSL\mathbb{Z};\mathbb{Z}) \cong H^4(BSL\mathbb{Z};\mathbb{Z})$ ist. Der nächste Satz wird diese Gruppe bestimmen.

b) Wie für $GL\mathbb{Z}$ definieren wir $c_j(SL\mathbb{Z})$ als die j-te Chern-Klasse der Inklusion $SL\mathbb{Z} \hookrightarrow GL\mathbb{C}$, $j \geqq 1$. Wegen $H^2(BSL\mathbb{Z};\mathbb{Z}) \cong H_1(BSL\mathbb{Z};\mathbb{Z})=0$ ist $c_1(SL\mathbb{Z})=0$. Sonst gelten die folgenden Resultate auch für $c_j(SL\mathbb{Z})$:

– Die Ordnung von $c_j(SL\mathbb{Z})$ ist gleich 2 für ungerade j, $j \geqq 3$.
– Die Ordnung von $c_j(SL\mathbb{Z})$ ist gleich E_j oder $2E_j$ für gerade j; insbesondere ist die Ordnung von $c_2(SL\mathbb{Z})$ gleich 12 oder 24.

Wir können nun den Hauptsatz dieses Abschnitts beweisen.

Satz 1.5. *$H^4(BSL\mathbb{Z};\mathbb{Z})$ ist eine zyklische Gruppe der Ordnung 24, erzeugt durch $c_2(SL\mathbb{Z})$.*

Beweis. a) Es bezeichne K den Raum $BSL\mathbb{Z}^+$. Weil K ein einfach zusammenhängender CW-Komplex ist, können wir folgende exakte Sequenz von Whitehead [16] verwenden:

$$\ldots \to H_{n+1}(K;\mathbb{Z}) \to \Gamma_n(K) \xrightarrow{\phi} \Pi_n K \xrightarrow{Hu} H_n(K;\mathbb{Z}) \to \ldots$$

$$\ldots \to \Gamma_3(K) \xrightarrow{\phi} \Pi_3 K \xrightarrow{Hu} H_3(K;\mathbb{Z}) \to 0 \to \Pi_2 K \xrightarrow{Hu} H_2(K;\mathbb{Z}) \to 0.$$

Hier bezeichnet $\Gamma_n(K)$ das Bild des durch die Inklusion $K^{n-1} \hookrightarrow K^n$ induzierten Homomorphismus $\Pi_n K^{n-1} \to \Pi_n K^n$, wobei K^n das n-Skelett von K ist; Hu ist der Hurewicz-Homomorphismus und ϕ wird durch die Inklusion $K^n \hookrightarrow K$ induziert. Um $H_3(K;\mathbb{Z})$ zu bestimmen, betrachten wir diese exakte Sequenz

$$\ldots \to \Gamma_3(K) \xrightarrow{\phi} \Pi_3 K \xrightarrow{Hw} H_3(K;\mathbb{Z})$$

und untersuchen das Bild von ϕ. Der Homomorphismus ϕ ist folgenderweise definiert (vgl. [16]): die Inklusionen $K^2 \overset{i}{\hookrightarrow} K^3 \overset{j}{\hookrightarrow} K$ induzieren

$$\Pi_3 K^2 \xrightarrow{i_*} \Pi_3 K^3 \xrightarrow{j_*} \Pi_3 K;$$

$\Gamma_3(K) = \text{Bild } i_*$ und ϕ ist die Zusammensetzung der Inklusion $\Gamma_3(K) \hookrightarrow \Pi_3 K^3$ mit j_*:

$$\begin{tikzcd}
\Pi_3 K^2 \arrow[r, "i_*"] \arrow[dr, "i_*"'] & \Pi_3 K^3 \arrow[r, two heads, "j_*"] & \Pi_3 K \\
& \Gamma_3(K) \arrow[u, hook] \arrow[ur, "\phi"'] &
\end{tikzcd}$$

bezeichnen wir mit α die Inklusion $K^2 \hookrightarrow K$, so gilt Bild $\phi = \text{Bild}(j_* \cdot i_*) = \text{Bild}\, \alpha_*$.

Weil $K = BSL\mathbb{Z}^+$ ein einfach zusammenhängender CW-Komplex ist, existiert eine homologische Zerlegung von K (vgl. [6]). Es ist möglich, die Homotopie-Äquivalenz $K^2 \simeq S^2$ für eine geeignete Zellenzerlegung von K zu beweisen. Deswegen ist $\Pi_3 K^2 \cong \Pi_3 S^2 \cong \mathbb{Z}$.

Es bezeichne $[\theta]$ eine beliebige Homotopieklasse von $\Pi_3 K^2$, repräsentiert durch $\theta\colon S^3 \to S^2$ und $[\alpha]$ die Klasse von $\Pi_2 K$, welche durch α repräsentiert wird. Da $\alpha_*\colon \Pi_3 K^2 \to \Pi_3 K$ durch α induziert ist, haben wir $\alpha_*([\theta]) = [\alpha] \cdot [\theta]$.

Es gilt die Gleichung $(2[\alpha]) \cdot [\theta] = 2([\alpha] \cdot [\theta]) + H_0([\theta])[[\alpha],[\alpha]]$, wobei $H_0([\theta])$ die Hopf-Invariante von $[\theta]$ und $[\,,\,]$ das Whitehead-Produkt bezeichnet (vgl. [15], S. 494). Wegen $\Pi_2 K \cong \mathbb{Z}/2\mathbb{Z}$ ist aber $2[\alpha] = 0$ und es gilt $[[\alpha],[\alpha]] = 0$, da K ein H-Raum ist (vgl. [15], S. 475). Folglich ist $2\alpha_*([\theta]) = 0$ in $\Pi_3 K$ für alle $[\theta] \in \Pi_3 K^2$, d.h. Bild $\phi = \text{Bild}\, \alpha_* \cong \mathbb{Z}/2\mathbb{Z}$ oder 0.

Die exakte Sequenz $\ldots \to \Gamma_3(K) \xrightarrow{\phi} \Pi_3 K \xrightarrow{Hu} H_3(K;\mathbb{Z})$ liefert uns dann wegen $\Pi_3 K \cong \mathbb{Z}/48\mathbb{Z}$, daß $H_3(K;\mathbb{Z})$, also auch $H^4(BSL\mathbb{Z};\mathbb{Z})$, eine zyklische Gruppe der Ordnung 24 oder 48 ist.

b) Schließlich betrachten wir die Reduktion modulo 2

$$\text{red}_2\colon H^4(BSL\mathbb{Z};\mathbb{Z}) \to H^4(BSL\mathbb{Z};\mathbb{Z}/2\mathbb{Z}),$$

deren Bild isomorph zu $\mathbb{Z}/2\mathbb{Z}$ ist. Die Chern-Klasse $c_2(SL\mathbb{Z}) \in H^4(BSL\mathbb{Z};\mathbb{Z})$ ist eigentlich so definiert: $c_2(SL\mathbb{Z}) := c_2(i^*(\gamma) \otimes \mathbb{C})$, wobei γ das reelle universelle Bündel und i die Abbildung $BSL\mathbb{Z} \to BGL\mathbb{R}$ bezeichnet. Bekanntlich gilt dann $\text{red}_2(c_2(SL\mathbb{Z})) = w_2^2(i^*(\gamma))$. Aus [13], S. 1011 ist $w_2^2(i^*(\gamma)) \neq 0$, also $\text{red}_2(c_2(SL\mathbb{Z})) \neq 0$ in $H^4(BSL\mathbb{Z};\mathbb{Z}/2\mathbb{Z})$. Sei nun ω ein erzeugendes Element von $H^4(BSL\mathbb{Z};\mathbb{Z})$ und sei $c_2(SL\mathbb{Z}) = l\omega$, so muß l eine ungerade Zahl sein.

Wir wissen aber, daß die Ordnung von $c_2(SL\mathbb{Z})$ gleich 12 oder 24 in $H^4(BSL\mathbb{Z};\mathbb{Z})$ ($\cong\mathbb{Z}/24\,\mathbb{Z}$ oder $\mathbb{Z}/48\,\mathbb{Z}$) ist. Deswegen ist $H^4(BSL\mathbb{Z};\mathbb{Z})$ eine zyklische Gruppe der Ordnung 24 und $c_2(SL\mathbb{Z})$ erzeugt diese Gruppe.

Aus diesem Satz folgt unmittelbar das folgende Korollar, welches die Frage über die Ordnung von $c_2(GL\mathbb{Z})$ beantwortet.

Korollar 1.6. *$H^4(BGL\mathbb{Z};\mathbb{Z})\cong\mathbb{Z}/24\,\mathbb{Z}\oplus 2\,\mathbb{Z}/2\,\mathbb{Z}$ und $c_2(GL\mathbb{Z})$ hat die Ordnung 24 in $H^4(BGL\mathbb{Z};\mathbb{Z})$.*

2. Die Ordnung von $c_{4k+2}(GL\mathbb{Z})$ in $H^{8k+4}(BGL\mathbb{Z};\mathbb{Z})$, $k\geqq 0$

In der Einleitung haben wir gesehen, daß $c_{4k+2}(GL\mathbb{Z})$ für alle $k\geqq 0$ die Ordnung E_{4k+2} oder $2E_{4k+2}$ besitzt. Aus der Zahlentheorie ist Folgendes bekannt: $E_{4k+2}=4S_{4k+2}$, wobei S_{4k+2} eine ungerade Zahl ist. Wir wollen nun bestimmen, ob die Ordnung von $c_{4k+2}(GL\mathbb{Z})$ gleich $4S_{4k+2}$ oder $8S_{4k+2}$ ist.

In diesem Abschnitt arbeiten wir mit Homologie- und Cohomologiegruppen mit Koeffizienten in $\mathbb{Z}/8\,\mathbb{Z}$. Wir betrachten die Reduktion modulo 8 red_8 und folgende Bezeichnungen

$$c_j^{(8)}:=\mathrm{red}_8(c_j)\in H^{2j}(BGL\mathbb{C};\mathbb{Z}/8\,\mathbb{Z}),$$
$$c_j^{(8)}(GL\mathbb{Z}):=\mathrm{red}_8(c_j(GL\mathbb{Z}))\in H^{2j}(BGL\mathbb{Z};\mathbb{Z}/8\,\mathbb{Z}).$$

Die übliche H-Raum-Struktur von $BGL\mathbb{C}$ induziert eine kommutative Ring-Struktur in $H_*(BGL\mathbb{C};\mathbb{Z})$. Der Ring $H_*(BGL\mathbb{C};\mathbb{Z})$ besitzt eine additive Basis, die aus Monomen

$$b_1^{\nu_1}\,b_2^{\nu_2}\,b_3^{\nu_3}\ldots(b_0=1, b_j\in H_{2j}(BGL\mathbb{C};\mathbb{Z}))$$

besteht, für welche das Folgende gilt: es seien die Elemente der dualen Basis von $H^*(BGL\mathbb{C};\mathbb{Z})$ mit

$$c_{(\nu_1,\nu_2,\nu_3,\ldots)}\in H^{2(\nu_1+2\nu_2+3\nu_3+\ldots)}(BGL\mathbb{C};\mathbb{Z})$$

bezeichnet (wobei nur endlich viele Zahlen $\nu_j\neq 0$ sind), dann ist $c_{(j,0,0,\ldots)}=c_j$, die j-te universelle Chern-Klasse (vgl. [1], S. 8). Die Monome $\beta_1^{\nu_1}\,\beta_2^{\nu_2}\,\beta_3^{\nu_3}\ldots$ bezeichnen schließlich $\mathrm{red}_8(b_1^{\nu_1}\,b_2^{\nu_2}\,b_3^{\nu_3}\ldots)$ in $H_{2(\nu_1+2\nu_2+3\nu_3+\ldots)}(BGL\mathbb{C};\mathbb{Z}/8\,\mathbb{Z})$; sie bilden eine additive Basis von $H_*(BGL\mathbb{C};\mathbb{Z}/8\,\mathbb{Z})$.

Es sei nun X ein Raum und R ein kommutativer Ring mit 1; für alle $n\geqq 0$ gibt es einen Homomorphismus $h\colon H^n(X;R)\to\mathrm{Hom}_R(H_n(X;R),R)$, welcher bezüglich X natürlich ist. Für alle $a\in H^n(X;R)$ bezeichne $\bar{a}$ das Bild $h(a)$ in $\mathrm{Hom}_R(H_n(X;R),R)$.

Bemerken wir an dieser Stelle, daß der Homomorphismus

$$h\colon H^n(BGL\mathbb{C};\mathbb{Z}/8\,\mathbb{Z})\to\mathrm{Hom}(H_n(BGL\mathbb{C};\mathbb{Z}/8\,\mathbb{Z}),\mathbb{Z}/8\,\mathbb{Z})$$

für alle $n\geqq 0$ ein Isomorphismus ist (vgl. [2], S. 17).

Lemma 2.1. *Die Elemente* $c_j^{(8)} \in H^{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$ *und* $\beta_1^j \in H_{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$ *sind für alle* $j \geqq 1$ *(streng) dual zueinander, d.h.* $\overline{c_j^{(8)}}(\beta_1^j)=1$ *und* $\overline{c_j^{(8)}}$ *nimmt auf allen anderen Basiselementen von* $H_{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$ *den Wert* 0 *an.*

Beweis. Wir betrachten das kommutative Diagramm

$$\begin{tikzcd}
H^{2j}(BGL\mathbb{C};\mathbb{Z}) \arrow[r, "h", "\cong"'] \arrow[d, "\mathrm{red}_8"] & \mathrm{Hom}(H_{2j}(BGL\mathbb{C};\mathbb{Z}),\mathbb{Z}) \arrow[d, "\mathrm{red}_8"] \\
H^{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z}) \arrow[r, "h'", "\cong"'] \arrow[dr, "h", "\cong"'] & \mathrm{Hom}(H_{2j}(BGL\mathbb{C};\mathbb{Z}),\mathbb{Z}/8\mathbb{Z}) \\
 & \mathrm{Hom}(H_{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z}),\mathbb{Z}/8\mathbb{Z}). \arrow[u, "\mathrm{red}_8^*"']
\end{tikzcd}$$

Dabei wird der Isomorphismus h' von dem universellen Koeffizienten-Theorem gegeben und der Homomorphismus red_8^* durch

$$\mathrm{red}_8\colon H_{2j}(BGL\mathbb{C};\mathbb{Z}) \to H_{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$$

induziert.

Die Elemente $c_j \in H^{2j}(BGL\mathbb{C};\mathbb{Z})$ und $b_1^j \in H_{2j}(BGL\mathbb{C};\mathbb{Z})$ sind (streng) dual zueinander, d.h. $\overline{c_j}(b_1^j)=1$ und $\overline{c_j}$ hat den Wert 0 auf allen anderen Basiselementen von $H_{2j}(BGL\mathbb{C};\mathbb{Z})$. Die Kommutativität des oberen Teils des Diagramms liefert:

$$h'(c_j^{(8)})(b_1^j)=\mathrm{red}_8(\overline{c_j}(b_1^j))=1 \in \mathbb{Z}/8\mathbb{Z}$$

und $h'(c_j^{(8)})$ nimmt den Wert 0 auf allen anderen Basiselementen an. Wegen der Kommutativität des unteren Teils des Diagramms gilt dann:

$$\begin{aligned}\overline{c_j^{(8)}}(\beta_1^j) &= \overline{c_j^{(8)}}(\mathrm{red}_8(b_1^j)) = \mathrm{red}_8^*(\overline{c_j^{(8)}})(b_1^j)\\ &= h'(c_j^{(8)})(b_1^j)=1 \in \mathbb{Z}/8\mathbb{Z}\end{aligned}$$

und $\overline{c_j^{(8)}}$ ist 0 auf allen anderen Basiselementen von $H_{2j}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$.

Um den Hauptsatz dieses Abschnitts zu beweisen, werden wir folgende Idee benutzen: $\beta_1^2 \in H_4(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$ ist im wesentlichen das Bild eines Elementes von $H_4(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z})$ unter dem durch die Inklusion $GL\mathbb{Z} \hookrightarrow GL\mathbb{C}$ induzierten Homomorphismus

$$i_*\colon H_*(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z}) \to H_*(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z}).$$

Genauer ist es möglich, das nächste Lemma mit Hilfe von Satz 1.5 zu beweisen (vgl. [2], S. 20).

Lemma 2.2. *Es gibt ein Element* $\xi \in H_4(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z})$, *so daß*

$$i_*(\xi)=\beta_1^2+l\,\beta_2 \in H_4(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z}) \quad \textit{mit } l=0 \textit{ oder } 4.$$

Lemma 2.3. *Seien X und Y zwei Räume, R ein kommutativer Ring mit 1, h der oben eingeführte Homomorphismus und f eine Abbildung $X \to Y$. Betrachten wir die durch f induzierte Homomorphismen $f_*\colon H_*(X;R) \to H_*(Y;R)$ und $f^*\colon H^*(Y;R) \to H^*(X;R)$ und die Elemente $a \in H^n(Y;R)$ und $b \in H_n(X;R)$. Dann gilt: $\overline{f^*(a)}(b) = \bar{a}(f_*(b))$.*

Beweis. Wegen der Natürlichkeit von h ist das folgende Diagramm kommutativ:

$$\begin{tikzcd} H^n(Y;R) \arrow[r, "h"] \arrow[d, "f^*"] & \operatorname{Hom}_R(H_n(Y;R),R) \arrow[d, "f^*"] \\ H^n(X;R) \arrow[r, "h"] & \operatorname{Hom}_R(H_n(X;R),R) \end{tikzcd}$$

d.h. $\overline{f^*(a)}(b) = (f^{\#}\,\bar{a})(b)$. Dabei ist $f^{\#}$ der durch $f_*\colon H_n(X;R) \to H_n(Y;R)$ induzierte Homomorphismus, d.h. $(f^{\#}\,\bar{a})(b) = \bar{a}(f_*(b))$. Daraus folgt die Behauptung.

Bemerkung. Die Elemente $a \in H^n(X;R)$ und $b \in H_n(X;R)$ heißen *schwach dual* zueinander, falls $\bar{a}(b) = 1 \in R$. Setzen wir $R = \mathbb{Z}/8\mathbb{Z}$; sind nun a und b schwach dual zueinander, so besitzen $a \in H^n(X;\mathbb{Z}/8\mathbb{Z})$ und $b \in H_n(X;\mathbb{Z}/8\mathbb{Z})$ die Ordnung 8.

Wir können nun den Hauptsatz des Abschnitts beweisen.

Satz 2.4. *$c_{2m}^{(8)}(GL\mathbb{Z})$ hat die Ordnung 8 in $H^{4m}(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z})$ für alle $m \geqq 1$.*

Beweis. Wir setzen $X = BGL\mathbb{Z}$, $Y = BGL\mathbb{C}$, $R = \mathbb{Z}/8\mathbb{Z}$ und $f = i\colon BGL\mathbb{Z} \to BGL\mathbb{C}$, die Abbildung, welche durch die Inklusion $GL\mathbb{Z} \hookrightarrow GL\mathbb{C}$ induziert wird. Nach Lemma 2.2 existiert $\xi \in H_4(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z})$ mit $i_*(\xi) = \beta_1^2 + l\beta_2$ ($l = 0$ oder 4). $H_*(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z}) \cong H_*(BGL\mathbb{Z}^+;\mathbb{Z}/8\mathbb{Z})$ besitzt eine kommutative Ring-Struktur, welche durch die H-Raum-Struktur von $BGL\mathbb{Z}^+$ induziert wird und der Homomorphismus

$$i_*\colon H_*(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z}) \to H_*(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$$

ist ein Ring-Homomorphismus.

Wir können also das Element $\xi^m \in H_{4m}(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z})$ betrachten. Dann ist

$$i_*(\xi^m) = (\beta_1^2 + l\beta_2)^m = \beta_1^{2m} + m\,l\,\beta_1^{2m-2}\beta_2 \in H_{4m}(BGL\mathbb{C};\mathbb{Z}/8\mathbb{Z})$$

wegen $l^2 \equiv 0 \bmod 8$.

Nach Lemma 2.1 sind $c_{2m}^{(8)}$ und β_1^{2m} (streng) dual zueinander; daraus folgt:

$$\overline{c_{2m}^{(8)}}(i_*(\xi^m)) = \overline{c_{2m}^{(8)}}(\beta_1^{2m} + m\,l\,\beta_1^{2m-2}\beta_2) = 1 \in \mathbb{Z}/8\mathbb{Z}.$$

Dann ist auch $\overline{i^*(c_{2m}^{(8)})}(\xi^m) = 1 \in \mathbb{Z}/8\mathbb{Z}$ nach Lemma 2.3. Das Element $i^*(c_{2m}^{(8)}) = c_{2m}^{(8)}(GL\mathbb{Z})$ hat also die Ordnung 8 in $H^{4m}(BGL\mathbb{Z};\mathbb{Z}/8\mathbb{Z})$ für alle $m \geqq 1$.

Korollar 2.5. *$c_{4k+2}(GL\mathbb{Z})$ besitzt die Ordnung $2E_{4k+2}$ in $H^{8k+4}(BGL\mathbb{Z};\mathbb{Z})$ für alle $k \geqq 0$.*

Beweis. Nach Satz 2.4 ist die Ordnung von $c_{2m}(GL\mathbb{Z})$ ein positives Vielfaches von 8 für alle $m \geqq 1$. Wir wissen schon, daß $c_{4k+2}(GL\mathbb{Z})$ ein Element der Ordnung $E_{4k+2}=4S_{4k+2}$ oder $2E_{4k+2}=8S_{4k+2}$ für alle $k\geqq 0$ ist, wobei S_{4k+2} eine ungerade Zahl ist. Das liefert die Behauptung.

Bemerkungen. a) Dasselbe Resultat gilt auch für die Chern-Klassen von $SL\mathbb{Z}$: $c_{4k+2}(SL\mathbb{Z})$ besitzt die Ordnung $2E_{4k+2}$ in $H^{8k+4}(BSL\mathbb{Z};\,\mathbb{Z})$, $k\geqq 0$.

b) Die Ordnung von $c_j(GL\mathbb{Z})$ in $H^{2j}(BGL\mathbb{Z};\mathbb{Z})$ ist also jetzt bekannt, falls $j\not\equiv 0 \mod 4$ ist. Es bleibt noch eine offene Frage: ist die Ordnung von $c_{4k}(GL\mathbb{Z})$ gleich E_{4k} oder $2E_{4k}$ $(k\geqq 1)$?

3. Chern-Klassen von Kongruenzuntergruppen

In diesem dritten Teil betrachten wir die *Kongruenzuntergruppen* Γ_m von $SL_n\mathbb{Z}$, welche so definiert sind: Γ_m ist der Kern der natürlichen Projektion $SL_n\mathbb{Z} \twoheadrightarrow SL_n\mathbb{Z}/m\mathbb{Z}$, $m\geqq 2$. Dabei ist n immer groß genug vorausgesetzt, so daß die singulären Homologie- und Cohomologiegruppen von $BSL_n\mathbb{Z}$, $BGL_n\mathbb{Z}$ und $BGL_n\mathbb{F}_p$ (für alle Primzahlen p) im Stabilitätsbereich liegen (vgl. [4]); diese Voraussetzung gilt für den ganzen Abschnitt. Es ist noch zu bemerken, daß die Gruppe Γ_m torsionsfrei ist, falls $m\neq 2$ ist. Es bezeichne i_m die Inklusion $\Gamma_m \hookrightarrow GL_n\mathbb{Z}$ und i_m^* den induzierten Homomorphismus $H^*(BGL_n\mathbb{Z};\mathbb{Z})\to H^*(B\Gamma_m;\mathbb{Z})$. Für $j\geqq 1$ definieren wir die j-te *Chern-Klasse der Kongruenzuntergruppe* Γ_m: $c_j(\Gamma_m):=i_m^*(c_j(GL_n\mathbb{Z}))$.

Für alle $m\geqq 2$ gilt zum Beispiel $c_1(\Gamma_m)=0$, weil

$$H^2(BSL\mathbb{Z};\mathbb{Z})\cong H_1(BSL\mathbb{Z};\mathbb{Z})=0$$

ist. Wir interessieren uns für die Ordnung dieser Chern-Klassen $c_j(\Gamma_m)$, $j\geqq 2$, und untersuchen in diesem Abschnitt die Ordnung von $c_2(\Gamma_m)$ für alle $m\geqq 2$.

Sei nun p eine Primzahl und $\mathbb{F}_p$ der Körper mit p Elementen. Für den nächsten Satz müssen wir den Hurewicz-Homomorphismus Hu: $K_3\mathbb{F}_p\to H_3(BGL\mathbb{F}_p^+;\mathbb{Z})$ kennen. Nach Quillen ist $K_3\mathbb{F}_p\cong\mathbb{Z}/(p^2-1)\mathbb{Z}$ und nach [11] gilt

$$H_3(BGL\mathbb{F}_p^+;\mathbb{Z})\cong\mathbb{Z}/(p^2-1)\mathbb{Z}\oplus 2\mathbb{Z}/(p-1)\mathbb{Z}.$$

Sei π: $BGL\mathbb{F}_p^+\to BU$ der Brauer-Lift und c_n die n-te universelle Chern-Klasse in $H^{2n}(BU;\,\mathbb{Z})$; wir definieren $\hat{c}_n:=\pi^*(c_n)\in H^{2n}(BGL\mathbb{F}_p^+;\,\mathbb{Z})$. Die Ordnung von $\hat{c}_n$ ist gleich p^n-1 (vgl. [9], S. 45 oder [11], Theorem B). Zum Beispiel hat $\hat{c}_2$ die Ordnung p^2-1.

Wegen

$$\begin{aligned}H^4(BGL\mathbb{F}_p^+;\mathbb{Z})&\cong\mathrm{Ext}(H_3(BGL\mathbb{F}_p^+;\mathbb{Z}),\mathbb{Z})\\&\cong\mathrm{Hom}(H_3(BGL\mathbb{F}_p^+;\mathbb{Z}),\,\mathbb{Q}/\mathbb{Z})\end{aligned}$$

können wir $\hat{c}_2$ als ein Element der Ordnung p^2-1 in $\mathrm{Hom}(H_3(BGL\mathbb{F}_p^+;\,\mathbb{Z}),\,\mathbb{Q}/\mathbb{Z})$ interpretieren. Sei $\hat{c}_2^*$ ein Element von $H_3(BGL\mathbb{F}_p^+;\,\mathbb{Z})$ mit

$$\hat{c}_2(\hat{c}_2^*)=\frac{1}{p^2-1}\in\mathbb{Q}/\mathbb{Z};$$

$\hat{c}_2^*$ besitzt natürlich auch die Ordnung p^2-1. Es gilt (vgl. [2], S. 29):

Lemma 3.1. *Es gibt ein erzeugendes Element* α *von* $K_3\mathbb{F}_p \cong \mathbb{Z}/(p^2-1)\mathbb{Z}$ *mit*

$$\mathrm{Hu}(\alpha) = \hat{c}_2^* + \eta \in H_3(BGL\mathbb{F}_p^+; \mathbb{Z}),$$

wobei $(p-1)\,\eta = 0$ *ist.*

Satz 3.2. $c_2(\Gamma_p) = 0$ *für alle Primzahlen* p *mit* $p \neq 2$, $p \neq 3$.

Beweis. a) Sei p eine Primzahl mit $p \neq 2$, $p \neq 3$; die natürliche Projektion π_p: $GL\mathbb{Z} \to GL\mathbb{F}_p$ induziert den Homomorphismus π_{p_*}: $K_3\mathbb{Z} \to K_3\mathbb{F}_p$. Es ist möglich zu zeigen (vgl. [2], S. 34), daß für ein beliebiges erzeugendes Element ε von $K_3\mathbb{Z} \cong \mathbb{Z}/48\mathbb{Z}$ die Ordnung von $\pi_{p_*}(\varepsilon)$ in $K_3\mathbb{F}_p \cong \mathbb{Z}/(p^2-1)\mathbb{Z}$ gleich 24 ist; sei $\lambda := \frac{p^2-1}{24} \in \mathbb{N}$, so ist $\pi_{p_*}(\varepsilon) = l\lambda\alpha$, wobei $(l, 24) = 1$ und α das erzeugende Element von $K_3\mathbb{F}_p$ ist, welches wir wie in Lemma 3.1 wählen.

b) Betrachten wir nun das kommutative Diagramm

$$\begin{array}{ccc}
K_3\mathbb{Z} & \xrightarrow{\pi_{p_*}} & K_3\mathbb{F}_p \\
\Big\downarrow{\scriptstyle \mathrm{Hu}} & & \Big\downarrow{\scriptstyle \mathrm{Hu}} \\
H_3(BGL\mathbb{Z}^+;\mathbb{Z}) & \xrightarrow{\pi_{p_*}} & H_3(BGL\mathbb{F}_p^+;\mathbb{Z}).
\end{array}$$

Wegen

$$\begin{aligned}
H^4(BGL\mathbb{Z}^+;\mathbb{Z}) &\cong \mathrm{Ext}(H_3(BGL\mathbb{Z}^+;\mathbb{Z}),\mathbb{Z}) \\
&\cong \mathrm{Hom}(H_3(BGL\mathbb{Z}^+:\mathbb{Z}),\mathbb{Q}/\mathbb{Z})
\end{aligned}$$

können wir $c_2(GL\mathbb{Z})$ als ein Element der Ordnung 24 in $\mathrm{Hom}(H_3(BGL\mathbb{Z}^+;\mathbb{Z}),\mathbb{Q}/\mathbb{Z})$ interpretieren. Sei $c_2^*(GL\mathbb{Z})$ ein Element von $H_3(BGL\mathbb{Z}^+;\mathbb{Z})$ mit

$$c_2(GL\mathbb{Z})(c_2^*(GL\mathbb{Z})) = \tfrac{1}{24} \in \mathbb{Q}/\mathbb{Z}.$$

Nach Satz 1.5 gibt es ein erzeugendes Element ε von $K_3\mathbb{Z}$ mit $\mathrm{Hu}(\varepsilon) = c_2^*(GL\mathbb{Z})$. Die Kommutativität liefert dann:

$$\pi_{p_*}(c_2^*(GL\mathbb{Z})) = \pi_{p_*}(\mathrm{Hu}(\varepsilon)) = \mathrm{Hu}(\pi_{p_*}(\varepsilon)) = \mathrm{Hu}(l\lambda\alpha) = l\lambda(\hat{c}_2^* + \eta)$$

wegen Lemma 3.1 ($(p-1)\eta = 0$). Also hat $\pi_{p_*}(c_2^*(GL\mathbb{Z}))$ die Ordnung 24 in $H_3(BGL\mathbb{F}_p^+;\mathbb{Z})$.

Wenn wir dann π_p^*: $H^4(BGL\mathbb{F}_p^+;\mathbb{Z}) \to H^4(BGL\mathbb{Z}^+;\mathbb{Z})$ betrachten, folgt daraus, daß $\pi_p^*(\hat{c}_2)$ in $H^4(BGL\mathbb{Z}^+;\mathbb{Z}) \cong \mathbb{Z}/24\mathbb{Z} \oplus 2\mathbb{Z}/2\mathbb{Z}$ ebenfalls die Ordnung 24 besitzt, d.h. $\pi_p^*(\hat{c}_2) = kc_2(GL\mathbb{Z}) + t$ mit $(k, 24) = 1$ und $2t = 0$. Folglich gilt $\pi_p^*(k\hat{c}_2) = c_2(GL\mathbb{Z}) + kt$ wegen $k^2 \equiv 1 \bmod 24$.

c) Wir betrachten schließlich die exakte Sequenz

$$\Gamma_p \xrightarrow{i_p} GL_n\mathbb{Z} \xrightarrow{\pi_p} GL_n\mathbb{F}_p.$$

Der Homomorphismus i_p^* ist eigentlich die Zusammensetzung

$$\begin{array}{ccccc}
H^4(BGL_n\mathbb{Z};\mathbb{Z}) & \longrightarrow & H^4(BSL_n\mathbb{Z};\mathbb{Z}) & \longrightarrow & H^4(B\Gamma_p;\mathbb{Z}). \\
\Big\| \wr & & \Big\| \wr & & \\
\mathbb{Z}/24\mathbb{Z} \oplus 2\mathbb{Z}/2\mathbb{Z} & & \mathbb{Z}/24\mathbb{Z} & &
\end{array}$$

Deswegen gilt $i_p^*(c_2(GL_n\mathbb{Z})+kt)=i_p^*(c_2(GL_n\mathbb{Z}))+0=c_2(\Gamma_p)$, also $i_p^*\cdot\pi_p^*(k\hat{c}_2)=c_2(\Gamma_p)$. Die Zusammensetzung $\pi_p\cdot i_p$ ist trivial, also auch $i_p^*\cdot\pi_p^*$, was uns die Behauptung liefert: $c_2(\Gamma_p)=0$.

(Dieser Satz folgt auch aus [5].)

Analog können wir folgenden Satz beweisen.

Satz 3.3.

$$\text{a) } 3c_2(\Gamma_3)=0,$$

$$\text{b) } 8c_2(\Gamma_2)=0.$$

Bemerkung. Da Γ_2 eine zyklische Untergruppe der Ordnung 2 enthält, ist die Ordnung von $c_2(\Gamma_2)$ in $H^4(B\Gamma_2;\mathbb{Z})$ ein positives Vielfaches von 2, also gleich 2, 4 oder 8.

Wir besitzen nun einige Informationen über die Ordnung von $c_2(\Gamma_m)$, wenn m eine Primzahl ist. Damit können wir Resultate über den allgemeinen Fall $m\geqq 2$ leicht bekommen (vgl. [2], S. 40–41), zum Beispiel:

Korollar 3.4. $c_2(\Gamma_m)=0$ *für alle natürlichen Zahlen m, welche keine Zweierpotenz und keine Dreierpotenz sind.*

Bemerkung. Bei dieser Untersuchung der Ordnung von $c_2(\Gamma_m)$, $m\geqq 2$, bleiben also noch zwei Fragen offen:
- Ist die Ordnung von $c_2(\Gamma_2)$ gleich 2, 4 oder 8 in $H^4(B\Gamma_2;\mathbb{Z})$?
- Ist $c_2(\Gamma_3)$ gleich 0 oder ein Element der Ordnung 3 in $H^4(B\Gamma_3;\mathbb{Z})$?

Die zweite Frage ist besonders interessant, weil Γ_3 eine torsionsfreie Gruppe ist. Wäre die Ordnung von $c_2(\Gamma_3)$ gleich 3, dann hätten wir ein Beispiel einer ganzzahligen treuen Darstellung einer torsionsfreien Gruppe, deren zweite Chern-Klasse nicht Null ist. Ein solches Beispiel würden wir auch bekommen, wenn für m eine Dreierpotenz oder eine von 2 verschiedene Zweierpotenz $c_2(\Gamma_m)\neq 0$ wäre.

Anhang: Chern-Klassen von rationalen Darstellungen diskreter Gruppen

Wir haben die Chern-Klassen der ganzzahligen Darstellungen diskreter Gruppen untersucht. Wir können uns auch für rationale Darstellungen von diskreten Gruppen interessieren und die beste obere Schranke für die Ordnung ihrer Chern-Klassen suchen.

Für alle $j\geqq 1$ definieren wir $c_j(GL\mathbb{Q})$ bzw. $c_j(SL\mathbb{Q})$ als die j-te Chern-Klasse der Inklusion $GL\mathbb{Q}\hookrightarrow GL\mathbb{C}$ bzw. $SL\mathbb{Q}\hookrightarrow GL\mathbb{C}$, wobei wir die Gruppen $GL\mathbb{Q}$ und $SL\mathbb{Q}$ als diskrete Gruppen auffassen; wir probieren Aussagen über die Ordnung dieser Chern-Klassen zu bekommen. Wie vorher ist es hier auch klar, daß die Chern-Klassen $c_j(GL\mathbb{Q})$ die Ordnung 2 besitzen, falls j eine ungerade Zahl ist.

Das Problem ist aber schwieriger, falls j gerade ist. Dazu betrachten wir die *profiniten Chern-Klassen* $\hat{c}_j(GL\mathbb{Q})$, welche folgenderweise definiert sind (vgl. [8]). Es bezeichne $\hat{\mathbb{Z}}$ den Ring der profiniten ganzen Zahlen ($\hat{\mathbb{Z}}=\varprojlim\mathbb{Z}/m\mathbb{Z}$)

und ϕ den natürlichen Homomorphismus $\mathbb{Z} \to \hat{\mathbb{Z}}$, welcher

$$\phi_*: H^{2j}(BGL\mathbb{Q}; \mathbb{Z}) \to H^{2j}(BGL\mathbb{Q}; \hat{\mathbb{Z}})$$

induziert. Dann ist $\hat{c}_j(GL\mathbb{Q}) := \phi_*(c_j(GL\mathbb{Q}))$. Nach [8] gilt $2E_j \hat{c}_j(GL\mathbb{Q}) = 0$, d.h. $2E_j c_j(GL\mathbb{Q}) \in \text{Kern } \phi_*$ für gerade j. Es ist einfach zu sehen, daß Kern ϕ_* die Menge aller ∞-divisiblen Elemente von $H^{2j}(BGL\mathbb{Q}; \mathbb{Z})$ ist. Wir bekommen also den folgenden Satz.

Satz 4.1. *Für alle geraden Zahlen j gilt: die Elemente $2E_j c_j(GL\mathbb{Q})$ in $H^{2j}(BGL\mathbb{Q}; \mathbb{Z})$ und $2E_j c_j(SL\mathbb{Q})$ in $H^{2j}(BSL\mathbb{Q}; \mathbb{Z})$ sind ∞-divisibel.*

Das Element 0 ist natürlich immer ∞-divisibel. Um die Ordnung von $c_{2k}(SL\mathbb{Q})$ und $c_{2k}(GL\mathbb{Q})$ zu kennen, untersuchen wir, ob die Gruppen $H^{4k}(BSL\mathbb{Q}; \mathbb{Z})$ und $H^{4k}(BGL\mathbb{Q}; \mathbb{Z})$ andere ∞-divisible Elemente besitzen $(k \geqq 1)$.

Lemma 4.2. *Sei T eine Torsionsgruppe. Dann enthält die Gruppe* $\text{Ext}(T; \mathbb{Z})$ *keine ∞-divisiblen Elemente außer* 0.

Beweis. Da T eine Torsionsgruppe ist, gilt $\text{Ext}(T, \mathbb{Z}) \cong \text{Hom}(T, \mathbb{Q}/\mathbb{Z})$. Sei f ein ∞-divisibles Element von $\text{Hom}(T, \mathbb{Q}/\mathbb{Z})$; für alle natürlichen Zahlen m existiert also ein $g_m \in \text{Hom}(T, \mathbb{Q}/\mathbb{Z})$, so daß $f = m g_m$. Sei nun x ein beliebiges Element der Torsionsgruppe T und s seine Ordnung: $s \cdot x = 0$. Wegen $f = s g_s$ nimmt f an der Stelle x den Wert $f(x) = s g_s(x) = g_s(sx) = g_s(0) = 0$ an. Es gilt also: $f(x) = 0$ für alle $x \in T$. Das einzige ∞-divisible Element von $\text{Hom}(T, \mathbb{Q}/\mathbb{Z}) \cong \text{Ext}(T, \mathbb{Z})$ ist 0.

Korollar 4.3. *Sei X ein Raum. Falls $H_{i-1}(X; \mathbb{Q}) = 0$, dann enthält $H^i(X; \mathbb{Z})$ keine ∞-divisiblen Elemente außer* 0.

Beweis. Wegen $H_{i-1}(X; \mathbb{Q}) = 0$ ist $H_{i-1}(X; \mathbb{Z})$ eine Torsionsgruppe. Nach dem universellen Koeffizienten-Theorem gilt

$$H^i(X; \mathbb{Z}) \cong \text{Hom}(H_i(X; \mathbb{Z}), \mathbb{Z}) \oplus \text{Ext}(H_{i-1}(X; \mathbb{Z}), \mathbb{Z}).$$

Die von 0 verschiedenen Elemente von $\text{Hom}(H_i(X; \mathbb{Z}), \mathbb{Z})$ nehmen ihre Werte in $\mathbb{Z}$ an und sind deshalb nicht ∞-divisibel. Die Behauptung folgt dann aus Lemma 4.2.

Also fragen wir uns: gibt es natürliche Zahlen k mit $H_{4k-1}(BSL\mathbb{Q}; \mathbb{Q}) = 0$? Dazu betrachten wir ein Resultat von Borel [3]: $H_*(BSL\mathbb{Q}; \mathbb{Q})$ ist eine äußere Algebra:

$$H_*(BSL\mathbb{Q}; \mathbb{Q}) = \wedge(x_5, x_9, x_{13}, \ldots, x_{4l+1}, \ldots)$$

mit Grad $x_{4l+1} = 4l+1$ für $l \geqq 1$.

Für $k = 1, 2, 3, 4, 5$ und 6 ist folglich $H_{4k-1}(BSL\mathbb{Q}; \mathbb{Q}) = 0$, also besitzt $H^{4k}(BSL\mathbb{Q}; \mathbb{Z})$ keine ∞-divisiblen Elemente außer 0; wegen Satz 4.1 ist dann $2E_{2k}\, c_{2k}(SL\mathbb{Q}) = 0$. Da die Ordnung von $c_{2k}(SL\mathbb{Q})$ notwendigerweise ein positives Vielfaches der Ordnung von $c_{2k}(SL\mathbb{Z})$ ist, liefern die zwei ersten Teile dieser Arbeit folgendes Resultat.

Satz 4.4. a) *Die Ordnung von* $c_{2k}(SL\mathbb{Q})$ *ist gleich* $2E_{2k}$ *für* $k=1, 3, 5$. b) *Die Ordnung von* $c_{2k}(SL\mathbb{Q})$ *ist gleich* E_{2k} *oder* $2E_{2k}$ *für* $k=2, 4, 6$.

Bemerkungen. a) Für $k \geqq 7$ ist $H_{4k-1}(BSL\mathbb{Q};\ \mathbb{Q}) \neq 0$, weil $\mathrm{Grad}(x_5 \wedge x_9 \wedge x_{4(k-4)+1}) = 4k-1$ ist. Unsere Methode liefert also keine obere Schranke für die Ordnung von $c_{2k}(SL\mathbb{Q})$, falls $k \geqq 7$ ist.

b) Als Hilfsmittel für die Untersuchung der Ordnung von $c_{2k}(GL\mathbb{Q})$ können wir analog zu Lemma 1.2 die Homotopie-Äquivalenz $BGL\mathbb{Q}^+ \simeq BSL\mathbb{Q}^+ \times B\mathbb{Q}^*$ beweisen. Damit ist es zum Beispiel möglich zu zeigen, daß die Ordnung von $c_2(GL\mathbb{Q})$ in $H^4(BGL\mathbb{Q};\ \mathbb{Z})$ gleich 24 ist (vgl. [2], S. 48).

Literatur

1. Adams, J.F.: Stable Homotopy and Generalised Homology. Chicago Lectures in Mathematics Series (1974)
2. Arlettaz, D.: Chern-Klassen von ganzzahligen und rationalen Darstellungen diskreter Gruppen. Dissertation Nr. 7301 ETH Zürich (1983)
3. Borel, A.: Cohomologie réelle stable de groupes S-arithmétiques classiques. C.R. Acad. Sci. Paris Sér A **274**, 1700–1702 (1972)
4. Charney, R.M.: Homology stability of GL_n of a Dedekind domain. Bull. Amer. Math. Soc. (N.S.) **1** 428–431 (1979)
5. Deligne, P., Sullivan, D.: Fibrés vectoriels complexes à groupe structural discret. C.R. Acad. Sci. Paris Sér. A **281**, 1081–1083 (1975)
6. Eckmann, B., Hilton, P.J.: On the homology and homotopy decomposition of continous maps. Proc. Nat. Acad. Sci. USA **45**, 372–375 (1959)
7. Eckmann, B., Mislin, G.: Chern classes of group representations over a number field. Compositio Math. **44**, 41–65 (1981)
8. Eckmann, B., Mislin, G.: Profinite Chern classes for group representations. Topological Topics. London Math. Soc. Lecture Note Ser. **86**. Cambridge: Cambridge University Press 1983
9. Fiedorowicz, Z., Priddy, S.: Homology of classical groups over finite fields and their associated infinite loop spaces. Lecture Notes in Math. **674**. Berlin-Heidelberg-New York: Springer 1978
10. Grothendieck, A.: Classes de Chern et représentations linéaires des groupes discrets. Dans: Dix exposés sur la cohomologie des schémas. Amsterdam-New York-Oxford: North-Holland 1968
11. Hübschmann, J.: The cohomology of $F\Psi^q$, the additive structure. Preprint, Forschungsinstitut für Mathematik ETH Zürich (1982)
12. Lee, R., Szczarba, R.H.: The group $K_3(\mathbb{Z})$ is cyclic of order forty-eight. Ann. of Math. **104**, 31–60 (1976)
13. Soulé, C.: Classes de torsion dans la cohomologie des groupes arithmétiques. C.R. Acad. Sci. Paris Sér. A, **284**, 1009–1011 (1977)
14. Wagoner, J.B.: Delooping classifying spaces in algebraic K-theory. Topology **11**, 349–370 (1972)
15. Whitehead, G.W.: Elements of Homotopy Theory. Graduate Texts in Math. **61**. Berlin-Heidelberg-New York: Springer 1978
16. Whitehead, J.H.C.: A certain exact sequence. Ann. of Math. **52**, 51–110 (1950)

Received November 25, 1983

Math. Z. 187, 61-73 (1984)

Mathematische Zeitschrift

Analytic Solutions of Some Nonlinear Diffusion Equations

Kyûya Masuda

Mathematical Institute, Tohoku University, Sendai, Japan, 980

1. Introduction

Consider the nonlinear diffusion equation of the form:

$$\frac{\partial u}{\partial t} = \Delta u + f(u), \qquad x \in \Omega,\ t > 0 \tag{1}$$

subject to the Neumann boundary condition

$$\partial_n u|_\Gamma = 0 \tag{2}$$

and with the initial condition

$$u|_{t=0} = u_0 \tag{3}$$

where Ω is a bounded domain of R^n with smooth (e.g. C^3-) boundary Γ, $f(z)$ is an entire function of z, and u_0 is a given function in $W_p^2(\Omega)$, $p > n$, satisfying the compatibility condition (2); Δ denotes the Laplace operator, ∂_n denotes the differentiation in the direction of external normal to Γ, $W_p^k(\Omega)$ is the Sobolev space (the norm is denoted by $\|\cdot\|_{k,p}$ and $\|\cdot\|_p$ denotes the L^p-norm).

There is an extensive literature on existence or nonexistence (blow-up) of solutions of (1) under various condition on f, u_0 (see H. Fujita [4], M. Tsutsumi [18], F. Weissler [10] and the references given in those papers). The solution of (1), (2), (3) exists, generally, only locally in time, and blows up at some finite positive $t = t^*$. To see this, let us consider f of the special form: $f(u) = u^2$.

$$\frac{\partial u}{\partial t} = \Delta u + u^2, \qquad x \in \Omega. \tag{4}$$

In this case, we know that if $u_0 \geqq 0$ ($u_0 \not\equiv 0$), then a solution u of (4), (2), (3) blows up at some $t = t^*$, $0 < t^* \leqq 1/a$, ($a = Pu_0$), where Pw is the average of w over Ω:

$$Pw = \frac{1}{|\Omega|} \int_\Omega w(x)\, dx. \tag{5}$$

This can be easily seen from the inequality:

$$dPu(t)/dt \geqq (Pu(t))^2, \qquad Pu(0)=a$$

($Pu(t)$ is the average of $u(\cdot, t)$ over Ω). In particular, if the initial function u_0 is a constant function, then the solution u of (4), (2), (3) is given by $u=y(t)$:

$$y(t)=1/(t_0-t), \qquad t_0=1/a, \tag{6}$$

(note $a=Pu_0=u_0$). Clearly the solution $y(t)$ of (4), (2), (3) is analytic in $(0, t_0)$, blows up at $t=t_0$, and can be, however, analytically continued in t into the infinite interval (t_0, ∞).

In this paper we are concerned with the problem of whether or not similar phenomena occur for solutions of (4), (2) with non-constant initial functions.

Let $0<\theta<\frac{\pi}{2}$, and a complex domain D_1 be as shown in Fig. 1 (δ is sufficiently small positive number); and D_2 is the complex conjugate of D_1. More precisely,

$$D_1=\{t\in\mathbb{C};\ |\arg t|<\theta, |\arg(t-t_0-i\delta\tan\theta)-\tfrac{\pi}{2}|<\theta+\tfrac{\pi}{2}\}.$$

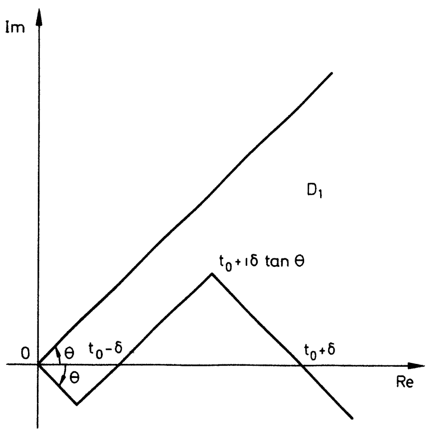


Fig. 1

Theorem 1.1. *Let u_0 be a non-negative function ($\not\equiv 0$) in $W_p^2(\Omega)$, $p>n$, and set $a=Pu_0$. If $\|\partial_x^2 u_0\|_p/|a|^2$ is sufficiently small, then there exists a unique solution u_j ($j=1, 2$) of (4), (2), (3) which is analytic in $t\in D_j$ as a $W_p^2(\Omega)$-valued function and converges to u_0 as $t\to 0$, $|\arg t|<\theta$, in the norm of $W_p^2(\Omega)$.*

If $t\in D_1\cap D_2$ with $\operatorname{Re} t<t_0-\delta$, then

$$u_1(\cdot, t)=u_2(\cdot, t) \tag{7}$$

by the uniqueness of solution of (4), (2), (3). (7) is also true for $\operatorname{Re} t>t_0+\delta$ if the initial function is constant as can be seen from (6). However it is not true for $\operatorname{Re} t>t_0+\delta$ if the initial function is not constant.

More precisely we can show:

Theorem 1.2. *Let u_j $(j=1,2)$ be as in Theorem 1.1. If $\|\partial_x^2 u_0\|_p/|a|^2$ is sufficiently small, and if (7) holds for some $t \in D_1 \cap D_2$ with $\operatorname{Re} t > t_0 + \delta$, then u_0 is a constant function.*

The outlines of the proofs of Theorems 1.1 and 1.2 are given in K. Masuda [14].

In the Sect. 2 we shall consider the problems posed above in the abstract setting. In the Sect. 3, as applications of the abstract theorems established in the Sect. 2, we shall prove Theorems 1.1 and 1.2, and also show some other examples: $f(u) = u + u^2$, $f(u) = e^u$.

2. Abstract Theory

2.1. We first introduce the notion of a θ-domain; θ is a real number such that $0 < \theta < \frac{\pi}{2}$. We shall call D a θ-domain if it is a simply connected domain in the complex plane and if any point t in D can be joined to the origin 0 by some C^1-curve $\Gamma(t)$, contained in D except for $t=0$, such that angle between the positive real axis and tangent to the curve $\Gamma(t)$ is less than or equal to θ; and $\Gamma(t)$, oriented from 0 to t, will be called a θ-path (or a θ-curve) in D. Throughout the present paper we shall fix θ, $0 < \theta < \frac{\pi}{2}$. $\| \ \|_Z$ denotes the norm of a Banach space Z.

In this section the nonlinear evolution equation in a complex Banach space X

$$dv/dt + Av = F(t, v), \tag{8}$$

$$v(0) = \phi \tag{9}$$

is considered on a θ-domain D. Here A and F are supposed to satisfy the following conditions.

Assumption 2.1. *$-A$ is the generator of holomorphic semi-groups $\{e^{-tA}\}$, $t \geqq 0$, such that e^{-tA} is holomorphic and bounded (in the operator norm) in the sector: $|\arg t| < \theta + \varepsilon$, ε being a positive constant;*

$$\|e^{-tA}\| \leqq M, \qquad |\arg t| < \theta + \varepsilon. \tag{10}$$

(Here and in what follows M stands for various constants).

Let γ be a fixed number, $0 \leqq \gamma < 1$. By the assumption above we can define the fractional power A^γ of A; for the basic properties of the fractional power of operators we refer to K. Yosida [20]. Let X_0, X_1 be the set of all elements of $D(A)$, $D(A^\gamma)$, respectively. ($D(S)$; domain of S). X_0, X_1 are Banach spaces, equipped with the graph norms.

Assumption 2.2. *$F(t, v)$ is a bounded analytic mapping of $D \times U \to X$ with*

$$F(t, 0) = 0; \qquad \partial_v F(t, 0) = 0 \tag{11}$$

where U is a convex open neighborhood of the origin in X_1. Furthermore, there is a constant M with

$$\int_{\Gamma(t)} (1+|t-s|^{-\gamma}) \sup_{v\in U} \|\partial_v F(s,v)\| \, |ds| \leqq M, \qquad t\in D, \tag{12}$$

where $\Gamma(t)$ is a θ-path in D. ($\partial_v F$: the Fréchet derivative of F with respect to v).

Our theorem on the existence of holomorphic solutions of (8), (9) reads as follows.

Theorem 2.1. *Let $\phi\in X_1$, and the assumptions* 2.1 *and* 2.2 *be satisfied. If $\|\phi\|_{X_1}$ is sufficiently small, then there is a unique X_0-valued holomorphic function $v=v(t)$ defined on D which satisfies* (8) *and converges to ϕ as $t\to 0$ $(t\in D)$ in the norm of X_1.*

Corollary. *Let $0<\beta\leqq\alpha\leqq 1$ such that $\gamma\leqq\alpha$. In addition to the assumptions of Theorem* 2.1, *assume that $F(t,v)\in D(A^\beta)$ and*

$$\|A^\beta F(t,v)\|_X \leqq \rho(\|v\|_{X_1}) \|A^\alpha v\|_X, \qquad v\in D(A^\alpha)\cap U, \ t\in D \tag{13}$$

with a continuous function $\rho=\rho(s)$ of s, $s\geqq 0$. If $\phi\in D(A^\alpha)$, then $v(t)$ converges to ϕ as $t\to 0$, $t\in D$, in the graph norm of A^α.

We next consider the equation (8), (9) on two different θ-domains D_j, $j=1,2$. Let v_j be solutions of (8), (9) on D_j. Then we are concerned with the value of $v_2(t)-v_1(t)$ for $t\in D_1\cap D_2$; the assumptions 2.1 and 2.2 are supposed to be satisfied with D replaced by D_j, $j=1,2$. We fix t in $D_1\cap D_2$. Let $\Gamma_j(t)$ be θ-paths in D_j such that $-\Gamma_1(t)+\Gamma_2(t)$ is a simple closed curve. Let D_0 be a domain bounded by $-\Gamma_1(t)+\Gamma_2(t)$, and set $D^*=D_0\cup D_1\cup D_2$. We make the following assumption.

Assumption 2.3. *There is a meromorphic function $G(s)$: $D^*\to L(X_1\times X_1; X)$ (the set of all bilinear continuous operators from $X_1\times X_1$ into X, with supremum norm) such that*

i) $G(s)=\partial_v^2 F(s,0)$, $s\in D_1\cup D_2$;

ii) *$G(s)$ has only a finite number of poles (say, $t_1, t_2, \ldots, t_N$) of at most order 2 in D_0:*

$$G(s)=\sum_{j=1}^N \frac{G_{2,j}}{(s-t_j)^2}+\sum_{j=1}^N \frac{G_{1,j}}{s-t_j}+G_0(s), \qquad s\in D_0 \tag{14}$$

where $G_{1,j}$, $G_{2,j}\in L(X_1\times X_1; X)$, and $G_0(s)$ is a holomorphic function on D^* with values in $L(X_1\times X_1; X)$.

We then show:

Theorem 2.2. *Let the assumptions* 2.1, 2.2, *and* 2.3 *be satisfied with D replaced by D_j, $j=1,2$. If $\phi(\in X_1)$ is sufficiently small in the norm of X_1, then*

$$\begin{aligned} v_2(t)-v_1(t) &= \pi i \sum_{j=1}^N \{A e^{-(t-t_j)A} G_{2,j}[\phi_j,\phi_j] - 2e^{-(t-t_j)A} G_{2,j}[\phi_j, A\phi_j]\} \\ &\quad + \pi i \sum_{j=1}^N e^{-(t-t_j)A} G_{1,j}[\phi_j,\phi_j] + R(t;\phi) \end{aligned} \tag{15}$$

with

$$R(t;\phi)=O(\|\phi\|_{X_1}^3),\ \text{as}\ \|\phi\|_{X_1}\to 0$$

where $\phi_j=e^{-t_jA}\phi$.

2.2. *Proof of Theorem 2.1.* To show the theorem, we prepare a lemma.

Lemma 2.1.

$$\|A^\gamma e^{-tA}\|\leqq M(1+|t|^{-\gamma}), \qquad |\arg t|<\theta. \tag{16}$$

Proof. Clearly (16) holds for real $t>0$. (See T. Kato [9], and also H. Tanabe [17]). For non-real t, we take a positive s so that $|\arg(t-s)|=\theta+\frac{1}{2}\varepsilon$. Then by a simple geometrical consideration

$$|t|\sin\tfrac{1}{2}\varepsilon<s<|t|. \tag{17}$$

Since $A^\gamma e^{-tA}=A^\gamma e^{-sA}e^{-(t-s)A}$, it follows from (16) with $t=s$, (10) and (17) that (16) holds for all t in $|\arg t|<\theta$.

Let Y (resp. Y_1) be the set of all bounded holomorphic X-(resp. X_1) valued functions on D; Y, Y_1 are Banach spaces, equipped with supremum norm. Let V be a convex open neighborhood of the origin in X_1 with its closure contained in U, and W be the set of all $v\in Y_1$ such that $v(t)\in V$, $t\in D$; W is an open neighborhood of the origin in Y_1. We now define the mapping $\mathscr{F}(\phi, w)$ by

$$\mathscr{F}(\phi, w)=w-e^{-tA}\phi-\mathscr{F}_0(w) \tag{18}$$

with

$$\mathscr{F}_0(w)(t)=\int_{\Gamma(t)} e^{-(t-s)A}F(s, w(s))\,ds \tag{19}$$

($\Gamma(t)$: a θ-path in D). We claim

i) $\mathscr{F}$ is well defined as a bounded analytic mapping of $V\times W$ into Y_1, and is independent of the path of integration as far as it is a θ-path in D.

ii) $\mathscr{F}(0,0)=0$; $\partial_w\mathscr{F}(0,0)=I$ (the identity operator).

If this claim is shown, then the classical implicit function theorem (see J. Dieudonne [2; p. 270]) is applicable to

$$\mathscr{F}(\phi, w)=0. \tag{20}$$

Consequently, there is an open neighborhood $V_0(\subset V)$ of the origin in X_1 and an analytic mapping $v=v(\cdot,\phi)$: $V_0\to W_1$ such that

$$\mathscr{F}(\phi, v(\cdot,\phi))=0(\phi\in V_0); \qquad v(\cdot,0)=0. \tag{21}$$

Moreover the v is the only X_1-valued continuous function defined on D that satisfies (21).

Now (21) implies that the v is a solution of the integral equation:

$$v(t)=e^{-tA}\phi+\mathscr{F}_0[v](t)(\equiv v^{(1)}(t)+v^{(2)}(t)), \tag{22}$$

(here we simply write $v(t)$ for $v(t,\phi)$).

Clearly $v^{(1)}(t)$ is holomorphic for $t \in D$ in the norm of X_0, satisfies $dv^{(1)}(t)/dt + Av^{(1)}(t) = 0$, and converges to ϕ as $t \to 0$, $t \in D$, in the norm of X_1. Since the composition of bounded analytic mappings is also bounded and analytic, $F(t, v(t))$ is bounded and analytic on D in the norm of X. In particular, it is Hölder continuous in t. Hence similarly to K. Masuda [12] (see also T. Kato [8; p. 491]), $v^{(2)}(t)$ is holomorphic in $t \in D$ in the norm of X, is in $D(A)$, and satisfies

$$dv^{(2)}/dt + Av^{(2)} = F(t, v(t)), \qquad t \in D. \tag{23}$$

The first term of the LHS and the RHS of (23) are holomorphic for $t \in D$ in the norm of X, and so is $Av^{(2)}$. This implies $v^{(2)}$ is holomorphic for $t \in D$ in the norm of X_0. Using Lemma 2.1, one finds

$$\|v^{(2)}(t)\|_{X_1} \leqq M \int_{\Gamma(t)} (1 + |t-s|^{-\gamma}) \|F(s, v(s))\|_X |ds|. \tag{24}$$

Since $F(s, v(s))$ is bounded for $s \in D$ in X, it follows from the above inequality (24) that $v^{(2)}(t)$ converges to zero as $t \to 0$, $t \in D$, in the norm of X_1. Thus $v(t) = v^{(1)}(t) + v^{(2)}$ is holomorphic in the norm of X_0, satisfies (8), and converges to ϕ as $t \to 0$ in the norm of X_1. The uniqueness of solution of (8), (9) follows from that of (22) (and (25)). Thus for the completion of the proof it remains only to show the claim. Set $\mathscr{F}_1[w] = w$, $\mathscr{F}_2[\phi] = e^{-tA}\phi$. Then it is easy to see that $\mathscr{F}_1$, $\mathscr{F}_2$ are analytic mappings of $V \times W \to Y_1$ with $\mathscr{F}_1[0] = \mathscr{F}_2[0] = 0$, $\partial_w \mathscr{F}_1 = I$, $\partial_w \mathscr{F}_2 = 0$. Hence we have only to show that $\mathscr{F}_0$ is a bounded analytic mapping of $W \to Y_1$, is independent of a path of integration, and $\partial_w \mathscr{F}_0(0) = 0$; it is clear that $\mathscr{F}_0(0) = 0$. As before we can see that for $w \in Y_1$ $\mathscr{F}_0(w)(t)$ is holomorphic for $t \in D$ in X_0, and so in X_1. By Taylor's Theorem (M. Berger [1; p. 75]),

$$\begin{aligned} F(t, w_2) - F(t, w_1) &= \partial_v F(t, w_1)[w_2 - w_1] \\ &\quad + \int_0^1 \{\partial_v F(t, w_1 + s(w_2 - w_1)) - \partial_v F(t, w_1)\}[w_2 - w_1]\, ds \end{aligned} \tag{25}$$

for $w_1, w_2 \in W$. In particular, setting $w_2 = w$, $w_1 = 0$, and using (11), we get

$$F(t, w) = \int_0^1 \partial_v F(t, sw)[w]\, ds$$

from which it follows that (12) holds with $\|\partial_v F(s, v)\|$ replaced by $\|F(s, v)\|_X$. Hence by Lemma 2.1, $\mathscr{F}_0(w)(t)$ is bounded on D in the norm of X_1, and so $\mathscr{F}_0(w) \in Y_1$:

$$\|\mathscr{F}_0(w)\|_{Y_1} \leqq M, \qquad w \in W.$$

Since $F(t, w)$ is continuously differentiable with respect to $t \in D$ and $w \in W$, it follows from (25) that $\mathscr{F}_0(w)$ is an analytic mapping of $W \to Y_1$ with

$$\partial_w \mathscr{F}_0(w)[h] = \int_{\Gamma(t)} e^{-(t-s)A} \partial_w F(s, w(s))[h(s)]\, ds, \qquad (h \in W).$$

(See Berger [1; p. 85]). Hence $\partial_w \mathscr{F}_0(0) = 0$ by (11). Since D is simply connected, and since $e^{-(t-s)A} F(s, w(s))$ is bounded and holomorphic for s in $|\arg(t-s)| < \theta$,

$s \in D$, by Cauchy's integral theorem, $\mathscr{F}_0(w)$ is independent of a path of integration. This completes the proof of Theorem 2.1.

Proof of Corollary of Theorem 2.1. By (18)

$$A^\alpha v(t) = e^{-tA} A^\alpha \phi + \int_{\Gamma(t)} A^{\alpha-\beta} e^{-(t-s)A} A^\beta F(s, v(s))\, ds,$$

from which the proof easily follows.

2.3. *Proof of Theorem 2.2.* Since $v_j(\cdot, \phi)$ is analytic in ϕ (in Y_1), $v_j(\cdot, \phi)$ can be expanded in the form:

$$v_j(\cdot, \phi) = v_j(\cdot, 0) + \partial_\phi v_j(\cdot, 0)[\phi] + \tfrac{1}{2}\partial_\phi^2 v_j(\cdot, 0)[\phi, \phi] + R_j(\cdot; \phi) \tag{26}$$

with $R_j(t; \phi) = O(\|\phi\|_{X_1}^3)$, uniformly in t, as $\|\phi\|_{X_1} \to 0$. We shall compute each term on the right. Clearly $v_j(\cdot, 0) = 0$. Hence, by (22) and (11)

$$\begin{aligned} \partial_\phi v_j(t, 0)[h] &= e^{-tA} h + \int_{\Gamma(t)} e^{-(t-s)A} \partial_v F(s, v_j(s, 0))[\partial_\phi v_j(s, 0)[h]]\, ds \\ &= e^{-tA} h. \end{aligned} \tag{27}$$

Since $v_j(\cdot, 0) = 0$, using (14) and (11) we have

$$\begin{aligned} \partial_\phi^2 F(s, v_j(s, 0))[\phi, \phi] &= \partial_v^2 F(s, v_j(s, 0))[\partial_\phi v_j(s, 0)[\phi], \partial_\phi v_j(s, 0)[\phi]] \\ &\quad + \partial_v F(s, v_j(s, 0))[\partial_\phi^2 v_j(s, 0)[\phi, \phi]] \\ &= G(s)[e^{-sA}\phi, e^{-sA}\phi] \qquad \text{(by (27))} \\ &= \sum_{j=1}^{N} \left\{ \frac{1}{(s-t_j)^2} g_{2,j}(s) + \frac{1}{s-t_j} g_{1,j}(s) \right\} + g_0(s) \end{aligned}$$

where

$$g_0(s) = G_0[e^{-sA}\phi, e^{-sA}\phi]; \qquad g_{k,j}(s) = G_{k,j}[e^{-sA}\phi, e^{-sA}\phi], \qquad k = 1, 2.$$

Hence

$$\begin{aligned} v_2(t) - v_1(t) &= \tfrac{1}{2} \sum_{j=1}^{N} \int_{\Gamma(t)} e^{-(t-s)A} \left\{ \frac{1}{(s-t_j)^2} g_{2,j}(s) + \frac{1}{s-t_j} g_{1,j}(s) \right\} ds \\ &\quad + \tfrac{1}{2} \int_{\Gamma(t)} e^{-(t-s)A} g_0(s)\, ds + R(t; \phi) \end{aligned}$$

with $R(t; \phi) = R_2(t; \phi) - R_1(t; \phi)$, and $\Gamma(t) = -\Gamma_1(t) + \Gamma_2(t)$.

Since $e^{-(t-s)A} g_{k,j}(s)$, $e^{-(t-s)A} g_0(s)$ are holomorphic for s in D_0, we obtain

$$\begin{aligned} v_2(t) - v_1(s) &= \pi i \sum_{j=1}^{N} \{A e^{-(t-t_j)A} g_{2,j}(t_j) + e^{-(t-t_j)A} g'_{2,j}(t_j) \\ &\quad + e^{-(t-t_j)A} g_{1,j}(t_j)\} + R(t; \phi) \end{aligned} \tag{28}$$

by Cauchy's integral formula, where $'$means the derivative in s. Since $\partial_v^2 F(s, 0)[h_1, h_2]$ is symmetric in h_1, h_2, it follows from analytic continuation that $G(s)$ is also symmetric. Hence $G_0(s)$, $G_{1,j}$, $G_{2,j}$ are all symmetric by the uniqueness of the coefficients of the Laurent expansion. Using the symmery of $G_{2,j}$, we have $g'_{2,j}(t_j) = -2G_{2,j}[\phi_j, A\phi_j]$, ($\phi_j = e^{-t_jA}\phi$). Hence by (28) we get the result.

3. Applications

Here are some applications of the abstract theory (Theorems 2.1 and 2.2) to solutions of nonlinear diffusion equations of the form (1).

Example 3.1. (The case $f(u)=u^2$.) Consider the initial-boundary value problem (4), (2), (3) on D_j, where D_j is as in Theorem 1.1. Let y be a function given in (6). If u_j is a solution of (4), (2), (3), then the function v_j defined by

$$v_j(\cdot, t)=(u_j(\cdot, t)-y(t))/f(y(t)) \tag{29}$$

satisfies the following equation on D_j:

$$\frac{\partial v}{\partial t}=\Delta v+g(t, v), \tag{30}$$

$$\partial_n v|_\Gamma=0; \qquad v|_{t=0}=(u_0(\cdot)-a)/f(a)(\equiv b)$$

where

$$g(t, w)=[f(y(t)+f(y(t))\,w)-f(y(t))-f'(y(t))\,f(y(t))\,w]/f(y(t)). \tag{31}$$

In this case a simple calculation gives: $g(t, v)=v^2/(t_0-t)^2$. Conversely a solution v_j of (30) determines a solution of (4), (2), (3) via (29). We like to apply Theorem 2.1 to the equation (30) on D_j. To this end we set $X=L^p(\Omega)$ $(p>n)$, and define an operator A in $L^p(\Omega)$ by:

$$D(A)=\{u\in W_p^2(\Omega);\ \partial_n u|_\Gamma=0\}, \qquad Au=-\Delta u.$$

Then the A has the following properties:

Lemma 3.1.–*A generates the holomorphic semi-groups $\{e^{-tA}\}$, $t\geqq 0$, which is bounded and holomorphic (in the operator norm) in the sector: $|\arg t|<\frac{\pi}{2}-\varepsilon$, ε being any positive constant.*

Lemma 3.2. *Let γ be such that* $1/2<\gamma<1$. Then

$$D(A^\gamma)\subset W_p^1(\Omega)\subset C(\bar\Omega), \tag{32}$$

the imbeddings being continuous. (For the proofs, see the appendix of the present paper).

Thus we see that the assumption 2.1 is satisfied. Let γ be a fixed number such that $1/2<\gamma<1$, and X_1 be the Banach space, consisting of all elements of $D(A^\gamma)$, and equipped with the graph norm of A^γ. Let U be the unit ball in X_1, centered at the origin. Set $F(t, v)=g(t, v)=v^2/(t_0-t)^2$. Then it follows from Lemma 3.2 that $F(t, v)$ is bounded and holomorphic on $D_j\times U$, and satisfies

$$\partial_v F(t, v)\,[h]=2\,v(\cdot)\,h(\cdot)/(t_0-t)^2, \qquad v, h\in X_1,\ t\in D_j$$

and so

$$\|\partial_v F(t, v)\|\leqq M/|t_0-t|^2, \qquad v\in U,\ t\in D_j.$$

Now the assumption 2.2 is easily verified. Since v^2 is in $D(A)$ and $\|Av^2\|_p \leqq M\|v\|_{1,p}\|v\|_{2,p}$ for $v \in D(A)$ by Lemma 3.2, it follows that (13) holds with $\alpha=\beta=1$. Thus the corollary of Theorem 2.1 is applicable to (30). Noting

$$\|b\|_{X_1} \leqq M\|\partial_x^2 b\|_p, \tag{33}$$

we have Theorem 1.1 by using (29). For the proof of (33), see the appendix 3.

We next show Theorem 1.2. We claim:

$$v_2(t)-v_1(t)=-4\pi i e^{-(t-t_0)A}(\nabla e^{-t_0 A} b)^2+R(t; b) \tag{34}$$

$(R(t; b)=O(\|b\|_{X_1}^3))$ for $t \in D_1 \cap D_2$ $\operatorname{Re} t>t_0+\delta$. Here and in what follows we simply write $v_j(t)$ for $v_j(\cdot, t)$ unless otherwise stated.

Since

$$\partial_v^2 F(t,0)[h,h]=2h(\cdot)^2/(t_0-t)^2,$$

it follows that the assumption 2.3 is satisfied with $G_0=G_1=0$, $G_2[h,h]=2h^2$. Consequently,

$$\begin{aligned}\text{the } LHS \text{ of } (34) = {} & 2\pi i A e^{-(t-t_0)A}(e^{-t_0 A} b)^2 \\ & -4\pi i e^{-(t-t_0)A}(A e^{-t_0 A} b)(e^{-t_0 A} b)+R(t; b)\end{aligned}$$

by Theorem 2.2. After a simple manipulation, one find the RHS of the above equality is equal to that of (34). This shows (34).

We next show the first term on the RHS of (34) is not zero, if u_0 is not constant. To this end we prepare a lemma.

Lemma 3.3. *Let $h \in L^p(\Omega)$. If $e^{-tA}h$ is constant, say c, at some $t=t^*$ in the sector: $|\arg t|<\theta$, then h must be equal to c.*

Proof. We have

$$e^{-tA}(h-c)=e^{-(t-t^*)A}(e^{-t^*A}h-c)=0, \quad |\arg(t-t^*)|<\theta$$

since $e^{-tA}c=c$. By the analytic continuation theorem in the complex function theory, $e^{-tA}(h-c)=0$ in the sector: $|\arg t|<\theta$. Letting $t \to 0$, we get $v=c$ as desired.

Now by the lemma just proved, $e^{-tA}b$ is not constant if b is not constant, i.e. if u_0 is not constant. Hence $\nabla e^{-tA}b \neq 0$ (as an element of $L^p(\Omega)$). Hence, applying Lemma 3.3 again we see $e^{-(t-t_0)A}(\nabla e^{-t_0 A}b)^2 \neq 0$; Here we note that $(\nabla e^{-tA}b)^2$ is in $L^p(\Omega)$, which can be seen from Lemma 3.2 and $e^{-tA}b \in D(A)$.

Let t_1 be a fixed number with $t_1 \in D_1 \cap D_2$, $\operatorname{Re} t_1>t_0+\delta$. If u_0 is not constant then the first term on the RHS of (34) dominates the second term at $t=t_1$, since the first term is not zero as just proved. Hence

$$v_2(t_1)-v_1(t_1) \neq 0 \tag{35}$$

if u_0 is not constant and if $\|\partial_x^2 u_0\|_p/|a|^2$, and so $\|b\|_{X_1}$, is sufficiently small. Suppose that $u_1(t_2)=u_2(t_2)$ for some t_2 in $D_1 \cap D_2$, $\operatorname{Re} t_2>t_0+\delta$. Then by (29) $v_1(t_2)=v_2(t_2)$. Hence $v_1(t)=v_2(t)$ for all $t \in D_1 \cap D_2$ such that $|\arg(t-t_2)|<\theta$ and

$|t-t_2|<\varepsilon$ (ε: small positive number), by the uniqueness of solution of (30) with initial moment t_2 and initial function $v(t_2)$, which can be proved in a routin manner. Since $v_1(t)-v_2(t)$ is analytic in $|\arg(t-t_0-\delta)|<\theta$, and vanishes in $|\arg(t-t_2)|<\theta$ and $|t-t_2|<\varepsilon$, it follows from the analytic continuation theorem in the complex function theory that $v_1(t)-v_2(t)$ vanishes in $|\arg(t-t_0-\delta)|<\theta$, which contradicts (35) if u_0 is not constant. This completes the proof of Theorem 1.2.

Remark. By (34), $u_2(\cdot, t)-u_1(\cdot, t)$ has the expansion of the form

$$u_2(t)-u_1(t)=4\pi i\frac{1}{a^2(1-at)^2}e^{-(t-t_0)A}(\nabla e^{-t_0 A}u_0)^2+R(t;b),$$
$$|\arg(t-t_0-\delta)|<\theta, \tag{36}$$

where $R(t;b)=O(\|\partial_x^2 b\|_p^3)$

Example 3.2. (The case $f(u)=u+u^2$.) Let u_0 be a non-negative function ($\neq 0$) in $W_p^2(\Omega)$, and set $a=Pu_0$ as before. Let

$$t_k=\log(1+a^{-1})+2\pi k i, \qquad k=0, \pm 1, \dots$$
$$\theta_k=\tan^{-1}(2k\pi/t_0), \qquad |\theta_k|<\tfrac{\pi}{2}.$$

We set

$$D_k=\{t\in\mathbb{C};\ |\arg t|<\theta_k-\delta,\ |\arg(t-t_{k-1}-\delta i)-\tfrac{\pi}{2}|<\theta_k-\delta+\tfrac{\pi}{2}\}$$

with δ sufficiently small. Suppose $|\theta_k|<\theta$. Then D_k is a θ-domain; note θ is arbitrary positive number with $0<\theta<\frac{\pi}{2}$, in the present case.

Now we consider the following equation on D_k:

$$\frac{\partial u}{\partial t}=\Delta u+u+u^2, \qquad x\in\Omega, \quad t\in D_k. \tag{37}$$

Let $y=y(t)$ be a function given by

$$y(t)=e^t/(c-e^t), \qquad c=1+a^{-1}.$$

Then the function v_k defined by (29) with $f(y(t))=ce^t/(c-e^t)^2$ satisfies the equation (30) with $g(t,v)=ce^t v^2/(c-e^t)^2$. Thus in a way similar to Example 3.1, we can show:

i) there exists a unique solution u_k of (37), (2), (3) which is analytic in D_k as a $W_p^2(\Omega)$-valued function and converges to u_0 as $t\to 0$, $t\in D_k$, in the norm of $W_p^2(\Omega)$, if $\|\partial_x^2 u_0\|_p/|a+a^2|$ is sufficiently small.

ii) for $t\in D_k\cap D_j$ ($|j|\leqq|k|$) with $\operatorname{Re} t>t_0$,

$$u_k(t)-u_j(t)=-\frac{4\pi i e^t}{a(a+1)(1+a-ae^t)^2}\sum_{s=j+1}^{k} e^{-(t-t_s)A}(\nabla e^{-t_s A}u_0)^2+R(t;b) \tag{38}$$

where $R(t,b)=O(\|\partial_x^2 b\|_p^3)$.

The special case $n=1$ and $\Omega=(0,\pi)$ is of some interest. In this case we have $e^{-(t-t_s)A}=e^{-(t-t_0)A}$, and hence, by (38),

$$u_k(t)-u_j(t)=-\frac{4\pi i(k-j)e^t}{a(1+a)(1+a-ae^t)^2}e^{-(t-t_0)A}(\nabla e^{-t_0A}u_0)^2+R(t;b)$$

since the eigenvalues of A are: $0,1,4,\dots,n^2,\dots$.

Thus in a way similar to the proof of Theorem 1.2, it can be shown that if $u_k(t)=u_j(t)$ for some $t\in D_j\cap D_k$, $\operatorname{Re} t>t_0$, then u_0 must be constant.

Example 3.3. (The case $f(u)=e^u$.) Consider the following equation on D_j:

$$\frac{\partial u}{\partial t}=\Delta u+e^u, \qquad x\in\Omega, \tag{39}$$

with the initial-boundary conditions (2), (3), where D_j is the same complex-domain as that in Example 3.1 in which we replace $t_0=1/a$ by $t_0=e^{-a}$. Let

$$y(t)=-\log(t_0-t), \qquad (t_0=e^{-a}).$$

The function v_j defined by (29) with $f(y(t))=1/(t_0-t)$ satisfies (30) with

$$g(t,v)=e^{v/(t_0-t)}-1-v/(t_0-t).$$

In ways similar to those of the preceding examples, we can show:

i) If $\|\partial_x^2 u_0\|_p/|e^a|$ is sufficiently small, then there exists a unique solution $u_j(t)$ of (39), (2), (3) which is analytic in D_j as a $W_p^2(\Omega)$-valued function, and converges to u_0 as $t\to 0$, $t\in D_j$, in the norm of $W_p^2(\Omega)$;

ii) we have

$$u_2(t)-u_1(t)=-2\pi i+\frac{2\pi i}{e^a-e^{2a}t}e^{-(t-t_0)A}(\nabla e^{-t_0A}u_0)^2+R(t;b)$$

for $t\in D_1\cap D_2$, $\operatorname{Re} t>t_0+\delta$.

Appendix 1

Proof of Lemma 3.1. It is known (see A. Friedman [3; p. 78]) that for any ω, $\frac{\pi}{2}<\omega<\pi$, there is a $\Lambda_0>0$ such that the resolvent set of A contains the set $\Sigma(\omega;\Lambda_0)$:

$$\Sigma(\omega;\Lambda_0)=\{\lambda\in\mathbb{C};\ |\arg\lambda|<\omega,\ |\lambda|>\Lambda_0\},$$

and that the inequality

$$\|(\lambda+A)^{-1}\|\leqq M/|\lambda|, \qquad \lambda\in\Sigma(\omega;\Lambda_0) \tag{A1}$$

holds. We also know that the spectrum of A consists only of a countable number of non-negative eigenvalues $\lambda_1<\lambda_2\leqq\dots$ with $\lambda_n\to\infty$ $(n\to\infty)$ and $\lambda_1=0$.

We claim that

$$\|(\lambda+A)^{-1}\| \leqq M/|\lambda|, \qquad |\arg \lambda| < \omega. \tag{A2}$$

Indeed, let P be the operator defined in (5), and set $Q = I - P$ (I: the identity operator). Then since

$$(\lambda+A)^{-1} = \lambda^{-1} P + (\lambda+A)^{-1} Q,$$

one find

$$\|(\lambda+A)^{-1}\| \leqq 1/|\lambda| + \|(\lambda+A)^{-1} Q\|.$$

Hence to show the claim, it suffices to show the estimate

$$\|(\lambda+A)^{-1} Q\| \leqq M/|\lambda|, \qquad |\arg \lambda| < \omega, \ |\lambda| \leqq \Lambda_0, \tag{A3}$$

by (A1). $(\lambda+A)^{-1} Q$ is holomorphic in $|\arg(\lambda+\lambda_1)| < \omega$, and so bounded on $|\arg \lambda| < \omega$, $|\lambda| \leqq \Lambda_0$. Hence (A3) clearly holds. Here we note that the P is the projection on the eigenspace corresponding to the first eigenvalue $\lambda_1 = 0$.

Now, by the general property of holomorphic semi-groups it follows from (A2) that e^{-tA} is holomorphic and bounded, in the operator norm, in the sector: $|\arg t| < \omega - \frac{\pi}{2}$.

Appendix 2

Proof of Lemma 3.2. We first show

$$\|\partial_x(\lambda+A)^{-1}\| \leqq M \lambda^{-1/2} \qquad \lambda \geqq 1. \tag{A4}$$

Since

$$\|\partial_x u\|_p \leqq M \, \|u\|_p^{1/2} \, \|\partial_x^2 u\|_p^{1/2}$$

by partial integration and the Schwarz inequality, we have

$$\|\partial_x(\lambda+A)^{-1}\| \leqq M \, \|(\lambda+A)^{-1}\|^{1/2} \, \|\partial_x^2(\lambda+A)^{-1}\|^{1/2}, \qquad \lambda > 0. \tag{A5}$$

Let X_0 be the set of all elements of $D(A)$; X_0 is a Banach space, equipped with the norm of $W_p^2(\Omega)$. Since the operator $I+A$ on X_0 is a bounded, invertible, and onto mapping: $X_0 \to L^p(\Omega)$, it follows that

$$\|u\|_{2,p} \leqq M \, \|(I+A)\, u\|_p$$

and hence that

$$\|\partial_x^2 (I+A)^{-1}\| \leqq M. \tag{A6}$$

Thus (A6), together with (A2) and (A5), gives the desired estimate (A4).

Since the fractional power $(I+A)^{-\gamma}$ of $I+A$ is given by

$$(I+A)^{-\gamma} = \frac{\sin \pi\gamma}{\pi} \int_0^\infty \lambda^{-\gamma} (\lambda+I+A)^{-1} \, d\lambda \tag{A7}$$

(see K. Yosida [20; p. 260]), it follows from (A4) that $\partial_x(I+A)^{-\gamma}$ is a bounded operator in $L^p(\Omega)$. This shows the first inclusion of (32); note that $A^\gamma - (I+A)^\gamma$

is a bounded operator in $L^p(\Omega)$. The second inclusion relation follows from the Sobolev imbedding theorem.

Appendix 3

Proof of the Inequality (33). By (A7) with γ replaced by $1-\gamma$, we see $(I+A)^\gamma (I+A)^{-1}$ is a bounded operator in $L^p(\Omega)$. Hence

$$\|A^\gamma u\|_p \leqq M(\|u\|_p + \|Au\|_p). \tag{A8}$$

On the other hand, as shown in the Appendix 1, AQ has a bounded inverse in $QL^p(\Omega)$. Therefore

$$\|Qu\|_p \leqq M\,\|Au\|_p \tag{A9}$$

where $Q=I-P$. Since $Qb=b$, (33) follows from (A8), and (A9).

References

1. Berger, M.S.: Nonlinearity and Functional Analysis. New York-San Francisco-London: Academic Press 1977
2. Dieudonne, J.: Foundations of Modern Analysis. New York-San Francisco-London: Academic Press 1967
3. Friedman, A.: Partial Differential Equations. New York-Chicago-London: Holt, Rinehart and Winston, Inc. 1967
4. Fujita, H.: On some nonexistence and nonuniqueness theorems for nonlinear parabolic equations. In: Proceedings of the Symposium on Pure Mathematics. Amer. Math. Soc. **18**, 105–113 (1970)
5. Fujiwara, D.: Concrete characterization of the domain of the fractional powers of some elliptic operators of the second order. Proc. Japan Acad. **43**, 82–86 (1967)
6. Furuya, K.: Analyticity of solutions of quasilinear evolution equations. Osaka J. Math. **18**, 669–698 (1981)
7. Hayden, T., Massey III, F.: Nonlinear holomorphic semigroups. Pac. J. Math. **57**, 423–439 (1975)
8. Kato, T.: Perturbation Theory for Linear Operators. Berlin-Heidelberg-New York: Springer 1966
9. Kato, T.: Fractional powers of dissipative operators. J. Math. Soc. Japan **13**, 246–274 (1961)
10. Kato, T., Tanabe, H.: On the abstract evolution operators. Osaka Math. J. **14**, 107–133
12. Masuda, K.: On the holomorphic evolution operators. J. Math. Anal. Appl. **39**, 706–711 (1972)
13. Masuda, K.: On the analyticity and the unique continuation theorems for solutions of the Navier-Stokes equations. Proc. Japan Acad. **43**, 827–832 (1967)
14. Masuda, K.: Blow-up of solutions of some nonlinear diffusion equations. To appear in: Proceedings of 1982 US-Japan Symposium on Nonlinear Partial Equations
15. Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Berlin-Heidelberg-New York: Springer 1983
16. Sobolevskii, P.E.: Parabolic equations in a Banach space with an unbounded variable operator, a fractional power of which has a constant domain of definition. Dokl. Akad. Nauk SSSR **138**, 59–62 (1961)
17. Tanabe, H.: Equations of Evolutions. London: Pitman 1979
18. Tsutsumi, M.: Existence and nonexistence of global solutions for nonlinear parabolic equations. Publ. RIMS, Kyoto Univ. **8**, 211–229 (1972/1973)
19. Weissler, F.B.: Local existence and nonexistence for semilinear equations in L^p. Indiana Univ. J. **29**, 79–102 (1980)
20. Yosida, K.: Functional Analysis. the Second Edition. Berlin-Heidelberg-New York: Springer 1968

Received February 13, 1984

Math. Z. 187, 75–80 (1984)

Mathematische Zeitschrift

Indecomposables Over Representation-Finite Algebras Are Extensions of an Indecomposable and a Simple

Klaus Bongartz

Mathematisches Institut B, Universität Stuttgart, Pfaffenwaldring 57, D-7000 Stuttgart 80, West Germany

To be more precise we prove the following:

Theorem. *Let A be a finite dimensional algebra over an algebraically closed field k, such that the number of isomorphism classes of finite dimensional indecomposable left A-modules is finite. Given any non-simple indecomposable U, there exist indecomposables U_1 and U_2 and an exact sequence $0 \to U_1 \to U \to U_2 \to 0$ such that U_1 or U_2 is simple.*

The theorem shows in particular that the lengths of the indecomposables form an interval $[1, n(A)] \cap \mathbb{N}$, where we have $n(A) \leqq 2 \dim A + 30(\dim A)^{1/2}$ by [3].

The theorem comes natural, at least to someone who is working on representation-finite algebras only. In fact, I know no counterexample in the general case. However, the proof depends heavily on the representation-finiteness assumption, because the existence of coverings is used to reduce the problem to the simply connected case. There the theorem follows from the geometric interpretation of extension groups given in [12] and some properties of modules over simply connected algebras.

1. A Lemma on Module Varieties

In this section, we give more details than later on because there are some representation theorists (e.g. the author), who are not familiar with algebraic geometry.

Let $a_1, a_2, \ldots, a_r$ be a basis of the finite dimensional associative k-algebra A with unit a_1. We suppose always that the field k is algebraically closed. To study the category $\operatorname{mod} A$ of finite dimensional unital left A-modules, we choose representatives $P_1, P_2, \ldots, P_s$ of isomorphism classes of indecomposable projectives. Hence the tops $S_i = P_i/\operatorname{rad} P_i$ form a list of simples. The map $M \mapsto \underline{\dim}\, M := (\dim_k \operatorname{Hom}_A(P_i, M)) \in \mathbb{N}^s$ induces an isomorphism between the Grothendieck-group $G(A)$ and $\mathbb{Z}^s$.

Given $t \in \mathbb{N}$, the variety $R(t)$ of t-dimensional A-modules is the closed subvariety of k^{rt^2} defined by

$$R(t) := \{m = (m(a_1), m(a_2), \ldots, m(a_r)) \mid m(a_i) \in k^{t,t},\; m(a_1) = E_t,\; m(a_i)\, m(a_j) = \sum_l a_{ijl}\, m(a_l)\}.$$

Here $k^{p,q}$ denotes the matrices with coefficients in k having p rows and q columns. E_t is the unit of the ring $k^{t,t}$ and the structure constants a_{ijl} are defined by $a_i a_j = \sum_l a_{ijl} a_l$. Of course, there is a bijection $m \leftrightarrow M$ between $R(t)$ and the set of unital A-module structures on k^t. The connected components of $R(t)$ with respect to the Zariski-topology are given by $R(\underline{d}) := \{m \mid \underline{\dim}\, M = \underline{d}\}$ where $\underline{d}$ is running through the elements of $\mathbb{N}^s$ satisfying $\sum_i \underline{d}(i) \dim_k S_i = t$ ([9], [12]). The general linear group $Gl_t(k)$ operates on each $R(\underline{d})$ such that the orbit of m corresponds to the isomorphism class of M.

Keeping the notation and assumptions, we have:

Proposition. *Let $M_i \leftrightarrow m_i \in R(\underline{d}_i)$ be given, $1 \leqq i \leqq 3$. Put $t_i = \sum_j \underline{d}_i(j) \dim_k S_j$ and $G = Gl_{t_2}(k)$. Suppose that:*

a) $\underline{d}_1 + \underline{d}_3 = \underline{d}_2$,

b) $\operatorname{Ext}^1_A(M_1 \oplus M_3, M_1 \oplus M_3) = \operatorname{Ext}^1_A(M_3, M_1)$,

c) $\overline{Gm_2} = R(\underline{d}_2)$, *where $\overline{Gm_2}$ denotes the closure of the orbit Gm_2. Then there is an exact sequence $0 \to M_1 \to M_2 \to M_3 \to 0$.*

Proof. Define

$$Z(m_1, m_3) := \{z = (z(a_1), \ldots, z(a_r)) \mid z(a_i) \in k^{t_1, t_3},\; \sum_l a_{ijl}\, z(a_l) = m_1(a_i)\, z(a_j) + z(a_i)\, m_3(a_j)\}.$$

This is a linear subspace of $k^{rt_1t_3}$, whence an irreducible algebraic variety. We look at the polynomial map $F\colon G \times Z(m_1, m_3) \to R(\underline{d}_2)$ given by

$$F((g, z)) := \left(\ldots, g \begin{bmatrix} m_1(a_i) & z(a_i) \\ 0 & m_3(a_i) \end{bmatrix} g^{-1}, \ldots\right).$$

We have to show that the image Y of F hits the orbit Gm_2. This follows by counting dimensions once we have calculated the differential f of F at the point $(1, 0)$. Using the functorial approach via dual numbers $k[\delta]/(\delta^2)$ (see e.g. [6], page 69) we obtain

$$(1)\quad F\left(\left(1 + \delta \begin{bmatrix} g_1 & g_2 \\ g_3 & g_4 \end{bmatrix}, \delta z\right)\right) = (B_1, \ldots, B_i, \ldots, B_r) \text{ with}$$

$$B_i = \begin{bmatrix} m_1(a_i) & 0 \\ 0 & m_3(a_i) \end{bmatrix} + \delta\left(\begin{bmatrix} g_1\, m_1(a_i) & g_2\, m_3(a_i) \\ g_3\, m_1(a_i) & g_4\, m_3(a_i) \end{bmatrix} - \begin{bmatrix} m_1(a_i)\, g_1 & m_1(a_i)\, g_2 \\ m_3(a_i)\, g_3 & m_3(a_i)\, g_4 \end{bmatrix} + \begin{bmatrix} 0 & z \\ 0 & 0 \end{bmatrix}\right).$$

Here $g_1 \in k^{t_1,t_1}$, $g_2 \in k^{t_1,t_3}$, $g_3 \in k^{t_3,t_1}$, $g_4 \in k^{t_3,t_3}$ and $z \in Z(m_1, m_3)$.

The restriction $F_1 = F|G \times \{0\}$ maps onto the orbit C of $m_1 \oplus m_3$ and it is smooth, so that the image of its differential f_1 at $(1,0)$ is the tangent space T_1 at $m_1 \oplus m_3$ in C. Denote the restriction of F to $\{1\} \times Z(m_1, m_3)$ by F_2 and its differential at $(1,0)$ by f_2. Using formula (1), we conclude that the images of f_1 and f_2 intersect in

$$B(m_1, m_3) := \{z \in Z(m_1, m_3) \mid z(a_i) = m_1(a_i)\, g_2 - g_2\, m_3(a_i) \text{ for some } g_2 \in k^{t_1,t_3}\}.$$

It is well-known that $\operatorname{Ext}_A^1(M_3, M_1)$ and $Z(m_1, m_3)/B(m_1, m_3)$ are isomorphic as k-vector spaces. By [12], p. 260, we have $\operatorname{Ext}_A^1(M_1 \oplus M_3,\ M_1 \oplus M_3) = T/T_1$, where T is the tangent space at $m_1 \oplus m_3$ in $R(\underline{d}_2)$. We calculate:

$$\begin{aligned}
\dim \operatorname{Im} f &= \dim \operatorname{Im} f_1 + \dim \operatorname{Im} f_2 - \dim(\operatorname{Im} f_1 \cap \operatorname{Im} f_2) \\
&= \dim T_1 + \dim Z(m_1, m_3) - \dim B(m_1, m_3) \\
&= \dim T_1 + \dim \operatorname{Ext}_A^1(M_3, M_1) \\
&= \dim T_1 + \dim \operatorname{Ext}_A^1(M_1 \oplus M_3, M_1 \oplus M_3) \\
&= \dim T \\
&\geqq \dim R(\underline{d}_2).
\end{aligned}$$

Let V be a non-empty open subset of Y consisting of smooth points. Since $G \times Z(m_1, m_3)$ is irreducible, $F^{-1}(V)$ has non-empty intersection with the open subset of points where the differential of F has rank not less than $\dim T$. This implies $\dim \bar{Y} \geqq \dim T \geqq \dim R(\underline{d}_2)$, whence $\bar{Y} = \overline{Gm_2}$ and $Y \cap Gm_2 \neq \varnothing$.

2. Indecomposables Over Simply Connected Algebras

An algebra A is simply connected ([5], p. 355), provided it is basic, finite dimensional and it has a finite simply connected Auslander-Reiten-quiver $\Gamma(A)$. If A is simply connected its Gabriel-quiver ([4]) contains no oriented cycle so that the Tits-form ([4]) q_A is defined on $G(A) = \mathbb{Z}^s$ (see section 1). Moreover, if the global dimension of A is not more than 2, q_A is induced from the bilinear form $B: G(A) \times G(A) \to \mathbb{Z}$ which maps the pair $(\underline{\dim}\, M,\ \underline{\dim}\, N)$ to $\dim_k \operatorname{Hom}_A(M, N) - \dim_k \operatorname{Ext}_A^1(M, N) + \dim_k \operatorname{Ext}_A^2(M, N)$ ([4], [10]).

The function $U \mapsto \underline{\dim}\, U$ gives rise to a bijection between the isomorphism classes of indecomposables and the positive integral vectors $\underline{d}$ satisfying $q_A(\underline{d}) = 1$. Such a vector is called a root. The canonical base vectors $\underline{e}_i$ are the simple roots.

Lemma. *Let A be a simply connected algebra such that $\underline{d}$ is a root of q_A. Then the following is true:*

a) *$R(\underline{d})$ is the closure of an orbit of an indecomposable $u \leftrightarrow U$.*

b) *Suppose $\underline{d} = \underline{d}_1 + \underline{d}_2$ for some roots $\underline{d}_i$. Let U_i be indecomposables such that $\underline{d}_i = \underline{\dim}\, U_i$. Then U is an extension of U_1 and U_2.*

c) *If $\underline{d}$ is not simple, there exists a root $\underline{d}'$ and an index i such that $\underline{d} = \underline{d}' + \underline{e}_i$.*

Proof. Let U be the indecomposable with $\underline{\dim}\, U=\underline{d}$. We may assume that all components of $\underline{d}$ are different from zero, because the support of U ([4]) is convex ([4]), whence again simply connected ([7]). Thus, we have $\operatorname{gldim} A \leqq 2$ ([2], [10]).

Part a) of the lemma is shown in [10], 8.6 by a degeneration-argument which is based on the fact that $\operatorname{Ext}^1(V, V)$ vanishes for all indecomposables V (*I* am grateful to *P*. Gabriel who pointed out an error in the preliminary version of this note).

Next, we prove part b). Using the Auslander-Reiten formula ([1])

$$\operatorname{Ext}_A^1(M, N) = D\,\underline{\operatorname{Hom}}_A(Tr\,DN, M) \tag{$*$}$$

one constructs an oriented cycle in $\Gamma(A)$ provided that two indecomposables $U_1, U_2 \in \Gamma(A)$ satisfy $\operatorname{Ext}_A^1(U_1, U_1) \neq 0$ or $\operatorname{Ext}_A^1(U_2, U_2) \neq 0$ or $\operatorname{Ext}_A^1(U_1, U_2) \neq 0 \neq \operatorname{Ext}_A^1(U_2, U_1)$. Therefore, we may assume $\operatorname{Ext}_A^1(U_1 \oplus U_2, U_1 \oplus U_2) = \operatorname{Ext}_A^1(U_2, U_1)$. The assertion follows from part a) and Sect. 1.

Finally, let S be the symmetric bilinear form corresponding to q_A. It is sufficient to find an index i such that $q_A(\underline{d}-\underline{e}_i)=1$. Otherwise, we have $q_A(\underline{d}-\underline{e}_i) = S(\underline{d}, \underline{d}) - 2S(\underline{d}, \underline{e}_i) + S(\underline{e}_i, \underline{e}_i) \geqq 2$ for all i, because q_A is positive on positive vectors ([4]). Thus, $1 = S(\underline{d}, \underline{d}) = \sum_i \underline{d}(i)\, S(\underline{d}, \underline{e}_i)$ with $\underline{d}(i) > 0$ and $S(\underline{d}, \underline{e}_i) \leqq 0$ for all i, a contradiction.

Remarks. a) Part b) has been shown in the special case of hereditary algebras in [8] using reflection functors.

b) Suppose all components of $\underline{d}$ are different from zero. Let v be a point in $R(\underline{d})$ and put $t = \sum_i \underline{d}(i)$, $G = Gl_t(k)$. Denote the tangent space at v in $R(\underline{d})$ resp. Gv by T resp. T'. Then we get:

$$\begin{aligned}
\dim R(\underline{d}) = \dim Gu = t^2 - 1 \\
&= t^2 - \dim \operatorname{End} V + \dim \operatorname{Ext}^1(V, V) - \dim \operatorname{Ext}^2(V, V) \\
&= t^2 - \dim \operatorname{End} V + \dim T - \dim T' - \dim \operatorname{Ext}^2(V, V) \\
&= t^2 - \dim \operatorname{End} V + \dim T - (t^2 - \dim \operatorname{End} V) - \dim \operatorname{Ext}^2(V, V) \\
&= \dim T - \dim \operatorname{Ext}^2(V, V).
\end{aligned}$$

Hence v is a smooth point if and only if $\operatorname{Ext}^2(V, V) = 0$. In particular, $R(\underline{d})$ is smooth if and only if A is hereditary.

c) The bilinear form B can be used to give another short proof of the following well-known result ([10]): *Let U and V be indecomposables with* $\underline{\dim}\, U = \underline{\dim}\, V$. *If U belongs to a preprojective component* ([10]) *we get* $U \xrightarrow{\sim} V$.

Proof. Without loss of generality $\underline{\dim}\, U$ has no non-zero component so that $\operatorname{gldim} A \leqq 2$. As is shown in [2], lemma 2.2, the formula $(*)$ implies $\operatorname{pdim} U \leqq 1$ and $\operatorname{idim} U \leqq 1$. We calculate:

$$\begin{aligned}
1 &= B(\underline{\dim}\, U, \underline{\dim}\, U) = \dim \operatorname{End} U \\
&= B(\underline{\dim}\, U, \underline{\dim}\, V) = \dim \operatorname{Hom}(U, V) - \dim \operatorname{Ext}^1(U, V) \\
&= B(\underline{\dim}\, V, \underline{\dim}\, U) = \dim \operatorname{Hom}(V, U) - \dim \operatorname{Ext}^1(V, U).
\end{aligned}$$

Therefore $\operatorname{Hom}(U, V) \neq 0 \neq \operatorname{Hom}(V, U)$, thus $U \xrightarrow{\sim} V$.

3. The Proof of the Theorem

In this section we look at right A-modules. For we interpret A as k-category ([5]), thus modules as functors, and we want to get a covariant Yoneda-embedding $a \mapsto A(?, a)$ from A to the category $\operatorname{mod} A$ of finitely generated contravariant k-linear functors from A to $\operatorname{mod} k$.

By Morita-equivalence, we can assume that A is basic. Furthermore, we can suppose that the Auslander-Reiten-quiver Γ of A is connected. Let $\tilde{\Gamma}$ be its universal cover ([5]). By [11] and [5], there is a covering functor $F: k(\tilde{\Gamma}) \to \operatorname{ind} A$ from the mesh-category ([11]) to the full subcategory of $\operatorname{mod} A$ whose objects are representatives of the isomorphism classes of the indecomposable modules. Let $\tilde{A}$ be the full subcategory of $k(\tilde{\Gamma})$ consisting of projective vertices. By [5], we can interpret $k(\tilde{\Gamma})$ as $\operatorname{ind} \tilde{A}$ so that we get a commutative diagram

$$\begin{tikzcd} \tilde{A} \arrow[r] \arrow[d, "G"] & \operatorname{ind} \tilde{A} \arrow[d, "F"] \\ A \arrow[r] & \operatorname{ind} A \end{tikzcd}$$

Here, the horizontal functors are Yoneda-embeddings and the vertical ones covering functors.

Given a non-simple $U \in \operatorname{ind} A$, there is a non-simple $\tilde{U} \in \operatorname{ind} \tilde{A}$ such that $F\tilde{U} = U$. The support B of $\tilde{U}$ in $\tilde{A}$ is a finite convex full subcategory ([4]), whence a simply connected algebra by [7]. Using section 2, we obtain – up to duality – an exact sequence of B-modules $0 \to S \to \tilde{U} \to V \to 0$ where S is simple and V indecomposable. Extension by zero makes it an exact sequence of $\tilde{A}$-modules such that FS is simple and FV indecomposable. It remains to show that $0 \to FS \to F\tilde{U} \to FV \to 0$ is exact, i.e. that $0 \to FS(a) \to F\tilde{U}(a) \to FV(a) \to 0$ is exact for any $a \in A$. This follows from the following commutative diagram whose upper row is exact since $0 \to S \to \tilde{U} \to V \to 0$ is exact.

$$\begin{tikzcd} 0 \arrow[r] & \bigoplus_{G\tilde{a}=a} \operatorname{Hom}_{\tilde{A}}(\tilde{A}(?, \tilde{a}), S) \arrow[r] \arrow[d, "\wr F"] & \bigoplus_{G\tilde{a}=a} \operatorname{Hom}_{\tilde{A}}(\tilde{A}(?, \tilde{a}), \tilde{U}) \arrow[r] \arrow[d, "\wr F"] & \bigoplus_{G\tilde{a}=a} \operatorname{Hom}_{\tilde{A}}(\tilde{A}(?, \tilde{a}), V) \arrow[r] \arrow[d, "\wr F"] & 0 \\ 0 \arrow[r] & \operatorname{Hom}_A(A(?, a), FS) \arrow[r] & \operatorname{Hom}_A(A(?, a), F\tilde{U}) \arrow[r] & \operatorname{Hom}_A(A(?, a), FV) \arrow[r] & 0 \end{tikzcd}$$

References

1. Auslander, M., Reiten, I.: Representation theory of artin algebras III. Comm. Algebra **3**, 239–294 (1975)
2. Bongartz, K.: Tilted algebras, In Springer Lecture Notes in Math. **903**, pp. 26–38. Berlin-Heidelberg-New York: Springer 1981
3. Bongartz, K.: Treue einfach zusammenhängende Algebren I and II. Comment. Math. Helv. **57**, 282–300 (1982) and Unpubl. Computer list (1982)
4. Bongartz, K.: Algebras and quadratic forms. To appear in J. London Math. Soc.
5. Bongartz, K., Gabriel, P.: Covering spaces in representation theory. Invent. math. **65**, 331–378 (1982)
6. Borel, A.: Linear algebraic groups. New York-Amsterdam: Benjamin 1969

7. Bretscher, O., Gabriel, P.: The standard form of a representation-finite algebra. Bull. Soc. Math. France **111**, 21–40 (1983)
8. Dlab, V., Ringel, C.M.: A module theoretic interpretation of properties of the root systems. Proceedings of a Conference on Ring Theory (Antwerp 1978). New York-Basel: Dekker 1979
9. Gabriel, P.: Finite representation type is open. Lecture Notes in Math. **488**, pp. 132–155. Berlin-Heidelberg-New York: Springer 1975
10. Happel, D., Ringel, C.M.: Tilted algebras. Trans. Amer. Math. Soc. **274**, 399–443 (1982)
11. Riedtmann, C.: Algebren, Darstellungen, Überlagerungen und zurück. Comment. Math. Helv. **55**, 199–224 (1980)
12. Voigt, D.: Induzierte Darstellungen in der Theorie der endlichen, algebraischen Gruppen. Lecture Notes in Math. **592**. Berlin-Heidelberg-New York: Springer 1977

Received November 23, 1983

Math. Z. 187, 81–88 (1984)

Mathematische
Zeitschrift

Brauer Trees in $GL(n,q)$

Paul Fong and Bhama Srinivasan*

Department of Mathematics, University of Illinois at Chicago, Box 4348, Chicago, Illinois 60680, USA

In this paper we give an explicit description of the Brauer trees which occur in the linear groups $GL(n,p^a)$. An announcement of these results appeared in [4]. In the case that the prime r for the modular representation theory is equal to the natural characteristic p of the group, Humphreys [6] has shown that the r-blocks of $GL(n,p^a)$ have full defect or zero defect. In particular, cyclic blocks only occur in $GL(2,r)$ and the trees of such blocks can be easily found. Here we consider the case that r and p are different. Feit [3] has given a description of the trees of cyclic blocks not only for the linear groups, but for the classical groups as well. However, the scope of his work differs from ours inasmuch as his description does not identify the vertices of the graph.

The main problem here is to construct projective indecomposable modules in the block. The techniques we use include:

(1) The classification of characters into r-blocks as given in [5].

(2) Induction of characters, chiefly from subparabolic subgroups.

The trees for the unitary groups can also be explicitly described. We have not included the unitary groups in this paper. Firstly, the arguments for the linear groups are simpler. Secondly, the trees for the unitary groups fall into patterns similar to those for orthogonal and symplectic groups. The trees for these groups will be treated together in a subsequent paper.

We shall use the notation of [5], which also contains references for the facts summarized below. V is a finite-dimensional vector space of dimension n over a field F of q elements and characteristic p, $G=GL(V)$ is its group of automorphisms, and $GL(n,F)$ and $GL(n,q)$ are the natural matrix representations of G on V. Let $F[X]$ be the polynomial ring in the indeterminate X over F, and let $\mathscr{F}$ be the subset of monic, irreducible polynomials different from X. The degree of a polynomial Δ in $\mathscr{F}$ will be denoted by d_Δ. Given Δ in $\mathscr{F}$ and a semisimple element s in G, let $m_\Delta(s)$ be the multiplicity of Δ as an elementary divisor of s. The primary decomposition of s will be denoted by s

* This work was supported by grants from the NSF

$=\prod_{\Delta\in\mathscr{F}} s_\Delta$, where s_Δ is the primary component of s corresponding to Δ. Then V and $C_G(s)$ decompose as $V=\sum V_\Delta$ and $C_G(s)=\prod C_G(s)_\Delta$. Here and elsewhere, unless stated otherwise, products and sums are over the elements Δ of $\mathscr{F}$. In particular, $m_\Delta(s)$ is the multiplicity of Δ as an elementary divisor of s_Δ and $C_G(s)_\Delta \cong GL(m_\Delta(s), q^{d_\Delta})$.

Let $\bar{F}$ be an algebraic closure of F, let $\bar{G}=GL(n,\bar{F})$, and let σ be the Frobenius endomorphism of $\bar{G}$ obtained by raising matrix entries to the q-th power. Then G is the group $\bar{G}^\sigma$ of fixed-points of σ on $\bar{G}$. A subgroup L of G is regular if $L=\bar{L}^\sigma$ for some σ-stable Levi subgroup $\bar{L}$ of a parabolic subgroup $\bar{P}$ of $\bar{G}$. If $\bar{P}$ itself is σ-stable, then L is called a subparabolic subgroup of G, and $P=\bar{P}^\sigma$ a parabolic subgroup of G. If s is a semisimple element and $\sum V_\Delta$ is the decomposition of V corresponding to the primary decomposition $\prod s_\Delta$ of s, then $\prod GL(V_\Delta)$ is a subparabolic subgroup of G, and $C_G(s)$ is a regular subgroup of $\prod GL(V_\Delta)$ and of G.

Let L be a regular subgroups of G, and let $X(L)$, $X(G)$ be the character rings of representations of L and G respectively over $\bar{Q}_\ell$, where $\bar{Q}_\ell$ is an algebraic closure of the ℓ-adic field Q_ℓ for some $\ell \neq p$. The Deligne-Lusztig operator R_L^G: $X(L)\to X(G)$ has the following properties:

(3) If $K\leqq L\leqq G$ are subgroups of G such that R_K^L, R_L^G, and R_K^G are defined, then $R_L^G(R_K^L(\theta))=R_K^G(\theta)$ for θ in $X(K)$.

(4) If L and P are corresponding subparabolic and parabolic subgroups of G, then $R_L^G=\mathrm{Ind}_P^G\circ\mathrm{Inf}_L^P$, where Inf_L^P is the inflation map from $X(L)$ to $X(P)$, and Ind_P^G is the induction map from $X(P)$ to $X(G)$. In particular, $R_L^G(\theta)$ is a character of G if θ is a character of L.

(5) The degree of $R_L^G(\theta)$ is $\varepsilon_G\varepsilon_L|G:L|_{p'}\theta(1)$, where ε_G, ε_L are signs depending respectively on G and L.

(6) If θ is a unipotent or a non-unipotent irreducible character of L, then the irreducible constituents of $R_L^G(\theta)$ are respectively unipotent or non-unipotent characters of G.

Since the Weyl group W_G of G is isomorphic to the symmetric group of degree n, the irreducible characters ϕ_μ of W_G are parameterized by partitions μ of n. In turn, the unipotent characters of G are in bijection with the ϕ_μ. We shall write χ_μ for the unipotent character corresponding to ϕ_μ. This notation will be extended to products of linear groups in the following situation: Let L be a regular subgroup of the form $L=\prod_i L_i$, where $L_i\cong GL(m_i, q^{d_i})$. Let ϕ_{μ_i} and χ_{μ_i} be the characters of W_{L_i} and L_i respectively corresponding to the partition μ_i of m_i. Set $\mu=\prod_i \mu_i$. We shall write ϕ_μ and χ_μ for the characters $\prod_i \phi_{\mu_i}$ and $\prod_i \chi_{\mu_i}$ of W_L and L respectively. If L is subparabolic as well, then W_L may be viewed as a Young subgroup of W_G.

The irreducible characters of G are constructed as follows. We fix an isomorphism of $\bar{F}^\times$ into $\bar{Q}_\ell^\times$, thereby inducing an isomorphism

$$Z(L)\cong \mathrm{Hom}(L/O^{p'}(L), \bar{Q}_\ell)$$

whenever $L = C_G(s)$ and s is a semisimple element of G. Here $O^{p'}(L) \geqq [L, L]$, and equality holds if there are no components L_Δ of type $GL(2,2)$ or $GL(2,3)$. The linear character of L corresponding to s will be denoted by $\hat{s}$. The irreducible characters of G are then in bijection with G-conjugacy classes of pairs (s,μ), where s is a semisimple element of G, $\mu = \prod \mu_\Delta$, and μ_Δ is a partition of $m_\Delta(s)$. A character χ and a pair (s,μ) correspond if $\chi = \varepsilon_G \varepsilon_L R_L^G(\hat{s}\chi_\mu)$ with $L = C_G(s)$. We shall write $\chi_{s,\mu}$ for χ. In particular, $\chi_{1,\mu}$ is the unipotent character χ_μ. This notation will also be extended to direct products of linear groups in the obvious way.

Lemma A. *Let H be a subparabolic subgroup of G of the form $\prod_i GL(V_i)$, where $\sum_i V_i$ is a direct sum decomposition of V. Suppose $\chi_{s,\mu}$ and $\chi_{t,\nu}$ are irreducible characters of G and H respectively.*

(i) *If $s = t = 1$, then $(\chi_\mu, R_H^G(\chi_\nu))_G = (\phi_\mu, \mathrm{Ind}_{W_H}^{W_G}(\phi_\nu))_{W_G}$.*

(ii) *If s and t are not G-conjugate, then $(\chi_{s,\mu}, R_H^G(\chi_{t,\nu}))_G = 0$.*

Proof. This is contained in [5] (1B) and (1C).

Let r be a prime different from p. A class function of G is a generalized projective character of G if it is an integral linear combination of projective characters of G.

Lemma B. *Let L be a regular subgroup of G, and let θ be a projective character of L. The following hold:*

(i) *$R_L^G(\theta)$ is a generalized projective character of G.*

(ii) *If L is subparabolic, $R_L^G(\theta)$ is a projective character of G.*

Proof. $R_L^G(\theta)$ is a generalized character of G. Now θ vanishes on r-singular elements of L, so by [5] (2A) and the formula for $R_L^G(\theta)$ preceding [5] (2C), $R_L^G(\theta)$ vanishes on r-singular elements of G. This implies (i) by [1] Theorem 17. If L is subparabolic and P is a corresponding parabolic subgroup of G containing L, then $R_L^G(\theta) = \mathrm{Ind}_P^G(\mathrm{Inf}_L^P(\theta))$ by (4). Since $|P:L|$ is an r'-number, $\mathrm{Inf}_L^P(\theta)$ is a projective character of P. This proves (ii).

Remark. The preceding lemma holds for all finite Lie groups G of the form $G = \bar{G}^\sigma$, where $\bar{G}$ is a connected, reductive algebraic group over $\bar{F}$. In such cases character formulas for $R_L^G(\theta)$ exist (see [2] (3.1)) and these imply that $R_L^G(\theta)$ vanishes on r-singular elements of G. The proofs of (i) and (ii) are as before.

For Δ in $\mathscr{F}$, let e_Δ be the multiplicative order of q^{d_Δ} modulo r. By [5] (5D) the r-blocks B of G are in bijection with the G-conjugacy classes of pairs (s,λ), where s is a semisimple r'-element of G, $\lambda = \prod \lambda_\Delta$, and λ_Δ is the e_Δ-core of a partition of $m_\Delta(s)$. Let B be a block with cyclic defect group R, and let (s,λ) be the pair corresponding to B. The following hold as well:

(7) s may be chosen in $C(R)$.

(8) There exists a unique elementary divisor Γ of s such that

$$m_\Delta(s) = \begin{cases} |\lambda_\Gamma| + e_\Gamma & \text{for } \Delta = \Gamma \\ |\lambda_\Delta| & \text{for } \Delta \neq \Gamma. \end{cases}$$

Set $\mathscr{F}_0 = \mathscr{F} - \{\Gamma\}$, $e = e_\Gamma$, $d = d_\Gamma$, $m = m_\Gamma(s)$. For any function $f = \prod f_\Delta$, let f_0 denote the restriction of f to $\mathscr{F}_0$. In particular, $\lambda_0 = \prod_{\Delta \neq \Gamma} \lambda_\Delta$.

(9) The inertial index of B is e. Let $\chi_{s,\mu_1}, \chi_{s,\mu_2}, \ldots, \chi_{s,\mu_e}$ be the e non-exceptional characters in B. Here $(\mu_i)_0 = \lambda_0$. For simplicity of notation, set $(\mu_i)_\Gamma = \tau_i$. The τ_i's are partitions of m having λ_Γ as e-core. We arrange the τ_i so that the unique e-hook ν_i in τ_i has leg length $i-1$. The $\frac{1}{e}(|R|-1)$ exceptional characters in B are the distinct characters of the form $\chi_{sy,\mu}$, where y is in $R^\#$, $\mu = \prod \mu_\Delta$, and μ_Δ has e_Δ-core λ_Δ for Δ in $\mathscr{F}$. Let Y be a set of representatives y in $R^\#$ corresponding to the distinct exceptional characters.

Remark. In [5] the prime r was assumed to be odd. Broué has since shown that the assumption is unnecessary.

Theorem C. *The tree of B is an open polygon*

$$\underset{(s,\mu_1)}{\circ} \!-\!\!-\!\!-\! \underset{(s,\mu_2)}{\circ} \!-\!\!-\!\!-\! \circ \cdots \underset{(s,\mu_e)}{\circ} \!-\!\!-\!\!-\! \underset{\text{exc}}{\circ}$$

Proof. We may assume $r > 2$ and $e > 1$; otherwise (C) is trivial.

Case 1. Suppose $s = s_\Gamma$ and $m = e$. Then $G = GL(de, q)$, $K = GL(e, q^d)$, $\lambda_\Gamma = \{-\}$ is the empty partition, and τ_i is the e-hook ν_i of leg length $i-1$. We consider the commutative diagram

$$\begin{array}{ccc} H & \xrightarrow{R_H^G} & G \\ \uparrow{\scriptstyle R_L^H} & & \uparrow{\scriptstyle R_K^G} \\ L & \xrightarrow[R_L^K]{} & K \end{array} \tag{10}$$

where $H = GL(d, q) \times GL(de-d, q)$, $L = GL(1, q^d) \times GL(e-1, q^d)$, and the inclusions are the obvious ones. We also choose H so that H contains s and $L = C_H(s)$. Since H and L are subparabolic subgroups of G and K respectively, the horizontal maps take projective characters into projective characters by (B). Let ν_i' be the $(e-1)$-hook of leg length $i-1$, let $\{\cdot\}$ be the unique partition of 1, and let ϕ_{ν_i} and $\phi_{\{\cdot\}\times\nu_i'}$ be the corresponding characters of W_K and W_L respectively. By the Littlewood-Richardson Theorem [7] (21.3)

$$\mathrm{Ind}_{W_L}^{W_K}(\phi_{\{\cdot\}\times\nu_i'}) = \phi_{\nu_i} + \phi_{\nu_{i+1}} + \xi,$$

where ξ is a character of W_K containing no $\phi_{\nu_1}, \phi_{\nu_2}, \ldots, \phi_{\nu_e}$ as a constituent. Hence by (6) and (A)

$$R_L^K(\chi_{\{\cdot\}\times\nu_i'}) = \chi_{\nu_i} + \chi_{\nu_{i+1}} + \Xi, \tag{11}$$

where Ξ is a sum of unipotent characters of K containing no $\chi_{\nu_1}, \chi_{\nu_2}, \ldots, \chi_{\nu_e}$ as a constituent. Set $\theta = \hat{s}\chi_{\{\cdot\}\times\nu_i'}$, where $\hat{s}$ stands for the restriction of the linear character $\hat{s}$ of L to K. Since $L = C_H(s)$, it follows by the classification of characters of H that

$$R_L^H(\theta)=\varepsilon\chi_{s,\{\cdot\}\times\nu_i'}$$

for some sign ε. Moreover, since L is an r'-group, $\chi_{s,\{\cdot\}\times\nu_i'}$ has degree divisible by $|H|_r$ by (5), and so is a projective character of H. Thus

$$R_L^G(\varepsilon\theta)=R_H^G(\chi_{s,\{\cdot\}\times\nu_i'})$$

is a projective character of G by (B). On the other hand,

$$\begin{aligned} R_L^G(\varepsilon\theta)&=R_K^G(R_L^K(\varepsilon\theta)) && \text{by (3) and (10)}\\ &=R_K^G(\varepsilon\hat{s}(\chi_{\nu_i}+\chi_{\nu_{i+1}}+\Xi)) && \text{by (11)} \end{aligned}$$

(The factor $\hat{s}$ can be pulled out by [5] (1.9) since s in $Z(K)$.) By (9), (11), and the classification of characters of G, no constituent of $R_K^G(\varepsilon\hat{s}\Xi)$ is in B. Thus the component of $R_L^G(\varepsilon\theta)$ in B is

$$(R_L^G(\varepsilon\theta))_B=\chi_{s,\nu_i}+\chi_{s,\nu_{i+1}}.$$

This component is a projective character and hence a projective indecomposable character. So

(12) $$\underset{(s,\mu_1)}{\circ}\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\underset{(s,\mu_2)}{\circ}\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\circ\ \cdots\ \underset{(s,\mu_e)}{\circ}$$

is a subgraph of the tree.

To finish the proof for this case, it will be enough to show (s,μ_e) is not an end-point of the tree. Now

(13) $$\chi_{s,\mu_i}(1)=|G:K|_{p'}\chi_{\nu_i}(1)$$

by (5), and

(14) $$\frac{\chi_{\nu_e}(1)}{\chi_{\nu_{e-1}}(1)}=\frac{q^{de-d}(q^d-1)}{q^{de-d}-1}>1$$

by the hook length formula [5] (1.15) for degrees of unipotent characters of G. If (s,μ_e) were an end-point, then $\chi_{\nu_{e-1}}(1)\geqq\chi_{\nu_e}(1)$, which is impossible by (13), (14).

Case 2. Suppose $s=1$. Then $G=GL(m,q)$. Let H be a subparabolic subgroup of G of the form $GL(m-e,q)\times GL(e,q)$, chosen so that R is contained in the second factor $GL(e,q)$. Let $\phi_{\lambda_\Gamma\times\nu_i}$ and ϕ_{τ_i} be the irreducible characters of W_H and W_G corresponding respectively to $\lambda_\Gamma\times\nu_i$ and τ_i. The Littlewood-Richardson Theorem implies

(15) $$\mathrm{Ind}_{W_H}^{W_G}(\phi_{\lambda_\Gamma\times\nu_i})=\phi_{\tau_i}+\xi,$$

where ξ is a character of W_G containing no $\phi_{\tau_1},\phi_{\tau_2},\dots,\phi_{\tau_e}$ as a constituent. Indeed, by [7] (21.5) the multiplicity of ϕ_ρ, where ρ is a partition of m, is zero if λ_Γ is not a subpartition of ρ or if $\rho-\lambda_\Gamma$ is not a union of skew hooks. In the remaining cases, the multiplicity of ϕ_ρ is the binomial coefficient $\binom{t-1}{c-e+i-1}$, where t is the number of skew hooks in $\rho-\lambda_\Gamma$, and c is the number of columns in the t skew hooks. When $\rho=\tau_j$, then $t=1$, $c=e-j+1$, and the binomial coefficient is δ_{ij}.

By (A) it follows that

$$R_H^G(\chi_{\lambda_\Gamma}\times\chi_{\nu_i})=\chi_{\tau_i}+\Xi_i, \tag{16}$$

where Ξ_i is a sum of unipotent characters of G containing no $\chi_{\tau_1}, \chi_{\tau_2}, \ldots, \chi_{\tau_e}$ as a constituent. Now χ_{λ_Γ} is a projective character of $GL(m-e, q)$, since $\chi_{\lambda_\Gamma}(1)_r = |GL(m-e, q)|_r$ by the hook length formula [5] (1.15). Also, $\chi_{\nu_i}+\chi_{\nu_{i+1}}$ is a projective character of $GL(e, q)$ by Case 1. Thus by (B)

$$R_H^G(\chi_{\lambda_\Gamma}\times(\chi_{\nu_i}+\chi_{\nu_{i+1}}))$$

is a projective character of G. By (9) and (16)

$$(R_H^G(\chi_{\lambda_\Gamma}\times(\chi_{\nu_i}+\chi_{\nu_{i+1}})))_B=\chi_{\tau_i}+\chi_{\tau_{i+1}},$$

so $\chi_{\tau_i}+\chi_{\tau_{i+1}}$ is necessarily a projective indecomposable character. Thus (12) is a subgraph of the tree. Now

$$\theta=\chi_{\lambda_\Gamma}\times(\chi_{\nu_e}+\sum_{y\in Y}\chi_{y,\{\cdot\}})$$

is a projective character of H, since the second factor is a projective character of $GL(e, q)$ by Case 1. So $R_H^G(\theta)$ is a projective character. Since χ_{τ_e} is the only non-exceptional constituent in $(R_H^G(\theta))_B$ by (6) and (16), the exceptional vertex must be linked to (s, μ_e).

Case 3. The general case. Let $\prod s_\Delta$ be the primary decomposition of s, and let $\sum V_\Delta$ and $\prod K_\Delta$ be the corresponding decompositions of V and $K=C_G(s)$. In particular, $K_\Gamma\cong GL(m, q^d)$. Let H_Γ be a subparabolic subgroup of $GL(V_\Gamma)$ of the form $GL(dm-de, q)\times GL(de, q)$, chosen so that H_Γ contains s_Γ, and so that

$$C_{H_\Gamma}(s_\Gamma)=GL(m-e, q^d)\times GL(e, q^d).$$

Let $L_\Gamma=C_{H_\Gamma}(s_\Gamma)$, $L=L_\Gamma\prod_{\Delta\neq\Gamma}K_\Delta$, and $H=H_\Gamma\prod_{\Delta\neq\Gamma}GL(V_\Delta)$. Then L and H are subparabolic subgroups respectively of K and G. We consider the commutative diagram (10) for these choices of H, K, and L.

Let b be the r-block of K corresponding to the pair $(1, \lambda)$. So $b=\prod b_\Delta$, where b_Δ is the r-block of K_Δ parametrized by $(1, \lambda_\Delta)$. The facts (7), (8), (9) apply to each b_Δ. In particular, b_Δ has defect 0 for $\Delta\neq\Gamma$, and b has defect group R and inertial index e. The non-exceptional characters in b have the form $\chi_{\mu_i}=\prod(\chi_{\mu_i})_\Delta$, where

$$(\chi_{\mu_i})_\Delta=\begin{cases}\chi_{\lambda_\Delta} & \text{for } \Delta\neq\Gamma\\ \chi_{\tau_i} & \text{for } \Delta=\Gamma.\end{cases}$$

We shall write χ_{μ_i} as $\chi_{\lambda_0}+\chi_{\tau_i}$, where χ_{λ_0} is the unipotent character of K_0 corresponding to λ_0. Case 2 applies to b with K in place of G, L in place of H. In particular, (16) becomes

$$(17) \qquad (R_L^K(\chi_{\lambda_0}\times\chi_{\lambda_\Gamma\times\nu_i}))_b=\chi_{\mu_i}.$$

Let $\theta_i=\chi_{\lambda_0}\times\chi_{\lambda_\Gamma\times\nu_i}$. Then

$$(18) \qquad \begin{aligned}(R_L^G(\hat{s}\theta_i))_B &= (R_K^G(\hat{s}R_L^K(\theta_i)))_B\\ &= R_K^G(\hat{s}(R_L^K(\theta))_b)\\ &= R_K^G(\hat{s}\chi_{\mu_i}) \quad \text{by (17)}\\ &= \varepsilon\chi_{s,\mu_i}\end{aligned}$$

for some sign ε independent of i. Here the classification of characters in B and in b is used for the second equality. On the other hand, since $L=C_H(s)$, the classification of characters of H implies

$$R_L^H(\hat{s}\theta_i)=\varepsilon'\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_i)}$$

for some sign ε'. Since $R_H^G(R_L^H(\hat{s}\theta_i))=R_K^G(R_L^K(\hat{s}\theta_i))$, the signs $\varepsilon, \varepsilon'$ are the same by (4) and (18). Thus

$$(19) \qquad R_L^H(\varepsilon\hat{s}\theta_i)=\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_i)}.$$

Let B' be the block of H corresponding to the pair $(s, \lambda_0\times(\lambda_\Gamma\times\{-\}))$. So $B'=\prod B'_\Delta$, where B'_Δ is the r-block of H_Δ parametrized by $(s_\Delta, \lambda_\Delta)$ for $\Delta\neq\Gamma$, and B'_Γ is the r-block of H_Γ parametrized by $(s_\Gamma, \lambda_0\times\{-\})$. In particular, B' has defect group R and inertial index e. The non-exceptional characters in B' are the $\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_i)}$. The exceptional characters in B' are the $\chi_{sy,\lambda_0\times\lambda_\Gamma\times\{\cdot\}}$, where the partition $\{\cdot\}$ is associated to the unique elementary divisor of $s_\Gamma y$. By Case 1

$$R_L^H(\varepsilon\hat{s}(\theta_i+\theta_{i+1}))=\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_i)}+\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_{i+1})}$$

is a projective character of H. Hence $R_L^G(s(\theta_i+\theta_{i+1}))$ is a projective character of G by (B). Since

$$(R_L^G(\varepsilon\hat{s}(\theta_i+\theta_{i+1})))_B=\chi_{s,\mu_i}+\chi_{s,\mu_{i+1}}$$

by (18), it follows as before that (12) is a subgraph of the tree. Again by Case 1

$$\theta=\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_e)}+\sum_{y\in Y}\chi_{sy,\lambda_0\times\lambda_\Gamma\times\{\cdot\}}$$

is a projective character of H. Thus $R_H^G(\theta)$ is a projective character of G. Now

$$(R_H^G(\chi_{s,\lambda_0\times(\lambda_\Gamma\times\nu_e)}))_B=\chi_{s,\mu_e}$$

by (18), (19), and χ_{s,μ_e} is the only non-exceptional constituent in $(R_H^G(\theta))_B$ by (A). Thus (s, μ_e) is linked to the exceptional vertex and (C) is proved.

References

1. Brauer, R.: A characterization of the characters of groups of finite order. Ann. of Math. **57**, 357–377 (1953)
2. Digne, F., Michel, J.: Theorie de Deligne-Lusztig et caractères des groupes linéaires et unitaires, Équipe de Rechèrche associée au CNRS no. 944, February 1983
3. Feit, W.: Possible Brauer trees. Illinois J. Math. **28**, 43–56 (1984)
4. Fong, P., Srinivasan, B.: Blocks with cyclic defect groups in $GL(n, q)$. Bull. Amer. Math. Soc. **3**, 1041–1044 (1980)
5. Fong, P., Srinivasan, B.: The blocks of finite general linear and unitary groups. Invent. Math. **69**, 109–153 (1982)
6. Humphreys, J.: Defect groups for finite groups of Lie type. Math. Z. **119**, 149–152 (1971)
7. James, G.: The representation theory of the symmetric groups. Lecture Notes in Math. **682**. Berlin-Heidelberg-New York: Springer 1978

Received January 24, 1984

Math. Z. 187, 89–96 (1984)

Mathematische Zeitschrift

On the Morse Complex

Xiao-Wei Peng

Mathematisches Institut der Universität Bonn, Wegelerstraße 10, D-5300 Bonn 1, Federal Republic of Germany

Let M be a riemannian manifold, the dimension of M may be finite or infinite. f is a Morse function on M which is bounded from below and has only non-degenerate critical points, every critical point has only finite index. For a non-degenerate critical point P^λ (we often use P^λ to denote critical point with index λ), we have a Morse-neighborhood $U(P^\lambda)$ of the point P^λ such that in this neighborhood the Morse function f can be represented by

$$f(x, y) = -\|x\|^2 + \|y\|^2 + f(P^\lambda)$$

(for the finite dimensional case to see Milnor [6] p. 6, for the infinite dimensional case to see Palais [8] p. 307). For the Morse function f we can introduce a "gradient-like" vector field ξ, that is $\xi(f) > 0$ except critical points and $\xi = (-x, y)$ on the Morse neighborhood.

Let φ_t be the flow of the gradient-like vector field ξ. To discuss the behaviour of the flow we can find some submanifolds of M, stable submanifolds $W_s(P^\lambda)$ and unstable submanifolds $W_u(P^\lambda)$ where P^λ is a critical point with index λ. The geometrical meaning of the submanifolds $W_s(P^\lambda)$, $W_u(P^\lambda)$ is that it consists of the points which go towards P^λ or go away from P^λ under the flow. The explicit definition is following:

$$W_s(P^\lambda) := \{P \in M \mid \varphi_t(P) \to P^\lambda \text{ as } t \to +\infty\}$$

$$W_u(P^\lambda) := \{P \in M \mid \varphi_t(P) \to P^\lambda \text{ as } t \to -\infty\}.$$

To see Klingenberg [2] p. 67 or Abraham and Robbin [1] p. 84. Klingenberg used the vector field $-\xi$. This yields the reverse definition of W_s and W_u.

The question is, how to use the stable submanifolds to describe the topological properties of the manifold M. We will define another complex, the Morse complex, using stable submanifolds which has a useful geometrical interpretation. It has the same homology groups as M. See also Klingenberg [4] p. 69.

Let f be the Morse function and ξ be a gradient-like vector field. One can find a new gradient-like vector field ξ_1 and a new Morse function f_1 such that

ξ_1 induces stable and unstable submanifolds having transverse intersection and f_1 self-indexing.

The method and results for the finite dimensional case belong to Milnor, cf. Milnor [7]. We use a simple generalization of Milnor's results and use these results to define the Morse complex.

I wish to thank Prof. Klingenberg for suggesting the problem and giving me useful advice.

1. The Existence of Stable and Unstable Submanifolds

First we like briefly to discuss the properties of W_s, W_u, which defined previously.

Theorem 1. *If ξ is a gradient-like vector field of f then W_s and W_u are submanifolds of M.*

Remark. In general we know that W_u and W_s are only immersed submanifolds. For the finite dimensional case see Abraham and Robbin [1] p. 84.

Proof. Since $\frac{d}{dt} f(\varphi(t, P)) = \xi(f) > 0$ (P is not a critical point of f), $f(\varphi(t, P))$ is a monotonously increasing function of variable t. It follows

$$\begin{aligned} W_s(P^\lambda) \cap U &= \{(x, 0) \mid (x, 0) \in U\} \\ W_u(P^\lambda) \cap U &= \{(0, y) \mid (0, y) \in U\}. \end{aligned} \tag{$*$}$$

Here U is the Morse neighborhood of P^λ. Cf. Fig. 1.

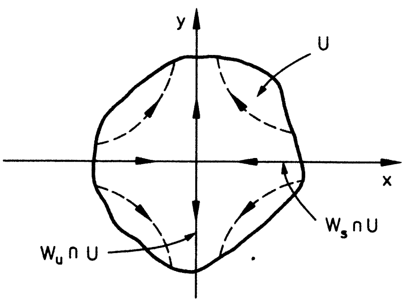


Fig. 1

Now we can define the global immersion

$$F_s\colon R^\lambda \to M, \qquad F_s(x) = \begin{cases} P^\lambda & x = 0 \\ \varphi\left(-\log \|x\|, \left(\frac{x}{\|x\|}, 0\right)\right), & x \neq 0 \end{cases}$$

$$F_u\colon H \to M, \qquad F_u(y) = \begin{cases} P^\lambda & y = 0 \\ \varphi\left(\log \|y\|, \left(0, \frac{y}{\|y\|}\right)\right), & y \neq 0. \end{cases}$$

It is easy to check that F_s and F_u are an injective immersion and $F_s(R^\lambda) = W_s(P^\lambda)$, $F_u(H) = W_u(P^\lambda)$, where H is model Hilbert space of M.

(*) means that W_u, W_s have the submanifold property near P^λ, and F_s, F_u are defined by a flow, hence, W_u, W_s have the submanifold property. This prove the theorem. □

2. Transversality

In general we don't know whether W_u, W_s have transversal intersection. But we can find another gradient-like vector field ξ_1 which determining W_u, W_s with transversal intersection.

Theorem 2. *There is a gradient-like vector field ξ_1 on M such that all stable and unstable submanifolds, which are induced by ξ_1, intersect transversally.*

To prove this theorem we need a Lemma.

Lemma 3. *If V, W are submanifolds of M and dimension of V is finite, codimension of W is also finite and V is compact. Then there is an isotopy h_t of M such that $h_1(V) \pitchfork W$.*

$h_1(V) \pitchfork W$ means $h_1(V)$ and W have transversal intersection.

Proof. For the finite dimensional case this was proved by Milnor [7], p. 40.

Let (U, ψ) be a tubular neighborhood of V in M, $\psi(U) = V \times H$, where H is a Hilbert space. This exists since V is compact.

We consider the following diagramm:

$$\begin{tikzcd} V \times H \arrow[r, "\pi"] & H \\ U \arrow[u, "\psi"] \arrow[ur, "F = \pi \circ \psi"'] & \end{tikzcd}$$

π is the projection of second functor. We put

$$W_1 = W \cap U, \quad F_1 = F \mid W_1.$$

Clearly,

1) $\dim \ker DF_1 \leqq \dim V < \infty$
2) $\dim \operatorname{coker} DF_1 \leqq \operatorname{codim} W < \infty$

this means that F_1 is a Fredholm mapping (generally, a Fredholm operator should satisfy another condition 3) $DF_1(T_P W_1)$ is closed in $T_{F_1(P)}H$. But we can show that 3) is a consequence of 1) and 2) without difficulty). We now use an infinite dimensional version of Sard's theorem which was proved by Smale, Smale [9]. It follows, that the regular values of F_1 form a residual set of H. Let x_0 be a regular value of F_1 in H. From the definition of the regular value we have

$$\psi^{-1}(V \times \{x_0\}) \pitchfork W.$$

Now we would like to construct the isotopy h_t. Let η be a vector field on H satisfying following conditions:

1) $\eta(x)=x_0, \quad \|x\| \leqq \|x_0\|$,
2) $\eta(x)=0, \quad \|x\| \leqq 2\|x_0\|$,
3) $0 \leqq \|\eta(x)\| \leqq \|x_0\|, \quad$ for all $x \in H$.

Let θ_t be the flow of η. Since η bounded and H is complete, θ_t also is complete that is to say θ_t is defined on all of $R \times H$. Now we can define the h_t with the help of θ_t:

$$h_t(w)=\begin{cases} w, & w \notin U \\ \psi^{-1}(P, \theta_t(x)), & w=\psi^{-1}(P, x) \in U. \end{cases}$$

h_t is a isotopy between h_1 and $h_0 = \mathrm{id}$ and we have

$$h_1(V)=\psi^{-1}(V \times \{x_0\})$$

which completes the proof. □

To prove Theorem 2 we can use Milnor's method, Lemma 3 and results of Palais (see Milnor [7], p. 42, Palais [8], p. 309) to modify the gradient-like vector field between two critical levels. Via induction we obtain the global construction. Here we omit the proof.

3. Self-Indexing Morse Function

In this section we make another hypothesis for the Morse function f. First, f shall satisfy the condition C. From the condition C we can obtain many informations about the critical set. For example, $f^{-1}[\alpha, \beta]$ consists of only finite critical points (see Palais [8], p. 313). Secondly, for every index λ there are only finitely many critical points with this index λ. If we consider only $f^{-1}[\alpha, \beta]$, where α, β are finite real numbers, this is not the whole manifold M. Then this assumption is not necessary, see our previous discussion.

Theorem. *Let f be a Morse function and ξ be a gradient-like vector field, which induces W_u, W_s with transversal intersection. Then there exists a new Morse function f_1 on M such that f_1 is self-indexing and f, f_1 has the same critical points and the same index for every critical point.*

Proof. Milnor [7], p. 37, has proved this for the finite dimensional case, but we can generalize the method of the proof to the infinite dimensional case without difficulty, since the construction is independent of the dimension. Only, we should pay attention to the fact that in the infinite dimensional case $f^{-1}[\alpha, \beta]$ is not compact. Instead of compactness we now use the condition C. Milnor's construction is only locally. By induction and limits we get a global construction. The new Morse function f_1 always satisfies condition C.

4. The Morse Complex

Let S_λ be free abelian group of the critical points with index λ, for $\lambda = 0, 1, 2, 3, \ldots$.

We use the result of 2, 3. Suppose f is self-indexing. Let

$$a_0 < a_1 < a_2 < a_3 < a_4 < \ldots < a_n < \ldots$$

be the middle value of two critical values, i.e., there is a critical value in $(a_{\lambda-1}, a_\lambda)$ and $f^{-1}[a_{\lambda-1}, a_\lambda]$ contains only critical points with index λ.

Let

$$M_\lambda = f^{-1}(-\infty, a_\lambda], \qquad C_\lambda = H_\lambda(M_\lambda, M_{\lambda-1}).$$

From the work of Palais [8], p. 335, we have

$$H_\lambda(M_\lambda, M_{\lambda-1}) \cong S_\lambda$$
$$H_k(M_\lambda, M_{\lambda-1}) = 0, \qquad k \neq \lambda.$$

There is a natural boundary homomorphism ∂_λ for C_λ, which is induced by the following diagram:

$$\underbrace{H_\lambda(M_\lambda, M_{\lambda-1}) \xrightarrow{\partial_*} H_{\lambda-1}(M_{\lambda-1}) \xrightarrow{i_*} H_{\lambda-1}(M_{\lambda-1}, M_{\lambda-2})}_{\partial_\lambda}.$$

Since of $C_\lambda \approx S_\lambda$, we also have a boundary homomorphism $\bar{\partial}_\lambda$ on S_λ which is induced by following diagram:

$$\begin{array}{ccc} C_\lambda & \xrightarrow{\approx} & S_\lambda \\ \downarrow{\scriptstyle \partial_\lambda} & & \downarrow{\scriptstyle \bar{\partial}_\lambda} \\ C_{\lambda-1} & \xrightarrow{\approx} & S_{\lambda-1} \end{array}$$

Theorem 5. $\qquad H_\lambda(S) = H_\lambda(M), \qquad \lambda = 0, 1, 2, \ldots.$

Proof. Clearly,

$$H_\lambda(S) = H_\lambda(C), \qquad \lambda = 0, 1, 2, \ldots$$

We need only to show

$$H_\lambda(S) = H_\lambda(M).$$

1) We note the following facts:

$$H_\lambda(M_k) = \begin{cases} 0, & \lambda > k \\ H_\lambda(M), & \lambda < k. \end{cases}$$

From $H_k(M_\lambda, M_{\lambda-1}) = 0$, $k \neq \lambda$, it follows

$$H_\lambda(M_{\lambda+1}) = H_\lambda(M_{\lambda+2}) = \ldots$$
$$H_\lambda(M_{\lambda-1}) = H_\lambda(M_{\lambda-2}) = \ldots = H_\lambda(M_0) = 0.$$

We use direct limits of algebraic groups,

$$H_\lambda(M) = \text{direct} \lim_{n\to\infty} H_\lambda(M_n)$$

(to see Massey [5] p. 381).

2) We consider the two triples $(M_{\lambda+1}, M_\lambda, M_{\lambda-2})$ and $(M_\lambda, M_{\lambda-1}, M_{\lambda-2})$, and construct from exact sequences of the homology groups of the tripel the diagram

$$\begin{tikzcd}
 & 0 \arrow[d] & & \\
H_{\lambda+1}(M_{\lambda+1}, M_\lambda) \arrow[r, "\alpha"] \arrow[dr, "\partial_{\lambda+1}"'] & H_\lambda(M_\lambda, M_{\lambda-2}) \arrow[r, "\beta"] \arrow[d, "\gamma"] & H_\lambda(M_{\lambda+1}, M_{\lambda-2}) \arrow[r] & 0 \\
 & H_\lambda(M_\lambda, M_{\lambda-1}) \arrow[d, "\partial_\lambda"] & & \\
 & H_{\lambda-1}(M_{\lambda-1}, M_{\lambda-2}) & &
\end{tikzcd}$$

γ is a monomorphism. β is an epimorphism. From the diagram we have the following properties:

$$\operatorname{Ker} \partial_\lambda = \operatorname{Im} \gamma \cong H_\lambda(M_\lambda, M_{\lambda-2})$$
$$\operatorname{Im} \partial_{\lambda+1} = \operatorname{Im} \gamma \circ \alpha = \beta \circ \operatorname{Im} \alpha \approx \operatorname{Im} \alpha$$
$$H_\lambda(C) = \operatorname{Ker} \partial_\lambda / \operatorname{Im} \partial_{\lambda+1} \cong H_\lambda(M_\lambda, M_{\lambda-2}) / \operatorname{Im} \alpha \cong H_\lambda(M_{\lambda+1}, M_{\lambda-2}).$$

Since clearly

$$H_\lambda(M_{\lambda+1}, M_{\lambda-2}) = H_\lambda(M_{\lambda+1}) = H_\lambda(M)$$

now we have our result

$$H_\lambda(S) = H_\lambda(M), \qquad \lambda = 0, 1, 2, \ldots .$$

If all $W_s(P^\lambda)$ are oriented, how can we describe the geometrical meaning of boundary homomorphism? For such a description we introduce some notations:

$$S_s(P^{\lambda+1}) = W_s(P^{\lambda+1}) \cap f^{-1}(a_\lambda),$$
$$S_u(P^\lambda) = W_u(P^\lambda) \cap f^{-1}(a_\lambda),$$
$$S_s \cap S_u = \{q_1, \ldots, q_j\}.$$

The U_i are small neighborhoods of the q_i in S_s and are pairwise disjoint. Let

$$D^{\lambda+1}(P^{\lambda+1}) = W_s(P^{\lambda+1}) \cap f^{-1}[a_\lambda, a_{\lambda+1}],$$
$$D^\lambda(P^\lambda) = W_u(P^\lambda) \cap f^{-1}[a_{\lambda-1}, a_\lambda],$$

be the generators of $H_{\lambda+1}(M_{\lambda+1}, M_\lambda)$ and $H_\lambda(M_\lambda, M_{\lambda-1})$, with given orientation. S_s has the orientation which is induced by $D^{\lambda+1}$, for U_i we take the same orientation as for S_s.

r is deformation retract from M_λ to $M_{\lambda-1} \cup D^\lambda$ (cf. Palais [8]). From the construction of the deformation r we know that

$$r_*: H_\lambda(U_i, U_i - \{q_i\}) \to H_\lambda(D^\lambda, D^\lambda - \{P^\lambda\})$$

is an isomorphism.

$r_*[U_i] = a_i[D^\lambda]$

$$a_i = \begin{cases} 1, & \text{if } r(U_i),\ D^\lambda \text{ have the same orientation,} \\ -1, & \text{if } r(U_i),\ D^\lambda \text{ have opposite orientation.} \end{cases}$$

Here $[U_i]$, $[D^\lambda]$ are generators of $H_\lambda(U_i, U_i - \{q_i\})$, $H_\lambda(D^\lambda, D^\lambda - \{P^\lambda\})$ with the given orientation. Let

$$A = \sum_{i=1}^{j} a_i.$$

Then we call A intersection number of $P^{\lambda+1}, P^\lambda$.

Let A_{lk} be intersection number of $P_l^{\lambda+1}, P_k^\lambda$, then the boundary homomorphism $\partial_{\lambda+1}$ are described by the matrix (A_{lk}), it means

$$\partial_{\lambda+1} P_l^{\lambda+1} = \sum_k A_{lk} P_k^\lambda$$

see Klingenberg [4].

The previous assertion we get from the following diagram:

$$\begin{tikzcd}
 & H_\lambda(S_s) \arrow[r, "3"] \arrow[d, "2"] & H_\lambda(S_s, S_s - \bigcup \{q_i\}) \arrow[r, "4"] & \sum_i H_\lambda(U_i, U_i - \{q_i\}) \arrow[d, "r_*"] \\
H_{\lambda+1}(M_{\lambda+1}, M_\lambda) \arrow[r, "\partial_*"] \arrow[dr, "\partial_{\lambda+1}"'] & H_\lambda(M_\lambda) \arrow[d, "1"] & & H_\lambda(D^\lambda, D^\lambda - \{P^\lambda\}) \arrow[d, "\approx\ 5"] \\
 & H_\lambda(M_\lambda, M_{\lambda-1}) & & H_\lambda(M_{\lambda-1} \cup D^\lambda, M_{\lambda-1}) \arrow[ll, "6\ \approx"']
\end{tikzcd}$$

1, 2, 3, are induced by inclusion, 4, 5, 6 are induced by excision. For this diagram we also suppose, $H_\lambda(M_\lambda, M_{\lambda-1})$, $H_{\lambda+1}(M_{\lambda+1}, M_\lambda)$ have only one generator.

References

1. Abraham, R., Robbin, J.: Transversal mappings and flows. New York and Amsterdam: Benjamin 1967
2. Klingenberg, W.: Lectures on closed geodesics. Grundlehren Math. Wiss. **203**. Berlin-Heidelberg-New York: Springer 1978

3. Klingenberg, W.: Riemannian geometry. De Gruyter Studies in Math. 1. Berlin-New York: de Gruyter 1982
4. Klingenberg, W.: Closed geodesics on Riemannian manifolds. Conf. Board of the Math. Sciences. Regional conf. series in Math. **53**. Providence: Amer. Math. Soc. 1983
5. Massey, W.: Homology and Cohomology Theory, An Approach Based on Alexander-Spanier Cochains. New York: Dekker 1978
6. Milnor, J.: Morse Theory. Ann. Math. Studies **51**. Princeton, N.J.: Princeton Univ. Press 1963
7. Milnor, J.: Lectures on the *h*-cobordism Theorem. Princeton Math. Notes. Princeton, N.J.: Princeton Univ. Press 1965
8. Palais, R.: Morse Theory on Hilbert manifolds. Topology **2**, 299–340 (1963)
9. Smale, S.: An infinite dimensional version of Sard's Theorem. Amer. J. Math. **87**, 861–866 (1965)

Received January 17, 1984

Math. Z. 187, 97–103 (1984)

Mathematische Zeitschrift

Dilation Theory and Systems of Simultaneous Equations in the Predual of an Operator Algebra. II

Hari Bercovici[1], Bernard Chevreau[2], Ciprian Foias[3], and Carl Pearcy[4]

[1] Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
[2] U.E.R. de Mathematique et d'Informatique, Universite de Bordeaux I, France
[3] Department of Mathematics, Indiana University, Bloomington, Indiana 47405, USA
[4] Department of Mathematics, The University of Michigan, 347 West Engineering Building, Ann Arbor, Michigan 48109, USA

1. This note is a continuation of our earlier paper [3], in which we developed a dilation theory for a certain class of contraction operators acting on a separable, infinite dimensional, complex Hilbert space $\mathscr{H}$. The notation and terminology in what follows is taken from [3]. For the convenience of the reader we recall a few pertinent definitions. The algebra of bounded linear operators on $\mathscr{H}$ is denoted by $\mathscr{L}(\mathscr{H})$. If $T\in\mathscr{L}(\mathscr{H})$, the ultraweakly closed algebra generated by T and $1_{\mathscr{H}}$ is denoted by $\mathscr{A}_T$; we recall that $\mathscr{A}_T$ can be identified with the dual space of the quotient space $Q_T=(\tau c)/{}^{\perp}\mathscr{A}_T$, where (τc) denotes the ideal of trace-class operators in $\mathscr{L}(\mathscr{H})$ and ${}^{\perp}\mathscr{A}_T$ is the pre-annihilator of $\mathscr{A}_T$ in (τc), under the pairing

$$\langle A, [L]\rangle=\operatorname{tr}(AL), \qquad A\in\mathscr{A}_T,\ [L]\in Q_T.$$

The open unit ball in $\mathbb{C}$ is denoted by $\mathbb{D}$, and we write $\mathbb{T}=\partial\mathbb{D}$. The class $\mathbb{A}(\mathscr{H})\subset\mathscr{L}(\mathscr{H})$ consists of all those absolutely continuous contractions T (i.e., all those contractions T whose unitary part is absolutely continuous or acts on the space (0)) such that the Sz.-Nagy-Foias functional calculus $\Phi_T\colon H^\infty(\mathbb{T})\to\mathscr{A}_T$ is an isometry. If $T\in\mathbb{A}(\mathscr{H})$ then Φ_T is the adjoint of an isometry ϕ_T of Q_T onto the predual $L^1(\mathbb{T})/H^1_0(\mathbb{T})$ of $H^\infty(\mathbb{T})$ (cf. [3, 5]), and via the pair $\{\phi_T,\Phi_T\}$, the pair of spaces $\{L^1(\mathbb{T})/H^1_0(\mathbb{T}), H^\infty(\mathbb{T})\}$ can be identified with the pair $\{Q_T,\mathscr{A}_T\}$.

If $x, y\in\mathscr{H}$, we write $x\otimes y$ for the rank-one operator defined, as usual, by $(x\otimes y)(u)=(u,y)x$, $u\in\mathscr{H}$. Of course, $x\otimes y\in(\tau c)$, and if some $T\in\mathscr{L}(\mathscr{H})$ is given, we write $[x\otimes y]_{Q_T}$ (or simply $[x\otimes y]$ when no confusion can result) for the image of $x\otimes y$ in Q_T. If n is any cardinal number satisfying $1\leqq n\leqq\aleph_0$, we denote by $\mathbb{A}_n(\mathscr{H})$ the set of all those T in $\mathbb{A}(\mathscr{H})$ for which every system of simultaneous equations

$$[x_i\otimes y_j]=[L_{ij}], \qquad 0\leqq i,\ j<n$$

(where the $[L_{ij}]$ are arbitrary elements from Q_T) has a solution $\{x_i\}_{0\leq i<n}$, $\{y_j\}_{0\leq j<n}$. When no confusion will result, we write simply $\mathbb{A}_n$ for $\mathbb{A}_n(\mathscr{H})$. In [3] we began the structure theory of the classes $\mathbb{A}_n$, and, in particular, the dilation theory of the class $\mathbb{A}_{\aleph_0}$. A primary motivation for the introduction of these classes in [3] was as follows. Let $(BCP)=(BCP)(\mathscr{H})$ denote the set of all those completely nonunitary contractions T in $\mathscr{L}(\mathscr{H})$ for which the intersection $\sigma_e(T)\cap\mathbb{D}$ of the essential spectrum of T with $\mathbb{D}$ is sufficiently large that almost every point of $\mathbb{T}$ is a non-tangential limit point of $\sigma_e(T)\cap\mathbb{D}$ (such sets are said to be dominating for $\mathbb{T}$). It was shown in [4] (and also in [7]) that $(BCP)\subset\mathbb{A}_{\aleph_0}$, so all of the results obtained in [3] for operators in $\mathbb{A}_{\aleph_0}$ apply, in particular, to (BCP)-operators. (In fact, in [4], an increasing family $\{(BCP)_\theta\}_{0\leq\theta\leq 1}$ of classes of contractions is introduced, with $(BCP)=(BCP)_0$, and it was shown there that $\bigcup_{0\leq\theta<1}(BCP)_\theta\subset\mathbb{A}_{\aleph_0}$.)

In [2] it was shown that all (BCP)-operators are reflexive, and the main purpose of this note is to show that all operators in the larger class $\mathbb{A}_{\aleph_0}$ are reflexive (Theorem 1.7). This is worthwhile because we show in the third paper [1] of this sequence that many familiar operators belong to $\mathbb{A}_{\aleph_0}$ and thus are reflexive. In particular, we will show in [1] on the basis of Theorem 1.7 that every weighted unilateral shift W that is a contraction whose spectrum satisfies $\sigma(W)\supset\mathbb{T}$ is reflexive, thus generalizing considerably the results on reflexivity of [8].

We write $\mathrm{Lat}(T)$ for the lattice of invariant subspaces of an operator T, and if $\mathscr{M},\mathscr{N}\in\mathrm{Lat}(T)$ with $\mathscr{M}\supset\mathscr{N}$, so $\mathscr{M}\ominus\mathscr{N}$ is a semi-invariant subspace of T, we write $T_{\mathscr{M}\ominus\mathscr{N}}$ for the compression of T to this semi-invariant subspace. We also write $P_{\mathscr{M}}$ for the (orthogonal) projection whose range is a subspace $\mathscr{M}$. Our principal tool is the following result of Robel [7, Proposition 6.1].

Proposition 1.1. *Suppose $T\in(BCP)(\mathscr{H})$, $y\in\mathscr{H}$, and $\varepsilon>0$. Then there exists a subspace $\mathscr{M}\subset\mathscr{H}$ such that $\mathscr{M}\in\mathrm{Lat}(T)$, $T|\mathscr{M}\in(BCP)(\mathscr{M})$, $T_{\mathscr{H}\ominus\mathscr{M}}\in(BCP)(\mathscr{H}\ominus\mathscr{M})$, and $\|P_{\mathscr{M}}y\|<\varepsilon$.*

We will also need the following easy lemma.

Lemma 1.2. *Suppose T is a completely nonunitary contraction in $\mathscr{L}(\mathscr{H})$ and $\{\lambda_n\}_{n=1}^\infty$ is a sequence in $\mathbb{D}$ that is dominating for $\mathbb{T}$. Suppose also that $\mathscr{M}\in\mathrm{Lat}(T)$ and $T|\mathscr{M}$ is a normal diagonal operator with the property that each λ_n is an eigenvalue of $T|\mathscr{M}$ of infinite multiplicity. Then $T\in(BCP)$.*

Proof. The hypothesis ensures that each λ_n belongs to $\sigma_{le}(T|\mathscr{M})$, and since $\sigma_{le}(T|\mathscr{M})\subset\sigma_{le}(T)$, we conclude that $\sigma_e(T)\cap\mathbb{D}$ is dominating for $\mathbb{T}$.

The following result is an easy consequence of Proposition 1.1 and Lemma 1.2.

Proposition 1.3. *Suppose $T\in\mathbb{A}_{\aleph_0}(\mathscr{H})$, $\{u_1,\ldots,u_n\}$ is any finite subset of $\mathscr{H}$, and $\varepsilon>0$. Then there exists $\mathscr{M}\in\mathrm{Lat}(T)$ such that*

(i) *both $T|\mathscr{M}$ and $T_{\mathscr{H}\ominus\mathscr{M}}$ are (BCP)-operators, and*

(ii) *$\|P_{\mathscr{M}}u_i\|<\varepsilon$ for $i=1,\ldots,n$.*

Proof. Let $\{\lambda_n\}_{n=1}^{\infty}\subset\mathbb{D}$ be dominating for $\mathbb{T}$, and let N be a normal diagonal operator of uniform infinite multiplicity whose eigenvalues constitute the sequence $\{\lambda_n\}_{n=1}^{\infty}$. By Proposition 4.2 of [3] there exist invariant subspaces $\mathcal{M}_0\supset\mathcal{K}$ for T such that $T_{\mathcal{M}_0\ominus\mathcal{K}}$ is unitarily equivalent to N. Thus N^* is the restriction to an invariant subspace of $(T|\mathcal{M}_0)^*$, and it follows from Lemma 1.2 that $(T|\mathcal{M}_0)^*$ (along with $T|\mathcal{M}_0$) belongs to $(BCP)(\mathcal{M}_0)$. Let y_1 be the orthogonal projection of u_1 onto $\mathcal{M}_0$. By Proposition 1.1 there exists $\mathcal{M}_1\in\mathrm{Lat}(T|\mathcal{M}_0)$ such that $(T|\mathcal{M}_0)|\mathcal{M}_1=T|\mathcal{M}_1$ is a (BCP)-operator and $\|P_{\mathcal{M}_1}y_1\|<\varepsilon$. Note that $\mathcal{M}_1\in\mathrm{Lat}(T)$ and that $\|P_{\mathcal{M}_1}u_1\|=\|P_{\mathcal{M}_1}y_1\|$. By an obvious finite induction argument we can find an invariant subspace $\mathcal{M}_n\subset\mathcal{M}_1$ for T such that $T|\mathcal{M}_n$ is a (BCP)-operator and such that $\|P_{\mathcal{M}_n}u_i\|<\varepsilon$, $i=1,\dots,n$. Since $T|\mathcal{M}_n\in\mathbb{A}_{\aleph_0}(\mathcal{M}_n)$, we may apply Proposition 4.2 of [3] to $T|\mathcal{M}_n$ and the operator $N\oplus N$ to conclude the existence of a decomposition

$$\mathcal{M}_n=\mathcal{N}_1\oplus\mathcal{N}_2\oplus\mathcal{N}_3\oplus\mathcal{N}_4,\qquad \text{where } \mathcal{N}_1 \text{ and } \mathcal{N}_1\oplus\mathcal{N}_2\oplus\mathcal{N}_3$$

belong to $\mathrm{Lat}(T|\mathcal{M}_n)$, and where $(T|\mathcal{M}_n)_{\mathcal{N}_2\oplus\mathcal{N}_3}$ is the operator $N\oplus N$ acting on $\mathcal{N}_2\oplus\mathcal{N}_3$ in the obvious way. We set $\mathcal{M}=\mathcal{N}_1\oplus\mathcal{N}_2$. Clearly $\mathcal{M}\in\mathrm{Lat}(T)$, and that $T|\mathcal{M}\in(BCP)(\mathcal{M})$ follows as before. Furthermore the restriction of $T_{\mathcal{H}\ominus\mathcal{M}}$ to the invariant subspace $\mathcal{N}_3$ is the operator N, so, once again by Lemma 1.2, $T_{\mathcal{H}\ominus\mathcal{M}}\in(BCP)(\mathcal{H}\ominus\mathcal{M})$. Finally, since $\mathcal{M}\subset\mathcal{M}_n$, it is obvious that $\|P_{\mathcal{M}}u_i\|<\varepsilon$, $i=1,\dots,n$, so the proof is complete.

The next corollary now follows from Proposition 1.3 by the same argument that Robel used to prove [7, Propositions 6.2 and 6.3] from [7, Proposition 6.1].

Corollary 1.4. *Suppose $T\in\mathbb{A}_{\aleph_0}(\mathcal{H})$. Then $\mathcal{H}$ admits a decomposition $\mathcal{H}=\bigoplus_{n=0}^{\infty}\mathcal{M}_n$ such that the operator matrix (T_{ij}) for T relative to this decomposition is in upper triangular form and satisfies $T_{nn}\in(BCP)(\mathcal{M}_n)$, $0\leqq n<\infty$. Furthermore $\mathcal{H}$ admits another decomposition $\mathcal{H}=\bigoplus_{n=-\infty}^{\infty}\mathcal{N}_n$ such that the operator matrix $(\tilde{T}_{ij})$ for T relative to this decomposition is in upper triangular form and satisfies $\tilde{T}_{nn}\in(BCP)(\mathcal{N}_n)$, $-\infty<n<\infty$.*

The following theorem shows that, for operators in $\mathbb{A}_{\aleph_0}$, finite systems of simultaneous equations can be solved with reasonable estimates on the distance from the initial data to the solution.

Theorem 1.5. *Suppose $T\in\mathbb{A}_{\aleph_0}(\mathcal{H})$, $\{[L_{ij}]\}_{1\leqq i,j\leqq n}$ is a finite set of elements of Q_T, $\{z_1,\dots,z_m\}$ is an arbitrary finite set of vectors from $\mathcal{H}$, and $\varepsilon>0$. Suppose also that $\{x_1^0,\dots,x_n^0\}$ and $\{y_1^0,\dots,y_n^0\}$ are sequences from $\mathcal{H}$ and $\delta>0$ is such that $\|[L_{ij}]-[x_i^0\otimes y_j^0]\|<\delta$ for $1\leqq i,j\leqq n$. Then there exist sequences $\{x_1,\dots,x_n\}$ and $\{y_1,\dots,y_n\}$ of vectors from $\mathcal{H}$ such that*

$$[L_{ij}]=[x_i\otimes y_j],\qquad 1\leqq i,\ j\leqq n, \tag{1}$$

$$\|x_i^0-x_i\|<n\delta^{1/2},\qquad \|y_i^0-y_i\|<n\delta^{1/2}, \tag{2}$$

and

$$\begin{aligned} &\|[(x_i^0-x_i)\otimes z_k]\|<\varepsilon, \quad &&\|[z_k\otimes(x_i^0-x_i)]\|<\varepsilon,\\ &\|[(y_i^0-y_i)\otimes z_k]\|<\varepsilon, \quad &&\|[z_k\otimes(y_i^0-y_i)]\|<\varepsilon,\\ & &&1\leqq i\leqq n,\ 1\leqq k\leqq m. \end{aligned} \tag{3}$$

Proof. Let $d_{ij}=\|[L_{ij}]-[x_i^0\otimes y_j^0]\|$, $1\leqq i,\ j\leqq n$, and let τ be a positive number such that

$$\tau<n(\delta^{1/2}-\max_{i,j}(d_{ij})^{1/2}). \tag{4}$$

Let $M>0$ be an upper bound for $\|x_i^0\|$, $\|y_j^0\|$, and $\|z_k\|$ for $1\leqq i,\ j\leqq n$ and $1\leqq k\leqq m$. We choose a positive number η such that

$$\eta<\min\{\tau/2,\ \varepsilon/3M,\ \varepsilon/3n\delta^{1/2}\} \tag{5}$$

and such that

$$0\leqq t,t' \quad\text{and}\quad |t'-t|<3M\eta \quad\text{imply}\ |\sqrt{t'}-\sqrt{t}|<\tau/2n. \tag{6}$$

(The reason for this choice of η will appear later. We choose it now to make it clear that η does not depend upon the choice of the upcoming vectors x_i and y_j.) It follows from Proposition 1.3 that there exists $\mathcal{M}\in\mathrm{Lat}(T)$ such that $T|\mathcal{M}$ and $S=T_{\mathcal{H}\ominus\mathcal{M}}$ are both (BCP)-operators and such that the norm of the (orthogonal) projection onto $\mathcal{M}$ of each of the $2n+m$ vectors $\{x_1^0,\dots,x_n^0\}$, $\{y_1^0,\dots,y_n^0\}$, and $\{z_1,\dots,z_m\}$ is less than η. We write $x_i'=P_{\mathcal{H}\ominus\mathcal{M}}x_i^0$, $1\leqq i\leqq n$, and define similarly y_j', $1\leqq j\leqq n$, and z_k', $1\leqq k\leqq m$. (The idea of the proof of this theorem should now be clear. We will transfer the equation solving problem to the semi-invariant subspace $\mathcal{H}\ominus\mathcal{M}$, using the fact that $S=T_{\mathcal{H}\ominus\mathcal{M}}$ is a (BCP)-operator to solve equations there with "good" bounds, and the smallness of the η we have chosen will then give us the estimates we desire.)

For $1\leqq i,\ j\leqq n$, let $[M_{ij}]\in Q_S$ be defined by $[M_{ij}]=\phi_S^{-1}\phi_T([L_{ij}])$, and note that the $[M_{ij}]$ are uniquely determined by the relations

$$\langle S^p,[M_{ij}]\rangle=\langle\lambda^p,\phi_S([M_{ij}])\rangle=\langle T^p,[L_{ij}]\rangle,\quad p=0,1,2,\dots. \tag{7}$$

In particular, since the $[L_{ij}]$ are arbitrary elements of Q, for $u,\ v\in\mathcal{H}\ominus\mathcal{M}$, we have

$$[u\otimes v]_{Q_S}=\phi_S^{-1}\phi_T([u\otimes v]_{Q_T}) \tag{8}$$

by virtue of (7), since

$$\langle S^p,[u\otimes v]_{Q_S}\rangle=(S^pu,v)=(T^pu,v)=\langle T^p,[u\otimes v]_{Q_T}\rangle,\quad 0\leqq p<\infty.$$

Let $\alpha=M\eta+\max_{i,j}d_{ij}'$, where $d_{ij}'=\|[M_{ij}]-[x_i'\otimes y_j']\|_{Q_S}$. It now follows from Corollary 6.13 and Remark 6.14 of [3] (applied with $\theta=0$ to the operator S) that there exist sequences $\{x_1,\dots,x_n\}$ and $\{y_1,\dots,y_n\}$ of vectors in $\mathcal{H}\ominus\mathcal{M}$ such that

$$[M_{ij}]=[x_i\otimes y_j]_{Q_S},\qquad 1\leqq i,\ j\leqq n, \tag{9}$$

$$\|x_i'-x_i\|<n\alpha^{1/2},\qquad \|y_i'-y_i\|<n\alpha^{1/2},\quad 1\leqq i\leqq n, \tag{10}$$

and

$$\begin{gathered}\|[(x_i'-x_i)\otimes z_k']\|_{Q_S}<\varepsilon/3,\qquad \|[z_k'\otimes(x_i'-x_i)]\|_{Q_S}<\varepsilon/3,\\ \|[(y_i'-y_i)\otimes z_k']\|_{Q_S}<\varepsilon/3,\qquad \|[z_k'\otimes(y_i'-y_i)]\|_{Q_S}<\varepsilon/3,\quad 1\leqq i\leqq n,\ 1\leqq k\leqq m.\end{gathered} \tag{11}$$

By applying $\phi_T\phi_S^{-1}$ to (9) and using (8), we see that (1) is satisfied. We will now prove (2) for the x_i's, recalling that ϕ_S and ϕ_T are isometries. We have from (5) and (10) that

$$\|x_i^0-x_i\|\leqq\|x_i^0-x_i'\|+\|x_i'-x_i\|<(\tau/2)+n\alpha^{1/2},\qquad 1\leqq i\leqq n. \tag{12}$$

Furthermore, from the inequalities

$$\begin{aligned}d_{ij}'&=\|[M_{ij}]-[x_i'\otimes y_j']\|_{Q_S}=\|[L_{ij}]-[x_i'\otimes y_j']\|_{Q_T}\\ &\leqq\|[L_{ij}]-[x_i^0\otimes y_j^0]\|_{Q_T}+\|[x_i^0\otimes y_j^0]-[x_i'\otimes y_j']\|_{Q_T},\end{aligned}$$

we obtain

$$\begin{aligned}d_{ij}'&\leqq d_{ij}+\|[x_i^0\otimes y_j^0]-[x_i'\otimes y_j']\|_{Q_T}\\ &\leqq d_{ij}+\|[x_i^0\otimes(y_j^0-y_j')]\|_{Q_T}+\|[(x_i^0-x_i')\otimes y_j']\|_{Q_T}<d_{ij}+2M\eta.\end{aligned}$$

Therefore

$$\alpha=M\eta+\max_{i,j}d_{ij}'<(\max_{i,j}d_{ij})+3M\eta,$$

and from (6) we obtain

$$\alpha^{1/2}<\max_{i,j}(d_{ij}^{1/2})+\tau/2n. \tag{13}$$

Hence from (12), (13), and (4) we conclude that

$$\|x_i^0-x_i\|<\tau/2+n(\max_{i,j}d_{i,j}^{1/2})+\tau/2<n\delta^{1/2},\qquad 1\leqq i\leqq n, \tag{14}$$

as desired. Of course this argument works equally well to prove that $\|y_i^0-y_i\|<n\delta^{1/2}$, $1\leqq i\leqq n$. To conclude the proof of the theorem we content ourselves with proving the first inequality in (3). For $1\leqq i\leqq n$ and $1\leqq k\leqq m$ we have

$$\begin{aligned}\|[(x_i-x_i^0)\otimes z_k]\|_{Q_T}\leqq{}&\|[(x_i-x_i')\otimes z_k']\|_{Q_T}+\|[(x_i-x_i')\otimes(z_k-z_k')]\|_{Q_T}\\ &+\|[(x_i'-x_i^0)\otimes z_k]\|_{Q_T},\end{aligned}$$

and using (11), (14), (5) and the fact that ϕ_S and ϕ_T are isometries, we obtain

$$\|[(x_i-x_i')\otimes z_k']\|_{Q_T}=\|[(x_i-x_i')\otimes z_k']\|_{Q_S}<\varepsilon/3,$$

$$\|[(x_i-x_i')\otimes(z_k-z_k')]\|_{Q_T}\leqq\|x_i-x_i'\|\cdot\|z_k-z_k'\|\leqq n\alpha^{1/2}\eta<n\delta^{1/2}\eta<\varepsilon/3,$$

and

$$\|[(x_i^0 - x_i') \otimes z_k]\|_{Q_T} \leqq \|x_i^0 - x_i'\| \cdot \|z_k\| \leqq \eta \cdot M < \varepsilon/3.$$

Thus $\|[(x_i^0 - x_i) \otimes z_k]\|_{Q_T} < \varepsilon$ as desired, and the proof is complete.

The special case of Theorem 1.5 when $n=1$ shows that [2, Proposition 1] is valid for all operators in $\mathbb{A}_{\aleph_0}$, and since the proof of [2, Corollary 1] only depends on [2, Proposition 1] we have the following.

Corollary 1.6. *Suppose* $T \in \mathbb{A}_{\aleph_0}(\mathscr{H})$, *and denote by* $\mathscr{W}_T$ *the smallest subalgebra of* $\mathscr{L}(\mathscr{H})$ *that contains* T *and* $1_{\mathscr{H}}$ *and is closed in the weak operator topology. Then* $\mathscr{W}_T = \mathscr{A}_T$ *and the weak operator and ultraweak operator topologies coincide on* $\mathscr{A}_T$.

It follows from this corollary and a result from [1] that every weighted unilateral shift operator W that is a contraction such that $\sigma(W) \supset \mathbb{T}$ satisfies $\mathscr{W}_W = \mathscr{A}_W$. This partly answers Question 5 of [8].

Theorem 1.5 also shows that [2, Proposition 2] is valid for all operators in $\mathbb{A}_{\aleph_0}$, and since the proof of the reflexivity of (BCP)-operators used only [2, Proposition 2], we also have the following corollary, which generalizes Theorems 3, 4, and 5 of [2].

Theorem 1.7. *Every operator in* $\mathbb{A}_{\aleph_0}(\mathscr{H})$ *is reflexive. In particular, all of the operators in the classes* $(BCP)_\theta$, $0 \leqq \theta < 1$, *defined in* [4] *are reflexive.*

As mentioned earlier, the utility of Theorem 1.7 will be greatly enhanced by the appearance of [1], because of the large number of operators that turn out to belong to $\mathbb{A}_{\aleph_0}$. For the moment we deduce the following corollary of Corollary 1.6 and Theorem 1.7.

Corollary 1.8. *Suppose* $T \in C_{00}$ *and also* $T \in \bigcap_{n=1}^{\infty} \mathbb{A}_n(\mathscr{H})$. *Then* T *is reflexive, the algebras* $\mathscr{W}_T$ *and* $\mathscr{A}_T$ *coincide, and the weak operator and ultraweak topologies agree on* $\mathscr{A}_T$.

Proof. Exner showed in [6] that $\left(\bigcap_{n=1}^{\infty} \mathbb{A}_n\right) \cap C_{00} \subset \mathbb{A}_{\aleph_0}$.

This corollary raises the interesting question whether operators in a fixed class $\mathbb{A}_n$ $(n < \aleph_0)$ and not in C_{00} have these same properties.

We also note that the upper bounds on $\|x_i^0 - x_i\|$ and $\|y_i^0 - y_i\|$ given by (2) in Theorem 1.5 for all operators in $\mathbb{A}_{\aleph_0}$ are better than those given in [4, Corollary 6.11] for $(BCP)_\theta$-operators, so Theorem 1.5 generalizes [4, Corollary 6.11].

We close this note with a further consequence of Theorem 1.5. If $n \in \mathbb{N}$, we denote by $\tilde{\mathscr{H}}_n$ the direct sum of n copies of the Hilbert space $\mathscr{H}$.

Corollary 1.9. *Suppose* $T \in \mathbb{A}_{\aleph_0}$, $n \in \mathbb{N}$, *and* $\{[L_{ij}]\}_{i,j=1}^n$ *is a doubly indexed sequence of elements in* Q_T. *Then the set of vectors* $(x_1, \ldots, x_n)$ *in* $\tilde{\mathscr{H}}_n$ *for which there exists a vector* $(y_1, \ldots, y_n)$ *in* $\tilde{\mathscr{H}}_n$ *satisfying* (1) *is dense in* $\tilde{\mathscr{H}}_n$.

Proof. Let $\tilde{x}_0 = (x_1^0, \ldots, x_n^0)$ be an arbitrary vector in $\tilde{\mathscr{H}}_n$, let τ be a positive number, and use as initial data in Theorem 1.5 the vectors $(\tau x_1^0, \ldots, \tau x_n^0)$ and

$(0,\dots,0)$ in $\tilde{\mathscr{H}}_n$. Then, according to that theorem, there exists a solution $\tilde{x}_\tau = (x_1^\tau, \dots, x_n^\tau)$, $\tilde{y}_\tau = (y_1^\tau, \dots, y_n^\tau)$ of (1) such that

$$\|x_i^\tau - \tau x_i^0\| < n\delta^{1/2}, \qquad \|y_i^\tau - 0\| < n\delta^{1/2}, \quad 1 \leqq i \leqq n, \tag{15}$$

where δ is any fixed positive number that exceeds $\max_{i,j} \|[L_{ij}]\|$. Thus, since for every $\tau > 0$, the pair $(1/\tau)\tilde{x}_\tau, \tau\tilde{y}_\tau$ is also a solution of (1), and since $\|(1/\tau)\tilde{x}_\tau - \tilde{x}_0\| \to 0$ by (15), the result follows. In fact, to obtain $\|(1/\tau)\tilde{x}_\tau - \tilde{x}_0\| < \varepsilon$, it suffices to take $\tau = n^2\delta^{1/2}/\varepsilon^{1/2}$, in which case the vector $\tau\tilde{y}_\tau$ satisfies $\|\tau\tilde{y}_\tau\| < n^4\delta/\varepsilon$.

References

1. Apostol, C., Bercovici, H., Foias, C., Pearcy, C.: Dilation theory, invariant subspaces, and dual algebras, Preprint.
2. Bercovici, H., Foias, C., Langsam, J., Pearcy, C.: (BCP)-operators are reflexive. Michigan Math. J. **29**, 371–379 (1982)
3. Bercovici, H., Foias, C., Pearcy, C.: Dilation theory and systems of simultaneous equations in the predual of an operator algebra. I., Michigan Math. J. **30**, 335–354 (1983)
4. Bercovici, H., Foias, C., Pearcy, C.: Factoring trace-class operator-valued functions with applications to the class $\mathbb{A}_{\aleph_0}$, to appear in J. Operator Theory
5. Brown, S., Chevreau, B., Pearcy, C.: Contractions with rich spectrum have invariant subspaces. J. Operator Theory **1**, 123–136 (1979)
6. Exner, G.: Systems of equations in the predual of an operator algebra, the classes $\mathbb{A}_n$, and related operators. Thesis, University of Michigan 1983
7. Robel, G.: On the structure of (BCP)-operators and related algebras. I. To appear in J. Operator Theory
8. Shields, A.: Weighted shift operators and analytic function theory, Topics in Operator Theory, pp. 49–128. Providence: Amer. Math. Soc. 1974
9. Sz.-Nagy, B., Foias, C.: Harmonic analysis of operators on Hilbert space. Budapest: Akademiai Kiado 1970

Received December 21, 1983

Math. Z. 187, 105–123 (1984)

Mathematische Zeitschrift

Über den Derivationenmodul und das Jacobi-Ideal von Kurvensingularitäten

Ernst Kunz und Rolf Waldi

Fakultät für Mathematik, Universität Regensburg, Universitätsstraße 31, D-8400 Regensburg, Bundesrepublik Deutschland

Einleitung

k sei ein algebraisch abgeschlossener Körper der Charakteristik 0 und $R = k[\![X_1, \ldots, X_n]\!]/I$ eine reduzierte algebroide Kurve über k.

Ist R Gorensteinsch, so gibt es im Derivationenmodul $\mathrm{Der}_k(R)$ von R über k einen ausgezeichneten freien Untermodul D von Rang 1, den „Modul der trivialen Derivationen" (für die Definition siehe § 3). Wir interessieren uns für den Quotienten $\mathrm{Der}_k(R)/D$ und schätzen seine Größe ab, indem wir seine Länge ℓ und die Länge μ eines kürzesten Erzeugendensystems betrachten. Wir werden zeigen:

Satz 1. *Es sei δ_R der Singularitätengrad und r_R die Anzahl der Zweige von R. Dann gilt*

$$\ell(\mathrm{Der}_k(R)/D) \leqq 2\delta_R - r_R + 1 \qquad (\textit{Milnor-Zahl}).$$

Gleichheit gilt genau dann, wenn R quasihomogen ist.

Dabei heißt R quasihomogen, wenn es eine Darstellung $R = k[\![X_1, \ldots, X_n]\!]/I$ gibt, wobei I von quasihomogenen Polynomen eines festen Typs erzeugt wird.

Sei c_R: $\Omega^1_{R/k} \to \omega^1_{R/k}$ der kanonische Homomorphismus des universell-endlichen Differentialmoduls in den Modul der regulären Differentiale von R/k. Wenn R Gorensteinsch ist, ist $\omega^1_{R/k} \cong R$ und daher Kokern $c_R \cong R/J$ mit einem Ideal $J \subset R$.

Satz 2. *Es gilt $\mu(\mathrm{Der}_k(R)/D) = r(R/J)$, wobei r den Cohen-Macaulay-Typ eines Rings bezeichnet. Ferner sind folgende Aussagen äquivalent:*

a) $\mathrm{Der}_k(R)/D$ *ist ein zyklischer R-Modul.*

b) R/J *ist Gorensteinring.*

c) R *ist quasihomogen.*

Ist Bedingung c) *erfüllt, so wird* $\mathrm{Der}_k(R)/D$ *vom Bild der Euler-Derivation von R erzeugt.*

Ist $R=k[\![X_1,\dots,X_n]\!]/(F_1,\dots,F_{n-1})=k[\![x_1,\dots,x_n]\!]$ ein vollständiger Durchschnitt, so ist $J=\mathfrak{d}_1(R/k)$ das Jacobiideal (die 1. Kählersche Differente) von R/k, d.h. $J=(\Delta_1,\dots,\Delta_n)$, wobei die Δ_i die $(n-1)$-Minoren der Jacobimatrix $\left(\frac{\partial F_i}{\partial x_j}\right)_{\substack{i=1,\dots,n-1\\ j=1,n}}$ sind.

Aus Satz 2 ergibt sich daher

Satz 3. *Ist R vollständiger Durchschnitt und $\mathfrak{d}_1(R/k)$ das Jacobiideal von R/k, so sind folgende Aussagen äquivalent:*

a) *R ist quasihomogen.*

b) *$R/\mathfrak{d}_1(R/k)$ ist ein Gorensteinring.*

Der Beweis der Sätze benutzt die lokale Dualitätstheorie von Grothendieck, für die wir generell auf [3] verweisen, und die Quasihomogenitätskriterien aus [2] und [5]. Ferner wird mit der Wertehalbgruppe von Kurven (mit mehreren Zweigen) gearbeitet, über die in § 1 Grundtatsachen zusammengestellt sind. § 2 enthält einige Aussagen über Gorensteinalgebren, die für die Beweise wichtig sind. § 3 beschäftigt sich mit dem Modul der trivialen Derivationen und dem Beweis von Satz 1. Die Sätze 2 und 3 werden schließlich in § 4 bewiesen.

Die Autoren danken Herrn Jürgen Herzog für die fruchtbaren Gespräche, die zur Entstehung dieser Arbeit beigetragen haben, und Herrn Michael Kühl für mehrere Verbesserungsvorschläge.

§ 1. Die Wertehalbgruppe einer algebroiden Kurve

Sei $(R,\mathfrak{m})$ eine algebroide Kurve im Sinne der Einleitung, wobei k zunächst beliebige Charakteristik haben darf. Unter den Zweigen von R verstehen wir die Restklassenringe $R_i=R/\mathfrak{p}_i$, wobei $\mathfrak{p}_i$ $(i=1,\dots,r)$ die minimalen Primideale von R durchläuft. Die ganze Abschließung $\bar{R}$ von R ist in kanonischer Weise das direkte Produkt der ganzen Abschließungen $\bar{R}_i$ der R_i und jedes $\bar{R}_i$ ist ein Potenzreihenring $\bar{R}_i=k[\![t_i]\!]$ über k.

Es sei $K:=Q(R)$, $K_i:=Q(R_i)$ $(i=1,\dots,r)$ und $v_i\colon K\to\mathbb{Z}\cup\{\infty\}$ die zu $\bar{R}_i$ gehörige normierte diskrete Bewertung. Wir schreiben in Zukunft $\mathbb{Z}_\infty$ für $\mathbb{Z}\cup\{\infty\}$ und $\mathbb{N}_\infty$ für $\mathbb{N}\cup\{\infty\}$. Das Bild von R bei der Abbildung $v:=v_1\times\dots\times v_r$ von K in $\mathbb{Z}_\infty^r$ liegt in $\mathbb{N}_\infty^r$, es heißt die *Wertehalbgruppe von R* und wird künftig mit H bezeichnet. Entsprechend sei H_i die Wertehalbgruppe von R_i und H_{ij} die von $R_{ij}:=R/\mathfrak{p}_i\cap\mathfrak{p}_j$ $(i\neq j)$. Die H_i und H_{ij} entstehen aus H durch Projektion. Ein Element $a\in R$ ist genau dann ein Nichtnullteiler von R, wenn $v(a)\in\mathbb{N}^r$. Ein Zweig R_i von R heißt regulär, wenn $R_i=\bar{R}_i$ ist. Dies ist genau dann der Fall, wenn $H_i=\mathbb{N}_\infty$ ist.

Die Wertehalbgruppen irreduzibler algebroider Kurven wurden in [4] ausführlich behandelt. Wir wollen uns hier auf den Fall von Kurven mit mehreren Zweigen konzentrieren und nur einige Eigenschaften aufführen, die uns im Beweis von Satz 2 nützlich sein werden.

Ein gebrochenes Ideal von H ist eine Teilmenge $I\subset\mathbb{Z}_\infty^r$ mit $I+H\subset I$. I heißt ganzes Ideal, wenn $I\subset H$. Für jedes gebrochene R-Ideal $\mathfrak{a}$ in K ist $v(\mathfrak{a})$

ein gebrochenes Ideal von H und $M := v(\mathfrak{m})$ ist das eindeutig bestimmte maximale Ideal von H. $v(\mathfrak{a})$ heißt das Werteideal von $\mathfrak{a}$.

Der Führer $\mathfrak{F}_{\bar{R}/R}$ von $\bar{R}$ nach R wird als $\bar{R}$-Ideal von einem Element der Form $(t_1^{c_1}, \ldots, t_r^{c_r})$ erzeugt $(c_i \in \mathbb{N})$. Das Werteideal von $\mathfrak{F}_{\bar{R}/R}$ ist daher $(c_1, \ldots, c_r) + \mathbb{N}_\infty^r$. Das r-Tupel $c := (c_1, \ldots, c_r)$ heißt der Führer von H.

Wir versehen $\mathbb{Z}_\infty^r$ mit der folgenden Anordnung: Es soll $(w_1, \ldots, w_r) \leqq (z_1, \ldots, z_r)$ genau dann gelten, wenn $w_i \leqq z_i$ für $i = 1, \ldots, r$. Der Führer c von H ist dann das (eindeutig bestimmte) kleinste Element von H mit $c + \mathbb{N}^r \subset H$.

Es läßt sich sofort die folgende einfache Regel angeben.

1.1. *Regel.* $I \subset \mathbb{Z}_\infty^r$ sei das Werteideal eines gebrochenen R-Ideals $\mathfrak{a}$ und $h = (h_1, \ldots, h_r)$, $\ell = (\ell_1, \ldots, \ell_r)$ seien zwei Elemente aus I. Dann gehört auch

$$\operatorname{Min}(h, \ell) := (\operatorname{Min}(h_1, \ell_1), \ldots, \operatorname{Min}(h_r, \ell_r))$$

zu I.

Zum Beweis wähle man $x, y \in \mathfrak{a}$ mit $v(x) = h$, $v(y) = \ell$. Für eine geeignete Linearkombination $\alpha x + \beta y$ mit $\alpha, \beta \in k$ gilt dann $v(\alpha x + \beta y) = \operatorname{Min}(h, \ell)$.

Das Werteideal I eines Ideals $\mathfrak{a} \subset R$ besitzt nach 1.1 immer ein eindeutig bestimmtes kleinstes Element. Es wird als Wert von jedem Element $x \in \bar{R}$ mit $\mathfrak{a}\bar{R} = x\bar{R}$ angenommen. Das kleinste Element von M wird in Zukunft mit $p = (p_1, \ldots, p_r)$ bezeichnet. Dabei ist $p_i \neq \infty$ für $i = 1, \ldots, r$, da $\mathfrak{m}$ einen Nichtnullteiler enthält. Ein Element $q \in H$ heißt *minimal außerhalb* $p\mathbb{Z}$, wenn q ein minimales Element der Menge $H \setminus p\mathbb{Z}$ ist. Im allgemeinen gibt es mehr als ein minimales Element außerhalb $p\mathbb{Z}$. Wenn R nur zwei Zweige besitzt, übersieht man die Verhältnisse noch recht gut. Im folgenden besitze R genau zwei Zweige.

Ein Element $(h_1, h_2) \in H \cap \mathbb{N}^2$ heiße *maximal über* h_1 (*maximal über* h_2), wenn es kein Element der Form $(h_1, \ell) \in H$ (bzw. $(\ell, h_2) \in H$) gibt mit $\ell > h_2$ (bzw. $\ell > h_1$). Ist $(h_1, h_2) \in H \cap \mathbb{N}^2$ nicht maximal über h_1, so gibt es nach 1.1 immer ein $(h_1, \ell) \in H \cap \mathbb{N}^2$ mit $\ell > h_2$.

1.2. *Bemerkung.* Ist $(h_1, h_2) \in H \cap \mathbb{N}^2$ maximal über h_1, so ist (h_1, h_2) auch maximal über h_2.

Angenommen, es gäbe ein Element $(\ell, h_2) \in H$ mit $\ell > h_1$. Wähle $x \in R$ mit $v(x) = (h_1, h_2)$ und $y \in R$ mit $v(y) = (\ell, h_2)$. Für geeignete $\alpha, \beta \in k$ ist dann $v(\alpha x + \beta y) = (h_1, h)$ mit $h > h_2$.

1.3. **Lemma.** *Jedes minimale Element außerhalb* $p\mathbb{Z}$ *ist von der Form* $\lambda p + \varepsilon$ *mit einem* $\lambda \in \mathbb{N} \setminus \{0\}$ *und einem* $\varepsilon \in \mathbb{N}^2$ *mit* $0 < \varepsilon < p$. *Es gibt höchstens zwei minimale Elemente außerhalb* $p\mathbb{Z}$. *Ferner gilt:*

a) *Besitzt* H *zwei verschiedene minimale Elemente außerhalb* $p\mathbb{Z}$, *so sind diese von der Form*

$$\lambda p + (\varepsilon_1, 0) \quad \textit{und} \quad \lambda p + (0, \varepsilon_2)$$

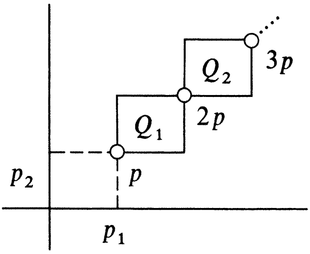
mit einem $\lambda \in \mathbb{N} \setminus \{0\}$ *und* $0 < \varepsilon_i \leqq p_i$ $(i = 1, 2)$.

b) *Ist* $H_1 = \mathbb{N}$, *so gibt es ein minimales Element außerhalb* $p\mathbb{Z}$ *von der Form*

$$\lambda p + (1, \varepsilon_2) \quad \textit{mit einem } \lambda \in \mathbb{N} \setminus \{0\} \textit{ und } 0 \leqq \varepsilon_2 < p_2$$

und es ist $\lambda p + (m, \varepsilon_2) \in H$ *für alle* $m \geqq 1$.

Beweis. Nach der ersten Aussage des Lemmas sind die minimalen Elemente außerhalb $p\mathbb{Z}$ Gitterpunkte in $\bigcup_{\ell=1}^{\infty} Q_\ell$, wobei $Q_\ell := \{w \in \mathbb{N}^2 \mid \ell p < w < (\ell+1)p\}$.



Man sieht dies wie folgt ein: Für jeden Punkt von H, der außerhalb $\bigcup Q_\ell$ liegt, erhält man durch Minimumbildung mit einem geeigneten Punkt aus $p\mathbb{Z}$ einen Punkt aus $H \cap \bigcup Q_\ell$ (nach 1.1).

Für $x \in Q_i$, $y \in Q_j$ mit $i < j$ ist auch $x < y$. Daher liegen alle minimalen Elemente außerhalb $p\mathbb{Z}$ in einem Q_ℓ mit festem ℓ. Sind $w, z \in Q_\ell$ zwei verschiedene solche Elemente, so muß $\operatorname{Min}(w, z) = \ell p$ sein. w und z sind daher von der in a) angegebenen Form und es folgt auch, daß es höchstens zwei verschiedene minimale Elemente außerhalb $p\mathbb{Z}$ geben kann.

Ist $H_1 = \mathbb{N}$, folglich $p_1 = 1$, so ist ein Gitterpunkt $q \in \bigcup Q_\ell$ entweder von der in b) gewünschten Form

$$q = \lambda p + (1, \varepsilon_2) \quad \text{mit } \lambda \in \mathbb{N} \setminus \{0\} \text{ und } 0 \leqq \varepsilon_2 < p_2 \tag{1}$$

oder es gilt:

$$q = \lambda p + (0, \varepsilon_2) \quad \text{mit } \lambda \in \mathbb{N} \setminus \{0\} \text{ und } 0 < \varepsilon_2 \leqq p_2. \tag{2}$$

Sei q wie in (2) und außerdem ein außerhalb $p\mathbb{Z}$ minimales Element von H. Dann ist λp nicht maximal über $\lambda = \lambda p_1$, also gibt es nach 1.2 ein $m \geqq 1$ mit $\lambda p + (m, 0) \in H$. Durch Minimumbildung mit $(\lambda+1)p$ ergibt sich, daß $\lambda p + (1, 0) \in H$ ist. Da $Q_\ell \cap H = \emptyset$ für $\ell < \lambda$, ist klar, daß $\lambda p + (1, 0)$ minimal außerhalb $p\mathbb{Z}$ ist.

In jedem Fall gibt es also in H ein minimales Element q außerhalb $p\mathbb{Z}$ von der Form (1). Ein solches Element ist nicht maximal über $\lambda + 1$. Nach 1.2 erhält man ein Element $\lambda p + (m, \varepsilon_2) \in H$ mit $m > 1$ und nach 1.1 ist dann auch $\lambda p + (\mu, \varepsilon_2) \in H$ für $\mu = 1, \ldots, m$. Ist $m = \infty$, so ist man fertig. Für $m \neq \infty$ ist $\lambda p + (m, \varepsilon_2)$ nicht maximal über $\lambda + m$, daher erhält man (durch Fortsetzung des obigen Verfahrens) schließlich, daß $\lambda p + (\mu, \varepsilon_2) \in H$ für alle $\mu \geqq 1$.

Es sei jetzt wieder R eine algebroide Kurve mit beliebig vielen Zweigen, $p = (p_1, \ldots, p_r)$ das kleinste Element von M und $q = (q_1, \ldots, q_r)$ ein minimales Element außerhalb $p\mathbb{Z}$. Wir wählen $x, y \in R$ mit $v(x) = p$, $v(y) = q$. Man hat in R die folgende „Division durch x mit Rest".

1.4. **Lemma.** *Zu jedem $a \in R$ gibt es ein Polynom $f \in k[T]$ mit* $\operatorname{Grad} f \cdot p < q$ *und ein $\tilde{a} \in R$ mit $v(\tilde{a}) \geqq q$, so daß*

$$a = f(x) + \tilde{a}.$$

Beweis. Wir setzen

$$\sigma(a) := \sum_{v_j(a) < q_j} q_j - v_j(a)$$

und schließen induktiv nach σ. Für $\sigma(a)=0$ können wir $f=0$ und $\tilde{a}=a$ setzen. Sei jetzt $\sigma(a)>0$, also $v_j(a)<q_j$ für mindestens ein j. Mit einem geeigneten $\alpha \in k$ gilt dann

$$v_j(a+\alpha y) = \operatorname{Min}(v_j(a), q_j) \qquad \text{für alle } j=1, \ldots, r.$$

Da $v(a+\alpha y)<q$ und q minimal außerhalb $p\mathbb{Z}$ ist, gilt $v(a+\alpha y)=np<q$ mit einem $n \in \mathbb{N}$, d.h. es ist

$$\begin{aligned} v_j(a) &= v_j(a+\alpha y) = n p_j, && \text{falls } v_j(a)<q_j \\ v_j(a+\alpha y) &= n p_j = q_j, && \text{falls } v_j(a) \geqq q_j. \end{aligned}$$

Insbesondere ist $\sigma(a+\alpha y)=\sigma(a)$. Mit einem geeigneten $\beta \in k$ ist $\sigma(\beta x^n + a + \alpha y) < \sigma(a)$. Durch Induktion ist man fertig.

Im folgenden sei nun k ein (algebraisch abgeschlossener) Körper der Charakteristik 0. Für $a=(a_1, \ldots, a_r) \in K = k((t_1)) \times \ldots \times k((t_r))$ sei

$$a' := \left(\frac{\partial a_1}{\partial t_1}, \ldots, \frac{\partial a_r}{\partial t_r}\right).$$

Wir schreiben manchmal auch $t := (t_1, \ldots, t_r)$ und $\frac{\partial a}{\partial t} := a'$ (Ableitung von a nach dem Parameter t). $R\mathfrak{m}'$ bezeichne das gebrochene R-Ideal, das von den Ableitungen der Elemente $a \in \mathfrak{m}$ erzeugt wird. Es hat auch folgende Bedeutung: Ist $\Omega^1_{R/k}$ der universell-endliche Differentialmodul von R/k, so ist das Bild der kanonischen Abbildung $\Omega^1_{R/k} \to K \otimes_R \Omega^1_{R/k} = K\,dt$ gerade der Modul $R\mathfrak{m}'\,dt$. Das Ideal $R\mathfrak{m}'$ interessiert uns hier, weil man nach [2], 2.1 das folgende Quasihomogenitätskriterium hat:

1.5. **Satz.** *Wenn R Gorensteinsch ist, so sind folgende Aussagen äquivalent:*

a) *R ist quasihomogen.*

b) *$M-1$ ist das Werteideal von $R\mathfrak{m}'$.*

Hierbei wurde $1 := (1, \ldots, 1)$ gesetzt. In [2], 2.1 wurde gezeigt, daß R genau dann quasihomogen ist, wenn $R\mathfrak{m}' = \mathfrak{m}'$ ist. Dies ist mit der Aussage in b) äquivalent: Wenn b) gilt, kann man jedes Element aus $R\mathfrak{m}'$ sukzessive durch Ableitungen approximieren bis man in den Führer gelangt.

Im allgemeinen Fall hat man folgende Aussagen:

1.6. **Lemma.** *Es ist stets $M-1 \subset v(R\mathfrak{m}')$. Ist q ein minimales Element von H außerhalb $p\mathbb{Z}$, so ist $q-1$ minimal in $v(R\mathfrak{m}') \setminus (p\mathbb{Z}-1)$. Mit anderen Worten: Ist $\omega \in R\mathfrak{m}'$ und $v(\omega)<q-1$, so ist $v(\omega)=n \cdot p-1$ mit einem $n \in \mathbb{N}$.*

Mit einem $x \in R$ mit $v(x)=p$ schreibt sich ω nach 1.4 in der Form

$$\omega = g(x) \cdot x' + \tilde{\omega} \qquad \text{mit } g \in k[T],\ v(\tilde{\omega}) \geqq q-1.$$

Wegen $v(\omega)<q-1$ folgt

$$v(\omega)=v(g(x))+v(x)-1\in p\mathbb{N}-1.$$

1.7. **Lemma.** *δ: $R\to R$ sei eine k-Derivation, $\bar{\delta}$ ihre Fortsetzung auf K und es sei*

$$v_i(\bar{\delta}t_i)\begin{cases}>1 & \textit{für } i=1,\dots,s\\ =1 & \textit{für } i=s+1,\dots,r.\end{cases}$$

Dann gilt in kanonischer Weise

$$R=R\Big/\bigcap_{i=1}^{s}\mathfrak{p}_i\times_k R\Big/\bigcap_{i=s+1}^{r}\mathfrak{p}_i\quad\text{(Faserprodukt über } k).$$

Beweis. Wir setzen $R^{(1)}:=R\big/\bigcap_{i=1}^{s}\mathfrak{p}_i$, $R^{(2)}:=R\big/\bigcap_{i=s+1}^{r}\mathfrak{p}_i$ und bezeichnen im folgenden für Elemente, Ideale etc. die jeweiligen Projektionen mit dem entsprechenden oberen Index. $H^{(i)}$ sei die Wertehalbgruppe von $R^{(i)}$ und $M^{(i)}$ ihr maximales Ideal ($i=1,2$). Es genügt zu zeigen, daß $M^{(1)}\times\{\infty\}^{r-s}\subset H$ ist, denn dann ist $\mathfrak{m}^{(1)}\times\{0\}\subset R$ und somit $R=R^{(1)}\times_k R^{(2)}$.

Sei $h^{(1)}$ ein beliebiges Element aus $M^{(1)}$. Es reicht aus, ein $y\in R$ mit $v^{(1)}(y)=h^{(1)}$, $v^{(2)}(y)\geqq c^{(2)}$ zu konstruieren, dann ist auch $(y^{(1)},0)\in R$. Wir gehen ähnlich wie in 1.4 vor und verringern schrittweise die Zahl $\varepsilon(y):=\sum_{\substack{j\in\{s+1,\dots,r\}\\ v_j(y)<c_j}} c_j-v_j(y)$. Wir beginnen mit einem beliebigen $y_0\in\mathfrak{m}$, für das $v^{(1)}(y_0)=h^{(1)}$ ist. Nach Voraussetzung ist dann

$$v^{(1)}(\delta y_0)-v^{(1)}(y_0)\geqq 1^{(1)}\quad\text{und}\quad v^{(2)}(\delta y_0)=v^{(2)}(y_0).$$

Ist $\varepsilon(y_0)>0$, so können wir ein $\alpha\in k$ so wählen, daß

$$v^{(1)}(y_0-\alpha\delta y_0)=h^{(1)}\quad\text{und}\quad\varepsilon(y_0-\alpha\delta y_0)<\varepsilon(y_0).$$

Durch Induktion sind wir fertig.

§ 2. Gorensteinalgebren

$(R,\mathfrak{m})$ sei ein lokaler Gorensteinring mit $\dim R=1$. Es sei $K:=Q(R)$ und $I\subset K$ ein gebrochenes R-Ideal, das eine Einheit von K enthält. Dann ist $I^{-1}:=\{x\in K\mid xI\subset R\}\cong\operatorname{Hom}_R(I,R)$. Da R Gorensteinring ist, gilt $(I^{-1})^{-1}=I$ und $\ell(I_1^{-1}/I_2^{-1})=\ell(I_2/I_1)$ für gebrochene R-Ideale $I_1\subset I_2$, wenn I_1 eine Einheit von K enthält. Speziell ist $\ell(\mathfrak{m}^{-1}/R)=1$ ([3], 1.46)).

2.1 **Lemma.** *Sei $I\subset R$ ein $\mathfrak{m}$-primäres Ideal. Dann ist I^{-1}/R ein kanonischer Modul von R/I.*

Durch Dualisieren der exakten Sequenz $0\to I\to R\to R/I\to 0$ bez. R erhält man eine exakte Sequenz $0\to R\to I^{-1}\to\operatorname{Ext}_R^1(R/I,R)\to 0$. Da $\operatorname{Ext}_R^1(R/I,R)$ ein kanonischer Modul von R/I ist, ergibt sich die Behauptung.

Ein Ideal $I\subset R$ heißt *Gorensteinideal*, wenn R/I ein Gorensteinring ist.

2.2. **Lemma.** *Für ein $\mathfrak{m}$-primäres Ideal I von R sind folgende Aussagen äquivalent:*

a) *I ist Gorensteinideal.*

b) *Es existiert eine Einheit $z \in K$, so daß $I = zR \cap R$.*

c) *Es gibt einen Nichtnullteiler $a \in R$ und ein $b \in R$, so daß*

$$I = (a) \underset{R}{:} (b) := \{r \in R / rb \in (a)\}.$$

d) *Für jeden Nichtnullteiler $a \in I$ gibt es ein $b \in R$ mit $I = (a) \underset{R}{:} (b)$.*

Beweis. a)$\leftrightarrow$b). Wenn I Gorensteinideal ist, so ist $I^{-1}/R \cong R/I$ nach 2.1. Es gibt daher ein $w \in I^{-1}$ mit $I^{-1} = (1, w)$. Indem man w notfalls um ein geeignetes Element von R abändert, kann man annehmen, daß w Einheit in K ist. Dann ergibt sich $I = (I^{-1})^{-1} = (R + Rw)^{-1} = R \cap Rw^{-1} = Rz \cap R$ mit $z := w^{-1}$.

Umgekehrt folgt aus $I = Rz \cap R$, daß $I^{-1} = \left(1, \frac{1}{z}\right)$ ist und daher $K_{R/I} = I^{-1}/R$ von einem Element erzeugt wird. Dann ist R/I ein Gorensteinring.

c)$\rightarrow$b). Mit einem geeigneten $c \in R$ ist $b' := b + ca$ ein Nichtnullteiler von R und $I = (a) \underset{R}{:} (b')$. Setze $z := \frac{a}{b'}$.

Da d)$\rightarrow$c) trivial ist, ist nur noch a)$\rightarrow$d) zu zeigen. $R/(a)$ ist ein Gorensteinring und $I/(a)$ ein Gorensteinideal von $R/(a)$. Nach [3], 1.44 und 5.20, ist $I/(a)$ der Annullator eines Hauptideals $(\bar{b})$ von $R/(a)$. Ist $b \in R$ ein Repräsentant von $\bar{b}$, so ergibt sich $I = (a) \underset{R}{:} (b)$.

2.3. *Bemerkung.* Ist $(R, \mathfrak{m})$ ein lokaler Gorensteinring beliebiger Dimension und I ein Cohen-Macaulay-Ideal der Höhe h von R, so zeigt man leicht, daß folgende Aussagen äquivalent sind:

a) I ist ein Gorensteinideal.

b) Es gibt eine reguläre Folge $\{a_1, \ldots, a_h\}$ in $\mathfrak{m}$ und ein $b \in R$, so daß $I = (a_1, \ldots, a_h) \underset{R}{:} (b)$.

c) Zu jeder regulären Folge $\{a_1, \ldots, a_h\}$ in I gibt es ein $b \in R$ mit $I = (a_1, \ldots, a_h) \underset{R}{:} (b)$.

Im folgenden sei R ein beliebiger Ring und S/R eine (kommutative) Algebra. S/R heißt *Gorensteinalgebra (lokal vollständiger Durchschnitt)*, wenn S als R-Modul flach ist und für jedes $\mathfrak{P} \in \operatorname{Spec}(S)$ mit $\mathfrak{p} := \mathfrak{P} \cap R$ gilt: $S_{\mathfrak{P}}/\mathfrak{p}S_{\mathfrak{P}}$ ist ein Gorensteinring (vollständiger Durchschnitt) im Sinne der lokalen Algebra. S/R heißt *vollständiger Durchschnitt*, wenn S flacher R-Modul ist und eine Darstellung $S = R[X_1, \ldots, X_n]/(t_1, \ldots, t_h)$ besitzt, wobei $(t_1, \ldots, t_h)$ eine quasireguläre Folge ist.

Eine Algebra S/R heißt *endlich (projektiv)*, wenn S als R-Modul endlich erzeugt (projektiv) ist.

2.4. **Lemma.** *S/R sei endlich und projektiv. Genau dann ist S/R eine Gorensteinalgebra, wenn $\operatorname{Hom}_R(S, R)$ ein projektiver S-Modul vom Rang 1 ist.*

Beweis. Für jedes $\mathfrak{p} \in \operatorname{Spec}(R)$ ist

$$\operatorname{Hom}_R(S, R)_{\mathfrak{p}}/\mathfrak{p} \operatorname{Hom}_R(S, R)_{\mathfrak{p}} \cong \operatorname{Hom}_{R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}}(S_{\mathfrak{p}}/\mathfrak{p} S_{\mathfrak{p}}, R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}),$$

da $S_{\mathfrak{p}}/R_{\mathfrak{p}}$ frei ist. Sind $\mathfrak{P}_1, \ldots, \mathfrak{P}_s$ die über $\mathfrak{p}$ liegenden Primideale von S, so ist ferner

$$\operatorname{Hom}_{R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}}(S_{\mathfrak{p}}/\mathfrak{p} S_{\mathfrak{p}}, R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}) = \prod_{i=1}^{s} \operatorname{Hom}_{R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}}(S_{\mathfrak{P}_i}/\mathfrak{p} S_{\mathfrak{P}_i}, R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}).$$

Dabei ist $\operatorname{Hom}_{R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}}}(S_{\mathfrak{P}_i}/\mathfrak{p} S_{\mathfrak{P}_i}, R_{\mathfrak{p}}/\mathfrak{p} R_{\mathfrak{p}})$ ein kanonischer Modul von $S_{\mathfrak{P}_i}/\mathfrak{p} S_{\mathfrak{P}_i}$. Genau dann ist S/R eine Gorensteinalgebra, wenn alle diese Moduln jeweils zu $S_{\mathfrak{P}_i}/\mathfrak{p} S_{\mathfrak{P}_i}$ isomorph sind. Wenn $\operatorname{Hom}_R(S, R)$ ein projektiver S-Modul vom Rang 1 ist, so ist dies jedenfalls richtig und S/R ist eine Gorensteinalgebra.

Ist umgekehrt S/R eine Gorensteinalgebra, so folgt nach Nakayama, daß $\operatorname{Hom}_R(S, R)_{\mathfrak{p}}$ für jedes $\mathfrak{p} \in \operatorname{Spec}(R)$ ein zyklischer $S_{\mathfrak{p}}$-Modul ist. Da $S_{\mathfrak{p}}/R_{\mathfrak{p}}$ frei ist, ist $\operatorname{Ann}_{S_{\mathfrak{p}}}(\operatorname{Hom}_{R_{\mathfrak{p}}}(S_{\mathfrak{p}}, R_{\mathfrak{p}})) = 0$ und es folgt $\operatorname{Hom}_R(S, R)_{\mathfrak{p}} \cong S_{\mathfrak{p}}$.

Ist S/R eine endliche Gorensteinalgebra und R ein noetherscher lokaler Ring, so ist $\operatorname{Hom}_R(S, R) = S\sigma$ mit einem Basiselement σ von $\operatorname{Hom}_R(S, R)$. Ein solches σ nennen wir eine *Spur* von S nach R. Man beachte, daß die kanonische Spur $\sigma_{S/R}: S \to R$ der Algebra S/R keine Spur in unserem jetzigen Sinn zu sein braucht. Zwei Spuren von S/R unterscheiden sich nur um eine Einheit von S. Im folgenden stellen wir einige Eigenschaften von Spuren zusammen.

Sei zunächst S/R eine beliebige endliche und projektive Algebra. Der Kern I der Abbildung

$$\mu: S \otimes_R S \to S \qquad (a \otimes b \mapsto ab)$$

wird bekanntlich von den Elementen $s \otimes 1 - 1 \otimes s$ ($s \in S$) erzeugt. Die beiden S-Modulstrukturen auf $S \otimes_R S$, die durch die Multiplikation im ersten bzw. zweiten Faktor gegeben werden, stimmen auf $\operatorname{Ann}_{S \otimes_R S}(I)$ wegen

$$(s \otimes 1 - 1 \otimes s) \cdot \operatorname{Ann}(I) = 0$$

überein. Da S ein projektiver R-Modul ist, hat man einen kanonischen Isomorphismus von R-Moduln

$$\Phi: S \otimes_R S \to \operatorname{Hom}_R(\operatorname{Hom}_R(S, R), S),$$

wobei für $x = \sum a_k \otimes b_k \in S \otimes_R S$ und $\ell \in \operatorname{Hom}_R(S, R)$ die Formel

$$\Phi(x)(\ell) = \sum \ell(a_k)\, b_k \tag{1}$$

gilt.

2.5. **Lemma.** *Φ induziert einen Isomorphismus von S-Moduln*

$$\Phi: \operatorname{Ann}(I) \xrightarrow{\sim} \operatorname{Hom}_S(\operatorname{Hom}_R(S, R), S).$$

Dies prüft man mit Hilfe der Formel (1) sofort nach.

Besitzt S/R eine Spur σ, so ist $\mathrm{Ann}(I)$ ein freier S-Modul vom Rang 1, erzeugt von einem Element Δ_σ mit $\Phi(\Delta_\sigma)(\sigma)=1$.

Ist ferner $\ell \in \mathrm{Hom}_R(S,R)$ gegeben und $\Phi(\Delta_\sigma)(\ell)=s\in S$, so ist $\ell=s\cdot\sigma$. Ist $\Delta_\sigma=\sum a_k\otimes b_k$ $(a_k, b_k\in S)$, so gilt nach (1)

$$s=\Phi(\Delta_\sigma)(\ell)=\sum \ell(a_k)b_k.$$

2.6. **Lemma.** *S/R besitze eine Spur σ und $\Delta_\sigma\in\mathrm{Ann}(I)$ sei das Element mit $\Phi(\Delta_\sigma)(\sigma)=1$. Ist $\{s_1,\dots,s_n\}$ eine Basis von S/R und schreibt man*

$$\Delta_\sigma=\sum_{k=1}^{n} s'_k\otimes s_k \qquad (s'_k\in S),$$

so gilt $\sigma(s'_i s_k)=\delta_{ik}$ $(i,k=1,\dots,n)$. Insbesondere ist auch $\{s'_1,\dots,s'_n\}$ eine Basis von S/R (die Dualbasis zu $\{s_1,\dots,s_n\}$ bez. σ).

Beweis. Aus $\Phi(\Delta_\sigma)(\sigma)=\sum_{i=1}^{n}\sigma(s'_i)s_i=1$ folgt $s_k=s_k\cdot\Phi(\Delta_\sigma)(\sigma)=\Phi(\Delta_\sigma)(s_k\sigma)=\sum_{i=1}^{n}\sigma(s'_i s_k)s_i$ und somit $\sigma(s'_i s_k)=\delta_{ik}$.

Unter den Voraussetzungen von 2.6 gilt

$$\sigma_{S/R}=\lambda\cdot\sigma \qquad \text{mit } \lambda:=\sum_{k=1}^{n}\sigma_{S/R}(s_k)s'_k. \tag{2}$$

Wir behalten diese Voraussetzung bei und setzen $K:=Q(R)$, $L:=Q(S)$. Dann ist $L=K\otimes_R S$ und $\sigma_{L/K}=\mathrm{id}_K\otimes\sigma_{S/R}$. Bekanntlich gilt $\mathrm{Hom}_K(L,K)=L\cdot\sigma_{L/K}$ genau dann, wenn L/K étale ist.

In diesem Fall ist der Dedekindsche Komplementärmodul $\mathfrak{C}(S/R):=\{y\in L/\sigma_{L/K}(yS)\subset R$ für alle $s\in S\}$ definiert. Mit $\sigma':=\mathrm{id}_K\otimes\sigma$ ist auch $\mathrm{Hom}_K(L,K)=L\sigma'$ und aus $\sigma_{L/K}=\lambda\sigma'$ ergibt sich, daß λ eine Einheit in L ist. Aus 2.6 erhält man

$$\sigma_{L/K}\left(\frac{s'_i}{\lambda}\cdot s_k\right)=\delta_{ik}$$

und daher

$$\mathfrak{C}(S/R)=R\cdot\frac{s'_1}{\lambda}\otimes\ldots\otimes R\cdot\frac{s'_n}{\lambda}=\frac{1}{\lambda}S.$$

Für die Dedekindsche Differente $\mathfrak{d}_D(S/R):=\mathfrak{C}(S/R)^{-1}$ ergibt sich daher

$$\mathfrak{d}_D(S/R)=\lambda\cdot S.$$

2.7. **Lemma.** *$(R,\mathfrak{m})$ sei ein noetherscher lokaler Ring und S/R eine endliche Gorensteinalgebra. Dabei sei $[S:R]$ eine Einheit von R und $Q(S)/Q(R)$ étale. Dann ist $\mathfrak{d}_D(S/R)\not\subset\mathfrak{m}S$.*

Beweis. Ist λ wie oben gegeben, so ist zu zeigen, daß $\lambda\notin\mathfrak{m}S$ ist. Da R lokal ist, kann man o.B.d.A. annehmen, daß in der Basis $\{s_1,\dots,s_n\}$ von S/R gilt: $s_1=1$.

Dann ist $\sigma_{S/R}(s_1)=[S:R]$ Einheit in R und weil $\{s'_1, \ldots, s'_n\}$ ebenfalls eine Basis von S/R ist, zeigt die Formel (2) für λ, daß $\lambda \notin \mathfrak{m}S$.

Die Kählersche Differente $\mathfrak{d}_0(S/R)$ zeigt hier ein etwas anderes Verhalten als die Dedekindsche Differente: Ist S/R endlich und neben R auch S ein lokaler Ring, so gilt

$$\mathfrak{d}_0(S/R) \not\subset \mathfrak{m}S \tag{3}$$

höchstens dann, wenn $S/\mathfrak{m}S$ ein vollständiger Durchschnitt ist. Dies zeigen die Ausführungen von Scheja-Storch [7] im Anschluß an (4.7), wo gezeigt wird, daß $\mathfrak{d}_0(S/\mathfrak{m}S/R/\mathfrak{m})=0$, wenn $S/\mathfrak{m}S$ kein vollständiger Durchschnitt ist (dies ist auch richtig, wenn $R/\mathfrak{m}$ eine Charakteristik $\neq 0$ besitzt). Da $\mathfrak{d}_0(S/\mathfrak{m}S/R/\mathfrak{m})$ das Bild von $\mathfrak{d}_0(S/R)$ in $S/\mathfrak{m}S$ ist, ergibt sich die Aussage.

Für einen reduzierten noetherschen lokalen Ring $(R, \mathfrak{m})$ der Dimension 1 ist bekanntlich $\ell(\mathfrak{m}^{-1}/R)=r(R)$ der Cohen-Macaulay-Typ von R (vgl. etwa [3], Beweis von 1.46). Wir betrachten hier auch die Invariante

$$\tilde{r}(R) := \ell(\mathfrak{m} \underset{Q(R)}{:} \mathfrak{m}/\mathfrak{m}).$$

Wenn R regulär ist, ist $\mathfrak{m} \underset{Q(R)}{:} \mathfrak{m}=R$ und daher $\tilde{r}(R)=1$. Ist R singulär, so ist $\mathfrak{m} \underset{Q(R)}{:} \mathfrak{m}=\mathfrak{m}^{-1}$: Für jedes $x\in\mathfrak{m}^{-1}$ gilt nämlich $x\mathfrak{m}\subset\mathfrak{m}$. Wäre xy für ein $y\in R$ eine Einheit in R, so wäre $\mathfrak{m}\subset yR$, was im singulären Fall nicht möglich ist. Daher ist $\tilde{r}(R)=\ell(\mathfrak{m}^{-1}/R)+1=r(R)+1$ im singulären Fall.

Es seien nun $(R_1, \mathfrak{m}_1)$ und $(R_2, \mathfrak{m}_2)$ zwei reduzierte algebroide Kurven über einem Körper k und $R:=R_1\times_k R_2$ ihr Faserprodukt über k. Es ist dann $Q(R)=Q(R_1)\times Q(R_2)$ und für das maximale Ideal $\mathfrak{m}$ von R gilt $\mathfrak{m}=\mathfrak{m}_1\times\mathfrak{m}_2$ und $\mathfrak{m} \underset{Q(R)}{:} \mathfrak{m}=(\mathfrak{m}_1 \underset{Q(R_1)}{:} \mathfrak{m}_1)\times(\mathfrak{m}_2 \underset{Q(R_2)}{:} \mathfrak{m}_2)$. Somit ist

$$\tilde{r}(R_1\times_k R_2)=\tilde{r}(R_1)+\tilde{r}(R_2).$$

Für den Typ $r(R_1\times_k R_2)$ bedeutet dies:

2.8. **Lemma.** *Es gilt*

$$r(R_1\times_k R_2)=\begin{cases} r(R_1)+r(R_2)+1, & \textit{falls } R_1 \textit{ und } R_2 \textit{ singulär sind,}\\ r(R_2)+1, & \textit{falls } R_1 \textit{ regulär und } R_2 \textit{ singulär ist,}\\ 1, & \textit{falls } R_1 \textit{ und } R_2 \textit{ regulär sind.}\end{cases}$$

2.9. **Korollar.** *Genau dann ist $R_1\times_k R_2$ Gorensteinsch, wenn R_1 und R_2 regulär sind.*

§ 3. Derivationenmoduln

Es sei jetzt k ein algebraisch abgeschlossener Körper und $R=k[x_1, \ldots, x_n]$ eine 1-dimensionale reduzierte affine k-Algebra oder $R=k[\![x_1, \ldots, x_n]\!]$ eine 1-dimensionale reduzierte komplette lokale k-Algebra mit dem Restklassenkörper k

(eine algebroide Kurve über k). Ferner sei $K := Q(R)$ und $c_R\colon \Omega^1_{R/k} \to \omega^1_{R/k}$ die kanonische Abbildung des (universell-endlichen) Differentialmoduls $\Omega^1_{R/k}$ in den Modul $\omega^1_{R/k}$ der regulären Differentiale von R/k. Ist $x \in R$ ein Element, so daß $R/k[x]$ (bzw. $R/k[\![x]\!]$) endlich ist und K/K_0 mit $K_0 := Q(k[x])$ (bzw. $K_0 := Q(k[\![x]\!])$) étale ist, so ist $\omega^1_{R/k} = \mathfrak{C}(R/k[x]) \cdot dx \subset K \otimes_R \Omega^1_{R/k}$ (unabhängig von x) und c_R wird durch die kanonische Abbildung $\Omega^1_{R/k} \to K \otimes_R \Omega^1_{R/k}$ induziert (entsprechend im kompletten Fall).

In der zugehörigen exakten Sequenz

$$0 \to \tau(R) \to \Omega^1_{R/k} \xrightarrow{c_R} \omega^1_{R/k} \to \tau'(R) \to 0 \tag{1}$$

sind $\tau(R) := \operatorname{Kern} c_R$ und $\tau'(R) := \operatorname{Kokern} c_R$ zwei R-Moduln endlicher Länge, da c_R für die $\mathfrak{p} \in \operatorname{Reg}(R)$ einen Isomorphismus induziert, und $\tau(R)$ ist die Torsion von $\Omega^1_{R/k}$.

Dualisiert man (1) bez. $\omega^1_{R/k}$, so erhält man eine exakte Folge

$$\begin{aligned} 0 &\to \operatorname{Hom}_R(\omega^1_{R/k}, \omega^1_{R/k}) \to \operatorname{Hom}_R(\Omega^1_{R/k}, \omega^1_{R/k}) \to \operatorname{Ext}^1_R(\tau'(R), \omega^1_{R/k}) \\ &\to \mathit{Ext}^1_R(\omega^1_{R/k}, \omega^1_{R/k}). \end{aligned} \tag{2}$$

Für jedes $\mathfrak{p} \in \operatorname{Spec}(R)$ ist $\operatorname{Ext}^1_R(\omega^1_{R/k}, \omega^1_{R/k})_\mathfrak{p} = \operatorname{Ext}^1_{R_\mathfrak{p}}(\omega^1_{R_\mathfrak{p}/k}, \omega^1_{R_\mathfrak{p}/k}) = 0$ nach [3], 6.1, denn $\omega^1_{R_\mathfrak{p}/k}$ ist ein kanonischer Modul von $R_\mathfrak{p}$. Somit ist $\operatorname{Ext}^1_R(\omega^1_{R/k}, \omega^1_{R/k}) = 0$. Nach [3], 6.1d) ist ferner $\operatorname{Hom}_{R_\mathfrak{p}}(\omega^1_{R_\mathfrak{p}/k}, \omega^1_{R_\mathfrak{p}/k}) = R_\mathfrak{p} \cdot \mathrm{id}$ und somit $\operatorname{Hom}_R(\omega^1_{R/k}, \omega^1_{R/k}) = R \cdot \mathrm{id}_{\omega^1_{R/k}}$. Aus (2) erhält man daher eine kanonische exakte Folge

$$\begin{aligned} 0 \to R &\to \operatorname{Hom}_R(\Omega^1_{R/k}, \omega^1_{R/k}) \to \operatorname{Ext}^1_R(\tau'(R), \omega^1_{R/k}) \to 0 \\ 1 &\longmapsto c_R \end{aligned} \tag{3}$$

Wir setzen jetzt voraus, daß $\omega^1_{R/k} = R \cdot \omega_0$ ein freier R-Modul vom Rang 1 ist. Im kompletten Fall ist dies äquivalent damit, daß R ein Gorensteinring ist. Mittels des Basiselements ω_0 erhält man einen Isomorphismus

$$\varepsilon_{\omega_0}\colon \omega^1_{R/k} \xrightarrow{\sim} R \qquad (r\omega_0 \mapsto r)$$

und daher Isomorphismen

$$\operatorname{Hom}_R(\Omega^1_{R/k}, \omega^1_{R/k}) \cong \operatorname{Hom}_R(\Omega^1_{R/k}, R) \cong \operatorname{Der}_k(R), \quad c_R \mapsto \varepsilon_{\omega_0} \circ c_R \mapsto \varepsilon_{\omega_0} \circ c_R \circ d$$

wobei $d\colon R \to \Omega^1_{R/k}$ die universelle (universell-endliche) Derivation von R/k ist.

Für zwei Basiselemente ω_0 und $\tilde{\omega}_0$ von $\omega^1_{R/k}$ unterscheiden sich ε_{ω_0} und $\varepsilon_{\tilde{\omega}_0}$ nur um eine Einheit von R. Daher ist das Bild D von R bei der zusammengesetzten Abbildung

$$R \to \operatorname{Hom}_R(\Omega^1_{R/k}, \omega^1_{R/k}) \xrightarrow{\sim} \operatorname{Der}_k(R), \quad 1 \mapsto c_R \mapsto \varepsilon_{\omega_0} \circ c_R \circ d$$

unabhängig von ω_0 und ein freier R-Modul von Rang 1. Wir nennen D den *Modul der trivialen Derivationen von R/k.*

Ferner ist unter der obigen Voraussetzung der Modul $\tau'(R) \cong R/J$ zyklisch mit einem Ideal $J \cong \operatorname{Bild} c_R$. Im kompletten Fall ist wegen (3) $\operatorname{Der}_k(R)/D$ kanonischer Modul von R/J und es gilt nach 2.1

$$\operatorname{Der}_k(R)/D \cong J^{-1}/R. \tag{4}$$

Wir beschreiben den Modul D in zwei Situationen noch etwas genauer:

a) D ist insbesondere definiert, wenn R/k eine Darstellung als vollständiger Durchschnitt besitzt:

$$R = k[X_1, \ldots, X_n]/(F_1, \ldots, F_{n-1}) = k[x_1, \ldots, x_n]. \tag{5}$$

Wir können dabei annehmen, daß K étale über $Q(k[x_1])$ ist. Setzt man

$$\Delta_i := \frac{\partial(F_1, \ldots, F_{n-1})}{\partial(x_1, \ldots, \hat{x}_i, \ldots, x_n)} \qquad (i = 1, \ldots, n),$$

so ist Δ_1 Einheit in K und es gilt in $K \otimes_R \Omega^1_{R/k}$

$$\omega^1_{R/k} = R\,\frac{dx_1}{\Delta_1}. \tag{6}$$

Wir können dann $\omega_0 := \dfrac{dx_1}{\Delta_1}$ in der obigen Konstruktion von D verwenden. D läßt sich dann auch wie folgt beschreiben:

Der Präsentation

$$\Omega^1_{R/k} = \bigoplus_{k=1}^{n} R\,dX_k \Big/ \left\langle \left\{ \frac{\partial F_i}{\partial x_1}\,dX_1 + \ldots + \frac{\partial F_i}{\partial x_n}\,dX_n \right\}_{i=1,\ldots,n-1} \right\rangle$$

des Differentialmoduls entspricht die exakte Folge

$$0 \to \operatorname{Der}_k(R) \to \operatorname{Hom}_R\left(\bigoplus_{k=1}^{n} R\,dX_k, R\right) \to \operatorname{Hom}_R\left(\left\langle\left\{\sum \frac{\partial F_i}{\partial x_k}\,dX_k\right\}\right\rangle, R\right).$$

Schreibt man $\operatorname{Hom}_R\left(\bigoplus_{k=1}^{n} R\,dX_k, R\right) = \bigoplus_{k=1}^{n} R\,\dfrac{\partial}{\partial X_k}$ mittels der Dualbasis $\left\{\dfrac{\partial}{\partial X_1}, \ldots, \dfrac{\partial}{\partial X_n}\right\}$ zu $\{dX_1, \ldots, dX_n\}$, so identifiziert sich $\operatorname{Der}_k(R)$ mit dem aus den Elementen

$$r_1\frac{\partial}{\partial X_1} + \ldots + r_n\frac{\partial}{\partial X_n} \quad \text{mit } (r_1, \ldots, r_n) \in R, \quad \sum_{k=1}^{n} r_k \frac{\partial F_i}{\partial x_k} = 0 \quad (i = 1, \ldots, n-1)$$

bestehenden Untermodul von $\bigoplus_{k=1}^{n} R\,\dfrac{\partial}{\partial X_k}$.

Ein spezielles Element dieser Art ist nach der Cramerschen Regel

$$\delta := \sum_{k=1}^{n} (-1)^{n-k} \Delta_k \frac{\partial}{\partial X_k}. \tag{7}$$

3.1. **Lemma.** *Unter den obigen Voraussetzungen ist δ ein Basiselement von D, d.h. jede triviale Derivation von R/k läßt sich eindeutig in der Form*

$$r\delta = \sum_{k=1}^{n} (-1)^{n-k} r\Delta_k \frac{\partial}{\partial X_k} \qquad (r \in R)$$

schreiben.

Aus den in $\Omega^1_{K/k}$ gültigen Gleichungen

$$\sum_{k=1}^{n} \frac{\partial F_i}{\partial x_k} dx_k = 0 \qquad (i = 1, \ldots, n-1)$$

folgt mittels der Cramerschen Regel

$$dx_k = (-1)^{n-k} \frac{\Delta_k}{\Delta_1} dx_1$$

und somit

$$\text{Bild}\, c_R = \left(R \frac{\Delta_1}{\Delta_1} + \ldots + R \frac{\Delta_n}{\Delta_1}\right) dx_1 = \mathfrak{d}_1(R/k) \frac{dx_1}{\Delta_1}$$

mit dem Jacobiideal $\mathfrak{d}_1(R/k)$ von R/k. Also ist

$$J = \mathfrak{d}_1(R/k), \quad \mathrm{Der}_k(R) \cong \mathfrak{d}_1(R/k)^{-1} \quad \text{und} \quad \mathrm{Der}_k(R)/D \cong \mathfrak{d}_1(R/k)^{-1}/R. \tag{8}$$

Mit $\omega_0 = \dfrac{dx_1}{\Delta_1}$ ist $(\varepsilon_{\omega_0} \circ c_R \circ d)(x_k) = (-1)^{n-k} \Delta_k$ $(k = 1, \ldots, n)$, d.h. $\varepsilon_{\omega_0} \circ c_R \circ d = \delta$ und somit ist δ ein Basiselement von D.

Alle diese Aussagen sind natürlich entsprechend auch im kompletten Fall richtig.

b) Es sei jetzt R eine reduzierte 1-dimensionale komplette k-Algebra, die zunächst noch nicht Gorensteinsch zu sein braucht. $\bar{R}$ sei die ganze Abschließung von R in K und $\mathfrak{F}_{\bar{R}/R}$ der Führer von $\bar{R}$ nach R. Wie in § 1 sei $\bar{R} = k[\![t_1]\!] \times \ldots \times k[\![t_r]\!]$. Mit $t := (t_1, \ldots, t_r)$, $dt := (dt_1, \ldots, dt_r)$ ist dann

$$\omega^1_{\bar{R}/k} = \Omega^1_{\bar{R}/k} = \bar{R}\, dt.$$

Ferner ist $\mathfrak{F}_{\bar{R}/R} = f\bar{R}$ mit einem Nichtnullteiler $f \in R$, da $\bar{R}$ Hauptidealring ist.

3.2. **Lemma.** $\omega^1_{\bar{R}/k} = \mathfrak{F}_{\bar{R}/R} \cdot \omega^1_{R/k}$.

Beweis. $x \in R$ sei so gewählt, daß $K := Q(R)$ über $Q(k[\![x]\!])$ étale ist. Es genügt dann mit $A := k[\![x]\!]$ die Formel

$$\mathfrak{C}(\bar{R}/A) = \mathfrak{F}_{\bar{R}/R} \cdot \mathfrak{C}(R/A)$$

zu beweisen. Aus der Definition des Komplementärmoduls folgt sofort die Regel

$$\mathfrak{C}(\bar{R}/A) = \mathfrak{C}(R/A) : \bar{R}.$$

Aus den allgemeinen Formeln

$$\mathfrak{a} : (\mathfrak{b} \cdot \mathfrak{c}) = (\mathfrak{a} : \mathfrak{b}) : \mathfrak{c} \quad \text{und} \quad \mathfrak{C}(R/A) : (\mathfrak{C}(R/A) : \mathfrak{a}) = \mathfrak{a}$$

(für gebrochene R-Ideale $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, welche Einheiten von K enthalten) ergibt sich

$$R:\mathfrak{F}_{\bar{R}/R}=R:\bar{R}\cdot fR=(R:\bar{R}):fR=f\bar{R}:fR=\bar{R}$$

und

$$\begin{aligned}\mathfrak{C}(\bar{R}/A)&=\mathfrak{C}(R/A):(R:\mathfrak{F}_{\bar{R}/R})\\&=\mathfrak{C}(R/A):(\mathfrak{C}(R/A):(\mathfrak{C}(R/A)\cdot\mathfrak{F}_{\bar{R}/R}))=\mathfrak{C}(R/A)\cdot\mathfrak{F}_{\bar{R}/R}.\end{aligned}$$

Ist R Gorensteinsch, so folgt aus 3.2

$$\omega^1_{R/k}=\frac{R}{\varepsilon f}dt \tag{9}$$

mit einer Einheit $\varepsilon\in\bar{R}$.

Wir fassen jetzt $\mathrm{Der}_k(R)$ als Untermodul von $\mathrm{Der}_k(K)=K\frac{\partial}{\partial t}$ auf. Dabei ist $\frac{\partial g}{\partial t}:=(g_1',\ldots,g_r')$ für $g=(g_1,\ldots,g_r)\in K$. Mit $\omega_0=\frac{1}{\varepsilon f}dt$ gilt für jedes $g\in R$

$$\varepsilon_{\omega_0}(c_R(dg))=\varepsilon_{\omega_0}\left(\frac{\partial g}{\partial t}dt\right)=\varepsilon_{\omega_0}\left(\frac{\partial g}{\partial t}\varepsilon f\omega_0\right)=\frac{\partial g}{\partial t}\varepsilon f=\left(\varepsilon f\frac{\partial}{\partial t}\right)(g).$$

Somit ist

$$D=R\varepsilon f\frac{\partial}{\partial t}. \tag{10}$$

Da mit f auch εf den Führer $\mathfrak{F}_{\bar{R}/R}$ als $\bar{R}$-Ideal erzeugt, erhalten wir:

3.3. **Lemma.** *Es gibt ein $f\in R$ mit $\mathfrak{F}_{\bar{R}/R}=f\bar{R}$, so daß*

$$D=Rf\frac{\partial}{\partial t}.$$

Ferner ist $J=Rf\mathfrak{m}'$, wobei $\mathfrak{m}'$ die Menge der Ableitungen nach t der Elemente aus dem maximalen Ideal $\mathfrak{m}$ von R ist.

Im Gegensatz zu $R\mathfrak{m}'$ hängt J nicht von der Wahl des Parameters t ab.

Beweis von Satz 1. Satz 1 ist i.w. ein Korollar zu [2], Satz 2.1. Nach (4) ist $\mathrm{Der}_k(R)/D\cong J^{-1}/R$, wenn $\tau'(R)\cong R/J$. Da R ein Gorensteinring ist, erhält man $\ell(\mathrm{Der}_k(R)/D)=\ell(\tau'(R))$. Nach [1], Proposition 1.2.1 ist $\ell(\omega^1_{R/k}/dR)=2\delta_R-r_R+1$ und somit $\ell(\tau'(R))\leqq 2\delta_R-r_R+1$. Nach [2], Satz 2.1 gilt $c_R(\Omega^1_{R/k})=dR$ genau dann, wenn R quasihomogen ist, somit gilt $\ell(\mathrm{Der}_k(R)/D)=2\delta_R-r_R+1$ genau dann, wenn R quasihomogen ist.

Für den affinen Fall erhält man durch Übergang zu den Komplettierungen der lokalen Ringe der Singularitäten:

3.4. **Korollar.** *k sei ein algebraisch abgeschlossener Körper der Charakteristik 0 und R eine reduzierte* 1*-dimensionale affine k-Algebra. R besitze eine Darstellung* (5) *als vollständiger Durchschnitt. Es sei s_R die Anzahl der Singularitäten von R, r_R die Summe der Anzahlen der Zweige in den Singularitäten von R und δ_R*

$:= \dim_k(\bar{R}/R)$ *der Singularitätengrad von R. Dann gilt*

$$\dim_k(\mathrm{Der}_k(R)/D) \leqq 2\delta_R - r_R + s_R.$$

Gleichheit gilt genau dann, wenn R an allen singulären Stellen $\mathfrak{p}$ quasihomogen ist (d.h. die Komplettierung $\hat{R}_\mathfrak{p}$ ist quasihomogen).

Nach (8) ist $\dim_k(\mathrm{Der}_k(R)/D) = \dim_k(\mathfrak{d}_1(R/k)^{-1}/R) = \dim_k(R/\mathfrak{d}_1(R/k))$, daher läßt sich das Korollar auch als Aussage über die Länge des Rings $R/\mathfrak{d}_1(R/k)$ verstehen.

§ 4. Beweis der Sätze 2 und 3

Wenn R regulär ist, so ist nichts zu zeigen. Im folgenden sei daher $(R, \mathfrak{m})$ singulär.

Beweis von Satz 2. Nach 2.2 ist Aussage b) in Satz 2 äquivalent mit

b′) Es gibt eine Einheit $z \in Q(R)$ mit $J = zR \cap R$.

Wie in § 3 gezeigt, ist $\mathrm{Der}_k(R)/D$ ein kanonischer Modul von R/J. Es bleiben daher nur noch die Behauptungen c) $\Rightarrow$ a) und b′) $\Rightarrow$ c) zu beweisen.

c) $\Rightarrow$ a). Der Epimorphismus π: $\Omega^1_{R/k} \to \mathfrak{m}$ mit $\pi(dr) = \mathrm{Grad}(r) \cdot r$ für alle homogenen $r \in R$ induziert einen Isomorphismus

$$\mathfrak{m}^{-1} \cong \mathrm{Hom}_R(\mathfrak{m}, R) \cong \mathrm{Hom}_R(\Omega^1_{R/k}/\tau(R), R) \cong \mathrm{Der}_k(R).$$

Sei x ein homogener Nichtnullteiler von R und $A = k[\![x]\!] \subset R$ die zugehörige noethersche Normalisierung. Ist $\mathfrak{d}_D(R/A) = (\lambda)$, $\lambda \in R$, so ist $\frac{\lambda}{x} \notin R$ nach Lemma 2.7. In dem Diagramm

$$\begin{array}{ccccc} \Omega^1_{R/k}/\tau(R) & \xrightarrow{c_R} & \omega^1_{R/k} = R\frac{1}{\lambda}dx & & dx \\ \wr\| \big\downarrow \pi & & \wr\| \big\downarrow & & \big\downarrow \\ \mathfrak{m} & \xrightarrow[\alpha]{} & R & & \lambda \end{array}$$

sei α so gewählt, daß das Diagramm kommutiert. α ist dann die Multiplikation mit einem Element aus $Q(R)$, nämlich mit $\frac{\lambda}{\mathrm{Grad}(x) \cdot x}$, wie man sofort sieht, wenn man dx betrachtet. Insbesondere ist $\delta := \alpha \circ \pi \circ d$ (d: $R \to \Omega^1_{R/k}$ die universell-endliche Derivation) eine triviale Derivation. Ferner ist gezeigt, daß $\frac{\lambda}{x}\mathfrak{m} \subset R$ ist, d.h. $\frac{\lambda}{x} \in \mathfrak{m}^{-1}$.

Da $\frac{\lambda}{x} \notin R$ und $\ell(\mathfrak{m}^{-1}/R) = 1$, ergibt sich $\mathfrak{m}^{-1} = \left(1, \frac{\lambda}{x}\right)$. Das Bild von 1 beim Isomorphismus $\mathfrak{m}^{-1} \cong \mathrm{Der}_k(R)$ ist die Eulerderivation e und das Bild von

$\frac{\lambda}{\mathrm{Grad}(x)\cdot x}$ ist δ. Damit ist $\mathrm{Der}_k(R)=Re+R\delta$ gezeigt und $\mathrm{Der}_k(R)/D$ wird vom Bild der Eulerderivation erzeugt.

b′) ⇒ c). Im folgenden sei $\bar{R}$ die ganze Abschließung von R in $Q(R)$. Wir verwenden die Begriffe und Bezeichnungen aus § 1. Da $J=fR\mathfrak{m}'$ mit einem $f\in R$, $f\bar{R}=\mathfrak{F}_{\bar{R}/R}$ ist (3.3), überträgt sich die Aussage von 1.6 wie folgt auf das Werteideal von J:

(1) $c-1+M\subset v(J)$, $c-1+p$ ist das kleinste Element von $v(J)$.

(2) Ist $q\in H$ minimal außerhalb $p\mathbb{Z}$, so ist auch $c-1+q$ ein minimales Element von $v(J)\setminus(c-1+p\mathbb{Z})$.

Die entsprechenden Aussagen gelten auch für $v_i(J)$ in H_i und für $v_{ij}(J)$ in H_{ij}, wenn v_{ij} die Einschränkung von v auf $k((t_i))\times k((t_j))$ bezeichnet: Es ist nämlich $v_i(J)=c_i+v_i(R_i\mathfrak{m}_i')$ mit dem maximalen Ideal $\mathfrak{m}_i$ von R_i und entsprechend $v_{ij}(J)=(c_i,c_j)+v_{ij}(R_{ij}\mathfrak{m}_{ij}')$ mit dem maximalen Ideal $\mathfrak{m}_{ij}$ von R_{ij}. Ist R_i ein regulärer Zweig, so ist natürlich $v_i(J)=c_i+\mathbb{N}_\infty$.

Nach [6], Thm. 1 ist $J^{-1}\cong\mathrm{Der}_k(R)$ kein Hauptideal; aus $J=zR\cap R$ folgt daher $z\notin J$ und wegen $J\subset\mathfrak{F}_{\bar{R}/R}$ ergibt sich

$$J=z\mathfrak{m}\cap R=z\mathfrak{m}\cap\mathfrak{F}_{\bar{R}/R}.$$

Da $c-1+p\in v(J)$, muß jedenfalls

$$v(z)\leqq c-1 \tag{3}$$

sein. Wenn $z\mathfrak{m}\subset R$ ist, so ist $J=z\mathfrak{m}\cong\mathfrak{m}$ und wir sind nach [5], Satz 2.1 fertig.

Im allgemeinen braucht aber $z\mathfrak{m}$ nicht in R enthalten zu sein, wie das folgende Beispiel zeigt:

$$R\cong k[\![X,Y]\!]/(X\cdot Y),\qquad J=\mathfrak{m}=zR\cap R\quad\text{mit}\quad z=\left(1,\frac{1}{t_2^i}\right)\quad (i\geqq 1).$$

Die späteren Ausführungen werden zeigen, daß unter der Voraussetzung $z\mathfrak{m}\not\subset R$ stets $R\cong k[\![X,Y]\!]/(X\cdot Y)$ gilt, womit ebenfalls nachgewiesen sein wird, daß R quasihomogen ist.

Ist $z\mathfrak{m}\not\subset R$, so ist die kleinste ganze Zahl λ mit

$$z\mathfrak{m}^{\lambda+1}\subset R$$

positiv. Wir wählen ein $x\in\mathfrak{m}$ mit $v(x)=p$. Dann ist $x^{\lambda+1}z\in R\cap zR=J$ und aus (1) ergibt sich $v(x^{\lambda+1}z)\geqq c-1+p$. Ferner ist $z\mathfrak{m}^\lambda\not\subset\mathfrak{F}_{\bar{R}/R}$ und $v(x^\lambda z)$ ist kleinstes Element von $v(z\mathfrak{m}^\lambda)$, folglich $v(x^\lambda z)\ngeqq c$. Sei o.B.d.A. $v_1(x^\lambda z)\leqq c_1-1$. Dann hat man

$$v(x^\lambda z)\geqq c-1,\qquad v_1(x^\lambda z)=c_1-1. \tag{4}$$

Wir werden jetzt nacheinander die folgenden Aussagen herleiten:

α) *Es ist* $v_j(x^\lambda z)\geqq c_j$ *für* $j=2,\dots,r$.

β) R_1 *ist regulär,* $v_1(z)\leqq c_1-2$.

γ) *Es ist* $v_j(z)=c_j-1$ *für* $j=2,\dots,r$.

Da R nicht regulär ist, muß dann $r \geqq 2$ sein. Da $R \cdot f \cdot \mathfrak{m}' = J = Rz \cap R$, ist $\delta := \frac{f}{z} \frac{\partial}{\partial t}$ eine k-Derivation von R in sich. Aus β) und γ) folgt, daß auf δ das Lemma 1.7 mit $s=1$ angewandt werden kann. Man erhält $R = R_1 \times_k R \Big/ \bigcap_{i=2}^{r} \mathfrak{p}_i$ und nach 2.9 ist dann $R \cong k[\![X, Y]\!]/(X \cdot Y)$, womit der Satz bewiesen ist.

Zu α). Angenommen, es sei $r \geqq 2$ und $v_{12}(x^\lambda z) = (c_1 - 1, c_2 - 1)$. Wir wählen ein $y \in \mathfrak{m}$ mit minimalem Wert $v_{12}(y) =: (q_1, q_2)$ außerhalb $(p_1, p_2)\mathbb{Z}$. O.B.d.A. sei $q_1 > p_1$, d.h. $v_1(yx^{\lambda-1}z) \geqq c_1$. Wir zeigen, daß die Elemente $x^\lambda z$ und $yx^{\lambda-1}z$ aus $\mathfrak{m}^{-1}$ k-linear unabhängig modulo R sind, im Widerspruch zu $\ell(\mathfrak{m}^{-1}/R) = 1$.

Nach Wahl von y hat jedes $w \in kyx^{\lambda-1}z + kx^\lambda z \setminus \{0\}$ Werte

$$v_1(w) = c_1 - 1 \quad \text{oder} \quad v_{12}(w) = (c_1, c_2) - 1 - (p_1, p_2) + (q_1, q_2).$$

Wir wenden (1) und (2) auf H_{12} an und erhalten in jedem Fall $w \notin J = zR \cap R$. Also ist, wie behauptet wurde, $w \notin R$.

Zu β). Ist R_1 singulär, so gibt es ein $y \in \mathfrak{m}$ mit $v_1(y) =: q_1 \in H_1$, das minimal außerhalb $p_1\mathbb{Z}$ ist. Nach α) ist $v(yx^{\lambda-1}z) \geqq c$, folglich $yx^{\lambda-1}z \in J$. Andererseits ist aber nach (2), angewandt auf H_1,

$$v_1(yx^{\lambda-1}z) = c_1 - 1 + q_1 - p_1 \notin v_1(J),$$

ein Widerspruch. Die zweite Aussage in β) folgt aus (4).

Zu γ). Wir zeigen zunächst, daß für jedes $j \neq 1$ gilt:

$$(m, p_j) \in H_{1j} \qquad \text{für alle } m \geqq 1. \tag{5}$$

Nach 1.3b) genügt es zu zeigen, daß $(2, p_j) \in H_{1j}$. Es sind $x^{\lambda+1}z$ und fx' Elemente von J mit Werten

$$v_{1j}(x^{\lambda+1}z) = (c_1, c_j) - 1 + (1, \mu), \qquad \mu \geqq p_j + 1$$

$$v_{1j}(fx') = (c_1, c_j) - 1 + (1, p_j).$$

Wendet man 1.2 auf $v_{1j}(J) \subset H_{1j}$ an, so ergibt

$$(c_1, c_j) - 1 + (m, p_j) \in v_{1j}(J) \text{ mit einem } m \geqq 2.$$

Ferner ist $(c_1, c_j) - 1 + (2, 2p_j)$ der Wert von $fxx' \in J$, mit 1.1 folgt

$$(c_1, c_j) - 1 + (2, p_j) \in v_{1j}(J) \setminus ((c_1, c_2) - 1 + (1, p_j)\mathbb{Z}).$$

Da H_{1j} nach 1.3b) ein minimales Element $(q_1, q_j) \geqq (2, p_j)$ außerhalb $(1, p_j)\mathbb{Z}$ besitzt, muß nach (2) somit $(q_1, q_j) = (2, p_j)$ sein und (5) folgt aus 1.3b).

Nach diesen Vorbereitungen verfahren wir ähnlich wie im Beweis von α). Angenommen, γ) sei falsch. Da $v(z) \leqq c - 1$ ist, ist dann die kleinste Zahl μ mit

$$v_j(x^\mu z) \geqq c_j - 1 \qquad \text{für alle } j = 2, \ldots, r \tag{6}$$

positiv und es gibt ein $j_0 \neq 1$ mit

$$v_{j_0}(x^{\mu-1}z) < c_{j_0} - 1. \tag{7}$$

Wegen (4) ist $\lambda > \mu - 1$ und wir können gemäß (5) Elemente $y_1, y_2 \in \mathfrak{m}$ finden mit

$$v_{j_0}(y_1) = v_{j_0}(y_2) = p_{j_0}, \quad v_1(y_1) = \lambda - \mu + 1, \quad v_1(y_2) = \lambda - \mu + 2. \tag{8}$$

Diese Wahl garantiert uns, daß $y_1 x^{\mu-1} z + R$ und $y_2 x^{\mu-1} z + R$ zwei k-linear unabhängige Elemente von $\mathfrak{m}^{-1}/R$ sind, wie man leicht mit Hilfe von (1), (6), (7) und (8) einsehen kann. Dieser Widerspruch zeigt, daß γ) richtig ist und vollendet den Beweis von Satz 2.

Da im Beweis von 3.1 gezeigt wurde, daß für vollständige Durchschnitte $J = \mathfrak{d}_1(R/k)$ ist, ist zu Satz 3 nichts mehr zu sagen. Es sei bemerkt, daß sich der Beweis von Satz 2 erheblich verkürzt, wenn man nur Kurven mit einem Zweig betrachtet.

Zum Schluß möchten wir noch zwei Anmerkungen für den affinen Fall machen, welche die Erzeugendenzahl von $\mathrm{Der}_k(R)$ betreffen. Sei k für den Rest von § 4 ein algebraisch abgeschlossener Körper der Charakteristik O und R der Koordinatenring einer reduzierten affinen Kurve über k. R besitze eine Darstellung als vollständiger Durchschnitt wie in (5), § 3.

Da $\mathrm{Der}_k(R)/D$ ein artinscher Modul ist, gilt nach dem Satz von Forster-Swan

$$\mu(\mathrm{Der}_k(R)/D) = \max_{\mathfrak{p} \in \mathrm{Sing}(R)} \{\mu_{\mathfrak{p}}(\mathrm{Der}_k(R)/D)\}.$$

Aus § 3 wissen wir, daß $(\mathrm{Der}_k(R)/D)_{\mathfrak{p}}$ für $\mathfrak{p} \in \mathrm{Sing}\, R$ ein kanonischer Modul von $R_{\mathfrak{p}}/\mathfrak{d}_1(R_{\mathfrak{p}}/k)$ ist. Nach [3], 5.12 gilt

$$\mu_{\mathfrak{p}}(\mathrm{Der}_k(R)/D) = r(R_{\mathfrak{p}}/\mathfrak{d}_1(R_{\mathfrak{p}}/k)) \tag{9}$$

und wir erhalten

4.1. **Korollar.** *R sei der Koordinatenring einer reduzierten Kurve in $\mathbb{A}_k^n$. R/k sei vollständiger Durchschnitt und*

$$r := \max_{\mathfrak{p} \in \mathrm{Sing}(R)} \{r(R_{\mathfrak{p}}/\mathfrak{d}_1(R_{\mathfrak{p}}/k))\}.$$

Dann ist

$$r \leqq \mu(\mathrm{Der}_k(R)) \leqq r + 1.$$

$\mathrm{Der}_k(R)/D$ *ist genau dann zyklisch, wenn alle Singularitäten von R quasihomogen sind.*

Schließlich wollen wir noch ebene Kurven betrachten.

4.2. *Bemerkung.* $R = k[X, Y]/(f) = k[x, y]$ sei der Koordinatenring einer reduzierten ebenen Kurve. Dann ist

$$\mu(\mathrm{Der}_k(R)) \leqq 2,$$

wobei für singuläre Kurven stets das Gleichheitszeichen gilt ([6]). R besitzt genau dann nur quasihomogene Singularitäten, wenn $\delta := f_y \frac{\partial}{\partial X} - f_x \frac{\partial}{\partial Y}$ ein basisches Element von $\mathrm{Der}_k(R)$ ist.

Nach Hilbert-Burch hat nämlich $R/\mathfrak{d}_1(R/k) = R/(f_x, f_y)$ eine freie Auflösung der Form

$$0 \to k[X,Y]^2 \to k[X,Y]^3 \xrightarrow{\begin{pmatrix} f \\ f_X \\ f_Y \end{pmatrix}} k[X;Y] \to R/\mathfrak{d}_1(R/k) \to 0$$

und somit wird $\ker(R^2 \xrightarrow{\binom{f_x}{f_y}} R)$ von zwei Elementen erzeugt. In §3 wurde gezeigt, daß $\mathrm{Der}_k(R) \cong \ker(R^2 \xrightarrow{\binom{f_x}{f_y}} R)$ ist. Die letzte Aussage in 4.2 folgt unmittelbar aus 4.1 und der in 3.1 gezeigten Tatsache, daß δ ein Basiselement von D ist.

Literatur

1. Buchweitz, R.-O., Greuel, G.-M.: The Milnor Number and Deformations of Complex Curve Singularities. Invent. Math. **58**, 241–281 (1980)
2. Greuel, G.-M., Martin, B., Pfister, G.: Numerische Charakterisierung quasihomogener Kurvensingularitäten. Preprint (1983)
3. Herzog, J., Kunz, E.: Der kanonische Modul eines Cohen-Macaulay-Rings. Lecture Notes in Math. **238**. Berlin-Heidelberg-New York: Springer 1971
4. Herzog, J., Kunz, E.: Die Wertehalbgruppe eines lokalen Rings der Dimension 1. Sitzungsber. Heidelb. Akad. Wiss. Math.-Nat. Kl. 2. 1971
5. Kunz, E., Ruppert, W.: Quasihomogene Singularitäten algebraischer Kurven. Manuscripta Math. **22**, 47–62 (1977)
6. Lipman, J.: Free derivation modules on algebraic varieties. Amer. J. Math. **87**, 874–898 (1965)
7. Scheja, G., Storch, U.: Über Spurfunktionen bei vollständigen Durchschnitten. J. Reine Angew. Math. **278/279**, 174–190 (1975)

Eingegangen am 9. November 1983

Math. Z. 187, 125–144 (1984)

Mathematische Zeitschrift

Centralizer Theorems for Hopf Type Galois Extensions

Yoichi Miyashita

Department of Mathematics, University of Tsukuba, Sakura-mura, Ibaraki, Japan

In this paper we will prove several centralizer theorems for Hopf type Galois extensions, in the form which are applicable to H-Galois extensions over a non-commutative R-ring or R-coring H. A commutative Hopf Galois extension was defined by Chase and Sweedler [6], and recently non-commutative Hopf Galois extensions were studied by Kreimer and Takeuchi [17], and Yokogawa [32]. Further, in [30], Ulbrich proved centralizer theorems of Hopf Galois extensions, in a separable extension in the sense of Hirata, and Yokogawa [33] may be connected with the last one. First we prove a centralizer theorem for H-weak Galois extensions over a finite R-module H (Proposition 1.3), where 'finite' means 'finitely generated and projective', and R is a non-commutative ring. Then we proceed to treat H-Galois extensions over a non-commutative R-ring or R-coring H, separately (cf. Lemma 1.6). In view of Hirata [12] and our [20], we can prove Theorems 1.7, 1.8, and 1.9, which give a correspondence between left H^*-Galois extensions over an R-ring (resp. R-coring) H^* and right H-Galois extensions over an R-coring (resp. R-ring) H, where H^* is the dual one of a finite right R-module H. As a corollary we obtain a centralizer theorem for Hopf Galois extensions. Since, for a group ring RG of a finite group G, a $(RG)^*$-Hopf Galois extension is nothing but a (generalized) crossed product of G, this may be considered as a generalization of [20; Theorem's 2.12 and 2.13], and is connected with Theorems in [30]. In 2, we state some remarks on Hopf Galois extensions, by using an isomorphism (# #), which is found in [17].

1. Weak Galois Extensions and Centralizer Theorems

Throughout this paper, every ring has $1(\neq 0)$, which is preserved by ring homomorphisms, inherited by subrings, and acts as the identity operator on modules. If M is a left A, right A'-module such that $(am)a'=a(ma')$ $(a\in A, m\in M, a'\in A')$ we call M an (A, A')-bimodule, and write ${}_AM_{A'}$. If $A=A'$ then we denote by M^A the submodule

$$\{m\in M \mid am=ma \text{ for all } a\in A\}.$$

By a ring extension A/B we means a ring homomorphism from B to A. Then A may be considered as a (B, B)-bimodule, and we call A^B the centralizer of B in A, which is a subring of A.

By R we denote a fixed ring, and the unadorned $\otimes$ means $\otimes_R$. By ${}_RH^*$ and H_R we denote a left R-module and a right R-module, which are finitely generated and projective over R. By Ψ, we denote a non-singular (R, R)-bilinear map from $H^* \times H$ to R, and we put $\Psi(h^*, h) = \langle h^*, h\rangle$ $(h^* \in H^*,\ h \in H)$. Namely Ψ satisfies the following conditions:

$$\begin{gathered}\langle rh^*, hr'\rangle = r\langle h^*, h\rangle r', \langle h^* + k^*, h\rangle = \langle h^*, h\rangle + \langle k^*, h\rangle,\\ \langle h^*, h+k\rangle = \langle h^*, h\rangle + \langle h^*, k\rangle,\\ H^* \xrightarrow{\sim} \mathrm{Hom}_R(H, R), \quad \text{by } h^* \mapsto (h \to \langle h^*, h\rangle), \quad \text{and}\\ H \xrightarrow{\sim} \mathrm{Hom}_R(H^*, R), \quad \text{by } h \mapsto (h^* \to \langle h^*, h\rangle),\\ \text{where } r, r' \in R,\ h^*, k^* \in H^*,\ h, k \in H.\end{gathered}$$

By T/R we denote a fixed ring extension of R. Let A/B, B'/A' be ring extensions such that both ${}_BA$ and $B'_{A'}$ are finitely generated and projective modules, and assume that there exist two given ring homomorphisms $A \to T$, $a \mapsto \bar{a}$, and $B' \to T$, $b' \mapsto \bar{b}'$ such that $\bar{A}' \subseteq T^A$ and $\bar{B} \subseteq T^{B'}$, where $\bar{A}'$ is the image of A' in T, under the map $A' \to B' \to T$. Take a left T-homomorphism

$$\alpha\colon {}_TT \otimes H^* \to {}_T\mathrm{Hom}(B'_{A'}, T_{A'}),$$

where $\mathrm{Hom}(B'_{A'}, T_{A'})$ is the module of all right A'-homomorphisms from B' to T, and $t'(t \otimes h^*) = (t't) \otimes h^*$, $(tf)(b') = t \cdot f(b')$ $(t, t' \in T,\ h^* \in H^*,\ f \in \mathrm{Hom}(B'_{A'}, T_{A'}),\ b' \in B')$. Then a bilinear map

$$H^* \times B' \to T, \quad (h^*, b) \mapsto \alpha(1_T \otimes h^*)(b') = h^*(b'),$$

is defined, and satisfies

$$(rh^*)(b') = r \cdot h^*(b'), \quad h^*(b'a') = h^*(b')a' \quad (r \in R,\ h^* \in H^*,\ b' \in B',\ a' \in A').$$

Applying $\mathrm{Hom}(-, {}_TT)$ to α, we get a right T-homomorphism

$$\alpha^+ : B' \otimes_{A'} T_T \to \mathrm{Hom}({}_RH^*, {}_RT)_T, \quad b' \otimes t \mapsto (h^* \to h^*(b')t).$$

Conversely, from α^+ we get $\alpha = (\alpha^+)^+$. Then α is an isomorphism if and only if so is α^+. In this case we call α a left H^*-weak Galois structure in T/R (or simply, in T). Symmetrically we can define a right H-Galois structure

$$\beta\colon H \otimes T_T \to \mathrm{Hom}({}_BA, {}_BT)_T$$

in T/R. From β we get a bilinear map

$$A \times H \to T, \quad (a, h) \mapsto \beta(h \otimes 1_T)(a) = [a]h,$$

which satisfies

$$[a](hr) = [a]h \cdot r, \quad [ba]h = b \cdot [a]h \quad (r \in R,\ h \in H,\ a \in A,\ b \in B).$$

Applying $\mathrm{Hom}(-, T_T)$ to β, we get a left T-isomorphism

$$\beta^+ : {}_T T \otimes_B A \to {}_T\mathrm{Hom}(H_R, T_R), \quad t\otimes a \mapsto (h \to t[a]h).$$

Conversely, from an isomorphism β^+, we get $\beta=(\beta^+)^+$. Further, there is a canonical right T-homomorphism

$$\delta: B' \otimes_{A'} T_T \to \mathrm{Hom}({}_B A, {}_B T), \quad b'\otimes t \mapsto (a \to b'(at)),$$

where $b' \in B'$, $t\in T$, $a\in A$. If the following diagram is a commutative one of isomorphisms we call it a right T-square of Galois structures:

$$(*) \qquad \begin{array}{ccc} B'\otimes_{A'} T & \xrightarrow[\alpha^+]{\sim} & \mathrm{Hom}({}_R H^*, {}_R T) \\ \delta \downarrow \wr & & \wr \downarrow \\ \mathrm{Hom}({}_B A, {}_B T) & \xleftarrow[\beta]{\sim} & H\otimes T. \end{array}$$

Applying $\mathrm{Hom}(-, T_T)$ to $(*)$, we get a "left T-square of Galois structures":

$$(**) \qquad \begin{array}{ccc} \mathrm{Hom}(B'_{A'}, T_{A'}) & \xleftarrow[\alpha]{\sim} & T\otimes H^* \\ \wr \uparrow \delta^+ & & \wr \uparrow \\ T\otimes_B A & \xrightarrow[\beta^+]{\sim} & \mathrm{Hom}(H_R, T_R), \end{array}$$

where $\delta^+(t\otimes a)=(b' \to (tb')a)$ $(t\in T,\ a\in A,\ b'\in B')$. For any $h\in H$, let $\sum_j b'_j\otimes y_j$ be the element of $B'\otimes_{A'} T$, corresponding to $h\otimes 1_T$ $(\in H\otimes T)$ in the diagram $(*)$. Then the commutativity of $(*)$ implies that

$$[a]h=\sum_j b'_j a y_j, \quad \sum_j h^*(b'_j)\, y_j=\langle h^*, h\rangle 1_T \quad (a\in A,\ h^*\in H^*).$$

For any $h^*\in H^*$, let $\sum_i x_i\otimes a_i$ be the element of $T\otimes_B A$, corresponding to $1\otimes h^*$ $(\in T\otimes H^*)$ in the diagram $(**)$. Then the commutativity of $(**)$ implies that

$$h^*(b')=\sum_i x_i b' a_i, \quad \sum_i x_i[a_i]h=\langle h^*, h\rangle 1_T \quad (b'\in B',\ h\in H).$$

If $A'=\{x\in B' \mid h^*(b'x)=h^*(b')x \text{ for all } h^*\in H^*,\ b'\in B'\}$ then we call B'/A' (more precisely, α) a left H^*-weak Galois extension in T/R. Symmetrically we can define a right H-weak Galois extension in T/R.

The above definition is a very primitive one, but it provides a good symmetric situation to discuss centralizer theorems for (non-commutative) Hopf type Galois extensions.

Now we explain a notation $X_A|Y_A$ for right A-modules X_A, Y_A, and state a well known lemma about this. If X_A is a direct summand of a finite direct sum of copies of Y_A we write $X_A|Y_A$. This is equivalent to that there are $\phi_1,\dots,\phi_v\in\mathrm{Hom}(X_A, Y_A)$, $\psi_1,\dots,\psi_v\in\mathrm{Hom}(Y_A, X_A)$ such that $\sum_i \psi_i\circ\phi_i=\mathrm{id}_X$. If $X_A|Y_A$ and $Y_A|X_A$ then we say that X_A and Y_A are similar (as A-modules) (cf. [22]). Then the following is easily seen.

Lemma 1.1. *Assume* $X_A | Y_A$, *and put* $A^* = \mathrm{End}(X_A)$ *and* $A^+ = \mathrm{End}(Y_A)$. *Then the following hold.*

(1) $\mathrm{Hom}(X_A, Y_A)$ *is finitely generated and projective as a left* A^+*-module, and* $\mathrm{End}({}_{A^+}\mathrm{Hom}(X_A, Y_A)) \xrightarrow{\sim} A^*$ *canonically. (Therefore, as is well known,* $\mathrm{Hom}(X_A, Y_A)$ *is a generator as a right* A^**-module.)*

(2) *The canonical map* $X \to \mathrm{Hom}({}_{A^+}\mathrm{Hom}(X_A, Y_A), {}_{A^+}Y)$, $x \mapsto (f \to f(x))$ $(x \in X, f \in \mathrm{Hom}(X_A, Y_A))$ *is an isomorphism.*

Proof. Cf. Hirata [11, 12], or [20].

The following lemma is also known (cf. Hirata [11, 12]). But we give its proof for completeness.

Lemma 1.2. *Let* A/B *and* T/A *be ring extensions, and put* $T^A = A^*$, $T^B = B^*$, $T^{A^*} = A^{**}$, *and* $T^{B^*} = B^{**}$. *Assume that* ${}_BA | {}_BB$ *(i.e.* ${}_BA$ *is finitely generated and projective) and that* ${}_TT \otimes_B A_A | {}_TT_A$ *as* (T, A)*-bimodules. Then* $B^*_{A^*} | A^*_{A^*}$, *and*

$$ {}_TT \otimes_B A_{A, B^*} \xrightarrow{\sim} {}_T\mathrm{Hom}(B^*_{A^*}, T_{A^*})_{A, B^*}, \qquad t \otimes a \mapsto (b^* \to t b^* a), $$

where $(t \otimes a) a' = t \otimes (a a')$, $(t \otimes a) b^* = (t b^*) \otimes a$, $(f a)(b^*) = f(b^*) a$, $(f d^*)(b^*) = f(d^* b^*)$ $(a, a' \in A, t \in T, d^*, b^* \in B^*, f \in \mathrm{Hom}(B^*_{A^*}, T_{A^*}))$. *Therefore, applying* $\mathrm{Hom}(-, {}_TT)$, *we obtain an isomorphism*

$$ {}_{A, B^*}B^* \otimes_{A^*} T_T \xrightarrow{\sim} {}_{A, B^*}\mathrm{Hom}({}_BA, {}_BT)_T, \qquad b^* \otimes t \mapsto (a \to b^* a t), $$

where operations of rings are right-left dual to the previous ones. Further, we have a canonical isomorphism

$$ B^{**} \otimes_B A \xrightarrow{\sim} A^{**}, \qquad b^{**} \otimes a \mapsto b^{**} a. $$

Proof. If we apply Lemma 1.1 to (T, A)-bimodules ${}_TT \otimes_B A_A$ and ${}_TT_A$ then we see that $B^*_{A^*}$ $(\xrightarrow{\sim} \mathrm{Hom}({}_BA_A, {}_BT_A)_{A^*} \xrightarrow{\sim} \mathrm{Hom}({}_TT \otimes_B A_A, {}_TT_A)_{A^*})$ is finitely generated and projective, and $T \otimes_B A \xrightarrow{\sim} \mathrm{Hom}(B^*_{A^*}, T_{A^*})$, $t \otimes a \mapsto (b^* \to t b^* a)$. Applying $\mathrm{Hom}(-, {}_TT)$, we obtain an isomorphism δ: $B^* \otimes_{A^*} T_T \xrightarrow{\sim} \mathrm{Hom}({}_BA, {}_BT)_T$, $b^* \otimes t \mapsto (a \to b^* a t)$, where $b^* \in B^*$, $t \in T$, $a \in A$. This is evidently a left B^*, A, right T-isomorphism. Since ${}_BA | {}_BB$ we have ${}_{B^*}B^* \otimes_{A^*} T_T | {}_{B^*}T_T$, and so we have an isomorphism $T \otimes_{B^{**}} A^{**} \xrightarrow{\sim} \mathrm{Hom}(B^*_{A^*}, T_{A^*})$, $t \otimes a^{**} \mapsto (b^* \to t b^* a^{**})$ $(t \in T, a^{**} \in A^{**}, b^* \in B^*)$, similarly. Combining this with the isomorphism

$$ T \otimes_B A \xrightarrow{\sim} \mathrm{Hom}(B^*_{A^*}, T_{A^*}), \quad t \otimes a \mapsto (b^* \to t b^* a), $$

we have isomorphisms

$$ (T \otimes_{B^{**}} A^{**})^{B^*} \xrightarrow{\sim} (\mathrm{Hom}(B^*_{A^*}, T_{A^*}))^{B^*} \xrightarrow{\sim} (T \otimes_B A)^{B^*}. $$

Noting that both ${}_{B^{**}}A^{**}$ and ${}_BA$ are finitely generated and projective we have $(T \otimes_{B^{**}} A^{**})^{B^*} = A^{**}$ and $(T \otimes_B A)^{B^*} = B^{**} \otimes_B A$. Thus $B^{**} \otimes {}_BA \xrightarrow{\sim} A^{**}$, $b^{**} \otimes a \mapsto b^{**} a$, completing the proof.

Remark. Under the assumption of Lemma 1.2, if ${}_BB | {}_BA$ (i.e. ${}_BA$ is a generator) then ${}_TT_{B^*} | {}_T\mathrm{Hom}(B^*_{A^*}, T_{A^*})_{B^*}$, or equivalently, ${}_{B^*}T_T | {}_{B^*}B^* \otimes_{A^*} T_T$. On the other hand, if ${}_TT_A | {}_TT \otimes_B A_A$ then $A^*_{A^*} | B^*_{A^*}$, by Lemma 1.1 (1). (Cf. Hirata [12].)

Let α be a left H^*-weak Galois structure in T/R, and assume that ${}_{B'}B' \otimes_{A'} T_T |_{B'} T_T$. For any $b' \in B'$ (resp. $a' \in A'$), $\bar{b}'$ (resp. $\bar{a}'$) is the image of b' (resp. a') in T, and put $\bar{B}' = \{\bar{b}' \mid b' \in B'\}$ and $\bar{A}' = \{\bar{a}' \mid a' \in A'\}$. If $\bar{b}' = 0$ in T, then $b' T = 0$, and so $b' \otimes T = 0$ in $B' \otimes_{A'} T$. Then, for any $h^* \in H^*$, we have $h^*(b') T = 0$, and so $h^*(b') = 0$. Therefore

$$h^*(\bar{b}') = h^*(b') \; (h^* \in H^*, \; b' \in B')$$

is well defined. The following proposition is fundamental.

Proposition 1.3. *Let* α: ${}_T T \otimes H^* \xrightarrow{\sim} {}_T\mathrm{Hom}(B'_{A'}, T_{A'})$ *be a left H^*-weak Galois structure in T/R, and assume that* ${}_{B'}B' \otimes_{A'} T_T |_{B'} T_T$. *Put* $T^{A'} = A$ *and* $T^{B'} = B$, *and let β be the isomorphism which renders the following diagram commutative:*

$$\begin{array}{ccc} B' \otimes_{A'} T & \xrightarrow[\alpha^+]{\sim} & \mathrm{Hom}({}_R H^*, {}_R T) \\ {\scriptstyle \delta} \big\downarrow \wr & & \wr \big\downarrow \\ \mathrm{Hom}({}_B A, {}_B T) & \xleftarrow[\beta]{\sim} & H \otimes T, \end{array}$$

where δ is defined by $\delta(b' \otimes t) = (a \to b' a t)$ $(b' \in B', t \in T, a \in A)$. *Then*

(1) *A/B (or β) is a right H-weak Galois extension in T/R, and* ${}_T T \otimes_B A_A |_T T_A$ *(or equivalently,* ${}_A\mathrm{Hom}({}_B A, {}_B T)_T |_A T_T$*).*

(2) *Therefore, if we put* $T^A = A^*$ *and* $T^B = B^*$ *then B^*/A^* is a left H^*-weak Galois extension in T/R, and* ${}_{B^*}B^* \otimes_{A^*} T_T |_{B^*} T_T$. *Furthermore,* $B' \otimes_{A'} A^* \xrightarrow{\sim} B^*$, $b' \otimes a^* \mapsto b' a^*$ $(b' \in B', a^* \in A^*)$.

(3) $\bar{B}'_{\bar{A}'} | \bar{A}'_{\bar{A}'}$, *and the following diagram is a commutative one of isomorphisms:*

$$\begin{array}{ccc} B' \otimes_{A'} T & \xrightarrow[\alpha^+]{\sim} & \mathrm{Hom}({}_R H^*, {}_R T), \\ {\scriptstyle \rho} \big\downarrow \wr & \nearrow_{\bar{\alpha}^+}^{\sim} & \\ \bar{B}' \otimes_{\bar{A}'} T & & \end{array}$$

where ρ is the canonical epimorphism. The Galois structure of $\bar{B}'/\bar{A}'$ coincides with the one induced by B^*/A^* (of (2)). If $A'_{A'} | B'_{A'}$ *then* $\bar{A}'_{\bar{A}'} | \bar{B}'_{\bar{A}'}$, *and $\bar{B}'/\bar{A}'$ is a left H^*-weak Galois extension in T/R.*

Proof. (1), (2) It suffices to prove that if $x \in A$, and $[xa]h = x[a]h$ for all $a \in A$, $h \in H$ then $x \in B$. Let h be any element of H, and let $\sum b'_j \otimes y_j$ be the element of $B' \otimes_{A'} T$ corresponding to $h \otimes 1_T$ $(\in H \otimes T)$ (cf. diagram (*)). Then, $[c]h = \sum b'_j c y_j$ for all $c \in A$, and so $\sum b'_j x a y_j y = \sum x b'_j a y_j y$ for all $a \in A$, $y \in T$. Since $H \otimes T = (H \otimes 1) T_T \xrightarrow{\sim} B' \otimes_{A'} T_T$, this implies that $b' x a y = x b' a y$ for all $b' \in B'$, $y \in T$. Thus $x \in T^{B'} = B$. By Lemma 1.2, we have ${}_T T \otimes_B A_A |_T T_A$. Similarly we obtain (2), by (1) and Lemma 1.2. (3) Any $f \in \mathrm{Hom}(B'_{A'}, T_{A'})$ is induced by some element of $T \otimes H^*$, and we know that $\bar{b}' = 0$ $(b' \in B')$ implies that $h^*(b') = 0$ for all $h^* \in H^*$. Therefore f canonically induces an element $\bar{f} \in \mathrm{Hom}(\bar{B}'_{\bar{A}'}, T_{\bar{A}'})$. Then the projective coordinate system for $B'_{A'}$ induces the one for $\bar{B}'_{\bar{A}'}$. Thus $\bar{B}'_{\bar{A}'} | \bar{A}'_{\bar{A}'}$, and $\bar{\alpha}$: $T \otimes H^* \to \mathrm{Hom}(\bar{B}'_{\bar{A}'}, T_{\bar{A}'})$ is induced by α. Then we have a commutative diagram:

$$\begin{tikzcd} B'\otimes_{A'}T \arrow[r, "\sim", "\alpha^+"'] \arrow[d, "\rho\ \wr"'] & \mathrm{Hom}({}_RH^*, {}_RT), \\ \bar{B}'\otimes_{\bar{A}'}T \arrow[ur, "\sim", "\bar{\alpha}^+"'] & \end{tikzcd}$$

where ρ is the canonical epimorphism. Thus ρ and $\bar{\alpha}^+$ are ismorphisms. For any $h^*\in H^*$, let $\sum x_i\otimes a_i$ be the element of $T\otimes_B A$ which corresponds to $1_T\otimes h^*$ $(\in T\otimes H^*)$. Then $h^*(\bar{b}')=\sum x_i\bar{b}' a_i$ for all $b'\in B'$, where $h^*(\ )$ means the operation induced by the one on B^*/A^*. On the other hand, $h^*((\bar{b}'))=\sum x_i \bar{b}' a_i$ for all $b'\in B'$, too, where $h^*((\ ))$ means the operation induced by the one of H^* defined on B'. Hence $h^*((\bar{b}'))=h^*(\bar{b}')$ for all $b'\in B'$. If $A'_{A'}|B'_{A'}$ then, as is well known, there is a right A'-homomorphism π from B' to A' such that $\pi|A'=\mathrm{id}_{A'}$. Then π induces $\bar{\pi}$: $\bar{B}'_{\bar{A}'}\to\bar{A}'_{\bar{A}'}$ such that $\bar{\pi}|\bar{A}'=\mathrm{id}_{\bar{A}'}$. Hence $\bar{A}'_{\bar{A}'}|\bar{B}'_{\bar{A}'}$. Let $\bar{x}$ be an element of $\bar{B}'$ such that $h^*(\bar{b}'\bar{x})=h^*(\bar{b}')\bar{x}$ for all $h^*\in H^*$, $\bar{b}'\in\bar{B}'$. For π, there exists an element $\sum t_s\otimes h_s^*$ in $T\otimes H^*$ such that $\sum t_s\cdot h_s^*(y)=\overline{\pi(y)}$ for all $y\in B'$. Then $\bar{A}'\ni\bar{\pi}(\bar{x})=\overline{\pi(x)}=\sum t_s\cdot h_s^*(x)=\sum t_s\cdot h_s^*(\bar{x})=\sum t_s\cdot h_s^*(\bar{1})\,\bar{x}=\bar{x}$. Hence $\bar{x}\in\bar{A}'$. This completes the proof.

Remark 1. If B'/A' is a left H^*-weak Galois extension in T/R, and T is a faithful left R-module, then $h^*(B')=0$ implies that $h^*=0$. In fact, $h^*(x)y=0$ for all $x\in B'$, $y\in T$, and so $\langle h^*,h\rangle\, 1_T=0$ for all $h\in H$, because $B'\otimes_{A'}T\xrightarrow{\sim}\mathrm{Hom}({}_RH^*, {}_RT)$. Then, by assumption, $\langle h^*,h\rangle=0$ for all $h\in H$. Thus $h^*=0$.

Remark 2. If we want to prove that $\bar{B}'_{\bar{A}'}|\bar{A}'_{\bar{A}'}$ and that if $A'_{A'}|B'_{A'}$ then $\bar{A}'_{\bar{A}'}|\bar{B}'_{\bar{A}'}$, the assumption concerning H^* is not necessary. In fact, if ${}_{B'}B'\otimes_{A'}T_T|{}_{B'}T_T$, $B'_{A'}|A'_{A'}$, $T^{B'}=B$, and $T^{A'}=A$ we have a commutative diagram of isomorphisms

$$\begin{tikzcd} B'\otimes_{A'}T \arrow[r, "\sim"] \arrow[d, "\wr"'] & \mathrm{Hom}({}_BA, {}_BT), \\ \bar{B}'\otimes_{\bar{A}'}T \arrow[ur, "\sim"'] & \end{tikzcd}$$

and then we have a commutative diagram of isomorphisms

$$\begin{tikzcd} \mathrm{Hom}(B'_{A'}, T_{A'}) & T\otimes_B A. \arrow[l, "\sim"'] \arrow[dl, "\sim"] \\ \mathrm{Hom}(\bar{B}'_{\bar{A}'}, T_{\bar{A}'}) \arrow[u, "\wr"] & \end{tikzcd}$$

Therefore any right A'-homomorphism $B'_{A'}\to T_{A'}$ induces a right $\bar{A}'$-homomorphism $\bar{B}'_{\bar{A}'}\to T_{\bar{A}'}$. Then, as in the proof of (3), we can do.

Remark 3. From the proof of (3), we can see the following: Let α: ${}_TT\otimes H^*\xrightarrow{\sim}{}_T\mathrm{Hom}(B'_{A'}, T_{A'})$ be a left H^*-weak Galois structure in T/R. Assume that $A'\subseteq B'\subseteq T$ and $A'_{A'}|B'_{A'}$. Then B'/A' is a left H^*-weak Galois extension in T/R.

Remark 4. Let B'/A' be a left H^*-weak Galois structure in T/R. Let x be an element of A' such that $h^*(x)=h^*(1)x$ for all $h^* \in H^*$. Then, for any $h^* \in H^*$, $y \in B'$, there exists an element $\sum t_j \otimes h_j^*$ of $T \otimes H^*$ such that $\sum t_j \cdot h_j^*(z)=h^*(yz)$ for all $z \in B'$. Then

$$h^*(yx)=\sum t_j \cdot h_j^*(x)=\sum t_j \cdot h_j^*(1)\,x=h^*(y)\,x.$$

Put

$$A^+=\{x \in B' \mid h^*(x)=h^*(1)\,x \quad \text{for all } h^* \in H^*\},$$

and assume that $A' \subseteq B' \subseteq T$. Then any element of $\operatorname{Hom}(B'_{A'}, A'_{A'})$ lies in

$$\operatorname{Hom}(B'_{A'}, T_{A'})=\operatorname{Hom}(B'_{A^+}, T_{A^+}),$$

and so, in $\operatorname{Hom}(B'_{A^+}, A^+_{A^+})$. Therefore, as in the proof of (3), B'/A^+ is a left H^*-weak Galois extension in T/R. Further, assume that $A' \subseteq T^R$. Then, for any $r \in R$, $(b' \to b'r)$ belongs to $\operatorname{Hom}(B'_{A'}, T_{A'})$ $(\subseteq \operatorname{Hom}(B'_{A^+}, T_{A^+}))$. Hence $A^+ \subseteq T^R$.

If B'/A' is a left H^*-weak Galois extension in "B'/R", we briefly call B'/A' a left H^*-weak Galois extension. A right H-weak Galois extension is defined similarly.

Theorem 1.4 (1) *Let B'/A' be a left H^*-weak Galois extension, and let T/B' be a ring extension such that ${}_{B'}B' \otimes_{A'} T_T \mid {}_{B'}T_T$. Put $T^{B'}=B$ and $T^{A'}=A$. Then A/B is a right H-weak Galois extension in T/R, $B \subseteq T^R$, ${}_TT \otimes_B A_A \mid {}_TT_A$, and*

$$[ab]\,h=[a]\,h \cdot b \quad \textit{for all } a \in A,\ b \in B,\ h \in H.$$

(2) *Conversely, let A/B be a right H-weak Galois extension in T/R such that $B \subseteq T^R$, and $[ab]\,h=[a]\,h \cdot b$ for all $a \in A$, $b \in B$, $h \in H$. If ${}_TT \otimes_B A_A \mid {}_TT_A$, $T^A=A'$, and $T^B=B'$, then B'/A' is a left H^*-weak Galois extension (in B'/R).*

$$\begin{tikzcd}[column sep=small]
& & T \arrow[dl, dash] \arrow[dr, dash] & & \\
& B' \arrow[dl, dash] \arrow[dr, dash] & & A \arrow[dr, dash] & \\
A' & & R \arrow[rr, dotted, dash] & & B
\end{tikzcd}$$

Proof. (1) It is evident that $B \subseteq T^R$. For any $h \in H$, let $\sum b'_j \otimes y_j$ be the element of $B' \otimes_{A'} B'$, corresponding to $h \otimes 1_{B'}$ $(\in H \otimes B')$ under the isomorphism

$$H \otimes B' \xrightarrow{\sim} \operatorname{Hom}({}_RH^*, {}_RB') \xrightarrow{\sim} B' \otimes_{A'} B'.$$

Then, for any $a \in A$, $b \in B$, there holds $[ab]\,h=\sum b'_j ab y_j=\sum (b'_j a y_j)\,b=[a]\,h \cdot b$, as desired. (2) For any $h^* \in H^*$, let $\sum x_i \otimes a_i$ be the element of $T \otimes_B A$, corresponding to $1 \otimes h^*$ $(\in T \otimes H^*)$. Then,

$$\sum x_i\,[a_i]\,h=\langle h^*, h\rangle\,1_T \quad \text{for all } h \in H,$$

and so, for any $b \in B$,

$$\sum b x_i\,[a_i]\,h=\sum x_i\,[a_i]\,h \cdot b=\sum x_i\,[a_i b]\,h$$

by assumption. Therefore $\sum b x_i \otimes a_i = \sum x_i \otimes a_i b$ in $T \otimes_B A$. Then, for any $b' \in B'$,

$$\sum b x_i b' \otimes a_i = \sum x_i b' \otimes a_i b$$

holds, and so

$$h^*(b') = \sum x_i b' a_i \in B' \qquad \text{for all } b' \in B',\ h^* \in H^*.$$

Then, as $B'_{A'} | A'_{A'}$, the isomorphism $B' \otimes_{A'} T \xrightarrow{\sim} \mathrm{Hom}({}_R H^*, {}_R T)$ induces the isomorphism

$$B' \otimes_{A'} B' = (B' \otimes_{A'} T)^B \xrightarrow{\sim} \mathrm{Hom}({}_R H^*, {}_R T)^B = \mathrm{Hom}({}_R H^*, {}_R B').$$

Hence B'/A' is a left H^*-weak Galois extension.

By the above theorem we get the following

Corollary 1.5. *Let B'/A' be a left H^*-weak Galois extension (in B'/R), and let T/B' be a ring extension such that ${}_{B'}B' \otimes_{A'} T_T | {}_{B'} T_T$. Assume that $A' \subseteq B'^R$ and*

$$h^*(a' b') = a' \cdot h^*(b') \quad \text{for all } h^* \in H^*, \quad a' \in A',\ b' \in B'.$$

Then A/B is a right H-weak Galois extension, $B \subseteq A^R$, ${}_T T \otimes_B A_A | {}_T T_A$, and

$$[a b]\, h = [a]\, h \cdot b \quad \text{for all } h \in H, \quad a \in A,\ b \in B.$$

$$\begin{tikzcd}[column sep=small] & & T \arrow[dl, dash] \arrow[dr, dash] & & \\ & B' \arrow[dl, dash] \arrow[dr, dash] & & A \arrow[dl, dash] \arrow[dr, dash] & \\ A' \arrow[rr, dotted, dash] & & R \arrow[rr, dotted, dash] & & B. \end{tikzcd}$$

Now we consider a special case. Take a right T-square of Galois structures (*) and its T-dual (**). We assume that $\bar{A}$ (the image of A in T)$\subseteq T^R$. It is easily seen that the isomorphism β^+ is a right A-isomorphism if and only if so is the isomorphism α^+, where $(fa)(h) = f(h)\, a$, $(ag)(h^*) = a \cdot g(h^*)$, and $a(b' \otimes t) = b' \otimes (at)$ ($f \in \mathrm{Hom}(H_R, T_R)$, $a \in A$, $h \in H$, $h^* \in H^*$, $g \in \mathrm{Hom}({}_R H^*, {}_R T)$, $b' \in B'$, $t \in T$). Furthermore the former is equivalent to that $[a]\, h = [1_A]\, h \cdot a$ for all $h \in H$, $a \in A$, and the latter is equivalent to that $h^*(b') \in T^A$ for all $h^* \in H^*$, $b' \in B'$. If β^+ is a right A-isomorphism, $T^A = A'$, and $T^B = B'$, then $(\alpha^+)^{-1}$ induces an isomorphism

$$H \otimes A'_{A'} \xrightarrow{\sim} B'_{A'}, \qquad h \otimes a' \mapsto [1_A]\, h \cdot a'.$$

Conversely if the isomorphism $H \otimes T \xrightarrow{\sim} (\mathrm{Hom}({}_R H^*, {}_R T) \xrightarrow{\sim})\ B' \otimes_{A'} T$ is induced by some isomorphism σ: $H \otimes A'_{A'} \xrightarrow{\sim} B'_{A'}$ then the isomorphism $H \otimes T \xrightarrow{\sim} B' \otimes_{A'} T$ is a left A-isomorphism, and $\beta^+(1_T \otimes 1_A) = (h \to \overline{\sigma(h \otimes 1)})$ holds, where $\overline{\sigma(h \otimes 1)}$ is the image of $\sigma(h \otimes 1)$ in T.

For example, let A/B be a left H-weak Galois extension (in A/R), and let T be the ring of all endomorphisms of A, acting on the "right" side of A. Then there are canonical ring homomorphisms $A \to T$, $a \mapsto (x \to xa)$ $(x, a \in A)$, and $R^0 \to T$, $r^0 \mapsto (x \to rx)$ $(r \in R, x \in A)$, where R^0 is the opposite ring of R. Then $A_B | B_B$ implies that $A \otimes_B A_A | A_A$. Assume ${}_B A | {}_B B$, and apply $\mathrm{Hom}(-, A)$. Then ${}_T T \otimes_B A_A | {}_T T_A$. Similarly, from the isomorphism

$$A \otimes_B A_A \xrightarrow{\sim} \mathrm{Hom}({}_R H, {}_R A)_A = \mathrm{Hom}(H_{R^0}, A_{R^0})_A,$$

we have a (T, A)-isomorphism

$$T \otimes_B A \xrightarrow{\sim} \mathrm{Hom}(H_{R^0}, T_{R^0}), \qquad 1 \otimes a \mapsto (h \to (x \to h(x)\, a) = [a]\, h),$$

where $a, x \in A$, $h \in H$. If $y \in A$, and $[ya]\, h = y[a]\, h$ for all $a \in A$, $h \in H$ then $y[1_A]\, h = [1_A]\, h \cdot y$ for all $h \in H$. Since $[1_A]\, h = (x \to h(x))$ $(x \in A)$, we have $y \in B$. Thus, if we assume that ${}_B A | {}_B B$ then A/B is a right H-weak Galois extension in T/R^0.

In the following we consider the case when H is an R-ring or R-coring over a (non-commutative) ring R.

If H and H^* are (R, R)-bimodules, and the non-singular (R, R)-bilinear map Ψ satisfies

$$\langle h^* r, h \rangle = \langle h^*, rh \rangle \quad (h^* \in H^*, r \in R, h \in H),$$

then Ψ is said to be a non-singular (R, R, R)-bilinear map. By an R-ring T/R we mean a ring homomorphism from R to T, and an R-ring homomorphism f from T/R to T'/R is a ring homomorphism which makes the following diagram commutative:

$$\begin{tikzcd} R \arrow[r] \arrow[dr] & T \arrow[d, "f"] \\ & T' \end{tikzcd}$$

Lemma 1.6. *Let H, H^* be (R, R)-bimodules, and Ψ a non-singular (R, R, R)-bilinear map. Let $B'/R \to T/R$ be an R-ring homomorphism, and consider a right T-square of Galois structures $(*)$. Then the following are equivalent:*

(1) $(h^* r)(b') = h^*(rb')$ *for all* $h^* \in H^*$, $r \in R$, $b' \in B'$.

(2) $[a](rh) = r[a]\, h$ *for all* $a \in A$, $r \in R$, $h \in H$.

Proof. As is easily seen, (1) is equivalent to that α^+ (in $(*)$) is a left R-isomorphism, where $r(b' \otimes t) = (rb') \otimes t$, $(rg)(h^*) = g(h^* r)$ $(r \in R$, $b' \in B'$, $t \in T$, $g \in \mathrm{Hom}({}_R H^*, {}_R T)$, $h^* \in H^*)$. On the other hand, (2) is equivalent to that β is a left R-isomorphism, where $(rf)(a) = r \cdot f(a)$, $r(h \otimes t) = (rh) \otimes t$ $(f \in \mathrm{Hom}({}_B A, {}_B T)$, $r \in R$, $h \in H$, $t \in T$, $a \in A)$. Then, since δ is a left R-isomorphism, we get the equivalence $(1) \Leftrightarrow (2)$.

Here we recall the relation between R-rings and R-corings. Assume that Ψ is a non-singular (R, R, R)-bilinear map. First we take a ring homomorphism u: $R \to H$, $r \mapsto \bar{r}$ such that $rh = \bar{r}h$, $hr = h\bar{r}$ for all $r \in R$, $h \in H$. By M_0 we denote the multiplication $H \otimes H \to H$. Applying $\mathrm{Hom}(-, R_R)$ to u and M_0, we obtain (R, R)-homomorphisms u^*: $H^* \to R$, M_0^*: $H^* \to H^* \otimes H^*$. We put $M_0^*(h^*) = \sum_{(h^*)} h^*_{(2)} \otimes h^*_{(1)}$, which is determined by the following

$$(***)\qquad \langle h^*, hk\rangle=\sum_{(h^*)}\langle h^*_{(2)}\langle h^*_{(1)}, h\rangle, k\rangle \quad (h^*\in H^*, h, k\in H).$$

Then H^* is a right H-module defined by $[h^*]h=\sum_{(h^*)} h^*_{(2)}\langle h^*_{(1)}, h\rangle$. Put $h=1_H$. Then we have $h^*=\sum_{(h^*)} h^*_{(2)}\langle h^*_{(1)}, 1_H\rangle$ for all $h^*\in H^*$. If $k\in H^R$ then $k(h^*)$ is defined by $\langle k(h^*), h\rangle=\langle h^*, hk\rangle$ $(h^*\in H^*, h\in H)$, and H^* is a left H^R-module over a ring H^R. It is easily seen that

$$k(h^*)=\sum_{(h^*)}\langle h^*_{(2)}, k\rangle h^*_{(1)} \quad (k\in H^R, h^*\in H^*).$$

Put $k=1_H$. Then we have

$$h^*=\sum_{(h^*)}\langle h^*_{(2)}, 1_H\rangle h^*_{(1)}.$$

Furthermore the associativity of M_0 implies the coassociativity of the comultiplication M_0^*. (We adopt (2)-(1) sigma notation instead of (1)-(2) sigma notation, in accordance with the one of the usual Hopf algebra theory.) Conversely, assume that H^* is an R-coring by (R, R)-homomorphisms ε: $H^*\to R$ and

$$\Delta: H^*\to H^*\otimes H^*, h^*\mapsto \sum_{(h^*)} h^*_{(2)}\otimes h^*_{(1)} \quad (h^*\in H^*).$$

Namely Δ is coassociative, and

$$h^*=\sum_{(h^*)}\varepsilon(h^*_{(2)})\, h^*_{(1)}=\sum_{(h^*)} h^*_{(2)}\,\varepsilon(h^*_{(1)}) \quad \text{for all } h^*\in H^*.$$

Then there is a unique element 1_H of H such that

$$\varepsilon(h^*)=\langle h^*, 1_H\rangle$$

for all $h^*\in H^*$. Applying $\mathrm{Hom}(-, {}_RR)$ to Δ we have an associative multiplication Δ^*: $H\otimes H\to H$, and we obtain a ring H. Further we can see that $hr=h(r\cdot 1_H)$, $rh=(r\cdot 1_H)h$, and $r\cdot 1_H=1_H\cdot r$ for all $r\in R$, $h\in H$. Thus the map $(r\to r\cdot 1_H)$ is a ring homomorphism, and 1_H is the identity of H. The multiplication of H is determined by $(***)$. Similarly an R-ring structure of H^* determines an R-coring structure of H, and conversely:

$$(****)\qquad \langle h^* k^*, h\rangle=\sum_{(h)}\langle h^*, \langle k^*, h_{(2)}\rangle h_{(1)}\rangle$$

where $\Delta(h)=\sum_{(h)} h_{(2)}\otimes h_{(1)}$, $h\in H$, $h^*, k^*\in H^*$.

For example, let G be a finite group with unit element e, and let R be a (non-commutative) ring. Let $H=RG$ be the group ring. Let $H^*/R=\bigoplus_{\sigma\in G} Rw_\sigma$ be a ring extension such that $w_\sigma w_\sigma=w_\sigma$, $w_\sigma w_\tau=0\,(\sigma\neq\tau)$, $w_\sigma r=rw_\sigma$ $(r\in R, \sigma, \tau\in G)$. Then $\sum_{\sigma\in G} w_\sigma$ is the identity of H^*, and a non-singular (R, R, R)-bilinear map Ψ:

$H^* \times H \to R$ is defined by $\Psi(w_\sigma, \sigma)=1$ and $\Psi(w_\sigma, \tau)=0\,(\sigma \neq \tau)$. The R-ring structure of H induces an R-coring structure of H^*:

$$\varepsilon(w_\sigma)=0\,(\sigma \neq e),\ \varepsilon(w_e)=1,\ \Delta(w_\sigma)=\sum_{\mu\nu=\sigma} w_\nu \otimes w_\mu .$$

The R-coring structure of H is the usual one:

$$\varepsilon(\sigma)=1 \quad \text{and} \quad \Delta(\sigma)=\sigma\otimes\sigma \qquad \text{for all } \sigma\in G.$$

Let H^* be an R-coring. If a left H^*-Galois extension B'/A' in T/R satisfies the following conditions we call B'/A' (or α) a left H^*-coring Galois extension in T/R:

(#) $(1)_l$ $h^*(1_{B'})=\varepsilon(h^*)\,1_T$ for all $h^*\in H^*$.

$(2)_l$ $h^*(xy)=\sum_{(h^*)} h^*_{(2)}(x)\,h^*_{(1)}(y)$ for all $h^*\in H^*$, $x, y\in B'$.

$(3)_l$ $T^R \supseteq \bar{A}'$ (the image of A' in T).

$(4)_l$ $(h^* r)(x)=h^*(x)\,r$ for all $h^*\in H^*$, $r\in R$, $x\in B'$.

Note that $h^*(a'b')=a'\cdot h^*(b')$ ($a'\in A'$, $b'\in B'$, $h^*\in H^*$) follows from (1), (2) and (3), in this case. Similarly we can define a right H-coring Galois extension, when H is an R-coring. Then $(2)_r$ (corresponding to $(2)_l$) is

$$[xy]\,h=\sum_{(h)}[x]\,h_{(2)}\cdot[y]\,h_{(1)}, \qquad \text{where } \Delta(h)=\sum_{(h)} h_{(2)}\otimes h_{(1)}.$$

Let H be an R-ring, and A/B a right H-weak Galois extension (in A/R). If A is a right H-module over a ring H such that

$$[a]\,(rh)=[ar]\,h$$

for all $a\in A$, $r\in R$, $h\in H$, then we call A/B a right H-ring Galois extension. Similarly we can define a left H^*-ring Galois extension, when H^* is an R-ring. Then we can prove the following

Theorem 1.7. *Let Ψ be a non-singular (R, R, R)-bilinear map, and H^* an R-coring (or equivalently, let H be an R-ring). Further, take a right T-square of Galois structures $(*)$.*

(1) *If B'/A' is a left H^*-coring Galois extension in T/R, ${}_{B'}B'\otimes_{A'}T_T|{}_{B'}T_T$, $T^{B'}=B$, and $T^{A'}=A$, then A/B is a right H-ring Galois extension (in A/R) such that ${}_T T\otimes_B A_A|{}_T T_A$.*

(2) *If A/B is a right H-ring Galois extension (in A/R), ${}_T T\otimes_B A_A|{}_T T_A$, $T^A=A'$, and $T^B=B'$, and if the ring homomorphism $A\to T$ is an R-ring homomorphism, then B'/A' is a left H^*-coring Galois extension in T/R such that ${}_{B'}B'\otimes_{A'}T_T|{}_{B'}T_T$.*

$$\begin{tikzcd}
 & & T \arrow[dl, dash] \arrow[dr, dash] & & \\
 & B' \arrow[dl, dash] & & A \arrow[dl, dash] \arrow[dr, dash] & \\
A' \arrow[rr, dotted, dash] & & R & & B
\end{tikzcd}$$

Proof. (1) By Theorem 1.4 (2), it suffices to prove that $[x]1_H=x$, and $[x](hk)=[[x]h]k$ for all $x\in A$, h, $k\in H$. Since $h^*(1_{B'})=\varepsilon(h^*)1_T=\langle h^*, 1_H\rangle 1_T$ for all $h^*\in H^*$, $1_{B'}\otimes 1_T$ corresponds to $1_H\otimes 1_T$ under the isomorphism

$$B'\otimes_{A'}T\xrightarrow{\sim}(\mathrm{Hom}({}_RH^*, {}_RT)\xrightarrow{\sim})\,H\otimes T.$$

Hence $[x]1_H=1_{B'}x1_T=x$. For any $h^*\in H^*$, let $\sum x_i\otimes a_i$, $\sum x'_j\otimes a'_j$, and $\sum x''_s\otimes a''_s$ be elements of $T\otimes_B A$ corresponding to $1_T\otimes h^*$, $1_T\otimes h^*_{(1)}$, $1_T\otimes h^*_{(2)}$, respectively. Then $(2)_1$ (in (#)) implies that

$$\sum_i x_i\,x\,y\,a_i=\sum_{(h^*)}\Big(\sum_s x''_s\,x\,a''_s\Big)\Big(\sum_j x'_j\,y\,a'_j\Big),$$

and so

$$\sum_i x_i\,x\otimes a_i=\sum_{(h^*)}\sum_{s,j} x''_s\,x\,a''_s\,x'_j\otimes a'_j \quad\text{in } T\otimes_B A.$$

Then, for any $h\in H$,

$$\sum_i x_i\,x\,[a_i]\,h=\sum_{(h^*)}\sum_{s,j} x''_s\,x\,a''_s\,x'_j\,[a'_j]\,h.$$

Noting that $[a_i]\,h\in A$ (Theorem 1.4 (2)) and

$$\sum_j x'_j\,[a'_j]\,h=\langle h^*_{(1)}, h\rangle\,1_T\in\bar{R}\subseteq A,$$

we have

$$\sum_i x_i\otimes[a_i]\,h\cdot c=\sum_{(h^*)}\sum_{s,j} x''_s\otimes a''_s\,x'_j\,[a'_j]\,h\cdot c$$

in $T\otimes_B A$, for all $h\in H$, $c\in A$. Then, as $\mathrm{Hom}({}_BA, {}_BA)\xrightarrow{\sim}H\otimes A$, we have that

$$\sum_i x_i\otimes[a_i]\,g=\sum_{(h^*)}\sum_{s,j} x''_s\otimes a''_s\,x'_j\,[a'_j]\,g$$

holds for all $g\in\mathrm{Hom}({}_BA, {}_BA)$. In particular, for any $f\in\mathrm{Hom}({}_BA, {}_BB)$,

$$\sum_i x_i\otimes 1\otimes[a_i]\,f=\sum_{(h^*)}\sum_{s,j} x''_s\otimes a''_s\,x'_j\otimes[a'_j]\,f$$

in $T\otimes_B T\otimes_B A$. Then, since ${}_BA\,|\,{}_BB$, we have

$$\sum_i x_i\otimes 1\otimes a_i=\sum_{(h^*)}\sum_{s,j} x''_s\otimes a''_s\,x'_j\otimes a'_j,$$

and so

$$\sum_i y\,x_i\,[[a_i]\,h]\,k=\sum_{(h^*)}\sum_{s,j} y\,x''_s\,[a''_s\,x'_j\,[a'_j]\,h]\,k \qquad \text{for all } y\in T,\ h,\ k\in H.$$

But

$$\begin{aligned}\sum_i x_i\,[a_i]\,(hk)&=\langle h^*, hk\rangle\,1_T=\langle [h^*]\,h, k\rangle\,1_T=\sum_{(h^*)}\langle h^*_{(2)}, \langle h^*_{(1)}, h\rangle\,k\rangle\,1_T\\&=\sum_{(h^*)}\sum_s x''_s\,[a''_s]\,(\langle h^*_{(1)}, h\rangle\,k)=\sum_{(h^*)}\sum_s x''_s\,[a''_s\,\langle h^*_{(1)}, h\rangle]\,k\\&=\sum_{(h^*)}\sum_{s,j} x''_s\,[a''_s\,x'_j\,[a'_j]\,h]\,k,\end{aligned}$$

by Lemma 1.6. Thus

$$\sum_i yx_i[a_i](hk)=\sum_i yx_i[[a_i]h]k \quad \text{for all } y\in T, h, k\in H.$$

Since ${}_T T\otimes_B A \xrightarrow{\sim} {}_T T\otimes H^*$, this implies that $y[a](hk)=y[[a]h]k$ for all $y\in T$, h, $k\in H$, and hence $[a](hk)=[[a]h]k$, as desired.

(2) For any h, $k\in H$, $x\in H^*$ there holds

$$\langle x, hk\rangle 1_T=\langle [x]h, k\rangle 1_T=\sum_{(x)}\langle x_{(2)}, \langle x_{(1)}, h\rangle k\rangle 1_T.$$

Let $\sum x'\otimes x''$, $\sum x_1'\otimes x_1''$, and $\sum x_2'\otimes x_2''$ be elements of $A\otimes_B A$, corresponding to $1_A\otimes x$, $1_A\otimes x_{(1)}$, and $1_A\otimes x_{(2)}$ respectively, under the isomorphism

$$A\otimes H^* \xrightarrow{\sim} \mathrm{Hom}(H_R, A_R) \xrightarrow{\sim} A\otimes_B A, \quad \text{where } \Delta(x)=\sum_{(x)} x_{(2)}\otimes x_{(1)}.$$

Then the last equality implies that

$$\begin{aligned}(\langle x, hk\rangle 1_T=)\sum_{(x)} x'[x''](hk)&=\sum_{(x)} x'[[x'']h]k=\sum_{(x)}\sum x_2'[x_2''\langle x_{(1)}, h\rangle]k\\ &=\sum_{(x)}\sum\sum x_2'[x_2''x_1'[x_1'']h]k,\end{aligned}$$

and so

$$\sum_{(x)} x'\otimes[x'']h=\sum_{(x)}\sum\sum x_2'\otimes x_2''x_1'[x_1'']h \quad \text{in } T\otimes_B A,$$

because $T\otimes_B A \xrightarrow{\sim} \mathrm{Hom}(H_R, T_R)$. Then, as in the proof of (1), we have

$$\sum_{(x)} x'\otimes 1\otimes x''=\sum_{(x)}\sum\sum x_2'\otimes x_2''x_1'\otimes x_1'' \quad \text{in } T\otimes_B T\otimes_B T.$$

Hence, for any y, $z\in B'$,

$$x(yz)=\sum_{(x)} x'yzx''=\sum_{(x)}\sum\sum x_2'yx_2''x_1'zx_1''=\sum_{(x)} x_{(2)}(y)x_{(1)}(z).$$

Furthermore,

$$x(1_{B'})=\sum_{(x)} x'x''=\sum_{(x)} x'[x'']1_H=\langle x, 1_H\rangle 1_T.$$

Combining these with Lemma 1.6, we can complete the proof.

From the above theorem we have the following (which corresponds to Corollary 1.5)

Theorem 1.8. *Let A/B be a right H-ring (resp. coring) Galois extension (in A/R), and let T/A be a ring extension such that ${}_T T\otimes_B A_A |{}_T T_A$ (or equivalently, ${}_A\mathrm{Hom}({}_B A, {}_B T)_T|{}_A T_T$). Assume that $\bar{B}\subseteq A^R$ and*

$$[ab]h=[a]h\cdot b$$

for all $a\in A$, $b\in B$, $h\in H$. Put $T^A=A'$ and $T^B=B'$. Then B'/A' is a left H^*-coring (resp. ring) Galois extension, $A'\subseteq T^R$, ${}_{B'}B'\otimes_{A'} T_T|{}_{B'}T_T$, and

$$h^*(a'b')=a'\cdot h^*(b')$$

for all $a'\in A'$, $b'\in B'$, $h^*\in H^*$.

In the following, we will state how to translate the preceding theorems into theorems concerning representation modules.

Let M be a right T^*-module over a ring T^*, and set $T=\mathrm{End}(M_{T^*})$ (which acts on the left side of M). Let B'/A', A/B be ring extensions such that $B'_{A'}|A'_{A'}$ and ${}_BA|{}_BB$, and take two ring homomorphisms $B'\to T$ and $A\to T$, or equivalently, two bimodules ${}_{B'}M_{T^*}$ and ${}_AM_{T^*}$. Furthermore, take an (R, T^*)-bimodule ${}_RM_{T^*}$. Then, from a right T^*-isomorphism

$$\mathrm{Hom}({}_BA, {}_BM)_{T^*} \xrightarrow{\sim} H\otimes M_{T^*},$$

we have a right T-isomorphism $\mathrm{Hom}({}_BA, {}_BT)_T \xrightarrow{\sim} H\otimes T_T$ by applying $\mathrm{Hom}(M_{T^*}, -)$. The converse is also done by applying $\otimes_T M$. Similarly, from a right T^*-isomorphism

$$B'\otimes_{A'} M_{T^*} \xrightarrow{\sim} \mathrm{Hom}({}_RH^*, {}_RM)_{T^*},$$

it follows a right T-isomorphism $B'\otimes_{A'} T_T \xrightarrow{\sim} \mathrm{Hom}({}_RH^*, {}_RT)_T$, and conversely. Evidently

$${}_{B'}B'\otimes_{A'} M_{T^*}|{}_{B'}M_{T^*}$$

is equivalent to that ${}_{B'}B'\otimes_{A'} T_T|{}_{B'}T_T$. Similarly

$${}_A\mathrm{Hom}({}_BA, {}_BM)_{T^*}|{}_AM_{T^*}$$

is equivalent to that ${}_A\mathrm{Hom}({}_BA, {}_BT)_T|{}_AM_T$. In the following we assume that

$$a(a'm)=a'(am), \quad \text{and} \quad b(b'm)=b'(bm)$$

for all $a\in A$, $a'\in A'$, $m\in M$, $b\in B$, $b'\in B'$. Then there is a canonical right T^*-homomorphism

$$D\colon B'\otimes_{A'} M_{T^*}\to \mathrm{Hom}({}_BA, {}_BM)_{T^*}, \qquad b'\otimes m\mapsto (a\to b'am),$$

where $b'\in B'$, $a\in A$, $m\in M$. And D is an isomorphism if and only if δ: $B'\otimes_{A'} T_T\to \mathrm{Hom}({}_BA, {}_BT)_T$, $b'\otimes t\mapsto (a\to b'at)$ is an isomorphism. Assume that the following diagram is a commutative one of right T^*-isomorphisms.

$$\begin{tikzcd} B'\otimes_{A'} M \arrow[r, "\sim"] \arrow[d, "\wr"'] \arrow[dr, phantom, "\cdot"] & \mathrm{Hom}({}_RH^*, {}_RM) \arrow[d, "\wr"] \\ \mathrm{Hom}({}_BA, {}_BM) & H\otimes M. \arrow[l, "\sim"'] \end{tikzcd}$$

Then we have a right T-square of Galois structures such as $(*)$, by applying $\mathrm{Hom}(M_{T^*}, -)$, and conversely.

For example, let B'/A' be a finite G-Galois extension with a finite group G of automorphisms of B' (cf. [7]). Then, as is well known,

$$B'\otimes_{A'}(A'G) \xrightarrow{\sim} \mathrm{End}(B'_{A'}),\ b'\otimes a'\sigma \quad (x\to b'a'\cdot\sigma(x))\ (b'\in B', a'\in A', \sigma\in G),$$

where $A'G$ denote the group ring of G over A'. But, on the left side, we can replace A' by $R=B'^{A'}$. Then $H^*=RG$ is a skew group ring (or, a trivial crossed product) defined by $\sigma x = \sigma(x)\sigma$ ($\sigma \in G$, $x \in R$). Then B'/A' is a left H^*-ring Galois extension such that $h^*(a'b') = a' \cdot h^*(b')$ ($h^* \in H^*$, $a' \in A'$, $b' \in B'$). As is well known, ${}_{A'}B'|_{A'}A'$, and so ${}_{B'}B' \otimes_{A'} B'|_{B'}B'$. If we put $T = \mathrm{End}(B')$ (the ring of all endomorphisms of B', acting on the left side of B'), then ${}_{B'}B' \otimes_{A'} T_T|_{B'}T_T$, where $B' \to T$, $b' \mapsto b'_1 = (x \to b'x)$ ($x, b' \in B'$). Evidently $\bar{A}' \subseteq T^R$, where $R \to T$, $r \mapsto (x \to rx)$ ($r \in R$, $x \in B'$). Thus we can apply Theorem 1.8 for this case. (Note that we put $\sigma(b'_1) = \sigma(b')_1$ and identify b' with b'_1.)

In the sequel of this paper, R is a commutative ring, and we treat R-algebras and R-homomorphisms. For any (R, R)-bimodule X, we always assume that $rx = xr$ for all $x \in X$, $r \in R$.

Here we rewrite the definition of a left H^*-coalgebra Galois extension. If a left H^*-weak Galois extension B'/A' satisfies the following condition, B'/A' is said to be a left H^*-coalgebra Galois extension:

$$h^*(1_{B'}) = \varepsilon(h^*)\, 1_{B'}, \qquad h^*(xy) = \sum_{(h^*)} h^*_{(2)}(x)\, h^*_{(1)}(y) \quad (h^* \in H^*,\ x, y \in B').$$

In this case, as we remarked before, $h^*(a'b') = a' \cdot h^*(b')$ ($h^* \in H^*$, $a' \in A'$, $b' \in B'$) follows from the above condition.

Then Theorem 1.8 yields the following

Theorem 1.9. *Let M be a right T^*-module over an R-algebra T^*, and let A, B' be R-algebras, and B, A' R-subalgebras of A and B' respectively. Put $\mathrm{End}(M_{T^*}) = T$, which acts on the left side of M, and take two R-algebra homomorphisms $A \to T$ and $B' \to T$.*

(1) *Let ${}_A\mathrm{Hom}({}_BA, {}_BM)_{T^*}|_AM_{T^*}$, and let A/B be a right H-algebra (resp. coalgebra) Galois extension such that $[ab]h = [a]h \cdot b$ for all $a \in A$, $b \in B$, $h \in H$. Assume that $A' = T^A$ $(= \mathrm{End}({}_AM_{T^*}))$ and $B' = T^B$. Then B'/A' is a left H^*-coalgebra (resp. algebra) Galois extension, ${}_{B'}B' \otimes M_{T^*}|_{B'}M_{T^*}$, and $h^*(a'b') = a' \cdot h^*(b')$ for all $a' \in A'$, $b' \in B'$, $h^* \in H^*$.*

(2) *Let ${}_{B'}B' \otimes_{A'} M_{T^*}|_{B'}M_{T^*}$, and let B'/A' be a left H^*-algebra (resp. coalgebra) Galois extension such that $h^*(a'b') = a' \cdot h^*(b')$ for all $a' \in A'$, $b' \in B'$, $h^* \in H^*$. Assume that $B = T^{B'}$ $(= \mathrm{End}({}_{B'}M_{T^*}))$ and $A = T^{A'}$. Then A/B is a right H-coalgebra (resp. algebra) Galois extension, ${}_A\mathrm{Hom}({}_BA, {}_BM)_{T^*}|_AM_{T^*}$, and $[ba]h = [b]h \cdot a$ for all $b \in B$, $a \in A$, $h \in H$.*

Remark. We may put $T^* = Z$ (the ring of integers). On the other hand, if we put $M = T^*_{T^*}$ we get a centralizer theorem in T^*, because $T^* \xrightarrow{\sim} \mathrm{End}(T^*_{T^*})$.

The above theorem yields a centralizer theorem concerning Hopf Galois extensions (cf. [17, 32, 30]). Therefore this may be considered as a generalization of [20; Theorem 2.12, Theorem 1.13], by [17; p. 682, example 2].

To see the case treated in [30], let A/R be an R-Azumaya algebra, and let Z be a commutative R-subalgebra such that $A^Z = Z$. Assume ${}_ZA|_ZZ$. Then, since Z is commutative, ${}_ZZ|_ZA$ holds. As is well known, ${}_AA \otimes_R A_A \sim {}_AA_A$ (similar) (cf. [2]), and so ${}_ZZ \otimes_R A_A \sim {}_ZA \otimes_R A_A \sim {}_ZA_A$. Then ${}_AA_A \sim {}_AA \otimes_Z Z \otimes_R A_A \sim {}_AA \otimes_Z A_A$. Therefore Theorem 1.9 (2) is applicable for Z/R or A/Z in A/R, when they are left H^*-algebra or coalgebra Galois extensions satisfying an additional condition. (Put $M = A_A$.) Since ${}_ZZ \otimes_R A_A \sim {}_ZA_A$ implies that $Z \otimes A \xrightarrow{\sim} \mathrm{End}({}_ZA)$,

$z\otimes a\mapsto(x\to zxa)$ we have ${}_A\mathrm{End}({}_ZA)_A\sim{}_AA_A$, where $[x](afa')=[xa]f\cdot a'$ $(x, a, a'\in A,\ f\in\mathrm{End}({}_ZA))$. Therefore Theorem 1.9 (1) is applicable for A/Z in A/R. (Put $M=A_A$.) Evidently A/R itself is a left A^{op}-algebra Galois extension, and A/Z is a Z-algebra Galois extension, where A^{op} is the opposite algebra of A.

Let U/R be an R-algebra such that $U_R\sim R_R$. Then $U^*=\mathrm{Hom}(U_R, R_R)$ has an R-coalgebra structure induced by the R-algebra structure of U. As is easily seen, U/R is a left U^{*t}-coalgebra Galois extension, where U^{*t} has the transposed comultiplication of the one of U^*.

Needless to say we have an example of a coalgebra Galois extension: Let L/K be a finite Galois extension field with Galois group G. Let F be a subgroup of G, and let S be the intermediate field corresponding to F. Let $G=\sigma_1F\cup\ldots\cup\sigma_rF$ be the coset decomposition. Then, as is well known, ${}_LL\otimes_K(\oplus_iK\sigma_i)\xrightarrow{\sim}{}_L\mathrm{Hom}(S_K, L_K)$ canonically. Hence S/K is a left $\oplus_iK\sigma_i$-coalgebra Galois extension in L/K, where $\oplus_iK\sigma_i$ is an R-subcoalgebra of KG.

2. Some Remarks on Hopf Galois Extensions

In this section, H is taken to be an R-Hopf algebra such that H is finitely generated and projective as an R-module. We put $P(H)=\{h\in H\mid kh=\varepsilon(k)h$ for all $k\in H\}$, where ε is the counit map of H. The following proposition corresponds to [20; Proposition 2.4 (1)].

Proposition 2.1. *Let A/B be a left H-Hopf Galois extension (i.e. a left H-algebra Galois extension in our sense, such that $h(1_A)=\varepsilon(h)1_A$ and $h(xy)=\sum_{(h)}h_{(1)}(x)h_{(2)}(y)$ for all $h\in H$, x, $y\in A$). Take an R-algebra homomorphism $A\to T$, and put $T^B=B^*$. Then the following are equivalent:*

(1) ${}_AA\otimes_BT_T|{}_AT_T$.

(2) *There are $b_i^*\in B^*$, $g_i\in\mathrm{Hom}({}_RP(H), {}_RB^*)$ $(i=1,\ldots,n)$ such that $\sum_i g_i(h)yb_i^*=h(y)1_T$ for all $h\in P(H)$, $y\in A$.*

Proof. We know that (1) is equivalent to that ${}_T\mathrm{Hom}(A_B, T_B)_A|{}_TT_A$. And it is evident that

$$B^*\xrightarrow{\sim}\mathrm{Hom}({}_AA_B, {}_AT_B)\xrightarrow{\sim}\mathrm{Hom}({}_TT_A, {}_T\mathrm{Hom}(A_B, T_B)_A),$$

by

$$b^*\mapsto(a\to ab^*)\mapsto(t\to(a\to tab^*)),$$

where $b^*\in B^*$, $a\in A$, $t\in T$. On the other hand, by [17; (1.7) Theorem], there is an isomorphism

$$(\#\,\#)\qquad P(H)\otimes A\xrightarrow{\sim}\mathrm{Hom}(A_B, B_B),\quad h\otimes a\mapsto(x\to h(ax)),$$

where $h\in P(H)$, a, $x\in A$, and then

$$\begin{aligned}\mathrm{Hom}({}_RP(H), {}_RB^*)&\xrightarrow{\sim}\mathrm{Hom}({}_RP(H), {}_R\mathrm{Hom}({}_BA_A, {}_BT_A))\\&\xrightarrow{\sim}\mathrm{Hom}({}_BP(H)\otimes A_A, {}_BT_A)\\&\xrightarrow{\sim}\mathrm{Hom}({}_TT\otimes_B(P(H)\otimes A)_A, {}_TT_A)\\&\xrightarrow{\sim}\mathrm{Hom}({}_TT\otimes_B\mathrm{Hom}(A_B, B_B)_A, {}_TT_A)\\&\xrightarrow{\sim}\mathrm{Hom}({}_T\mathrm{Hom}(A_B, T_B)_A, {}_TT_A).\end{aligned}$$

In fact, for any $g \in \operatorname{Hom}({}_R P(H), {}_R B^*)$,

$$\begin{aligned} g = (h \to g(h)) &\mapsto (h \to (a \to g(h)\, a)) \\ &\mapsto (h \otimes a \to g(h)\, a) \\ &\mapsto (t \otimes (h \otimes a) \to t \cdot g(h)\, a) \\ &\mapsto (t \otimes (a' \to h(a a')) \to t \cdot g(h)\, a) \\ &\mapsto ((a' \to t \cdot h(a a')) \to t \cdot g(h)\, a), \end{aligned}$$

which is the image of g. Then we can easily see the equivalence (1)$\Leftrightarrow$(2).

Remark. If we put $T = A$, then the above yields a characterization of that ${}_A A \otimes_B A_A | {}_A A_A$ (i.e. separable in the sense of Hirata).

Proposition 2.2. (Cf. Yokogawa [33].) *Under the same assumptions as in Proposition 2.1, let* ${}_A A \otimes_B T_T | {}_A T_T$, *and put*

$$W = \{v \in \operatorname{Hom}({}_R H, {}_R T) \mid v(h)\, a = \sum_{(h)} h_{(1)}(a)\, v(h_{(2)}) \text{ for all } h \in H,\ a \in A\}.$$

Then we have an isomorphism

$$W \otimes P(H) \xrightarrow{\sim} B^*, \qquad v \otimes h \mapsto v(h) \quad (v \in W,\ h \in P(H)).$$

Proof. Under the isomorphism $A \otimes_B T \xrightarrow{\sim} \operatorname{Hom}({}_R H, {}_R T)$, $a \otimes t \mapsto (h \to h(a)\, t)$, $(A \otimes_B T)^A$ corresponds to W. By [17], we have (# #): $P(H) \otimes A \xrightarrow{\sim} \operatorname{Hom}(A_B, B_B)$, $h \otimes a \mapsto (x \to h(ax))$. Taking left-right dual, we have an isomorphism

$$P(H) \otimes A \xrightarrow{\sim} \operatorname{Hom}({}_B A, {}_B B), \qquad h \otimes a \mapsto (x \to h(x a)),$$

where $h \in P(H)$, x, $a \in A$. (Apply $\operatorname{Hom}(-, {}_B B)$ to (# #), and multiply $P(H) \otimes$ on the left side.) Therefore

$$P(H) \otimes A \otimes_B T \xrightarrow{\sim} \operatorname{Hom}({}_B A, {}_B B) \otimes_B T \xrightarrow{\sim} \operatorname{Hom}({}_B A, {}_B T).$$

Since $(A \otimes_B T)^A \xrightarrow{\sim} \operatorname{Hom}({}_A T_T, {}_A A \otimes_B T_T)$ and ${}_R P(H) | {}_R R$ (cf. [26]), we have a sequence of isomorphisms:

$$\begin{aligned} (A \otimes_B T)^A \otimes P(H) &\xrightarrow{\sim} \operatorname{Hom}({}_A T_T, {}_A A \otimes_B T_T) \otimes P(H) \\ &\xrightarrow{\sim} \operatorname{Hom}({}_A T_T, {}_A P(H) \otimes A \otimes_B T_T) \\ &\xrightarrow{\sim} \operatorname{Hom}({}_A T_T, {}_A \operatorname{Hom}({}_B A, {}_B T)_T) \\ &\xrightarrow{\sim} \operatorname{Hom}({}_B A_A, {}_B T_A) \xrightarrow{\sim} B^*. \end{aligned}$$

In fact, $(\sum_i a_i \otimes t_i) \otimes h$ (in $(A \otimes_B T)^A \otimes P(H)$) corresponds to $\sum_i h(a_i)\, t_i$ (in B^*).

Proposition 2.3. *If A/B is a left H-Hopf Galois extension then* $\operatorname{End}(A_B)$ *is a projective Frobenius extension of B (= the left multiplications by elements of B).*

Proof. Put $\Omega = \operatorname{End}(A_B)$. By (# #), we have an isomorphism

$A \otimes_B (P(H) \otimes A) \xrightarrow{\sim} A \otimes_B \operatorname{Hom}(A_B, B_B) \xrightarrow{\sim} \Omega$, and so $\Omega_B | B_B$. We identify Ω with $A \otimes_B (P(H) \otimes A)$. Noting that $P(H)$ is a projective R-module of rank 1 (cf.

[26]), we have a sequence of isomorphisms

$$\begin{aligned}\Omega &\xrightarrow{\sim} \operatorname{Hom}(A\otimes P(H)_B, A\otimes P(H)_B)\\ &\xrightarrow{\sim} \operatorname{Hom}(A\otimes_B(P(H)\otimes A)_A, A\otimes P(H)_A)\\ &\xrightarrow{\sim} \operatorname{Hom}(A\otimes(P(H)\otimes A)_A, \operatorname{Hom}(A_B, B_B)_A)\\ &\xrightarrow{\sim} \operatorname{Hom}(A\otimes(P(H)\otimes A)_B, B_B)=\operatorname{Hom}(\Omega_B, B_B).\end{aligned}$$

In fact, for any $f\in\Omega$,

$$\begin{aligned} f&\mapsto(x\otimes h\to f(x)\otimes h)\\ &\mapsto(x\otimes(h\otimes y)\to f(x)\,y\otimes h)\\ &\mapsto(x\otimes(h\otimes y)\to(a\to h(f(x)\,y\,a)))\\ &\mapsto(x\otimes(h\otimes y)\to h(f(x)\,y)).\end{aligned}$$

Now we define

$$F:\Omega\to B,\text{ by } F(x\otimes(h\otimes y))=h(x\,y)\quad (h\in P(H),\ x,\ y\in A).$$

Then F is a (B, B)-homomorphism, and

$$F(f\cdot(x\otimes(h\otimes y)))=F(f(x)\,h\,y)=h(f(x)\,y)\quad\text{for any } f, x\otimes(h\otimes y)\in\Omega.$$

Thus Ω/B is a Frobenius extension with a Frobenius homomorphism F.

Let A/B be a left H-Hopf Galois extension. Then, from the proof of Theorem 2.3, we know that

$$A\otimes P(H)\xrightarrow{\sim}\operatorname{Hom}({}_BA, {}_BB),\ a\otimes h\mapsto(x\to h(x\,a))\quad (a, x\in A,\ h\in P(H)).$$

Therefore

$$(A\otimes P(H))\otimes_B A\xrightarrow{\sim}\operatorname{End}({}_BA),\qquad (a\otimes h)\otimes a'\mapsto(x\to h(x\,a)\,a'),$$

where $a, a', x\in A$, $h\in P(H)$. Combining this with the isomorphism

$$((A\otimes P(H))\otimes_B A\xrightarrow{\sim})\ A\otimes_B(P(H)\otimes A)\xrightarrow{\sim}\operatorname{End}(A_B),$$

we get an isomorphism

$$\theta\colon \operatorname{End}(A_B)\xrightarrow{\sim}\operatorname{End}({}_BA).$$

Then we can see the following

Proposition 2.4. *Let A/B be a left H-Hopf Galois extension. Then θ is an R-algebra isomorphism such that $\theta(x\to a\,x)=(x\to x\,a)$ $(x, a\in A)$. Therefore if*

$$\sum_i r_i\otimes(h_i\otimes l_i)\leftrightarrow \mathrm{id}_A\quad\textit{under}\quad A\otimes_B(P(H)\otimes A)\xrightarrow{\sim}\operatorname{End}(A_B),$$

Then

$$\sum_i (r_i\otimes h_i)\otimes l_i\leftrightarrow \mathrm{id}_A\quad\textit{under}\quad (A\otimes P(H))\otimes_B A\xrightarrow{\sim}\operatorname{End}({}_BA).$$

Furthermore, A/B is separable (i.e. ${}_AA_A|{}_AA\otimes_B A_A$) if and only if

$$\sum_i r_i\cdot f(h_i)\,l_i=1 \quad \textit{for some } f\in\mathrm{Hom}({}_RP(H),{}_R(A^B)).$$

To prove the above we need the following

Lemma 2.5. *For any $h, k\in P(H)$, $h^*\in H^*$, $h\langle h^*, k\rangle=k\langle h^*, h\rangle$ holds.*

Proof. Noting that $P(H)$ is a projective R-module of rank 1, we can easily prove this by localization.

Corollary 2.6. *If A/B is a left H-Hopf Galois extension, then $h(a)\otimes k=k(a)\otimes h$ in $A\otimes H$, for all $h, k\in P(H)$, $a\in A$.*

Proof. From the lemma, it follows that $h(a)\langle h^*, k\rangle=k(a)\langle h^*, h\rangle$ for all $h^*\in H^*$. As ${}_RH|{}_RR$, this implies that $h(a)\otimes k=k(a)\otimes h$ in $A\otimes H$.

Proof of Theorem 2.4. Using the above corollary, we can easily see that θ is an R-algebra isomorphism such that $\theta(x\to ax)=(x\to xa)$ $(x, a\in A)$. We put $\mathrm{End}(A_B)=\Omega$. Then, as ${}_A\Omega_A\xrightarrow{\sim}{}_AA\otimes_B(P(H)\otimes A)_A$, A/B is separable if and only if ${}_AA_A|{}_A\Omega_A$. But A is a subring of Ω, and hence ${}_AA_A|{}_A\Omega_A$ implies that ${}_AA_A$ is a direct summand of ${}_A\Omega_A$. (In fact, there are $\delta_i\in\Omega^A$, f_i: ${}_A\Omega_A\to{}_AA_A$ such that $\sum_i f_i(\delta_i)=1$. Then $(\omega\to\sum_i f_i(\omega\delta_i))$ $(\omega\in\Omega)$ is a projection from ${}_A\Omega_A$ to ${}_AA_A$.) Now

$$\begin{aligned}\mathrm{Hom}({}_RP(H),{}_R(A^B))&\xrightarrow{\sim}\mathrm{Hom}({}_BP(H)\otimes A_A,{}_BA_A)\\&\xrightarrow{\sim}\mathrm{Hom}({}_AA\otimes_B(P(H)\otimes A)_A,{}_AA_A),\end{aligned}$$

and $(x\to ax)\mapsto\sum_i ar_i\otimes(h_i\otimes l_i)$ under the isomorphism $\Omega\xrightarrow{\sim}A\otimes_B(P(H)\otimes A)$. Therefore ${}_AA_A|{}_A\Omega_A$ if and only if $\sum_i r_i\cdot f(h_i)\,l_i=1_A$ for some $f\in\mathrm{Hom}({}_RP(H),{}_R(A^B))$.

References

1. Auslander, M., Goldman, O.: Maximal orders. Trans. Amer. Math. Soc. **97**, 1–24 (1960)
2. Auslander, M., Goldman, O.: The Brauer group of a commutative ring. Trans. Amer. Math. Soc. **97**, 367–409 (1960)
3. Azumaya, G.: Completely faithful modules and self injective rings. Nagoya Math. J. **27**, 697–708 (1966)
4. Bourbaki, N.: Algebre Commutative. Paris: Hermann 1961
5. Chase, S.U., Harrison, D.K., Rosenberg, A.: Galois theory and Galois cohomology of commutative rings. Mem. Amer. Math. Soc. **52**, 15–33 (1965)
6. Chase, A.U., Sweedler, M.E.: Hopf algebras and Galois theory. Lecture Notes in Math. **97**. Berlin-Heidelberg-New York: Springer 1969
7. DeMeyer, F.R.: Some notes on the general Galois theory of rings. Osaka J. Math. **2**, 117–127 (1965)
8. Endo, S.: Completely faithful modules and quasi-Frobenius algebras. J. Math. Soc. Japan **19**, 437–456 (1967)
9. Hattori, A.: Semisimple algebras over a commutative ring. J. Math. Soc. Japan **15**, 404–419 (1963)
10. Hirata, K., Sugano, K.: On semisimple extensions and separable extensions over non-commutative rings. J. Math. Soc. Japan **18**, 360–373 (1966)

11. Hirata, K.: Some types of separable extensions of rings. Nagoya Math. J. **33**, 107–115 (1968)
12. Hirata, K.: Separable extensions and centralizers of rings. Nagoya Math. J. **35**, 31–45 (1969)
13. Hochschild, G.: Note on relative homological algebra. Trans. Amer. Math. Soc. **82**, 246–269 (1956)
14. Kanzaki, T.: On commutor rings and Galois theory of separable algebras. Osaka J. Math. **2**, 137–145 (1965)
15. Kanzaki, T.: On generalized crossed product and Brauer group. Osaka. J. Math. **5**, 175–188 (1968)
16. Kasch, F.: Projective Frobenius-Erweiterungen. Sitzungsber. Heidelb. Akad. Wiss. Math.-Natur. Kl. 1960/1961, 89–109
17. Kreimer, H.F., Takeuchi, M.: Hopf algebras and Galois extensions of an algebra. Indiana Univ. Math. J. **30**, 675–692 (1981)
18. Larson, R.G., Sweedler, M.E.: An associative orthogonal bilinear form for Hopf algebras. Amer. J. Math. **91**, 75–94 (1967)
19. Miyashita, Y.: Finite outer Galois theory of non-commutative rings. J. Fac. Sci. Hokkaido Univ. Ser. I **19**, 114–134 (1966)
20. Miyashita, Y.: On Galois extensions and crossed products. J. Fac. Sci. Hokkaido Univ. Ser. I **21**, 97–121 (1970)
21. Morita, K.: Duality for modules and its application to the theory of rings with minimum condition. Sci. Rep. Tokyo Kyoiku Daigaku Sect. A **6**, 83–142 (1958)
22. Morita, K.: Adjoint pairs of functors and Frobenius extensions. Sci. Rep. Tokyo Kyoiku Daigaku Sect. A **9**, 40–71 (1965)
23. Morita, K.: The endomorphism ring theorem for Frobenius extensions. Math. Z. **102**, 385–404 (1967)
24. Müller, B.: Quasi-Frobenius Erweiterungen. Math. Z. **85**, 345–368 (1964)
25. Nakayama, T.: On a generalized notion of Galois extensions of a ring. Osaka J. Math. **15**, 11–13 (1963)
26. Pareigis, B.: When Hopf algebras are Frobenius algebras. J. Algebra **18**, 588–596 (1971)
27. Sugano, K.: Note on semi-simple extensions and separable extensions. Osaka J. Math. **4**, 265–270 (1967)
28. Sweedler, M.E.: Hopf algebras. New York: Benjamin 1969
29. Ulbrich, K.H.: Vollgraduierte Algebren. Abh. Math. Sem. Univ. Hamburg **51**, 136–148 (1981)
30. Ulbrich, K.H.: Galoiserweiterungen von nicht-kommutativen Ringen. Comm. Algebra **10**(6), 655–672 (1982)
31. Yokogawa, K.: On $S \otimes_R S$-module structure of S/R-Azumaya algebras. Osaka J. Math. **12**, 673–690 (1975)
32. Yokogawa, K.: Non-commutative Hopf Galois extensions. Osaka J. Math. **18**, 63–73 (1981)
33. Yokagawa, K.: A pair of subalgebras in an Azumaya algebra. To appear in Osaka J. Math.

Received May 9, 1983

Math. Z. 187, 145–150 (1984)

Mathematische Zeitschrift

Unfolding Bifurcations of an Elliptic Boundary Value Problem

H.-O. Peitgen[1] and A.J. Tromba[2]

[1] Forschungsschwerpunkt “Dynamische Systeme”, FB Mathematik, Universität Bremen, D-2800 Bremen 33, Federal Republic of Germany
[2] Department of Mathematics, University of California, Santa Cruz, CA95064, USA and Max-Planck-Institut für Mathematik, D-5300 Bonn-Beuel, Federal Republic of Germany

The object of this note is to show that if one considers the two parameter family of elliptic equations on a domain $\Omega \subset \mathbb{R}^n$ with say C^∞ boundary $\partial\Omega$, given by

$$\text{(1)} \qquad \begin{aligned} \Delta u + \lambda f(u) &= \mu g \quad \text{in } \Omega \\ u &= \mu\phi \quad \text{on } \partial\Omega, \end{aligned}$$

$(\lambda, \mu) \in \mathbb{R}^2$, f, g: $\mathbb{R} \to \mathbb{R}$ smooth, $f(0)=0$, and $\phi \not\equiv 0$ smooth, then under some mild restrictions on f a bifurcation around the solution $u \equiv 0$ near (λ_0, μ_0), where λ_0 is a simple eigenvalue of the linearized equation

$$\text{(2)} \qquad \begin{aligned} \Delta w + \lambda_0 f'(0) w &= 0 \quad \text{in } \Omega \\ w &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

and $\mu_0 = 0$, can be described by the cusp-catastrophe of Thom.

We begin by developing some analytical tools necessary to prove the main theoretical result. For this we follow [7].

Let G be a Banach space with a continuous symmetric bilinear pairing $\langle\ ,\ \rangle$ with the property that $\langle w, w\rangle \geqq 0$ with equality holding only when $w=0$. Let $U \subset G$ be open with E: $U \to \mathbb{R}$ a C^r, $r \geqq 1$, real valued map. Then E is called C^r-*acceptable* with respect to $\langle\ ,\ \rangle$ if for each $p \in U$ there is a C^{r-1} vector field X: $U \to G$ such that the differential of E, $\mathrm{d}E$ satisfies

$$\text{(3)} \qquad \mathrm{d}E(p)[h] = \langle X(p), h\rangle$$

for all $h \in G$ and $p \in U$. X is the gradient of E with respect to $\langle\ ,\ \rangle$. If X exists it is necessarily unique. The following notion will be crucial.

Definition. Let E: $U \to \mathbb{R}$ be a C^r-acceptable map. A critical point $q \in U$ is *non-degenerate* if the Frechét derivative of the *associated vector field* X at q, say $\mathrm{D}X(q)$: $G \to G$ is an isomorphism.

Now let $E_{\lambda,\mu}$ be a two parameter family of maps $((\lambda,\mu)\in\mathbb{R}^2)$ with $(\lambda,\mu,p)\to E_{\lambda,\mu}(p)$ C^r, $r>1$. If each $E_{\lambda,\mu}$ is C^r-acceptable and $(\lambda,\mu,p)\to X_{\lambda,\mu}(p)$ is C^{r-1}, we say that $E_{\lambda,\mu}$ is a C^r-acceptable family.

The following normal form theorem from [1] is crucial for us:

Theorem 1. *Suppose that $E_{\lambda,\mu}$: $U\to\mathbb{R}$ is a C^∞-acceptable family with $q\in U$ a critical point of E_{λ_0,μ_0}. Assume further that the kernel J of the Hessian $D^2E_{\lambda_0,\mu_0}(q)$: $G\times G\to\mathbb{R}$ is a one dimensional subspace spanned by $e\in G$. Let $J^\perp$ denote a complement of J in G; $G=J\oplus J^\perp$. Let E denote E_{λ_0,μ_0} and assume further that the following partial derivatives all evaluated at (λ_0,μ_0,q) satisfy the conditions*

(i) $$\frac{\partial^3 E}{\partial e^3}=0;$$

(ii) $$\frac{\partial^4 E}{\partial e^4}\neq 0;$$

(iii) $$\frac{\partial^3 E_{\lambda,\mu}}{\partial\lambda\,\partial e^2}\neq 0;$$

(iv) $$\frac{\partial^2 E_{\lambda,\mu}}{\partial\mu\,\partial e}\neq 0.$$

After a suitable translation we may assume that $q=0$. Then there are neighborhoods W of q in U and V of (λ_0,μ_0) in $\mathbb{R}^2$ and an origin (in G) preserving diffeomorphism

$$\Psi\colon V\times W\to\mathbb{R}^2\times G=\mathbb{R}^2\times J\times J^\perp$$

such that $E(\lambda,\mu,p)=E_{\lambda,\mu}(p)$ satisfies

(a) $E\circ\Psi(\lambda,\mu,t,p)=\frac{1}{4}t^4-\zeta t^2+\gamma t+\frac{1}{2}D^2E_{\lambda,\mu}(q)[p,p]+E_{\lambda,\mu}(q)$,
where ζ and γ depend linearly on λ and μ and where $D^2E_{\lambda,\mu}(q)[p,p]$ denotes the second derivative of $E_{\lambda,\mu}$ along $J^\perp$ evaluated at q. Thus, for fixed (λ,μ)

$$p\mapsto D^2E_{\lambda,\mu}(q)[p,p]$$

is a quadratic form on G, which is "non-degenerate" in the weak sense that if for all h, $D^2E_{\lambda,\mu}(q)[p,h]=0$, then $p=0$.

(b) *If* $\dfrac{\partial^2 E_{\lambda,\mu}}{\partial\lambda\,\partial e}=0$, $\dfrac{\partial^4 E}{\partial e^4}>0$ *and* $\dfrac{\partial^3 E_{\lambda,\mu}}{\partial\lambda\,\partial e^2}<0$,

then in the above normal form one can take $\zeta=\lambda$ and $\gamma=\mu$, and the diffeomorphism $\Psi=(\Psi_1,\Psi_2,\Psi_3,\Psi_4)$ so that at (λ_0,μ_0,q)

$$\frac{\partial\Psi_2}{\partial\lambda}>0.$$

These last conditions are designed to locate the axis and direction of the cusp and hence the direction of bifurcation. In this last coordinate representation the set of critical points (locally, of course) is the set $(\lambda,\mu,t,0)$ in $\mathbb{R}^2\times J\times J^\perp$ with

$$t^3-\lambda t+\mu=0,$$

i.e. the cusp of Thom.

We now apply these ideas to the study of bifurcation of solutions to (1). Let $s>\frac{n}{2}+2$ and introduce the space

$$\mathscr{H}_\mu=\{u\in H^s(\Omega,\mathbb{R}): u=\mu\phi \text{ on } \partial\Omega\}.$$

This is an affine subspace of the Sobolev space $H^s(\Omega,\mathbb{R})$ of s-times differentiables functions with (distributional) derivatives which are square integrable. Let

$$\mathscr{H}=\bigcup_{\mu\in\mathbb{R}}\mathscr{H}_\mu.$$

Then $\mathscr{H}$ has the structure of a C^∞ trivial vector bundle over $\mathbb{R}$ with trivialization

$$\mathscr{H}_0\times\mathbb{R}\to\mathscr{H} \tag{4}$$

given by $(v,\mu)\to v+\mu\psi$, where ψ is the unique solution of $(\phi\not\equiv 0)$

$$\begin{aligned}\Delta\psi&=0 \quad \text{in } \Omega\\ \psi&=\phi \quad \text{on } \partial\Omega\end{aligned} \tag{5}$$

Consider the energy functional $E_{\lambda,\mu}\colon\mathscr{H}\to\mathbb{R}$ given by

$$E_{\lambda,\mu}(u)=\tfrac{1}{2}\int_\Omega\nabla u\cdot\nabla u\,dx-\lambda\int_\Omega F(u)dx+\mu\int_\Omega g(x)u\,dx, \tag{6}$$

where $F\colon\mathbb{R}\to\mathbb{R}$ is C^∞ and $F'(s)=f(s)$.

Note that the parameter μ in (6) also enters in the space $\mathscr{H}$. This is not uncommon in complicated variational problems (e.g. see [1, 7]).

By using the trivialization (4) and Eq. (5) we may view $E_{\lambda,\mu}$ as a function on $\mathscr{H}_0$ given by

$$\begin{aligned}E_{\lambda,\mu}(v)=\tfrac{1}{2}\int_\Omega\nabla(v+\mu\psi)\cdot\nabla(v+\mu\psi)dx-\lambda\int_\Omega F(v+\mu\psi)dx\\ +\mu\int_\Omega g(x)(v+\mu\psi)dx.\end{aligned} \tag{7}$$

The Euler-Lagrange operator associated to (7) at a critical point v is given by (for fixed λ and μ)

$$v\mapsto-\Delta v-\lambda f(v+\mu\psi)+\mu g.$$

From the Sobolev embedding theorem it follows that $H^s(\Omega)$, $s>n/2+2$, may be viewed as a subspace of the C^2 functions on Ω. One consequence of this is that $(\lambda,\mu,v)\to E_{\lambda,\mu}(v)$ is C^∞ as a functional on $\mathbb{R}^2\times H^s(\Omega)$. If, however, one worked with the space $H^1(\Omega)$, which might seem natural, we would not in general obtain that E is C^∞. But with $H^s(\Omega)$, with $s>n/2+2$, all the critical points of $E_{\lambda,\mu}$ for fixed (λ,μ) must be degenerate in the classical sense that the Hessian induces an isomorphism between the space $\mathscr{H}_0$ and its dual. Moreover the generalizations of classical singularity theory to Banach spaces as carried out by Magnus [4] and Chillingworth [2] do not apply to these problems (see for example the introduction to [7]).

The second author developed an approach to singularity theory on Banach manifolds which would apply to problems in the classical calculus of variations. This was first presented in abstract form in [6], 1976 where it was applied to normal forms for the energy whose critical points are geodesics on a Riemannian manifold. Somewhat later less abstract versions were given in [1] and [7], where applications to minimal surfaces were given. Recent generalizations of this approach have been given by Golubitsky and Marsden [3].

Our main result is another application of this approach.

Theorem 2. *Let f, g: $\mathbb{R}\to\mathbb{R}$ and ϕ be C^∞ and let λ_0 be a simple eigenvalue of* (2) *with eigenfunction e (necessarily C^∞). We make the following assumptions on e, f, g and ψ, where ψ is the unique function given in* (5):

(H1) $$f(0)=0, \quad f'(0)=1, \quad f''(0)=0, \quad f'''(0)\neq 0$$

(H2) $$\int_\Omega(\lambda_0\psi-g)e\,dx\neq 0.$$

We then may conclude that the bifurcation of solutions to (1) *about $u\equiv 0$ for λ near λ_0 and μ near $\mu_0=0$ can be described by the cusp in the sense of Theorem* 1.

Proof. Solutions of (1) are critical points of $E_{\lambda,\mu}$. To apply Theorem 1 one first must show that $E_{\lambda,\mu}$ is a C^∞-acceptable family and then verify conditions (i) through (iv). Define $\langle\ ,\ \rangle$: $\mathscr{H}_0\times\mathscr{H}_0\to\mathbb{R}$ by

$$\langle v,w\rangle=\int_\Omega\nabla v\cdot\nabla w\,dx.$$

Define $X_{\lambda,\mu}$: $\mathscr{H}_0\to\mathscr{H}_0$ as follows. Let $v\in\mathscr{H}_0$ and define $X_{\lambda,\mu}$ by the partial differential equation

$$X_{\lambda,\mu}(v)\colon \Omega\to\mathbb{R}, \qquad X_{\lambda,\mu}(v)=0 \quad \text{on } \partial\Omega$$

and

$$-\Delta X_{\lambda,\mu}(v)=-\Delta v-\lambda f(v+\mu\psi)+\mu g \quad \text{in } \Omega.$$

One readily checks by elliptic regularity theory that $X_{\lambda,\mu}(v)\in\mathscr{H}_0$ and that

$$(\lambda,\mu,v)\mapsto X_{\lambda;\mu}(v)$$

is C^∞. Moreover it is trivial to see that

$$dE_{\lambda,\mu}(v)[h]=\langle h,X_{\lambda,\mu}(v)\rangle.$$

Consequently it remains to compute the derivatives (i) through (iv) as given in Theorem 1. Since $u\equiv 0$ satisfies (1) for $\mu=0$ it is a critical point of E_{λ_0,μ_0}, which for notational convenience we now simply denote by E.

We have that

$$\begin{aligned} D^2E(0)[h,k]&=\int_\Omega\nabla h\nabla k\,dx-\lambda_0\int_\Omega hk\,dx\\ &=\int_\Omega(-\Delta h-\lambda_0 h)k\,dx.\end{aligned}$$

Since λ_0 is a simple eigenvalue of (2), $D^2E(0)$ has a kernel spanned by $e \in C^\infty(\Omega, \mathbb{R})$. Finally, we obtain by simple computation at the point $q = (\lambda_0, 0, 0) \in \mathbb{R}^2 \times \mathscr{H}_0$ using (H_1), (H_2):

(i′) $$\frac{\partial^3 E}{\partial e^3} = -\lambda_0 \int_\Omega f''(0) e^3 dx = 0;$$

(ii′) $$\frac{\partial^4 E}{\partial e^4} = -\lambda_0 \int_\Omega f'''(0) e^4 dx \neq 0;$$

(iii′) $$\frac{\partial^3 E_{\lambda,\mu}}{\partial \lambda \partial e^2} = -\int_\Omega f'(0) e^2 dx \neq 0;$$

(iv′) $$\frac{\partial^2 E_{\lambda,\mu}}{\partial \mu \partial e} = \int_\Omega \nabla e \nabla \psi dx - \lambda_0 \int_\Omega f'(0) e \psi dx + \int_\Omega g e dx \neq 0,$$

where the first term is zero because of (5) and (2) (integrate by parts). This concludes the proof.

Remarks. (a) The assumption of C^∞ on f is not necessary. One must only choose f to be somewhat smoother than that required by the standard normal form theorem for two parameter families of functions of one variable.

(b) In [5] a weak form of Theorem 2 was obtained for $g \equiv 0$ using continuation techniques under less restrictive assumptions on f. More precisely, there it was shown that the sections of constant λ of solutions to (1) provide μ-continua of solutions, which fit together in a "cusp-like" picture.

(c) Assumption (H1) appears to be quite natural for a singularity approach. We want to comment on assumption (H2). Consider the problem

(8) $$\begin{aligned} &u'' + \lambda f(u) = 0 \\ &u(0) = \mu = u(\pi). \end{aligned}$$

and assume that f satisfies (H1). It is well known, that $\lambda_n = n^2$, $n = 1, 2, \ldots$ are simple eigenvalues of

(9) $$\begin{aligned} &u'' + \lambda u = 0 \\ &u(0) = 0 = u(\pi) \end{aligned}$$

with eigenfunctions $e_n(x) = \sin(nx)$ and that each λ_n is a point of bifurcation to solutions of (8) near $u \equiv 0$ and $\mu_0 = 0$.

Note that

$$\int_0^1 e_n dx \begin{cases} \neq 0 & \text{for } n \text{ odd} \\ = 0 & \text{for } n \text{ even.} \end{cases}$$

Thus, according to Theorem 2 solutions of (8) near $u \equiv 0$, $\mu_0 = 0$ and λ_n, n odd, form a cusp. The question is, what is the picture near $u \equiv 0$, $\mu_0 = 0$ and λ_n, n even? This is answered in [5] where it is shown by continuation techniques, that nontrivial solutions locally look like a "paraboloid", i.e. not a cusp. In this sense assumption (H2) seems to be a necessary assumption.

(d) Theorem 2 can be generalized in various directions, e.g. one can formulate a theorem for systems of non-linear elliptic variational problems, which we omit here. We want to present here, however, a generalization which appears to be noteworthy:

$$E_{\lambda,\mu}(v) = \int_\Omega H(\nabla v)\,\mathrm{d}x + \lambda \int_\Omega F(v+\mu 1)\,\mathrm{d}x, \tag{10}$$

where $E_{\lambda,\mu}$: $\mathscr{H}_0 \to \mathbb{R}$. Assume that F: $\mathbb{R} \to \mathbb{R}$ and H: $\mathbb{R}^2 \to \mathbb{R}$ are C^∞ and $D^2 H(p)$: $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is positive definite for all $p \in \mathbb{R}^2$. Moreover assume that $DH(0)=0$, $D^3 H(0)=0$ and $D^4 H(0)=0$. Write $H = H(p_1, p_2)$ and

$$A = \frac{\partial^2 H}{\partial p_1^2}(0), \quad B = \frac{\partial^2 H}{\partial p_1 \partial p_2}(0) \quad \text{and} \quad C = \frac{\partial^2 H}{\partial p_2^2}(0).$$

Theorem 3. *If λ_0 is a simple eigenvalue of the elliptic equation*

$$A\frac{\partial^2 h}{\partial p_1^2} + 2B\frac{\partial^2 h}{\partial p_1 \partial p_2} + C\frac{\partial^2 h}{\partial p_2} + \lambda h = 0 \quad \text{in } \Omega$$

$$h = 0 \quad \text{on } \partial\Omega$$

with eigenfunction e satisfying $\int_\Omega e\,\mathrm{d}x \neq 0$. Furthermore assume that $f(s) = F'(s)$ satisfies (H1). *Then the two parameter bifurcation from the critical point $v \equiv 0$ near $\mu_0 = 0$ and λ_0 is described by the cusp catastrophe.*

The proof is strictly analogous to the proof above.

References

1. Beeson, M.J., Tromba, A.J.: The cusp catastrophe of Thom in the bifurcation of minimal surfaces. To appear, preprint no. 462, SFB 72, University of Bonn
2. Chillingworth, D.: A global genericity theorem for bifurcations in variational problems. J. Functional Analysis **35**, 251–278 (1980)
3. Golubitsky, M., Marsden, J.: The Morse lemma in infinite dimensions via singularity theory. To appear
4. Magnus, R.J.: On universal unfoldings of certain real functions on a Banach space. Math. Proc. Cambridge Philos. Soc. **81**, 91–95 (1976)
5. Peitgen, H.-O., Schmitt, K.: Global analysis of elliptic two-parameter eigenvalue problems. To appear in: Trans. Amer. Math. Soc.
6. Tromba, A.J.: Almost Riemannian structures on Banach manifolds, the Morse lemma, and the Darboux theorem. Canadian J. Math. **28**, 640–652 (1976)
7. Tromba, A.J.: A sufficient condition for a critical point of a functional to be a minimum and its application to Plateau's problem. Math. Ann. **263**, 303–312 (1983)

Received October 4, 1983

Math. Z. 187, 151–164 (1984)

Mathematische Zeitschrift

An Elementary Proof for a Compact Imbedding Result in Generalized Electromagnetic Theory*

R. Picard

Abteilung für math. Methoden der Physik, Institut für angewandte Mathematik der Universität, Wegelerstr. 10, D-5300 Bonn, West Germany

0. Introduction

In the investigation of generalized curl and div operations d and δ in the framework of alternating differential forms, Hilbert space ideas have been introduced by many authors (see e.g. [25, 7, 6, 11] to mention a few). This appears to be quite natural, taking into account that the exterior derivative d and its co-derivative δ can be regarded as formally skew-adjoint operations relative to a certain inner product (see below). In particular the study of harmonic differential forms has also been attacked with Hilbert space methods. An estimate, first stated and proved for compact, oriented, Riemannian manifolds by Gaffney [7], has played a key role in this methodological context. Gaffney's inequality for manifolds with boundary involves certain boundary conditions, e.g. [6, 14].

Let M be a bounded, smooth, oriented, m-dimensional, Riemannian manifold with boundary. Then the regular, alternating differential q-forms over M can be regarded as an inner product space with inner product $(\omega, \eta) = \int_M \omega \wedge *\eta$, where ω and η are regular q-forms, $q \in \{0, \ldots, m\}$, $*$ is the Hodge star operator, and $\wedge$ the alternating product (see e.g. [8, 12, 24]). Its completion with respect to the induced norm $\|\cdot\|$ is a Hilbert space denoted by $L_2^q(M)$, $q \in \{0, \ldots, m\}$. The exterior derivative d may be extended to a closed operator from a domain in $L_2^q(M)$ into $L_2^{q+1}(M)$. In order to keep the resemblance to electromagnetic theory which historically speaking has provided the main source of stimulation, we prefer to denote this operator by curl, (compare [25, 22]). Its formal adjoint will be denoted by $-\mathrm{div}$. Gaffney's inequality may now be formulated as:

$$(0.1) \qquad \|\omega\|_1^2 \leqq C(\|\mathrm{curl}\,\omega\|^2 + \|\mathrm{div}\,\omega\|^2 + \|\omega\|^2)$$

* This paper is a publication of results obtained by the author in his habilitation thesis accepted by the Faculty of Mathematics and Science of the University of Bonn

for some uniform constant $C>0$ and all regular q-forms ω satisfying the boundary condition $\omega=0$ on ∂M (or $*\omega=0$ on ∂M by $*$-duality). In (0.1) the norm $\|\cdot\|_1$ is an inner product norm built up from the $L_2^q(M)$-norm and the $L_2^0(M)$-norms of all partial derivatives of the coefficients of say ω with respect to a fixed covering with coordinate systems, see e.g. [6], [12], p. 289. The corresponding Hilbert space $H_1^q(M)$ thus coincides locally as far as coefficients are concerned with the usual Sobolev space $H_1(M)$ of functions with measurable, square integrable zeroth and first derivatives.

By a density argument, (0.1) shows that if we regard those $\omega\in L_2^q(M)$ with $\operatorname{curl}\omega\in L_2^{q+1}(M)$, $\operatorname{div}\omega\in L_2^{q-1}(M)$, satisfying the boundary condition $\omega=0$ on ∂M (in a generalized sense) as another Hilbert space $\mathring{R}^q(M)\cap D^q(M)$ with norm provided by the right-hand side of (0.1), then Gaffney's inequality gives continuity of the imbedding of $\mathring{R}^q(M)\cap D^q(M)$ in $H_1^q(M)$, i.e.

$$(0.2)\qquad \mathring{R}^q(M)\cap D^q(M)\hookrightarrow H_1^q(M).$$

Applying Rellich's selection theorem, which gives compactness of the imbedding $H_1^q(M)\hookrightarrow L_2^q(M)$, we get

$$(0.3)\quad \mathring{R}^q(M)\cap D^q(M)\hookrightarrow\hookrightarrow L_2^q(M),\ q\in\{0,\dots,m\},$$

(for '$\hookrightarrow\hookrightarrow$' read 'compactly imbedded in'). The basic importance of a (0.3) type result in the Hilbert space approach to the theory of harmonic differential forms has been stressed by G. Fichera, [5], Later we shall point out some of the consequences of (0.3) in this connection (see Chap. 2).

Comparing (0.1) and (0.3) it is essential to realize that there are examples for domains with non-smooth boundaries, where (0.1) fails to be correct, but (0.3) is still valid (compare e.g. [3]). This gives rise to the problem of characterizing a class of domains with non-smooth boundaries for which (0.3) is satisfied, without making use of (0.1) in the above way. The first solution in this direction was given by N. Weck [22, 23], who was able to show that actually

$$(0.4)\qquad \mathring{R}^q(M)\cap\varepsilon^{-1}\{E\in D^q(M)|\operatorname{div}E=0\}\hookrightarrow\hookrightarrow L_2^q(M), q\in\{0,\dots,m\},$$

holds, where ε is a positive definite, bounded linear transformation of $L_2^q(M)$ into itself. An easy consequence of his result is

$$(0.5)\qquad \mathring{R}^q(M)\cap\varepsilon^{-1}D^q(M)\hookrightarrow\hookrightarrow L_2^q(M),\quad q\in\{0,\dots,m\}, [13].$$

Weck's work also contains an important observation, namely that (0.4) (and thus (0.5)) is essentially independent of ε, in the sense that it suffices to show the result for only one special ε, say $\varepsilon=1$. In Chap. 2 we shall give a simple proof of this fact. This will be the basis of a reduction scheme allowing us to use (0.1) in a more indirect way in order to prove (0.5) for a more simply characterized class of manifolds with non-smooth boundaries.

A comparable result has been achieved in [18], based on an idea of C. Weber [21], who gave a proof for (0.5) in the case of a bounded, open domain G in $\mathbb{R}^3$ and $q=1,2$, assuming that G has the restricted cone property (for a definition see e.g. [1]). As it turns out the desired result can be proved in an even more elementary way as well as in more generality. Only the following tools will be employed:

a) basic functional analysis

b) the possibility of transforming the problem in question to the case where M is a ball in $\mathbb{R}^m$; (a situation already covered by the arguments leading to (0.3)).

The latter is based on an assumption to be imposed in Chap. 3, restricting the class of admissible manifolds to those with Lipschitz type boundaries. Since we intend to prove (0.5) in the general form (i.e. on manifolds and for arbitrary order q), there is some terminological effort involved. This will be the main concern of the following Chap. 1. In Chap. 2 we shall make precise what we mean by the term 'essentially independent of ε' used above, and we shall prove the according result. The final chapter then will contain the actual statement and proof of our main result.

1. Basic Definition

We first recall the basic notations used in similar contexts (see e.g. [22, 14, 15]). Let M be a m-dimensional, oriented, Riemannian $C_{1,1}$-manifold, $m \in \mathbb{N}$, i.e. a manifold with Lipschitz-continuously differentiable atlas. For an open subset $G \subset M$ we define $\mathring{C}^q_{0,1}(G)$ as the subset of $C^q_{0,1}(G)$ – the set of all Lipschitz continuous, alternating differential forms of degree $q, q \in \{0, \ldots, m\}$, ('$q$-forms'), – having compact support in G. We restrict our considerations to 'even' forms 'odd' forms could be treated similarly. Setting $C^q_{0,1}(G) = \{0\}$ (i.e. the set consisting of the zero tensor field of degree $|q|$) if $q < 0$ or $q > d$, we may let q range in $\mathbb{Z}$, the set of integers.

Thus for $q \in \mathbb{Z}$ the Hodge star operator induces a map

$$*: C^q_{0,1}(G) \to C^{m-q}_{0,1}(G). \tag{1.1}$$

With

$$(\omega, \eta) := \int_G \omega \wedge *\eta, \quad \omega, \eta \in \mathring{C}^q_{0,1}(G), \tag{1.2}$$

the set $\mathring{C}^q_{0,1}(G)$ can be regarded as a linear space with inner product $(\cdot, \cdot)$, and we denote its completion with respect to the corresponding norm $\|\cdot\|$ by $L^q_2(G)$. $L^q_2(G)$ consists exactly of those q-forms with L_2-coefficients with respect to any coordinate system (see [12]). For later use we define

(1.3) $L^{q,\mathrm{loc}}_2(\bar{G})$ as the completion of $\mathring{C}^q_{0,1}(G)$ with respect to the following notion of convergence:

$$\varphi_n \to \varphi \text{ in } L^{q,\mathrm{loc}}_2(\bar{G}) \text{ as } n \to \infty \text{ if } \forall_{\psi \in \mathring{C}^0_{0,1}(\bar{G})}: \psi\varphi_n \to \psi\varphi \text{ in } L^q_2(G) \tag{1.4}$$

as $n \to \infty$, (following [18]).

We now define the operator

$$\begin{aligned} \text{curl}: \mathring{C}^q_{0,1}(G) \subset L^q_2(G) &\to L^{q+1}_2(G) \\ \omega &\to \text{curl}_0\, \omega, \quad q \in \mathbb{Z}. \end{aligned} \tag{1.5}$$

Here $\mathrm{curl}_0\,\omega$ is the uniquely defined $(q+1)$-form in $L_2^{q+1}(G)$ satisfying

$$\psi^{-1*}\mathrm{curl}_0\,\omega = \mathrm{d}\psi^{-1*}\omega \quad \text{in } L^{q+1}(\psi(U\cap G))$$

for any coordinate system (ψ, U). The exterior derivative d on the right hand side has to be interpreted in the sense of distributions over $\psi(U\cap G)$; ψ^{-1*} denotes - as usual (e.g. [8, 24]) - the lifting operation, which assigns a local coordinate representation to a tensor field on G. For convenience we do not denote the dependence of curl_0 on $q\in\mathbb{Z}$.

Obviously, curl is densely defined and we denote the domain of its adjoint as

(1.6) $$D^{q+1}(G) := D((\mathrm{curl}_0)^*), \qquad q\in\mathbb{Z}.$$

With $\mathrm{curl}_0{}^* = -\mathrm{div}\colon D^{q+1}(G)\subset L_2^{q+1}(G)\to L_2^q(G)$, we have using Stoke's theorem (after suitable mollification) and since $\mathring{C}_{0,1}^{q+1}(G)\subset D^{q+1}(G)$,

$$\mathrm{div}\restriction_{\mathring{C}_{0,1}^{q+1}(G)} = (-1)^{(q-1)(q-m)} *\mathrm{curl}_0 *, \qquad q\in\mathbb{Z}.$$

Thus div is a generalized differential operator, ('$\restriction$' read 'restricted to').

Since $-\mathrm{div}$ is densely defined, we know that its adjoint exists and is equal to the closure $\overline{\mathrm{curl}_0}$ of curl_0 (see e.g. [18], compare [16]). We set

(1.7) $$\mathring{R}^q(G) := D(-\mathrm{div}^*) \equiv D(\overline{\mathrm{curl}_0}).$$

Similarly, let

$$\mathrm{div}_0 := \mathrm{div}\restriction_{\mathring{C}_{0,1}^{q+1}(G)}$$

and

(1.8) $$D(\mathrm{div}_0{}^*) =: R^q(G),$$

(1.9) $$\mathring{D}^q(G) := D(\overline{\mathrm{div}_0}) \equiv D(\mathrm{div}_0{}^{**}).$$

Since $\mathring{C}_{0,1}^q(G)\subset R^q(G)$, $\mathrm{div}_0{}^*\restriction_{\mathring{C}_{0,1}^q(G)} = -\mathrm{curl}_0$, it is reasonable to denote

$$\mathrm{curl} = -\mathrm{div}_0{}^*.$$

We note the following inclusions:

(1.10) $$\mathring{R}^q(G)\subset R^q(G), \quad \mathring{D}^{q+1}(G)\subset D^{q+1}(G), \; q\in\mathbb{Z}.$$

As an immediate consequence of (1.10) we have

(1.11) $$\begin{aligned} \mathrm{curl}\restriction\mathring{R}^q(G) &= \overline{\mathrm{curl}_0} \\ \mathrm{div}\restriction\mathring{D}^{q+1}(G) &= \overline{\mathrm{div}_0}. \end{aligned}$$

Furthermore, since curl, div are closed operators on $\mathring{R}^q(G)$, $R^q(G)$, $\mathring{D}^{q+1}(G)$, $D^{q+1}(G)$ respectively, $q\in\mathbb{Z}$, these domains are Hilbert spaces with respect to the corresponding graph norms.

From the definitions of $D^q(G)$, $\mathring{R}^q(G)$ we see that

$$u \in \mathring{D}^q(G) \Leftrightarrow \forall\, v \in R^{q-1}(G) : (v, \operatorname{div} u) + (\operatorname{curl} v, u) = 0,$$
$$u \in \mathring{R}^q(G) \Leftrightarrow \forall\, v \in D^{q+1}(G) : (v, \operatorname{curl} u) + (\operatorname{div} v, u) = 0. \tag{1.12}$$

This shows that these spaces generalize boundary conditions. In fact, assuming sufficient regularity the left hand sides in the equations are boundary terms of the type $\int_{\partial G} v \wedge *u$, $\int_{\partial G} *v \wedge u$ respectively. Thus it is legitimate to introduce the following 'façon de parler'

$$*u = 0 \text{ on } \partial G \text{ in the generalized sense iff } u \in \mathring{D}^q(G),$$
$$u = 0 \text{ on } \partial G \text{ in the generalized sense iff } u \in \mathring{R}^q(G),\ q \in \mathbb{Z}. \tag{1.13}$$

These conditions are usually referred to as Neumann's or Dirichlet's boundary condition, respectively.

In our first Lemma we restate in the framework developed so far the analogue of the classical fact that exact forms are closed (Poincaré's Lemma).

Lemma 1. *We have*

$$\operatorname{curl} R^q(G) \subset R^{q+1}(G) \quad \text{and} \quad \operatorname{curl}\operatorname{curl} R^q(G) = \{0\},$$
$$\operatorname{div} D^q(G) \subset D^{q-1}(G) \quad \text{and} \quad \operatorname{div}\operatorname{div} D^q(G) = \{0\},\ q \in \mathbb{Z}.$$

Furthermore,

$$\operatorname{curl} \mathring{R}^q(G) \subset \mathring{R}^{q+1}(G)$$
$$\operatorname{div} \mathring{D}^q(G) \subset \mathring{D}^{q-1}(G), \qquad q \in \mathbb{Z}. \tag{1.14}$$

Proof. Noting that

$$*R^q(G) = D^{m-q}(G) \quad \text{and} \quad ** = (-1)^{q(m-q)}, \tag{1.15}$$

we realize that it is sufficient to deal with $R^q(G)$-spaces only; the results for the $D^q(G)$-spaces then follow by $*$-duality (compare [15, 17]). So in order to prove the Lemma it remains to show that we have

$$\forall\, \phi \in \mathring{C}^{q+2}_{0,1}(G)\ \forall\, u \in R^q(G) : (\operatorname{curl} u, \operatorname{div} \phi) = 0. \tag{1.16}$$

By a density argument (1.16) implies

$$\operatorname{curl} R^q(G) \perp \operatorname{div} \mathring{D}^{q+2}(G).$$

From (1.8), (1.9), this in turn yields the statements in the Lemma.

In order to show (1.16) we first remark that

$$(\operatorname{curl} u, \operatorname{div} \phi) = (-1)^{q+1} \int_G \operatorname{curl} u \wedge \operatorname{curl} *\phi$$

for $u \in R^q(G)$, $\phi \in \mathring{C}^{q+2}_{0,1}(G)$. Thus (1.16) reduces to

$$\forall\, \phi \in \mathring{C}^{m-q-2}_{0,1}(G)\ \forall\, u \in R^q(G) : \int_G \operatorname{curl} u \wedge \operatorname{curl} \phi = 0. \tag{1.17}$$

Without loss of generality (partition of unity!) we may assume that ϕ has its support in some coordinate neighbourhood U of a coordinate system (ψ, U). This leaves us with the investigation of

$$I = \int_{\psi(U \cap G)} d(\psi^{-1*} u \wedge d\psi^{-1*}\phi).$$

By use of a mollification it is clear that $I=0$ is implied by Stoke's theorem in the form:

$$\forall \omega \in \mathring{C}_{\infty}^{m-1}(\mathbb{R}^m)\colon \int_{\mathbb{R}^m} d\omega = 0. \tag{1.18}$$

□

Lemma 1 gives rise to the following decompositions (compare [5, 12, 15, 18]):

$$\begin{aligned} L_2^p(G) &= \overline{\operatorname{curl} \mathring{R}^{q-1}(G)} \oplus \mathscr{H}^q_{R,\varepsilon_q}(G) \oplus \varepsilon_q^{-1} \overline{\operatorname{div} D^{q+1}(G)} \\ &= \varepsilon_q^{-1} \overline{\operatorname{div} \mathring{D}^{q+1}(G)} \oplus \mathscr{H}^q_{D,\varepsilon_q}(G) \oplus \overline{\operatorname{curl} R^{q-1}(G)}, \; q \in \mathbb{Z}, \end{aligned} \tag{1.19}$$

with

$$\mathscr{H}^q_{D,\varepsilon_q} = \{E \in \varepsilon_q^{-1} \mathring{D}^q(G) \mid \operatorname{curl} E = 0, \quad \operatorname{div} \varepsilon_q E = 0\},$$
$$\mathscr{H}^q_{R,}\varepsilon_q = \{E \in \mathring{R}^q(G) \mid \operatorname{curl} E = 0, \quad \operatorname{div} \varepsilon_q E = 0\}.$$

Here ε_q is a symmetric, bounded, positive definite, linear mapping from $L_2^q(G)$ into $L_2^q(G)$, $q \in \mathbb{Z}$. Such ε_q, $q \in \mathbb{Z}$, will henceforth be called admissible. The decompositions (1.19) are orthogonal in the sense of the inner product $(\cdot, \varepsilon_q \cdot)$ and follow as simple applications of the projection theorem.

Since we now have the means, let us formulate the compact imbedding property for which we want to give a simple sufficient condition in Chap. 3.

2. The Compact Imbedding Property and Some of Its Consequences

We shall investigate the following compact imbedding property:

$$\mathring{R}^q(G) \cap \varepsilon_q^{-1} D^q(G) \hookrightarrow\hookrightarrow L_2^{q,\mathrm{loc}}(\bar{G}) \tag{2.1}$$

for all admissible ε_q, $q \in \mathbb{Z}$.

Here the term on the left hand side is to be understood as a Hilbert space with the inner product

$$(\phi, \psi) + (\operatorname{curl} \phi, \operatorname{curl} \psi) + (\operatorname{div} \varepsilon_q \phi, \operatorname{div} \varepsilon_q \psi),$$

$\phi, \psi \in \mathring{R}^q(G) \cap \varepsilon_q^{-1} D^q(G)$.

Remark 1. a) *If* $G \subset\subset M$ *(i.e.* G *is open,* $\bar{G}$ *compact and* $\bar{G} \subset M$*) then* (2.1) *reduces to*

$$\mathring{R}^q(G) \cap \varepsilon_q^{-1} D^q(G) \hookrightarrow\hookrightarrow L_2^q(G) \quad \textit{for all admissible } \varepsilon_q,\ q \in \mathbb{Z}. \tag{2.2}$$

b) *We observe that by the assumptions on* ε_q, $q \in \mathbb{Z}$, *we have*

$$R^q(G) \cap *\varepsilon_{m-q} * \mathring{D}^q(G) \hookrightarrow *\varepsilon_{m-q} * (\mathring{D}^q(G) \cap *\varepsilon_{m-q}^{-1} * R^q(G))$$

(i.e. continuously imbedded) for admissible ε_{m-q}, $q \in \mathbb{Z}$. *Now, if* ε_q *is admissible, so is* ε_q^{-1}, $*\varepsilon_q*$, $*\varepsilon_q^{-1}*$ *for* $q \in \mathbb{Z}$. *Thus* (2.1) *implies by* $*$*-duality*

$$R^q(G) \cap \varepsilon_q^{-1} \mathring{D}^q(G) \hookrightarrow\hookrightarrow L_2^{q,\mathrm{loc}}(\bar{G}) \tag{2.3}$$

for all admissible ε_q, $q \in \mathbb{Z}$.

To illustrate the importance of property (2.1) let us point out some consequences.

If we assume in addition that $G \subset\subset M$, then

a) curl $R^q(G)$, curl $\mathring{R}^q(G)$, div $D^q(G)$, div $\mathring{D}^q(G)$ are closed subspaces of $L_2^q(G)$.

b) $\mathscr{H}_{R,\varepsilon_q}^q(G)$, $\mathscr{H}_{D,\varepsilon_q}^q(G)$ are finite-dimensional.

c) The spectrum of the (generalized) Maxwell operator, which can be written (in an obvious notation) as

$$\mathscr{M}_q = i \begin{pmatrix} 0 & \mathrm{div} \\ -\mathrm{div}^* & 0 \end{pmatrix},$$

(thus being selfadjoint on $H = L_2^q(G) \oplus L_2^{q+1}(G)$), consists of isolated points on the real axis, and the eigenvalues $\lambda_n \neq 0$, $n \in \mathbb{N}$, have finite multiplicity.

In the case that G is unbounded we assume that M, outside a compact set containing $M \setminus \bar{G}$, is isometrically $C_{1,1}$-diffeomorphic to the exterior of a ball as an m-dimensional submanifold of Euclidean $\mathbb{R}^m$. As a consequence of (2.1) it is known that for every sequence $(\varepsilon_n)_n \subset \mathbb{R}$, with $\varepsilon_n \to \pm 0$ as $n \to \infty$, if $\lambda \in \mathbb{R} \setminus \{0\}$, and $f \in H$ with compact support, $((\mathscr{M}_q - \lambda - i\varepsilon_n)^{-1} f)_n$ converges in the sense of $L_2^{q,\mathrm{loc}}(\bar{G}) \oplus L_2^{q+1,\mathrm{loc}}(\bar{G})$, i.e. the principle of limiting absorption holds (see [4, 22, 14, 18]). This in turn implies absolute continuity of the continuous spectrum (see [11], [18]) as well as the time asymptotic behaviour of the solution of the corresponding evolution equation

$$V(t) = e^{itM_q} V_0 \to V_{0,s} \quad \text{in} \quad L_2^{q,\mathrm{loc}}(\bar{G}) \oplus L_2^{q+1,\mathrm{loc}}(\bar{G}) \quad \text{as} \quad t \to \infty,$$

where $V_{0,s}$ denotes the static part of V_0, i.e. its projection on the null space of $\mathscr{M}_q$, $q \in \mathbb{Z}$, ([14], compare [26]). (For a), b) compare [5, 11, 16], see [15, 18]; for c) see e.g. [22, 23, 14].)

The first important observation on our way to investigating (2.1) is that the compactness of the imbedding is essentially independent of the choice of ε_q, $q \in \mathbb{Z}$.

We state this (see Remark 1.a) more precisely as

Lemma 2. *Let* $G \subset\subset M$; *if there are admissible* ε_q *such that*

$$\mathring{R}^q(G) \cap \varepsilon_q^{-1} D^q(G) \hookrightarrow\hookrightarrow L_2^q(G), q \in \mathbb{Z}, \qquad \textit{then (2.2) is valid.}$$

Proof. Let us write ε instead of ε_q and let μ be another admissible mapping from $L_2^q(G)$ into $L_2^q(G)$ and let $q \in \mathbb{Z}$ be fixed. Then, comparing the decompositions

$$\begin{aligned} L_2^q(G) &= \overline{\mathrm{curl}\, \mathring{R}^{q-1}(G)} \oplus \varepsilon^{-1} \{E \in D^q(G) | \mathrm{div}\, E = 0\}, \\ &= \overline{\mathrm{curl}\, \mathring{R}^{q-1}(G)} \oplus \mu^{-1} \{E \in D^q(G) | \mathrm{div}\, E = 0\}, \end{aligned} \tag{2.4}$$

orthogonal in the sense of $(\cdot, \varepsilon\cdot)$, $(\cdot, \mu\cdot)$ respectively, we get the following relations between the corresponding projections P_ε, Q_ε, P_μ, Q_μ:

$$(2.5) \qquad P_\mu P_\varepsilon = P_\varepsilon, \; Q_\mu \mu^{-1}\varepsilon Q_\varepsilon = \mu^{-1}\varepsilon Q_\varepsilon.$$

Using $P_\varepsilon + Q_\varepsilon = P_\mu + Q_\mu = \mathrm{id}$ we get from (2.5)

$$(2.6) \qquad Q_\mu Q_\varepsilon = Q_\mu, \; P_\mu \mu^{-1}\varepsilon P_\varepsilon \varepsilon^{-1}\mu = P_\mu.$$

First we have

$$Q_\mu \restriction \mathring{R}^q(G) = Q_\mu Q_\varepsilon \restriction \mathring{R}^q(G),$$

$Q_\mu \restriction \mathring{R}^q(G)$, $Q_\varepsilon \restriction \mathring{R}^q(G)$ are mappings from $\mathring{R}^q(G)$ into $L_2^q(G)$, and

$$P_\mu \restriction \mu^{-1} D^q(G) = P_\mu \mu^{-1}\varepsilon P_\varepsilon \restriction \varepsilon^{-1} D^q(G)\, \varepsilon^{-1}\mu \restriction \mu^{-1} D^q(G),$$

$P_\mu \restriction \mu^{-1} D^q(G)$, $P_\varepsilon \restriction \varepsilon^{-1} D^q(G)$ are mappings from $\mu^{-1} D^q(G)$, $\varepsilon^{-1} D^q(G)$ respectively into $L_2^q(G)$, by (2.6). Now boundedness of

$$\begin{aligned} Q_\mu &: L_2^q(G) \to L_2^q(G), \\ P_\mu \mu^{-1}\varepsilon &: L_2^q(G) \to L_2^q(G), \\ \varepsilon^{-1}\mu &: \mu^{-1} D^q(G) \to \varepsilon^{-1} D^q(G), \end{aligned}$$

and the assumed compactness of

$$Q_\varepsilon \restriction \mathring{R}^q(G), \; P_\varepsilon \restriction \varepsilon^{-1} D^q(G),$$

implies that

$$(2.7) \qquad Q_\mu \restriction \mathring{R}^q(G) \quad \text{and} \quad P_\mu \restriction \mu^{-1} D^q(G) \quad \text{are compact,}$$

(observe that the ranges $R(Q_\varepsilon \restriction \mathring{R}^q(G))$, $R(P_\varepsilon \restriction \varepsilon^{-1} D^q(G))$ are according to Lemma 1 (boundedly) imbedded in $\mathring{R}^q(G) \cap \varepsilon^{-1} D^q(G)$). *A*-fortiori,

$$(2.8) \qquad \begin{aligned} &Q_\mu \restriction \mathring{R}^q(G) \cap \mu^{-1} D^q(G), \\ &P_\mu \restriction \mathring{R}^q(G) \cap \mu^{-1} D^q(G), \end{aligned}$$

are compact operators; but the sum of these is just the imbedding

$$\begin{aligned} j: \mathring{R}^q(G) \cap \mu^{-1} D^q(G) &\to L_2^q(G) \\ u &\to u \end{aligned}$$

Thus the Lemma is proved. □

Furthermore we have the following:

Corollary 1. *Let* $G \subset\subset M$. *Then* (2.2) *is also independent of the Riemannian metric tensor.*

Proof. Using Lemma 2 we may characterize (2.2) by the case $\varepsilon_q = \mathrm{id}$, $q \in \mathbb{Z}$.

So let

$$\overset{\circ}{R}^q(G) \cap D^q(G) \hookrightarrow\hookrightarrow L_2^q(G), \quad q \in \mathbb{Z}.$$

Different Riemannian metric tensors give rise to different but metrically equivalent L_2-spaces.

If we indicate reference to a different Riemannian metric tensor by '$\sim$' we have

$$\begin{aligned}\mathrm{div} &= (-1)^{(q-1)(m-q)} * \mathrm{curl} * \\ &= (-1)^{qm-1} * \tilde{*}\tilde{*}\, \mathrm{curl}\, \tilde{*}\tilde{*} * \\ &= (-1)^{m-1} * \tilde{*}\, \mathrm{div}\, \tilde{*} * = (-1)^{(q-1)(m-1)} * \tilde{*}\, \mathrm{div} (-1)^{q(m-q)} \tilde{*} *\end{aligned}$$

where $\varepsilon = (-1)^{q(m-q)} \tilde{*} *$ admissible (with respect to $L_2^q(G)$, and so is $(-1)^{(q-1)(m-q+1)} * \tilde{*}$ with respect to $L^{q-1}(G)$). Thus

$$\tilde{\overset{\circ}{R}}{}^q(G) \cap \varepsilon^{-1} \tilde{D}^q(G) \hookrightarrow \overset{\circ}{R}^q(G) \cap D^q(G) \hookrightarrow\hookrightarrow L_2^q(G) \hookrightarrow \tilde{L}_2^q(G), \quad q \in \mathbb{Z},$$

i.e. (2.2) with respect to a different Riemannian metric tensor (again using Lemma 2). □

Since (2.1) is a local property (partition of unity!) we get the following from Lemma 2 and Corollary 1:

Corollary 2. *Property* (2.1) *is independent of the special choice of Riemannian metric tensor and admissible* ε_q, $q \in \mathbb{Z}$.

Remark 2. Since a change in the choice of a Riemannian metric tensor or an admissible ε_q, $q \in \mathbb{Z}$, *can be regarded as a mere change of the meaning of orthogonality in* $L_2^q(G)$, *we see another interesting fact, namely that the dimension numbers* $\beta^q = \dim \mathcal{H}_{R,\varepsilon_q}^q$, $\beta^{m-q} = \dim \mathcal{H}_{D,\varepsilon_q}^q$ *are actually independent of these data, too. (The definition of* β^q, $q \in \mathbb{Z}$, *is noncontradictory because of* $*$*-duality!.)*

This can easily be concluded from the first relation in (2.6). *From* (2.6) *and* (1.19) *we see that*

$$\mathcal{H}_{R,\mu}^q = Q_\mu \mathcal{H}_{R,\mu}^q, \quad = Q_\mu Q_\varepsilon \mathcal{H}_{R,\mu}^q$$

and

$$Q_\varepsilon \mathcal{H}_{R,\mu}^q \subset \mathcal{H}_{R,\varepsilon}^q, \quad \textit{implying}$$

$$\mathcal{H}_{R,\mu}^q \subset Q_\mu \mathcal{H}_{R,\varepsilon}^q, \quad \textit{thus} \tag{2.9}$$

$$\dim \mathcal{H}_{R,\mu}^q \leqq \dim \mathcal{H}_{R,\varepsilon}^q.$$

Interchanging the role of ε *and* μ *we get indeed*

$$\dim \mathcal{H}_{R,\mu}^q = \dim \mathcal{H}_{R,\varepsilon}^q, \textit{(compare [20])}. \tag{2.10}$$

The independence of the Riemannian metric tensor follows along the same line as in the proof of Corollary 1. The argument does not use the special assumption that $G \subset\subset M$ and is thus also valid in the exterior case described above ($\varepsilon = \mu = \mathrm{id}$ outside the mentioned compact set).

We now turn to the proof of our main result.

3. The Compact Imbedding Result

In this final chapter we will give a relatively simple condition under which (2.1) can be proved valid in an elementary way. The condition in question is the following additional assumption:

(3.1) *G is a $C_{0,1}$-manifold with boundary having a compatible structure with respect to M.*

By 'compatible structure', we mean that if $((\psi_\alpha | V_\alpha))_\alpha$ is a (locally finite) covering of G with $C_{0,1}$-charts and $((\varphi_\beta | U_\beta))_\beta$ is a (locally finite) covering of G with $C_{1,1}$-charts (according to the $C_{1,1}$-structure of M) then

$$\varphi_\beta \circ \psi_\alpha^{-1} \colon \psi_\alpha(V_\alpha \cap U_\beta) \to \varphi_\beta(V_\alpha \cap U_\beta)$$

are $C_{0,1}$-homeomorphisms for α, β in the respective index sets. Without loss of generality we may assume that the index sets coincide, $U_\alpha = V_\alpha$ for all α and that $\psi_\alpha(V_\alpha \cap G)$ is a sphere, and we shall do so in the following. Thus we have a situation as indicated in the following diagram (id_α denotes the identity on U_α).

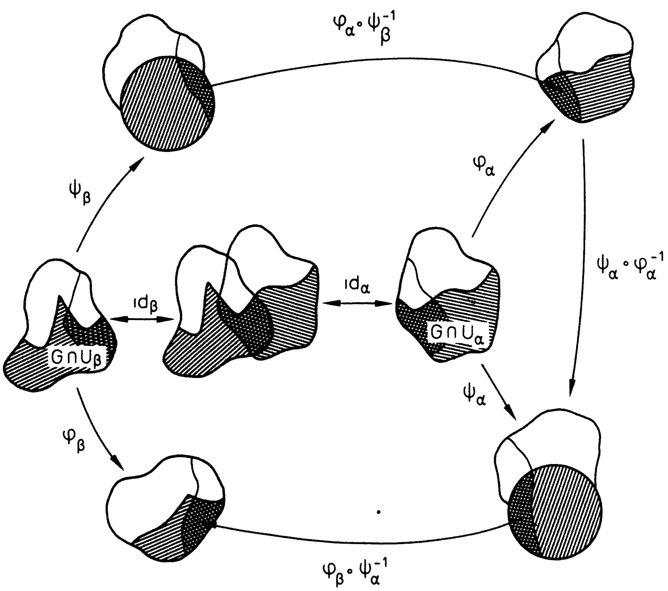


Fig. 1

Now we state our theorem:

Theorem. *Under the additional assumption* (3.1), *property* (2.1) *is valid.*

Proof. First we observe that by the consideration that led to Corollary 2 and by the use of local coordinates (3.1) leads to the following simplification. It suffices to show

(3.2) $$\mathring{R}^q(G) \cap D^q(G) \hookrightarrow\hookrightarrow L_2^q(G),$$

where $G \subset\subset \mathbb{R}^m$ (as an Euclidean manifold) is $C_{0,1}$-homeomorphic to the unit ball in $\mathbb{R}^m$. We shall prove (3.2) by transformation to the unit ball B. Let $\phi: B \to G$ be the resulting $C_{0,1}$-homeomorphism (which is a diffeomorphism in a generalized sense with L_∞-partial derivatives), so transformation of differential forms may be defined as follows:

$$\phi^* f = f \circ \phi,$$
$$\phi^* \operatorname{curl} f = \mathrm{d}\phi^* f = \mathrm{d}(f \circ \phi), \quad \text{for all } f \in \overset{\circ}{C}{}^0_{0,1}(G). \tag{3.3}$$

Here the exterior derivative d is again to be understood in the sense of distributions over B (coefficient-wise).

Definition (3.3) can be extended inductively by setting

$$\phi^*(\omega \wedge \eta) = \phi^* \omega \wedge \phi * \eta, \quad \omega \in \overset{\circ}{C}{}^p_{0,1}(G),\ \eta \in \overset{\circ}{C}{}^q_{0,1}(G),\ p, q \in \mathbb{Z}. \tag{3.4}$$

An easy calculation shows that

$$\phi^* \operatorname{curl} \omega = \operatorname{curl} \phi^* \omega, \quad \omega \in \overset{\circ}{C}{}^q_{0,1}(G). \tag{3.5}$$

Formulae (3.3), (3.4) extend naturally to the L_2-case by closure, leading to a bijection between L_2-spaces:

$$\phi^*: L_2^q(G) \to L_2^q(B). \tag{3.6}$$

Extending (3.5) we see that even

$$\phi^*: \mathring{R}^q(G) \to \mathring{R}^q(B) \tag{3.7}$$

is one-to-one and onto. Furthermore, the mapping (3.7) is a homeomorphism (as well as (3.6)).

This is implied by the estimate

$$C'(\phi^* \omega \wedge * \phi^* \omega) \leqq \phi^*(\omega \wedge * \omega) \leqq C(\phi^* \omega \wedge * \phi^* \omega) \quad \text{a.e. in } B, \tag{3.8}$$

for some constants C', $C > 0$ and all $\omega \in L_2^q(G)$.

To see (3.8) we first calculate $\phi^* \omega \wedge * \phi^* \omega$ in Cartesian coordinates. We have a representation of the type

$$\omega = \frac{1}{q!} \omega_\alpha dx^\alpha \equiv \frac{1}{q!} \omega_\alpha dx^{\alpha_1} \wedge \ldots \wedge dx^{\alpha_q}, \tag{3.9}$$

where the summation convention ranges over $\alpha \equiv (\alpha_1, \ldots, \alpha_q) \in \{1, \ldots, m\}^q$, and ω_α are alternating respective to permutations (see [12], Chap. 7.2). Substituting (3.9) in $\phi^* \omega \wedge * \phi^* \omega$, we get

$$\phi^* \omega \wedge * \phi^* \omega = \frac{1}{q!} (\omega_\alpha \circ \phi) \frac{\partial \phi_\alpha}{\partial y_\beta} \frac{\partial \phi_\gamma}{\partial y_\beta} (\omega_\gamma \circ \phi), \tag{3.10}$$

where $\dfrac{\partial \phi_\alpha}{\partial y_\beta}$ is short for $\dfrac{\partial \phi_{\alpha_1}}{\partial y_{\beta_1}} \cdots \dfrac{\partial \phi_{\alpha_q}}{\partial y_{\beta_q}}$.

Since $\frac{\partial \phi_i}{\partial y_j}$, $i,j\in\{1,\dots,m\}$, are uniformly bounded a.e., the estimate

$$\phi^*\omega\wedge *\phi^*\omega \leqq C\phi^*(\omega\wedge *\omega) \quad \text{a.e. in } B$$

for some $C>0$ and all $\omega\in L_2^q(G)$, is obvious.

In order to show the remaining inequality we observe that the Jacobian $J = \left(\frac{\partial \phi_i}{\partial y_j}\right)_{i,j}$ is regular in the sense that

$$JJ' \geqq C_0 > 0 \quad \text{a.e.;}$$

it follows that $\phi^*\omega\wedge *\phi^*\omega \geqq C_0^q|\omega_\alpha\circ\phi|^2 = C_0^q\phi^*(\omega\wedge *\omega)$, (summation convention!). This in fact shows that (3.7) is a homeomorphism.

Furthermore, we have

$$\begin{aligned} *\phi^* * \operatorname{div} u &= (-1)^{(q-1)(m-q)}(-1)^{(q-1)(m-q+1)} * \operatorname{curl}\phi^* * u, \\ &= (-1)^{q-1}(-1)^{q(m-q)} * \operatorname{curl} ** \phi^* * u, \\ &= (-1)^{qm+1}\operatorname{div} * \phi^* * u, \\ &= (-1)^{qm+1}\operatorname{div}[*\phi^* * \phi^{*-1}]\phi^* u. \end{aligned}$$

Since $\varepsilon = (-1)^{q(m-q)} * \phi^* * \phi^{-1*}$ is admissible as a mapping of $L_2^q(B)\to L_2^q(B)$, we see that

$$\phi^*\colon \mathring{R}^q(G)\cap D^q(G) \to \mathring{R}^q(B)\cap \varepsilon^{-1}D^q(B)$$

is a homeomorphism.

Thus our problem is reduced to showing

$$\mathring{R}^q(B)\cap\varepsilon^{-1}D^q(B) \hookrightarrow\!\!\hookrightarrow L_2^q(B).$$

According to Lemma 2, it is enough to show

$$\mathring{R}^q(B)\cap D^q(B) \hookrightarrow\!\!\hookrightarrow L_2^q(B). \tag{3.11}$$

But (3.11) is well-known (B is the unit ball!) and we may quote arguments using Gaffney's inequality, e.g. [5, 7, 12], to conclude the proof of our theorem. □

Having proved our main result, we end our investigation with some remarks.

Remark 3. a) *With respect to remark 2 it is clear from the argument used in the proof of the above theorem that the numbers β^q, $q\in\mathbb{Z}$, are in fact invariant with respect to global Lipschitz-homeomorphisms. In the light of cohomology theory this is quite satisfactory since β^q, $q\in\mathbb{Z}$, are expected to be related to topological characteristics (Betti numbers). In* [15–17] *even less regular boundaries have been characterized for which the β^q, $q\in\mathbb{Z}$, can be related to Betti numbers. $C_{0,1}$-invariance generalizes further the class of admissible boundaries.*

b) *The theorem generalizes the compactness result achieved in* [22, 23] *considerably. In* [21] *the case of $G\subset M=\mathbb{R}^3$ has been treated. Since there are* *examples known of $C_{0,1}$-manifolds in the above sense not satisfying the so-called restricted cone property (see e.g.* [1, 2] *for definition) and since on the other hand the restricted cone property 'almost' implies that ∂G is a Lipschitz boundary, (see* [10] *for details), in view of the $C_{0,1}$-invariance of* (2.1) *we have achieved a generalization in this case, too.*

c) *It is easy to think of counter-examples to property* (2.1). *In* [17] *an explicit construction of linear independent elements in $\mathscr{H}^q_{R,\varepsilon_q}$ is given, which shows that e.g. a boundary of a domain $G \subset \mathbb{R}^3$ including an accumulation point of cavities or handles cannot have property* (2.1). *The corresponding β^q, $q \in \{1, 2\}$, will not be finite. This contradicts the fact that compactness of a sphere (here in $\mathscr{H}^q_{R,\varepsilon_q}$) is a criterion for finite dimensionality, (compare the argument in* [18]*).*

References

1. Adams, R.A.: Sobolev Spaces. New York: Academic Press 1975
2. Agmon, S.: Lectures on Elliptic Boundary Value Problems. Princeton: Van Nostrand 1965
3. Breinlinger, K.: Randwertaufgaben der Maxwellschen Gleichungen des $\mathbb{R}^n$ für inhomogene Medien. Bonner Math. Schriften Nr. 37, Univ. of Bonn 1969
4. Eidus, D.M.: Das Prinzip der Grenzabsorption. Mat. Sb. (N.S.) **58** (100), 65–86 (1962) (Russian)
5. Fichera, G.: Teoria assiomatica delle forme armoniche. Rendiconti di Matematica, Roma **20**, 147–171 (1961)
6. Friedrichs, K.O.: Differential Forms on Riemannian Manifolds. Comm. Pure Appl. Math. **8**, 551–590 (1955)
7. Gaffney, M.: The harmonic operator for exterior differential forms. Proceedings of the National Science, U.S.A., **37**, 48–50 (1951)
8. Hicks, N.J.: Notes on Differential Geometry, Princeton: Van Nostrand 1965
9. Kato, T.: Perturbation Theory for Linear Operators. Berlin-Heidelberg-New York: Springer 1966
10. König, H., Loch, H.: Eine abstrakte Theorie der Kegel- und Segmentbedingungen. Math. Meth. Appl. Sci. **3** (1981)
11. Morrey, C.B.: A Variational Method in the Theory of Harmonic Integrals II. Amer. Jour. Math. **78**, 137–170 (1956)
12. Morrey, C.B. Jr.: Multiple Integrals in the Calculus of Variations. Berlin-Heidelberg-New York: Springer 1966
13. Picard, R.: Ein Kompaktheitsresultat in der Theorie der verallgemeinerten Maxwell-Gleichung. SFB 72, Preprint Nr. 120, Univ. of Bonn 1977
14. Picard, R.: Zur Existenz des Wellenoperators bei Anfangsrandwertproblemen vom Maxwell-Typ. Math. Z. **156**, 175–185 (1980)
15. Picard, R.: Randwertaufgaben der verallgemeinerten Potentialtheorie. Math. Meth. Appl. Sci. **3**, 218–228 (1981)
16. Picard, R.: On Boundary Value Problems of Electro- and Magnetostatics. Proc. Roy. Soc. Edin. **92 A**, 165–174 (1982)
17. Picard, R.: Ein Hodge-Satz für Mannigfaltigkeiten mit nichtglattem Rand. Math. Meth. Appl. Sci. **5**, 153–161 (1983)
18. Picard, R.: Ein vereinheitlichter Zugang für eine Klasse linearer Wellenausbreitungs-Phänomene. Habilitation thesis, SFB 72, Preprint Nr. 489, Univ. of Bonn 1982
19. Rinkens, H.D.: Zur Theorie der Maxwellschen Gleichungen in der Ebene. Bonner Math. Schriften, Nr. 38, Univ. of Bonn 1969
20. Saranen, J.: On Electric and Magnetic Problems for Vector Fields in Anisotropic Non-homogeneous Media. J. Math. Anal. Appl. **91**, 254–275 (1983)
21. Weber, C.: A local compactness theorem for Maxwell's equations. Math. Meth. Appl. Sci. **2**, 12–25 (1980)

22. Weck, N.: Eine Lösungstheorie für die Maxwellschen Gleichungen auf Riemannschen Mannigfaltigkeiten mit nicht-glattem Rand. Habilitation thesis, Univ. of Bonn 1972
23. Weck, N.: Maxwell's Boundary Value Problem on Riemannian Manifolds with Nonsmooth Boundaries. J. Math. Anal. and Appl. **46**, 410–437 (1974)
24. Westenholz, C. von: Differential Forms in Mathematical Physics. Studies in Mathematics and its Applications, Amsterdam, North-Holland Publ. Comp. 1978
25. Weyl, H.: The Method of Orthogonal Projection in Potential Theory. Duke Math. J. **7**, 411–444 (1940)
26. Weyl, H.: Die natürlichen Randwertaufgaben im Außenraum für Strahlungsfelder beliebiger Dimension und beliebigen Ranges. Math. Z. **56**, 105–119 (1952)
27. Wilcox, C.H.: Scattering Theory for the D'Alembert Equation in Exterior Domains. Berlin-Heidelberg-New York: Springer 1975

Received November 4, 1983

Math. Z. 187, 165–170 (1984)

Mathematische Zeitschrift

Some Examples of Power Series with Non-Hadamard Gaps

Ch. Pommerenke*

Fachbereich Mathematik, Technische Universität, 1000 Berlin 12

1. Introduction and Results

Throughout this paper, we consider analytic functions f of the form

(1.1) $$f(z)=\sum_{k=0}^{\infty} a_k z^{n_k}, \qquad 0 \leqq n_0 < n_1 < \ldots,$$

in the unit disk $\mathbb{D}$. The Hadamard gap condition is

(1.2) $$\inf_k n_{k+1}/n_k > 1.$$

We study the question whether some complex analytic results that are known for Hadamard gaps remain true under weaker gap conditions.

Let first f have Hadamard gaps. Binmore [1] has proved

(1.3) f has a finite asymptotic value $\Rightarrow$ $a_k \to 0$ $(k \to \infty)$,

and Gnuschke and the author [3] have shown

(1.4) $\int_C |f'(z)|\,|dz| < \infty$ for some curve to $\partial\mathbb{D}$ $\Rightarrow$ $\sum_k |a_k| < \infty$;

the corresponding results for radii are due to Hardy and Littlewood [5] and to Zygmund [12]. An example of Rudin [10] shows that each of these implications actually characterizes Hadamard gaps:

If (n_k) is not an Hadamard sequence then there exist a_k such that

(1.5) $$\int_0^1 |f'(x)|\,dx < \infty, \qquad \limsup_{k\to\infty} |a_k| = \infty.$$

T. Murai [7] has proved, for Hadamard gaps, the Paley conjecture

(1.6) $$f(z) \neq 0 \ (z \in \mathbb{D}) \ \Rightarrow \ \sum_k |a_k| < \infty;$$

* This research was in part supported by the British SERC

this is also a consequence of (1.4). The next example shows that (1.2) cannot be replaced by a weaker condition of the quotient type.

Theorem 1. *Suppose that* (ε_k) *is decreasing to zero. Then there exist* (n_k) *with*

$$\frac{n_{k+1}}{n_k} > 1 + \varepsilon_k \quad \textit{for } k \geqq k_0 \tag{1.7}$$

and (a_k) *with* $a_k \to 0$ $(k \to \infty)$,

$$\sum_k |a_k|^q = \infty \quad \textit{for every positive } q, \tag{1.8}$$

such that

$$\operatorname{Re} f(z) > 0 \quad \textit{for } z \in \mathbb{D}. \tag{1.9}$$

Note that (1.7) implies

$$n_{k+1}/n_k > 1 + \varepsilon_{n_k} \quad (k \geqq k_0) \tag{1.10}$$

because $n_k \geqq k$. This condition was considered by Kennedy [6] in a different connection. We shall discuss the connection with Sidon sets in the final section.

T. Murai [8] has also proved

(1.11) (a_k) unbounded $\Rightarrow f$ has asymptotic value ∞ at every $\zeta \in \partial \mathbb{D}$

for Hadamard gaps; see [4] for another proof.

Theorem 2. *Suppose that* (ε_k) *decreases to zero. Then there exist* (n_k) *such that*

(i) $n_{k+1}/n_k > 1 + \varepsilon_k$ *for* $k \geqq k_0$,

(ii) (n_k) *is the union of four Hadamard sequences,*

and furthermore an unbounded sequence (a_k) *such that* f *does not have the asymptotic value* ∞ *at* 1.

I want to thank D. Gnuschke and W.K. Hayman for our discussions on these problems.

2. Construction of the Examples

Proof of Theorem 1. We may assume that $\varepsilon_\nu < 1$. For $\nu = 0, 1, \ldots$, we set

$$\eta_\nu = \varepsilon_\nu^{1/3}, \quad r_\nu = 1 - \eta_\nu^2, \quad p_\nu = \left[\frac{1}{\eta_\nu^3}\right] = \left[\frac{1}{\varepsilon_\nu}\right] \tag{2.1}$$

and choose m_ν such that

$$m_{\nu+1} \geqq 2 p_\nu m_\nu. \tag{2.2}$$

Since $\varepsilon_\nu \to 0$ and thus $\eta_\nu \to 0$ as $\nu \to \infty$, we can choose $q_\nu \in \mathbb{N}$ such that

$$q_\nu \to \infty, \quad \eta_\nu^{1/q_\nu} \to 0 \quad \text{as} \quad \nu \to \infty. \tag{2.3}$$

Finally we can find an infinite subset $I \subset \mathbb{N}$ such that

$$\sum_{\nu \in I} \eta_\nu^{1/q_\nu} < \infty. \tag{2.4}$$

We write $b_\nu = \eta_\nu^{1/q_\nu}$ for $\nu \in I$ and $b_\nu = 0$ otherwise. We define

$$\begin{aligned} f(z) &= \sum_{\nu=N}^{\infty} b_\nu \frac{1-(r_\nu z^{m_\nu})^{p_\nu+1}}{1-r_\nu z^{m_\nu}} \\ &= \sum_{\nu=N}^{\infty} b_\nu + \sum_{\nu=N}^{\infty} \sum_{j=1}^{p_\nu} b_\nu r_\nu^j z^{j m_\nu}, \end{aligned} \tag{2.5}$$

where N will be chosen below. The sum converges because of (2.4). We write f in the form (1.1); note that the exponents in the double sum are all distinct, by (2.2).

Let $k = 1, 2, \ldots$. There exists a unique ν such that

$$n_k = j m_\nu, \quad 1 \leqq j \leqq p_\nu \tag{2.6}$$

because of (2.5). If $j = p_\nu$ then $n_{k+1} = m_{\nu+1}$ and thus $n_{k+1}/n_k \geqq 2$ by (2.2); if $j < p_\nu$ then, by (2.6) and (2.1),

$$\frac{n_{k+1}}{n_k} = \frac{(j+1)m_\nu}{j m_\nu} = 1 + \frac{1}{j} \geqq 1 + \frac{1}{p_\nu} \geqq 1 + \varepsilon_\nu. \tag{2.7}$$

Since n_k is the k^{th} exponent, we have

$$k = p_1 + \ldots + p_{\nu-1} + j \geqq p_{[\nu/2]} + \ldots + p_{\nu-1}.$$

If ν is large then $p_{[\nu/2]} \geqq 3$ and thus $k \geqq 3(\nu-1)/2 > \nu$. Hence (2.7) shows that

$$\frac{n_{k+1}}{n_k} \geqq 1 + \varepsilon_\nu \geqq 1 + \varepsilon_k \quad \text{for } k \geqq k_0.$$

It follows from (2.1) that, for $z \in \mathbb{D}$,

$$\begin{aligned} \operatorname{Re} \frac{1 - r_\nu^{p_\nu} z^{p_\nu m_\nu}}{1 - r_\nu z^{m_\nu}} &\geqq \operatorname{Re} \frac{1}{1 - r_\nu z^{m_\nu}} - \frac{r_\nu^{p_\nu}}{1 - r_\nu} \\ &\geqq \frac{1}{2} - \frac{1}{\eta_\nu^2}(1-\eta_\nu^2)^{p_\nu} \\ &\geqq \frac{1}{2} - \frac{1}{\eta_\nu^2} \exp\left[-\frac{1}{2\eta_\nu}\right] > \frac{1}{4} \end{aligned}$$

for $\nu \geqq N$ if N is chosen large enough. Hence we see from (2.5) that $\operatorname{Re} f(z) > 0$ for $z \in \mathbb{D}$.

Let now $0 < q < \infty$. We choose $\nu \in I$ so large that $q_\nu > q$. Then it follows from (2.5), (2.6) and (2.1) that

$$\sum_{k=0}^{\infty} a_k^q \geqq \sum_{j=1}^{p_\nu} (b_\nu r_\nu^j)^q \geqq \eta_\nu \frac{1-r_\nu^{p_\nu+1}}{1-r_\nu^q}$$
$$\geqq \eta_\nu \frac{1-(1-\eta_\nu^2)^{1/\eta_\nu^3}}{q\eta_\nu^2} > \frac{1}{2q\eta_\nu}$$

if ν is large enough. Since $\eta_\nu \to 0$ as $\nu \to \infty$ it follows that (1.8) holds.

Proof of Theorem 2. We may assume that $\varepsilon_1 < 1/2$. We define η_ν recursively by $\eta_0 = 1$ and

$$\eta_\nu = \max(\tfrac{1}{2}\eta_{\nu-1}, (2\varepsilon_{4\nu})^{1/4}) \qquad (\nu = 1, 2, \ldots). \tag{2.8}$$

Then (η_ν) decreases to zero because the same holds for (ε_k). We set

$$p_\nu = [2^\nu \eta_\nu^2], \qquad q_\nu = [2^\nu/\eta_\nu^2] + 1 \quad (\nu = 0, 1, \ldots) \tag{2.9}$$

and consider the function

$$f(z) = \sum_{\nu=0}^{\infty} \frac{1}{\eta_\nu}(1-z^{p_\nu})(1-z^{2^\nu})z^{q_\nu} \tag{2.10}$$
$$= \sum_{\nu=0}^{\infty} \frac{1}{\eta_\nu}(z^{q_\nu} - z^{q_\nu+p_\nu} - z^{q_\nu+2^\nu} + z^{q_\nu+2^\nu+p_\nu}).$$

Rearranging this series in the form (1.1) we see that $a_k \to \infty$ as $k \to \infty$.

It follows from (2.9) and (2.10) that $n_{k+1}/n_k - 1$ has one of the following forms (as $\nu \to \infty$):

$$\frac{p_\nu}{q_\nu} \sim \eta_\nu^4, \qquad \frac{2^\nu - p_\nu}{q_\nu + p_\nu} \sim \eta_\nu^2, \qquad \frac{p_\nu}{q_\nu + 2^\nu} \sim \eta_\nu^4,$$

or finally

$$\frac{q_{\nu+1} - q_\nu - 2^\nu - p_\nu}{q_\nu + 2^\nu + p_\nu} \sim \frac{2\eta_\nu^2}{\eta_{\nu+1}^2} - 1 \geqq 1.$$

We conclude, by (2.8), that

$$\frac{n_{k+1}}{n_k} > 1 + \frac{1}{2}\eta_\nu^4 \geqq 1 + \varepsilon_{4\nu} \geqq 1 + \varepsilon_k$$

for $4\nu \leqq k \leqq 4\nu + 3$ if k is sufficiently large. Hence (i) is satisfied, and (ii) follows immediately from (2.9) and (2.10).

We write $\zeta_m = e^{2\pi i/2^m}$, $m = 1, 2, \ldots$. Since $|1-w| \leqq 1 - |w| + |\arg w| (|w| < 1)$, we see that, for $0 < r < 1$,

$$|1-(r\zeta_m)^{p_\nu}| \leqq 1 - r^{p_\nu} + \frac{2\pi p_\nu}{2^m} \leqq \left(1 - r + \frac{2\pi}{2^m}\right) p_\nu.$$

It follows that

$$|1-(r\zeta_m)^{p_\nu}|\,|1-(r\zeta_m)^{2^\nu}| \leqq \begin{cases} 2\left(1 - r + \dfrac{2\pi}{2^m}\right) p_\nu & \text{for } \nu < m, \\ 2(1-r^{2^\nu}) & \text{for } \nu \geqq m \end{cases} \tag{2.11}$$

because $\zeta_m^{2^\nu} = 1$ for $\nu \geqq m$.

Suppose now that $r_1 \leqq r < 1$ where $r_1 = \exp[-1/(\eta_1 q_1)]$. We choose j such that

$$\exp[-1/(\eta_j q_j)] \leqq r < \exp[-1/(\eta_{j+1} q_{j+1})] \tag{2.12}$$

and obtain from (2.10) and (2.11) that (empty sums being zero)

$$\begin{aligned}|f(r\zeta_m)| \leqq \sum_{\nu=0}^{m-1} \frac{2p_\nu}{\eta_\nu}\left(\frac{1}{\eta_j q_j} + \frac{2\pi}{2^m}\right) + \sum_{\nu=m}^{j} \frac{2^{\nu+1}}{\eta_\nu \eta_j q_j} \\ + \sum_{\nu=j+1}^{\infty} \frac{2}{\eta_\nu} \exp[-q_\nu/(\eta_{j+1} q_{j+1})].\end{aligned} \tag{2.13}$$

We see from (2.9) that the first two sums together are bounded by

$$\sum_{\nu=0}^{m-1} \frac{2\eta_\nu \eta_j}{2^{j-\nu}} + \sum_{\nu=0}^{m-1} \frac{4\pi\eta_\nu}{2^{m-\nu}} + \sum_{\nu=m}^{j} \frac{2\eta_j}{2^{j-\nu}\eta_\nu} \leqq 4 + 4\pi$$

because $\eta_\nu \geqq \eta_j$ for $\nu \leqq j$. It also follows from (2.9) that

$$\frac{q_\nu}{q_{j+1}} \geqq \frac{2^\nu}{\eta_\nu^2} \frac{\eta_{j+1}^2}{2 \cdot 2^{j+1}} \geqq 2^{\nu-j-2} \qquad \text{for } \nu \geqq j+1.$$

Since $\eta_\nu \geqq 2^{j-\nu}\eta_{j+1}$ $(\nu \geqq j+1)$ by (2.8) we conclude that the last sum in (2.13) is bounded by

$$\sum_{\nu=j+1}^{\infty} \frac{2^{\nu-j+1}}{\eta_{j+1}} \exp[-2^{\nu-j-2}/\eta_{j+1}] \leqq 8 \sum_{\nu=j+1}^{\infty} \exp[-2^{\nu-j-3}/\eta_{j+1}].$$

Thus it follows from (2.1) that, for $m = 1, 2, \ldots$ and $r_1 \leqq r < 1$,

$$|f(r\zeta_m)| \leqq M \equiv 4 + 4\pi + 8 \sum_{n=1}^{\infty} \exp[-2^{n-3}]$$

and thus also $|f(r\bar{\zeta}_m)| = |f(r\zeta_m)| \leqq M$. If we let $m \to \infty$ we obtain that $|f(r)| \leqq M$. Hence there is no (connected) curve $C \subset \mathbb{D} \cup \{1\}$ ending at 1 such that $|f(z)| \to \infty$ as $z \in C$, $z \to 1$.

3. Concluding Remarks

The sequence (n_k) is called a Sidon set if

$$f \in H^\infty \Rightarrow \sum_k |a_k| < \infty \tag{3.1}$$

where f has the form (1.1); see Rudin [9], or [2, p. 237]. The union of a finite number of Hadamard sequences is an important example. We see from (1.5) and Theorem 2 that the implications (1.3), (1.4) and (1.11) are not true for Sidon instead of Hadamard sets.

On the other hand, Theorem 1 is related to the theory of Sidon sets. If (n_k) is not a Sidon set then, by (3.1), we can construct a function $f \in H^\infty$ such that

$f(z) \neq 0$ $(z \in \mathbb{D})$ and $\sum |a_k| = \infty$. It remains open whether (1.6) is true for Sidon sets that are not Hadamard sets.

Zygmund [11] has proved

(3.2) $$f \in H^p \text{ for some } p>0 \;\Rightarrow\; \sum_k |a_k|^2 < \infty$$

if f has Hadamard gaps. Since a function of positive real part belongs to H^p for $0<p<1$, we see that the Hadamard condition (1.2) cannot be replaced by (1.7). To prove the last fact, we could have chosen any sequence satisfying (1.7) that contains arbitrarily long arithmetic progressions; see [9, Theorem 4.1] [2, p. 291].

The problem under what gap conditions the implication

(3.3) $$f(z) \neq 0 \quad (z \in \mathbb{D}) \;\Rightarrow\; (a_k) \text{ bounded}$$

holds is completely open. It follows from the theory of theta functions that

$$\sum_{k=0}^{\infty} (-1)^k (2k+1) z^{k(k+1)} = \prod_{n=1}^{\infty} (1-z^{2n})^3 \neq 0 \qquad \text{for } z \in \mathbb{D},$$

and it seems that no zero-free function with substantially larger gaps and unbounded coefficients is known.

References

1. Binmore, K.B.: Analytic functions with Hadamard gaps. Bull. London Math. Soc. **1**, 211–217 (1969)
2. Edwards, R.E.: Fourier series, vol. 2, 2nd ed. Berlin-Heidelberg-New York: Springer 1982
3. Gnuschke, D., Pommerenke, Ch.: On the absolute convergence of power series with Hadamard gaps. Bull. London Math. Soc. **15**, 507–512 (1983)
4. Gnuschke, D., Pommerenke, Ch.: On the growth of functions with Hadamard gaps. J. London Math. Soc., to appear
5. Hardy, G.H., Littlewood, J.E.: A further note on the converse of Abel's theorem. Proc. London Math. Soc. **25**, 219–236 (1926)
6. Kennedy, P.B.: Remark on a theorem of Zygmund. J. London Math. Soc. **33**, 71–72 (1958)
7. Murai, T.: The value-distribution of lacunary series and a conjecture of Paley. Ann. Inst. Fourier, Grenoble **31**, 135–156 (1981)
8. Murai, T.: The boundary behaviour of Hadamard lacunary series. Nagoya Math. J. **89**, 65–76 (1983)
9. Rudin, W.: Trigonometric series with gaps. J. Math. Mech. (Indiana) **9**, 203–227 (1960)
10. Rudin, W.: A converse of the high indices theorem. Proc. Amer. Math. Soc. **17**, 434–435 (1966)
11. Zygmund, A.: On the convergence of lacunary trigonometric series. Fund. Math. **16**, 90–107 (1930)
12. Zygmund, A.: On certain integrals. Trans. Amer. Math. Soc. **55**, 170–204 (1944)

Received November 14, 1983; received in final form March 7, 1984

Math. Z. 187, 171-183 (1984)

Mathematische Zeitschrift

On the Dirichlet Problem for Semi-Linear Elliptic Equation with L^2-Boundary Data

J. Chabrowski

University of Queensland, Department of Mathematics, 4067 St. Lucia, Brisbane, Australia

1. Introduction and Preliminaries

Let Q be a bounded domain in R^n with the boundary of class C^2. Let

$$Lu = -\sum_{i,j=1}^{n} D_i(a_{ij}(x) D_j u) + \sum_{i=1}^{n} b_i(x) D_i u + c(x) u$$

be a second order elliptic operator on Q, $a_{ij} = a_{ji}$ $(i, j = 1, \ldots, n)$. Let $x \in Q$, and let $r(x)$ denote the distance from x to the boundary ∂Q.

Throughout this paper we make the following assumption

(A) There exists a number $\gamma > 0$ such that

$$\gamma^{-1} |\xi|^2 \leqq \sum_{i,j=1}^{n} a_{ij}(x) \xi_i \xi_j \leqq \gamma |\xi|^2$$

for all $\xi \in R^n$ and $x \in Q$. Moreover we assume that b_i, $c \in L^\infty(Q)$ $(i = 1, \ldots, n)$, $a_{ij} \in C^1(Q)$ and $|D a_{ij}(x)| \leqq K r(x)^{-\alpha}$ on Q $(i, j = 1, \ldots, n)$, where K and α are positive constants, $0 < \alpha < 1$.

In this note we investigate the existence of a generalized solution in $W^{1,2}_{\mathrm{loc}}(Q)$ of the Dirichlet problem.

$$Lu = f(x, u, Du) \quad \text{in } Q \tag{1}$$

$$u = \phi \quad \text{on } \partial Q. \tag{2}$$

Here the boundary data ϕ is a function in $L^2(\partial Q)$. The function $f(x, s, p)$ defined for $x \in Q$, $s \in R$ and $p \in R^n$ satisfies the Caratheodory conditions, i.e.,

(i) for each $(s, p) \in R^{n+1}$, the function $x \to f(x, s, p)$ is measurable in Q,

(ii) for each $x \in Q$ (a.e.), the function $(s, p) \to f(x, s, p)$ is continuous on R^{n+1}.

We also assume that there exists a positive constant K_1 and a non-negative function $b \in L^2(Q)$ such that

(B) $$|f(x,s,p)| \leqq K_1(|s|+|p|)+b(x)$$

for all $(s,p) \in R^{n+1}$ and almost all $x \in Q$.

Under these assumptions the Niemytski mapping F given by $(Fu)(x) = f(x,u(x),Du(x))$ is bounded and continuous from $W^{1,2}_{\mathrm{loc}}(Q)$ to $L^2_{\mathrm{loc}}(Q)$.

We recall briefly that a function $u(x)$ is said to be a generalized solution of Eq. (1) if $u \in W^{1,2}_{\mathrm{loc}}(Q)$ and u satisfies

(3) $$\int_Q \left[\sum_{i,j=1}^{n} a_{ij} D_i u D_j v + \sum_{i=1}^{n} b_i D_i u \cdot v + cuv\right] dx = \int_Q F(u) v \, dx$$

for every $v \in W^{1,2}(Q)$ with compact support in Q.

We also recall that if the boundary function $\phi \in L^2(\partial Q)$ is a trace of some function ϕ_1 from $W^{1,2}(Q)$ then a generalized solution in $W^{1,2}(Q)$ of (1) is said to be a solution of the Dirichlet problem with the boundary condition (2) if and only if $u - \phi_1 \in \mathring{W}^{1,2}(Q)$. This definition is rather restrictive, because not every function in $L^2(\partial Q)$ is a trace of some function belonging to $W^{1,2}(Q)$. In connection with the results obtained in Sect. 2 it makes possible to consider the Dirichlet problem with a boundary data in $L^2(\partial Q)$. Sections 3, 4 and 5 are devoted to constructing a solution of the Dirichlet problem in $W^{1,2}_{\mathrm{loc}}(Q)$ with a boundary condition in $L^2(\partial Q)$. The proofs of the existence theorems are based on the energy estimate, given in Sect. 2, and the results of De Figueiredo [4] and Hess [7]. Finally we point out that in the case when ϕ is a trace of a function ϕ_1 from $W^{1,2}(Q)$, the transformation $v = u - \phi_1$ reduces the problem (1), (2) to the Dirichlet problem in $\mathring{W}^{1,2}(Q)$. The latter problem has an extensive literature (see for example [4–8] and [10] and historical references given there). For the Dirichlet problem with L^2-boundary data for linear elliptic equations we refer to papers [1–3, 14, 15] and [16].

2. Traces of Solutions in $W^{1,2}_{\mathrm{loc}}(Q)$ and the Energy Estimate

Before stating the main result of this section we begin with some definitions.

It follows from the regularity of the boundary ∂Q that there is a number $\delta_0 > 0$ such that for $\delta \in (0, \delta_0]$ the domain

$$Q_\delta = Q \cap \{x;\ \min_{y \in \partial Q} |\dot{x} - y| > \delta\}$$

with the boundary Q possesses the following property: to each $x_0 \in \partial Q$ there is a unique point $x_\delta(x_0) \in \partial Q_\delta$ such that

$$x_\delta(x_0) = x_0 - \delta \nu(x_0),$$

where $\nu(x_0)$ is the outward normal to ∂Q at x_0. The above relation gives a one-to-one mapping, of class C^1, of ∂Q onto ∂Q_δ.

According to Lemma 1 in [17] p. 382, the distance $r(x)$ belongs to $C^2(\bar{Q} - Q_{\delta_0})$ if δ_0 is sufficiently small. Denote by $\rho(x)$ the extension of the function $r(x)$ into $\bar{Q}$ satisfying the following properties $\rho(x) = r(x)$ for $x \in \bar{Q} - Q_{\delta_0}$,

$\rho\in C^2(\bar{Q})$, $\rho(x)\geqq\frac{3\delta_0}{4}$ in Q_{δ_0}, $\gamma_1^{-1}r(x)\leqq\rho(x)\leqq\gamma_1 r(x)$ in Q for some positive constant γ_1, $\partial Q_\delta=\{x;\rho(x)=\delta\}$ for $\delta\in(0,\delta_0]$ and finally $\partial Q=\{x;\rho(x)=0\}$.

We will use the surface integrals

$$M_1(\delta)=\int_{\partial Q} u(x_\delta(x))^2\,dS_x \quad \text{and} \quad M(\delta)=\int_{\partial Q_\delta} u(x)^2\,dS_x,$$

where $u\in W_{\mathrm{loc}}^{1,2}(Q)$ and the values of $u(x_\delta)$ on ∂Q and $u(x)$ on ∂Q_δ are understood in the sense of trace (see [9], Chap. 6).

Proposition 1 and Theorem 1 below have been established in [1] under stronger assumption $a_{ij}\in C^1(\bar{Q})$. In our situation the proofs are essentially the same and therefore they are omitted (see also the proofs of the corresponding results for linear equations Theorems 1 and 4 in [2]).

Proposition 1. *Let u be a solution of* (1) *belonging to $W_{\mathrm{loc}}^{1,2}(Q)$, then the following conditions are equivalent*

I $\qquad M(\delta)$ *is bounded on* $(0,\delta_0]$,

II $\qquad \int_Q |Du(x)|^2 r(x)\,dx<\infty$,

III $\qquad M_1(\delta)$ *is continuous on* $[0,\delta_0]$.

Theorem 1. *Let $u\in W_{\mathrm{loc}}^{1,2}(Q)$ be a solution of* (1) *such that one of the conditions* I, II *or* III *holds. Then there is a function ζ belonging to $L^2(\partial Q)$ such that $u(x_\delta)$ converges to ζ in $L^2(\partial Q)$.*

Theorem 1 plays an essential part in the ensuing treatment of the Dirichlet problem in $W_{\mathrm{loc}}^{1,2}(Q)$.

Consider now the equation

$$Lu+\lambda u=f(x,u,Du), \tag{1'}$$

where λ is a real parameter. Theorem 1 suggests the following definition of the Dirichlet problem in $W_{\mathrm{loc}}^{1,2}(Q)$.

Let $\phi\in L^2(\partial Q)$. A generalized solution u in $W_{\mathrm{loc}}^{1,2}(Q)$ of (1) (or (1')) is a solution of the Dirichlet problem with the boundary condition (2) if

$$\lim_{\delta\to 0}\int_{\partial Q}[u(x_\delta(x))-\phi(x)]^2\,dS_x=0.$$

Theorem 2. *Suppose $\phi\in L^2(\partial Q)$. Let u be a solution of the Dirichlet problem* (1'), (2) *in $W_{\mathrm{loc}}^{1,2}(Q)$. Then there exists positive constants λ_0, d and C independent of u such that*

$$\int_Q |D^2u(x)|^2 r(x)^3\,dx+\int_Q |Du(x)|^2 r(x)\,dx+\int_Q u(x)^2 r(x)\,dx+\sup_{0<\delta\leqq d} M(\delta)$$
$$\leqq C(1+\lambda^2)\Big[\int_{\partial Q}\phi(x)^2\,dS_x+\int_Q b(x)^2\,dx\Big]$$

for all $\lambda\geqq\lambda_0$.

Proof. We follow the proof of Lemma 1 in [3] (see also the proof of Theorem 5 in [2]).

Put

$$v(x)=\begin{cases} u(x)(\rho(x)-\delta) & \text{for } x\in Q_\delta, \\ 0 & \text{for } x\in Q-Q_\delta, \end{cases}$$

where $0<\delta<\delta_0$. It is clear that v is an admissible test function in the corresponding integral equation defining a generalized solution of (1′) and that

$$\begin{aligned} (4)\qquad & \int_{Q_\delta}\sum_{i,j=1}^{n} a_{ij}D_iuD_ju(\rho-\delta)\,dx+\int_{Q_\delta}\sum_{i,j=1}^{n} a_{ij}D_iu\cdot uD_j\rho\,dx \\ &\quad +\int_{Q_\delta}\sum_{i=1}^{n} b_iD_iu\cdot u(\rho-\delta)\,dx \\ &\quad +\int_{Q_\delta} cu^2(\rho-\delta)\,dx+\lambda\int_{Q_\delta} u^2(\rho-\delta)\,dx \\ &= \int_{Q_\delta} f(x,u,Du)u(\rho-\delta)\,dx. \end{aligned}$$

By Green's formula we have

$$\begin{aligned} (5)\qquad \int_{Q_\delta}\sum_{i,j=1}^{n} a_{ij}D_iu\cdot uD_j\rho\,dx &= -\tfrac{1}{2}\int_{\partial Q_\delta}\sum_{i,j=1}^{n} a_{ij}D_i\rho D_j\rho u^2\,dx \\ &\quad -\tfrac{1}{2}\int_{Q_\delta}\sum_{i,j=1}^{n} D_i(a_{ij}D_j\rho)u^2\,dx. \end{aligned}$$

Using Hölder's inequality we deduce from the assumptions (A) and (B) that

$$\begin{aligned} (6)\qquad & \int_{Q_\delta}|Du|^2(\rho-\delta)\,dx+\lambda\int_{Q_\delta}u^2(\rho-\delta)\,dx \\ &\leqq C_1[\int_{\partial Q_\delta}u^2\,dS_x+\int_{Q_\delta}u^2\,dx+\int_{Q_\delta}u^2\rho^{-\alpha}\,dx+\int_{Q_\delta}b^2\,dx, \end{aligned}$$

where C_1 is a positive constant independent of u and δ. Similarly from (4) and (5) we have

$$\begin{aligned} (7)\qquad \int_{\partial Q_\delta}u^2\,dS_x &\leqq C_2[\int_{Q_\delta}|Du|^2(\rho-\delta)\,dx+\lambda\int_{Q_\delta}u^2(\rho-\delta)\,dx \\ &\quad +\int_{Q_\delta}u^2\,dx+\int_{Q_\delta}u^2\rho^{-\alpha}\,dx+\int_{Q_\delta}b^2\,dx]. \end{aligned}$$

By the boundary condition (2), $\int_{\partial Q_\delta}u^2\,dS_x$ is bounded on $(0,\delta_0]$ and consequently $\int_{Q_\delta}u^2\rho^{-\alpha}\,dx$ is bounded independently of δ (see Lemma 5 in [1]).

On the other hand, we have

$$\int_{\partial Q_\delta}u(x)^2\,dS_x=\int_{\partial Q}u(x_\delta)^2\mathbf{J}_{x_\delta}\,dS_x,$$

where $\mathbf{J}_{x_\delta}$ denotes the Jacobian of the mapping x_δ and $\lim_{\delta\to 0}\mathbf{J}_x=1$ uniformly on ∂Q (see [15]).

Hence letting $\delta \to 0$ we deduce from (6) that

$$\int_Q |Du|^2 \rho\, dx + \lambda \int_Q u^2 \rho\, dx \leqq C_1 \left[\int_Q \phi^2 dS_x + \int_Q u^2 dx + \int_Q u^2 \rho^{-\alpha} dx + \int_Q b^2 dx \right]. \tag{8}$$

It follows from (7) and (8) that

$$\sup_{0<\delta\leqq d} \int_{\partial Q_\delta} u^2 dS_x \leqq \tilde{C}_2 \left[\int_{\partial Q} \phi^2 dx + \int_Q u^2 dx + \int_Q u^2 \rho^{-\alpha} dx + \int_Q b^2 dx \right]. \tag{9}$$

Now

$$\int_Q u^2 \rho^{-\mu} dx \leqq \frac{d^{1-\mu}}{1-\mu} \sup_{0<\delta\leqq d} \int_{\partial Q_\delta} u^2 dS_x + \frac{1}{m_d^{\mu+1}} \int_Q u^2 \rho\, dx, \tag{10}$$

where $m_d = \inf_{Q_d} \rho(x)$. Taking λ_0 sufficiently large and d sufficiently small we deduce from (8), (9) and (10) that

$$\int_Q |Du|^2 \rho\, dx + \int_Q u^2 \rho\, dx + \sup_{0<\delta\leqq d} M(\delta) \leqq C_3 \left[\int_{\partial Q} \phi^2 dS_x + \int_Q b^2 dx \right]. \tag{11}$$

To proceed further we note that under our assumptions $u \in W^{2,2}_{\text{loc}}(Q)$. Let v be a function in $W^{2,2}(Q)$ with compact support in Q. Taking $D_k v$ as a test function and integrating by parts we obtain

$$\begin{aligned} \int_Q \sum_{i,j=1}^n a_{ij} D_{ik} u D_j v\, dx &+ \int_Q \sum_{i,j=1}^n D_k a_{ij} D_i u D_j v\, dx \\ &- \int_Q \sum_{i=1}^n b_i D_i u D_k v\, dx - \int_Q (c+\lambda) u D_k v\, dx \\ &= -\int_Q f(x, u, Du) D_k v\, dx. \end{aligned}$$

Now put

$$v(x) = \begin{cases} D_k u(x)(\rho(x)-\delta)^3 & \text{for } x \in Q_\delta, \\ 0 & \text{for } x \in Q - Q_\delta \end{cases}$$

thus

$$\begin{aligned} &\int_{Q_\delta} \sum_{i,j=1}^n a_{ij} D_{ik} u D_{jk} u (\rho-\delta)^3 dx + 3 \int_{Q_\delta} \sum_{i,j=1}^n a_{ij} D_{ik} u D_k u (\rho-\delta)^2 D_j \rho\, dx \\ &\quad + 3 \int_{Q_\delta} \sum_{i,j=1}^n D_k a_{ij} D_i u D_k u (\rho-\delta)^2 D_j \rho\, dx + \int_{Q_\delta} \sum_{i,j=1}^n D_k a_{ij} D_i u D_{jk} u (\rho-\delta)^3 dx \\ &\quad - \int_{Q_\delta} \sum_{i=1}^n b_i D_i u D_{kk} u (\rho-\delta)^3 dx - 3 \int_{Q_\delta} \sum_{i=1}^n b_i D_i u D_k u (\rho-\delta)^2 D_k \rho\, dx \\ &\quad - \int_{Q_\delta} (c+\lambda) u D_{kk} u (\rho-\delta)^3 dx - 3 \int_{Q_\delta} (c+\lambda) u D_k u (\rho-\delta)^2 D_k \rho\, dx \\ &= -\int_{Q_\delta} f(x, u, Du) D_{kk} u (\rho-\delta)^3 dx - 3 \int_{Q_\delta} f(x, u, Du) D_k u D_k \rho (\rho-\delta)^2 dx. \end{aligned}$$

Applying the ellipticity condition to the first integral on the left hand side and Young's inequality, the assumptions (A) and (B) to the remaining integrals we easily arrive at the estimate

$$\int_Q |D^2u|^2 \rho^3 dx \leqq C_4\left(\int_Q |Du|^2 \rho\, dx + \int_Q u^2 \rho\, dx + \lambda^2 \int_Q u^2 \rho\, dx\right), \tag{12}$$

where C_4 is a positive constant independent of u. Combining (11) and (12) we obtain the full energy estimate.

To proceed further let us introduce the Hilbert space $\tilde{W}^{1,2}(Q)$ of all functions in $W^{1,2}_{\mathrm{loc}}(Q)$ such that

$$\|u\|^2_{\tilde{W}^{1,2}(Q)} = \int_Q u(x)^2\, dx + \int_Q |Du(x)|^2 r(x)\, dx < \infty.$$

Similarly $\tilde{W}^{2,2}(Q)$ denotes the space of all functions in $W^{2,2}_{\mathrm{loc}}(Q)$ such that

$$\|u\|^2_{\tilde{W}^{2,2}(Q)} = \int_Q |D^2u(x)|^2 r(x)^3\, dx + \|u\|^2_{\tilde{W}^{1,2}(Q)} < \infty.$$

It is obvious that

$$\|u\|^2_{\tilde{W}^{2,2}(Q)} \leqq \max(1, d, d^{-1})\Big[\int_Q |D^2u(x)|^2 r(x)^3\, dx + \int_Q |Du(x)|^2 r(x)\, dx + \int_Q u(x)^2 r(x)\, dx + \sup_{0<\delta\leqq d} M(\delta)\Big].$$

Proposition 2. *Let $\{u_m\}$ be a bounded sequence in $\tilde{W}^{2,2}(Q)$. Then there exist $u \in \tilde{W}^{2,2}(Q)$ and subsequence $\{u_{m_\nu}\}$ such that*

(a) $\lim_{\nu\to\infty} u_{m_\nu} = u \quad$ *in* $L^2(Q)$,

(b) $\lim_{\nu\to\infty} \int_Q |Du_{m_\nu}(x) - Du(x)|\, r(x)^{\frac{1}{2}}\, dx = 0$,

(c) $\lim_{\nu\to\infty} \int_Q |f(x, u_{m_\nu}(x), Du_{m_\nu}(x)) - f(x, u(x), Du(x))|\, r(x)^{\frac{1}{2}}\, dx = 0$.

Proof. First we choose a subsequence, denoted again by $\{u_m\}$, converging weakly in $\tilde{W}^{2,2}(Q)$ to a function u. By Theorem 4.11 in [12], $\tilde{W}^{1,2}(Q)$ is compactly embedded in $L^2(Q)$ and therefore we may assume that u_m converges to u in $L^2(Q)$ and a.e. in Q. We note that for every $0<\delta\leqq\delta_0$, u_m is bounded in $W^{2,2}(Q_\delta)$. Hence given $0<\delta\leqq\delta_0$, by Rellich's theorem (see [13], p. 75) we can select a subsequence $\{u_{m_\delta}\}$ such that $\lim_{m_\delta\to\infty} u_{m_\delta} = u$ in $W^{1,2}(Q_\delta)$. Let $\{\delta_s\}$ be a sequence in $(0,\delta_0)$ converging to 0 as $s\to\infty$. It is obvious that there exists a subsequence $\{u_{m_\nu}\}$ of $\{u_m\}$ such that $\lim_{\nu\to\infty} u_{m_\nu} = u$ in $W^{1,2}(Q_{\delta_s})$ for every s. To prove (b) observe that

$$\int_Q |Du_{m_\nu} - Du|\, \rho(x)^{\frac{1}{2}}\, dx \leqq \int_{Q_{\delta_s}} |Du_{m_\nu} - Du|\, \rho(x)^{\frac{1}{2}}\, dx + [\mathrm{meas}(Q - Q_{\delta_s})]^{\frac{1}{2}} \left(\sup_\nu \|Du_{m_\nu}\rho^{\frac{1}{2}}\|_{L^2} + \|Du\rho^{\frac{1}{2}}\|_{L^2}\right)$$

and the assertion (b) easily follows and we also may assume that $\lim_{\nu\to\infty} D_i u_{m_\nu}(x) = D_i u(x)$ $(i=1,\dots,m)$ a.e. in Q. Finally it follows from (b) and (a) that there exist non-negative functions $v(x)$ and $z(x)$ such that

$$\int_Q v(x)\, r(x)^{\frac{1}{2}}\, dx < \infty \quad \text{and} \quad \int_Q z(x)^2\, dx < \infty,$$
$$|Du_{m_\nu}(x)| \leqq v(x) \quad \text{and} \quad |u_{m_\nu}(x)| \leqq z(x)$$

for all ν and almost everywhere in Q. Now (c) follows from (B) and the Lebesgue Dominated Convergence Theorem.

We are now in a position to establish the existence of a solution in $W^{1,2}_{\text{loc}}(Q)$ to the Dirichlet problem (1'), (2).

Theorem 3. *Let $\phi \in L^2(\partial Q)$. There exists a positive constant $\tilde{\lambda}$ such that for every $\lambda \geqq \tilde{\lambda}$ the Dirichlet problem* (1'), (2) *admits a solution in $W^{1,2}_{\text{loc}}(Q)$.*

Proof. Let $\{\phi_m\}$ be a sequence in $C^1(\partial Q)$ converging to ϕ in $L^2(\partial Q)$. It is easy to see that the Dirichlet form on $W^{1,2}(Q) \times W^{1,2}(Q)$ defined by

$$a(u, v) = \int_Q \left[\sum_{i,j=1}^n a_{ij} D_i u D_j v + \sum_{i=1}^n b_i D_i v + (c+\lambda) uv - f(x, u, Du) v \right] dx$$

is coercive, that is $\displaystyle\lim_{\|v\|_{W^{1,2}(Q)} \to \infty} \frac{a(v, v)}{\|v\|_{W^{1,2}(Q)}} = \infty$, provided λ is sufficiently large, say $\lambda \geqq \bar{\lambda}$, (see Theorem 2.1, p. 171 in [11]). Hence for every m the Dirichlet problem

$$Lu + \lambda u = f(x, u, Du) \quad \text{in } Q, \qquad u = \phi_m \quad \text{on } \partial Q$$

admits a solution u_m in $W^{2,2}_{\text{loc}}(Q) \cap W^{1,2}(Q)$. Let $\tilde{\lambda} = \max(\lambda_0, \bar{\lambda})$. By Theorem 2 for every $\lambda \geqq \tilde{\lambda}$ we have

$$\int_Q |D^2 u_m|^2 r^3\, dx + \int_Q |Du_m|^2 r\, dx + \int_Q u_m^2 r\, dx + \sup_{0<\delta \leqq d} M(\delta) \leqq C\Big(\int_{\partial Q} \phi_m^2\, dS_x + \int_Q b^2\, dx\Big)$$

for $m = 1, 2, \ldots$. Hence $\{u_m\}$ is a bounded sequence in $\tilde{W}^{2,2}(Q)$. By Proposition 2 we can select a subsequence $\{u_{m_\nu}\}$ converging weakly in $\tilde{W}^{2,2}(Q)$ to a function u and having properties (a), (b) and (c). Since u_{m_ν} satisfies the equation (1') we have

$$\int_Q \left[\sum_{i,j=1}^n a_{ij} D_i u_{m_\nu} D_j \Psi + \sum_{i=1}^n b_i D_i u_{m_\nu} \Psi + (c+\lambda) u_{m_\nu} \Psi \right] dx = \int_Q f(x, u_{m_\nu}, Du_{m_\nu}) \Psi\, dx$$

for all $\Psi \in C_0^1(Q)$. Letting $\nu \to \infty$, it follows from (a), (b), (c) and the density of $C_0^1(Q)$ in the space of all functions in $W^{1,2}(Q)$ with compact supports in Q that u is a generalized solution of (1'). Since $\int_Q |Du|^2 r\, dx < \infty$, by Theorem 1 there exists a function ζ in $L^2(\partial Q)$ such that $\lim_{\delta \to 0} u(x_\delta) = \zeta$ in $L^2(\partial Q)$. We now show that $\zeta = \phi$ a.e. on ∂Q. Let Ψ be an arbitrary function in $C^1(\bar{Q})$. Put

$$v(x) = \begin{cases} \Psi(x)(\rho(x) - \delta) & \text{for } Q_\delta, \\ 0 & \text{for } Q - Q_\delta. \end{cases}$$

Using v a test function in the integral equation defining a generalized solution of (1'), applying Green's theorem and letting $\delta \to 0$ we obtain

$$\int_{\partial Q} \phi_{m_\nu} \Psi \sum_{i,j=1}^{n} a_{ij} D_i \rho D_j \rho \, dx$$
$$= \int_Q \Big[-\sum_{i,j=1}^{n} D_i(a_{ij} \Psi D_j \rho) u_{m_\nu} + \sum_{i,j=1}^{n} a_{ij} D_i u_m D_j \Psi \rho$$
$$+ \sum_{i=1}^{n} b_i D_i u_{m_\nu} \Psi \rho + (c+\lambda) u_{m_\nu} \Psi \rho - f(x, u_{m\nu}, D u_{m_\nu}) \Psi \rho \Big] dx.$$

Similarly

$$\int_{\partial Q} \zeta \Psi \sum_{i,j=1}^{n} a_{ij} D_i \rho D_j \rho \, dx$$
$$= \int_Q \Big[-\sum_{i,j=1}^{n} D_i(a_{ij} \Psi D_j \rho) u + \sum_{i,j=1}^{n} a_{ij} D_i u D_j \Psi \rho$$
$$+ \sum_{i=1}^{n} b_i D_i u \Psi \rho + (c+\lambda) u \Psi \rho - f(x, u, Du) \Psi \rho \Big] dx.$$

Now using Proposition 2 we deduce from the last two equations that

$$\int_{\partial Q} \zeta \Psi \sum_{i,j=1}^{n} a_{ij} D_i \rho D_j \rho \, dS_x = \int_{\partial Q} \zeta \Psi \sum_{i,j=1}^{n} a_{ij} D_i \rho D_j \rho \, dS_x,$$

for every $\Psi \in C^1(\bar{Q})$ and this completes the proof.

3. The Dirichlet Problem when there is no Resonance

In this section we need following assumption on f (see [4]).

($\mathrm{A_N}$) Let $\lambda_1 < \lambda_2$ be two consecutive eigenvalues of L. Let μ and κ be such that $\lambda_1 < \mu < \lambda_2$ and $0 < \kappa < \min(\mu - \lambda_1, \mu - \lambda)$. We suppose that there exist positive constants C_1 and $0 < k < 1$ and a function $\bar{b} \in L^2(Q)$ such that

$$|f(x, s, p) - \mu s| \leqq \kappa |s| + C_1(|s|^k + |p|^k) + \bar{b}(x)$$

for all $(s, p) \in Q \times R^{n+1}$ and almost all $x \in Q$.

Since Q is bounded set, ($\mathrm{A_N}$) implies (B).

Theorem 4. *Let $Q \in L^2(\partial Q)$. There exists at least one solution in $W^{1,2}_{\mathrm{loc}}(Q)$ to the Dirichlet problem* (1), (2).

Proof. Let $\lambda > \tilde{\lambda}$. Then by Theorem 3 there exists a solution $u \in W^{1,2}_{\mathrm{loc}}(Q)$ to the problem (1′), (2). Now consider the Dirichlet problem in $\mathring{W}^{1,2}(Q)$.

$$Lv = F(x, v, Dv) \quad \text{in } Q, \tag{11}$$
$$v = 0 \quad \text{on } \partial Q, \tag{12}$$

where

$$F(x, s, p) = -\lambda u(x) + f(x, u(x), Du(x)) + f(x, s - u(x), p - Du(x)).$$

By the assumption (A_N) we have

$$|F(x,s,p)-\mu s| \leqq \kappa |s| + 2^k C_1(|p|^k + |s|^k) + B(x),$$

where

$$B(x) = 2\bar{b}(x) + (\kappa+\mu+\lambda)|u(x)| + C_1(1+2^k)(|u(x)|^k + |Du(x)|^k).$$

Since $\int_Q |Du|^2 r\,dx < \infty$ and $0<k<1$, we verify using Holder's inequality that

$$\int_Q |Du|^{2k} dx \leqq \Big(\int_Q |Du|^2 \rho\, dx\Big)^{\frac{k}{2}} \Big(\int_Q \rho^{-\frac{k}{2-k}} dx\Big)^{\frac{2-k}{2}} < \infty.$$

Consequently $B \in L^2(Q)$. It follows from Theorem 1 in [4] that the Dirichlet problem (11), (12) admits a solution $v \in \mathring{W}^{1,2}(Q)$. It is easy to see that $v-u$ is a solution in $W^{1,2}_{\mathrm{loc}}(Q)$ to the Dirichlet problem (1), (2).

4. The Dirichlet Problem at Resonance

Following De Figueiredo [4] we introduce the assumption

(A_R) Let λ_1 be an eigenvalue of the operator L. Suppose that there exist constants C_1, $0 \leqq k < 1$ and a non-negative function $d \in L^2(Q)$ such that

$$|f(x,s,p) - \lambda_1 s| \leqq d(x) + C_1(|s|^k + |p|^k)$$

for all $(s,p) \in R^{n+1}$ and almost all $x \in Q$.

In Theorem 5 below, using the idea originating in [8] (see also [5] and [6]) we approximate the Eq. (1) by nonresonant equations of the type

$$Lu - \frac{1}{m}u = f(x,u,Du) \qquad m=1,2,\ldots. \tag{13}$$

Theorem 5. *Suppose that $\phi \in L^2(\partial Q)$. Then the Dirichlet problem* (13), (2) *in $W^{1,2}_{\mathrm{loc}}(Q)$ has a solution provided m is sufficiently large. Moreover, either*

(i) *$\{\|u_m\|_{\tilde{W}^{1,2}}\}$ is a bounded sequence and then the Dirichlet problem* (1), (2) *in $W^{1,2}_{\mathrm{loc}}(Q)$ admits a solution, or*

(ii) *there is a subsequence of $\{\|u_m\|_{\tilde{W}^{1,2}}\}$, which we denote also by $\{\|u_m\|_{\tilde{W}^{1,2}}\}$, tending to ∞ as $m \to \infty$. In this case, there exists a subsequence of $v_m = u_m \cdot \|u_m\|^{-1}_{\tilde{W}^{1,2}}$ converging in $\tilde{W}^{1,2}$ to a non-zero λ_1-eigenfunction $v \in \mathring{W}^{1,2}(Q)$ of the operator L.*

Proof. Let λ_2 be the first eigenvalue for L larger than λ_1. Then $\lambda_1 - \frac{1}{m}$ and $\lambda_2 - \frac{1}{m}$ are consecutive eigenvalues of the eigenvalue problem

$$Lu - \frac{1}{m}u = \mu u \quad \text{in } Q, \qquad u=0 \quad \text{on } Q$$

in $\mathring{W}^{1,2}(Q)$. It is clear that $\lambda_1-\frac{1}{m}<\lambda_1<\lambda_2-\frac{1}{m}$ for m sufficiently large and the assumption (A_N) holds with $\mu=\lambda_2$ and $\kappa=0$. The existence of a solution in $W^{1,2}_{\text{loc}}(Q)$ of the Dirichlet problem (13), (2) follows from Theorem 4.

If the sequence $\{\|u_m\|_{\tilde{W}^{1,2}}\}$ is bounded, it follows from the proof of Theorem 2 that

$$\|D^2 u_m \rho^{\frac{3}{2}}\|_{L^2(Q)} \leqq \text{Const}\, \|u_m\|_{\tilde{W}^{1,2}} \tag{14}$$

(see (12)). Hence the sequence $\{u_m\}$ is bounded in $\tilde{W}^{2,2}(Q)$ and the result follows from Proposition 2 (see the argument of Theorem 3).

Finally assume that $\lim_{m\to\infty} \|u_m\|_{\tilde{W}^{1,2}}=\infty$.

By Theorem 4.11 in [12] there exists a subsequence of $v_m=u_m\|u_m\|^{-1}_{\tilde{W}^{1,2}}$, denote it again by v_m, such that $v_m\to v$ weakly in $\tilde{W}^{1,2}(Q)$, $v_m\to v$ in $L^2(Q)$ and a.e. in Q as $m\to\infty$. Note that

$$\begin{aligned}\int_Q \sum_{i,j=1}^n a_{ij} D_i v_m D_j z\,dx + \int_Q \sum_{i=1}^n b_i D_i v_m z\,dx + \int_Q \left(c-\frac{1}{m}-\lambda_1\right) v_m z\,dx \\ = \|u_m\|^{-1}_{\tilde{W}^{1,2}} \int_Q [f(x,u_m,Du_m)-\lambda_1 u_m] z\,dx\end{aligned} \tag{15}$$

for every $z\in W^{1,2}(Q)$ with compact support in Q. By virtue of the assumption (A_R)

$$\lim_{m\to\infty} \|u_m\|^{-1}_{\tilde{W}^{1,2}} \int_Q [f(x,u_m,Du_m)-\lambda_1 u_m] z\,dx=0. \tag{16}$$

Hence letting $m\to\infty$ in (15) we see that v is a solution in $\tilde{W}^{1,2}(Q)$ of the equation

$$Lv-\lambda_1 v=0 \quad \text{in } Q. \tag{17}$$

It then follows from (15) and (17) that

$$\begin{aligned}\int_Q \sum_{i,j=1}^n a_{ij} D_i(v_m-v) D_j z\,dx + \int_Q \sum_{i=1}^n b_i D_i(v_m-v) z\,dx \\ + \int_Q (c-\lambda_1)(v_m-v) z\,dx - \frac{1}{m}\int_Q v_m z\,dx \\ = \|u_m\|^{-1}_{\tilde{W}^{1,2}} \int_Q [f(x,u_m,Du_m)-\lambda_1 u_m] z\,dx\end{aligned} \tag{18}$$

for every $z\in W^{1,2}(Q)$ with compact support. Let w be an arbitrary function in $C^1(\bar{Q})$ and take

$$z(x)=\begin{cases} w(x)(\rho(x)-\delta) & \text{on } Q_\delta, \\ 0 & \text{on } Q-Q_\delta, \end{cases}$$

in the integral identity (18). Since $v\in\tilde{W}^{1,2}(Q)$ is a solution of (17), it has a boundary value $\zeta\in L^2(\partial Q)$ in the sense of L^2-convergence (see Theorem 1). Hence by Green's theorem we obtain

$$\int_Q \sum_{i,j=1}^n a_{ij} D_i(v_m - v) w D_j\rho\, dx = -\int_{\partial Q_\delta} (v_m - v) w \sum_{i,j=1}^n a_{ij} D_i\rho D_j\rho\, dS_x$$
$$- \int_{Q_\delta} \sum_{i,j=1}^n D_i(a_{ij} D_j\rho \cdot w)(v_m - v)\, dx$$

and letting $\delta \to 0$ we get

$$\int_Q \sum_{i,j=1}^n a_{ij} D_i(v_m - v) w D_j\rho\, dx = -\int_{\partial Q} (\phi \|u_m\|_{\tilde{W}^{1,2}}^{-1} - \zeta) w \sum_{i,j=1}^n a_{ij} D_i\rho D_j\rho\, dS_x \tag{19}$$
$$- \int_Q \sum_{i,j=1}^n D_i(a_{ij} D_j\rho\, w)(v_n - v)\, dx.$$

Now it is clear that letting $\delta \to 0$ and $m \to \infty$ we deduce from (16), (18) and (19) that

$$\int_{\partial Q} w\zeta \sum_{i,j=1}^n a_{ij} D_i\rho D_j\rho\, dS_x = 0$$

for every $w \in C^1(\bar{Q})$ and consequently $\zeta = 0$ a.e. on ∂Q. (Here we have used (16) with $z = w\rho$.) Therefore v is a solution in $\tilde{W}^{1,2}(Q)$ of (17) with the boundary condition

$$\lim_{\delta \to 0} \int_{\partial Q} v(x_\delta(x))^2\, dS_x = 0.$$

That is $v \in \mathring{W}^{1,2}(Q)$. To show that $v_m \to v$ in $\tilde{W}^{1,2}(Q)$ we take

$$z(x) = \begin{cases} (v_m(x) - v(x))(\rho(x) - \delta) & \text{on } Q_\delta, \\ 0 & \text{on } Q - Q_\delta, \end{cases}$$

as a test function in (18). Applying the ellipticity condition, Young inequality and letting $\delta \to 0$ we obtain

$$\int_Q |Dv_m - Dv|^2 \rho\, dx \leqq C[\|u_m\|_{\tilde{W}^{1,2}}^{-1} \int_Q |f(x, u_m) Du_m) - \lambda_1 u_m|\, dx$$
$$+ \|u_m\|_{\tilde{W}^{1,2}}^{-1} \int_{\partial Q} \phi^2\, dS_x + \int_Q |v_m - v|^2\, dx],$$

where C is a positive constant independent of m. Then convergence $v_m \to v$ in $\tilde{W}^{1,2}(Q)$ and the fact that $\|v_m\|_{\tilde{W}^{1,2}} = 1$ for all m, imply that $v \not\equiv 0$.

To prove the existence theorem in the resonant case we need more restrictive assumption on f. For simplicity we consider the case when the nonlinearity f is independent of Du.

Theorem 6. *Let $\phi \in L^2(\partial Q)$. Suppose that the assumption* (A_R) *holds and that there exist functions $h_+ \in L^{\frac{2}{1-k}}(Q)$ and $h_- \in L^{\frac{2}{1-k}}(Q)$ such that*

$$\lim_{s \to \pm\infty} \frac{f(x, s) - \lambda_1 s}{|s|^k} = h_\pm(x) \tag{20}$$

for almost all $x \in Q$ and for all eigenfunctions $v \in \mathring{W}^{1,2}(Q)$ of L corresponding to λ_1 we have

$$\int_{v>0} h_+ |v|^{1+k}\, dx - \int_{v<0} h_- |v|^{1+k}\, dx > 0.$$

Then there exists a solution in $W^{1,2}_{\mathrm{loc}}(Q)$ *to the Dirichlet problem* (1), (2).

The proof follows lines of the proof of Theorem 4. One can construct a solution in the form $v-u$, where u and v are solutions to the problems (1′), (2) and (11), (12) in $W^{1,2}_{\mathrm{loc}}(Q)$ and $\mathring{W}^{1,2}_{\mathrm{loc}}(Q)$ respectively. We only point out that the right hand side of the Eq. (11) takes the form

$$F(x,s)=-\lambda u(x)+f(x,u(x))+f(x,s-u(x))$$

and since $0<k<1$, it is easy to see that F satisfies ($\mathrm{A_R}$) and (20) and hence Theorem 5-3 in [4] is applicable.

5. The Dirichlet Problem with a Monotone Non-Linearity

In Theorem 7 below we make use of a upper and lower solution associated with (1) (for definitions see [7]). We assume that $f(x,s,p)$ is independent of p.

Theorem 7. *Let* $\phi\in L^2(\partial Q)$ *and let* $c(x)\geqq 0$ *on* Q. *If* $f(x,s)$ *is non-increasing in* s *for almost all* $x\in Q$, *then the Dirichlet problem* (1), (2) *admits at least one solution in* $W^{1,2}_{\mathrm{loc}}(Q)$.

Proof. Let u be a solution to the problem (1′), (2) in $W^{1,2}_{\mathrm{loc}}(Q)$ $(\lambda\geqq\tilde\lambda)$. We now consider the problem in $\mathring{W}^{1,2}(Q)$:

$$Lv=-\lambda u+f(x,u)-f(x,u-v)\quad \text{in } Q, \tag{21}$$

$$v=0\quad \text{on } \partial Q. \tag{22}$$

Let $\Psi\in\mathring{W}^{1,2}(Q)$ be a solution to the problem

$$L\Psi=\lambda|u|\quad\text{in } Q,\qquad \Psi=0\quad\text{on } \partial Q.$$

By the maximum principle $\Psi\geqq 0$ and hence

$$L\Psi=\lambda|u|\geqq-\lambda u+f(x,u)-f(x,u-\Psi),$$

that is, Ψ is an upper solution of problem (21), (22). Similarly a solution $\phi\in\mathring{W}^{1,2}(Q)$ to the problem

$$L\phi=-\lambda|u|\quad\text{in } Q,\qquad \phi=0\quad\text{on } \partial Q$$

is a lower solution of the problem (21), (22). Theorem 1 in [7] yields the existence at least one solution v in $\mathring{W}^{1,2}(Q)$ of (21), (22). It is clear that $u-v$ is a solution in $W^{1,2}_{\mathrm{loc}}(Q)$ of (1), (2).

References

1. Chabrowski, J., Thompson, B.: On traces of solutions of a semi-linear partial differential equation of elliptic type. Ann. Polon. Math. **62**, 47–73 (1982)
2. Chabrowski, J., Thompson, B.: On the boundary values of the solutions of linear elliptic equations. Bull. Austral. Math. Soc. **27**, 1–30 (1983)

3. Chabrowski, J.: Note on the Dirichlet problem with L^2-boundary data. Manuscripta Math. **40**, 91–108 (1982)
4. De Figueiredo, D.G.: The Dirichlet problem for nonlinear elliptic equations; Hilbert space approach, Lecture Notes in Mathematics **446**, pp. 144–165. Berlin-Heidelberg-New York: Springer 1975
5. De Figueiredo, D.G.: Semilinear elliptic problems with nonlinearities near the first eigenvalue, Proc. Univ. Florida International Symposium (A.R. Bednarek and L. Cesari, Eds.) pp. 59–71. New York: Academic Press 1977
6. De Figueiredo, D.G., Gossez, J.P.: Nonlinear perturbations of a linear elliptic problem near its first eigenvalue. J. Differential Equations **30**, 1–19 (1978)
7. Hess, P.: On the solvability of nonlinear elliptic boundary value problems. Indiana Univ. Math. J. **25**, 461–466 (1976)
8. Hess, P.: On a Theorem by Landesman and Lazar, Indiana Univ. Math. J. **23**, 827–829 (1974)
9. Kufner, A., John, O., Fučik, S.: Function spaces. Leyden, Noordhoff; Prague: Academia 1977
10. Landesman, E.M., Lazer, A.C.: Nonlinear perturbations of linear elliptic boundary value problems at resonance. J. Math. Mech. **19**, 609–623 (1970)
11. Lions, J.L.: Quelques méthodes de résolution des problèmes aux limites non linéaires. Paris: Dunod, Gauthier-Villars 1969
12. Meyer, R.D.: Some embedding theorems for generalized Sobolev spaces and applications to degenerate elliptic differential operators. J. Math. Mech. **16**, 739–760 (1967)
13. Morrey, Ch.B., Jr.: Multiplie integrals in the calculus of variations. Berlin-Heidelberg-New York: Springer 1966
14. Mikhailov, V.P.: On the boundary values of the solutions of elliptic equations. Appl. Math. Optim. **6**, 193–199 (1980)
15. Mikhailov, V.P.: Boundary values of the solutions of elliptic equations in domains with a smooth boundary. Mat. Sb. **101** (143), 163–188 (1976)
16. Mikhailov, Yu.A.: Solvability of the Dirichlet problem in the space H^1_{loc} for second order elliptic equations (Russian) Differential equations and operator theory, Mat. Issled. Akademija Nauk Moldavskoĭ SSR, **67**, 109–123 (1982)
17. Gilbarg, D., Trudinger, N.S.: Elliptic partial differential equations. Berlin-Heidelberg-New York: Springer 1977

Received January 23, 1984

Math. Z. 187, 185–194 (1984)

Mathematische Zeitschrift

Odd Order Hall Subgroups of $GL(n,q)$ and $\mathrm{Sp}(2n,q)$

Fletcher Gross

University of Utah, Department of Mathematics, Salt Lake City, UT 84112, USA

1. Introduction

Assume that r and s are odd primes with $r<s$ and let q be a power of some prime other than r or s. Then E.L. Spitznagel in [9] determined necessary and sufficient conditions for the group $GL(n,q)$ to have a Hall $\{r,s\}$-subgroup. Using these conditions, it can be shown that such a Hall subgroup, if it exists, always has a normal Sylow s-subgroup and any two Hall $\{r,s\}$-subgroups of $GL(n,q)$ are conjugate. More generally, suppose $GL(n,q)$ has a Hall ω-subgroup where ω is a set of prime divisors of the order of $GL(n,q)$ and assume ω contains neither 2 nor the characteristic of the field $GF(q)$. Let r be the smallest prime in ω and let $\tau=\omega-\{r\}$. Then we show that any ω-subgroup, whether a Hall subgroup or not, contained in $GL(n,q)$ is an extension of a normal abelian τ-subgroup by an r-group. Using this result it then is easy to show that any two Hall ω-subgroups of $GL(n,q)$ are conjugate.

Since the above structure theorem applies to all ω-subgroups and not just the Hall subgroups, it is tempting to try to show that each ω-subgroup of $GL(n,q)$ is contained in some Hall ω-subgroup. Indeed, Philip Hall [7, p. 288] conjectured that if a finite group G has a Hall ω-subgroup with $2\notin\omega$, then all Hall ω-subgroups of G are conjugate and every ω-subgroup of G is contained in a Hall ω-subgroup. Hall's conjecture is false however, and we show that if ω is any finite set of at least two odd primes then there is a positive integer n and a finite field $GF(q)$ such that ω does not contain the characteristic of $GF(q)$, each member of ω divides the order of $GL(n,q)$, $GL(n,q)$ has a Hall ω-subgroup, and $GL(n,q)$ contains an ω-subgroup which is not even isomorphic to any subgroup of a Hall ω-subgroup. Further counter-examples to Hall's conjecture as well as a discussion of the related (and probably correct) conjecture that the existence of a Hall ω-subgroup with $2\notin\omega$ always implies that all Hall ω-subgroups are conjugate may be found in [6].

Whether $Sp(2n,q)$ has Hall subgroups is also considered in [9] but the results are not as complete. With r and s as before, necessary and sufficient conditions are obtained only when $r>3$. This gap is filled using a simple property of the group $Sp(2n,q)$ which was suggested originally by the fact that $Sp(2n,q)$ and $GL(n,q^2)$ have very similar orders. Specifically, we show that if H is any subgroup of $Sp(2n,q)$ and H has order prime to both 2 and q, then H is isomorphic to a subgroup of $GL(n,q^2)$. It follows from this that, with ω as before, if $Sp(2n,q)$ has a Hall ω-subgroup then so does $GL(n,q^2)$. We then obtain the same results about ω-subgroups of $Sp(2n,q)$ as we previously had for $GL(n,q)$.

2. Notation

If n is a positive integer and ω is a set of primes, then n_ω is the largest integer t such that t divides n and all primes dividing t belong to ω. If ω consists of a single prime p, we write simply n_p. In particular, $|G|_p$ is the order of a Sylow p-subgroup of the group G. If G is a group and p is a prime, then $O_p(G)$ is the largest normal p-subgroup of G. We say that G is p-closed if $O_p(G)$ is a Sylow p-subgroup of G. The largest normal nilpotent subgroup of G is denoted by $F(G)$. If P is a p-group, then $\Omega_1(P)$ is the subgroup generated by all elements of order p. For each n, Z_n denotes a cyclic group of order n. If V is a symplectic space and U is a subspace of V, then $U^\perp$ is the subspace of V consisting of all vectors which are orthogonal to each vector in U.

If G is a group and ω is a set of primes, then G satisfies E_ω if G has a Hall ω-subgroup. If G satisfies E_ω and all Hall ω-subgroups are conjugate, then we say that G satisfies C_ω. Finally, G satisfies D_ω if G satisfies C_ω and if each ω-subgroup of G is contained in some Hall ω-subgroup of G.

The following three lemmas are straightforward and are included for the sake of completeness.

2.1 Lemma. *Let r be an odd prime and q an integer >1 which is not divisible by r. Let e be the smallest positive integer such that $q^e \equiv 1 \pmod r$. If n is any positive integer, then*

$$(q^n-1)_r = \begin{cases} (q^e-1)_r(n/e)_r & \text{if } e \text{ divides } n. \\ 1 & \text{if } e \text{ does not divide } n. \end{cases}$$

This is proved in [10, p. 530].

2.2 Lemma. *Let r be an odd prime. Then there is an integer a such that*

$$\left(\prod_{i=2}^{r}(a^i-1)\right)_r = r.$$

Proof. Since the multiplicative group of $GF(r)$ is cyclic of order $(r-1)$, there must be an integer b such that

$$b\prod_{i=1}^{r-2}(b^i-1) \not\equiv 0 \pmod r.$$

Then $\left(\prod_{i=2}^{r}(b^i-1)\right)=(b^{r-1}-1)_r>1$. If r^2 does not divide $(b^{r-1}-1)$, then set $a=b$. If r^2 does divide $(b^{r-1}-1)$, then set $a=b+r$. Then

$$(a^{r-1}-1)=(b+r)^{r-1}-1\equiv(b^{r-1}-1)+r(r-1)b^{r-2}\not\equiv 0 \pmod{r^2}.$$

Then $\left(\prod_{i=2}^{r}(a^i-1)\right)_r=r$.

2.3 Lemma. *Let R be an extra-special r-group of order r^{2n+1} and let F be a field whose characteristic is not r. Let k be the degree of a primitive r-th root of unity over F. If U is a faithful FR-module, then $\dim(U)\geqq kr^n$.*

Proof. Let E be the extension of F obtained by adjoining a primitive r-th root of unity. Let T be the Galois group of E over F and let $V=U\otimes_F E$. Then V is a faithful ER-module and $(V{:}E)=(U{:}F)$. Since the characteristic of E does not divide $|R|$, V must contain an irreducible ER-module V_0 such that R' does not act trivially on V_0. It follows from [5, Theorem 5.5.5] that $(V_0{:}E)=r^n$. Also a generator of R' is represented by a scalar matrix on V_0. Since a primitive r-th root of unity has k distinct conjugates under T, V_0 must have k distinct conjugates under the action of T. But then

$$(U{:}F)=(V{:}E)\geqq k(V_0{:}E)=kr^n.$$

3. Hall subgroups of $GL(n,q)$

We begin by analyzing the structure of $\{r,s\}$-subgroups of $GL(n,q)$ if a Hall $\{r,s\}$-subgroup exists.

3.1. Theorem. *Let H be an $\{r,s\}$-subgroup of $GL(n,q)$ where r and s are primes such that $r<s$, $(rs,2q)=1$, and r and s both divide $|GL(n,q)|$. If $GL(n,q)$ has a Hall $\{r,s\}$-subgroup, then H has a normal abelian Sylow s-subgroup.*

Proof. Assume we have a counter-example with $|H|$ as small as possible. Let $G=GL(n,q)$, let V be the vector space on which G acts, let R be a Sylow r-subgroup of H, and let S be a Sylow s-subgroup of H. Let a and b denote the multiplicative orders of q modulo r and s, respectively (i.e., a and b are the smallest positive integers such that $q^a\equiv 1 \pmod r$ and $q^b\equiv 1 \pmod s$). We now proceed in a series of steps.

1. *S is abelian and one of the following is true:*

(i) $a=b>n/s$.

(ii) $a=r-1$, $b=1$, $n<s$, $r\leqq n<r(r-1)$.

(iii) $a=r-1$, $b=r$, $n<rs$, $r\leqq n<r(r-1)$.

Proof. Since G satisfies $E_{r,s}$, it follows from Corollary 2.1.6 and Theorem 2.1.3 of [9] that one of (i), (ii) or (iii) hold. In all 3 cases, $n<bs$ and so it follows from [10] that a Sylow s-subgroup of G is the direct product of $[n/b]$ copies of

a Sylow s-subgroup of the multiplicative group of $GF(q^b)$. Hence S must be abelian.

2. (i) *H is not s-closed but every proper subgroup of H is s-closed.*

(ii) *Neither R nor S is the identity.*

(iii) *S is cyclic generated by some x with $x^s \in Z(H)$.*

(iv) *R is a special r-group of exponent r such that $R \lhd H$, $R' \leqq Z(H)$, and R/R' is faithfully and irreducibly transformed by $\langle x \rangle / \langle x^s \rangle$.*

Proof. Since we are assuming we have a counter-example and since S is abelian, S must not be normal in H. Then H is not s-closed but the minimality of $|H|$ implies that every proper subgroup of H is s-closed. If either R or S were the identity, then certainly H would be s-closed. Hence we have proved (i) and (ii).

Since $S \neq O_s(H), S$ must contain an element x such that $x \notin O_s(H)$ but $x^s \in O_s(H)$. Now $F(H) = O_r(H) O_s(H)$ and $C_H(F(H)) \leqq F(H)$. Since S is abelian and $x \notin F(H)$, x cannot centralize $O_r(H)$. Then $\langle x \rangle O_r(H)$ is not s-closed and so we must have $H = \langle x \rangle O_r(H)$. Then $S = \langle x \rangle$ and $R = O_r(H) \lhd H$. Since $x^s \in O_s(H)$, we see that $[R, x^s] = 1$ and thus $x^s \in Z(H)$.

Suppose R_1 is a proper subgroup of R such that x normalizes R_1. Then $R_1 \langle x \rangle$, being a proper subgroup of H, must be s-closed. This implies that x centralizes R_1. Theorem C of [8] now yields that R is a special r-group, $[R, x] = 1$, and R/R' is transformed faithfully and irreducibly by $\langle x \rangle / \langle x^s \rangle$. Since $r > 2, x$ cannot centralize $\Omega_1(R)$ [5, Theorem 5.3.10]. Since x certainly normalizes $\Omega_1(R)$, we must have $R = \Omega_1(R)$. Since r is odd and R has class at most 2, it follows that R has exponent r [5, Lemma 5.3.9]. This completes the proof of step 2.

3. *$R' \neq 1$ and $a = b > n/s$.*

Proof. Suppose R is abelian and let

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_m$$

be the decomposition of V into the homogeneous $GF(q)R$-components. Without loss of generality, we may assume that R does not act trivially on V_1. Since R has exponent r, since V_1 is a homogeneous module for R, and since we are assuming that R is abelian, we must have $|R/C_R(V_1)| = r$. Now x acts irreducibly on R (since $R' = 1$). Since $r < s$, then certainly $|R| > r$. It follows from this that x cannot normalize $C_R(V_1)$. But then x does not fix V_1. Since x must permute the subspaces V_i among themselves, we see that

$$n = \dim(V) \geqq s \dim(V_1).$$

Since V_1 contains a faithful module for the group $R/C_R(V_1)$, the dimension of V_1 must be at least a. Therefore, $n \geqq sa > ra$. A check of the possibilities in step 1 shows that this is impossible.

Hence $R' \neq 1$. Then a Sylow r-subgroup of $GL(n, q)$ is non-abelian. It then follows from [10] that $[n/a] \geqq r$. Then $n \geqq ar$ and it follows from step 1 that $a = b > n/s$.

4. *Contradiction.*

Proof. V must contain an irreducible $GF(q)H$-submodule U such that R' does not act trivially on U. Let $C=C_R(U)$. Then $R' \nleqq C$. Since $R' \leqq Z(H)$, and since U is an irreducible module for H, $R'C/C$ must be cyclic. Hence $|R'C/C|=r$. Now $\langle x \rangle$ transforms R/R' irreducibly and so $R'C$ must be either R' or R. Since $R'C=R$ would imply $C=R \geqq R'$, we must have $R'C=R'$. Then $C \leqq R'$. Since $|R'/C|=r$, it follows that R/C is an extra-special group. Then $|R/C|=r^{2l+1}$ for some integer l and $|R/R'|=r^{2l}$. Since $\langle x \rangle/\langle x^s \rangle$ acts faithfully and irreducibly on R/R', $2l$ must be the multiplicative order of r modulo s. Then s divides $(r^{2l}-1)$ but s does not divide (r^l-1). It follows from this that s divides (r^l+1). Since r and s are both odd, we must have $r^l+1 \geqq 2s$. From Lemma 2.3 we obtain $\dim(U) \geqq ar^l$. But then $n=\dim(V) \geqq \dim(U) \geqq ar^l \geqq a(2s-1) > as$ which contradicts step 3.

3.2. Theorem. *Suppose $GL(n,q)$ satisfies E_ω where each prime in ω divides $|GL(n,q)|$. Assume that ω contains neither 2 nor the characteristic of $GF(q)$. Let r be the smallest prime in ω and let $\tau=\omega-\{r\}$. Then the following are true:*

(1) *$GL(n,q)$ has an abelian Hall τ-subgroup.*

(2) *Each ω-subgroup of $GL(n,q)$ is an extension of a normal abelian τ-subgroup by an r-group.*

(3) *If G is any subgroup of $GL(n,q)$ such that G satisfies E_ω, then G satisfies C_ω.*

Proof. First, note that (3) includes the assertion that both $GL(n,q)$ and $SL(n,q)$ satisfy C_ω. If $|\tau| \leqq 1$, then (1) follows from the previous theorem. Assume then that $|\tau|=l \geqq 2$. Let H be a Hall ω-subgroup of $GL(n,q)$. Let $\tau=\{r_1,\ldots,r_l\}$ with $r_1<r_2<\ldots<r_l$ and set $r_0=r$. H must have a Sylow system $\{R_0,R_1,\ldots,R_l\}$ [5, Theorem 6.43] since H is solvable [4], i.e., R_i is a Sylow r_i-subgroup of H and $R_iR_j=R_jR_i$ for all i and j. If $i \geqq 1$, then R_0R_i is a Hall $\{r_0,r_i\}$-subgroup of $GL(n,q)$. The previous theorem implies that R_i is abelian. Since $R_1R_2\ldots R_l$ is a Hall τ-subgroup of G, (1) will follow once we show that $[R_i,R_j]=1$ if $1 \leqq i<j$. Accordingly assume $1 \leqq i<j$. Set $s=r_i$, $t=r_j$, $S=R_i$, $T=R_j$, and $R=R_0$. Then consideration of RS, RT, and ST shows that $GL(n,q)$ satisfies $E_{r,s}$, $E_{r,t}$, and $E_{s,t}$. Let a, b, and c be the smallest positive integers such that $q^a \equiv 1 \pmod r$, $q^b \equiv 1 \pmod s$, and $q^c \equiv 1 \pmod t$. Using the conditions given in [9] for $GL(n,q)$ to satisfy each of $E_{r,s}$, $E_{r,t}$ and $E_{s,t}$, we find that we must have

$$b=c>n/s>n/t.$$

But then, by Lemma 2.1,

$$|GL(n,q)|_{s,t}=((q^b-1)_{s,t})^{[n/b]}.$$

Now $GL(n,q)$ contains a subgroup A which is the direct product of $[n/b]$ copies of the multiplicative group of $GF(q^b)$. Then A is an abelian group and $|A|_{s,t}=|GL(n,q)|_{s,t}$. Therefore, $GL(n,q)$ has an abelian Hall $\{s,t\}$-subgroup. It now follows from [11] that every $\{s,t\}$-subgroup of $GL(n,q)$ is abelian. This implies that $[S,T]=1$ and so (1) is proved.

Now suppose K is any ω-subgroup of $GL(n,q)$. As before, let $\{R_0, R_1, \ldots, R_l\}$ be a Sylow system of K. Now for $i \geqq 1$, $R_0 R_1$ is a $\{r, r_i\}$-subgroup of $GL(n,q)$ and $GL(n,q)$ satisfies E_{r,r_i}. Theorem 3.1 now implies that R_0 normalizes R_i for all $i \geqq 1$. Then $R_1 R_2 \ldots R_i$ is a normal τ-subgroup of $K = R_0 R_1 \ldots R_l$ and $K/(R_1 \ldots R_l)$ is isomorphic to R_0. Finally, it follows from (1) and [11] that all τ-subgroups of $GL(n,q)$ are abelian. Thus (2) is proved.

Now suppose L and M are Hall ω-subgroups of G where $G \leqq GL(n,q)$. It follows from (2) and from Theorem A1 of [7] that L and M must be conjugate in G.

Our final result in this section provides counter-examples to the conjecture of Philip Hall mentioned in the introduction.

3.3. Theorem. *Let ω be a finite set of at least two different odd primes. Then there is a prime p and a positive integer n such that the following all are true:*

(1) *Every prime in ω divides $|GL(n,p)|$.*

(2) $p \notin \omega$.

(3) *$GL(n,p)$ satisfies E_ω*

(4) *$GL(n,p)$ contains an ω-subgroup which is not isomorphic to any subgroup of any Hall ω-subgroup of $GL(n,p)$.*

Proof. Perhaps it should be pointed out that Arad and Ward [1] have verified Hall's conjecture when ω is the set of *all* odd primes. Thus we assume here that ω is finite. If ω consists of a single prime, then G satisfies D_ω because of the Sylow theorems. Hence we need to assume that $|\omega| \geqq 2$.

Now let r be the smallest prime in ω and let $\tau = \omega - \{r\}$. By Lemma 2.2, there is an integer a such that

$$\left(\prod_{i=2}^{r} (a^i - 1)\right)_r = r.$$

The Chinese Remainder Theorem together with Dirichlet's Theorem about primes in an arithmetic progression imply that there is a prime p such that

$$p \equiv a (\operatorname{mod} r^2) \quad \text{and} \quad p \equiv 1 (\operatorname{mod} s)$$

for all $s \in \tau$. Now let $G = GL(r,p)$. Certainly $p \notin \omega$ but $|G|$ is divisible by each member of ω. Since $r < s$ for all $s \in \tau$ and since $p \equiv 1 (\operatorname{mod} s)$, it follows from Lemma 2.1 that

$$|G|_\tau = ((p-1)_\tau)^r.$$

Since $p \equiv a (\operatorname{mod} r^2)$, $|G|_r = r$. Now the subgroup D of all diagonal matrices in G is an abelian group of order $(p-1)^r$. Hence D contains a Hall τ-subgroup T of G. Let x be a permutation matrix corresponding to an r-cycle. Then x normalizes T and so $T\langle x\rangle$ is a Hall ω-subgroup of G. Now it is straightforward to verify that $C_T(x)$ contains only scalar matrices and so $C_T(x)$ is cyclic. This implies that $T\langle x\rangle$ contains no subgroup which is isomorphic to

$$Z_r \times Z_s \times Z_s$$

with $s \in \tau$. By Theorem 3.2, all Hall ω-subgroups of G are conjugate and so the proof of 3.3 will be complete if we can show that G contains a subgroup of the form $Z_r \times Z_s \times Z_s$ for some $s \in \tau$.

But $GL(r,p)$ contains $GL(r-1,p) \times GL(1,p)$. It follows that $GL(r,p)$ contains $A \times B$ where A and B are isomorphic to the multiplicative groups of $GF(p^{r-1})$ and $GF(p)$, respectively. Since $|\omega| \geqq 2$, τ contains some prime s. Then r and s both divide $|A|$ and s divides $|B|$. Hence $A \times B$ contains a subgroup isomorphic to $Z_r \times Z_s \times Z_s$. The theorem is proved.

Remark. As an example of this last theorem, suppose $\omega = \{3, 5\}$. Then $r=3$ and we could choose $a=2$ $\left(\text{since } \prod_{i=2}^{3} (2^i - 1)_3 = 3\right)$. Then we could want a prime p satisfying $p \equiv 2 \pmod 9$ and $p \equiv 1 \pmod 5$. Obviously $p=11$ is an acceptable choice. Then $GL(3, 11)$ satisfies E_ω (and, by Theorem 3.2, C_ω) but it does not satisfy D_ω.

4. Subgroups of $Sp(2n, q)$

Our results for $Sp(2n,q)$ depend upon the following embedding theorem.

4.1. Theorem. *Suppose $H \leqq Sp(2n,q)$ with $(2q, |H|) = 1$. Then H is isomorphic to a subgroup of $GL(n, q^2)$.*

Proof. Assume we have a counter-example with n minimal. Let V be the $2n$-dimensional vector space on which $Sp(2n,q)$ acts, let f be the associated symplectic form, and let $F = GF(q)$. We proceed in a series of steps.

1. *If U is a non-zero FH-submodule of V and the restriction of f to U is non-singular, then $U = V$.*

Proof. Suppose $U \neq V$ and let $W = U^\perp$. Since the restriction of f to U is non-singular,

$$V = U \oplus W$$

and both U and W have a non-singular sympletic form fixed by H. Then $\dim(U) = 2a$ and $\dim(W) = 2b$ for some integers a and b such that $a+b=n$. Now $H/C_H(U)$ and $H/C_H(W)$ are isomorphic to subgroups of $Sp(2a,q)$ and $Sp(2b,q)$, respectively. The minimality of n and the fact that

$$C_H(U) \cap C_H(W) = C_H(V) = 1$$

imply that H is isomorphic to a subgroup of

$$GL(a, q^2) \times GL(b, q^2).$$

But $GL(a,q^2) \times GL(b,q^2)$ is contained in $GL(a+b,q^2) = GL(n,q^2)$. This concludes step 1.

2. *V is an irreducible FH-module.*

Proof. Let U be an irreducible FH-submodule of V and assume $U \neq V$. Step 1 implies that $U \cap U^{\perp} \neq (0)$. Since U is irreducible, we must have $U \subseteq U^{\perp}$. By Maschke's Theorem, $V = U^{\perp} \oplus W$ for some FH-submodule W. Now f is non-singular and so $W \neq (0)$. If $u \in U$ and $0 \neq w \in W$, then $u + w \notin U^{\perp}$ and so $f(u', u+w) \neq 0$ for some $u' \in U$. If $0 \neq u \in U$, then $f(u, w) \neq 0$ for some $w \in W$ (otherwise, $f(u, v) = 0$ for all $v \in V$). It now follows that the restriction of f to $U \oplus W$ is non-singular. From step 1, we conclude that $V = U \oplus W$. Since $V = U^{\perp} \oplus W$ and $U \subseteq U^{\perp}$, we must have $U = U^{\perp}$. This implies that

$$\dim(U) = \dim(W) = n.$$

Now suppose $x \in C_H(U)$. Let $u \in U$ and $w \in W$. Then

$$f(u, w) = f(ux, wx) = f(u, wx).$$

This implies that $w - wx \in W \cap U^{\perp} = (0)$. Hence $x \in C_H(U) \cap C_H(W) = 1$. Therefore, U is a faithful FH-module. But then H is isomorphic to a subgroup of $GL(n, q)$ which is a subgroup of $GL(n, q^2)$. Hence we have proved step 2.

3. *Contradiction.*

Proof. Let $E = \mathrm{Hom}_{FH}(V, V)$. From Schur's Lemma, it follows that $E = GF(q^a)$ for some integer a. If b is the dimension of V over E, then $ab = 2n$. Now V is an absolutely irreducible EH-module and so b is the degree of an absolutely irreducible representation of H over a field whose characteristic does not divide $|H|$. It follows from this that b must divide $|H|$ [2, p. 600]. Hence b is odd. This implies, since $ab = 2n$, that $a = 2c$ for some integer c. Then $bc = n$. Since V is a faithful EH-module, H is isomorphic to a subgroup of $GL(b, q^{2c})$. Since $GL(b, q^{2c})$ is isomorphic to a subgroup of $GL(bc, q^2) = GL(n, q^2)$ the theorem is proved.

4.2. Corollary. *Suppose $Sp(2n, q)$ satisfies E_ω where each prime in ω divides $|Sp(2n, q)|$. Assume that ω contains neither 2 nor the characteristic of $GF(q)$. Let r be the smallest prime in ω and let $\tau = \omega - \{r\}$. Then the following are true:*

1. *$GL(n, q^2)$ satisfies E_ω.*
2. *$Sp(2n, q)$ has an abelian Hall τ-subgroup.*
3. *Each ω-subgroup of $Sp(2n, q)$ is an extension of a normal abelian τ-subgroup by an r-group.*
4. *$Sp(2n, q)$ satisfies C_ω.*

Proof. The order of $Sp(2n, q)$ is simply q^n times the order of $GL(n, q^2)$. Thus $|Sp(2n, q)|_\omega = |GL(n, q^2)|_\omega$. If H is a Hall ω-subgroup of $Sp(2n, q)$, then H is isomorphic to a subgroup K of $GL(n, q^2)$ and K must be a Hall ω-subgroup of $GL(n, q^2)$. Theorem 3.2 implies that K contains an abelian Hall τ-subgroup and every ω-subgroup of $GL(n, q^2)$ is extension of a normal abelian τ-subgroup by an r-group. Then (2) and (3) follow at once. As in the proof of 3.2, the conjugacy of the Hall ω-subgroups of $Sp(2n,q)$ follows from [7, Theorem A].

Remarks. 1. If p is the characteristic of $GF(p)$, then $|GL(n,q^2)|_p<|Sp(2,nq)|_p$. Thus it is necessary in Theorem 4.1 to assume that $(q,|H|)=1$. If $p>2$ and $n=1$, then a Sylow 2-subgroup of $Sp(2n,q)$ is non-abelian whereas $GL(n,q^2)$ is cyclic. Thus we need to assume that $(2,|H|)=1$ in 4.1.

2. The "reverse" of 4.1 does not hold, i.e., if H is a subgroup of $GL(n,q^2)$ with $(2q,|H|)=1$, it does not follow that H is isomorphic to a subgroup of $Sp(2n,q)$. For example, $GL(1,4^2)$ is cyclic of order 15 which is not isomorphic to any subgroup of $Sp(2,4)$. ($Sp(2,4)$ is the simple group of order 60.)

3. If $GL(n,q^2)$ satisfies E_ω, it is not necessarily true that $Sp(2n,q)$ satisfies E_ω. For example, if $\omega=\{3,5\}$, $n=1$, and $q=4$, then $GL(n,q^2)$ satisfies E_ω but $Sp(2n,q)$ does not.

Using 4.1 and 4.2, we now can give necessary and sufficient conditions for $S(2n,q)$ to satisfy $E_{r,s}$ where $r<s$ and $(2q,rs)=1$. Spitznagel [9] obtained this result but under the added assumption that $r>3$.

4.3. Theorem. *Let r and s be primes dividing $|Sp(2n,q)|$ with $r<s$ and $(2q,rs)=1$. Let a and b be the smallest positive integers such that $q^a\equiv 1 \pmod r$ and $q^b\equiv 1 \pmod s$. Then $Sp(2n,q)$ satisfies $E_{r,s}$ if, and only if, $a=b$ and one of the following is true:*

(1) *a is odd and $a>n/s$.*
(2) *a is even and $a>2n/s$.*

Proof. Let $G=Sp(2n,q)$, $L=GL(n,q^2)$, and $\omega=\{r,s\}$. If either (1) or (2) (together with $a=b$) holds, then in [9] an explicit construction is given of a Hall ω-subgroup of G. Thus we assume that G satisfies E_ω. Then L also satisfies E_ω. Let a' and b' be the multiplicative orders of q^2 modulo r and s, respectively. Then, since both r and s are odd and q^2 is a square, a' and b' must divide $(r-1)/2$ and $(s-1)/2$, respectively. Since L satisfies E_ω, it follows from [9, Theorem 2.1.6] that $a'=b'>n/s$. It is easy to see that $a'=a/2$ if a is even and $a'=a$ if a is odd. A similar result holds for b'. If $a\equiv b \pmod 2$, then certainly $a=b$ (since $a'=b'$) and we obtain (1) or (2).

Suppose now that $a\not\equiv b \pmod 2$. Although it is not difficult to handle this case in general, to simplify matters we deal only with the case not covered in [9], namely $r=3$. Therefore assume $r=3$. Certainly then $q^2\equiv 1 \pmod r$ and so $a'=1$. Then $b'=1$. Since we are assuming that $a\not\equiv b \pmod 2$, one of a and b is 1 and the other is 2. Then for some sign $e=\pm 1$, $q\equiv e \pmod r$ and $q\equiv -e \pmod s$.

All Hall ω-subgroups of L are conjugate by Theorem 3.2. Now the Hall ω-subgroup of $GL(n,q^2)$ constructed in [9] when $a'=b'=1$ has a normal subgroup which is the direct product of n copies of a Hall ω-subgroup of the multiplicative group of $GF(q^2)$. It follows then that a Hall ω-subgroup of G contains such a subgroup. In particular, G contains a subgroup A which is the direct product of n copies of a cyclic group of order rs. The proof will be finished by showing that no such subgroup exists.

Changing our labels if necessary, we may assume that $q\equiv 1 \pmod r$ and $q\equiv -1 \pmod s$. (Of course we no longer assume that $r=3$ or $r<s$.) Let V be the

symplectic space on which G acts and let $V=V_1\oplus\ldots\oplus V_k$ be the decomposition of V into the direct sum of irreducible $GF(q)A$-submodules. Now rs divides q^2-1 and so $GF(q^2)$ is a splitting field for A. This implies that $\dim(V_i)\leqq 2$ for all i. Since $q\equiv -1 \pmod s$, all elements of order s in A must act trivially on any V_i of dimension 1. Thus, since $A/C_A(V_i)$ must be cyclic for all i and since A contains an elementary abelian group of order s^n, the number of V_i of dimension 2 must be at least n. Since $\dim(V)=2n$, we conclude that $\dim(V_i)=2$ for each i. It now follows that $A/C_A(V_i)$ must be cyclic of order rs for all i.

Since V_i is an irreducible $GF(q)A$-module, we must have $V_i\cap V_i^\perp=V_i$ or (0). But if $V_i\cap V_i^\perp=(0)$, then $A/C_A(V_i)$ is contained in $Sp(2,q)=SL(2,q)$. Now every element of $SL(2,q)$ has order dividing one of the numbers $2p$ (here p is the characteristic of $GF(q)$), $q+1$, or $q-1$ [3, Sect. 240]. Hence $SL(2,q)$ has no element of order rs. Therefore $V_i\cap V_i^\perp=V_i$ for all i. Now $V_1^\perp \neq V$ and so $V_1^\perp$ does not contain V_i for some $i>1$. Without loss of generality we may assume that $V_2 \nsubseteq V_1^\perp$. The irreducibility of V_2 forces $V_1^\perp\cap V_2=(0)$. As in step 2 of the proof of Theorem 4.1, we find that $V_1\oplus V_2$ is a non-degenerate symplectic space and $C_A(V_1)=C_A(V_2)$. Then if $W_1=V_1\oplus V_2$, we obtain $V=W_1\oplus W_1^\perp$, $\dim(W_1)=4$, and $A/C_A(W_1)=A/C_A(V_1)$ which is contained in $GL(2,q)$. Repeating the same argument on $W_1^\perp$, we eventually find that

$$V=W_1\oplus\ldots\oplus W_t$$

where $\dim(W_i)=4$ and $A/C_A(W_i)\leqq GL(2,q)$ for $1\leqq i\leqq t$. Then, since $\dim(V)=2n$, $t=n/2$ and so A is isomorphic to a subgroup of the direct product of $(n/2)$ copies of $GL(2,q)$. But the construction in [10] shows that a Sylow s-subgroup of $GL(2,q)$ is cyclic. Since a Sylow s-subgroup of A is elementary abelian, we would have $|A/C_A(W_i)|_s\leqq s$ for all i. This would lead to $|A|_s\leqq s^{(n/2)}$ which contradicts the fact that A is the direct product of n copies of a cyclic group of order rs. The proof is complete.

References

1. Arad, Z., Ward, M.B.: New criteria for the solvability of finite groups. J. Algebra **77**, 234–246 (1982)
2. Curtis, C.W., Reiner, I.: Representation theory of finite groups and associative algebras. New York: Wiley (Interscience) 1962
3. Dickson, L.E.: Linear groups with an exposition of Galois field theory. New York: Dover 1962
4. Feit, W., Thompson, J.: Solvability of groups of odd order, Pacific J. Math. **13**, 755–1029 (1963)
5. Gorenstein, D.: Finite groups. New York: Harper and Row 1968
6. Gross, F.: On a conjecture of Philip Hall. In preparation
7. Hall, P.: Theorems like Sylow's. Proc. London Math. (3) **6**, 286–304 (1956)
8. Hall, P., Higman, G.: On the p-length of p-soluble groups and reduction theorems for Burnside's problem. Proc. London Math. Soc. (3) **6**L, 1–42 (1956)
9. Spitznagel, E.L.: Hall subgroups of certain families of finite groups, Math. Z. **97**, 259–290 (1967)
10. Weir, A.J.: Sylow p-subgroups of the classical groups over finite fields with characteristic prime to p. Proc. Amer. Math. Soc. **6**, 529–533 (1955)
11. Wielandt, H.: Zum Satz von Sylow. Math. Z. **60**, 407–408 (1954)

Received March 15, 1984

Math. Z. 187, 195–208 (1984)

Mathematische Zeitschrift

Isoparametric Hypersurfaces in the Pseudo-Riemannian Space Forms

Jörg Hahn

Mathematisches Institut der Universität Bonn,
Wegelerstr. 10, D-5300 Bonn, Federal Republic of Germany

The purpose of the present paper is to extend the notion of isoparametric hypersurfaces to the pseudo-riemannian space forms, to investigate the properties of isoparametric families and to describe some examples, in particular those examples that arise by applying the construction of Ferus, Karcher and Münzner [2] to the standard space S_s^n. We give a great number of complicated examples of hypersurfaces in the pseudo-riemannian space forms that can be described quite explicitely.

The main intention of the first section is to fix our notations. Section 2 gives the definition of isoparametric hypersurfaces and some first examples. We sketch a general theory of isoparametric families in the pseudo-riemannian space forms including isoparametric functions, curvature foliations, focal varieties and Cartan's identity.

The main part of this paper is Sect. 3. We introduce Clifford systems on $\mathbb{R}_s^{2l}$ and thus construct isoparametric functions on S_s^{2l-1}. In general their levels form three isoparametric families. One family has four distinct real principal curvatures and two focal varieties and the other two families have two real and two complex principal curvatures and only one focal variety that coincides with a focal variety of the first family. The level hypersurfaces are diffeomorphic to De Sitter sphere bundles over the focal varieties. The level surface can be disconnected and the components are again isoparametric. Thus we obtain isoparametric families with up to five focal varieties.

In Sect. 4 we give an example of an isoparametric hypersurface with only one principal curvature whose geometric multiplicity is not constant.

This paper is a short version of the author's "Diplomarbeit" (Bonn, January 1983). The author gratefully acknowledges the constant encouragement and motivation given to him by Prof. H. Karcher.

1. Preliminaries

Let $\mathbb{R}_s^n$ be an n-dimensional real vector space provided with an inner product given by $\langle x, x\rangle = -\sum_{i=1}^{s} x_i^2 + \sum_{i=s+1}^{n} x_i^2$. We have the standard spaces of signature

(n, s): $\mathbb{R}^n_s$, and the De Sitter spheres

$$S^n_s := \{x \in \mathbb{R}^{n+1}_s \mid \langle x, x\rangle = +1\}, \qquad H^n_s := \{x \in \mathbb{R}^{n+1}_{s+1} \mid \langle x, x\rangle = -1\}$$

with constant curvature $\kappa = 0$, $+1$ and -1 resp. (cf. [7]).

In these spaces the geodesic through a point x in direction $v(\langle v, v\rangle \in \{0, +1, -1\})$ is given by $\exp_x tv = c_\delta(t)\,x + s_\delta(t)\,v$, where $\delta := \kappa\langle v, v\rangle$ and $c_\delta(t)$, $s_\delta(t)$ stands for 1, t (in case $\delta = 0$), $\cos t$, $\sin t$ ($\delta = +1$) or $\cosh t$, $\sinh t$ ($\delta = -1$), resp.

Let $\bar{M}$ be a pseudo-riemannian space form of curvature $\kappa \in \{0, +1, -1\}$ with metric $\langle\, ,\,\rangle$ and $M \subset \bar{M}$ be a nondegenerate submanifold (i.e. the induced metric is nondegenerate).

We then have the orthogonal decomposition of the tangent bundle $T\bar{M} = TM \oplus \perp M$. The Levi-Civita covariant derivative of $\bar{M}$ will be denoted by $\bar{D}$, that of M by D and the curvature tensor of M by R. Let N be a normal vector field on M. We have the shape operator associated to N given by $S_N X = -(\bar{D}_X N)^{\tan}$ and S_N is symmetric, i.e. $\langle S_N X, Y\rangle = \langle X, S_N Y\rangle$ for vector fields X, Y on M. If M is a hypersurface with unit normal field ($\langle N, N\rangle = \pm 1$) we write $S := S_N$ and have the equations of *Gauß:*

$$R(X, Y)Z = \kappa(\langle Y, Z\rangle X - \langle X, Z\rangle Y) + \langle N, N\rangle(\langle SY, Z\rangle SX - \langle SX, Z\rangle SY)$$

and *Codazzi* $(D_X S)\,Y = (D_Y S)\,X$ for vector fields X, Y, Z on M. The - possibly complex - eigenvalues of S are called the *principal curvatures*. If M is indefinite the algebraic and geometric multiplicity of a principal curvature need not coincide. If they coincide we speak simply of the multiplicity.

At least locally and for sufficiently small $r \in \mathbb{R}$ the map $\varphi_r\colon M \to \bar{M}$, $p \to \exp_p rN_p$ describes the parallel surface M_r of the hypersurface M at distance r. For fixed $p \in M$ we denote the parallel translation along the geodesic $c\colon r \to c(r) = \varphi_r(p)$ by $\tau_r\colon T_p\bar{M} \to T_{c(r)}\bar{M}$. We write τ_{-r} for $(\tau_r)^{-1}$. Similarly we define the parallel translation of operators, e.g. of the shape operator $\tau_r(S)\colon \tau_r(T_p M) \to \tau_r(T_p M)$, $X \to \tau_r(S\tau_{-r}X)$. The shape operator S_r of M_r is

$$S_r = \tau_r[(\delta + \mathrm{ct}_\delta(r)\,S)\,(\mathrm{ct}_\delta(r) - S)^{-1}],$$

where $\delta := \kappa\langle N, N\rangle = (\dim M)^{-1}\,\mathrm{Ric}(N, N) \in \{0, +1, -1\}$ and $\mathrm{ct}_\delta(r) := \dfrac{c_\delta(r)}{s_\delta(r)}$. If the principal curvatures of M at p are k_j, $j = 1, \ldots, g$ then M_r has the principal curvatures $k_{r,j}$, $j = 1, \ldots, g$ with the same algebraic and geometric multiplicities and $k_{r,j} = k_j$ if $k_j^2 + \delta = 0$ and $k_{r,j} = \mathrm{ct}_\delta(t_j - r)$ if $k_j = \mathrm{ct}_\delta(t_j)$. The eigenspaces are parallel along c.

2. Isoparametric Families

We now introduce isoparametric hypersurfaces in the pseudo-riemannian space forms, state some basic facts and give first examples. Most proofs are similar to the riemannian case (cf. [1], [5]) and therefore omitted.

Let $\bar{M}$ be a pseudo-riemannian space form of curvature $\kappa \in \{0, \pm 1\}$ and $M \subset \bar{M}$ be a nondegenerate hypersurface with unit normal field N.

2.1. Proposition. *The following are equivalent:*

(1) *All parallel surfaces M_r of M for r sufficiently close to zero have constant mean curvature.*

(2) *The principal curvatures and their algebraic multiplicities are constant on M.*

(3) *The shape operator has the same characteristic polynominal at all points of M.*

A *hypersurface* satisfying these conditions is called *isoparametric*, the family of its parallel surfaces is an *isoparametric family*. The number $\delta := \kappa\langle N, N\rangle = \operatorname{sign} \operatorname{Ric}(N, N) \in \{0, \pm 1\}$ is called the *type* of the hypersurface or family.

Replacing the metric of $\bar{M}$ with its negative, we get a space of curvature $-\kappa$ and M remains isoparametric with the same type. Therefore it suffices to study isoparametric hypersurfaces in the space forms of curvature $\kappa = 0, +1$.

A *function* $f: \bar{M} \to \mathbb{R}$ is called *isoparametric* if the length of its gradient vector field $\bar{\nabla} f$ and the Laplacian $\bar{\Delta} f$ are constant on the level hypersurfaces, i.e. there are functions $\Phi, \Psi: f(\bar{M}) \to \mathbb{R}$ such that $\langle \bar{\nabla} f_{(p)}, \bar{\nabla} f_{(p)}\rangle = \Phi \circ f_{(p)}$ and $\bar{\Delta} f_{(p)} = \Psi \circ f_{(p)}$ for $p \in \bar{M}$.

We set $W_{RN}(f) := \{c \in f(\bar{M}) \mid \Phi(c) \neq 0\}$.

2.2. Proposition. *If $f: \bar{M} \to \mathbb{R}$ is isoparametric and $I \subset W_{RN}(f)$ is connected, then the level hypersurfaces $(f_{(c)}^{-1})_{c \in I}$ form an isoparametric family. A normal field on $M := f_{(c)}^{-1}$ is $N := |\langle \bar{\nabla} f, \bar{\nabla} f\rangle|^{-1/2} \bar{\nabla} f$ and the shape operator in $p \in M$ is $S_p X = -|\langle \bar{\nabla} f, \bar{\nabla} f\rangle|^{-1/2} \bar{D}_X \bar{\nabla} f$ with $\operatorname{trace} S_p = \zeta'(c) - |\Phi(c)|^{-1/2} \Psi(c)$ where $\zeta(c) := \sqrt{|\Phi(c)|} \operatorname{sign} \Phi(c)$.*

Quadratic Examples

Despite this similiarity there are remarkable differences to the riemannian case that occur in simple examples. We determine the quadratic isoparametric functions on $\mathbb{R}_s^n$ and those quadratic functions on $\mathbb{R}_s^{n+1}$ whose restriction to S_s^n is isoparametric. The level hypersurfaces can have complex principal curvatures or non-diagonalizable shape operators.

Now d, ∇, Δ denote the euclidean derivative, gradient, Laplacian in $\mathbb{R}_s^n$ or $\mathbb{R}_s^{n+1}$ and $\bar{D}, \bar{\nabla}, \bar{\Delta}$ denote covariant derivative, gradient, Laplacian in S_s^n.

$\operatorname{Sym}(\mathbb{R}_s^n)$ denotes the symmetric endomorphisms of $\mathbb{R}_s^n$, i.e. $A \in \operatorname{Sym}(\mathbb{R}_s^n)$ iff $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in \mathbb{R}_s^n$.

2.3. Proposition. *Let $f: \mathbb{R}_s^n \to \mathbb{R}$ be a quadratic function, i.e. $f(x) = \langle Ax, x\rangle + 2\langle a, x\rangle$ with $0 \neq A \in \operatorname{Sym}(\mathbb{R}_s^n)$, $a \in \mathbb{R}_s^n$. Then:*

(1) *f is isoparametric iff there is an $\varepsilon \in \mathbb{R}$ such that $(A - \varepsilon)A = 0$ and $(A - \varepsilon)a = 0$.*

(2) *If $\varepsilon = 0$ and $\langle a, a\rangle = 0$ then $W_{RN}(f) = \varnothing$. If $\varepsilon \neq 0$ (case I) or $\varepsilon = 0$ and $\langle a, a\rangle \neq 0$ (case II), then $W_{RN}(f) \neq \varnothing$ and the level sets $f_{(c)}^{-1}$, $c \in W_{RN}(f)$ are isoparametric hypersurfaces of type $\delta = 0$.*

(3) *In case I the level hypersurfaces are (hyperbolic) cylinders. The shape operator of* $M := f_{(c)}^{-1}$, $c \in W_{RN}(f)$ *is diagonalizable with at most one non-zero principal curvature.*

(4) *In case II the shape operator* S *is nilpotent with* $S^2 = 0$.

Proof. We have for $x \in \mathbb{R}^n_s$: $\nabla f(x) = 2Ax + 2a$, $\langle \nabla f(x), \nabla f(x) \rangle = 4\langle A^2 x, x\rangle + 8\langle Aa, x\rangle + 4\langle a, a\rangle$ and $\Delta f(x) = 2$ trace $A = \text{const.}$

If f is isoparametric and $W_{RN}(f) = \emptyset$, then $A^2 = 0$, $Aa = 0$ and $\langle a, a\rangle = 0$.

If f is isoparameteric, $c \in W_{RN}(f) \neq \emptyset$ and $x \in M := f_{(c)}^{-1}$, we have $T_x M = [Ax + a]^{\perp}$ and $T_x M$ is an invariant subspace of A. Since A is symmetric, $Ax + a$ is an eigenvector, i.e. there is $\varepsilon(x) \in \mathbb{R}$ such that $(A - \varepsilon(x))(Ax + a) = 0$. $U := \{x \in \mathbb{R}^n_s \mid f(x) \in W_{RN}(f)\}$ is open, the map $U \to \mathbb{R}$, $x \to \varepsilon(x)$ is continuous, takes only finitely many values and is therefore constant $= \varepsilon$ on an open set U_0. Hence $(A - \varepsilon)Ax + (A - \varepsilon)a = 0$ for $x \in U_0$ and comparing coefficients yields $(A - \varepsilon)A = 0$, $(A - \varepsilon)a = 0$.

If on the other hand $(A - \varepsilon)A = 0$ and $(A - \varepsilon)a = 0$, then $\langle \nabla f(x), \nabla f(x)\rangle = 4\varepsilon f(x) + 4\langle a, a\rangle$, i.e. f is isoparametric. □

Example. $f: \mathbb{R}^3_1 \to \mathbb{R}$, $(x_1, x_2, x_3) \to x_1^2 - 2x_1 x_2 + x_2^2 - 2x_3$ is isoparametric (case II) and $W_{RN}(f) = \mathbb{R}$. The level surfaces are parabolic cylinders.

2.4. Proposition. *Let* $f: S^n_s \to \mathbb{R}$, $x \to \langle Ax, x\rangle$, $A \in \mathrm{Sym}(\mathbb{R}^{n+1}_s)$ *be a quadratic function. Then:*

(1) f *is isoparametric with* $W_{RN}(f) \neq \emptyset$ *iff the minimal polynomial of* A *is* $\mu_A(t) = t^2 + \alpha t + \beta$, $\alpha, \beta \in \mathbb{R}$.

(2) *If* f *is isoparametric,* $c \in W_{RN}(f)$, *then* $\mu_A(c) \neq 0$ *and the level surface* $M := f_{(c)}^{-1}$ *is of type* $\delta = +1$ *for* $\mu_A(c) < 0$ *and of type* $\delta = -1$ *for* $\mu_A(c) > 0$.

The shape operator in $x \in M$ *is* $S = (-\delta \mu_A(c))^{-1/2} (c - A)|_{T_x M}$.

If $\alpha^2 - 4\beta > 0$, *then* S *is diagonalizable with (at most) two distinct real principal curvatures.*

If $\alpha^2 - 4\beta = 0$, *then* M *is of type* $\delta = -1$, S *is not diagonalizable* $\left(\textit{except}\ \text{rank}\left(A + \frac{\alpha}{2}\right) = 1\right)$ *and the only principal curvature is* ± 1.

If $\alpha^2 - 4\beta < 0$, *then* M *is of type* $\delta = -1$, $n + 1 = 2s$, S *is diagonalizable with two complex principal curvatures.*

Proof. We have for $x \in S^n_s$: $\bar{\nabla} f(x) = 2Ax - 2\langle Ax, x\rangle x$, $\langle \bar{\nabla} f(x), \bar{\nabla} f(x)\rangle = 4\langle A^2 x, x\rangle - 4f(x)^2$, $\bar{\Delta} f(x) = 2\,\mathrm{trace}\,A - 2(n+1) f(x)$ and for $v \in T_x S^n_s$: $\bar{D}_v \bar{\nabla} f = 2Av - 2\langle Av, x\rangle x - 2f(x) v$.

(1) If f is isoparametric, $c \in W_{RN}(f)$ and $x \in M := f_{(c)}^{-1}$ we have $T_x M = [x, Ax]^{\perp}$ (orthogonal complement of the span of x, Ax) and $T_x M$ is an invariant subspace of A. By the symmetry of A $[x, Ax]$ is also invariant. We now set $A_{1,x} := A|_{[x, Ax]}: [x, Ax] \to [x, Ax]$ and $A_{2,x} := A|_{T_x M}: T_x M \to T_x M$. Then the shape operator at x is $S_x = \rho(c)^{-1}(-A_{2,x} + c)$ (by 2.2) where $\rho(c) := |\langle A^2 x, x\rangle - c^2|^{-1/2}$ is independent of $x \in M$, because f is isoparametric. One easily computes the characterictic polynomial of $A_{2,x}$:

$$\chi_{2,x}(t) = (-\rho(c))^{n-1} \det(S_x - \rho(c)^{-1}(c - t))$$

which depends only on $f(x) = c$, not on $x \in M$ by (2.1).

Therefore the characteristic polynomial of $A_{1,x}$ is independent of $x\in M$, i.e. there are $\alpha, \beta\in\mathbb{R}$ such that $\chi_{1,x}(t)=t^2+\alpha t+\beta$ for all $x\in M$. This yields $(A^2+\alpha A+\beta)|_V=0$, where $V:=\operatorname{span}_{x\in M}(x, Ax)$. But $V=\mathbb{R}^{n+1}_s$ because $x\in M\subset V$ gives $T_xM\subset V$ and $\mathbb{R}^{n+1}_s=[x, Ax]\oplus T_xM\subset V$. Therefore we have $A^2+\alpha A+\beta=0$ and since f is not constant we get the conclusion.

If on the other hand the minimal polynomial of A is of the given form, then one easily checks that $\langle\bar{\nabla}f(x), \bar{\nabla}f(x)\rangle=-4\mu_A(f(x))$. Therefore f is isoparametric and $W_{\mathrm{RN}}(f)\neq\varnothing$.

(2) The formula for the shape operator follows from (2.2) since $\langle Av, x\rangle=0$ for all $v\in T_xM$. $\square$

2.5. *Example*. B-scroll over a null curve in $\mathbb{R}^3_1$ (cf. [3]).

Let $t\to x(t)$ be a null curve in $\mathbb{R}^3_1$ with Cartan frame (A, B, C), i.e. A, B, C are vector fields along $x(t)$ satisfying the following conditions:

$$\langle A, A\rangle=\langle B, B\rangle=0, \qquad \langle A, B\rangle=-1$$
$$\langle A, C\rangle=\langle B, C\rangle=0, \qquad \langle C, C\rangle=+1$$

and

$$\dot{x}=A, \qquad \dot{A}=ac \qquad (a=a(t))$$
$$\dot{B}=bC(b=\text{const.}), \qquad \dot{C}=bA+aB.$$

Then the B-scroll immersion $\psi\colon(s, t)\to x(t)+sB(t)$ describes (at least locally) an isoparametric surface in $\mathbb{R}^3_1$ with shape operator $S=\begin{bmatrix}-b & -a(t)\\ 0 & -b\end{bmatrix}$ relative to the basis $\left(\frac{\partial\psi}{\partial s}, \frac{\partial\psi}{\partial t}\right)$.

If we have $b=0$ it is easier to describe the so given surfaces as cylinders over plane curves: Let (A, B, C) be a fixed basis of $\mathbb{R}^3_1$ that satisfies the same metric conditions as before. If $t\to y(t)$ is a curve in the $[A, C]$-plane: $y(t)=\alpha(t)A+\gamma(t)C$ with $\dot{\alpha}(t)\neq 0$, then the cylinder given by $\psi\colon(s, t)\to y(t)+sB$ is isoparametric.

For example take $A=\frac{1}{\sqrt{2}}(1, -1, 0)$, $B=\frac{1}{\sqrt{2}}(1, 1, 0)$, $C=(0, 0, 1)$, $\alpha(t)=t$, $\gamma(t)=\sin t$.

By this we have isoparametric surfaces where the geometric multiplicity of the principal curvature is not constant, the minimal polynomial of the shape operator changes its degree. The surface in the example above is obviously not algebraic. This is a remarkable difference to the riemannian case where all isoparametric hypersurfaces are level sets of homogeneous polynomials (cf. [4]).

From now on let $M\subset\bar{M}$ be an isoparametric hypersurface of type δ, signature (n, s), with normal field $\langle N, N\rangle=:\nu=\pm 1$ and shape operator S.

2.6. Proposition: Curvature Foliations. *Let k be a* real *principal curvature of constant geometric multiplicity m_k. Then the eigenspace distribution $T_k:=\operatorname{Ker}(S-k)$ is autoparallel in M. If k is of multiplicity m_k (i.e. also algebraic multiplicity $=m_k$), then the maximal integral manifold B_k through $p\in M$ is a nondegenerate manifold of dimension m_k. B_k is totally geodesic in M and umbilical as a submanifold of $\bar{M}$.*

We now study the situation when the map φ_r (defined in Sect. 1) becomes singular. This happens if $\mathrm{ct}_\delta(r)=k$ is a principal curvature. This is only possible for $k\in\mathbb{R}$ if $\delta=+1$, $k\in\mathbb{R}\setminus\{0\}$ if $\delta=0$, $k\in]-\infty, -1[\cup]1,\infty[$ if $\delta=-1$. The proof of the following is similar to [4], [6].

2.7. Proposition: Focal Varieties. *Let* $\mathrm{ct}_\delta(r)=k$, $p\in M$ *and set* $q:=\varphi_r(p)$, $N_{q,p}:=\tau_r(N_p)$.

(1) *If k is of multiplicity m_k, φ_r is (at least in a neighborhood of p) a submersion onto a nondegenerate submanifold of codimension m_k+1 in $\bar{M}$, that will be denoted by M_k, the* focal variety associated to k.

If (m_k, s_k) is the signature of T_k, then M_k has signature $(n-m_k, s-s_k)$. φ_r maps the leaf B_k of T_k through p into the single point q.

If the shape operator at p is given by $S=\begin{bmatrix} k & 0 \\ 0 & A_p \end{bmatrix}$ *relative to the orthogonal decomposition* $T_pM=T_{k(p)}\oplus T_{k(p)}^\perp$ *then the shape operator of M_k at q associated to $N_{q,p}$ is*

$$S_{N_{q,p}}=\tau_r(\delta+kA_p)(k-A_p)^{-1}.$$

If we set $\perp M_k(v):=\{X\in\perp M_k \mid \langle X,X\rangle=v\}$, *then the shape operators of M_k have the same characteristic polynomial on the components of $\perp M_k(v)$, i.e. the map* $\chi\colon \perp M_k(v)\to\mathbb{R}[t]$, $N\to\chi_N(t):=\det(S_N-t)$ *is constant on the components.*

(2) *If the geometric multiplicity of k is constant but different from the algebraic multiplicity, then φ_r is (locally) a submersion onto a submanifold with degenerate metric.*

(3) *If on the other hand $\varphi_r(M)=:M_k$ is a submanifold with nondegenerate metric of signature $(n-m_k, s-s_k)$ $(m_k>0)$, then k is a principal curvature of M of multiplicity m_k and M_k is the focal variety associated to k. The eigenspace $T_{k(p)}$ has signature (m_k, s_k). The shape operator at p is given by* $S=\begin{bmatrix} k & 0 \\ 0 & A \end{bmatrix}$ *relative to the orthogonal decomposition* $T_pM=T_{k(p)}\oplus T_{k(p)}^\perp$, *where* $A=\tau_{-r}(k+S_{N_{q,p}})^{-1}(kS_{N_{q,p}}-\delta)$.

If $S_{N_{q,p}}$ has the eigenvalues $\mathrm{ct}_\delta(t_j)$, $j=2,\dots,g$, then M has the principal curvatures $k=\mathrm{ct}_\delta(r)$ with multiplicity m_k and $\mathrm{ct}_\delta(t_j+r)$, $j=2,\dots,g$ with the same algebraic and geometric multiplicities.

We now select a *real* principal curvature k whose algebraic and geometric multiplicity coincide and a geodesic c tangent to the leaf B_k (2.6). We can describe the operator $A(t):=((S-k)|_{T_{k(c(t))}^\perp})^{-1}$ along $c(t)$. From this one obtains the basic identity of Cartan. Although we get a somewhat different result, the proof of Karcher and Ferus [1] from the riemannian case also applies to the indefinite case.

2.8. Proposition. *Let $c(t)$ be a geodesic in B_k with $\langle \dot c, \dot c\rangle=:\gamma\in\{0,\pm1\}$. Then:* $\frac{D}{dt}\frac{D}{dt}A(t)+\gamma v(\delta+k^2)A(t)+\gamma v k=0$. *If $\gamma=\pm1$ and $\delta+k^2\neq0$ we set* $\varepsilon:=\operatorname{sign}(\gamma v(\delta+k^2))$ *and obtain the explicit solution:*

$$(\delta+k^2)A(t)+k=\tau_t[\mathrm{c}_\varepsilon(\sqrt{\varepsilon\gamma v(\delta+k^2)}\,t)((\delta+k^2)A(0)+k) + \mathrm{s}_\varepsilon(\sqrt{\varepsilon\gamma v(\delta+k^2)}\,t)(\delta+k^2)A'(0)].$$

2.9. Theorem. Basic identity of Cartan. *If the – possibly complex – principal curvatures of an isoparametric hypersurface of type δ are k_j, $j=1,\dots,g$ with algebraic multiplicities m_j and if for $i\in 1,\dots,g$ the principal curvature k_i is* real *and its algebraic and geometric multiplicity coincide, then:*

$$\sum_{j\neq i} m_j \frac{\delta+k_i k_j}{k_i-k_j}=0.$$

This yields some conclusions for isoparametric hypersurfaces with *real-diagonalizable* shape operator: Hypersurfaces of type $\delta\in\{0,-1\}$ have at most two distinct principal curvatures k_1, k_2 and these satisfy $\delta+k_1 k_2=0$ (cf. [1], [5]).

If the hypersurface is of type $\delta=+1$ and if in addition no eigenspace is negative definite in case $\nu=+1$ and no eigenspace is positive definite in case $\nu=-1$, then the principal curvatures are $k_j=\cot\left(\alpha+(j-1)\frac{\pi}{g}\right)$, $j=1,\dots,g$ with $\alpha\in\left]0,\frac{\pi}{g}\right[$ and the multiplicities satisfy $m_j=m_{j+2}$ (indices mod g) (cf. [1]).

The condition on the signature of the eigenspaces is needed in order to achieve $\varepsilon=+1$ in (2.8) and then use the periodicity of cos, sin.

The condition holds iff in (2.7.1) $\perp M_k(\nu)$ is connected and then the conclusion can also be obtained by investigating the characteristic polynomials of the shape operators of M_k as in [4].

3. The Clifford Examples

We now introduce Clifford systems and construct isoparametric hypersurfaces in S_s^{2l-1}. The general reference for this section is [2].

3.1. Clifford Systems. A *Clifford system* of signature (m,r) on $\mathbb{R}_s^{2l}$ is an m-tuple $(P_1,\dots,P_m)$ $(m\geqq 2)$ such that $P_i\in\mathrm{Sym}(\mathbb{R}_s^{2l})$ and $P_iP_j+P_jP_i=2\eta_{ij}$ for $i,j=1,\dots,m$, where $\eta_{ij}=-1$ $(i=j\leqq r)$, $=+1$ $(r<i=j\leqq m)$ and $=0$ otherwise.

On $\mathrm{Sym}(\mathbb{R}_s^{2l})$ we have a pseudo metric given by $\langle A,B\rangle := (2l)^{-1}\operatorname{trace}(A\cdot B)$. The subspace spanned by the Clifford system will be called its Clifford span $\Sigma := [P_1,\dots,P_m]\subset\mathrm{Sym}(\mathbb{R}_s^{2l})$. We set $\Sigma(\delta) := \{P\in\Sigma\,|\,\langle P,P\rangle=\delta\}$, $(\delta=\pm 1)$ and $[P]^\perp := \{Q\in\Sigma\,|\,\langle P,Q\rangle=0\}$, $P\in\Sigma$.

We give some important properties of Clifford systems:

(1) $(P_1,\dots,P_m)$ is an ordered orthonormal basis of Σ and conversely every ordered orthonormal basis of Σ is a Clifford system.

(2) For $P,Q\in\Sigma$ we have $PQ+QP=2\langle P,Q\rangle$ and $\langle Px,Qx\rangle=\langle P,Q\rangle\langle x,x\rangle$ for all $x\in\mathbb{R}_s^{2l}$.

(3) $P\in\Sigma(+1)$ has eigenvalues $+1$, -1 with multiplicities l and $P\in\Sigma(-1)$ has eigenvalues $+i$, $-i$. In particular trace $P=0$, $P\in\Sigma$.

If $P\in\Sigma(+1)$, $Q\in[P]^\perp$, $\langle Q,Q\rangle\neq 0$, then Q interchanges the eigenspaces $E_\pm(P) := \mathrm{Ker}(P\mp 1)$.

(4) We have $r\leqq l$, and $s\neq 0$ implies $m-r\leqq s$ and $s\neq 2l$ implies $m-r\leqq 2l-s$. If $r>0$ then $s=l$ and if $r\leqq m-2$ then there exists $s_1\in\mathbb{N}_0$ such that $s=2s_1$.

(5) If $r \leqq m-2$ then all eigenspaces $E_+(P)$, $P \in \Sigma(+1)$ have signature (l, s_1). If $r=m-1$, then $\Sigma(+1)$ has two components Σ_i $(i=1,2)$ and there exists s_i such that $E_+(P)$ has signature (l, s_i) for all $P \in \Sigma_i$ and $s_1+s_2=s$.

(6) The function $H: \mathbb{R}^{2l}_s \to \mathbb{R}$, $H(x)=\sum_{j=1}^{m} \eta_{jj}\langle P_j x, x\rangle^2$ depends only on Σ, not on $(P_1, \ldots, P_m)$.

3.2. Proposition. *The function*

$$f: S^{2l-1}_s \to \mathbb{R}, \quad f(x)=\langle x, x\rangle^2 - 2\sum_{j=1}^{m} \eta_{jj}\langle P_j x, x\rangle^2$$

is isoparametric.

Proof. Extending f to a homogeneous function F of degree 4 on $\mathbb{R}^{2l}_s$ and using the formulas relating gradient and Laplacian on $\mathbb{R}^{2l}_s$ and on S^{2l-1}_s, we compute

(1) $\bar{\nabla} f(x) = \nabla F(x) - 4f(x)x = 4(1-f(x))x - 8\sum_j \eta_{jj}\langle P_j x, x\rangle P_j x.$

(2) $\langle \bar{\nabla} f(x), \bar{\nabla} f(x)\rangle = \langle \nabla F(x), \nabla F(x)\rangle - 16 f(x)^2$
$= 16(1-f(x)^2) \qquad (3.1.2)$

(3) $\bar{\Delta} f(x) = \Delta F(x) - 8(l+1)f(x)$
$= 8(l+1-2m-(l+1)f(x)) \qquad (3.1.2., 3.) \quad \square$

We now set $M_c := f^{-1}(c)$, $c \in W_{RN}(f)$. Then $c \neq \pm 1$ and M_c is an isoparametric hypersurface of type $\delta = -1$ (if $c < -1$), $\delta = +1$ $(-1 < c < +1)$ and $\delta = -1$ $(c > +1)$ resp. We need the following

3.3. Lemma. *Let $x \in M_c$, $c \in W_{RN}(f)$. Then:*

(1) *The normal vector at x is:*

$$N_x = (\delta(1-c^2))^{-1/2}((1-c)x - 2\sum_j \eta_{jj}\langle P_j x, x\rangle P_j x).$$

(2) *For $P \in \Sigma$ we have $\langle Px, N_x\rangle = -(\delta(1-c^2))^{-1/2}(1+c)\langle Px, x\rangle$ and*

$$\langle PN_x, N_x\rangle = -\delta\langle Px, x\rangle.$$

Proof. (1) is clear from (3.2), (2.2). For (2) it suffices to assume $P=P_i$, $i \in \{1, \ldots, m\}$ and an easy computation yields the result. $\square$

We now investigate the set where the gradient of f vanishes. It turns out that these are the focal varieties (as is to be expected from the riemannian case). This enables us to determine the topological structure of the hypersurfaces. We set $M_\pm := \{x \in S^{2l-1}_s \mid \bar{\nabla} f(x) = 0 \text{ and } f(x) = \pm 1\}$,

$$\perp M_\pm(\delta) := \{(x, v) \mid x \in M_\pm, v \in \perp_x M_\pm, \langle v, v\rangle = \delta\}, \qquad \delta = \pm 1$$

and $m_1 := m-1$, $m_2 := l-m$. Note that $M_\pm \neq f^{-1}(\pm 1)$ if the metric is indefinite, i.e. the focal set is a proper subset of the irregular level set.

3.4. Proposition: Characterisation of M_-. (1) *We have $M_- = \{x \in S^{2l-1}_s \mid x = Px$ for some $P \in \Sigma\}$.*

(2) *If* $m_2<0$, *then* $f\equiv -1$ *and* $M_-=S_s^{2l-1}$.
(3) *If* $r=m$, *then* $M_-=\varnothing$.
In the following we suppose $l\geqq m>r$:
(4) *For* $P\in\Sigma(+1)$ *we set* $S(P):=E_+(P)\cap S_s^{2l-1}$.

$$\Sigma^*:=\{P\in\Sigma(+1)\,|\,S(P)\neq\varnothing\},\qquad \Gamma:=\{(P,x)\,|\,P\in\Sigma^*,\ x\in S(P)\}.$$

Then $\pi\colon \Gamma\to\Sigma^*$, $(P,x)\to P$ *is a De Sitter sphere bundle over the "Clifford sphere"* Σ^*.

If $r=m-1$ *and* $E_+(P_m)$ *has signature* $(l,0)$ *(resp.* (l,l)*), then* Σ^* *is the component of* $\Sigma(+1)$ *through* P_m *(*$-P_m$ *resp.). In all other cases* $\Sigma^*=\Sigma(+1)$.

(5) M_- *is diffeomorphic to the De Sitter sphere bundle* Γ, *where the diffeomorphism is given by* $(P,x)\to x$. M_- *is a submanifold of codimension* m_2+1 *in* S_s^{2l-1}. *If* $m_2=0$ *then* M_- *is a hypersurface.*

(6) *The normal space of* $M_-\subset S_s^{2l-1}$ *at* $x=Px$, $P\in\Sigma$, *is* $\perp_x M_-=E_-(P)\cap([P]^\perp\cdot x)^\perp$. *The tangent space is* $T_xM_-=(E_+(P)\cap[x]^\perp)\oplus[P]^\perp\cdot x$.

M_- *is a nondegenerate submanifold and* T_xM *has signature* $(l+m-2, s_1+r)$ *if* $E_+(P)$ *has signature* (l,s_1).

If $r=m-1$, $s_1\neq s_2$, s_1, $s_2\neq 0$ *(cf. 3.1.5) then* M_- *has no constant signature. We then set* $M_{-,i}:=\{x\in S_s^{2l-1}\,|\,x\in E_+(P), P\in\Sigma_i\}$, $i=1,2$ *and* $M_{-,i}$ *has signature* $(l+m-2, s_i+r)$.

(7) *If* $-1\neq c<+1$ *and* $t\in\mathbb{R}$ *such that* $c=-\mathrm{c}_\delta(4t)$, $\mathrm{s}_\delta(t)<0$, *then* $\Phi_t\colon \perp M_-(\delta)\to M_c$, $(x,v)\to \exp_x tv$ *is a diffeomorphism and the inverse is* $\Phi_t^{-1}(y)=(\varphi_t(y),\tau_t(N_y))$, *where* φ_t, τ_t *are as in Sect.* 1.

If $m_2>0$, *then* M_- *is the focal variety associated to the principal curvature* $\mathrm{ct}_\delta(t)$.

Proof. (1) If $x\in M_-$, then $x\in E_+(P)$, where $P:=\sum_j \eta_{jj}\langle P_jx,x\rangle P_j$. If $P\in\Sigma$, $x\in S_s^{2l-1}$ and $Px=x$, then $\langle P,P\rangle=1$ and P can be extended to a Clifford system. We can assume $P=P_m$ and obtain $f(x)=1-2\eta_{mm}\langle P_mx,x\rangle^2=-1$ and $\bar\nabla f(x)=8x-8\eta_{mm}\langle P_mx,x\rangle P_mx=0$.

(2) From (3.1.4) we have $r\leqq l<m$ and $P_m\in\Sigma(+1)$. For $x\in S_s^{2l-1}$ let $x=x_++x_-$ be the decomposition into eigenvectors of P_m, i.e. $x_\pm\in E_\pm(P_m)$. Without loss of generality assume $\langle x_+,x_+\rangle\neq 0$. Then the $(P_jx_+)_{j=1,\dots,m-1}$ form a basis of $E_-(P_m)$ and one computes $f(x)=-1$.

(3) If $r=m$ then $f(x)\geqq 1$.

(4) Setting $U:=\{Q\in\Sigma^*\,|\,\langle P,Q\rangle>0\}$ for $P\in\Sigma^*$ we find that $U\times S(P)\to\pi^{-1}(U)$, $(Q,x)\to(Q,(2+2\langle P,Q\rangle)^{-1/2}(x+Qx))$ is a local trivialisation near P.

(5) We already know that $\Gamma\to M_-$, $(P,x)\to x$ is surjective and it is also a submersion. Using (3.1.2) we see that it is also injective and therefore a diffeomorphism.

(6) Since $E_+(P)\cap S_s^{2l-1}\subset M_-$ we have $\perp_x M_-\subset E_-(P)$ and we find $Qx\in T_xM_-$ for $Q\in[P]^\perp$: Suppose $\langle Q,Q\rangle=:\varepsilon\in\{0,\pm 1\}$ and set $P(t):=\mathrm{c}_\varepsilon(t)P+\mathrm{s}_\varepsilon(t)Q$, then $x(t):=P(t)x$ is a curve in M_-, because $x(t)\in E_+(P(2t))$. But also $[P]^\perp\cdot x\subset E_-(P)$ and a dimension argument yields the assertion concerning the normal space from which we conclude the rest.

(7) Suppose $x \in E_+(P)$, $P \in \Sigma(+1)$. We can assume $P = P_m$ and obtain by (6):

$$\begin{aligned} f(\Phi_t(x,v)) &= 1 - 2\eta_{mm}(\mathrm{c}_\delta^2(t)\langle Px, x\rangle + 2\,\mathrm{c}_\delta(t)\,\mathrm{s}_\delta(t)\langle Px, v\rangle + \mathrm{s}_\delta^2(t)\langle Pv, v\rangle)^2 \\ &= 1 - 2(\mathrm{c}_\delta^2(t) - \delta\,\mathrm{s}_\delta^2(t))^2 = -\mathrm{c}_\delta(4t). \end{aligned}$$

Therefore $\Phi_t(\perp M_-(\delta)) \subset M_c$. With (3.3.1) one computes for $y \in M_c$:

$$\varphi_t(y) = (2\,\mathrm{c}_\delta(t))^{-1}\,y + (2\,\mathrm{c}_\delta(t)\,\mathrm{c}_\delta(2t))^{-1} \sum_j \eta_{jj}\langle P_j y, y\rangle P_j y,$$

$$\tau_t(N_y) = (2\,\mathrm{s}_\delta(t))^{-1}\,y - (2\,\mathrm{s}_\delta(t)\,\mathrm{c}_\delta(2t))^{-1} \sum_j \eta_{jj}\langle P_j y, y\rangle P_j y$$

and for $(x,v) \in \perp M_-(\delta)$ (assume $x \in E_+(P_m)$):

$$N_{\Phi_t(x,v)} = \delta\,\mathrm{s}_\delta(t)\,x - \mathrm{c}_\delta(t)\,v.$$

Setting $P := (\mathrm{c}_\delta(2t))^{-1}\sum_j \eta_{jj}\langle P_j y, y\rangle P_j$ we find $\varphi_t(y) \in E_+(P)$, $\tau_t(N_y) \in E_-(P)$ and $\langle \tau_t(N_y), Q\varphi_t(y)\rangle = 0$ for all $Q \in [P]^\perp$ by (3.3.2). Therefore $\Phi_t^{-1}(M_c) \subset \perp M_-(\delta)$. Now one easily verifies $\Phi_t \circ \Phi_t^{-1}(y) = y$ and $\Phi_t^{-1} \circ \Phi_t(x,v) = (x,v)$. If $m_2 > 0$, then φ_t is singular and the assertion follows from (2.7.3). □

3.5. Proposition: Characterisation of M_+. (1) *We have* $M_+ = \{x \in S_s^{2l-1} \mid \langle P_j x, x\rangle = 0, j = 1, \ldots, m\}$.

If $M_+ \neq \varnothing$, then M_+ is a nondegenerate submanifold of codimension $m_1 + 1$ in S_s^{2l-1}.

The normal bundle is trivial: $\perp_x M_+ = \Sigma \cdot x$.

(2) *If $-1 < c \neq +1$ and $t \in \mathbb{R}$ such that $c = \mathrm{c}_\delta(4t)$, $\mathrm{s}_\delta(4t) > 0$, then $\Phi_t\colon \perp M_+(\delta) \to M_c$, $(x,v) \to \exp_x tv$ is a diffeomorphism and the inverse is $\Phi_t^{-1}(y) = (\varphi_t(y), \tau_t(N_y))$.*

M_+ is the focal variety associated to the principal curvature $\mathrm{ct}_\delta(t)$.

Proof. (1) is clear from (3.2.1) and the regular value theorem since the $(P_j x)_{j=1,\ldots,m}$ are orthonormal (3.1.2).

(2) We have $v = Px$ for some $P \in \Sigma(\delta)$ and P can be extended to a Clifford system. We can assume $P = P_i$, $\eta_{ii} = \delta$ and obtain

$$\begin{aligned} f(\Phi_t(x,v)) &= 1 - 2\sum_j \eta_{jj} \\ &\quad \cdot (\mathrm{c}_\delta^2(t)\langle P_j x, x\rangle + 2\,\mathrm{c}_\delta(t)\,\mathrm{s}_\delta(t)\langle P_j x, P_i x\rangle + \mathrm{s}_\delta^2(t)\langle P_j P_i x, P_i x\rangle)^2 \\ &= 1 - 2\eta_{ii}(\mathrm{s}_\delta(2t)\,\eta_{ii}\langle x, x\rangle)^2 = \mathrm{c}_\delta(4t). \end{aligned}$$

Therefore $\Phi_t(\perp M_+(\delta)) \subset M_c$. Using (3.3.2) we verify $\langle P_j \varphi_t(y), \varphi_t(y)\rangle = 0$, i.e. $\varphi_t(M_c) \subset M_+$. One now computes with formula (3.3.1) for $y \in M_c$:

$$\tau_t(N_y) = -\delta\,\mathrm{s}_\delta(t)(\mathrm{c}_\delta(2t))^{-1}\,y - (2\,\mathrm{s}_\delta(t)\,\mathrm{c}_\delta(2t))^{-1} \sum_j \eta_{jj}\langle P_j y, y\rangle P_j y,$$

$$\varphi_t(y) = \mathrm{c}_\delta(t)(\mathrm{c}_\delta(2t))^{-1}\,y - (2\,\mathrm{c}_\delta(t)\,\mathrm{c}_\delta(2t)^{-1} \sum_j \eta_{jj}\langle P_j y, y\rangle P_j y$$

and for $(x,v) \in \perp M_+(\delta)$: $N_{\Phi_t(x,v)} = \delta\,\mathrm{s}_\delta(t)\,x - \mathrm{c}_\delta(t)\,v$.

Setting $P:=(s_\delta(2t))^{-1}\sum_j \eta_{jj}\langle P_j y,y\rangle P_j$ we find $\tau_t(N_y)=P\varphi_t(y)$, i.e. $\tau_t(N_y)\in\perp_{\varphi_t(y)}M_+(\delta)$.

Now one easily verifies $\Phi_t\circ\Phi_t^{-1}(y)=y$ and $\Phi_t^{-1}\circ\Phi_t(x,v)=(x,v)$. Since φ_t is singular the assertion follows from (2.7.3). □

By this time we know the topological structure of the level hypersurfaces. (3.4.7), (3.5.2) permit to evaluate the principal curvatures of the level surfaces M_c, $-1<c<+1$. In order to compute also the principal curvatures of the levels M_c, $c<-1$ or $c>+1$, we investigate the shape operators of the focal varieties M_+, M_-.

3.6. Lemma: Shape Operator on M_+. *Let $x\in M_+$ and $N\in\perp_x M_+(\delta)$, i.e. $N=Px$ for some $P\in\Sigma(\delta)$, $\delta=\pm1$. If $\delta=+1$, then the shape operator S_N associated to N is real-diagonalizable and the eigenvalues are (with multiplicities): $1(m_2)$, $-1(m_2)$, $O(m_1)$.*

If $\delta=-1$, then S_N is complex-diagonalizable and the eigenvalues are (with multiplicities): $i(m_2)$, $-i(m_2)$, $O(m_1)$.

Proof. We have the orthogonal decomposition of the tangent space $T_xM_+=[P]^\perp\cdot N\oplus(\Sigma\cdot x\oplus\Sigma\cdot N)^\perp$ and from $S_N v=-(Pv)^{\tan M_+}$ we obtain $\operatorname{Ker}S_N=[P]^\perp\cdot N$ and $S_N|_{(\Sigma\cdot x\oplus\Sigma\cdot N)^\perp}=-P|_{(\Sigma\cdot x\oplus\Sigma\cdot N)^\perp}$, since $\Sigma x\oplus\Sigma N$ is an invariant subspace of P. □

3.7. Lemma: Shape Operator on M_-. *Let $x\in M_-$, i.e. $x\in E_+(P)$ for some $P\in\Sigma$ and $N\in\perp_x M_-(\delta)$, $\delta=\pm1$. If $\delta=+1$, then the shape operator S_N associated to N is real-diagonalizable and the eigenvalues are (with multiplicities): $1(m_1)$, $-1(m_1)$, $O(m_2)$. If $\delta=-1$, then S_N is complex-diagonalizable and the eigenvalues are (multiplicities): $i(m_1)$, $-i(m_1)$, $O(m_2)$.*

Proof. From (3.4.6) we obtain the orthogonal decomposition of the tangent space $T_xM_-=[P]^\perp\cdot x\oplus[P]^\perp\cdot N\oplus K$, where $K:=E_+(P)\cap[x]^\perp\cap([P]^\perp\cdot N)^\perp$. If $v\in K$ we have by (3.4.6) $\langle d_vN,Qx\rangle=-\langle N,Qv\rangle=0$ for all $Q\in[P]^\perp$, i.e. $d_vN\in\perp_xM_-$ and $S_Nv=0$.

Now let $Q\in[P]^\perp$, $\langle Q,Q\rangle=:\varepsilon=\pm1$. As in the proof of (3.4.6) $x(t):=c_\varepsilon(t)x+s_\varepsilon(t)Qx$ is a curve in direction Qx in M_- and one proves that $N(t):=c_\varepsilon(t)N-s_\varepsilon(t)QN$ is a normal field along $x(t)$. Then $S_NQx=QN$.

Similarly we find $S_NQN=\delta Qx$. Therefore $S_N^2|_{[Qx,QN]}=\delta$ and the eigenvalues are ±1, $\pm i$ resp. □

We now collect our results:

3.8. Theorem. (1) *The level hypersurface M_c, $-1\neq c<+1$ is diffeomorphic to the De Sitter sphere bundle $\perp M_-(\delta)$, where $\delta=+1$ if $-1<c<+1$ and $\delta=-1$ if $c<-1$.*

(2) *The level hypersurface M_c, $-1<c\neq+1$ is diffeomorphic to the De Sitter sphere bundle $\perp M_+(\delta)$, where $\delta=+1$ if $-1<c<+1$ and $\delta=-1$ if $c>+1$.*

(3) *If $f(S_s^{2l-1})\cap]-1,+1[\neq\emptyset$, then the level surfaces M_c, $-1<c<+1$ form an isoparametric family of type $\delta=+1$ with 4 distinct real principal curvatures. The hypersurface M_c, $c=\cos 4t_0$, $0<t_0<\frac{\pi}{4}$, has the principal curvatures $k_j=\cot(t_0-(j-1)\frac{\pi}{4})$, $j=1,\dots,4$ with multiplicities (m_1,m_2,m_1,m_2). M_+ is*

the focal variety associated to k_1, k_3 *and* M_- *is associated to* k_2, k_4. *(If* $m_2=0$ *we have only the focal variety* M_+ *and the two principal curvatures* k_1, k_3*).*

(4) *If* $f(S_s^{2l-1}) \cap]+1, \infty[\neq \varnothing$, *then the level surfaces* $M_c, c>1$ *form an isoparametric family of type* $\delta=-1$ *with 4 distinct principal curvatures. The hypersurface* M_c, $c=\cosh 4t_0$, $t_0>0$ *has the real principal curvatures* $k_1=\coth t_0$ *and* $k_3=\tanh t_0$, *both with multiplicity* m_1, *and the complex principal curvatures* $k_2=\coth(t_0+\frac{\pi}{4}i)$ *and* $k_4=\coth(t_0+\frac{3}{4}\pi i)$, *both with multiplicity* m_2. M_+ *is the focal variety associated to* k_1. *(If* $m_2=0$ *we have only the real principal curvatures* k_1, k_3*)*

(5) *If* $f(S_s^{2l-1}) \cap]-\infty, -1[\neq \varnothing$, *then the level surfaces* M_c, $c<-1$ *form an isoparametric family of type* $\delta=-1$ *with 4 distinct principal curvatures. The hypersurface* M_c, $c=-\cosh 4t_0$, $t_0<0$ *has the real principal curvatures* $k_1 =\coth t_0$ *and* $k_3=\tanh t_0$, *both with multiplicity* m_2, *and the complex principal curvatures* $k_2=\coth(t_0+\frac{\pi}{4}i)$ *and* $k_4=\coth(t_0+\frac{3}{4}\pi i)$, *both with multiplicity* m_1. M_- *is the focal variety associated to* k_1.

(If $m_2=0$ *we have only the two complex principal curvatures* k_2, k_4 *without focal variety).*

Proof. (1), (2) were already proved in (3.4.7), (3.5.2).

(3) Similarly as in (3.5.2) we compute $\Phi_t \circ \Phi_{t_0}^{-1}(x)=\varphi_{t_0-t}(x)$ for $x\in M_c$, $t\in\mathbb{R}$. Therefore $\varphi_{t_0-t}(M_c)=M_+$ for $t=0, \frac{\pi}{2}$ and $\varphi_{t_0-t}(M_c)=M_-$ for $t=\frac{\pi}{4}$, $\frac{3}{4}\pi$. The assertion then follows from (2.7.3).

(4) We now have $\varphi_{t_0}(M_c)=M_+$ (3.5.2) and conclude with (3.6), (2.7.3).

(5) is proved similarly to (4), but with (3.4.7), (3.7) and (2.7.3). □

3.9. *Note.* From (3.4), (3.8) we can compute - depending on the signature of the Clifford system - the range of f and the number of components of M_c. We give only the cases in which the family M_c, $-1<c<+1$ occurs (see Table 1).

Table 1

Cliffordsystem	number of components of		
	M_-	M_+	M_c, $-1<c<+1$
$r=0$, $s=2l-2m$	1	2	2
$r=0$, $s<2l-2m$	1	1	1
$1\leqq r\leqq m-2$, $2m-2r=l$	1	. 2	2
$1\leqq r\leqq m-2$, $2m-2r<l$	1	1	1
$r=m-1$, $2\leqq s_1, s_2\leqq l-2$	2	1	2
$r=1$, $m=2$, $l>2$, $\{s_1, s_2\}=\{1, l-1\}$	3	2	4

Every component of M_c is again isoparametric.

If the focal variety M_- has no constant signature, then the signature of the eigenspace of the associated principal curvature is not constant on the hypersurface, which is therefore disconnected. In case $-1<c<+1$ there are two principal curvatures associated to M_- and the signatures of their eigenspaces are different. A component of the hypersurface is again isoparametric and the focal varieties of the component associated to these principal curvatures are connected and have different signatures and are therefore distinct. This component yields an isoparametric family with at least 3 distinct focal varieties!

3.10. *Examples.* (1) If the Clifford system satisfies $l>m$, $r=m-1$, $2\leqq s_1$, $s_2\leqq l-2$, then M_- and M_c, $-1<c<+1$ have 2 components and M_+ is connected. From (3.5) we have the diffeomorphism ($t_0\in]0,\frac{\pi}{4}[$) Ψ_{t_0}: $M_+\times\Sigma(+1)\to M_{\cos 4t_0}$, $(x,P)\to\cos t_0\, x+\sin t_0\, Px$ and $M_{c,1}:=\Psi_{t_0}(M_+\times\Sigma_1)$ is a connected isoparametric hypersurface with three focal varieties: As in (3.8) we have

$$\varphi_{t_0-t}(x)=\Psi_t\circ\Psi_{t_0}^{-1}(x) \text{ and find}$$

$$\begin{aligned}\varphi_{t_0-t}(M_{c,1})&=M_+ &&\text{for } t=n\tfrac{\pi}{2},\quad n\in\mathbb{Z},\\ &=M_{-,1} &&\text{for } t=(4n+1)\tfrac{\pi}{4}\\ &=M_{-,2} &&\text{for } t=(4n+3)\tfrac{\pi}{4}.\end{aligned}$$

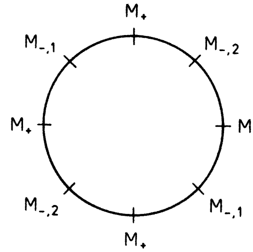


Fig. 1

M_+ is the focal variety associated to the principal curvatures $k_1=\cot t_0$ and $k_3=\cot(t_0-\frac{\pi}{2})$, $M_{-,1}$ is associated to $k_2=\cot(t_0-\frac{\pi}{4})$ and $M_{-,2}$ is associated to $k_4=\cot(t_0-\frac{3}{4}\pi)$.

(2) We now display an isoparametric family with five focal varieties, where the normal circle through any point of the hypersurface consecutively meets all these focal varieties. If the Clifford system satisfies $m=2$, $r=1$, $l>2$, $s_1=1$, $s_2=l-1$, then the bundle Γ (cf. 3.4) has 3 components. Let Γ_1 denote the component over Σ_1 and Γ_2,Γ_3 the components over Σ_2.

Now define $M_{-,i}:=\{x\in S_s^{2l-1}\mid x\in E_+(P),\,(P,x)\in\Gamma_i\}$ $(i=1,2,3)$ and

$$M_{+,1}:=\left\{x\in M_+\,\middle|\,\left(-P,\frac{1}{\sqrt{2}}(x-Px)\right)\in\Gamma_2 \text{ for } P\in\Sigma_1\right\},$$

$$M_{+,2}:=\left\{x\in M_+\,\middle|\,\left(-P,\frac{1}{\sqrt{2}}(x-Px)\right)\in\Gamma_3 \text{ for } P\in\Sigma_1\right\}.$$

Similarly as before we set $M_{c,1}:=\Psi_{t_0}(M_{+,1}\times\Sigma_1)$ and find

$$\begin{aligned}\varphi_{t_0-t}(M_{c,1})&=M_{+,1} &&\text{for } t=0,\tfrac{3}{2}\pi\\ &=M_{+,2} &&\text{for } t=\tfrac{\pi}{2},\pi\\ &=M_{-,1} &&\text{for } t=\tfrac{\pi}{4},\tfrac{5}{4}\pi\\ &=M_{-,2} &&\text{for } t=\tfrac{7}{4}\pi\\ &=M_{-,3} &&\text{for } t=\tfrac{3}{4}\pi.\end{aligned}$$

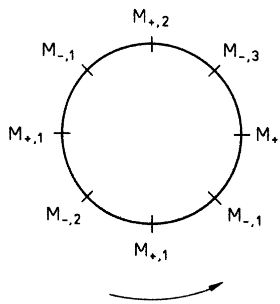


Fig. 2

Now it is not possible to speak of *the* focal variety associated to a principal curvature. The principal curvatures $k_1 = \cot t_0$, $k_3 = \cot(t_0 - \frac{\pi}{2})$ both lead to the focal varieties $M_{+,1}$, $M_{+,2}$, $k_2 = \cot(t_0 - \frac{\pi}{4})$ leads to $M_{-,1}$ and $k_4 = \cot(t_0 - \frac{3}{4}\pi)$ leads to $M_{-,2}$, $M_{-,3}$.

4. Totally Isotropic Clifford Systems

We now modify the definition of Clifford systems and obtain isoparametric hypersurfaces with only one principal curvature whose geometric multiplicity is not constant.

A *totally isotropic Clifford system* on $\mathbb{R}_s^{n+1}$ is an m-tuple $(P_1, \ldots, P_m)$ such that $P_i \in \mathrm{Sym}(\mathbb{R}_s^{n+1})$ and $P_i P_j + P_j P_i = 0$ for $i, j = 1, \ldots, m$. e.g. every finite subset of a totally isotropic subspace of a Clifford span is a totally isotropic Clifford system.

Proposition. *The function* $f\colon S_s^n \to \mathbb{R}$, $f(x) = \sum_{j=1}^{m} \langle P_j x, x\rangle^2$ *is isoparametric. The level surfaces* $M_c := f^{-1}(c)$, $c > 0$ *form an isoparametric family of type* $\delta = -1$. *The only principal curvature is* $k = +1$ *independent of* $c > 0$. *Its geometric multiplicity can vary.*

Proof. By the homogeneity of f we find $\langle \bar{\nabla} f(x), \bar{\nabla} f(x)\rangle = -16 f(x)^2$ since $\langle P_i x, P_j x\rangle = 0$ and $\bar{\Delta} f(x) = -4(n+3) f(x)$ since trace $P_j = 0$.

Therefore f is isoparametric and (2.2) shows that the mean curvature of M_c is $=1$, independent of c. Since the principal curvatures are determined by the mean curvature of M_c as c varies (see the proof of (2.1) given in [1], [5]) the only principal curvature is $k = +1$. From the following example we see that the geometric multiplicity can vary. $\square$

Example. Let $f\colon S_2^3 \to \mathbb{R}$, $f(x) = (x_1 + x_3)^4 + (x_2 + x_4)^4$. We compute that the geometric multiplicity of the principal curvature $k = 1$ at $x \in M_c (c > 0)$ is $=2$ if $x_1 + x_3 = 0$ or $x_2 + x_4 = 0$ and $=1$ otherwise.

References

1. Ferus, D.: Notes on Isoparametric Hypersurfaces. Escola de Geometria Differencial, Universidade Estadual de Campinas (1980)
2. Ferus, D., Karcher, H., Münzner, H.F.: Cliffordalgebren und neue isoparametrische Hyperflächen. Math. Z. **177**, 479–502 (1981)
3. Graves, L.K.: Codimension one isometric immersions between Lorentz spaces. Trans. Amer. Math. Soc. **252**, 367–392 (1979)
4. Münzner, H.F.: Isoparametrische Hyperflächen in Sphären I. Math. Ann. **251**, 57–71 (1980)
5. Nomizu, K.: Elie Cartan's work on isoparametric families of Hypersurfaces. Proc. Symposia in Pure Math (AMS) **27**, 191–200 (1975)
6. Nomizu, K.: On isoparametric hypersurfaces in the Lorentzian space forms. Japan J. of Math. **7**, 217–226 (1981)
7. Wolf, J.: Spaces of constant curvature. New York: Mc Graw Hill 1967

Received October 20, 1983

Math. Z. 187, 209–219 (1984)

Mathematische Zeitschrift

Über 2-codimensionale Untermannigfaltigkeiten vom Grad 7 in $\mathbb{P}^4$ und $\mathbb{P}^5$

Christian Okonek

Mathematisches Institut der Universität, Bunsenstr. 3–5, D-3400 Göttingen, Bundesrepublik Deutschland

0. Einleitung

Sei k ein algebraisch abgeschlossener Körper der Charakteristik $0, \mathbb{P}^n = \mathbb{P}^n_k$ der n-dimensionale projektive Raum über k. Es wird vermutet [4], daß 2-codimensionale Untermannigfaltigkeiten im $\mathbb{P}^n$ für $n \geqq 6$ stets vollständige Durchschnitte sind. In dieser Arbeit betrachten wir die Dimensionen $n=4, 5$. In [10], [11] haben wir alle 2-codimensionalen Untermannigfaltigkeiten in $\mathbb{P}^4$ und $\mathbb{P}^5$ vom Grad $d \leqq 6$ klassifiziert (vgl. auch [6]). Es wurde ferner eine Vermutung von Hartshorne erwähnt, nach der es keine glatten, rationalen Flächen im $\mathbb{P}^4$ vom Grad $d > 6$ geben soll. Diese Vermutung war der Anstoß zu der vorliegenden Arbeit. Wir zeigen, daß für den Grad d einer glatten, rationalen Fläche $Y \subset \mathbb{P}^4$ mit $h^1(\mathcal{O}_Y(1)) = 0$ notwendig $d \leqq 12$ gilt. Für $d=7$ konstruieren wir dann eine Familie Y_A solcher Flächen und widerlegen damit die oben erwähnte Vermutung. Man erhält Y_A durch Aufblasung von 11 Punkten x_i des $\mathbb{P}^2$; die Einbettung in den $\mathbb{P}^4$ wird durch das lineare System

$$\left|6L - \sum_{i=1}^{6} 2x_i - \sum_{j=7}^{11} x_j\right|$$

der Kurven vom Grad 6 gegeben, die durch diese Punkte gehen, und zwar mit Multiplizität 2 durch 6 davon.

Anschließend klassifizieren wir die restlichen glatten Flächen vom Grad 7. In [8] hat Lanteri die Struktur der möglichen Flächen bestimmt, ohne allerdings in jedem Fall ihre Existenz zu beweisen (vgl. auch [3]). Es gibt noch zwei weitere Familien, Flächen vom Typ B und C. Flächen vom Typ B entstehen durch Projektion eines vollständigen Durchschnittes $S_{(2,2,2)} \subset \mathbb{P}^5$ von einem generischen Punkt auf $S_{(2,2,2)}$, sind also $K3$ Flächen. Flächen Y_C vom Typ C sind elliptisch mit Kodaira-Dimension $\kappa = 1$; es gibt eine Ebene E, so daß $Y_C \cup E$ der vollständige Durchschnitt einer Quadrik und einer Quartik ist.

Wir bestimmen die lokal freien Auflösungen der Idealgarben dieser Familien.

Man kann nun diese Auflösungen auf $\mathbb{P}^5$ fortsetzen. Mit Hilfe eines Satzes von Kleiman [7] bekommt man so 3-Mannigfaltigkeiten im $\mathbb{P}^5$, deren generische Hyperebenenschnitte zu einer der 3 oben genannten Familien gehören. Wir zeigen, daß es genau 3 Typen von 3-Mannigfaltigkeiten $X \subset \mathbb{P}^5$ vom Grad 7 gibt und untersuchen ihre Struktur. Die einfachste Familie X_A bekommt man als projektive Bündel $\mathbb{P}(E)$ einer lokal freien Garbe E vom Rang 2 auf einer nichtsingulären Kubik im $\mathbb{P}^3$. Schließlich konstruieren wir zu diesen 3-Mannigfaltigkeiten stabile bzw. semistabile reflexive Garben vom Rang 2 mit homologischer Dimension 1 auf $\mathbb{P}^5$.

1. Rationale Flächen im $\mathbb{P}^4$

Sei $Y \subset \mathbb{P}^4$ eine glatte Fläche, H ein Hyperebenendivisor, K ein kanonischer Divisor. Ein generischer Hyperebenenschnitt

$$C = Y \cap \mathbb{P}^3$$

ist eine glatte, zusammenhängende Kurve, deren geometrisches Geschlecht wir mit π bezeichnen. Ist $d = H^2$ der Grad von Y, p_a das arithmetische Geschlecht von Y, so schreibt sich das Hilbertpolynom von Y in der Form

$$p_Y(l) = \tfrac{1}{2} d\, l^2 + (\tfrac{1}{2} d + 1 - \pi)\, l + 1 + p_a.$$

Die Adjunktionsformel [1] liefert

$$H \cdot (H + K) = 2\pi - 2,$$

also

$$H \cdot K = 2\pi - 2 - d.$$

Wie üblich schreiben wir $p_g = h^0(\mathcal{O}_Y(K))$ für das geometrische Geschlecht und $q = h^1(\mathcal{O}_Y)$ für die Irregularität von Y. Es ist

$$q = p_g - p_a.$$

Es gilt nun [5], p. 434:

$$d^2 - 10d - 5H \cdot K - 2K^2 + 12(1 + p_a) = 0.$$

Nach Severi [12] ist die einzige glatte Fläche im $\mathbb{P}^4$, die nicht linear normal ist, die Veronese Fläche vom Grad 4.

Für alle anderen Flächen ist das lineare System $|H|$ vollständig. Wir interessieren uns zunächst für rationale Flächen.

Proposition 1. *Sei $Y \subset \mathbb{P}^4$ eine nicht entartete, rationale Fläche mit $h^1(\mathcal{O}_Y(H)) = 0$. Dann gilt für den Grad d von Y*

$$d \leqq 12.$$

Beweis. Da die Veronese Fläche den Grad 4 hat, dürfen wir annehmen, daß Y linear normal ist. Die Riemann-Roch Formel für H lautet:

$$\tfrac{1}{2} H \cdot (H - K) + 1 + p_a = h^0(\mathcal{O}_Y(H)) - h^1(\mathcal{O}_Y(H)) + h^2(\mathcal{O}_Y(H)).$$

Da Y nach Voraussetzung rational ist, gilt $p_a=0$ und

$$h^2(\mathcal{O}_Y(H))=h^0(\mathcal{O}_Y(K-H)\leqq p_g=0.$$

Aus der Voraussetzung $h^1(\mathcal{O}_Y(H))=0$ bekommt man deshalb

$$\tfrac{1}{2}H\cdot(H-K)+1=h^0(\mathcal{O}_Y(H))=5 \quad \text{oder} \quad H\cdot K=d-8.$$

Ferner haben wir die Gleichung [5]

$$d^2-10d-5H\cdot K-2K^2+12=0. \tag{1}$$

Wir setzen $K^2=9-s$ mit $s\geqq 0$. Es folgt

$$\frac{d(d-15)}{2}=-(17+s). \tag{2}$$

Insbesondere gilt $d\leqq 12$.

Bemerkung 2. Wegen $s\geqq 0$ hat (2) nur endlich viele Lösungen. Wir haben sie in folgender Tabelle zusammengestellt:

d	3	4	5	6	7	8	9	10	11	12
s	1	5	8	10	11	11	10	8	5	1

Die Klassifikation der Flächen im $\mathbb{P}^4$ bis zum Grad 6 zeigt, daß die Paare (d, s) mit $d\leqq 6$ realisiert werden können.

Wir konstruieren als nächstes eine glatte Fläche $Y\subset\mathbb{P}^4$ vom Grad 7, die isomorph ist zur Aufblasung von $\mathbb{P}^2$ in 11 Punkten.

Proposition 3. *Es gibt eine glatte, rationale Fläche vom Grad 7 in $\mathbb{P}^4$.*

Beweis. Sei Y der Abhängigkeitsort von 4 generischen Schnitten in $\mathcal{O}(1)\oplus\Omega^1_{\mathbb{P}^4}(2)$. Nach [7] ist Y glatt. Die Invarianten von Y sind $d=7$, $\pi=4$ und $p_a=0$. Aus der Adjunktionsformel folgt $H\cdot K=-1$, aus (1) $K^2=-2$. Y ist also eine rationale Regelfläche.

Das Linearsystem $|K+H|$ hat keine Basispunkte [10], wir können Y daher mit Hilfe des Morphismus

$$\Phi=\Phi_{|K+H|}\colon Y\to\mathbb{P}^3$$

untersuchen.

Es ist $(K+H)^2=3$, also $\bar{Y}=\Phi(Y)\subset\mathbb{P}^3$ eine Kubik. Sei $C\subset Y$ ein generischer Hyperebenenschnitt. Die Einschränkung von Φ auf C ist durch das kanonische Linearsystem auf C gegeben, $\bar{C}=\Phi(C)\subset\mathbb{P}^3$ ist eine kanonische Kurve (da C nach [8] nicht hyperelliptisch sein kann).

Nach [8] Proposition 4.2 ist $\bar{Y}$ nicht singulär, $\bar{C}$ ein vollständiger Durchschnitt von $\bar{Y}$ mit einer Quadrik.

Der Morphismus

$$\Phi\colon Y\to\bar{Y}$$

ist dann eine Aufblasung von $\bar{Y}$ in 5 Punkten y_i. Zu $|H|$ gehört unter Φ ein System von kanonischen Kurven vom Geschlecht 4 auf $\bar{Y}$ mit 5 Basispunkten y_i, ausgeschnitten durch Quadriken. Stellt man $\bar{Y}$ als Aufblasung von $\mathbb{P}^2$ in

6 Punkten dar, so schneiden die kanonischen Kurven jede der zugehörigen 6 exzeptionellen Geraden in 2 Punkten. Auf $\mathbb{P}^2$ wird $|H|$ also dargestellt durch das Linearsystem

$$\left|6L-\sum_{i=1}^{6} 2E_i-\sum_{j=7}^{11} E_j\right|$$

der Kurven vom Grad 6 durch 11 Punkte, die 6 Doppelpunkte haben.

Theorem 4. *Sei $Y \subset \mathbb{P}^4$ eine glatte Fläche vom Grad* 7 *mit $\pi=4$. Dann ist Y isomorph zur Aufblasung von $\mathbb{P}^2$ in* 11 *Punkten, eingebettet durch*

$$\left|6L-\sum_{i=1}^{6} 2E_i-\sum_{j=7}^{11} E_j\right|.$$

Jede dieser Flächen hat die Auflösung

$$0 \to \mathcal{O}^{\oplus 4} \to \mathcal{O}(1) \oplus \Omega^1_{\mathbb{P}^4}(2) \to \mathscr{I}_Y(4) \to 0.$$

Beweis. Es bleibt zu zeigen, daß die Idealgarben dieser Flächen stets durch die Auflösung

$$0 \to \mathcal{O}^{\oplus 4} \to \mathcal{O}(1) \oplus \Omega^1_{\mathbb{P}^4}(2) \to \mathscr{I}_Y(4) \to 0$$

gegeben sind. Nach [10], Theorem 2.5 existiert jedenfalls eine exakte Sequenz

$$0 \to \mathcal{O} \to F(2) \to \mathscr{I}_Y(4) \to 0$$

mit einer semistabilen, reflexiven Garbe F mit $hd(F) \leqq 1$ und den Chernklassen

$$c_1=0, \qquad c_2=3, \qquad c_3=6, \qquad c_4=15.$$

Man zeigt leicht, daß F stabil sein muß (es wäre sonst $hd(F)=2$).

Ferner gilt $h^1(F_H(-1))=0$ für eine generische Hyperebene. Mit Hilfe von [2], Theorem 3.5 sieht man nun sofort, daß F_H das Spektrum $(-1, -1, -1)$ haben muß und findet für $F(2)$ die Auflösung

$$0 \to \mathcal{O}^{\oplus 3} \to \mathcal{O}(1) \oplus \Omega^1_{\mathbb{P}^4}(2) \to F(2) \to 0.$$

2. Flächen vom Grad 7 im $\mathbb{P}^4$

In diesem Abschnitt klassifizieren wir die glatten Flächen vom Grad 7 im $\mathbb{P}^4$.

Lemma 5. *Sei $Y \subset \mathbb{P}^4$ eine Fläche vom Grad* 7. *Dann gilt für das Geschlecht π eines glatten Hyperebenenschnittes*

$$4 \leqq \pi \leqq 6.$$

Ist $\pi \geqq 5$, so ist $p_g > 0$.

Beweis. Siehe [8].

Wir können nun die Korrespondenz von Flächen und reflexiven Garben verwenden, um die Idealgarben zu bestimmen [9]. Sei $Y \subset \mathbb{P}^4$ eine glatte Fläche mit $d=7$, $\pi \geqq 5$. Dann gilt $p_g > 0$, es gibt also eine Extension

$$0 \to \mathcal{O} \xrightarrow{s} F(3) \longrightarrow \mathscr{I}_Y(5) \to 0. \tag{3}$$

Hier ist F eine reflexive Garbe vom Rang 2 mit $hd(F) \leqq 1$ und den Chernklassen

$$c_1 = -1, \quad c_2 = 1, \quad c_3 = 2\pi - 9, \quad c_4 = K^2 + 6\pi - 27.$$

Wir unterscheiden 2 Fälle;

i) F ist nicht stabil: Dann ist Y eine minimale Fläche mit Kodaira Dimension $\kappa = 1$, wie wir in [10]. Theorem 3.1 gezeigt haben; speziell gilt $\pi = 6$.

ii) F ist stabil: Dann folgt [9] $c_3 \leqq 1$, $c_4 = 2$, also $\pi = 5$, $K^2 = -1$, $p_a = 1$.

In diesem Fall erhält man aus der Klassifikation [9] dieser Garben die Sequenz

$$0 \to \mathcal{O}(1) \to \mathcal{O}(2)^{\oplus 3} \to F(3) \to 0.$$

Zusammen mit (3) liefert dies das kommutative Diagramm

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \mathcal{O}(1) \arrow[r] \arrow[d, equal] & \mathcal{O} \oplus \mathcal{O}(1) \arrow[r] \arrow[d] & \mathcal{O} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{O}(1) \arrow[r] & \mathcal{O}(2)^{\oplus 3} \arrow[r] \arrow[d] & F(3) \arrow[r] \arrow[d] & 0 \\
 & & \mathcal{J}_Y(5) \arrow[r, equal] \arrow[d] & \mathcal{J}_Y(5) \arrow[d] & \\
 & & 0 & 0 &
\end{tikzcd}$$

F ist – bis auf Isomorphie – festgelegt durch die Singularitätenmenge $S(F)$, eine Gerade im $\mathbb{P}^4$. Der Schnitt s ist durch 3 Quadriken gegeben. Wählt man diese Quadriken generisch, so trifft ihre Durchschnittskurve $S(F)$ nicht, der zugehörige Schnitt s hat also ein glattes Nullstellenschema.

Man sieht leicht, daß ein effektiver Divisor $E \subset |K|$ eine Gerade ist $(H \cdot K = 1)$, die Adjunktionsformel liefert $E^2 = -1$.

Die mittlere Spalte des Diagrammes liefert die Auflösung

$$0 \to \mathcal{O} \oplus \mathcal{O}(1) \to \mathcal{O}(2)^{\oplus 3} \to \mathcal{J}_Y(5) \to 0.$$

Durch Dualisieren erhält man

$$0 \to \mathcal{O}(-5) \to \mathcal{O}(-2)^{\oplus 3} \to \mathcal{O} \oplus \mathcal{O}(-1) \to \omega_Y \to 0.$$

Es folgt $h^0(\mathcal{O}_Y(K)) = 1$, $\operatorname{supp}(K) = E = S(F)$. Durch Zusammenblasen von E erhält man also eine nicht singuläre Fläche Y_0 mit $1 = p_a(Y_0) = p_g(Y_0)$, $q(Y_0) = 0$ und $\omega_{Y_0} \cong \mathcal{O}_{Y_0}$. Y_0 enthält keine exzeptionellen Geraden, ist daher eine minimale $K3$-Fläche.

Wir betrachten den Morphismus

$$\Phi = \Phi_{|K+H|}\colon Y \to \mathbb{P}^5.$$

Φ kontrahiert E, faktorisiert also über Y_0:

$$\begin{tikzcd} Y \arrow[r, "\Phi"] \arrow[d, "\pi"'] & \mathbb{P}^5 \\ Y_0 \arrow[ur, "\Phi_0"'] & \end{tikzcd}$$

Das Bild ist eine Fläche vom Grad $8 = (K+H)^2$ im $\mathbb{P}^5$. Es ist klar, daß $|K+H|$ sehr ampel ist außerhalb von E, Φ_0 ist daher eine Einbettung. Lanteri [8] hat gezeigt, daß $\Phi_0(Y_0)$ ein vollständiger Durchschnitt von 3 Quadriken im $\mathbb{P}^5$ ist, und Y durch Projizieren von einem Punkt dieses vollständigen Durchschnittes entsteht.

Geht man umgekehrt von einem generischen vollständigen Durchschnitt $Y_0 = S_{(2,2,2)}$ von 3 Quadriken im $\mathbb{P}^5$ aus, so ist Y_0 eine $K3$-Fläche. Wie in [1] Proposition IV.16 sieht man, daß es nur endlich viele Geraden auf Y_0 gibt. Wählt man einen Punkt $y_0 \in Y_0$, der nicht auf diesen Geraden liegt, so liefert die Projektion eine glatte Fläche Y im $\mathbb{P}^4$ vom Grad 7 ([1], Lemma IV.4).

Wir haben gezeigt (für C) vgl. [10], Theorem 3.1.):

Theorem 6. *Sei $Y \subset \mathbb{P}^4$ eine nicht entartete glatte Fläche vom Grad 7. Dann gehört Y zu einem der drei folgenden Typen:*

A) *$Y_A \cong \tilde{\mathbb{P}}^2(x_1, \ldots, x_{11})$ ist die Aufblasung von $\mathbb{P}^2$ in 11 Punkten x_i. Die Einbettung ist durch $\left|6L - \sum_{i=1}^{6} 2x_i - \sum_{j=7}^{11} x_j\right|$ gegeben.*

B) *$Y_B = \tilde{Y}_0(y_0)$ ist die Aufblasung einer $K3$ Fläche $Y_0 \subset \mathbb{P}^5$ vom Grad 8 in einem Punkt. Die Einbettung ist gegeben durch Projektion von $y_0 \in Y_0$.*

C) *Y_C ist eine elliptische Fläche über $\mathbb{P}^1$, ohne exzeptionelle Kurven mit Kodaira Dimension $\kappa = 1$.*

Wir fassen die numerischen Invarianten der drei Familien in einer Tabelle zusammen:

Flächentyp	d	π	p_g	p_a	q	K^2	$K \cdot H$	χ	Klassifikation
Y_A	7	4	0	0	0	-2	-1	14	rational
Y_B	7	5	1	1	0	-1	1	25	$K3$
Y_C	7	6	2	2	0	0	3	36	$\kappa = 1$

Die Idealgarben dieser Typen haben folgende lokalfreie Auflösungen:

A) $$0 \to \mathcal{O}^{\oplus 4} \longrightarrow \mathcal{O}(1) \oplus \Omega^1_{\mathbb{P}^4}(2) \longrightarrow \mathcal{J}_{Y_A}(4) \to 0$$

B) $$0 \to \mathcal{O} \oplus \mathcal{O}(1) \longrightarrow \mathcal{O}(2)^{\oplus 3} \longrightarrow \mathcal{J}_{Y_B}(5) \to 0$$

C) $$0 \to \mathcal{O}^{\oplus 2} \longrightarrow \mathcal{O}(1)^{\oplus 2} \oplus \mathcal{O}(3) \longrightarrow \mathcal{J}_{Y_C}(5) \to 0.$$

Die Auflösung C) erhält man aus [10], Theorem 3.1.

3. Beispiele für 3-Mannigfaltigkeiten im $\mathbb{P}^5$

In diesem Abschnitt zeigen wir, daß die 3 Familien von glatten Flächen vom Grad 7 in $\mathbb{P}^4$, die wir gefunden haben, Hyperebenenschnitte von 3-Mannigfaltigkeiten im $\mathbb{P}^5$ sind. Anschließend konstruieren wir zu diesen 3-Mannigfaltigkeiten stabile und semistabile, reflexive Garben vom Rang 2 und bestimmen deren numerische Invarianten. Wir verwenden die Bezeichnungen aus [11].

Proposition 7. *Es gibt eine 3-Mannigfaltigkeit $X_A \subset \mathbb{P}^5$ vom Grad 7, deren generischer Hyperebenenschnitt $X_A \cap \mathbb{P}^4$ eine Fläche vom Typ A ist. X_A hat folgende Invarianten:*

$$H^2 K = -8, \qquad HK^2 = 7, \qquad K^3 = -2.$$

X_A ist das projektive Bündel $\mathbb{P}(E)$ eines Vektorbündels E vom Rang 2 auf einer Aufblasung von $\mathbb{P}^2$ in 6 Punkten.

Beweis. Wir gehen von der Auflösung

A) $$0 \to \mathcal{O}^{\oplus 4} \to \mathcal{O}(1) \oplus \Omega^1_{\mathbb{P}^4}(2) \to \mathcal{J}_{Y_A}(4) \to 0$$

einer Fläche Y_A vom Typ A im $\mathbb{P}^4$ aus. Nun gilt

$$\Omega^1_{\mathbb{P}^5}(2)|_{\mathbb{P}^4} \cong \mathcal{O}(1) \oplus \Omega^1_{\mathbb{P}^4}(2).$$

Wählt man also in dem global erzeugten 5-Bündel $\Omega^1_{\mathbb{P}^5}(2)$ auf $\mathbb{P}^5$ 4 generische Schnitte, so erhält man als Abhängigkeitsort eine 3-Mannigfaltigkeit X_A mit $X_A \cap H \cong Y_A$ für generische Hyperebenen $H \subset \mathbb{P}^5$. X_A hat die Auflösung

(4) $$0 \to \mathcal{O}^{\oplus 4} \to \Omega^1_{\mathbb{P}^5}(2) \to \mathcal{J}_{X_A}(4) \to 0.$$

Durch Dualisieren folgt

$$0 \to \mathcal{O}(-4) \to T_{\mathbb{P}^5}(-2) \to \mathcal{O}^{\oplus 4} \to \omega_X(2) \to 0.$$

Wir betrachten den Morphismus

$$\Phi = \Phi_{|K+2H|} \colon X \to \mathbb{P}^3$$

zu dem Linearsystem $|K+2H|$. Es gilt

$$(K+2H)^3 = c_5(\Omega^1_{\mathbb{P}^5}(2)) = 0,$$

das Bild $\bar{X} = \Phi(X) \subset \mathbb{P}^3$ ist also niederdimensional. Die Einschränkung Φ_Y von Φ auf einen generischen Hyperebenenschnitt Y stellt Y als Aufblasung einer Kubik $\bar{Y} \subset \mathbb{P}^3$ dar. Man hat folgende Situation:

$$\begin{array}{ccccc}
\mathbb{P}^5 \supset & X & \xrightarrow{\ \Phi\ } & \bar{X} & \subset \mathbb{P}^3 \\
\uparrow\!\!\cup & \uparrow\!\!\cup & & \uparrow\!\!\cup & \| \\
H \supset & Y & \xrightarrow{\ \Phi_Y\ } & \bar{Y} & \subset \mathbb{P}^3
\end{array}$$

Es folgt $\bar{X}=\bar{Y}$. Um den Grad einer Faser $X_{\bar{y}}$ über $\bar{y}\in\bar{Y}$ zu berechnen, wählen wir eine generische Gerade $L\subset\mathbb{P}^3$. L schneidet $\bar{Y}$ in 3 Punkten $\bar{y}_i$, es folgt

$$\sum_{i=1}^{3} \deg X_{\bar{y}_i} = H\cdot(K+2H)^2 = c_4(\Omega^1_{\mathbb{P}^5}(2)) = 3.$$

Daher ist $\Phi\colon X\to\bar{Y}$ ein $\mathbb{P}^1$-Bündel, $X=\mathbb{P}(E)$ mit

$$E\cong\Phi_*\mathcal{O}_X(1)$$

auf $\bar{Y}\cong\tilde{\mathbb{P}}^2(x_1,\dots,x_6)$.

Projiziert man die Sequenz

$$0\to\mathcal{O}_X\to\mathcal{O}_X(1)\to\mathcal{O}_Y(1)\to 0$$

mittels Φ auf $\bar{Y}$, so bekommt man

$$0\to\mathcal{O}_{\bar{Y}}\to E\to\mathcal{J}_Z\otimes\mathcal{O}_{\bar{Y}}\left(6L-\sum_{i=1}^{6}2E_i\right)\to 0.$$

Hier besteht $Z\subset\bar{Y}$ aus den 5 Punkten, durch deren Aufblasung Y entsteht.

Man kann zeigen, daß E nicht Pullback eines 2-Bündels auf $\mathbb{P}^2$ ist.

Aus der Auflösung (4) von $\mathcal{J}_X(4)$ entnimmt man sofort die behaupteten Invarianten.

Proposition 8. *Es gibt eine 3-Mannigfaltigkeit $X_B\subset\mathbb{P}^5$ vom Grad 7, deren generischer Hyperebenenschnitt $X_B\cap\mathbb{P}^4$ eine Fläche vom Typ B ist. X_B hat folgende Invarianten:*

$$H^2K=-6,\qquad HK^2=4,\qquad K^3=0.$$

X_B ist die Aufblasung eines vollständigen Durchschnittes von 3 Quadriken im $\mathbb{P}^6$.

Beweis. Wie oben kann man eine 3-Mannigfaltigkeit X_B mit Auflösung

$$0\to\mathcal{O}\oplus\mathcal{O}(1)\to\mathcal{O}(2)^{\oplus 3}\to\mathcal{J}_{X_B}(5)\to 0 \tag{5}$$

konstruieren. Daraus ergeben sich unmittelbar die Invarianten. Durch Dualisieren von (5) sieht man, daß $|K+2H|$ keine Basispunkte hat. Wir erhalten einen Morphismus

$$\Phi=\Phi_{|K+2H|}\colon X_B\to\mathbb{P}^6.$$

Es gilt $(K+2H)^3=8$, $\Phi(X_B)$ ist also 3-dimensional vom Grad 8 in $\mathbb{P}^6$. Wir wollen zeigen, daß $\Phi(X_B)$ glatt ist.

Es gilt $h^0(\mathcal{O}_{X_B}(K+H))=1$, wegen $H^2\cdot(K+H)=1$ ist ein effektiver Divisor in $|K+H|$ eine Ebene $E\subset X_B$. Die exakte Sequenz

$$0\to\mathcal{O}_{X_B}(H)\to\mathcal{O}_{X_B}(K+2H)\to\mathcal{O}_E(K+2H)\to 0$$

liefert $h^0(\mathcal{O}_E(K+2H))=1$, also

$$\mathcal{O}_E(K+2H)\cong\mathcal{O}_E.$$

Der Morphismus Φ kontrahiert E zu einem Punkt $e\in\mathbb{P}^3$ und ist außerhalb eine Einbettung.

Daher ist $\Phi(X_B) \subset \mathbb{P}^6$ glatt.

Analog zu [8] zeigt man, daß X_B die Aufblasung eines vollständigen Durchschnittes von 3 Quadriken in einem Punkt ist.

Projiziert man umgekehrt einen generischen vollständigen Durchschnitt von 3 Quadriken im $\mathbb{P}^6$ von einem Punkt, der nicht auf einer Trisekante liegt, so erhält man eine 3-Mannigfaltigkeit vom Typ B im $\mathbb{P}^5$.

Proposition 9. *Es gibt eine 3-Mannigfaltigkeit $X_C \subset \mathbb{P}^5$ vom Grad 7, deren generischer Hyperebenenschnitt $X_C \cap \mathbb{P}^4$ eine Fläche vom Typ C ist. X_C hat folgende Invarianten:*

$$H^2 K = -4, \qquad HK^2 = 1, \qquad K^3 = 2.$$

X_C ist gefasert über $\mathbb{P}^1$ mit generischer Faser

$$X_{C,t} \cong \tilde{\mathbb{P}}^2(x_1, \ldots, x_6).$$

Beweis. Man findet eine 3-Mannigfaltigkeit X_C als Abhängigkeitsort von 2 generischen Schnitten in $\mathcal{O}(1)^{\oplus 2} \oplus \mathcal{O}(3)$, bekommt also folgende Auflösung:

$$0 \to \mathcal{O}^{\oplus 2} \to \mathcal{O}(1)^{\oplus 2} \oplus \mathcal{O}(3) \to \mathcal{J}_{X_C}(5) \to 0. \tag{6}$$

Durch Dualisieren folgt, daß $|K+H|$ keine Basispunkte hat. Der zugehörige Morphismus

$$\Phi \colon X_C \to \mathbb{P}^1$$

hat Fasern vom Grad $H^2(K+H) = c_3(\mathcal{O}(1)^{\oplus 2} \oplus \mathcal{O}(3)) = 3$; generische Hyperebenenschnitte sind elliptische Kurven, es gilt also

$$X_{C,t} \cong \tilde{\mathbb{P}}^2(x_1, \ldots, x_6)$$

für die generische Faser.

Außer diesen 3 Familien von 3-Mannigfaltigkeiten gibt es keine weiteren 3-Mannigfaltigkeiten vom Grad 7 in $\mathbb{P}^5$.

Theorem 10. *Jede nicht entartete 3-Mannigfaltigkeit X vom Grad 7 im $\mathbb{P}^5$ gehört zu einer der drei Familien*

A) *$X_A = \mathbb{P}(E)$, E lokal frei vom Rang 2 auf $\tilde{\mathbb{P}}^2(x_1, \ldots, x_6)$.*

B) *$X_B = \tilde{S}_{(2,2,2)}(x_0)$ Aufblasung eines vollständigen Durchschnittes $S_{(2,2,2)} \subset \mathbb{P}^6$.*

C) *$X_C \to \mathbb{P}^1$ mit $\tilde{\mathbb{P}}^2(x_1, \ldots, x_6)$ als generischer Faser.*

Beweis. Zum Beweis genügt es, zu zeigen, daß eine 3-Mannigfaltigkeit $X \subset \mathbb{P}^5$ vom Grad 7 stets eine der folgenden 3 Typen von Auflösungen für ihre Idealgarbe hat:

$$\begin{array}{llcccc}
\text{A)} & 0 \to & \mathcal{O}^{\oplus 4} & \longrightarrow & \Omega^1_{\mathbb{P}^5}(2) & \longrightarrow \mathcal{J}_{X_A}(4) \to 0 \\
\text{B)} & 0 \to & \mathcal{O} \oplus \mathcal{O}(1) & \longrightarrow & \mathcal{O}(2)^{\oplus 3} & \longrightarrow \mathcal{J}_{X_B}(5) \to 0 \\
\text{C)} & 0 \to & \mathcal{O}^{\oplus 2} & \longrightarrow & \mathcal{O}(1)^{\oplus 2} \oplus \mathcal{O}(3) & \longrightarrow \mathcal{J}_{X_C}(5) \to 0.
\end{array}$$

Dies folgt in allen drei Fällen wie in [11] aus den entsprechenden Auflösungen für die möglichen Hyperebenenschnitte $Y = X \cap \mathbb{P}^4$.

In [9] wurde der Zusammenhang zwischen 2-codimensionalen Cohen Macauley Unterschemata des $\mathbb{P}^n$ und reflexiven Garben vom Rang 2 mit homologischer Dimension $\leqq 1$ erklärt. Für $n=5$ bedeutet dies speziell:

Ist $X \subset \mathbb{P}^5$ eine 3-Mannigfaltigkeit, $\varepsilon \in H^0(\omega_X(6-c_1))$ ein nicht trivialer Schnitt, so wird durch (X, ε) eine Extension

$$\varepsilon^*: 0 \to \mathcal{O} \xrightarrow{s} F \longrightarrow \mathcal{J}_X(c_1) \to 0$$

definiert, die X als Nullstellenschema $(s)_0$ eines Schnittes $s \in H^0(F)$ in einer reflexiven Garbe F vom Rang 2 mit $hd(F) \leqq 1$ darstellt. Die Chernklassen von F berechnen sich folgendermaßen [11]:

$$\begin{aligned}
&c_1(F) = c_1\\
&c_2(F) = H^3\\
&c_3(F) = H^2(K + (6-c_1)H)\\
&c_4(F) = H(K + (6-c_1)H)^2\\
&c_5(F) = (K + (6-c_1)H)^3.
\end{aligned}$$

Um diese Korrespondenz anwenden zu können, muß man nur feststellen, für welche $c_1 \in \mathbb{Z}$ man $h^0(\omega_X(6-c_1)) > 0$ hat.

Beispiel A. Sei $X_A \subset \mathbb{P}^5$ eine 3-Mannigfaltigkeit vom Typ A. Dualisiert man die Auflösung (4), so erhält man

$$0 \to \mathcal{O}(-4) \to T(-2) \to \mathcal{O}^{\oplus 4} \to \omega_{X_A}(2) \to 0.$$

Also ist $h^0(\omega_{X_A}(6-c_1))$ genau dann ungleich Null, wenn $c_1 \leqq 4$ ist. Wir geben die Chernklassen der zugehörigen normierten, stabilen und semistabilen Garben in der nachstehenden Tabelle an.

c_1	F	$c_i(F_{\text{norm}})$
0	semistabil	(0, 7, 34, 163, 772)
1	stabil	(−1, 7, 27, 129, 509)
2	stabil	(0, 6, 20, 75, 276)
3	stabil	(−1, 5, 13, 48, 174)
4	stabil	(0, 3, 6, 15, 36)

Beispiel B. Für eine 3-Mannigfaltigkeit $X_B \subset \mathbb{P}^5$ vom Typ B ist $h^0(\omega_{X_B}(6-c_1))$ für $c_1 \leqq 5$ ungleich Null. Die Chernklassen der zugehörigen normierten, stabilen und semistabilen Garben sind:

c_1	F	$c_i(F_{\text{norm}})$
0	semistabil	(0, 7, 36, 184, 936)
1	stabil	(−1, 7, 29, 148, 752)
2	stabil	(0, 6, 22, 90, 366)
3	stabil	(−1, 5, 15, 61, 247)
4	stabil	(0, 3, 8, 24, 72)
5	stabil	(−1, 1, 1, 2, 4)

Beispiel C. Für 3-Mannigfaltigkeiten $X_C \subset \mathbb{P}^5$ vom Typ C ist $h^0(\omega_{X_C}(6-c_1)) \neq 0$ für $c_1 \leqq 5$. Man bekommt folgende Garben:

c_1	F	$c_i(F_{\text{norm}})$
0	semistabil	(0, 7, 38, 205, 1100)
1	stabil	(−1, 7, 31, 167, 895)
2	stabil	(0, 6, 24, 105, 456)
3	stabil	(−1, 5, 17, 74, 320)
4	semistabil	(0, 3, 10, 33, 108)
5	instabil	(−1, 1, 3, 9, 27)

Literatur

1. Beauville, A.: Surfaces algébriques somplexes. Astérisque 54. Paris: Soc. Math. de France, France 1978
2. Chang, M.: Stable rank-2 reflexive sheaves on $\mathbb{P}^3$ with small c_2 and applications. Trans. A.M.S. (to appear)
3. Ein, L.: Nondegenerate surfaces of degree $n+3$ in $\mathbb{P}^n_c$ (preprint)
4. Hartshorne, R.: Varieties of small codimension in projective space. Bull. Amer. Math. Soc. **80**, 1017-1032 (1974)
5. Hartshorne, R.: Algebraic Geometry Graduate Texts in Math. **52**. New York: Springer 1977
6. Ionescu, P.: An enumeration of all smooth projective varieties of degree 5 and 6. INCREST Preprint Series Math. No. 74 (1981)
7. Kleiman, S.L.: Geometry on grassmannians and applications to splitting bundles and smoothing cycles. Inst. Hautes Études Sci. Publ. Math. **36**, 282-298 (1969)
8. Lanteri, A.: Sulle sperfici di grado sette. Instituto Lombardo (Rend. Sc.) A **115** (1981)
9. Okonek, C.: Reflexive Garben auf $\mathbb{P}^4$. Math. Ann. **269**, 211-237 (1982)
10. Okonek, C.: Moduli reflexiver Garben und Flächen von kleinem Grad in $\mathbb{P}^4$. Math. Z. **184**, 459-572 (1983)
11. Okonek, C.: 3-Mannigfaltigkeiten im $\mathbb{P}^5$ und ihre zugehörigen stabilen Garben. manuscripta math. **38**, 175-199 (1982)
12. Severi, F.: Intorno ai punti doppi impropi di una superficie generale dello spazio a quatro dimensioni, e a suoi punti tripli apparenti. Rend. Circ. Math. Palermo **15**, 33-51 (1901)

Eingegangen am 15. Februar 1984; in endgültiger Form am 4. April 1984

Math. Z. 187, 221-225 (1984)

Mathematische Zeitschrift

Zur Realisierbarkeit endlicher Gruppen als Automorphismengruppen algebraischer Funktionenkörper

Henning Stichtenoth

Fachbereich Mathematik der Universität Gesamthochschule Essen, Universitätsstr. 3, D-4300 Essen 1, Bundesrepublik Deutschland

§1. Kürzlich bewiesen D.J. Madden und R.C. Valentini [5], daß jede endliche Gruppe G als volle Automorphismengruppe eines geeigneten Funktionenkörpers E über einem vorgegebenen algebraisch abgeschlossenen Körper K auftritt, d.h. $G \simeq \mathrm{Aut}(E/K)$. Über den Fixkörper von $\mathrm{Aut}(E/K)$ können sie dabei jedoch nur wenig aussagen; dessen Geschlecht ist bei der Konstruktion in [5] mindestens gleich 13, meist aber erheblich größer.

Eine genaue Analyse des Beweises von Madden und Valentini führt – in Verbindung mit einigen Zusatzüberlegungen – zu viel schärferen Ergebnissen. Man kann nämlich den Fixkörper der Automorphismengruppe – unter sehr schwachen Voraussetzungen an G – weitgehend beliebig vorschreiben. Dies wird in der vorliegenden Arbeit durchgeführt.

Zunächst einige für die gesamte Arbeit gültige Bezeichnungen: K ist ein *algebraisch abgeschlossener* Körper, seine Charakteristik $\mathrm{char}(K)$ ist beliebig. Ein *Funktionenkörper* F/K ist ein algebraischer Funktionenkörper einer Variablen über K, d.h. eine endlich erzeugte Körpererweiterung vom Transzendenzgrad 1. Das *Geschlecht* von F/K wird mit g_F bezeichnet. Für eine Körpererweiterung L/M sei $[L:M]$ der Grad der Erweiterung und $\mathrm{Aut}(L/M)$ die Gruppe aller Automorphismen von L, die auf M trivial operieren. Ist L/M endlich galoissch, so bezeichne ich die Automorphismengruppe von L/M auch als $\mathrm{Gal}(L/M)$.

Das Hauptergebnis dieser Arbeit lautet:

Satz 1. *Sei $E_0/K(x)$ eine endliche separable Erweiterung des rationalen Funktionenkörpers $K(x)$ mit $[E_0: K(x)]>1$. Dann gibt es zu jedem Funktionenkörper F/K vom Geschlecht $g_F \geqq 2$ eine separable Erweiterung E/F mit $[E:F] = [E_0:K(x)]$ und*

$$\mathrm{Aut}(E/K)=\mathrm{Aut}(E/F)\simeq \mathrm{Aut}(E_0/K(x)).$$

Die Betonung in Satz 1 liegt dabei auf der Aussage, daß alle Automorphismen von E/K auf F trivial operieren, d.h. F liegt im Fixkörper von $\mathrm{Aut}(E/K)$.

Von besonderem Interesse ist ein Spezialfall von Satz 1. Dazu wird folgende Sprechweise eingeführt: Eine endliche Gruppe G heißt *realisierbar* über dem

Funktionenkörper F/K, wenn eine Galoiserweiterung E/F existiert mit $\mathrm{Gal}(E/F) \simeq G$. Die Gruppe G heißt *exakt realisierbar* über F, wenn eine Galoiserweiterung E/F existiert mit $\mathrm{Aut}(E/F) = \mathrm{Gal}(E/F) \simeq G$.

Als Spezialfall von Satz 1 ergibt sich sofort:

Satz 2. *Sei $G \neq 1$ eine endliche Gruppe, welche über dem rationalen Funktionenkörper $K(x)$ realisierbar ist. Dann ist G über jedem Funktionenkörper F/K vom Geschlecht $g_F \geqq 2$ exakt realisierbar.*

Im Falle $K = \mathbb{C}$ wurde von Greenberg [4] mit Methoden aus der Theorie der Riemannschen Flächen gezeigt, daß jede Gruppe $G \neq 1$ über jedem Funktionenkörper $F/\mathbb{C}$ exakt realisierbar ist. Dies folgt aus Satz 2 bei $g_F \geqq 2$ für jeden Körper K mit $\mathrm{char}(K) = 0$, denn nach Douady [3] ist dann jede endliche Gruppe G über $K(x)$ realisierbar. Ob dies auch bei positiver Charakteristik der Fall ist, ist mir nicht bekannt. Man weiß aber z.B. folgendes: Falls die Charakteristik von K kein Teiler der Ordnung von G ist, kann man G über $K(x)$ realisieren (Popp [6], 13. Vorlesung). Auch jede auflösbare Gruppe ist bei beliebiger Charakteristik von K über $K(x)$ realisierbar (D'Mello und Madan [2]).

Die Gruppe $G = 1$ läßt sich natürlich nicht über jedem Funktionenkörper F/K vom Geschlecht $g_F \geqq 2$ exakt realisieren, denn $\mathrm{Aut}(F/K)$ kann nichttrivial sein. Man kann jedoch mit Satz 1 leicht zeigen:

Satz 3. *Zu jedem Funktionenkörper F/K mit $g_F \geqq 2$ und jeder natürlichen Zahl $k \geqq 3$ gibt es eine Erweiterung E/F mit $[E:F] = k$ und* $\mathrm{Aut}(E/K) = 1$.

Schließlich bekommt man auch das Ergebnis von Madden und Valentini [5] als einfache Folgerung aus Satz 1, denn es gilt:

Satz 4. *Sei G eine endliche Gruppe. Dann gibt es eine endliche separable Erweiterung $E_0/K(x)$ mit* $\mathrm{Aut}(E_0/K(x)) \simeq G$.

Das entscheidende Hilfsmittel zum Beweis von Satz 1 ist eine Ungleichung von Castelnuovo über das Geschlecht eines Funktionenkörpers, der als Kompositum zweier Teilkörper dargestellt ist (s. Stichtenoth [8]). Die Bedeutung dieser Ungleichung wurde zuerst von D'Mello und Madan [2] erkannt; sie spielt auch bei Madden und Valentini [5] eine zentrale Rolle.

§ 2. Zum Beweis von Satz 1 werden 2 Lemmata benötigt:

Lemma 1. *Sei F/K ein Funktionenkörper und E/F eine endliche Erweiterung. Für jeden Körper M mit $F \subsetneqq M \subseteq E$ gelte die Ungleichung*

$$g_M > 2[M:F]g_F + ([M:F]-1)^2.$$

Dann folgt $\sigma(F) = F$ für jeden Automorphismus $\sigma \in \mathrm{Aut}(E/K)$.

Beweis. Angenommen, es sei $\sigma(F) \neq F$. Dann ist $\sigma(F) \cdot F = M \supsetneqq F$, und es gilt $g_F = g_{\sigma(F)}$ sowie $[M:F] = [M:\sigma(F)] > 1$. Nach Voraussetzung folgt

$$g_M > 2[M:F]g_F + ([M:F]-1)^2.$$

Andererseits besagt die Ungleichung von Castelnuovo [8] gerade

$$g_M \leqq 2[M:F]g_F + ([M:F]-1)^2.$$

Dieser Widerspruch beweist das Lemma.

Lemma 2. *Sei $E_0/K(x)$ eine endliche separable Erweiterung des rationalen Funktionenkörpers $K(x)$ und $C>0$ eine reelle Zahl. Dann gibt es eine separable Erweiterung $E_1/K(x)$ mit folgenden Eigenschaften:*

a) $[E_1:K(x)]=[E_0:K(x)]$

b) $\mathrm{Aut}(E_1/K(x)) \simeq \mathrm{Aut}(E_0/K(x))$

c) *Für jeden Körper M_1 mit $K(x) \subsetneqq M_1 \subseteq E_1$ ist $g_{M_1} \geqq C$.*

Beweis. Sei $\Omega \supseteq E_0$ ein algebraisch abgeschlossener Körper und $\tilde{E}_0 \subseteq \Omega$ die galoissche Hülle von $E_0/K(x)$. In der Galoiserweiterung $\tilde{E}_0/K(x)$ verzweigen nur endlich viele Stellen von $K(x)$; man kann daher ohne weiteres annehmen, daß Nullstelle und Pol von x in $\tilde{E}_0/K(x)$ unverzweigt sind. Ich wähle $t \in \Omega$ mit $t^m = x$, wobei $m \geqq C+1$ eine natürliche Zahl ist, welche nicht durch $\mathrm{char}(K)$ teilbar ist. In der zyklischen Erweiterung $K(t)/K(x)$ sind Nullstelle und Pol von x voll verzweigt, daher sind $\tilde{E}_0$ und $K(t)$ linear disjunkt über $K(x)$. Ich setze $\tilde{E}_1 := \tilde{E}_0 \cdot K(t)$; dann ist $\tilde{E}_1/K(t)$ galoissch und $\mathrm{Gal}(\tilde{E}_1/K(t)) \simeq \mathrm{Gal}(\tilde{E}_0/K(x))$. Für den Körper $E_1 := E_0 \cdot K(t)$ gilt nach Galoistheorie: $E_1/K(t)$ ist separabel, $[E_1:K(t)]=[E_0:K(x)]$ und $\mathrm{Aut}(E_1/K(t)) \simeq \mathrm{Aut}(E_0/K(x))$. Es bleibt für einen Körper M_1 mit $K(t) \subsetneqq M_1 \subseteq E_1$ das Geschlecht g_{M_1} abzuschätzen. Dazu setze man $M_0 := M_1 \cap \tilde{E}_0$. Die Erweiterung M_1/M_0 ist zyklisch vom Grad m, und es gilt $[M_0:K(x)]=[M_1:K(t)] \geqq 2$. In $M_0/K(x)$ sind Nullstelle und Pol von x voll zerlegt, also verzweigen in M_1/M_0 mindestens $2[M_0:K(x)] \geqq 4$ Stellen voll. Die Riemann-Hurwitzsche Geschlechtsformel für M_1/M_0 liefert nun

$$2g_{M_1} - 2 \geqq m(2g_{M_0} - 2) + 4(m-1) \geqq -2m + 4(m-1),$$

also $g_{M_1} \geqq m-1 \geqq C$ nach Wahl von m. Die Erweiterung $E_1/K(t)$ besitzt somit alle in Lemma 2 behaupteten Eigenschaften.

Beweis von Satz 1. Vorgegeben ist ein Funktionenkörper F/K vom Geschlecht $g_F \geqq 2$. Dann ist die Gruppe $H := \mathrm{Aut}(F/K)$ endlich (H.L. Schmid [7]). Sei n die Ordnung von H und $T \subseteq F$ der Fixkörper von H. Die Erweiterung F/T ist galoissch und $H = \mathrm{Gal}(F/T)$. Ich wähle eine Primzahl $q \neq \mathrm{char}(K)$ mit $q \geqq 2g_F$ und $q \nmid n$. Sei $\mathfrak{P}$ eine Stelle von F/K. Nach dem Satz von Riemann-Roch existiert ein Element $z \in F$ mit dem genauen Poldivisor $q\mathfrak{P}$. Für den rationalen Funktionenkörper $K(z) \subseteq F$ gilt also

i) $F/K(z)$ ist separabel vom Grad q, und in $F/K(z)$ ist die Stelle $\mathfrak{p} := \mathfrak{P}|_{K(z)}$ voll verzweigt.

Wegen $q \nmid n$ ist F das Kompositum von T und $K(z)$. Es gibt deshalb eine Stelle $\mathfrak{q}$ von T/K mit folgenden Eigenschaften (s. [8], Satz 3):

ii) $\mathfrak{q}$ ist in F/T unverzweigt, d.h. $\mathfrak{q}$ besitzt in F genau n paarweise verschiedene Fortsetzungen $\mathfrak{P}_1, \ldots, \mathfrak{P}_n$.

iii) Die Einschränkungen $\mathfrak{p}_i := \mathfrak{P}_i|_{K(z)}$ von $\mathfrak{P}_i$ auf $K(z)$ sind paarweise verschieden.

iv) $\mathfrak{p}_1$ ist in $F/K(z)$ unverzweigt.

Nach Lemma 2 findet man eine separable Erweiterung $E_1/K(z)$, so daß gilt:

v) $\mathfrak{p}_1$ ist in $E_1/K(z)$ verzweigt, aber $\mathfrak{p}_2, \ldots, \mathfrak{p}_n$ sind in $E_1/K(z)$ unverzweigt. (Man beachte hierzu, daß in einer echten Erweiterung eines rationalen Funktionenkörpers mindestens eine Stelle verzweigt.)

vi) Die Stelle $\mathfrak{p}$ ist in $E_1/K(z)$ unverzweigt.

vii) $[E_1:K(z)]=[E_0:K(x)]$ und $\mathrm{Aut}(E_1/K(z)) \simeq \mathrm{Aut}(E_0/K(x))$.

viii) Für jeden Körper M_1 mit $K(z) \subsetneqq M_1 \subseteq E_1$ ist

$$g_{M_1} > 2g_F \cdot [E_0:K(x)] + ([E_0:K(x)]-1)^2.$$

Wie im Beweis von Lemma 2 betrachte ich die galoissche Hülle $\tilde{E}_1$ von $E_1/K(z)$. Wegen vi) ist $\mathfrak{p}$ auch in $\tilde{E}_1/K(z)$ unverzweigt, und wegen i) sind F und $\tilde{E}_1$ linear disjunkt über $K(z)$. Die Erweiterung $\tilde{E}/F$ mit $\tilde{E} := \tilde{E}_1 \cdot F$ ist also galoissch mit $\mathrm{Gal}(\tilde{E}/F) \simeq \mathrm{Gal}(\tilde{E}_1/K(z))$, und es folgt für den Körper $E := E_1 \cdot F$:

$$E/F \text{ ist separabel, } [E:F]=[E_1:K(z)]=[E_0:K(x)]$$

und

$$\mathrm{Aut}(E/F) \simeq \mathrm{Aut}(E_1/K(z)) \simeq \mathrm{Aut}(E_0/K(x).$$

Satz 1 ist vollständig bewiesen, wenn noch gezeigt wird, daß $\mathrm{Aut}(E/K) = \mathrm{Aut}(E/F)$ gilt. Sei also $\sigma \in \mathrm{Aut}(E/K)$. Für einen Körper M mit $F \subsetneqq M \subseteq E$ setze ich $M_1 := M \cap E_1$ und erhalte wegen $[M:F]=[M_1:K(z)]$ und viii) die Ungleichung

$$\begin{aligned} g_M \geqq g_{M_1} &> 2[E_0:K(x)]g_F + ([E_0:K(x)]-1)^2 \\ &\geqq 2[M:F]g_F + ([M:F]-1)^2. \end{aligned}$$

Nach Lemma 1 folgt $\sigma(F)=F$, d.h. die Einschränkung σ_0 von σ auf F ist ein Automorphismus von F/K, also $\sigma_0 \in H = \mathrm{Gal}(F/T)$.

Angenommen, $\sigma_0 \neq \mathrm{id}_F$. Dann ist $\sigma_0(\mathfrak{P}_1)=\mathfrak{P}_i$ für ein $i \in \{2, \ldots, n\}$ wegen ii). Weil $\mathfrak{P}_1$ in E/F verzweigt (nach iv) und v)), muß auch $\mathfrak{P}_i$ in E/F verzweigen. Das ist aber nicht der Fall, denn in $E_1/K(z)$ ist $\mathfrak{p}_i = \mathfrak{P}_i|_{K(z)}$ nach v) unverzweigt. Daher muß $\sigma_0 = \mathrm{id}_F$ sein, und der Beweis von Satz 1 ist beendet.

Beweis von Satz 3. Wegen Satz 1 braucht man nur eine separable Erweiterung $E_0/K(x)$ vom Grad k anzugeben mit $\mathrm{Aut}(E_0/K(x))=1$. Dazu setze man

$$E_0 := K(y) \quad \text{mit} \quad y^{k-1}(y-1)=x.$$

Diese Gleichung ist irreduzibel und separabel, also ist $E_0/K(x)$ separabel vom Grad k. Die Nullstelle von x in $K(x)$ hat in $K(y)$ genau 2 Fortsetzungen, von denen eine mit der Verzweigungsordnung $k-1$ verzweigt. Daraus folgt, daß es keinen nichttrivialen Automorphismus von $K(y)/K(x)$ geben kann.

Beweis von Satz 4. Ist $q \neq \mathrm{char}(K)$ eine Primzahl, so läßt sich die symmetrische Gruppe S_q über dem rationalen Funktionenkörper $K(t)$ realisieren (s. z.B. Birch und Swinnerton-Dyer [1] oder Madden und Valentini [5], Lemma 3). Da jede endliche Gruppe G isomorph zu einer Untergruppe von S_q ist (für $q \geqq |G|$), folgt daraus: Es gibt eine Galoiserweiterung E_0/F_0 von algebraischen Funktionenkörpern über K mit $\mathrm{Gal}(E_0/F_0) \simeq G$. Sei nun $2 < l \neq \mathrm{char}(K)$ eine Primzahl mit $l \geqq 2g_{F_0}$. Ich wähle Stellen $\mathfrak{p}_1$, $\mathfrak{p}_2$ von F_0/K, die in E_0/F_0 nicht verzweigen.

Nach Riemann-Roch gibt es ein $x \in F_0$ mit dem genauen Poldivisor $\mathfrak{p}_1 + (l-1)\mathfrak{p}_2$. Die Erweiterung $E_0/K(x)$ ist dann offensichtlich separabel, und es gilt $\mathrm{Aut}(E_0/K(x)) = \mathrm{Aut}(E_0/F_0) \simeq G$.

§ 3. Es sollen noch einige abschließende Bemerkungen angefügt werden.

i) Ich vermute, daß die Voraussetzung $g_F \geqq 2$ in Satz 1 und 2 nicht erforderlich ist, kann das allerdings bisher nicht beweisen. In Satz 3 kann man auf die Voraussetzung $g_F \geqq 2$ verzichten - der Beweis soll hier jedoch nicht ausgeführt werden.

ii) Im Falle $g_F \leqq 1$ kann man natürlich zuerst eine geeignete Erweiterung F_1/F vornehmen und dann die Sätze 1 und 2 auf F_1 anwenden.

iii) Geht man von zwei nichtisomorphen Funktionenkörpern F bzw. $\tilde{F}$ mit $g_F = g_{\tilde{F}} = :g$ aus, so sind die in Satz 1 (bzw. in Satz 3) konstruierten Erweiterungskörper $E \supseteq F$ bzw. $\tilde{E} \supseteq \tilde{F}$ ebenfalls nicht isomorph; denn F (bzw. $\tilde{F}$) ist der einzige Teilkörper von E (bzw. $\tilde{E}$) vom Geschlecht g und Kograd $[E_0 : K(x)]$ (bzw. vom Kograd k in Satz 3). Dies folgt genau wie Lemma 1 aus der Ungleichung von Castelnuovo. Insbesondere hat man damit folgendes Resultat: Zu jeder endlichen Gruppe G gibt es unendlich viele paarweise nichtisomorphe Funktionenkörper E/K mit $\mathrm{Aut}(E/K) \simeq G$ (vgl. auch [5]).

iv) Im Beweis von Satz 4 werde benutzt, daß für eine Primzahl $q \neq \mathrm{char}(K)$ die symmetrische Gruppe S_q über $K(t)$ realisierbar ist. In der Tat ist für jedes $n \geqq 1$ die symmetrische Gruppe S_n über $K(t)$ realisierbar (bei beliebiger Charakteristik). Man erhält die S_n für $n \geqq 2$ etwa als Galoisgruppe der Gleichung

$$f(y) := y^2 \cdot \prod_{i=1}^{n-2} (y - a_i) - t(y - a_{n-1})^{n-1}(y - a_n)$$

über $K(t)$; dabei sind $a_1, \ldots, a_n \in K$ paarweise verschieden zu wählen. Zum Beweis überlegt man sich, daß die Gruppe von f primitiv auf den Wurzeln von f operiert und eine Transposition enthält; daraus folgt, daß die Gruppe isomorph zu S_n ist.

Literatur

1. Birch, B.J., Swinnerton-Dyer, H.P.F.: Note on a Problem of Chowla. Acta arith. **5**, 417–423 (1959)
2. D'Mello, J.G., Madan, M.L.: Algebraic Function Fields with Solvable Automorphism Group in Characteristic *p*. Communications in Algebra **11**, 1187–1236 (1983)
3. Douady, A.: Determination d'un Groupe de Galois, C.R. Acad. Sci. Paris **258**, 5305–5308 (1964)
4. Greenberg, L.: Maximal Groups and Signatures. Ann. of Math. Studies **79**, 207–226 (1974)
5. Madden, D.J., Valentini, R.C.: The Group of Automorphisms of Algebraic Function Fields. J. Reine Angew. Math. **343**, 162–168 (1983)
6. Popp, H.: Fundamentalgruppen algebraischer Mannigfaltigkeiten, Lecture Notes in Mathematics **176**. Berlin-Heidelberg-New York: Springer 1970
7. Schmid, H.L.: Über die Automorphismen eines algebraischen Funktionenkörpers von Primzahlcharakteristik. J. Reine Angew. Math. **179**, 5–15 (1938)
8. Stichtenoth, H.: Die Ungleichung von Castelnuovo. J. Reine Angew. Math. **348**, 197–202 (1984)

Eingegangen am 23. März 1984

Math. Z. 187, 227–257 (1984)

Mathematische Zeitschrift

Über Räume konformer Selbstabbildungen ebener Gebiete

Dieter Gaier

Mathematisches Institut der Universität, Arndtstraße 2,
D-6300 Gießen, Bundesrepublik Deutschland

Herrn Professor Helmut Grunsky zum 80. Geburtstag gewidmet

Einleitung

Die vorliegende Abhandlung entwickelte sich aus einer Fragestellung der komplexen Approximationstheorie, welche im Herbst 1981 von Paul Gauthier so skizziert wurde. Gegeben sei ein einfach zusammenhängendes Gebiet G, und f sei regulär in G. Um f durch Polynome, ganze Funktionen oder rationale Funktionen in G zu approximieren, kann man $f \circ \Psi$ statt f betrachten, wobei $\varphi = \Psi^{-1}$ eine konforme Abbildung (KA) von G auf sich [oder ein Teilgebiet von G] ist, und zunächst jene Funktion $f \circ \Psi$ zu approximieren suchen. Hernach ist dann die identische Abbildung id durch φ zu approximieren, und es stellt sich nunmehr die rein theoretische *Frage, ob es für jedes einfach zusammenhängende Gebiet G möglich ist, die identische Abbildung* id *durch konforme Selbstabbildungen* (SA) $\varphi \neq \mathrm{id}$ *beliebig gut zu approximieren.*

Zur Präzisierung unserer Aufgabe sei nun G einfach zusammenhängend und beschränkt, und $\Sigma(G)$ bezeichne die Menge aller konformen SA von G auf G. In $\Sigma(G)$ kann bekanntlich eine *Multiplikation* eingeführt werden:

$$(\varphi_1 \cdot \varphi_2)(z) := \varphi_1(\varphi_2(z)) \qquad (z \in G).$$

Daneben machen wir aber $\Sigma(G)$ noch zu einem *metrischen Raum:*

$$d(\varphi_1, \varphi_2) := \sup\{|\varphi_1(z) - \varphi_2(z)| : z \in G\}.$$

Metrische Räume von KA sind gelegentlich studiert worden (Hornich [12, 13]), doch scheint $\Sigma(G)$ bisher nicht untersucht.

Teil I bringt nun ganz losgelöst von der ursprünglichen Fragestellung eine Untersuchung der Struktur von $\Sigma(G)$ in verschiedener Hinsicht. So wird die Vollständigkeit und Kompaktheit von $\Sigma(G)$ untersucht, und der Frage nachgegangen, wann $\Sigma(G)$ eine topologische Gruppe ist, ferner werden Zusammenhangsfragen in $\Sigma(G)$ studiert. Manche feineren Probleme scheinen in die Primenden-Theorie hinein zu führen und müssen offen gelassen werden. Am Schluß der Arbeit geben wir eine Zusammenstellung ungelöster Fragen, die uns wichtig erscheinen.

In Teil II kommen wir auf die ursprüngliche Frage zurück: Für welche Gebiete G ist id Häufungspunkt in $\Sigma(G)$? Eine Charakterisierung solcher Gebiete durch einfache geometrische Eigenschaften wird nicht möglich sein; vielmehr zeigen wir an drei Klassen von Gebieten, was vorkommen kann. Beim allbekannten Kammgebiet ist id tatsächlich isoliert, wie Gauthier vermutete, während dies bei Schlangengebieten und anderen Kammgebieten nicht der Fall ist. Durch diese Beispiele wird eine Vermutung von Gauthier ([9], S. 565) widerlegt.

Für die Beweise werden Hilfsmittel und Sätze aus der Theorie der KA verwendet, die auch unabhängig von den hier behandelten Fragen interessant sein dürften.

Teil I: Untersuchung der Struktur der Räume $\Sigma(G)$

Im folgenden sei stets G ein beschränktes, einfach zusammenhängendes Gebiet, und $\Sigma(G)$ bezeichne die Menge aller konformen Abbildungen (KA) φ von G auf sich. Unser allgemeines Ziel ist das Studium des Raumes $\Sigma(G)$, wobei $\Sigma(G)$ mit algebraischer und/oder metrischer Struktur versehen ist.

§ 1. $\Sigma(G)$ als Gruppe

Erklärt man in $\Sigma(G)$ eine Multiplikation durch Komposition von Abbildungen, $\varphi_1 \cdot \varphi_2 := \varphi_1 \circ \varphi_2$, so wird $\Sigma(G)$ zu einer *Gruppe*. Im Sonderfall $G = \mathbb{D} = \{w: |w| < 1\}$ entsteht $\Sigma(\mathbb{D})$, die Gruppe der linearen Abbildungen L von $\mathbb{D}$ auf sich. In $\Sigma(\mathbb{D})$ zeichnen wir besonders aus:

$$L_\varepsilon\colon\ L_\varepsilon(w) = e^{i\varepsilon} w, \qquad \text{wo } \varepsilon \in \mathbb{R},$$

$$L_h\colon\ L_h(w) = \frac{w-h}{1-hw}, \qquad \text{wo } h \in (-1, +1).$$

Die Drehungen $\{L_\varepsilon\colon \varepsilon \in \mathbb{R}\}$ und die Abbildungen $\{L_h\colon h \in (-1, +1)\}$ bilden *Abelsche Untergruppen* von $\Sigma(G)$, und jedes $L \in \Sigma(\mathbb{D})$ läßt sich darstellen in der Form

$$L = L_{\varepsilon_1} \cdot L_h \cdot L_{\varepsilon_2} \tag{1.1}$$

für gewisse reelle $\varepsilon_1, \varepsilon_2$ und $h \in (-1, +1)$. Die Darstellung ist eindeutig, wenn L keine Drehung ist und $h > 0$ verlangt wird.

Zwischen $\Sigma(G)$ und $\Sigma(\mathbb{D})$ kann ein *Isomorphismus* hergestellt werden. Dazu sei f_0 eine im folgenden festgehaltene KA von $\mathbb{D}$ auf G und $g_0 = f_0^{-1}$. Jede SA $\varphi \in \Sigma(G)$ läßt sich dann in der Form

$$\varphi = f_0 L g_0 \tag{1.2}$$

schreiben für ein geeignetes $L \in \Sigma(\mathbb{D})$, und die Umkehrabbildung wird durch $L = g_0 \varphi f_0$ geliefert. Diese Abbildung zwischen $\Sigma(\mathbb{D})$ und $\Sigma(G)$ ist ein Isomorphismus; denn mit $\varphi_1 = f_0 L_1 g_0$, $\varphi_2 = f_0 L_2 g_0$ folgt

$$\varphi_1 \varphi_2 = (f_0 L_1 g_0)(f_0 L_2 g_0) = f_0(L_1 L_2) g_0.$$

Für später notieren wir die Beziehungen

$$g_0 \varphi = L g_0 \quad \text{und} \quad \varphi f_0 = f_0 L. \tag{1.3}$$

Die ausgezeichneten Elemente L_ε und L_h von $\Sigma(\mathbb{D})$ gehen bei Abbildung nach $\Sigma(G)$ über in

$$\varphi_\varepsilon := f_0 L_\varepsilon g_0 \quad \text{bzw.} \quad \varphi_h := f_0 L_h g_0;$$

diese bilden jeweils eine Abelsche Untergruppe von $\Sigma(G)$. Die SA φ_ε nennen wir (mit Rodin [22]) *konforme Drehungen* von G um den Fixpunkt $f_0(0)$ der Abbildung, während die SA φ_h (nach Jaenisch) als *konforme Translationen* bezeichnet werden können. Und aus (1.1) folgt, daß jedes $\varphi \in \Sigma(G)$ durch spezielle Elemente darstellbar ist:

$$\varphi = \varphi_{\varepsilon_1} \cdot \varphi_h \cdot \varphi_{\varepsilon_2};$$

dabei sind ε_1, ε_2 reell und $h \in (-1, +1)$.

§ 2. $\Sigma(G)$ als metrischer Raum; Vollständigkeit und Kompaktheit

Für den Moment vergessen wir die Gruppenstruktur von $\Sigma(G)$ wieder, führen jedoch eine *Metrik* in $\Sigma(G)$ ein:

$$d(\varphi_1, \varphi_2) = \sup\{|\varphi_1(z) - \varphi_2(z)| : z \in G\}. \tag{2.1}$$

Die Konvergenz $\varphi_n \to \varphi_0$ in $\Sigma(G)$ bedeutet also die gleichmäßige Konvergenz $\varphi_n(z) \rightrightarrows \varphi_0(z)$ in ganz G. Zunächst interessieren wir uns für die Vollständigkeit und Kompaktheit des Raumes $\Sigma(G)$ bei Verwendung dieser Metrik. (Über die Anfangsgründe konformer SA unterrichtet Behnke-Sommer [3], Kap. IV, § 4–6.)

Satz 1. a) *$\Sigma(G)$ ist nie kompakt.*
b) *$\Sigma(G)$ ist stets vollständig.*

Beweis. a) Bekanntlich (z.B. Franz [6], S. 116) ist ein metrischer Raum genau dann kompakt, wenn jede Punktfolge eine konvergente Teilfolge besitzt. Sei nun $z_0 \in G$ fest, $z_n \to \partial G$, und $\varphi_n \in \Sigma(G)$ so, daß $\varphi_n(z_0) = z_n$. Dann gibt es kein $\varphi \in \Sigma(G)$, für das $d(\varphi_n, \varphi) \to 0$ für $n \to \infty$ durch eine Teilfolge. Denn sonst wäre notwendig $|\varphi_n(z_0) - \varphi(z_0)| \to 0$; dies ist aber unmöglich, weil $\varphi_n(z_0) = z_n \to \partial G$.

b) Es sei $\{\varphi_n\}$ eine Cauchy-Folge in $\Sigma(G)$, also $\sup\limits_G |\varphi_n(z) - \varphi_k(z)| \to 0$ für $n, k \to \infty$. Dann hat man jedenfalls gleichmäßige Konvergenz der φ_n in G: Es gilt

$$\sup_G |\varphi_n(z) - \varphi(z)| \to 0 \qquad (n \to \infty)$$

für eine in G analytische Funktion φ. Wir zeigen $\varphi \in \Sigma(G)$.

Es ist entweder φ konstant $=a$ oder φ eine KA von G. Im ersten Fall wäre $|\varphi_n(z)-a|<\varepsilon$ $(n>N,\ z\in G)$, was offenbar nicht sein kann. Also ist φ eine KA von G und $\varphi(G)\subset\bar{G}$, da $\varphi_n(z)\in G$ für alle $z\in G$. Zu zeigen ist $\varphi(G)=G$.

Aber $\varphi(z)=\zeta\in\partial G$ für ein $z\in G$ ist nicht möglich, weil $\varphi_n(z)-\zeta\neq 0$ ist für $z\in G$ und alle n. Folglich ist jedenfalls $\varphi(G)\subset G$.

Angenommen, $a\in G$ wäre nicht in $\varphi(G)$. Wähle dann $\varepsilon>0$ so, daß $|z-a|\leqq\varepsilon$ in G liegt, und n so, daß

$$|\varphi_n(z)-\varphi(z)|<\varepsilon \qquad \text{für alle } z\in G; \tag{2.2}$$

dann wird n festgehalten. Ferner sei Γ eine Jordankurve in G so nahe am Rand von G, daß $\varphi_n(\Gamma)$ die ε-Scheibe um a umläuft; dieses Γ wird festgehalten. Jetzt betrachten wir $\varphi(\Gamma)$: Dies ist eine Jordankurve, die wegen (2.2) den Punkt a umläuft: $a\in\operatorname{int}\varphi(\Gamma)$. Somit haben wir: φ bildet Γ auf $\varphi(\Gamma)$ ab, und $\operatorname{int}\Gamma$ auf $\operatorname{int}\varphi(\Gamma)$. Also nimmt φ in G den Wert a an, und insgesamt ist $\varphi\in\Sigma(G)$, $\Sigma(G)$ ist vollständig.

Ob $\Sigma(G)$ stets *lokal kompakt* ist, können wir nicht sagen. *Jedoch ist* $\Sigma(\mathbb{D})$ *lokal kompakt.* Jedes $\varphi\in\Sigma(\mathbb{D})$ läßt sich nämlich eindeutig darstellen als $\varphi(z) = c\dfrac{z-z_0}{1-\bar{z}_0 z}$ mit $z_0\in\mathbb{D}$, $c\in\partial\mathbb{D}$. Die Abbildung

$$H:\quad \mathbb{D}\times\partial\mathbb{D}\to\Sigma(\mathbb{D}), \tag{2.3}$$

durch die Vorschrift

$$H(z_0,c)=\varphi \quad \text{mit} \quad \varphi(z)=c\frac{z-z_0}{1-\bar{z}_0 z}$$

erklärt, ist also bijektiv. Man zeigt leicht, daß sowohl H als auch H^{-1} stetige Abbildungen sind, sodaß H den Torus $\mathbb{D}\times\partial\mathbb{D}$ homöomorph auf $\Sigma(\mathbb{D})$ abbildet.

Mit $\Sigma(\mathbb{D})$ ist auch $\Sigma(G)$ lokal kompakt in all den Fällen, in denen $\Sigma(\mathbb{D})$ zu $\Sigma(G)$ homöomorph ist. Nehmen wir das Ergebnis von Satz 2 vorweg, so kommt der

Zusatz. *Der Raum* $\Sigma(G)$ *ist sicher dann lokal kompakt, wenn* G *nur punktförmige Primenden hat.*

Man beachte jedoch, daß $\Sigma(G)$ auch in anderen Fällen lokal kompakt sein kann, etwa wenn id isoliert liegt und damit $\Sigma(G)$ total unzusammenhängend ist.

§ 3. Untersuchung der Abbildung zwischen $\Sigma(\mathbb{D})$ und $\Sigma(G)$

Für ein beliebiges (beschränktes und einfach zusammenhängendes) Gebiet G sei der Raum $\Sigma(G)$ mit der Metrik $d(\varphi_1,\varphi_2)=\sup\{|\varphi_1(z)-\varphi_2(z)|:\ z\in G\}$ gebildet, andererseits der Raum $\Sigma(\mathbb{D})$ mit der Metrik $d(L_1,L_2)=\sup\{|L_1(w)-L_2(w)|:\ w\in\mathbb{D}\}$. Mit Hilfe einer KA f_0 von $\mathbb{D}$ auf G und ihrer Umkehrung $g_0=f_0^{-1}$ haben wir in § 1 eine Abbildung

$$\varphi=f_0 L g_0, \qquad L=g_0\varphi f_0 \tag{3.1}$$

zwischen den Elementen $L \in \Sigma(\mathbb{D})$ und denen $\varphi \in \Sigma(G)$ hergestellt. Wir fragen: Wann ist diese Abbildung stetig in der einen oder anderen Richtung?[1]

Satz 2. a) *Die Abbildung $\Sigma(G) \to \Sigma(\mathbb{D})$ ist immer stetig.*

b) *Die Abbildung $\Sigma(\mathbb{D}) \to \Sigma(G)$ ist stetig genau dann, wenn f_0 gleichmäßig stetig in $\mathbb{D}$ ist, d.h. wenn G nur punktförmige Primenden hat.*

Die Abbildung (3.1) stellt daher einen *Homöomorphismus* zwischen $\Sigma(\mathbb{D})$ und $\Sigma(G)$ her genau dann, wenn G nur punktförmige Primenden hat. - Zum Beweis von a) und auch später benötigen wir den einfachen

Hilfssatz 1. *Es seien L_n lineare* SA *von $\mathbb{D}$, mit $L_n(w) \Rightarrow w$ in kompakten Teilen von $\mathbb{D}$. Dann gilt $L_n(w) \Rightarrow w$ sogar in $\overline{\mathbb{D}}$.*

Beweis. Schreibt man $L_n(w) = e^{i\alpha_n}\dfrac{w - w_n}{1 - \bar{w}_n w}$, so folgt $w_n \to 0$ aus $L_n(0) \to 0$, und $L_n(x) \to x$ $(0 < x < 1)$ bringt $\alpha_n \to 0$. Dann aber ist $L_n(w) - w = o(1)$ $(n \to \infty)$, gleichmäßig in $\overline{\mathbb{D}}$.

Beweis zu a). Es seien $\varphi_n, \varphi_0 \in \Sigma(G)$ mit $\varphi_n \to \varphi_0$, und L_n, L_0 seien ihre Bilder gemäß (3.1). Dann gilt $\varphi_n(z) \Rightarrow \varphi_0(z)$ $(n \to \infty)$ in kompakten Teilen von G, folglich $g_0 \varphi_n(z) \Rightarrow g_0 \varphi_0(z)$, das heißt wegen (1.3) $L_n g_0(z) \Rightarrow L_0 g_0(z)$ oder $L_n(w) \Rightarrow L_0(w)$ $(n \to \infty)$ in kompakten Teilen von $\mathbb{D}$. Mit Hilfssatz 1 folgt $L_n(w) \Rightarrow L_0(w)$ sogar in $\overline{\mathbb{D}}$, das heißt $d(L_n, L_0) \to 0$ $(n \to \infty)$.

Beweis zu b). Zunächst sei f_0 in $\mathbb{D}$ gleichmäßig stetig, so daß

$$\omega_{f_0}(\delta) = \sup\{|f_0(w_1) - f_0(w_2)| : |w_1 - w_2| \leqq \delta;\ w_1, w_2 \in \mathbb{D}\} \to 0 \qquad (\delta \to 0).$$

Dann ist für zwei beliebige $L_1, L_2 \in \Sigma(\mathbb{D})$ und ihre Bilder $\varphi_1, \varphi_2 \in \Sigma(G)$

$$\begin{aligned} d(\varphi_1, \varphi_2) &= \sup_G |\varphi_1(z) - \varphi_2(z)| = \sup_{\mathbb{D}} |f_0 L_1(w) - f_0 L_2(w)| \\ &\leqq \omega_{f_0}(\sup_{\mathbb{D}} |L_1(w) - L_2(w)|) = \omega_{f_0}(d(L_1, L_2)). \end{aligned}$$

Daher ist die Abbildung von $\Sigma(\mathbb{D})$ nach $\Sigma(G)$ stetig.

Zum Beweis des Umkehrschlusses benötigen wir zwei Hilfssätze.

Hilfssatz 2 (Tamrazov [23], S. 726; [24], S. 166; [25], S. 110). *Es sei f in $\overline{\mathbb{D}}$ stetig, in $\mathbb{D}$ regulär, und ω_f sei der Stetigkeitsmodul von f auf $\partial\mathbb{D}$. Dann gilt*

$$|f(w_1) - f(w_2)| \leqq A\,\omega_f(|w_1 - w_2|) \qquad \text{für } w_1, w_2 \in \mathbb{D}$$

mit $A = 108$.

Hiermit beweisen wir

Hilfssatz 3. *Es sei f in $\mathbb{D}$ regulär, und es gelte für jedes $\varepsilon > 0$*

$$|f(e^{i\varphi} w) - f(w)| < \varepsilon \qquad \text{für } w \in \mathbb{D},\ \text{sobald } |\varphi| < \delta(\varepsilon).$$

Dann ist f in $\mathbb{D}$ gleichmäßig stetig.

---

[1] Nimmt man in $\Sigma(G)$ die Topologie der gleichmäßigen Konvergenz auf kompakten Teilen von G, so liefert (3.1) stets einen Homöomorphismus zwischen $\Sigma(G)$ und $\Sigma(\mathbb{D})$. Wesentlich ist jedoch hier, daß wir in $\Sigma(G)$ die Metrik (2.1) verwenden

Beweis. Ohne Einschränkung kann angenommen werden, daß $\delta(\varepsilon)$ für $\varepsilon>0$ positiv, monoton wachsend und stetig ist; $\varepsilon(\delta)$ mit $\varepsilon(\delta)\to 0$ $(\delta\to 0)$ sei die Umkehrfunktion. Dann ist

$$|f(e^{i\varphi}w)-f(w)|<\varepsilon(\delta) \qquad \text{für } w\in\mathbb{D}, \text{ sobald } |\varphi|<\delta.$$

Sind nun $w_1, w_2\in\mathbb{D}$ und $r\in(0,1)$ so, daß $|w_1|\leqq|w_2|\leqq r$ ist, so betrachten wir

$$F_r(w)=f(rw) \qquad \text{für } w\in\overline{\mathbb{D}}.$$

Falls $\omega_1, \omega_2\in\partial\mathbb{D}$ sind mit $|\omega_1-\omega_2|\leqq\delta$, so haben wir $|\arg\omega_1-\arg\omega_2|\leqq\frac{\pi}{2}\delta$ und daher

$$|F_r(\omega_1)-F_r(\omega_2)|=\left|f(r\omega_1)-f\left(r\omega_1\cdot\frac{\omega_2}{\omega_1}\right)\right|\leqq\varepsilon(\tfrac{\pi}{2}\delta)$$

nach Voraussetzung, so daß der Stetigkeitsmodul von F_r auf $\partial\mathbb{D}$ $\omega_{F_r}(\delta)\leqq\varepsilon(\frac{\pi}{2}\delta)$ ist. Nach Hilfssatz 2 gilt daher

$$|F_r(w_1)-F_r(w_2)|\leqq A\,\omega_{F_r}(|w_1-w_2|)\leqq A\,\varepsilon(\tfrac{\pi}{2}|w_1-w_2|).$$

Läßt man hierin $r\to 1$, so folgt

$$|f(w_1)-f(w_2)|\leqq A\,\varepsilon(\tfrac{\pi}{2}|w_1-w_2|),$$

gültig für alle $w_1, w_2\in\mathbb{D}$. Das ergibt unsere Behauptung.

Bemerkung. Die Verwendung von Hilfssatz 2 läßt sich vermeiden, wenn man weiß, daß f f.ü. radiale Grenzwerte besitzt, wie dies z.B. bei beschränktem oder schlichtem f der Fall ist.

Für Satz 2b) ist nun noch zu zeigen, daß f_0 sicher dann gleichmäßig stetig ist in $\mathbb{D}$, wenn die Abbildung (3.1) von $\Sigma(\mathbb{D})$ nach $\Sigma(G)$ stetig ist. Wir zeigen einiges mehr. Es bezeichne L_ε die Drehungen von $\mathbb{D}$ in der w-Ebene, $L_\varepsilon(w)=e^{i\varepsilon}w$, und

$$\varphi_\varepsilon := f_0 L_\varepsilon g_0$$

ihre Bilder in $\Sigma(G)$, d.h. die *konformen Drehungen* von G um $f_0(0)$. Man beachte, daß

$$d(L_\varepsilon, id)=\sup_{\mathbb{D}}|e^{i\varepsilon}w-w|=|e^{i\varepsilon}-1|\to 0 \qquad (\varepsilon\to 0).$$

Über das Geforderte hinaus zeigen wir sogar

Satz 3. *Es gilt* $\varphi_\varepsilon\to\mathrm{id}$ $(\varepsilon\to 0)$ *in* $\Sigma(G)$ *genau dann, wenn* f_0 *in* $\mathbb{D}$ *gleichmäßig stetig ist.*

Beweis. Die eine Richtung der Aussage folgt aus dem schon Bewiesenen:

$$d(\varphi_\varepsilon, \mathrm{id})\leqq\omega_{f_0}(d(L_\varepsilon, \mathrm{id}))=\omega_{f_0}(|e^{i\varepsilon}-1|).$$

Und ist andererseits

$$d(\varphi_\varepsilon, \mathrm{id})=\sup_G|f_0(e^{i\varepsilon}g_0(z))-z|=\sup_{\mathbb{D}}|f_0(e^{i\varepsilon}w)-f_0(w)|\to 0 \qquad (\varepsilon\to 0),$$

so zeigt Hilfssatz 3, daß f_0 in $\mathbb{D}$ notwendig gleichmäßig stetig sein muß.

Die Sätze 2 und 3 sind damit vollständig bewiesen.

Zusätzliche Bemerkungen. 1. Es stellt sich die Frage, ob auch die Stetigkeit der Abbildung $\Sigma(\mathbb{D}) \to \Sigma(G)$ längs *anderer* Abbildungen L hinreicht, um die gleichmäßige Stetigkeit von f_0 in $\mathbb{D}$ zu erschließen. Im allgemeinen ist dies nicht der Fall: Später geben wir Schlangengebiete (mit nicht punktförmigen Primenden) an, für die trotzdem $\varphi_h \to \mathrm{id}\,(h \to 0)$ gilt.

2. In Satz 3 war $\varphi_\varepsilon \to \mathrm{id}$ für *alle* $\varepsilon \to 0$ gefordert worden. Wir fragen: Reicht für die Gültigkeit von Satz 3 die schwächere Forderung, daß $\varphi_{\varepsilon_n} \to \mathrm{id}$ gilt für eine Nullfolge $\{\varepsilon_n\}$? Mit anderen Worten: Folgt aus $\varphi_{\varepsilon_n} \to \mathrm{id}$ für eine Nullfolge $\{\varepsilon_n\}$ immer $\varphi_\varepsilon \to \mathrm{id}$ $(\varepsilon \to 0)$? Das heißt konkret: Folgt aus $f_0(e^{i\varepsilon_n} w) \Rightarrow f_0(w)$ $(n \to \infty,\ w \in \mathbb{D})$ immer $f_0(e^{i\varepsilon} w) \Rightarrow f_0(w)$ $(\varepsilon \to 0,\ w \in \mathbb{D})$? (Letzteres ist wegen Hilfssatz 3 äquivalent damit, daß G nur punktförmige Primenden hat.)

Die gestellte Frage können wir nicht allgemein beantworten. Folgendes kann jedoch gezeigt werden.

a) Im Hilfssatz 6 von § 5 zeigen wir: Wenn die konformen Drehungen φ_{ε_n} von G gegen id konvergieren, so gilt für den Durchmesser $d(P_{\alpha_0})$ jedes Primendes P_{α_0} von G (welches dem Punkt $e^{i\alpha_0}$ entspricht) die Beziehung

$$d(P_{\alpha_0}) = \limsup_{\alpha \to \alpha_0^+} d(P_\alpha) = \limsup_{\alpha \to \alpha_0^-} d(P_\alpha). \tag{3.2}$$

G kann also zum Beispiel kein isoliertes, nicht punktförmiges Primende haben.

b) Die Funktion f_0 ist in $\mathbb{D}$ beschränkt und schlicht. Verzichtet man darauf, so ist unsere Frage zu verneinen. Dies zeigt folgendes

Beispiel. Es sei $\{a_k\}$ eine Nullfolge, für die Σa_k divergiert, und

$$f(w) = \sum_{k=0}^{\infty} a_k w^{2^k}$$

gesetzt. Dann ist f in $\mathbb{D}$ regulär. aber für $w \to 1-0$ nicht stetig (high indices theorem), also f sicher nicht gleichmäßig stetig in $\mathbb{D}$. Mit $\varepsilon_n = 2\pi/2^n$ $(n = 1, 2, \ldots)$ ist jedoch

$$\begin{aligned} f(w) - f(w\,e^{i\varepsilon_n}) &= \sum_{k=0}^{\infty} a_k [1 - e^{i\varepsilon_n \cdot 2^k}]\, w^{2^k} \\ &= \sum_{k<n} a_k [\quad]\, w^{2^k}, \end{aligned}$$

somit

$$|f(w) - f(w\,e^{i\varepsilon_n})| \leqq C \sum_{k<n} |a_k|\, 2^{k-n} \qquad \text{für } w \in \mathbb{D}.$$

Rechts wird eine Nullfolge mit einer positiven Dreiecksmatrix transformiert, deren Zeilensummen $\sum_{k<n} 2^{k-n} < 1$ sind, und in deren Spalten Nullfolgen stehen. Also gilt

$$f(w) - f(w\,e^{i\varepsilon_n}) \Rightarrow 0 \qquad (n \to \infty;\ w \in \mathbb{D}).$$

§ 4. $\Sigma(G)$ als topologische Gruppe

Da in $\Sigma(G)$ einerseits eine Gruppenstruktur vorliegt, andererseits in $\Sigma(G)$ eine Metrik erklärt ist, liegt die *Frage* nahe: Unter welchen Annahmen über G ist $\Sigma(G)$ eine topologische Gruppe (TG)? Dazu müssen in $\Sigma(G)$ folgende Regeln gelten:

$$\text{Aus } \varphi_n \to \varphi, \quad \psi_n \to \psi \quad \text{folgt } \varphi_n \psi_n \to \varphi\psi \quad (n \to \infty); \tag{4.1}$$

$$\text{aus } \varphi_n \to \varphi \quad \text{folgt } \varphi_n^{-1} \to \varphi^{-1} \quad (n \to \infty). \tag{4.2}$$

Diese beiden Bedingungen werden zunächst auf eine einfachere Forderung reduziert.

4.1. Vorbereitungen

Hilfssatz 4. *Die Rechtsmultiplikation mit einem Element aus $\Sigma(G)$ ist eine isometrische Abbildung von $\Sigma(G)$:*

$$d(\varphi_1 \varphi_3, \varphi_2 \varphi_3) = d(\varphi_1, \varphi_2) \quad \textit{für alle } \varphi_1, \varphi_2, \varphi_3 \textit{ aus } \Sigma(G).$$

Beweis. Es ist

$$d(\varphi_1 \varphi_3, \varphi_2 \varphi_3) = \sup_G |\varphi_1(\varphi_3(z)) - \varphi_2(\varphi_3(z))| = \sup_G |\varphi_1(z) - \varphi_2(z)| = d(\varphi_1, \varphi_2).$$

Setzt man $\varphi_2 = \mathrm{id}$, $\varphi_3 = \varphi_1^{-1}$, so erhält man die

Folgerung. *Es gilt*

$$d(\varphi, \mathrm{id}) = d(\varphi^{-1}, \mathrm{id}) \quad \textit{für alle } \varphi \textit{ aus } \Sigma(G).$$

Weniger harmlos als die Rechtsmultiplikation ist die Linksmultiplikation. Selbst im Fall $G = \mathbb{D}$ ist sie *keine isometrische Abbildung*. Wir wählen dazu $G = \mathbb{D}$, $\varphi_1(z) = a\,z(|a| = 1, a \neq 1)$, $\varphi_2 = \mathrm{id}$, während φ_3 eine SA von $\mathbb{D}$ sein soll mit $\varphi_3(1) = 1$, $\varphi_3(a) = -1$. Dann ist

$$d(\varphi_3 \varphi_1, \varphi_3 \varphi_2) = d(\varphi_3 \varphi_1, \varphi_3) \geqq |\varphi_3(\varphi_1(1)) - \varphi_3(1)| = |\varphi_3(a) - \varphi_3(1)| = 2,$$

während

$$d(\varphi_1, \varphi_2) = \max_{|z|=1} |a\,z - z| = |a - 1|$$

beliebig klein gemacht werden kann.

Schließlich bemerken wir noch, daß (4.1) im Sonderfall $\varphi = \mathrm{id}$ immer gilt:

$$\textit{Aus } \varphi_n \to \mathrm{id}, \ \psi_n \to \psi \quad \textit{folgt } \varphi_n \psi_n \to \psi \quad (n \to \infty). \tag{4.3}$$

Denn wenn für $\varepsilon > 0$ gilt

$$|\varphi_n(z) - z| < \varepsilon \quad (n > N, z \in G) \qquad |\psi_n(z) - \psi(z)| < \varepsilon \quad (n > N, z \in G),$$

so gilt für jedes solche n, $z=\psi_n(\zeta)$ $(\zeta\in G)$ gesetzt, auch

$$|\varphi_n(\psi_n(\zeta))-\psi_n(\zeta)|<\varepsilon \qquad (n>N,\ \zeta\in G),$$

zusammen also $d(\varphi_n\psi_n,\psi)<2\varepsilon$ $(n>N)$. – Insbesondere zeigt (4.3), daß die Multiplikation eine an der Stelle id stetige Operation ist.

Wir kommen nun zu dem angekündigten Kriterium dafür, daß $\Sigma(G)$ eine TG ist.

Satz 4. *Es ist $\Sigma(G)$ eine TG genau dann, wenn gilt:*

$$\textit{Aus } \varphi_n\to\mathrm{id} \textit{ folgt } \psi\varphi_n\to\psi \ (n\to\infty) \quad \textit{für alle } \psi\in\Sigma(G). \tag{4.4}$$

Beweis. Daß (4.4) notwendige Bedingung ist, ist klar. Nun sei (4.4) erfüllt, und $\varphi_n\to\varphi,\psi_n\to\psi$. Dann gilt $\varphi_n\varphi^{-1}\to\mathrm{id}$, $\psi_n\psi^{-1}\to\mathrm{id}$, also wegen (4.4) $\varphi\cdot\psi_n\psi^{-1}\to\varphi$, und wegen (4.3) $\varphi_n\varphi^{-1}\cdot\varphi\psi_n\psi^{-1}\to\varphi$ $(n\to\infty)$. Das heißt $\varphi_n\psi_n\psi^{-1}\to\varphi$, und da Rechtsmultiplikation eine Isometrie ist $\varphi_n\psi_n\to\varphi\psi$ $(n\to\infty)$; das ist (4.1).

Um (4.2) nachzuweisen, sei $\varphi_n\to\varphi$ gegeben, also $\varphi_n\varphi^{-1}\to\mathrm{id}$, folglich $\varphi\varphi_n^{-1}\to\mathrm{id}$ nach der Folgerung aus Hilfssatz 4, und da (4.4) gelten soll, darf mit φ^{-1} links multipliziert werden: $\varphi_n^{-1}\to\varphi^{-1}$ $(n\to\infty)$; das ist (4.2).

Ob Satz 4 richtig bleibt, wenn (4.4) nur für alle konformen Drehungen $\psi=\psi_\varepsilon\in\Sigma(G)$ gefordert wird, muß offen bleiben. – Aus Satz 4 erhalten wir sofort die

Folgerung. *Der Raum $\Sigma(\mathbb{D})$ ist eine TG.*

Denn hier ist

$$\varphi_n(z)=L_n(z)=e^{i\alpha_n}\frac{z-z_n}{1-\bar{z}_n z} \qquad \text{mit } \alpha_n\to 0,\ z_n\to 0$$

und

$$\psi(z)=L(z)=e^{i\alpha}\frac{z-z_0}{1-\bar{z}_0 z}.$$

Leichte Rechnung zeigt $LL_n(z)\Rightarrow L(z)$ $(n\to\infty, z\in\mathbb{D})$, also $\psi\varphi_n\to\psi$ $(n\to\infty)$.

4.2. *Wann ist $\Sigma(G)$ eine topologische Gruppe?*

Nun kommen wir zu unserer Ausgangsfrage zurück. In zwei wichtigen Fällen ist sofort beweisbar, daß $\Sigma(G)$ eine TG ist.

Satz 5. *Es ist $\Sigma(G)$ sicher dann eine TG, wenn*

a) *G nur punktförmige Primenden hat, oder wenn*
b) *das Element id in $\Sigma(G)$ isoliert liegt.*

Beweis. Im Fall a) ist ja $\Sigma(G)$ isomorph und homöomorph zu $\Sigma(\mathbb{D})$, und da $\Sigma(\mathbb{D})$ eine TG ist, gilt dasselbe für $\Sigma(G)$. Und im Fall b) ist das Kriterium (4.4) trivial erfüllt.

Ob $\Sigma(G)$ nur dann eine TG ist, wenn a) oder b) gilt, kann nicht entschieden werden. Wir geben aber nun ein Kriterium an, wann $\Sigma(G)$ sicher *keine* TG ist. Oder anders: Wir geben eine notwendige Bedingung an dafür, daß $\Sigma(G)$ eine TG ist. Darin bezeichne $d(P_\alpha)$ den *Durchmesser* der Projektion des Primendes (PE) P_α von G, welches unter der Abbildung f_0 dem Punkt $e^{i\alpha}$ entspricht.

Satz 6. *Das Element* id *sei in $\Sigma(G)$ nicht isoliert, und $\Sigma(G)$ sei eine TG. Dann gilt für die Primenden P_α von G notwendig*

$$d(P_{\alpha_0}) = \limsup_{\alpha\to\alpha_0+} d(P_\alpha) = \limsup_{\alpha\to\alpha_0-} d(P_\alpha), \qquad \textit{für alle } \alpha_0\in\mathbb{R}. \tag{4.5}$$

Daraus ziehen wir gleich eine Folgerung, die einen häufig auftretenden Fall abdeckt.

Folgerung. *Das Element* id *sei in $\Sigma(G)$ nicht isoliert, dagegen besitze G ein isoliertes, nicht punktförmiges Primende P_{α_0}. Dann ist $\Sigma(G)$ keine TG.*

Dabei nennen wir ein nicht punktförmiges PE P_{α_0} isoliert, wenn alle PE P_α mit $0<|\alpha-\alpha_0|<\varepsilon$ punktförmig sind. Offenbar ist dann (4.5) verletzt, $\Sigma(G)$ kann keine TG sein. Beispiele für solche Gebiete bringen wir in Teil II, § 3. Gebiete (mit nicht punktförmigen PE), die (4.5) erfüllen, lassen sich leicht angeben; Modifikation der bei Collingwood-Lohwater ([4], S. 186) angegebenen Konstruktion. Gebiete, für die $d(P_\alpha)=\text{const}>0$ ist, haben Denjoy und Piranian angegeben.

Hilfsmittel zum Beweis von Satz 6 ist

Hilfssatz 5. *Es gebe eine Folge linearer Abbildungen $L_n\in\Sigma(\mathbb{D})$ mit $L_n(1)\to 1$, aber $L_n(1)\neq 1$, für die*

$$f_0\,L_n(w)\Rightarrow f_0(w) \qquad (n\to\infty;\, w\in\mathbb{D}) \tag{4.6}$$

gelte. Dann gilt für das $w=1$ entsprechende Primende P_0 von G notwendig

$$d(P_0) = \limsup_{\alpha\to 0+} d(P_\alpha) = \limsup_{\alpha\to 0-} d(P_\alpha). \tag{4.7}$$

Insbesondere ist P_0 kein isoliertes, nicht punktförmiges Primende von G.

Beweis. Zunächst sei angenommen, die Zahlen $L_n(1)\in\partial\mathbb{D}$ rücken von *oben* her gegen $w=1$. Zu $\varepsilon>0$ sei N so bestimmt, daß

$$|f_0\,L_n(w)-f_0(w)|<\varepsilon \qquad (n>N;\, w\in\mathbb{D}).$$

Läßt man hierin $w\to 1$, so erkennt man, daß sich die cluster sets von f_0 an $w=1$ und an $w=L_n(1)$ höchstens um ε unterscheiden. Dies liefert sofort

$$d(P_0) \leqq \limsup_{\alpha\to 0+} d(P_\alpha),$$

und da andererseits $\geqq$ stets richtig ist, muß das Gleichheitszeichen stehen.

Aus (4.6) folgt aber auch

$$f_0\,L_n^{-1}(w)\Rightarrow f_0(w) \qquad (n\to\infty;\, w\in\mathbb{D}),$$

und jetzt rücken die Zahlen $L_n^{-1}(1) \in \partial \mathbb{D}$ von *unten* her gegen $w=1$. Schließt man analog wie oben, so folgt vollends die Behauptung (4.7).

Beweis von Satz 6. Die Aussage von Satz 6 ist unabhängig von der Wahl der Normalabbildung f_0. Wir denken f_0 so gewählt, daß $w=1$ dem ausgezeichneten PE P_{α_0} entspricht, also $P_{\alpha_0}=P_0$ ist. Da id nicht isoliert sein soll, gibt es SA $\varphi_n \to \mathrm{id}$ mit $\varphi_n \neq \mathrm{id}$; ihre Bilder in $\Sigma(\mathbb{D})$ seien L_n, wobei für sie gilt $L_n \to \mathrm{id}$ (wegen Satz 2a)) und $L_n \neq \mathrm{id}$.

Jedes dieser L_n hat höchstens zwei Fixpunkte. Folglich gibt es eine Menge E vom Maß 2π so, daß

$$L_n(e^{i\beta}) \neq e^{i\beta} \quad \text{für alle } n \text{ und } \beta \in E. \tag{4.8}$$

Ein solches β sei gewählt, $l(w)=e^{-i\beta}w$ gesetzt, und die Eigenschaft (4.4) ($\Sigma(G)$ ist TG) nur für ein einziges ψ verwendet, nämlich für $\psi=f_0 l g_0$; dies ist eine gewisse konforme Drehung von G. Es soll also gelten $\psi \varphi_n \to \psi$ in $\Sigma(G)$, das heißt $f_0 l g_0 f_0 L_n g_0 \to f_0 l g_0$ oder $f_0 l L_n \to f_0 l$ in $\Sigma(\mathbb{D})$, das heißt

$$f_0 M_n(w) \Rightarrow f_0(w) \qquad (n \to \infty;\, w \in \mathbb{D}),$$

wobei wir $M_n = l L_n l^{-1}$ gesetzt haben. Diese linearen $M_n \in \Sigma(\mathbb{D})$ genügen wegen $L_n \to \mathrm{id}$ und (4.8)

$$M_n \to \mathrm{id}, \text{ insbesondere } M_n(1) \to 1 \ (n \to \infty) \text{ und } M_n(1) \neq 1.$$

Nun wird Hilfssatz 5 anwendbar, und die Behauptung (4.5) folgt für $\alpha_0=0$.

§ 5. Zusammenhangsfragen in $\Sigma(G)$

Wie jeder topologische Raum zerfällt auch $\Sigma(G)$ in Zusammenhangskomponenten; $C(\varphi)$ bezeichne die φ enthaltende Komponente. In $\Sigma(G)$ stellt jedoch die Rechtsmultiplikation mit dem festen Element φ eine Isometrie dar, die id nach φ schafft, und man sieht leicht, daß dabei $C(\mathrm{id})$ in $C(\varphi)$ übergeht: *Alle Zusammenhangskomponenten von $\Sigma(G)$ sind zu* $C(\mathrm{id})$ *isometrisch.* Wie $C(\mathrm{id})$ allgemein aussieht, muß noch untersucht werden; Spezialfälle werden in Teil II, § 3 behandelt. Hier wird untersucht, wann $\Sigma(G)$ zerfällt oder sogar total unzusammenhängend ist.

5.1. *Einfache Fälle*

Der folgende Satz 7 behandelt Extremfälle.

Satz 7. a) *Hat G nur punktförmige Primenden, so ist $\Sigma(G)$ zusammenhängend.*
b) *Liegt* id *in $\Sigma(G)$ isoliert, so ist $\Sigma(G)$ total unzusammenhängend.*

Beweis. Für a) ziehen wir den Homöomorphismus H von (2.3) heran, der den Torus $\mathbb{D} \times \partial \mathbb{D}$ auf $\Sigma(\mathbb{D})$ abbildet, sowie den Homöomorphismus von $\Sigma(\mathbb{D})$ nach $\Sigma(G)$ (vgl. Satz 2). Die Aussage b) ist nach dem oben Gesagten klar. Die Frage, wann id in $\Sigma(G)$ isoliert liegt, ist das Thema von Teil II.

5.2. Unzusammenhängende Räume $\Sigma(G)$

Ob die Umkehrung der Aussage von Satz 7a) gilt, können wir nicht entscheiden. (Eventuell reicht schon der Zusammenhang der Menge der konformen Drehungen φ_ε von G hin dafür, daß G nur punktförmige Primenden besitzt.) Jedoch zeigt der unten folgende Satz 8, daß im Falle des Zusammenhangs von $\Sigma(G)$ entweder G nur punktförmige Primenden oder aber eine recht komplizierte Randstruktur hat.

Zunächst benötigen wir folgenden

Hilfssatz 6. *Das Gebiet G besitze ein Primende, etwa P_0, für welches die Beziehung* (4.7) *nicht gelte. Es ist also*

$$d(P_0)>\limsup_{\alpha\to 0+} d(P_\alpha) \quad \textit{oder} \quad d(P_0)>\limsup_{\alpha\to 0-} d(P_\alpha) \tag{5.1}$$

oder beides. Ferner seien $\varphi_n\in\Sigma(G)$ gegeben mit $\varphi_n\to \mathrm{id}\,(n\to\infty)$. Dann läßt φ_n das Primende P_0 fest, sobald n hinreichend groß ist.

Dabei sagen wir, daß die SA $\varphi=f_0 L g_0$ das PE P von G *fest läßt*, wenn die lineare Abbildung L den Fixpunkt $g_0(P)$ hat. Man beachte, daß (5.1) sicher dann zutrifft, wenn P_0 ein isoliertes nicht punktförmiges PE ist.

Beweis. Hilfssatz 6 ist im wesentlichen eine Umformung von Hilfssatz 5. Wir schreiben $\varphi_n=f_0 L_n g_0$, mit $L_n\to\mathrm{id}$, folglich $L_n(1)\to 1$ $(n\to\infty)$. Außerdem ist

$$\sup_G|\varphi_n(z)-z|=\sup_{\mathbb{D}}|f_0\,L_n(w)-f_0(w)|\to 0 \qquad (n\to\infty).$$

Wäre $L_n(1)\neq 1$ unendlich oft, so könnte Hilfssatz 5 auf eine Teilfolge der L_n angewendet werden, und (4.7) würde gelten. Also ist $L_n(1)=1$ von einer Stelle an, d.h. φ_n läßt P_0 fest von dieser Stelle an.

Um nun Gebiete zu finden, für die $\Sigma(G)$ nicht zusammenhängt, benötigen wir noch eine topologische Vorbemerkung. Die Zusammenhangskomponente $C(x)$ eines topologischen Raumes Σ ist bekanntlich die maximale zusammenhängende Teilmenge von Σ, die x enthält. In einem metrischen Raum kann daneben eine *Verkettungskomponente* $V(x)$ eingeführt werden (vgl. etwa Franz [6], S. 119). Zwei Punkte x,y heißen ε-verkettet, wenn endlich viele Punkte $x_0=x,x_1,x_2,\dots,x_n=y$ aus Σ existieren mit $d(x_{j-1},x_j)<\varepsilon$. Weiter heißt

$$V_\varepsilon(x)=\{y: y \text{ ist mit } x\ \varepsilon\text{-verkettet}\}$$

die ε-Komponente, die x enthält, und schließlich $V(x)=\bigcap\{V_\varepsilon(x):\varepsilon>0\}$ die *Verkettungskomponente*, die x enthält. Die Mengen $V_\varepsilon(x)$ sind stets offen-abgeschlossen, weshalb $C(x)\subset V(x)$ gilt.

Es seien nun $\varphi_1,\varphi_2\in\Sigma(G)$ ε-verkettet für ein $\varepsilon>0$, so daß es SA $\psi_0=\varphi_1,\psi_1,\dots,\psi_n=\varphi_2$ gibt mit $d(\psi_{j-1},\psi_j)<\varepsilon$. Schreibt man

$$\varphi_2=(\varphi_2\psi_{n-1}^{-1})(\psi_{n-1}\psi_{n-2}^{-1})\cdots(\psi_2\psi_1^{-1})(\psi_1\psi_0^{-1})\,\varphi_1,$$

so sieht man: *φ_1,φ_2 sind ε-verkettet genau dann, wenn $\alpha_1,\alpha_2,\dots,\alpha_n\in\Sigma(G)$ existieren mit*

$$\varphi_2=\alpha_1\alpha_2\dots\alpha_n\varphi_1, \quad \textit{wobei} \quad d(\alpha_j,\mathrm{id})<\varepsilon \qquad (j=1,2,\dots,n). \tag{5.2}$$

Nun sind wir in der Lage, eine notwendige Bedingung für G anzugeben, damit $\Sigma(G)$ zusammenhängend ist.

Satz 8. *Das Gebiet G besitze ein Primende, etwa P_0, für welches* (5.1) *gelte, und φ sei aus der Zusammenhangskomponente $C(\mathrm{id})$ von $\Sigma(G)$. Dann läßt φ das Primende P_0 fest.*

Die konformen Drehungen $\varphi_\varepsilon(\varepsilon \neq 0)$ von G liegen demnach nicht in $C(\mathrm{id})$, $\Sigma(G)$ ist nicht zusammenhängend. Daraus ergeben sich

Folgerung 1. *Ist $\Sigma(G)$ zusammenhängend, so gilt notwendig* (4.5):

$$d(P_{\alpha_0}) = \limsup_{\alpha \to \alpha_0 +} d(P_\alpha) = \limsup_{\alpha \to \alpha_0 -} d(P_\alpha) \qquad \text{für alle } \alpha_0 \in \mathbb{R}.$$

Folgerung 2. *Besitzt G ein isoliertes, nicht punktförmiges Primende, so ist $\Sigma(G)$ nicht zusammenhängend.*

Beweis von Satz 8. Ist $\varphi \in C(\mathrm{id})$, so ist φ mit id ε-verkettet für jedes $\varepsilon > 0$, also gilt wegen (5.2)

$$\varphi = \alpha_1 \alpha_2 \ldots \alpha_n \quad \text{mit } d(\alpha_j, \mathrm{id}) < \varepsilon \quad (j = 1, 2, \ldots, n).$$

Ist ε hinreichend klein gewählt, so läßt jede SA α_j das PE P_0 fest (Hilfssatz 6), so daß φ dieses PE fest läßt, und alles ist bewiesen.

Teil II: Wann liegt id isoliert in $\Sigma(G)$?

Unsere weiteren Untersuchungen des Raumes $\Sigma(G)$ beschäftigen sich vorwiegend mit der Frage, wann das Element id von $\Sigma(G)$ isoliert liegt. Dann liegen alle Elemente isoliert, $\Sigma(G)$ is total unzusammenhängend. Verschiedene Gebietsklassen werden daraufhin untersucht: Schlangengebiete haben nicht punktförmige PE, trotzdem liegt id nicht isoliert. Dasselbe gilt für Kammgebiete 2. Art; hingegen ist bei Kammgebieten 1. Art id isoliert, wie Gauthier vermutete.

§ 1. Zusammenstellung von Hilfsmitteln

Im folgenden werden einige Abschätzungen harmonischer Maße benötigt, die wir zunächst zusammenstellen.

1.1. Harmonisches Maß und extremale Länge; Abschnittssatz

Wir nehmen Bezug auf die bekannte Definition des harmonischen Maßes ω eines Randbogens $B \subset \partial G$ und des Moduls $m(\Gamma)$ einer Kurvenfamilie Γ; siehe etwa Ahlfors ([1], S. 37 ff.) oder Fuchs ([7], S. 67 ff.). Unser Ziel ist zunächst, ω durch geometrische Größen *nach oben* abzuschätzen.

Satz 1. *Es sei G ein einfach zusammenhängendes Gebiet mit $0 \in G$, $B \subset \partial G$ ein Randbogen von G, und q ein Querschnitt in G, der B von 0 trennt. Der abgeschnittene*

Teil A von G habe die Fläche F, und $d := \mathrm{dist}(B, q) > 0$. *Dann gilt für das harmonische Maß* $\omega(B)$ *von B bezüglich* 0

$$\omega \leqq \sin \tfrac{\pi}{2}\omega \leqq 4 \exp\left(-\pi \frac{d^2}{F}\right). \tag{1.1}$$

Rechts kann die Zahl π *durch keine größere ersetzt werden, und* 4 *kann durch keine Zahl* <2 *ersetzt werden.*

Die Abschätzung eignet sich besonders, wenn der Querschnitt q eine „lange Tasche“ von G abschneidet; dann ist d^2/F groß, folglich ω klein. Ist die Tasche etwa ein Rechteck mit den Seiten 1 und h, und B die ferner liegende schmale Seite des Rechtecks, so ist $d=1$, $F=h$, und (1.1) liefert

$$\omega \leqq 4 \exp(-\pi/h). \tag{1.2}$$

Solche Abschätzungen sind nicht neu; eine Abschätzung bei Nevanlinna ([19], S. 77) würde $\omega \leqq \exp(-4/(\pi h))$ geben. Doch ist (1.1) allgemeiner, und die Konstanten sind hier besser.

Beweis. Es bezeichne Γ die Familie der (lokal rektifizierbaren) Kurvenbögen $\gamma \subset G$, deren Endpunkte auf B liegen und die 0 vom komplementären Bogen $\partial G \setminus B$ trennen. Zwischen dem Modul $m = m(\Gamma)$ dieser Kurvenfamilie und dem harmonischen Maß $\omega = \omega(B)$ besteht dann folgender Zusammenhang (Hersch [11], S. 320):

$$m = \frac{1}{2} \frac{K(\sin \frac{\pi}{2}\omega)}{K(\cos \frac{\pi}{2}\omega)},$$

wobei $K(r)$ $(0<r<1)$ das vollständige elliptische Integral 1. Art zum Parameter r bedeutet. Verwendet man

$$\mu(r) := \frac{\pi}{2} \frac{K'(r)}{K(r)} \leqq \log \frac{4}{r} \qquad (0<r<1)$$

(siehe Hersch [11], S. 318 oder Lehto-Virtanen [15], S. 64), so erhält man sofort

$$\sin \tfrac{\pi}{2}\omega \leqq 4 \exp\left(-\frac{\pi}{4} \cdot \frac{1}{m}\right), \tag{1.3}$$

und es bleibt übrig, $m(\Gamma)$ nach oben abzuschätzen.

Dazu führen wir in $\mathbb{C}$ eine Metrik ρ ein: Es sei $\rho = \dfrac{1}{2d}$ in dem durch q abgeschnittenen Teil A von G, und $\rho = 0$ sonst in $\mathbb{C}$. Für jeden Bogen $\gamma \in \Gamma$ gilt dann

$$\int_\gamma \rho\, ds = \frac{1}{2d} \int_\gamma ds \geqq \frac{1}{2d} \cdot 2d = 1,$$

das heißt ρ ist eine für Γ zulässige Metrik. Daher ist

$$m(\Gamma) \leqq \iint_{\mathbb{C}} \rho^2\, db = \iint_A \rho^2\, db = \frac{1}{4d^2} F,$$

und (1.1) folgt nun aus (1.3).

Zur Beurteilung der Güte von (1.1) betrachten wir den Halbstreifen

$$G=\{z\colon \operatorname{Re} z> -n, |\operatorname{Im} z|<\tfrac{\pi}{2}\};$$

B sei der vertikale Teil von ∂G, q die Strecke $[-\frac{\pi}{2}i, +\frac{\pi}{2}i]$. Dann wird $d=n$ und $F=n\pi$. Durch konforme Abbildung von G auf die rechte Halbebene findet man ferner $\omega(B)=\dfrac{4}{\pi}e^{-n}+O(e^{-2n})$, und daraus ergibt sich die in Satz 1 gemachte Behauptung.

1.2. Abschätzung eines harmonischen Maßes nach unten

Als nächstes stellen wir eine Abschätzung des harmonischen Maßes eines Randbogens $B\subset\partial G$ nach unten bereit. Dabei stützen wir uns wesentlich auf ein neueres Ergebnis von Lesley [17]. Wir sagen, das Gebiet G erfülle eine $(\alpha\pi)$-*Keilbedingung* $(0<\alpha\leqq 1)$, wenn es zu jedem Punkt $P\in\partial G$ einen Sektor S vom Öffnungswinkel $\alpha\pi$ und von festem Durchmesser gibt, dessen Spitze in P liegt und mit $S^0\subset G$. Jedes konvexe Gebiet erfüllt eine solche Bedingung für ein geeignetes $\alpha>0$.

Hilfssatz 1 (Lesley [17]). *Es sei G ein Jordan-Gebiet mit $0\in G$, welches einer $(\alpha\pi)$-Keilbedingung genüge $(0<\alpha\leqq 1)$, und f bilde $\mathbb{D}$ konform auf G ab mit $f(0)=0$. Dann gilt*

$$|f(w_1)-f(w_2)|\leqq c_1|w_1-w_2|^\alpha \qquad (w_1, w_2\in\mathbb{D}), \tag{1.4}$$

wobei c_1 nur von G abhängt.

Weitere neuere Arbeiten, die sich mit der Aussage $f\in\operatorname{Lip}\alpha$ befassen, sind die von Näkki und Palka [18], Lesley [16] und Becker und Pommerenke [2].

Aus Hilfssatz 1 folgern wir

Hilfssatz 2. *Es sei G ein Gebiet wie in Hilfssatz 1, und $B\subset\partial G$ sei ein Randbogen vom Durchmesser h. Dann gilt für das harmonische Maß von B bezüglich 0*

$$\omega(B)\geqq c_2 h^{1/\alpha}, \tag{1.5}$$

wobei c_2 nur von G abhängt.

Beweis. Es sei B' ein Teilbogen von B, dessen Enden z_1, z_2 den Abstand h haben; es sei $z_j=f(w_j)$ mit $w_j\in\partial\mathbb{D}$ $(j=1,2)$. Dann gilt wegen (1.4)

$$h=|z_1-z_2|=|f(w_1)-f(w_2)|\leqq c_1|w_1-w_2|^\alpha\leqq c_1[l(w_1,w_2)]^\alpha,$$

wenn $l(w_1,w_2)$ die Länge des Bogens von w_1 nach w_2 bezeichnet. Da

$$l(w_1,w_2)=2\pi\,\omega(B')\leqq 2\pi\,\omega(B)$$

ist, folgt die Behauptung.

1.3. Harmonisches Maß in schmalen Rechtecken

Zur Untersuchung der konformen Abbildung gewisser Kammgebiete werden wir nun das harmonische Maß in schmalen Rechtecken näher betrachten. R_h

bezeichne ein Rechteck der Höhe 2 und der Breite $h<1$, B sei eine schmale Seite von R_h und B_1, B_2 die an B anschließenden vertikalen Strecken der Länge h.

Zunächst gilt für das *harmonische Maß* $\omega(B)$ *von* B *bezüglich des Mittelpunktes von* R_h

$$\omega(B)=\frac{2}{\pi}\left(\frac{1}{\operatorname{Cos}\alpha}-\frac{1}{3\operatorname{Cos}3\alpha}+\frac{1}{5\operatorname{Cos}5\alpha}-+\dots\right), \quad \text{wo } \alpha=\frac{\pi}{2}\cdot\frac{2}{h}=\frac{\pi}{h} \tag{1.6}$$

ist. Diese Darstellung erhält man, wenn man $\omega(B)$ als Lösung eines Dirichlet-Problems mit der bekannten Methode des Reihenansatzes ermittelt. Daraus folgt $\omega(B)\simeq\frac{4}{\pi}e^{-\pi/h}(h\to 0)$ und daher

$$\omega(B)>e^{-\pi/h}, \quad \text{sobald } h \text{ hinreichend klein ist.} \tag{1.7}$$

Diese Abschätzung gilt auch für B_1 *und* B_2. Um dies zu sehen, betrachten wir die von B_1, B_2 erzeugten harmonischen Maße $\omega_1(z), \omega_2(z)$ in R_h. Es ist nicht schwer zu sehen, daß $\omega_1(z)+\omega_2(z)\geqq\frac{1}{2}$ ist auf der Verbindungsstrecke B_3 der Mitten von B_1, B_2. Also ist $\omega_1(z)+\omega_2(z)\geqq\frac{1}{2}\omega_3(z)$ in dem Rechteck der Höhe $2-h$ und Breite h, dessen schmale Seite B_3 ist. Nimmt man dies im Mittelpunkt M von R_h und beachtet $\omega_1(M)=\omega_2(M)$, d.h. $\omega(B_1)=\omega(B_2)$, so kommt

$$\omega(B_1)\geqq\frac{1}{4}\omega(B_3)\simeq\frac{1}{4}\cdot\frac{4}{\pi}\exp\left(-\frac{\pi}{2}\frac{2-h}{h}\right)=\frac{e^{\pi/2}}{\pi}e^{-\pi/h},$$

also gilt auch

$$\omega(B_1)=\omega(B_2)>e^{-\pi/h}, \quad \text{sobald } h \text{ hinreichend klein ist.} \tag{1.8}$$

Mit (1.7) und (1.8) beweisen wir nun

Hilfssatz 3. *Die Punkte* z', z'' *mögen den (kleineren) Randbogen* $\widehat{z', z''}$ *von* R_h *begrenzen. Für hinreichend kleine* h *gilt dann:*

$$\textit{Aus } \omega(\widehat{z', z''})<e^{-\pi/h} \quad \textit{folgt } |z'-z''|<\sqrt{2}h.$$

Im Prinzip handelt es sich hier, wie in 1.2, um die Abschätzung eines harmonischen Maßes nach unten, jedoch ist hier das Gebiet von h abhängig.

Beweis. 1. Fall: z' und z'' liegen auf einer schmalen Seite von R_h. Dann ist nichts zu zeigen. 2. Fall: z' und z'' liegen auf einer langen Seite von R_h. Dann überlegt man zunächst, daß das harmonische Maß von $\widehat{z', z''}$ monoton abnimmt, wenn man die Strecke $\widehat{z', z''}$ bis zu einer Ecke von R_h schiebt (Anwendung der Schwarz-Christoffel-Formel). Wegen (1.8) muß daher $|z'-z''|<h$ sein. 3. Fall: z' und z'' liegen auf verschiedenen Seiten von R_h. Für kleine h ist dies nur dann möglich, wenn $z'\in B$ und $z''\in B_1$. Nach dem Gesagten muß daher $|z'-z_E|<h$ und $|z''-z_E|<h$ sein, wenn z_E eine Ecke von R_h ist. Dann ist $|z'-z''|<\sqrt{2}h$.

1.4. Harmonisches Maß und Länge einer Strecke

Es sei $H=\{z\colon \operatorname{Re} z>0\}$, und b und B seien zwei Intervalle auf dem Rand ∂H:

$$b=(A\,i,(A+\Delta)\,i), \qquad B=(A\,i,\infty i);$$

ihre harmonischen Maße bezüglich $z=1$ seien $\omega(b)$ und $\omega(B)$. Der nachfolgende Hilfssatz drückt in harmonischen Maßen aus, wann die Länge Δ von b klein ist.

Hilfssatz 4. a) *Es sei*

$$\text{(i)}\ A\geqq 1\ \textit{und}\ \text{(ii)}\ \omega(b)\leqq\frac{\pi}{8}\,\omega(B).$$

Dann gilt

$$\Delta\leqq\frac{16}{\pi^2}\,\frac{\omega(b)}{[\omega(B)]^2}. \tag{1.9}$$

b) *Ist $A\geqq 1$, so gilt $\Delta\to 0$ genau dann, wenn $\omega(b)/[\omega(B)]^2\to 0$.*

Man bemerkt, daß (1.9) ohne die Annahme (ii) falsch wird: Dazu hält man B fest und läßt Δ groß werden.

Beweis. a) Wir verwenden, daß das harmonische Maß eines Intervalls auf ∂H bezüglich $z=1$ gleich Sehwinkel: π ist; dies ergibt zunächst

$$\pi\,\omega(B)\leqq \operatorname{tg}(\pi\,\omega(B))=\frac{1}{A} \quad \text{und} \quad \operatorname{tg}(\pi\,\omega(b))=\frac{\Delta}{1+A\Delta+A^2}. \tag{1.10}$$

Aus der letzten Beziehung folgt wegen $A\geqq 1$, $\omega(b)\leqq\frac{1}{4}$

$$\Delta=\operatorname{tg}(\pi\,\omega(b))(1+A\Delta+A^2)\leqq 4\,\omega(b)(1+A\Delta+A^2). \tag{1.11}$$

Aus (ii) und (1.10) folgt $\omega(b)\leqq\frac{\pi}{8}\,\omega(B)\leqq\frac{1}{8A}$, also gilt

$$\Delta\leqq\frac{1}{2A}(1+A\Delta+A^2)\leqq\frac{1}{2A}(2A^2+A\Delta)=A+\frac{\Delta}{2},$$

mithin $\Delta\leqq 2A$, so daß aus (1.11) jetzt

$$\Delta\leqq 4\,\omega(b)(1+2A^2+A^2)\leqq 16\,A^2\,\omega(b)$$

folgt. Mit (1.10) ergibt sich (1.9).

b) Ist $\omega(b)/[\omega(B)]^2\to 0$, so ist (ii) von einer Stelle an erfüllt, so daß $\Delta\to 0$ aus (1.9) folgt. Ist umgekehrt $\Delta\to 0$ und $A\geqq 1$, so ergibt (1.10)

$$\operatorname{tg}(\pi\,\omega(b))=\frac{1}{1+\Delta/A+A^{-2}}\,\Delta\,[\operatorname{tg}(\pi\,\omega(B))]^2\leqq\Delta\,[\operatorname{tg}(\pi\,\omega(B))]^2,$$

folglich $\operatorname{tg}(\pi\,\omega(b))/[\operatorname{tg}(\pi\,\omega(B))]^2\to 0$. Wegen $\omega(b)\leqq\omega(B)\leqq\frac{1}{4}$ folgt daraus die Behauptung, daß $\omega(b)/[\omega(B)]^2\to 0$.

1.5. Schlichte Funktionen in der Halbebene

Schließlich stellen wir noch ein Ergebnis über schlichte Funktionen in der Halbebene $H=\{z\colon \operatorname{Re} z>0\}$ bereit.

Hilfssatz 5. *Es sei f in H regulär und schlicht, und es sei $G=f(H)$. Dann gilt*

$$|f'(z)| \leqq \frac{2}{x} \operatorname{dist}(f(z), \partial G \qquad \text{für } z=x+iy \in H. \tag{1.12}$$

Ist das Bildgebiet G beschränkt, so gilt daher

$$f'(x_0+iy)\to 0 \qquad \text{für } y\to \pm\infty,$$

für jedes feste $x_0>0$.

In (1.12) entsteht Gleichheit zum Beispiel für $f(z)=z^2$ und $z=x>0$.

Beweis. Wir führen (1.12) auf ein entsprechendes Ergebnis für schlichte Funktionen in $\mathbb{D}$ zurück. Durch $w=\frac{1-z}{1+z}$, $z=\frac{1-w}{1+w}$ bilden wir $\mathbb{D}$ auf H ab; die Rechnung ergibt

$$\frac{1}{1+z}=\frac{1+w}{2} \quad \text{und} \quad 1-|w|^2=x\,|1+w|^2.$$

Setzt man $f(z)=f(z(w))=F(w)$, $w\in\mathbb{D}$, so wird

$$f'(z)=F'(w)\cdot\frac{-2}{(1+z)^2}=-\frac{1}{2}(1+w)^2\,F'(w)$$

und also $|f'(z)|=\frac{1}{2x}(1-|w|^2)\,|F'(w)|$.

Nach Pommerenke ([21], S. 22) ist $(1-|w|^2)|F'(w)|\leqq 4\operatorname{dist}(F(w),\partial G)$, also gilt (1.12).

§ 2. Konforme Abbildung von Kamm- und Schlangengebieten

Wie schon bemerkt, steht im Zentrum des weiteren Interesses die Frage, wann das Element $\mathrm{id}\in\Sigma(G)$ isoliert liegt. Sie wird an Kamm- und Schlangengebieten G studiert werden. Daher sollen jetzt einige feinere Untersuchungen über die KA f einer Halbebene H oder eines Streifens S auf ein solches Gebiet vorausgestellt werden.

Entspricht dabei ∞ dem nicht punktförmigen PE von G, so ist natürlich $\lim_{w\to\infty} f(w)$ nicht vorhanden. Unsere Ergebnisse besagen aber, daß der Grenzwert doch vorhanden ist bei gewisser (tangentieller) Annäherung an ∞, oder daß $f(w)$ für $w\to\infty$ in gewissem Sinne langsam schwankt.

2.1. Konforme Abbildung von Schlangengebieten

Wir beginnen mit zwei Definitionen.

Definition 1 *(Schlangengebiet). Es sei G ein einfach zusammenhängendes, beschränktes Gebiet mit einem nicht punktförmigen* PE P, *alle anderen* PE *seien* *punktförmig; ferner sei* $0 \in G$. *Dann heißt* G *ein Schlangengebiet, wenn es zu jedem* $z \in G$, $z \neq 0$, *einen Querschnitt* $q(z)$ *durch* z *gibt, der* 0 *von* P *trennt, so daß* $\operatorname{diam} q(z) \to 0$ *strebt für* $z \to P$.

Darunter fallen alle Außenschlangen und Innenschlangen, d.h. Gebiete, die sich von außen oder innen an den Einheitskreis heranwinden, sowie andere schlangenförmige Gebiete und gewisse Kammgebiete. Notwendig ist jedenfalls, daß das PE P aus lauter Hauptpunkten besteht, jedoch ist diese Bedingung nicht hinreichend.

Definition 2 (Langsam schwankende Funktion). *Eine im Streifen* $S = \{w: 0 < \operatorname{Im} w < 1\}$ *erklärte Funktion* f *heißt langsam schwankend für* $\operatorname{Re} w \to +\infty$, *wenn für jedes* $M > 0$ *gilt:*

$$|f(w_1) - f(w_2)| \to 0, \qquad \text{falls } w_1, w_2 \in S,\ \min(\operatorname{Re} w_1, \operatorname{Re} w_2) \to +\infty, |w_1 - w_2| \leqq M. \tag{2.1}$$

Aus der Definition folgt sofort: Ist f in S stetig, ferner stetig bei Annäherung an jeden endlichen Randpunkt von S, und langsam schwankend für $w \to +\infty$ und für $w \to -\infty$, so ist f in S *gleichmäßig stetig:*

Zu $\varepsilon > 0$ gibt es $\delta > 0$ so, daß $|f(w_1) - f(w_2)| < \varepsilon$, sobald $|w_1 - w_2| < \delta$, $w_1, w_2 \in S$. – Unser Ziel ist nun

Satz 2. *Es sei* G *ein Schlangengebiet mit dem nicht punktförmigen* PE P, *und* f *bezeichne eine* KA *des Parallelstreifens* $S = \{w: 0 < \operatorname{Im} w < 1\}$ *auf* G, *für die* $f(w) \to P$ *strebt für* $\operatorname{Re} w \to +\infty$. *Dann ist* f *für* $\operatorname{Re} w \to +\infty$ *langsam schwankend.*

Da f unter unseren Annahmen stetig ist bei Annäherung an die übrigen Randpunkte von S, so folgt aus Satz 2 sofort der

Zusatz. *Die Abbildungsfunktion* f *von Satz 2 ist in* S *gleichmäßig stetig.*

Beweis. Angenommen, die Aussage von Satz 2 sei falsch. Dann gibt es $M > 0$ und Punkte $w_n, w_n' \in S$ mit $w_n, w_n' \to \infty$ und $|w_n - w_n'| \leqq M$, für die

$$|f(w_n) - f(w_n'| \geqq 2\delta > 0.$$

Ihre Bilder $z_n = f(w_n)$ und $z_n' = f(w_n')$ streben dann gegen P, aber $|z_n - z_n'| \geqq 2\delta$. Dann gilt für die in Definition 1 genannten Querschnitte $q(z_n)$ und $q(z_n')$ von einer Stelle an

$$\operatorname{dist}(q(z_n), q(z_n')) \geqq \delta.$$

Wir zeigen: Die extremale Distanz von $q(z_n)$ nach $q(z_n')$ strebt gegen ∞ für $n \to \infty$; das von $q(z_n)$, $q(z_n')$ und einem Teil von ∂G berandete Viereck ist vom Typ „langer Schlauch". Dazu sei Γ_n die Klasse der stetigen Kurvenbögen γ, die $q(z_n)$ in G mit $q(z_n')$ verbinden. Jedes γ hat eine Länge $\geqq \delta$, so daß $\rho = \frac{1}{\delta}$ eine für Γ_n zulässige Metrik ist. Damit gilt für den Modul

$$M(\Gamma_n) \leqq \iint \rho^2 \, db = \delta^{-2} \cdot F_n,$$

wenn F_n die Fläche des Abschnitts von G zwischen $q(z_n)$ und $q(z'_n)$ ist. Diese Fläche strebt gegen 0 für $n\to\infty$, also gilt $M(\Gamma_n)\to 0$ oder $\lambda(\Gamma_n)=1/M(\Gamma_n)\to\infty$ $(n\to\infty)$.

Nun betrachten wir die Urbilder in S, das heißt $e_n=f^{-1}(q(z_n))$ und $e'_n=f^{-1}(q(z'_n))$. Sie schieben sich nach $+\infty$ hinaus, und da sie $+\infty$ von $f^{-1}(0)$ trennen, verlaufen sie für $n>N$ jeweils vom unteren zum oberen Rand von S; dabei ist $w_n\in e_n$ und $w'_n\in e'_n$, und e'_n liege rechts von e_n. Wir setzen wie üblich

$$\alpha_n=\sup\{\operatorname{Re} w\colon w\in e_n\} \quad \text{und} \quad \beta_n=\inf\{\operatorname{Re} w\colon w\in e'_n\}.$$

Wegen der konformen Invarianz gilt für die extremale Distanz $d(e_n, e'_n)$ zwischen e_n und e'_n

$$d(e_n, e'_n)=\lambda(\Gamma_n)\to\infty \qquad (n\to\infty).$$

Nach dem Ahlforsschen Verzerrungssatz gilt aber

$$d(e_n, e'_n)\leqq 2\Lambda(e^{\pi(\beta_n-\alpha_n)}) \tag{2.2}$$

(siehe etwa Ahlfors [1], S. 76–77); dabei ist $\Lambda(R)$ der Modul des Teichmüllerschen Ringgebiets G_R, somit $\Lambda(R)\nearrow\infty$ für $R\nearrow\infty$. Daher folgt aus (2.2) $\beta_n-\alpha_n\to\infty$ und insbesondere $|w_n-w'_n|\to\infty$ $(n\to\infty)$. Dies widerspricht unserer Annahme, also ist f für $w\to+\infty$ langsam schwankend, und Satz 2 ist bewiesen.

2.2. Konforme Abbildung von Kammgebieten 1. Art

Das wohl geläufigste Gebiet mit einem nicht punktförmigen PE entsteht etwa dadurch, daß man in einem Rechteck unendlich viele Schlitze einführt, die sich an einer Rechteckseite häufen. Etwas allgemeiner betrachten wir jetzt ein Gebiet G der in Fig. 1 beschriebenen Art:

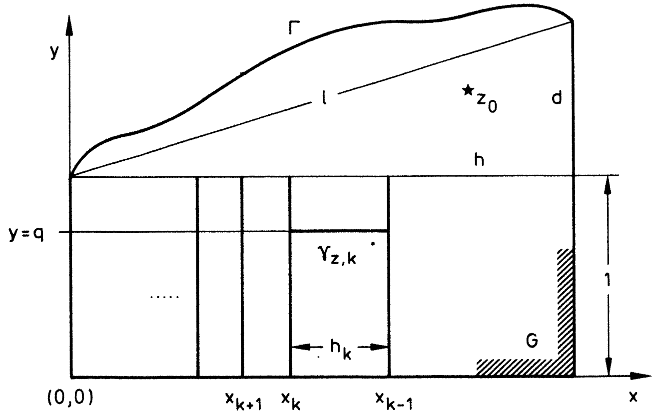


Fig. 1. Kammgebiet 1. Art

Die Gerade l sei von der Form $y=1+cx$ mit $c>0$; der obere Teil von ∂G sei ein Jordanbogen Γ, der oberhalb l verläuft. Später ist wichtig, daß das von l, h und d begrenzte Dreieck G_1 in G liegt. Die Schlitze der Länge 1 häufen sich gegen die Strecke $0\ldots i$, von ihren Abständen h_k wird zunächst nichts weiteres verlangt. z_0 sei ein Bezugspunkt in G_1.

Schließlich bezeichne g die KA von G auf die Halbebene $H=\{w\colon \operatorname{Re} w>0\}$, bei der $g(z_0)=1$ ist und wo das PE P, dessen Projektion die Strecke $0\dots i$ ist, nach ∞ abgebildet wird. Die Umkehrung von g heiße f. Wir interessieren uns vor allem für das Verhalten von $f(w)$, wenn w in der Nähe der imaginären Achse nach ∞ strebt.

Dazu betrachten wir zunächst für *festes* q mit $0<q<1$ die Querschnitte

$$\gamma_{z,k}=\{z=x+iq\colon x_k\leqq x\leqq x_{k-1}\} \qquad k=1,2,\dots;$$

ihre Bilder $\gamma_{w,k}=g(\gamma_{z,k})$ sind Jordanbögen in $\bar{H}$, welche zwei Punkte $A_k i$ und $B_k i$ verbinden, die auf ∂H liegen. Grundlegend für alles weitere ist nun

Satz 3. *Mit den vorgenannten Bezeichnungen gilt* $\operatorname{diam}\gamma_{w,k}\to 0$ $(k\to\infty)$.

Bezeichnet g_k den „Sack“, den $\gamma_{z,k}$ von G abtrennt, so folgt aus Satz 3, daß auch $\operatorname{diam} g(g_k)\to 0$ strebt für $k\to\infty$; denn es ist $\operatorname{diam} g(g_k)=\operatorname{diam}\gamma_{w,k}$. – Interessant ist in diesem Zusammenhang die Frage, was geschieht, wenn $q=1$ gewählt wird. Dann trennt $\gamma_{z,k}$ einen Sack ab, dessen seitlicher und unterer Rand $\Gamma_{z,k}$ heiße; sein Bild $g(\Gamma_{z,k})$ sei $\Gamma_{w,k}\subset\partial H$. *Wie verhalten sich die Längen* $|\Gamma_{w,k}|$?

Wählt man $h_k=\text{const.}\cdot k^{-p}$ $(p>1)$, so zeigt grobe Abschätzung, daß vermutlich $|\Gamma_{w,k}|\to\infty$ $(k\to\infty)$ für $p>\frac{3}{2}$, $|\Gamma_{w,k}|\to 0$ $(k\to\infty)$ für $p<\frac{3}{2}$, während $|\Gamma_{w,k}|$ gegen einen positiven Grenzwert konvergiert, wenn $p=\frac{3}{2}$ ist.

Beweis von Satz 3. Wir arbeiten wieder mit dem harmonischen Maß ω, zunächst des Querschnitts $\gamma_{z,k}$ und der Strecke $\Gamma_{z,k}$, jeweils bezüglich $z=z_0$. Dabei soll $\Gamma_{z,k}$ jetzt das rechte Ufer von $\{z=x_k+iy\colon 1-h_k\leqq y\leqq 1\}$ sein. *Wir zeigen zunächst*

$$\omega(\gamma_{z,k})\leqq a\,e^{-b/h_k} \tag{2.3}$$

$$\omega(\Gamma_{z,k})\geqq c\,h_k^{\alpha} \tag{2.4}$$

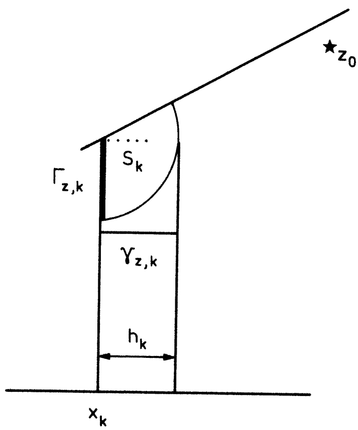


Fig. 2.

für gewisse Konstanten a, b, c und $\alpha > 0$. Die Abschätzung (2.3) folgt sofort aus Satz 1 von § 1; man hat dort $F = (1-q) h_k$ und $d = (1-q)$ zu wählen. Um (2.4) zu sehen, betrachte man obenstehende Skizze; die schräg laufende Gerade hat dieselbe Steigung $c > 0$ wie l. Die durch $\Gamma_{z,k}$ erzeugte Maßfunktion ist im Kreissektor ($\subset G$!) mindestens so groß wie diejenige harmonische Funktion, welche auf $\Gamma_{z,k}$ gleich 1 und sonst am Rande des Sektors 0 ist. Damit ist sie $\geqq \tau > 0$ auf der punktierten Strecke s_k, und daher gilt für das stets in z_0 gemessene harmonische Maß

$$\omega(\Gamma_{z,k}, G) \geqq \tau\, \omega(s_k, G_1);$$

G_1 ist das oben eingeführte Dreieck.

Nun kann Hilfssatz 2 von § 1 angewendet werden; er liefert

$$\omega(s_k, G_1) \geqq \text{const} \left(\frac{h_k}{2}\right)^{\alpha} \tag{2.5}$$

für ein positives α, und somit gilt auch (2.4).

Jetzt gehen wir in die Halbebene H: Dort liegt das Bild $\gamma_{w,k}$ von $\gamma_{z,k}$, welches ein Jordanbogen von $A_k i$ nach $B_k i$ ist; wir können $1 \leqq A_k < B_k$ annehmen. Es bezeichne b_k die von $w = A_k i$ aus genommene zirkulare Projektion von $\gamma_{w,k}$ auf die imaginäre Achse. b_k ist demnach ein Intervall $[A_k i, (A_k + \Delta_k) i]$ der Länge Δ_k. Da alle Punkte von $\gamma_{w,k}$ in $|w - A_k i| \leqq \Delta_k$ liegen, gilt

$$\operatorname{diam} \gamma_{w,k} \leqq 2 \Delta_k.$$

Um das harmonische Maß von b_k (bezüglich $w = 1$) abzuschätzen, wenden wir das Hall'sche Lemma an; siehe Gaier [8], S. 97 oder Fuchs [7], S. 82. Und zwar wird es zunächst auf einen von $A_k i$ ausgehenden Teilbogen b'_k von b_k angewendet; sodann läßt man b'_k gegen b_k rücken und erhält, wieder bei Messung im Punkt $w = 1$,

$$\omega(\gamma_{w,k}) \geqq \omega(b_k).$$

Insbesondere ist daher wegen (2.3) auch

$$\omega(b_k) \leqq a e^{-b/h_k}, \tag{2.6}$$

während das harmonische Maß des Halbstrahls $B_k = [A_k i, \infty i)$ wegen (2.4) sicher $\geqq c h_k^{\alpha}$ ist. Daher wird Hilfssatz 4 von § 1 anwendbar und bringt leicht $\Delta_k \to 0$, und dies wiederum beweist Satz 3.

Wir kommen nun zu unserem Hauptergebnis bezüglich Kammgebieten 1. Art.

Satz 4. *Es sei G ein Kammgebiet der oben eingeführten Art, und f sei die* KA *der Halbebene $H = \{w : \operatorname{Re} w > 0\}$ auf G, bei der $f(1) = z_0$ ist und wo ∞ dem* PE *P von G entspreche. Dann ist das Bild von $\{w\colon\ w = u_0 + iv,\ v \geqq 0\}$ für jedes $u_0 > 0$ ein Jordanbogen, d.h. es existiert $\lim\limits_{v \to +\infty} f(u_0 + iv)$ für jedes $u_0 > 0$.*

Ob der Jordanbogen sogar rektifizierbar ist, muß offen bleiben.

1. Beweis. Es bezeichne B_z den lokal rektifizierbaren Bogen $\{f(u_0+iv)\colon v\geqq 0\}$. Er konvergiert jedenfalls gegen das ausgezeichnete PE P von G. Wir müssen zeigen, daß er nur gegen den Hauptpunkt i von P konvergiert, daß er also nicht in die Taschen des Gebiets "einsinkt". Angenommen, für ein $\delta>0$ gebe es in der k-ten Tasche Punkte $P_{z,k}$, $Q_{z,k}$ auf B_z mit $\operatorname{Im} P_{z,k}=1-\delta$, $\operatorname{Im} Q_{z,k}=1-2\delta$. Ihre Bilder in H seien $P_{w,k}$ und $Q_{w,k}$. Nach Satz 3 gilt $|P_{w,k}-Q_{w,k}|\to 0$, und Hilfssatz 5 von §1 zeigt daher, daß der Teilbogen von $P_{z,k}$ nach $Q_{z,k}$ eine Länge $\to 0$ hat, $k\to\infty$. Dies widerspricht $|P_{z,k}-Q_{z,k}|\geqq\delta$.

2. Beweis. Wir zeichnen den Querschnitt $\gamma_{z,k}$ auf der Höhe $1-\delta$ ein. Satz 3 besagt, daß $\operatorname{diam}\gamma_{w,k}\to 0$ $(k\to\infty)$, also müssen die Realteile von $P_{w,k}$ und $Q_{w,k}$ gegen Null streben. Dies widerspricht $\operatorname{Re} P_{w,k}=\operatorname{Re} Q_{w,k}=u_0>0$.

Abschließend wollen wir noch auf eine mögliche *Verallgemeinerung* hinweisen. Sie besteht darin, daß wir als obere Begrenzung Γ von G einen Bogen zulassen, der sich der Geraden $y=1$ im Punkt $z=i$ *tangentiell* nähert. Während (2.3) unverändert richtig bleibt, wird (2.4) problematisch: Dafür wurde benötigt, daß der dort genannte Kreissektor in G liegt, außerdem war für (2.5) die Konvexität von G_1 ausgenützt.

Unser Beweis läßt sich jedoch modifizieren, wenn wir fordern:

(i) Die obere Begrenzung Γ von G erfülle $y(x)\geqq 1+cx^\alpha$ mit $c>0$, $\alpha>1$;
(ii) Für die Abstände h_k zwischen den Schlitzen gelte $h_k=O(x_k^\alpha)$ $(k\to\infty)$.

Die Aussagen der Sätze 3 und 4 gelten dann unverändert weiter. Während im Fall $\alpha=1$ über die Lage der Schlitze nichts gefordert wurde, verlangt (ii), daß die Schlitze dicht genug liegen müssen. Zum Beispiel erhält man für $x_k=\frac{1}{k^p}$ $(p>0)$ die Bedingung $p\leqq\frac{1}{\alpha-1}$, während (ii) für die Folge $x_k=\frac{1}{k!}$ nie erfüllt ist.

2.3. *Konforme Abbildung von Kammgebieten 2. Art*

Wir stellen noch eine dritte Klasse von Gebieten vor. Bei diesen Kammgebieten wird, im Gegensatz zu den vorgenannten, das nichtpunktförmige PE durch die Einschnitte nahezu abgetrennt. Das wird zur Folge haben, daß bei ihnen die Abbildung id in $\Sigma(G)$ nicht isoliert liegt. Es ist günstig, G symmetrisch zu $\mathbb{R}$ anzunehmen.

1. Es bezeichne f die KA des Streifens $S=\{w\colon|\operatorname{Im} w|<\frac{\pi}{2}\}$ auf das Gebiet G, wobei $f(0)=0$ und $f'(0)>0$ sei. Wir zeigen, daß f auf ∂S in gewissem Sinne langsam schwankt, wenn die Öffnungen ε_k zwischen gegenüberliegenden Schlitzen klein sind im Vergleich zum Abstand h_k der Schlitze.

Satz 5. *Für die Größen ε_k, h_k des Kammgebiets von Fig. 3 gelte*

$$\prod_{j=1}^{k-1}\frac{\varepsilon_j}{2h_j}<e^{-\pi/h_k}\quad\textit{für } k>k_0,\quad\textit{sowie}\quad\frac{\varepsilon_k}{h_k}\to 0\ (k\to\infty),\tag{2.7}$$

während ε_0 so klein sei, daß $\varepsilon_0\leqq 2h_0$ und $e^{-\alpha_0}\leqq\frac{1}{32}$ ist. Dann gilt

$$f(w)-f(w')\to 0,\quad\textit{falls } w,w'\in\partial S,\quad |w-w'|<e^{-\operatorname{Re} w}\quad\textit{und}\quad w,w'\to\infty.\tag{2.8}$$

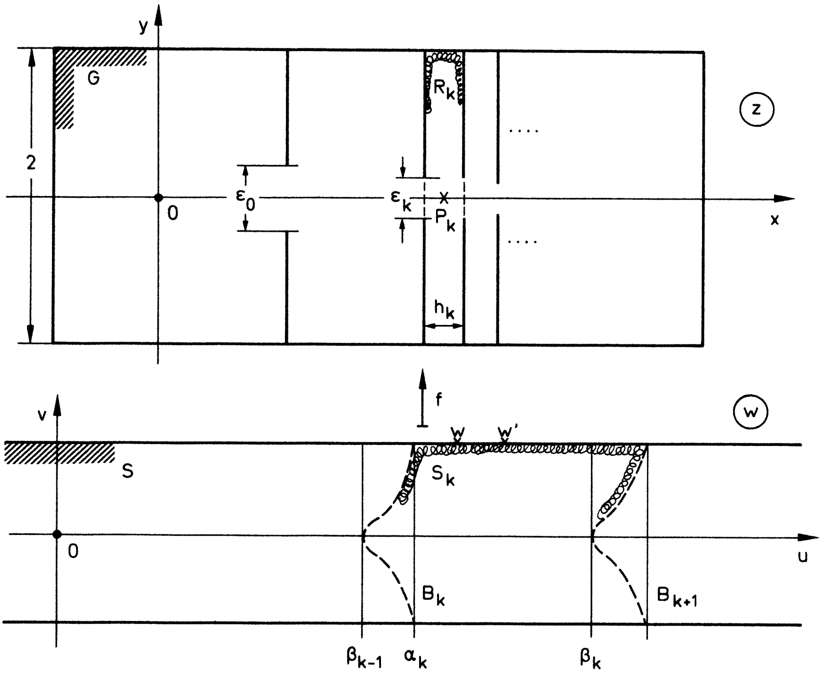


Fig. 3. Kammgebiet 2. Art

Beweis. Wir zerlegen G in die Teilrechtecke R_k; ihre Mittelpunkte seien P_k, und ihre Bilder unter f^{-1} heißen $S_k \subset S$. Die Bilder der Querschnitte der Länge ε_k sind Jordanbögen $B_k \subset \bar{S}$, die 0 von ∞ trennen, und es sei wie üblich

$$\alpha_k = \sup\{\operatorname{Re} w \colon w \in B_k\}, \qquad \beta_k = \inf\{\operatorname{Re} w \colon w \in B_{k+1}\}.$$

Da $\beta_k - \alpha_k \to \infty$ [2], liegen w und w' auf dem Rand desselben Abschnitts S_k oder auf dem Rand zweier angrenzender S_k. Es genügt offenbar, (2.8) zu zeigen, wenn w, w' auf dem Rand *eines* S_k liegen. Ist $d = |w - w'| < e^{-\operatorname{Re} w}$, so gilt für das harmonische Maß von $\widehat{ww'}$ bezüglich eines beliebigen Punktes $P \in \mathbb{R}$

$$\omega(P, S, \widehat{ww'}) < d < e^{-\operatorname{Re} w} \leqq e^{-\beta_{k-1}}; \tag{2.9}$$

die erste Ungleichung folgt am einfachsten, wenn man den Streifen S auf eine Halbebene erweitert. Unser nächstes Ziel ist daher, β_{k-1} nach unten abzuschätzen.

Für die extremalen Distanzen d gilt (siehe Ahlfors [1], S. 55)

$$d(B_0, B_k) \geqq \sum_{j=0}^{k-1} d(B_j, B_{j+1}) \geqq \frac{1}{\pi} \sum_{j=0}^{k-1} \log \frac{2h_j}{\varepsilon_j}. \tag{2.10}$$

Hier ist $d(B_j, B_{j+1}) \geqq \frac{1}{\pi} \log \frac{2h_j}{\varepsilon_j}$ noch zu begründen. Für $\varepsilon_j \geqq 2h_j$ ist dies sicher

[2] Dies gilt wegen $\varepsilon_k/h_k \to 0$ $(k \to \infty)$, weil dann die extremale Distanz $d(B_k, B_{k+1}) \to \infty$ strebt (siehe unten)

richtig. Ist aber $\frac{\varepsilon_j}{2} < h_j$, so zeichnen wir die beiden Halbkreise k_j und k'_j der Radien $\frac{\varepsilon_j}{2}$ und h_j um den Mittelpunkt des Querschnitts der Länge ε_k. Sicher ist dann $d(B_j, B_{j+1}) \geqq d(k_j, k'_j)$ (Vergleichsprinzip!), und letztere Distanz ist $\frac{1}{\pi} \log \frac{2h_j}{\varepsilon_j}$, wie man nach einer log-Abbildung sieht.

Auf der anderen Seite ist nach dem Ahlforsschen Verzerrungssatz ([1], S. 76–77)

$$d(B_0, B_k) \leqq 2\Lambda(R) \quad \text{mit} \quad R = e^{\beta_{k-1} - \alpha_0}.$$

Dies zusammen mit (2.10) gibt

$$\prod_{j=0}^{k-1} \frac{2h_j}{\varepsilon_j} \leqq e^{2\pi \Lambda(R)} \leqq 16(R+1)$$

([1], S. 76). Da von einer Stelle an $R \geqq 1$, so haben wir mit unserer Wahl von ε_0

$$\prod_{j=1}^{k-1} \frac{2h_j}{\varepsilon_j} \leqq \prod_{j=0}^{k-1} \frac{2h_j}{\varepsilon_j} \leqq 32R = 32e^{-\alpha_0} \cdot e^{\beta_{k-1}} \leqq e^{\beta_{k-1}},$$

also gilt von einer Stelle an

$$e^{-\beta_{k-1}} \leqq \prod_{j=1}^{k-1} \frac{\varepsilon_j}{2h_j}.$$

Nun seien z, z' die Bildpunkte von w, w' unter f; sie liegen auf ∂R_k, jedoch nicht auf den beiden Querschnitten von G. Aus (2.9) folgt erst recht

$$\omega(P_k, R_k, \widehat{zz'}) < e^{-\beta_{k-1}} \leqq \prod_{j=1}^{k-1} \frac{\varepsilon_j}{2h_j},$$

jedenfalls für hinreichend große k. Nun hatten wir (2.7) verlangt, so daß

$$\omega(P_k, R_k, \widehat{zz'}) < e^{-\pi/h_k},$$

und der zu diesem Zweck bereitgestellte Hilfssatz 3 zeigt, daß $|z-z'| < \sqrt{2}h_k$ ist, falls k hinreichend groß ist. Dies ergibt die Behauptung (2.8).

2. Wichtiger als die Streifenabbildung f ist für uns die KA der rechten Halbebene $H = \{w\colon \operatorname{Re} w > 0\}$ auf G. Dabei sei $F(1) = 0$ und $F'(1) > 0$.

Satz 6. *Für das Kammgebiet G der Fig. 3 sei die Bedingung* (2.7) *erfüllt. Dann gilt für die* KA F *von* H *auf* G

$$F(w+ih) - F(w) \Rightarrow 0 \qquad (w \in H, h \to 0). \tag{2.11}$$

Beweis. Wir betrachten zwei Punkte $w = iv$, $w' = iv'$ mit $1 \leqq v \leqq v'$, $v' - v \leqq 1$, auf der imaginären Achse. Ihre Bilder unter $\omega = \log w$ fallen auf den Rand des in Satz 5 auftretenden Streifens S, und zwar nach ω, ω' mit $|\omega - \omega'| < e^{-\operatorname{Re}\omega}$. Zu gegebenem $\varepsilon > 0$ gibt es nach Satz 5 ein $R > 0$ so, daß $|F(w+ih) - F(w)| < \varepsilon$

ausfällt, sofern $h \leqq 1$ und $w = iv$, $v \geqq R$ ist; entsprechendes gilt für $w = iv$, $v \leqq -R$ (Symmetrie). Da aber F auf ∂H stetig ist, gilt

$$|F(w+ih) - F(w)| < \varepsilon \quad \text{für } w \in \partial H, \text{ sofern } 0 < h \leqq h_0 \text{ ist.} \tag{2.12}$$

Und weil F in H beschränkt ist, läßt sich für jedes solche h das Maximumprinzip anwenden, und (2.12) gilt sogar für $w \in H$.

3. Wir zeigen jetzt, daß das Bild von $\{w: w = u_0 + iv\}$ unter F für jedes feste $u_0 > 0$ zwischen den Einschnitten des Kammgebiets beliebig weit "einsinkt".

Satz 7. *Für das Kammgebiet G der Fig. 3 sei die Bedingung* (2.7) *erfüllt. Dann gilt für die* KA *F von H auf G*

$$\overline{\lim_{v \to +\infty}} \operatorname{Im}\{F(u_0 + iv)\} = 1 \quad \text{für jedes feste } u_0 > 0. \tag{2.13}$$

Beweis. Wir nehmen zwei Punkte w, w' auf ∂H vom Abstand 1 her, die sich zusammen nach ∞ bewegen, und setzen $w'' = u_0 + \frac{1}{2}(w + w')$. Das harmonische Maß der Strecke ww' bezüglich w'' ist dann konstant > 0. Andererseits haben wir im Beweis von Satz 6 gesehen, daß $F(w) - F(w') \to 0$ strebt. Wählt man w, w' so, daß $z = F(w)$ und $z' = F(w')$ am oberen Rand einer Tasche von G liegen, so sieht man, daß das harmonische Maß von $\widehat{zz'}$ bezüglich $F(w'')$ nur dann $\geqq c > 0$ sein kann, wenn der Meßpunkt $F(w'')$ selbst nahe z, z' liegt; anderenfalls Anwendung des Abschnittssatzes aus § 1. Also gibt es eine Folge von Punkten w'' auf $\operatorname{Re} w = u_0$, für die $\operatorname{Im} F(w'')$ gegen 1 strebt.

4. Nun studieren wir noch die Bedingung (2.7). Zunächst ist klar, daß es zu jeder Folge $\{h_j\}$ eine Folge $\{\varepsilon_j\}$ gibt so, daß (2.7) gilt.

Weiter ist (2.7) erfüllbar für jede Nullfolge $\left\{\frac{\varepsilon_j}{h_j}\right\}$, wenn $\{h_j\}$ hinreichend langsam gegen 0 strebt. Allerdings kann es vorkommen, daß dann $\Sigma h_j = \infty$ wird und folglich G unbeschränkt wird. Dies ist z.B. der Fall, wenn $h_j = \frac{1}{j}$ und $\varepsilon_j = \frac{1}{j^\alpha}$ $(\alpha > 1)$ gewählt wird.

Jedoch wird G beschränkt (was wir immer gefordert hatten), wenn etwa

$$h_j = \frac{1}{j^\alpha} \quad (\alpha > 1) \quad \text{und} \quad \varepsilon_j = e^{-j^\beta} \quad (\beta > \alpha - 1)$$

gewählt wird. Die "Blenden-Öffnungen" ε stehen dann zum Abstand r der Blenden gegen das Primende in der Beziehung $\varepsilon \approx \exp\left\{-\left(\frac{1}{r}\right)^\gamma\right\}$ für ein $\gamma > 0$; das PE wird also von den Einschnitten stark abgeschnürt.

§ 3. Wann ist id in $\Sigma(G)$ isoliert?

Wir kommen nun auf unsere eingangs gestellte Frage zurück und behandeln das Problem, wann das Element id isoliert liegt. Notwendige und hinreichende geometrische Bedingungen für G können wir nicht angeben, vielmehr wird die Frage an den in § 2 studierten Gebietsklassen diskutiert. Grob kann man jedoch sagen, daß id immer dann isoliert liegt, wenn G tiefe Taschen aufweist, welche „gut erreichbar" sind. Das ist z.B. der Fall bei unseren Kammgebieten 1. Art.

Zunächst erinnern wir an Hilfssatz 6 von § 5.2. Hat G ein isoliertes nicht punktförmiges PE P_0, und ist die SA $\varphi \in \Sigma(G)$ hinreichend nahe an id, so wird P_0 durch φ festgehalten; das heißt, die lineare Abbildung L von $\mathbb{D}$ nach $\mathbb{D}$ in der Darstellung $\varphi = f_0 L g_0$ hat den Fixpunkt $g_0(P_0)$. *Hat folglich G drei isolierte nicht punktförmige* PE, *so liegt* id *in $\Sigma(G)$ isoliert*, ganz unabhängig von der Art der PE. Denn die Abbildung L hat dann drei Fixpunkte auf $\partial\mathbb{D}$, so daß $L = \mathrm{id}$ ist, sobald φ hinreichend nahe an id liegt.

Nach Denjoy [5] und Piranian [20] gibt es Gebiete G, deren PE alle dieselbe Projektion haben, für die also $d(P_\alpha)$ konstant >0 ist. Ob für diese id isoliert liegt, muß offen bleiben.

3.1. Der Fall der Schlangengebiete

Hat G nur punktförmige PE, so ist $\Sigma(G)$ zu $\Sigma(\mathbb{D})$ homöomorph, also liegt id in $\Sigma(G)$ nicht isoliert. Die Schlangengebiete (zur Def. siehe § 2.1) zeigen uns zunächst, daß es Gebiete mit nicht punktförmigen PE gibt, für die id trotzdem nicht isoliert liegt. Wir zeigen etwas mehr.

Satz 8. *Es sei G ein Schlangengebiet, P_0 sein nicht punktförmiges* PE, *und $\Sigma'(G)$ bezeichne die Menge aller* SA *φ von G, welche P_0 fest lassen. Dann ist $\Sigma'(G)$ homöomorph zur Halbebene $H = \{a+ib\colon a>0,\ b \in \mathbb{R}\}$, wobei* $\mathrm{id} \in \Sigma'(G)$ *dem Punkt $1 \in H$ entspricht. Folglich ist* id *in $\Sigma'(G)$ nicht isoliert.*

Beweis. Es sei f eine KA von $H_w = \{w\colon \operatorname{Re} w > 0\}$ auf G, bei der ∞ dem PE P_0 entspreche, und $g = f^{-1}$. Jede SA $\varphi \in \Sigma'(G)$ ist dann von der Form

$$\varphi = \varphi_L = f L g, \qquad \text{wo } L(w) = aw + ib \text{ ist}, \qquad a>0,\ b \in \mathbb{R}.$$

Hierdurch werden die Paare (a, b) mit $a>0$, $b \in \mathbb{R}$ bijektiv auf die SA $\varphi_L \in \Sigma'(G)$ abgebildet. Wir zeigen, daß dies ein Homöomorphismus ist.

Zunächst hängt φ_L stetig von (a, b) ab. Denn

$$\begin{aligned} d(\varphi_L, \varphi_{L'}) &= \sup_G |\varphi_L(z) - \varphi_{L'}(z)| = \sup_G |f L g(z) - f L' g(z)| \\ &= \sup_{H_w} |f L(w) - f L'(w)| \\ &= \sup_{H_w} |f(aw+ib) - f(a'w+ib')|. \end{aligned} \tag{3.1}$$

Weil nun die KA eines Parallelstreifens S auf G für $w \to +\infty$ langsam schwankt (Satz 2), ist hier

$$f(w_1) - f(w_2) \to 0, \qquad \text{falls } |w_1| \to \infty \quad \text{und} \quad 0 < m \leqq \left|\frac{w_1}{w_2}\right| \leqq M < \infty.$$

Hält man (a,b) fest und wählt (a',b') mit $|a-a'|\leqq\frac{a}{2}$, $|b-b'|\leqq 1$, so wird daher die rechte Seite von $(3.1)<\varepsilon$, sobald $|w|>R=R(\varepsilon)$ ist. Für $|w|\leqq R$ wird diese rechte Seite $<\varepsilon$, sobald $|a-a'|$ und $|b-b'|$ hinreichend klein sind; denn G hat nur ein nichtpunktförmiges PE, und daher ist f auf ∂H_w stetig.

Die Umkehrabbildung $\Sigma'(G)\to H$ ist immer stetig. Sei etwa $\varphi_{L'}\to\varphi_L$ im Sinne der Metrik von $\Sigma(G)$, also (siehe oben)

$$f(a'w+ib')\Rightarrow f(aw+ib) \qquad \text{für } w\in H_w$$

und gewisse variable Paare (a',b'). Dann gilt dies insbesondere in einer kompakten Scheibe von H_w, und die Anwendung von g auf einen kompakten Teil von G bringt $a'w+ib'\Rightarrow aw+ib$ auf einer Scheibe in H_w. Also muß (a',b') gegen (a,b) streben. Damit ist Satz 8 vollständig bewiesen.

Wir studieren noch kurz die Menge $\Sigma''(G)$ aller SA φ von G, die außer P_0 noch ein weiteres, von φ unabhängiges PE fest lassen. $L(w)=aw+ib$ hat dann außer ∞ noch einen Fixpunkt $w_0=iv_0$, $v_0\in\mathbb{R}$, so daß für die zugeordneten Paare $(a,b)\in H$ noch $av_0+b=v_0$ gilt. Bei festem v_0 ist dies eine Gerade durch den Punkt (1,0). *Aus Satz 8 folgt, daß* $\Sigma''(G)$ *zu dem Strahl* $H\cap\{(a,b)\colon av_0+b=v_0\}$ *homöomorph ist,* wobei id auf (1,0) abgebildet wird. Daher ist id sogar in $\Sigma''(G)$ nicht isoliert.

3.2. Der Fall der Kammgebiete 1. Art

Die wohl einfachsten Gebiete G, für die id in $\Sigma(G)$ isoliert liegt, sind die Kammgebiete 1. Art. Wir beweisen die Vermutung von Gauthier:

Satz 9. *Ist G ein Kammgebiet 1. Art, so liegt* id *in $\Sigma(G)$ isoliert.*

Beweis. (i) Da G nur ein PE P hat, läßt jede SA φ, die hinreichend nahe an id liegt, P fest (Hilfssatz 6 von § 5.2); aus $z\to P$ folgt $\varphi(z)\to P$. Darüber hinaus zeigt eine einfache Überlegung: Läuft z am Rande ∂G gegen P, so auch $\varphi(z)$, und zwar liegen z und $\varphi(z)$ in derselben „Tasche“, wenn etwa z am Boden, auf der Höhe $y=0$, liegt, sobald φ nahe genug an id liegt (Methode „Eisenbahnzug“). Solch eine SA φ sei jetzt betrachtet.

(ii) Wie in § 2.2 bezeichne g die KA. von G auf $H=\{w\colon \operatorname{Re} w>0\}$, bei der $g(z_0)=1$ ist und wo das PE P nach ∞ abgebildet wird, und es sei $f=g^{-1}$. Jeder SA φ von G, welche P fest läßt, entspricht eine SA L von H, die ∞ fest läßt: $L(w)=aw+ib$ mit $a>0$, $b\in\mathbb{R}$. Wir zeigen, daß notwendig $L(w)=w$, folglich $\varphi=\text{id}$ ist.

(iii) Dazu betrachten wir eine Folge $\{z_n\}$, $z_n\in\partial G$, $\operatorname{Im} z_n=0$, $0<\operatorname{Re} z_n\to 0$ und die zugehörigen Punkte $z'_n=\varphi(z_n)$, $w_n=g(z_n)$, $w'_n=g(z'_n)=L(w_n)\in\partial H$. Wir messen jetzt das harmonische Maß der Strecke w_n, w'_n von einem „fahrenden Beobachter“ P_n aus: Dieser liege auf dem Halbstrahl $\Gamma_w=\{w\colon \operatorname{Re} w=1, \operatorname{Im} w\geqq 0\}$ mit der Ordinate $\frac{1}{2}(\operatorname{Im} w_n+\operatorname{Im} w'_n)$.

Das Bild dieses Halbstrahls ist ein Jordanbogen $\Gamma_z\subset G$, der bei $z=i\in\partial G$ einmündet (Satz 4), und es sei $Q_n:=f(P_n)$. Da z_n, z'_n in derselben Tasche stecken, strebt das harmonische Maß des Randbogens z_n, z'_n, gemessen in Q_n, gegen

Null, und wegen der konformen Invarianz des harmonischen Maßes gilt dasselbe für das Maß der Strecke w_n, w'_n, gemessen in P_n.

Dies impliziert aber $w_n - w'_n \to 0$, d.h. $w_n - L(w_n) \to 0$ oder

$$(a-1)w_n + ib \to 0 \qquad (n \to \infty).$$

Da $w_n \to \infty$ $(n \to \infty)$, folgt $a=1$, $b=0$. Es ist also $L(w)=w$ und daher $\varphi = \mathrm{id}$.

Es können noch andere Gebiete konstruiert werden, für die id in $\Sigma(G)$ isoliert liegt. Dies kann einmal durch Verallgemeinerung der Idee geschehen, die dem Kammgebiet 1. Art zugrunde liegt. Eine andere Konstruktionsmethode geht von einer beliebigen Jordankurve aus, der abzählbar viele Taschen aufgesetzt werden, welche durch eingezogene Schlitze nachher fast wieder abgetrennt werden. Wir verzichten auf eine genauere Darstellung.

3.3. Selbstabbildungen mit zwei Fixpunkten

Jetzt sei G ein Gebiet, welches eine Folge von Taschen besitze (vgl. Skizze), die durch Querschnitte q_n mit $\operatorname{diam} q_n \to 0$ von G abgetrennt werden können. Betrachtet werden SA φ von G, welche *zwei* von φ unabhängige PE fest lassen; die Menge aller dieser SA heiße $\Sigma''(G)$. (Zum Beispiel gehören dazu alle hinreichend nahe an id gelegenen φ, wenn G zwei isolierte nicht punktförmige PE besitzt.)

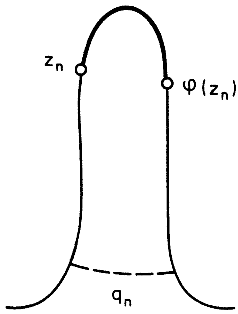


Fig. 4.

Wir zeigen: id *liegt isoliert in* $\Sigma''(G)$. Andernfalls wähle man Punkte z_n, wie es die Skizze zeigt, und ihre Bilder $\varphi(z_n)$ mit φ nahe id. Die extremale Distanz von q_n zum Randbogen $z_n, \varphi(z_n)$ strebt dann $\to \infty$ $(n \to \infty)$.

Nun bilden wir G konform auf einen Parallelstreifen S ab: $S = g(G)$, $G = f(S)$, wobei die beiden ausgezeichneten PE nach $\pm\infty$ fallen sollen. Jeder SA $\varphi \in \Sigma''(G)$ entspricht dann eine SA von S, die $\pm\infty$ festhält und also eine *Translation* ist. Die extremale Distanz von $g(q_n)$ zum Bild von z_n, $\varphi(z_n)$ unter g – dies ist ein Intervall der Länge h auf ∂S – muß ebenfalls gegen ∞ streben. Sie ist aber höchstens gleich der des Intervalls zur gegenüberliegenden Seite von ∂S, kann also nicht beliebig groß werden. Dieser Widerspruch zeigt, daß man φ nicht beliebig nahe an id finden kann.

3.4. Der Fall der Kammgebiete 2. Art

Die Ergebnisse der letzten Abschnitte legen die Vermutung nahe, daß id in $\Sigma(G)$ immer dann isoliert liegt, wenn G eine Folge schmaler werdender Taschen besitzt. Dies ist jedoch nicht der Fall, wenn die Taschen selbst „schwer zugänglich" sind.

Satz 10. *Ist G ein Kammgebiet 2. Art, für welches die Bedingung* (2.7) *gilt, so liegt* id *in $\Sigma(G)$ nicht isoliert.*

Beweis. Entsprechend den Bezeichnungen in Satz 6 sei F die KA von $H=\{w: \operatorname{Re} w>0\}$ auf G mit $F(1)=0$, $F'(1)>0$. Dabei sei (2.7) erfüllt, und zu $\varepsilon>0$ sei $h>0$ so gewählt, daß

$$|F(w+ih)-F(w)|<\varepsilon \quad \text{ausfällt für } w\in H.$$

Setzt man $z=F(w)$, so bedeutet dies

$$|F(F^{-1}(z)+ih)-z|<\varepsilon \quad \text{für } z\in G,$$

also $\operatorname{dist}(\varphi, \mathrm{id})<\varepsilon$ für die SA $\varphi(z)=F(F^{-1}(z)+ih)$ von G auf sich. Damit ist schon alles gezeigt.

§ 4. Abschließende Bemerkungen

1. Paul Gauthier hat in einer problem session in Varna (September 1981) vorgeschlagen, Gebiete G, für die id in $\Sigma(G)$ isoliert liegt, „starr" zu nennen. Da der Begriff der Starrheit (rigidity) aber schon von vielen Autoren (z.B. Aumann, Carathéodory, Heins, Reich, Rudin) für verschiedene Eigenschaften verwendet worden ist, haben wir verzichtet, ihn einzuführen.

2. Man kann fragen, ob sich unsere Untersuchungen auf p-fach zusammenhängende Gebiete übertragen lassen, $2\leqq p\leqq\infty$. Nach dem Theorem von Klein und Poincaré (siehe Weyl [26], S. 156) ist $\Sigma(G)$ in den Fällen $2<p\leqq\infty$ stets eine diskontinuierliche Gruppe; id und daher alle Elemente von $\Sigma(G)$ liegen folglich isoliert. Nach Koebe ([14], S. 324) läßt überdies ein p-fach zusammenhängendes Gebiet G, $2<p<\infty$, höchstens $L(p)<\infty$ konforme SA zu; die Zahl $L(p)$ wurde von Heins [10] ermittelt.

Interessant sind also nur einfach und zweifach zusammenhängende Gebiete. Im letzteren Fall besteht $\Sigma(G)$ nur aus konformen Drehungen und Stürzungen, so daß sich die Struktur von $\Sigma(G)$ wesentlich von der hier vorgestellten unterscheiden wird.

3. Auch die Einführung anderer Metriken in $\Sigma(G)$ dürfte zu anderen Ergebnissen führen. Yang Lo schlägt vor, statt unserer Metrik

$$d(\varphi_1, \varphi_2)=\Big(\iint_G |\varphi_1(z)-\varphi_2(z)|^2\, db\Big)^{1/2}$$

zu verwenden.

4. Die im Verlauf der Arbeit aufgetretenen offenen Fragen sollen hier nochmals kurz zusammengestellt werden.

4.1. Ist der metrische Raum $\Sigma(G)$ stets lokal kompakt? (Zu I, § 2)

4.2. Es sei f in $\mathbb{D}$ schlicht und beschränkt, und es sei $f(e^{i\varepsilon_n}w)\Rightarrow f(w)$ $(w\in\mathbb{D})$ für eine Nullfolge $\{\varepsilon_n\}$. Gilt dann $f(e^{i\varepsilon}w)\Rightarrow f(w)$ $(w\in\mathbb{D})$ für $\varepsilon\to 0$? Äquivalent dazu: Hat dann G nur punktförmige Primenden? (Zu I, § 3)

4.3. Ist $\Sigma(G)$ schon dann eine topologische Gruppe, wenn (4.4) gilt für alle konformen Drehungen $\psi = \psi_\varepsilon$ von $\Sigma(G)$? (Zu I, § 4)

4.4. Ist $\Sigma(G)$ nur dann eine topologische Gruppe, wenn G nur punktförmige Primenden hat oder wenn id in $\Sigma(G)$ isoliert liegt? (Zu I, § 4)

4.5. Es sei $\Sigma(G)$ zusammenhängend. Hat dann G nur punktförmige Primenden? (Zu I, § 5)

4.6. Es sei f die KA der Halbebene H auf ein Kammgebiet 1. Art (Teil II, § 2). Ist das Bild von $\{w\colon w = u_0 + iv\}$ $(u_0 > 0)$ unter der Abbildung f rektifizierbar?

Bemerkung bei der Korrektur. Dr. G. Schmieder (Hannover) hat inzwischen die Fragen 4.1 und 4.2 beantwortet. $\Sigma(G)$ ist stets lokal kompakt. Zu 4.2 kann er ein Gegenbeispiel angeben.

Literatur

1. Ahlfors, L.V.: Conformal invariants. New York-Toronto: McGraw-Hill 1973
2. Becker, J., Pommerenke, Ch.: Hölder continuity of conformal mappings and non-quasiconformal Jordan curves. Comment. Math. Helv. **57**, 221–225 (1982)
3. Behnke, H., Sommer, F.: Theorie der analytischen Funktionen einer komplexen Veränderlichen. Berlin-Göttingen-Heidelberg: Springer 1955
4. Collingwood, E.F., Lohwater, A.J.: The theory of cluster sets. Cambridge: University Press 1966
5. Denjoy, A.: Sur la représentation conforme. C.R. Acad. Sci. Paris **213**, 15–17 (1941)
6. Franz, W.: Topologie I. Berlin-New York: Walter de Gruyter 1973
7. Fuchs, W.H.J.: Topics in the theory of functions of one complex variable. Princeton: Van Nostrand 1967
8. Gaier, D.: A note on Hall's lemma. Proc. Amer. Math. Soc. **37**, 97–99 (1973)
9. Gauthier, P.: Problem 7.43 in: Aspects of contemporary complex analysis. Ed. D.A. Brannan and J.G. Clunie. New York: Academic Press 1980
10. Heins, M.: On the number of 1–1 directly conformal maps which a multiply-connected plane region of finite connectivity $p(>2)$ admits onto itself. Bull. Amer. Math. Soc. **52**, 454–457 (1946)
11. Hersch, J.: Longueurs extrémales et théorie des fonctions. Comment. Math. Helv. **29**, 301–337 (1955)
12. Hornich, H.: Über einen Banachraum analytischer Funktionen. Manuscripta Math. **1**, 79–86 (1969)
13. Hornich, H.: Über den Zusammenhang von schlichten Funktionen im Banachraum B. Monatsh. Math. **75**, 239–243 (1971)
14. Koebe, P.: Abhandlungen zur Theorie der konformen Abbildung. IV. Acta math. **41**, 305–344 (1918)
15. Lehto, O., Virtanen, K.I.: Quasikonforme Abbildungen. Berlin-Heidelberg-New York: Springer 1965
16. Lesley, F.D.: Hölder continuity of conformal mappings via the strip method. To appear
17. Lesley, F.D.: Conformal mappings of domains satisfying a wedge condition. Preprint 1983
18. Näkki, R., Palka, B.: Quasiconformal circles and Lipschitz classes. Comment. Math. Helv. **55**, 485–498 (1980)
19. Nevanlinna, R.: Eindeutige analytische Funktionen. Berlin-Göttingen-Heidelberg: Springer 1953
20. Piranian, G.: Persönliche Mitteilung
21. Pommerenke, Ch.: Univalent functions. Göttingen: Vandenhoeck und Ruprecht 1975
22. Rodin, B.: Intrinsic rotations of simply connected regions. Complex Variables **2**, 319–326 (1984)
23. Tamrazov, P.M.: Contour and solid properties of holomorphic functions in a complex domain. Soviet Math. Dokl. **13**, 725–730 (1972)
24. Tamrazov, P.M.: Contour and solid structure properties of holomorphic functions of a complex variable. Russian Math. Surveys **28**, 141–173 (1973)
25. Tamrazov, P.M.: Smoothness and polynomial approximations. Russisch. Moskau: Naukova Dumka 1975
26. Weyl, H.: Die Idee der Riemannschen Fläche, 3. Aufl. Stuttgart: Teubner 1955

Eingegangen am 18. November 1983

Math. Z. 187, 259–267 (1984)

Mathematische Zeitschrift

On Gupta Representations of Central Extensions

Ralph Stöhr

Akademie der Wissenschaften der DDR, Institut für Mathematik, Mohrenstr. 39, DDR-1086 Berlin, German Democratic Republic

§ 1. Introduction

Let F be a non-cyclic free group, R a normal subgroup of F and $\gamma_c R$ the c-th term of the lower central series of R ($c \geqq 2$). In this paper we investigate the group

$$F/[\gamma_c R, F]. \tag{1}$$

In view of the exact sequence

$$1 \to \gamma_c R/[\gamma_c R, F] \to F/[\gamma_c R, F] \to F/\gamma_c R \to 1,$$

the group (1) is a central extension of $F/\gamma_c R$. In [4] C.K. Gupta and N.D. Gupta obtained a matrix representation

$$\varphi^*: F/[\gamma_c R, F] \to T(\Omega),$$

where $T(\Omega)$ is a ring of $(c+1) \times (c+1)$ matrices over a certain ring (for details see [4]). Using this representation, which we call the Gupta representation, they proved that $\gamma_c R/[\gamma_c R, F]$ is precisely the centre of (1). Moreover, it turned out that the centre decomposes into the direct sum of a free abelian group D and the kernel of φ^*:

$$\gamma_c R/[\gamma_c R, F] = \operatorname{Ker} \varphi^* \oplus D.$$

Using another method, Baumslag, Strebel and Thomson came in [1] to the same decomposition. They proved that the free abelian group D is of infinite rank if F/R is infinite. Our actual aim is to study the kernel of the Gupta representation. This kernel is known in some special cases, particulary for $c=2$. C.K. Gupta pointed out [3] that the kernel of the representation φ^* for the free centre-by-metabelian group ($R=F'$) of rank n is trivial for $n=2, 3$ and an elementary abelian 2-group of rank $\binom{n}{4}$ for $n \geqq 4$. Another proof of this result was given by J.V. Kusmin [6] who used homological considerations instead of commutator calculus. J. Lewin and T. Lewin have shown [9] that φ^* is faithful if R/R' is a projective F/R-module.

All these results are covered by a recent theorem of Kusmin [8]. He proved that if F/R has no 2-torsion, then the kernel of the Gupta representation is characterized by the following exact sequence:

$$0 \to H_4(F/R) \otimes \mathbb{Z}_2 \to \operatorname{Ker} \varphi^* \to \operatorname{Hom}(\mathbb{Z}_2, H_3(F/R)) \to 0.$$

Moreover, a complete description of $\operatorname{Ker}\varphi^*$ for F/R a one-relator group (including such with 2-torsion) was given in Kusmin's paper [7].

Concerning $c>2$ we mention that $\operatorname{Ker}\varphi^*$ has been calculated for $c=3$ and $R=F'$ ([11]).

A considerable advancement in studying the kernel of φ^* has been made by M.W. Thomson [12]. Using the approach of [1] and additional homological considerations he obtained more detailed information about $\operatorname{Ker}\varphi^*$. In fact, he proved that there is an exact sequence

$$0 \to J_c(R) \to \operatorname{Ker}\varphi^* \to H_{2c}(F/R), \tag{2}$$

where $H_{2c}(F/R)$ as usual denotes the $2c$-th integral homology group of F/R whereas $J_c(R)$ is a special functor which is known to be a $\mathbb{Z}_c$-module, i.e. a abelian group of exponent c. Thomson proved that $J_c(R)$ vanishes in some cases and stated as a theorem that $\operatorname{Ker}\varphi^*$ is then embedded in the homology group $H_{2c}(F/R)$.

In this paper we focus on the image of $\operatorname{Ker}\varphi^*$ in $H_{2c}(F/R)$. In [7] J.V. Kusmin has shown that in the case $c=2$ the image of $\operatorname{Ker}\varphi^*$ in $H_4(F/R)$ is a $\mathbb{Z}_2$-module. We strengthen the result of Thomson by generalizing that of Kusmin, i.e. we prove that the homology group $H_{2c}(F/R)$ in the exact sequence (2) can be replaced by $H_{2c}(F/R)[c]$, where, given an abelian group A, $A[n]$ denotes the subgroup $\{a \in A;\ na=0\}$ (Theorem 1). Although stated as theorems, our further results are easy consequences of Theorem 1. We prove that the kernel of the Gupta representation consists precisely of all elements of finite order of the group (1) (Theorem 2). This proves simultaneously a conjecture of C.K. Gupta and N.D. Gupta [4] concerning the relation between $[\gamma_c R, F]$ and another normal subgroup of F which is naturally defined by a certain ideal in the integral group ring $\mathbb{Z}F$ (Theorem 3). Generalizing the corresponding ($c=2$)-result of Kusmin [7], we prove that any. element of finite order in (1) has an order dividing c^2 (Theorem 4). Finally, we prove that the Gupta representation is faithful if F/R is a finite group and its order is prime to c (Theorem 5).

The arrangement of the paper is as follows. The exact sequence (2), the starting point of our further considerations, will be introduced in Sect. 2. In this Section we confine ourselves to a very condensed presentation. A detailed presentation of the matter, including a new proof of the Gupta representation Theorem for (1), can be found in M.W. Thomson's paper [12]. Generally, the acquaintance with [12] is in any way conducive to the understanding of the present paper. In Sect. 3 we prove a technical result, our main Lemma, which will be exploited in Sect. 4 to prove the theorems.

Most of our notations are standart and largely in line with those of [12]. Occasionally, well-known facts will be drawn from homological algebra without references being given; these however can easily be found, e.g. in Hilton and Stammbach [5].

§ 2. Preliminaries

Denote F/R by G and the relation module R/R' by M. Let $\mathfrak{L}M$ be the free Lie ring over M, $\mathfrak{L}_c M$ its c-th homogeneous component considered as a right G-module with naturally defined action. Then there is an isomorphism

$$\gamma_c R/[\gamma_c R, F] \cong \mathfrak{L}_c M \otimes_G \mathbb{Z},$$

where $\mathbb{Z}$ denotes the integers considered as a trivial G-module. The Lie ring $\mathfrak{L}M$ is embedded in the tensor Lie ring of the tensor ring TM over M (TM is actually the universal enveloping ring of $\mathfrak{L}M$). The restriction of this embedding to $\mathfrak{L}_c M$ gives an embedding v_c of $\mathfrak{L}_c M$ into $T_c M = M \otimes M \otimes \ldots \otimes M$ (c times, tensoring over $\mathbb{Z}$), the c-th homogeneous component of TM, which is considered as a G-module with diagonal action:

$$v_c \colon \mathfrak{L}_c M \to T_c M. \tag{3}$$

The tensor power $T_c M$ may also be viewed as a module for S_c, the full symmetric group of degree c, by defining $(a_1 \otimes \ldots \otimes a_c)\sigma^{-1} = a_{1\sigma} \otimes \ldots \otimes a_{c\sigma}$ ($a_1, \ldots, a_c \in M$, $\sigma \in S_c$). We need a special element of the integral group ring $\mathbb{Z}S_c$. Define $\varepsilon_1 = 1$ and for $c \geqq 2$

$$\varepsilon_c = (1-(1,2))(1-(1,2,3))\ldots(1-(1,\ldots,c)).$$

Now the image of $\mathfrak{L}_c M$ in $T_c M$ can be simply described by

$$\mathfrak{L}_c M v_c = T_c M \varepsilon_c. \tag{4}$$

The element ε_c has the following useful property:

$$\varepsilon_c^2 = c\,\varepsilon_c. \tag{5}$$

Recall the Magnus embedding μ. It maps the module M injectively into a free G-module P. Moreover, there is an exact sequence (well-known as the relation sequence)

$$0 \to M \to P \to \mathfrak{g} \to 0, \tag{6}$$

where $\mathfrak{g}$ denotes the augmentation ideal of the integral group ring $\mathbb{Z}G$, i.e. the kernel of the canonical homomorphism $\varepsilon \colon \mathbb{Z}G \to \mathbb{Z}$. The monomorphism μ induces an embedding

$$\mu^c \colon T_c M \to T_c P. \tag{7}$$

Using (3) and (7) we get the following commutative triangle:

$$\begin{tikzcd}[column sep=large]
\mathfrak{L}_c M \otimes_G \mathbb{Z} \arrow[rr, "\nu_c \mu^c \otimes 1"] \arrow[dr, "\nu_c \otimes 1"'] & & T_c P \otimes_G Z \\
& T_c M \otimes_G \mathbb{Z} \arrow[ur, "\mu^c \otimes 1"'] &
\end{tikzcd} \tag{8}$$

Denote the kernel of the homomorphism $\nu_c \otimes 1$ by $J_c(R)$. As it was mentioned in the introduction, $J_c(R)$ is a $\mathbb{Z}_c$-module. Indeed, besides ν_c we have a natural epimorphism $\rho_c\colon T_c M \to \mathfrak{L}_c M$ defined by $a_1 \otimes \ldots \otimes a_c \to (a_1 \otimes \ldots \otimes a_c)\, \varepsilon_c \nu_c^{-1}$. Because of (5), it turns out that the composite $\nu_c \rho_c$ is the c-th multiple map on $\mathfrak{L}_c M$. Now it follows immediately that $J_c(R) = \operatorname{Ker} \nu_c \otimes 1$ is a $\mathbb{Z}_c$-module.

Since $T_c P$ is a free G-module, $T_c P \otimes_G \mathbb{Z}$ is a free abelian group and so is the image of $\mathfrak{L}_c M \otimes_G \mathbb{Z}$ in $T_c P \otimes_G \mathbb{Z}$.

In [12] M.W. Thomson proved that the kernel of the Gupta representation φ^* is equivalent to the kernel of the horizontal map in the commutative triangle (8). Hence, the image of the homomorphism $\nu_c \mu^c \otimes 1$ is the free abelian group D which was mentioned in Sect. 1. On the other hand, for $\operatorname{Ker} \varphi^* = \operatorname{Ker}(\nu_c \mu^c \otimes 1)$ one has the exact sequence

$$0 \to J_c(R) \to \operatorname{Ker} \varphi^* \to \operatorname{Ker} \mu^c \otimes 1.$$

Finally, to get the exact sequence (2) one has to prove the following

Lemma 1 (M.W. Thomson [12]). *The kernel of the canonical homomorphism $\mu^c \otimes 1$ is $H_{2c}(G)$.*

We give a proof because some of its details will be needed in the next Section.

Proof of Lemma 1. The canonical commutative triangle

$$\begin{tikzcd}[column sep=huge]
& T_c P \otimes_G \mathbb{Z} \\
T_c M \otimes_G \mathbb{Z} \arrow[ur, "\mu^c \otimes 1"] \arrow[r, "(\mu \otimes 1^{c-1}) \otimes 1"'] & (P \otimes T_{c-1} M) \otimes_G \mathbb{Z} \arrow[u, "(1 \otimes \mu^{c-1}) \otimes 1"']
\end{tikzcd}$$

defines a decomposition of the homomorphism $\mu^c \otimes 1$. The Lemma will follow once we show that the horizontal map has the kernel $H_{2c}(G)$ and the vertical map is injective. Combining the exact sequence $0 \to \mathfrak{g} \to \mathbb{Z}G \to \mathbb{Z} \to 0$ with the relation sequence (6) we obtain the following exact sequence of G-modules:

$$0 \to M \to P \to \mathbb{Z}G \to \mathbb{Z} \to 0. \tag{9}$$

Since the i-th tensor power $T_i M$ is a free $\mathbb{Z}$-module, the exactness of (9) implies, for every non-negative integer i, the exactness of the canonical sequence

$$0 \to T_{i+1} M \to P \otimes T_i M \to \mathbb{Z}G \otimes T_i M \to T_i M \to 0$$

(diagonal action). Combining these exact sequences with differentials defined in the obvious way, we obtain an exact sequence

$$\begin{aligned} 0 \to T_c M \to P \otimes T_{c-1} M \to \mathbb{Z}G \otimes T_{c-1} M \to P \otimes T_{c-2} M \to \dots \\ \to \mathbb{Z}G \otimes M \to P \to \mathbb{Z}G \to \mathbb{Z} \to 0, \end{aligned} \tag{10}$$

where $P \otimes T_i M$, $\mathbb{Z}G \otimes T_i M$ $(0 \leqq i < c-1)$ are clearly free G-modules. By definition we have then

$$H_{2c}(G) = \operatorname{Ker}(T_c M \otimes_G \mathbb{Z} \to (P \otimes T_{c-1} M) \otimes_G \mathbb{Z}).$$

On the other hand, we have the exact sequence of G-modules

$$0 \to P \otimes T_{c-1} M \to P \otimes T_{c-1} P \to P \otimes (T_{c-1} P / T_{c-1} M) \to 0$$

(diagonal action). Since $P \otimes (T_{c-1} P / T_{c-1} M)$ is a free G-module, after tensoring with $\mathbb{Z}$ the sequence remains exact. Hence, the vertical map in our triangle is injective and the Lemma follows.

§ 3. The Main Lemma

The canonical S_c-action on $T_c M$ induces a canonical S_c-module-structure on $T_c M \otimes_G \mathbb{Z}$. Our main Lemma refers to the S_c-action on $\operatorname{Ker} \mu^c \otimes 1$, which is obviously a S_c-submodule of $T_c M \otimes_G \mathbb{Z}$.

Lemma 2. *The full symmetric group S_c acts trivially on* $\operatorname{Ker} \mu^c \otimes 1$.

Proof. Take a free resolution of the trivial G-module $\mathbb{Z}$

$$\mathscr{P}_1 \colon \dots \to P_3 \to P_2 \to P_1 \to P_0 \to \mathbb{Z},$$

where $P_0 = \mathbb{Z}G$ and $P_1 = P$, the free module from the relation sequence. Then it follows from the exactness of (9) that the image of P_2 in P $(= P_1)$ is our module M. Hence, the complex

$$\mathscr{P}_2 \colon \dots \to P_4 \to P_3 \to P_2 \xrightarrow{\alpha} M$$

forms a free G-resolution of the relation module M. Consider the c-fold tensor power $T_c \mathscr{P}_2$, i.e.

$$T_c \mathscr{P}_2 \colon \dots \to (T_c \mathscr{P}_2)_2 \to (T_c \mathscr{P}_2)_1 \to T_c P_2 \xrightarrow{\alpha^c} T_c M,$$

where $(T_c \mathscr{P}_2)_k$ is the direct sum of all c-fold tensor products $P_{i_1} \otimes \dots \otimes P_{i_c}$ with $i_j \geqq 2$ $(j = 1, \dots, c)$ and $\sum_{j=1}^{c} i_j = 2c + k$ $(k = 0, 1, 2, \dots)$ and the differentials are defined as usual for tensor products of complexes. Since M is a free $\mathbb{Z}$-module it follows by the Künneth Theorem that $T_c \mathscr{P}_2$ is a free $\mathbb{Z}$-resolution of $T_c M$. Moreover, $T_c \mathscr{P}_2$ forms even a free G-resolution of $T_c M$ ($T_c M$ and all terms of $T_c \mathscr{P}_2$ are endowed with diagonal G-action).

Now, combining (10) and $T_c\mathscr{P}_2$ we get a new G-resolution $\mathscr{P}_3$ of the trivial G-module $\mathbb{Z}$:

$$\mathscr{P}_3\colon \ldots \to (T_c\mathscr{P}_2)_1 \to T_c P_2 \to P \otimes T_{c-1} M \to \mathbb{Z}G \otimes T_{c-1} M \to \ldots$$
$$\ldots \to \mathbb{Z}G \otimes M \to P \to \mathbb{Z}G \to \mathbb{Z}.$$

The homology group $H_{2c}(G)$ can be computed by using the resolution $\mathscr{P}_3$. For, we have to tensor $\mathscr{P}_3$ with the trivial G-module $\mathbb{Z}$ and then we have to take the homology of the complex $\mathscr{P}_3 \otimes_G \mathbb{Z}$ in dimension $2c$. The connecting homomorphism $H_{2c}(G) \to T_c M \otimes_G \mathbb{Z}$ mapping the homology group onto the kernel of $\mu^c \otimes 1$ can be computed by using the following commutative diagramm.

$$\begin{tikzcd}
\ldots \arrow[r] & (T_c\mathscr{P}_2)_1 \otimes_G \mathbb{Z} \arrow[r] & T_c P_2 \otimes_G \mathbb{Z} \arrow[r] \arrow[d, "\alpha^c \otimes 1"] & (P \otimes T_{c-1} M) \otimes_G \mathbb{Z} \arrow[r] & \ldots \\
& & T_c M \otimes_G \mathbb{Z} \arrow[ur] & &
\end{tikzcd}$$

Clearly, $\operatorname{Ker} \mu^c \otimes 1$, the image of $H_{2c}(G)$ in $T_c M \otimes_G \mathbb{Z}$, is the image of the kernel of the canonical homomorphism

$$T_c P_2 \otimes_G \mathbb{Z} \to (P \otimes T_{c-1} M) \otimes_G \mathbb{Z}$$

under the action of the epimorphism $\alpha^c \otimes 1$, i.e.

$$\operatorname{Ker} \mu^c \otimes 1 = (\operatorname{Ker}(T_c P_2 \otimes_G \mathbb{Z} \to (P \otimes T_{c-1} M) \otimes_G \mathbb{Z}))\, \alpha^c \otimes 1.$$

It is clear that the S_c-action on $T_c M \otimes_G \mathbb{Z}$ is induced by the similary defined S_c-action on $T_c P_2 \otimes_G \mathbb{Z}$. Therefore, to prove that S_c acts trivially on $\operatorname{Ker} \mu^c \otimes 1$, it is sufficient to verify that any transposition $\tau = (i, i+1) \in S_c$ $(1 \leqq i < c)$ acting on $T_c P_2$ induces the identical action on $H_{2c}(G) = H_{2c}(\mathscr{P}_3 \otimes_G \mathbb{Z})$.

Return to the free resolution $\mathscr{P}_1$. Since $\mathbb{Z}$ is a free $\mathbb{Z}$-module, the Künneth Theorem implies that the tensor power $T_c\mathscr{P}_1$ forms a free $\mathbb{Z}$-resolution for $T_c\mathbb{Z}$:

$$T_c\mathscr{P}_1\colon \ldots \to (T_c\mathscr{P}_1)_2 \to (T_c\mathscr{P}_1)_1 \to T_c\mathbb{Z}G \to T_c\mathbb{Z},$$

where $(T_c\mathscr{P}_1)_k$ $(k=0, 1, 2, \ldots)$ is the direct sum of all tensor products $P_{i_1} \otimes \ldots \otimes P_{i_c}$ with $i_j \geqq 0$ and $\sum_{j=1}^{c} i_j = k$. Moreover, $T_c\mathscr{P}_1$ (as before $T_c\mathscr{P}_2$) can be taken as a G-resolution (all terms endowed with diagonal action). Because of the isomorphism

$$T_c\mathbb{Z} = \mathbb{Z} \otimes \ldots \otimes \mathbb{Z} \cong \mathbb{Z},$$

$T_c\mathscr{P}_1$ is a free G-resolution of the trivial G-module $\mathbb{Z}$. Now we construct a special chain map $f\colon T_c\mathscr{P}_1 \to \mathscr{P}_3$ inducing the identical mapping $\mathbb{Z} \to \mathbb{Z}$. We define f as follows. In dimension $2k+1$ $(0 \leqq k < c)$ f maps the direct summand

$$T_{c-k}\mathbb{Z}G \otimes P \otimes T_k P_2 \subseteq (T_c\mathscr{P}_1)_{2k+1}$$

via

$$r_1 \otimes \ldots \otimes r_{c-k} \otimes p \otimes m_1 \otimes \ldots \otimes m_k \to \left(\prod_{i=1}^{c-k} r_i \varepsilon\right)(p \otimes m_1 \alpha \otimes \ldots \otimes m_k \alpha),$$

$m_1, \dots, m_k \in P_2$, $p \in P$, $r_1, \dots, r_{c-k} \in \mathbb{Z}G$, onto $P \otimes T_k M = (\mathscr{P}_3)_{2k+1}$, whereas the other direct summands of $(T_c \mathscr{P}_1)_{2k+1}$ are sent to zero. Similary, in dimension $2k$ $(1 \leqq k < c)$ f maps the direct summand

$$T_{c-k} \mathbb{Z}G \otimes T_k P_2 \subseteq (T_c \mathscr{P}_1)_{2k}$$

via

$$r_1 \otimes \dots \otimes r_{c-k} \otimes m_1 \otimes \dots \otimes m_k \to \left(\prod_{i=1}^{c-k-1} r_i \varepsilon \right) (r_{c-k} \otimes m_1 \alpha \otimes \dots \otimes m_k \alpha)$$

onto $\mathbb{Z}G \otimes T_k M = (\mathscr{P}_3)_{2k}$, whereas the remaining direct summands of $(T_c \mathscr{P}_1)_{2k}$ are sent to zero.

To define the chain map f in higher dimensions, we note that for $k \geqq 2c$ $(\mathscr{P}_3)_k = (T_c \mathscr{P}_2)_{k-2c}$ may be viewed as a direct summand of $(T_c \mathscr{P}_1)_k$. Indeed, every direct summand of $(T_c \mathscr{P}_2)_{k-2c}$ $(k \geqq 2c)$ also occurs as a direct summand of $(T_c \mathscr{P}_1)_k$. Now we define f in dimension $k \geqq 2c$ simply as the projection of $(T_c \mathscr{P}_1)_k$ onto $(\mathscr{P}_3)_k$. Thus, the mapping f is completely defined. The verification that f is a chain map, i.e. that it commutes with differentials, is straightforward and will be omitted. The chain map f induces an isomorphism of the homology:

$$H(f \otimes 1) \colon H(T_c \mathscr{P}_1 \otimes_G \mathbb{Z}) \to H(\mathscr{P}_3 \otimes_G \mathbb{Z}).$$

Now, to any transposition $\tau = (i, i+1) \in S_c$ there is a chain map $g_\tau \colon T_c \mathscr{P}_1 \to T_c \mathscr{P}_1$ defined by

$$\begin{aligned} &(m_1 \otimes \dots \otimes m_i \otimes m_{i+1} \otimes \dots \otimes m_c) g_\tau \\ &\quad = (-1)^{k_i k_{i+1}} m_1 \otimes \dots \otimes m_{i+1} \otimes m_i \otimes \dots \otimes m_c, \end{aligned}$$

where $m_j \in P_{k_j}$, $j = 1, 2, \dots, c$, $k_j \geqq 0$ (see [2], Chap. VI, §1). The chain map g_τ induces the identical mapping $\mathbb{Z} \to \mathbb{Z}$ (indeed, g_τ acts identically on $T_c \mathbb{Z} = \mathbb{Z}$). Hence, $g_\tau \otimes 1$ induces the identical isomorphism of the homology, i.e.

$$H(g_\tau \otimes 1) \colon H(T_c \mathscr{P}_1 \otimes_G \mathbb{Z}) \to H(T_c \mathscr{P}_1 \otimes_G \mathbb{Z})$$

is the identical chain map.

Obviously, in dimension $2c$ the chain map g_τ induces via f the canonical action of $\tau \in S_c$ on $T_c P_2$ $(g_\tau f = f \tau)$. It follows that $\tau \otimes 1$ acts identically on $H_{2c}(G) = H_{2c}(\mathscr{P}_3 \otimes_G \mathbb{Z})$. Indeed, for any $t \in H_{2c}(\mathscr{P}_3 \otimes_G \mathbb{Z})$ there is a $t_1 \in H_{2c}(T_c \mathscr{P}_1 \otimes_G \mathbb{Z})$ such that $t = t_1 H(f \otimes 1)$. Now, since $H(g_\tau \otimes 1)$ is the identical isomorphism, we have

$$t(\tau \otimes 1) = t_1 H(f \otimes 1)(\tau \otimes 1) = t_1 H(g_\tau \otimes 1) H(f \otimes 1) = t_1 H(f \otimes 1) = t.$$

This completes the proof of the Lemma.

§4. Theorems

Theorem 1. *Let φ^* be the Gupta representation for $F/[\gamma_c R, F]$. Then there is an exact sequence*

$$0 \to J_c(R) \to \operatorname{Ker} \varphi^* \to H_{2c}(G)[c],$$

where $J_c(R)$ is the kernel of the canonical homomorphism $\mathfrak{L}_c M \otimes_G \mathbb{Z} \to T_c M \otimes_G \mathbb{Z}$ and $H_{2c}(G)[c]$ is the subgroup of all elements of order dividing c in the $2c$-th homology group of G.

Proof. The kernel of the Gupta representation is equivalent to the kernel of the horizontal homomorphism in (8); the kernel of $\mu^c \otimes 1$ is $H_{2c}(G)$ (see § 2). Hence, to verify the claim of the Theorem, it is sufficient to show that the intersection $\operatorname{Ker} \mu^c \otimes 1 \cap \operatorname{Im} \nu_c \otimes 1$ is a $\mathbb{Z}_c$-module. Let $m \in \operatorname{Ker} \mu^c \otimes 1 \cap \operatorname{Im} \nu_c \otimes 1$. Then it follows from (4) that there is an element $m_1 \in T_c M \otimes_G \mathbb{Z}$ such that $m_1(\varepsilon_c \otimes 1) = m$. In view of (5) we have then on the one hand

$$m(\varepsilon_c \otimes 1) = m_1(\varepsilon_c \otimes 1)(\varepsilon_c \otimes 1) = m_1(\varepsilon_c^2 \otimes 1) = c m_1(\varepsilon_c \otimes 1) = cm.$$

On the other hand, since S_c acts trivially on $\operatorname{Ker} \mu^c \otimes 1$ (Lemma 2),

$$m(\varepsilon_c \otimes 1) = m((1-(1,2))\dots(1-(1,\dots,c))) \otimes 1 = 0.$$

Hence, we have $cm = 0$ and this completes the proof of the Theorem.

Theorem 2. *The kernel of the Gupta representation for $F/[\gamma_c R, F]$ consists precisely of all elements of finite order in $F/[\gamma_c R, F]$.*

Proof. Since both, $J_c(R)$ and $H_{2c}(G)[c]$ are $\mathbb{Z}_c$-modules, the exact sequence from Theorem 1 implies that $\operatorname{Ker} \varphi^*$ is a periodic group. On the other hand, since $F/\gamma_c R$ is torsionfree for any normal subgroup R (see [10]), elements of finite order may occur only in $\gamma_c R/[\gamma_c R, F]$. But the later is the direct sum of $\operatorname{Ker} \varphi^*$ and the free abelian group D.

Consider the integral group ring $\mathbb{Z}F$. Let $\mathfrak{r} \subseteq \mathbb{Z}F$ be the kernel of the canonical homomorphism $\mathbb{Z}F \to \mathbb{Z}G$ and let $\mathfrak{f}$ be the augmentation ideal of $\mathbb{Z}F$. Define a normal subgroup $K_c(R)$ of F by

$$K_c(R) = F \cap (1 + \mathfrak{f}\mathfrak{r}^{c-2}\mathfrak{f}).$$

It is easy to see that $[\gamma_{c-1} R, F] \subseteq K_c(R)$ for all $c \geqq 3$. In [4], C.K. Gupta and N.D. Gupta conjectured that $K_c(R) = I_R([\gamma_{c-1} R, F])$, where for $S \subset R$ and R/S nilpotent $I_R(S) = \{r \in R;\, r^{k(r)} \in S \text{ for some } k(r) \geqq 1\}$ is the isolator of S in R. Now we prove this conjecture.

Theorem 3. *For $c \geqq 3$, $K_c(R) = I_R([\gamma_{c-1} R, F])$.*

Proof. In [4] C.K. Gupta and N.D. Gupta have actually proved that φ^* is a faithful representation for $F/K_{c+1}(R)$ $(c = 2, 3, \dots)$. The Theorem follows since $\operatorname{Ker} \varphi^*$ is periodic.

In [7], J.V. Kusmin proved that every element of finite order in $F/[R', F]$, i.e. in the case $c = 2$, has order 2 or 4. A corresponding result is true for all $c \geqq 2$.

Theorem 4. *Every element of finite order in $F/[\gamma_c R, F]$ has an order dividing c^2.*

Proof. By Theorem 2, every element of finite order is in $\operatorname{Ker} \varphi^*$. But the later is, by Theorem 1, an extension of the $\mathbb{Z}_c$-module $J_c(R)$ by a certain subgroup of the $\mathbb{Z}_c$-module $H_{2c}(G)[c]$.

In [7], J.V. Kusmin has given an example of a normal subgroup $R \lhd F$ such that $F/[R', F]$ contains an element of order exactly 4. This means that, in general, the extension mentioned in the proof of Theorem 4 does not split.

Finally, we strenghten a result of M.W. Thomson [12] concerning the faithfulness of the Gupta representation in the case when $G = F/R$ is a finite group.

Theorem 5. *Let R be a normal subgroup of F such that G is finite and the order of G is prime to c. Then the Gupta representation for $F/[\gamma_c R, F]$ is faithful.*

Proof. M.W. Thomson [12] has shown that under the conditions of the Theorem $J_c(R) = 0$. On the other hand, it is well-known that $|G|\, H_{2c}(G) = 0$ (see, e.g. [2], Chap. XII, § 3). Because of $(|G|, c) = 1$ it follows that $H_{2c}(G)[c] = 0$ and, thus, our Theorem is a consequence of Theorem 1.

References

1. Baumslag, G., Strebel, R., Thomson, M.W.: On the multiplicator of $F/\gamma_c R$. J. Pure Appl. Algebra **16**, 121–132 (1980)
2. Cartan, H., Eilenberg, S.: Homological Algebra. Princeton: Princeton University Press 1956
3. Gupta, C.K.: The free centre-by-metabelian groups. J. Austral. Math. Soc. **16**, 294–300 (1973)
4. Gupta, C.K., Gupta, N.D.: Generalized Magnus embeddins and some applications. Math. Z. **160**, 75–87 (1978)
5. Hilton, J.P., Stammbach, U.: A course in homological algebra. Graduate Texts in Mathematics **4**. Berlin-Heidelberg-New York: Springer 1971
6. Kusmin, J.V.: Free centre-by-metabelian groups, Lie algebras and D-groups. Izv. Akad. Nauk SSSR Ser. Mat. **41**, 3–33 (1977) [Russian]; English Transl.: Math. USSR-Izv. **11**, 1–30 (1977)
7. Kusmin, J.V.: On elements of finite order in free groups of some varieties. Mat. Sb. **119**, 119–131 (1982) [Russian]
8. Kusmin, J.V.: The structure of the free groups of some varieties. Mat. Sb. (to appear) [Russian]
9. Lewin, J., Lewin, T.: On center by abelian by one-relator groups. Comm. Pure Appl. Math. **26**, 767–774 (1973)
10. Šmelkin, A.L.: Wreath products and varieties of groups. Izv. Akad. Nauk SSSR Ser. Mat. **29**, 149–170 (1965) [Russian]
11. Stöhr, R.: On free central extensions of free nilpotent-by-abelian groups. Preprint, Berlin 1983
12. Thomson, M.W.: Representations of certain verbal wreath products by matrices. Math. Z. **167**, 239–257 (1979)

Received January 24, 1984

Math. Z. 187, 269–271 (1984)

Mathematische
Zeitschrift

All Knot Groups Are Metric

Bradd E. Clark and Victor P. Schneider

University of Southwestern Lousiana, Department of Mathematics and Statistics, Lafayette, Louisiana 70504, USA

Markov [6] asked which groups admit nontrivial Hausdorff topologies and Sharma [9] showed that the torus knot groups along with many other groups admit such a topology. The purpose of this paper is to show that all of the classical knot groups can be made into topological groups with a topology homeomorphic to the rationals. Furthermore, this topology when placed on an individual knot group can be directly related to the finite sheeted regular covering spaces of the associated knot manifold. In addition to this, we shall consider several other topologies which can be placed on knot groups.

Let $N_1, N_2 \subset G$ be normal subgroups of finite index. By considering the natural map from G to $G/N_1 \times G/N_2$ we see that $N_1 \cap N_2$ is also a normal subgroup of finite index. Hence we note from Hewitt and Ross [5] that $\{gN \mid g \in G$ and N is a normal subgroup of G with finite index$\}$ forms a basis for a topology τ on G which makes G into a topological group. G is said to be a residually finite group if and only if for every $g \in G$ with $g \neq e$ there exists a normal subgroup N of finite index in G with $g \notin N$. Obviously (G, τ) will be a Hausdorff space if and only if G is a residually finite group. We call τ the *residually finite topology* for G.

Suppose that $k \subset S^3$ is a tamely embedded knot and that $K = S^3 - k$. We call K a knot manifold and $\pi_1(K)$ a knot group.

Theorem A. *If $\pi_1(K)$ is the fundamental group of a knot manifold (Haken manifold), then $(\pi_1(K), \tau)$ is a topological group homeomorphic to the rationals.*

Proof. As we have already seen, $(\pi_1(K), \tau)$ is a topological group. Thurston [11] has shown that all knot groups are residually finite and hence $(\pi_1(K), \tau)$ is Hausdorff.

Seifert and Threlfall [8] have shown that every normal subgroup of $\pi_1(K)$ corresponds with a regular covering space of K and that there are only countably many finite sheeted regular covering spaces for K. Thus τ is first countable at e and hence τ is a metric topology.

As we noted earlier, the intersection of two normal subgroups of finite index is also a normal subgroup of finite index. Thus τ cannot be the discrete topology. Since $(\pi_1(K), \tau)$ must therefore be a dense-in-itself space, we know by [10] that $(\pi_1(K), \tau)$ is homeomorphic as a space to the rationals.

Several other Hausdorff topologies have been found which make some knot groups into topological groups. The topology described by Sharma [9] can be made metric by choosing a metric topology to place on the center of the group. However, the topology will be the discrete topology except when the knot group is a torus knot group since only the torus knot groups have an infinite center [1]. If K is a fibered knot, then by [7] we know that $[\pi_1(K), \pi_1(K)]$ is a free group. Thus by [4] we know that a nontrivial Hausdorff topology can be placed on $\pi_1(K)$. However, this topology fails to be metric.

Let N be any infinite normal subgroup of $\pi_1(K)$ (e.g., $[\pi_1(K), \pi_1(K)]$) and let $V=\{V_\alpha | V_\alpha = U_\alpha \cap N$ and $U_\alpha \in \tau\}$ be the relative topology on N. If N' is a normal subgroup of finite index in $\pi_1(K)$, then $N' \cap N \neq \{e\}$ and hence V cannot be the discrete topology. $\tau_N = \{g V_\alpha | V_\alpha \in V$ and $g \in \pi_1(K)\}$ forms a topology for $\pi_1(K)$. In general such a topology will not be compatible with the group operations. However, since $(\pi_1(K), \tau)$ extends the topology (N, V) we know by [3] that $(\pi_1(K), \tau_N)$ is a topological group and that $\tau \subset \tau_N$. Since (N, V) is metric, we know by [2] that $(\pi_1(K), \tau_N)$ is metric. Certainly, $\tau_N \neq \tau$ whenever N is of infinite index in $\pi_1(K)$. Hence many nontrivial metric topologies can be placed on $\pi_1(K)$.

Suppose that M is a regular covering space of K. We say that g_1 and g_2 *lift alike* if and only if any representative loops of g_1 and g_2 in K, g_1^* and g_2^*, either both lift to loops in M or both lift to paths in M.

Theorem B. *If $g \in \pi_1(K)$ and N is an infinite normal subgroup then there exists a nonconstant sequence $\{g_i\}$ in gN which converges to g with the property that g and all but finitely many elements of $\{g_i\}$ life alike.*

Proof. Since (N, V) is a countable topological group with a nontrivial metric topology we know that (N, V) is homeomorphic to the rationals. Hence (gN, gV) where $gV = \{gV_\alpha | V_\alpha \in V\}$ is also homeomorphic to the rationals. Let $\{g_i\}$ be a nonconstant sequence in gN which converges to g in gV.

Let M be an arbitrary finite sheeted regular covering space of K. If N' is the normal subgroup of $\pi_1(K)$ associated with $\pi_1(M)$ we know that $gN' \in \tau \subset \tau_N$. Hence all but finitely many elements of $\{g_i\}$ lie in gN'. But g_i lifts to a loop in M if and only if $g_i N' = N'$.

In a topological group G the sequence $\{g_i\}$ converges to g if and only if $\{g_i g^{-1}\}$ converges to e. This is because multiplication by a fixed element is a homeomorphism from G to G.

Corollary C. *If the nonconstant sequence $\{g_i\}$ converges to g then all but finitely many elements of $\{g_i\}$ and g will lift alike if and only if all but finitely many elements of $\{g_i g^{-1}\}$ lie in every normal subgroup of $\pi_1(K)$ of finite index.*

References

1. Burde, G., Zieschang, H.: Eine Kennzeichnung der Torusknoten. Math. Z. **167**, 169–176 (1966)
2. Clark, B., Dooley, M., Schneider, V.: Extending topologies from subgroups to groups. (To appear)

3. Clark, B., Schneider, V.: The extending topologies. (To appear)
4. Hall, M., Jr.: A topology for free groups and related groups. Ann. of Math. **52**, 127–139 (1950)
5. Hewitt, E., Ross, K.: Abstract Harmonic Analysis I. New York: Springer 1963
6. Markov, A.: On free topological groups. Izv. Akad. Nauk. SSSR **9**, 3–64 (1945) [Russian]; English Transl.: Amer. Math. Soc. Transl. **30**, 11–88 (1950)
7. Neuwirth, L.: Knot Groups. Princeton, New Jersey: Princeton University Press 1965
8. Seifert, H., Threlfall, W.: Lehrbuch der Topologie. New York: Chelsea 1947
9. Sharma, C.: Hausdorff topologies on Groups I. Math. Japon. **26**, 555–556 (1981)
10. Sierpinski, W.: Sur une propriete topologique des ensemble denombrable denses en soi. Fund. Math. **1**, 11–14 (1920)
11. Thurston, W.: Three Dimensional Manifolds, Kleinian Groups and Hyperbolic Geometry. Bull. Amer. Math. Soc. **6**, 357–381 (1982)

Received December 12, 1983; received in final form March 7, 1984

Math. Z. 187, 273–283 (1984)

Mathematische Zeitschrift

Total Minimality of the Unitary Groups

L. Stojanov

Institute of Mathematics, Bulgarian Academy of Sciences, 1090 Sofia, Bulgaria

Introduction

Dierolf and Schwanengel ([7]) showed that if X is an infinite diescrete space, then the group $F(X)$ of all bijections f: $X \to X$ provided with the topology of pointwise convergence is a totally minimal topological group. That was the first example of a (totally) minimal group which is not precompact. Other examples of such groups can be found in [8] (see also [21, § 2]). Infinite groups which does not admit non-discrete Hausdorff group topologies were constructed first by Shelah ([22], assuming CH), and then by Hesse ([14]), and A. Ol'šhanskiĭ (who noted that a quotient of the Adian's group $A(m,n)$ has the property in question, cf. [1, §13.4]). Note that all the examples mentioned above are examples of non-Abelian groups. Prodanov ([19]) established the totally minimal Abelian groups are precompact, and recently Prodanov and the author ([20]) proved that all minimal Abelian groups are precompact.

The main purpose of the present paper is to show that the unitary group of every real or complex Hilbert space provided with the strong operator topology is a totally minimal topological group. This result gives an affirmative answer to a question posed by I. Prodanov.

In Sects. 2 and 3 we study the equivariant (with respect to the action of the unitary group) compactifications of the unit sphere S of an infinite-dimensional Hilbert space. It is shown in Sect. 2 that the unit ball endowed with the weak topology is the greatest equivariant compactification of S. This fact is used in Sect. 3 to describe all equivariant compactification of S. The main theorem is proved in Sect. 4. The proof uses the scheme of the proof of [7, (1)], a generalization of which is discussed in Proposition 4.6.

Thanks are due to the referee for his helpful remarks and suggestions.

1. Preliminaries

1.1. A Hausdorff topological group G is said to be *minimal* if it does not admit a strictly coarser Hausdorff group topology. If G/N is minimal for every closed normal subgroup N of G, then G is called *totally minimal*. A great deal of information concerning minimal and totally minimal groups can be found in the surveys of Comfort and Grant ([6]) and Dikranjan ([9]).

1.2. Let G be a topological group and X be a G-space (we always assume G-spaces to be left G-spaces). A continuous mapping φ: $X \to Y$ is called a *G-compactification* of X if Y is a compact G-space, $\varphi(X)$ is dense in Y, and φ is *equivariant*, i.e. $\varphi(gx) = g\varphi(x)$ for any $g \in G$ and any $x \in X$. If moreover φ is a homeomprphism between X and $\varphi(X)$, then φ is called a *compact G-extension* of X. For the terminology and elementary properties of G-spaces we refer the reader to [24].

There is a description of the G-compactifications of a G-space X by means of algebras of functions on X (see [2, Theorem S3], [18, Theorem 2, where universal algebras instead of G-spaces are considered] or [25, Propositions 2.6 and 2.7]). Let π: $G \times X \to X$ be the action of G on X. A function f: $X \to \mathbb{C}$, where $\mathbb{C}$ is the complex plane, is called *π-uniform* ([2] or [25]) if for each $\varepsilon > 0$ there exists a neighbourhood V of the neutral element of G such that $|f(x) - f(gx)| < \varepsilon$ whenever $g \in V$ and $x \in X$. A particular result from the basic theory is the following.

1.3. If φ: $X \to Y$ is a compact G-extension of X, then φ is the greatest one (up to equivalence) iff for every continuous bounded π-uniform function f: $X \to \mathbb{C}$ there is a continuous function h: $Y \to \mathbb{C}$ with $h\varphi = f$.

1.4. Let $\mathscr{H}$ be a real or complex Hilbert space. By $\langle .,. \rangle$ we denote the inner product in $\mathscr{H}$, and by $\|\cdot\|$ – the induced norm in $\mathscr{H}$. Let S and B, respectively, be the *unit sphere* and the *unit ball* of $\mathscr{H}$ both endowed with the weak topology (which coincides with the norm topology on S). A net $\{x_\alpha\}_\alpha$ tends to x for this topology (we write briefly $x_\alpha \xrightarrow{w} x$) if and only if $\langle x_\alpha, y \rangle \xrightarrow[\alpha]{} \langle x, y \rangle$ for any $y \in \mathscr{H}$. It is easy to see that for every $x \in B$ a basic neighbourhood of x in B is given by

$$\{z \in B : |\langle e, x - z \rangle| < \varepsilon \text{ for each } e \in F\},$$

where $\varepsilon > 0$, and F is a finite subset of a fixed orthonormal basis E of $\mathscr{H}$. It is well known that B is compact, and if $\mathscr{H}$ is infinite-dimensional, then S is dense in B.

1.5. Denote by $\mathbf{U}(\mathscr{H})$ the *group of the unitary operators* on $\mathscr{H}$. It is known (cf. [3, 68.26] or [23, p. 84]) that the strong, the weak, and the ultraweak topologies coincide on $\mathbf{U}(\mathscr{H})$, and equipped with this topology $\mathbf{U}(\mathscr{H})$ is a topological group. It is known also (cf. loc. cit.) that $\mathbf{U}(\mathscr{H})$ is complete with respect to the uniform structure induced by the strong topology. In fact the uniformity in question coincides with the two-sided uniformity (cf. [5, ch. III, §3, Exercise 6]) on $\mathbf{U}(\mathscr{H})$ (the latter being considered as a topological group). So $\mathbf{U}(\mathscr{H})$ is complete with respect to its two-sided uniformity; following [12] we call such groups *sup-complete*.

Below we always assume $\mathbf{U}(\mathscr{H})$ to be equipped with the strong topology. So $T_\alpha \xrightarrow[\alpha]{} T$ in $\mathbf{U}(\mathscr{H})$ if and only if $\|T_\alpha x - Tx\| \xrightarrow[\alpha]{} 0$ whenever $x \in \mathscr{H}$. A basic neighbourhood of the *identical operator* I in $\mathbf{U}(\mathscr{H})$ is given by

$$\{T\in\mathbf{U}(\mathscr{H}):\|Te-e\|<\varepsilon \text{ for each } e\in F\},$$

where $\varepsilon>0$, and F is a finite subset of a fixed orthonormal basis E of $\mathscr{H}$.

If $\mathscr{H}$ is finite-dimensional, then $\mathbf{U}(\mathscr{H})$ is compact and therefore totally minimal. Below we restrict our attention to the infinite-dimensional case.

1.6. The *center* Z of $\mathbf{U}(\mathscr{H})$ coincides with the group of the operators λI, λ being scalars with $|\lambda|=1$. More precisely, $Z=\{I, -I\}$ in the real case, and in the complex one Z is topologically isomorphic to the *unit circle* $\mathbf{S}^1$ in the complex plane. For $x\in\mathscr{H}$ denote by $St(x)$ the *isotropy group* of x in $\mathbf{U}(\mathscr{H})$, i.e. the group of those $T\in\mathbf{U}(\mathscr{H})$ such that $Tx=x$.

It is known (and in fact follows easily from the three-dimensional case) that the groups of the type $St(x)\cdot Z(x\neq 0)$ are maximal subgroups of $\mathbf{U}(\mathscr{H})$. Consequently, one can find that any proper closed normal subgroup of $\mathbf{U}(\mathscr{H})$ is contained in Z. Another way to establish this is to describe the proper closed (for the operator norm topology) normal subgroups of $\mathbf{U}(\mathscr{H})$. The latter can be done using Banach-Lie group theory (cf. [13]), and especially the connection between Banach-Lie groups and their Banach-Lie algebras.

1.7. There is a natural action of $\mathbf{U}(\mathscr{H})$ on B, defined by $Tx=T(x)$ for $T\in\mathbf{U}(\mathscr{H})$ and $x\in B$. A straightforward verification shows that this action is jointly continuous. In other words it makes B (and hence S) a $\mathbf{U}(\mathscr{H})$-space.

The *projective space* P can be obtained from S identifying x with λx for any $x\in S$ and any λ with $|\lambda|=1$. There is a natural action of $\mathbf{U}(\mathscr{H})$ on P so that the quotient map $p\colon S\to P$ is equivariant. Taking $x\in S$, it is easy to see that the natural map $\varphi\colon \mathbf{U}(\mathscr{H})/St(x)\to S$, defined by $\varphi(T\cdot St(x))=T(x)$, is a homeomorphism. Let $q\colon \mathbf{U}(\mathscr{H})/St(x)\to\mathbf{U}(\mathscr{H})/St(x)Z$ be the quotient map. There is a unique bijection $\psi\colon \mathbf{U}(\mathscr{H})/St(x)Z\to P$ with $\psi q=p\varphi$. A simple verification shows that ψ is an equivariant homeomorphism.

Throughout this paper all topological spaces are assumed to be Hausdorff.

2. π-Uniform Functions on the Unit Sphere

Let $\mathscr{H}$ be an infinite-dimensional real or complex Hilbert space, and $\pi\colon \mathbf{U}(\mathscr{H})\times B\to B$ be the action of $\mathbf{U}(\mathscr{H})$ on B. In this section we characterize the π-uniform functions on S.

2.1. Theorem. *A function $f\colon S\to\mathbb{C}$ is π-uniform if and only if there is a continuous function $\hat{f}\colon B\to\mathbb{C}$ which coincides with f on S.*

Proof. The sets of the type

$$V_F(\varepsilon)=\{(x,y)\colon x,y\in B \text{ and } |\langle x-y,e\rangle|<\varepsilon \text{ for each } e\in F\},$$

where $\varepsilon>0$, and F is a finite subset of a fixed orthonormal basis E of $\mathscr{H}$, form a base (cf. [5, ch. II, §1.1]) for a uniformity $\mathscr{V}$ on B. Clearly the topology generated by $\mathscr{V}$ coincides with the weak topology on B.

Let $f\colon S\to\mathbb{C}$ be π-uniform. We are going to prove that f is uniformly continuous with respect to the uniformity induced by $\mathscr{V}$ on S ([5, ch. II, §§2.1 and 2.4]). Take $\varepsilon>0$. Since f is π-uniform, there exist a subset $F=\{e_1,\dots,e_n\}$ of E and $\gamma>0$ such that if $T\in\mathrm{U}(\mathscr{H})$ and $\|Te_j-e_j\|<\gamma$ for each $j=1,\dots,n$, then $|f(x)-f(Tx)|<\varepsilon$ whenever $x\in S$. Choose δ with $0<\delta<\gamma$, and such that u, $v\in[0,1]$ and $|u-v|<\delta$ imply $|\sqrt{u}-\sqrt{v}|<\gamma/2$.

Suppose $(x,y)\in(S\times S)\cap V_F(\delta/2n)$. We are to show that $|f(x)-f(y)|<\varepsilon$. There are scalars $a_1,\dots,a_n$, $b_1,\dots,b_n$ and elements x' and y' of F such that $x=a_1e_1+\dots+a_ne_n+x'$ and $y=b_1e_1+\dots+b_ne_n+y'$. By $(x,y)\in V_F(\delta/2n)$

$$|a_j-b_j|<\delta/2n \qquad (j=1,\dots,n). \tag{1}$$

Take an arbitrary element e of $E\setminus F$ and set

$$z=a_1e_1+\dots+a_ne_n+\sqrt{1-|a_1|^2-\dots-|a_n|^2}\cdot e,$$

and

$$t=b_1e_1+\dots+b_ne_n+\sqrt{1-|b_1|^2-\dots-|b_n|^2}\cdot e.$$

It follows from $\|x\|=1$ that $\|x'\|=\sqrt{1-|a_1|^2-\dots-|a_n|^2}$, hence there is $T_1\in\mathrm{U}(\mathscr{H})$ with $T_1e_j=e_j$ for $j=1,\dots,n$ and $T_1(x')=\sqrt{1-|a_1|^2-\dots-|a_n|^2}\cdot e$. Thus $T_1x=z$. Similarly, there is $T_2\in\mathrm{U}(\mathscr{H})$ with $T_2e_j=e_j$ for $j=1,\dots,n$ and such that $T_2t=y$.

On the other hand $\|z-t\|<\gamma$. Indeed, by (1) we have

$$|(1-|a_1|^2-\dots-|a_n|^2)-(1-|b_1|^2-\dots-|b_n|^2)|\leqq 2\sum_{j=1}^{n}|a_j-b_j|<\delta,$$

and the choice of δ implies

$$|\sqrt{1-|a_1|^2-\dots-|a_n|^2}-\sqrt{1-|b_1|^2-\dots-|b_n|^2}|<\gamma/2.$$

Combining the latter with (1) and using $\delta<\gamma$, we obtain

$$\|z-t\|<\sum_{j=1}^{n}|a_j-b_j|+\gamma/2<\delta/2+\gamma/2<\gamma.$$

So there exists $L\in\mathrm{U}(\mathscr{H})$ with $\|L-I\|<\gamma$ such that $Lz=t$. Set $T=T_2LT_1$. Then $T\in\mathrm{U}(\mathscr{H})$ and $Tx=y$. For each $j=1,\dots,n$ we have

$$\|Te_j-e_j\|=\|T_2Le_j-e_j\|=\|T_2Le_j-T_2e_j\|=\|Le_j-e_j\|\leqq\|L-I\|<\gamma,$$

and therefore $|f(u)-f(Tu)|<\varepsilon$ whenever $u\in S$. In particular

$$|f(x)-f(y)|=|f(x)-f(Tx)|<\varepsilon.$$

Hence f is uniformly continuous with respect to the uniformity induced by $\mathscr{V}$ on S. Since S is dense in B, f can be extended to a (uniformly) continuous function $\hat{f}\colon B\to\mathbb{C}$ ([5, ch. II, §3.6]).

If $h\colon B\to\mathbb{C}$ is a continuous function, a straightforward verification shows that the restriction of h on S is π-uniform. $\square$

We shall use the above theorem in the following equivalent form.

2.2. Corollary. *The identical embedding i: $S \to B$ is the greatest (up to equivalence) compact* $\mathrm{U}(\mathscr{H})$*- extension of* S.

Proof. Combine Theorem 2.1 with 1.3. □

3. $\mathrm{U}(\mathscr{H})$-Compactifications of the Unit Sphere

Let $\mathscr{H}$ be again an infinite-dimensional Hilbert space. In this section we shall describe (up to equivalence) all $\mathrm{U}(\mathscr{H})$-compactifications of S. For clarity we shall consider separately the real and the complex case.

Suppose first that $\mathscr{H}$ is a real Hilbert space. Taking two reals a and b with $0 \leqq b \leqq a \leqq 1$ consider the following relation $\sim$ on B.

$$x \sim y \quad \text{if and only if} \quad \begin{cases} x = y, & \text{or} \\ \|x\| \leqq a & \text{and} \quad y = -x, \text{ or} \\ \|x\| \leqq b & \text{and} \quad \|y\| \leqq b. \end{cases}$$

Clearly $\sim$ is a closed equivalence relation on B which is invariant with respect to the action of $\mathrm{U}(\mathscr{H})$ on B. Then the quotient space $B_{a,b} = B/\sim$ has a natural structure of a $\mathrm{U}(\mathscr{H})$-space such that the quotient map $p_{a,b}$: $B \to B_{a,b}$ is equivariant. Denote by $\varphi_{a,b}$ the restriction of $p_{a,b}$ on S. Clearly $\varphi_{a,b}$ is a $\mathrm{U}(\mathscr{H})$-compactification of S, and it is equivalent to $\varphi_{c,d}$ if and only if $a = c$ and $b = d$.

3.1. Theorem. *Let $\mathscr{H}$ be an infinite-dimensional real Hilbert space and φ: $S \to Y$ be a* $\mathrm{U}(\mathscr{H})$*-compactification of S. Then there are unique real numbers a and b with $0 \leqq b \leqq a \leqq 1$ such that φ is equivalent to $\varphi_{a,b}$.* □

We omit the proof since actually it is a part of the proof of Theorem 3.2 below.

Let $\mathscr{H}$ be a complex Hilbert space and m be a non-negative integer or ∞. For a symbol α of the type

$$\alpha = (a_0, a_1, \ldots, a_m, b; H_0, H_1, \ldots, H_m), \tag{2}$$

where

$$1 = a_0 > a_1 > \ldots > a_m \geqq b \geqq 0 \tag{3}$$

are reals with $a_m = \inf\{a_j : j = 1, 2, \ldots\}$ for $m = \infty$, and

$$H_0 \subsetneqq H_1 \subsetneqq \ldots \subsetneqq H_m \subset \mathbf{S}^1 \tag{4}$$

are closed subgroups of the unit circle $\mathbf{S}^1$, consider the following relation $\sim$ on B.

$$x \sim y \quad \text{if and only if} \quad \begin{cases} \text{there are an index } j \leqq m \text{ and } \lambda \in H_j \text{ with} \\ \|x\| \leqq a_j \quad \text{and} \quad y = \lambda x, \text{ or} \\ \|x\| \leqq b \quad \text{and} \quad \|y\| \leqq b. \end{cases}$$

Set $B_\alpha = B/\sim$ and denote by φ_α the restriction of the quotient map p_α: $B \to B_\alpha$ on S. Then φ_α is a $\mathbf{U}(\mathscr{H})$-compactification of S, and it is equivalent to φ_β iff α is identical to β.

3.2. Theorem. *Let $\mathscr{H}$ be an infinite-dimensional complex Hilbert space and φ: $S \to Y$ be a $\mathbf{U}(\mathscr{H})$-compactification of S. Then there exists a unique symbol α of the type* (2) *with* (3) *and* (4) *such that φ is equivalent to φ_α.*

Proof. By Corollary 2.2 there is a continuous extension ψ: $B \to Y$ of φ. Consider the following relation $\sim$ on B: $x \sim y$ if and only if $\psi(x) = \psi(y)$. Clearly $\sim$ is a closed equivalence relation on B which is invariant under the action of $\mathbf{U}(\mathscr{H})$ on B.

We begin with the following observation. If x is a non-zero element of B and $x \sim y$ for some $y \in B$ with $y \notin \mathbf{S}^1 \cdot x$, then $x \sim z$ for each $z \in B$ with $\|z\| \leqq \|x\|$. To prove this suppose first that $y = \lambda x$ for some scalar λ. Then $|\lambda| \neq 1$, and we may assume $|\lambda| < 1$ (otherwise we could replace x by y). For every $z \in B$ with $\|z\| = \|x\|$ there is $T \in \mathbf{U}(\mathscr{H})$ such that $Tx = z$, and therefore $z = Tx \sim T(\lambda x) = \lambda Tx = \lambda z$. On the other hand the unit sphere $\|z\| = \|x\|$ is dense in the ball $\|z\| \leqq \|x\|$, so $z \sim \lambda z$ for every z with $\|z\| \leqq \|x\|$. A simple induction shows now that $x \sim \lambda^n x$ for each natural n. Since $\lambda^n x \xrightarrow{w} 0$, we obtain $x \sim 0$. The above arguments shows that $z \sim 0$ for every $z \in B$ with $\|z\| \leqq \|x\|$, and then $z \sim x$.

Suppose now that $y \notin \mathbf{C} \cdot x$. Denote by G the set of those $T \in \mathbf{U}(\mathscr{H})$ such that $Tx \sim x$. Clearly G is a closed subgroup of $\mathbf{U}(\mathscr{H})$ and $St(x) \subset G$. Moreover $G \not\subset St(x) \cdot Z$. Indeed, there is $T \in \mathbf{U}(\mathscr{H})$ with $Ty = y$ and such that Tx is not collinear with x. Since $x \sim y$, we have $Tx \sim Ty = y$, and then $Tx \sim x$, i.e. $T \in G$. Clearly T does not belong to $St(x) \cdot Z$. Thus $St(x) \cdot Z$ is a proper subgroup of $G \cdot Z$, and then $G \cdot Z = \mathbf{U}(\mathscr{H})$ (cf. 1.6). Latter shows that G is a normal subgroup of $\mathbf{U}(\mathscr{H})$ and, since $St(x) \subset G$, $G = \mathbf{U}(\mathscr{H})$ (cf. 1.6 again). Thus $z \sim x$ for every z with $\|z\| = \|x\|$, and therefore all elements of the ball $\|z\| \leqq \|x\|$ are equivalent to x.

We are going to construct α. Denote by b the supremum of $\|x\|$ for $x \in B$ with $x \sim 0$. Then $z \sim 0$ for every z with $\|z\| \leqq b$. If $b = 1$, then Y contains exactly one point. In this case set $\alpha = (1)$.

Suppose $b < 1$. It follows from above that if $x \in B$ and $\|x\| > b$, then $x \sim y$ implies $y = \lambda x$ for some λ with $|\lambda| = 1$. Let a be a real number with $b < a \leqq 1$, and $\|x\| = a$. Denote by H_a the set of those $\lambda \in \mathbf{S}^1$ such that $x \sim \lambda x$. Clearly H_a is a closed subgroup of $\mathbf{S}^1$ which does not depend on the choice of x. It is easy to check that $H_a \subset H_c$ for $c \leqq a \leqq 1$. Hence there exist $m = 1, 2, \ldots$ or ∞ and real numbers

$$1 = a_0 > a_1 > \ldots > a_m \geqq b$$

such that $H_a = H_{a_j}$ for $j = 0, 1, \ldots, m-1$ and $a_j \geqq a > a_{j+1}$ and $H_a = H_{a_m}$ for $a_m \geqq a > b$. It is clear also that $a_m = \inf_j a_j$ if $m = \infty$. Set $H_j = H_{a_j}$ for $j = 0, 1, \ldots$, and define α by (2). We see now that ψ is equivalent to the quotient map p_α. Hence φ is equivalent to φ_α. □

Combining Theorems 3.1 and 3.2 we obtain the following.

3.3. Corollary. *If $\mathscr{H}$ is an infinite-dimensional real or complex Hilbert space and φ: $S \to Y$ is a $\mathbf{U}(\mathscr{H})$-compactification of S, then φ: $S \to \varphi(S)$ is an open map.*

Proof. It is sufficient to prove the assertion for $\varphi=\varphi_{a,b}$ in the real case, and for $\varphi=\varphi_\alpha$ in the complex one. Both cases follow from [5, ch. I, § 3.6]. □

4. The Main Theorem

The first part of the section contains the scheme which will be used to prove our main result.

4.1. Definition. Let G be a topological group and X be a G-space. We call X a *weakly minimal G-space* if every injective G-compactification φ: $X\to Y$ of X is an embedding, i.e. φ is a homeomorphism between X and $\varphi(X)$.

The latter means that X does not admit a strictly coarser Hausdorff G-topology with respect to which X has a compact G-extension. Obviously every compact G-space is weakly minimal. The converse is not true. For example, if X is an infinite discrete space, then X is a weakly minimal $F(X)$-space ([7, (1)]) (for the definition of $F(X)$ see the introduction).

The following proposition is an immediate consequence from Corollary 3.3.

4.2. Proposition. *The unit sphere S and the projective space P* (1.6) *of every real or complex Hilbert space $\mathscr{H}$ are weakly minimal $\mathrm{U}(\mathscr{H})$-spaces.* □

4.3. Corollary. *If $\mathscr{H}$ is an infinite-dimensional Hilbert space, then for each $x\in S$, $\mathrm{U}(\mathscr{H})/St(x)Z$ is a weakly minimal $\mathrm{U}(\mathscr{H})$-space.*

Proof. By 1.6 $\mathrm{U}(\mathscr{H})/St(x)Z$ is equivariantly homeomorphic to P. □

Notice that there exist completely regular G-spaces without compact G-extensions (see, for example, [2, Lemma S 5]). A sufficient condition for existence of compact G-extensions is given by Ludesher and de Vries ([17, Theorem 3.3]). A straightforward consequence from [17] is the following.

4.4. Lemma. *Let G be a topological group and H be a closed subgroup of G. Then the left coset G-space G/H has a compact G-extension.* □

4.5. Definition. Let G be a sup-complete topological group (cf. 1.5). Suppose there exists a collection $\mathscr{L}$ of closed subgroups of G so that the following two conditions are fulfilled:

(i) the sets of the type $U\cdot H$, where U is a neighbourhood of the neutral element e of G, and $H\in\mathscr{L}$, form a subbase of neighbourhoods of e in G;

(ii) for every natural n and every subset $\{H_1,\ldots,H_{n-1}\}$ of $\mathscr{L}$ the groups of the type $G'\cap H$, where $G'=\bigcap_{j=1}^{n-1} H_j$, $H\in\mathscr{L}$ and $H\neq H_j\,(j=1,\ldots,n-1)$, are maximal closed subgroups of G', and $G'/(G'\cap H)$ are weakly minimal G'-spaces.

Then G will be called an *M-group*.

Note that the condition (ii) above includes the case $n=1$, so all elements of $\mathscr{L}$ are conjugate maximal closed subgroups of G.

If X is a discrete space, then $F(X)$ is an M-group ([7, (1)]). We shall show below that for every infinite-dimensional Hilbert space $\mathscr{H}$ the quotient group $\mathrm{U}(\mathscr{H})/Z$ is an M-group.

The following proposition generalizes the arguments from the proof of [7, (1)].

4.6. Proposition. *Every M-group is a minimal topological group.*

Proof. We follow the proof of [7, (1)]. Let G be an M-group. Denote by τ the topology of G, and suppose there exists a Hausdorff group topology σ on G strictly weaker than τ.

Denote by $\mathscr{L}$ a collection of closed subgroups of (G, τ) such that (i) and (ii) of 4.5 are fulfilled. We are going to prove by induction on n that the intersection of any n elements of $\mathscr{L}$ is σ-dense in G. We show first that there are no elements of $\mathscr{L}$ which are σ-closed in G. Assume the contrary. Then all elements of $\mathscr{L}$, being conjugate, are σ-closed. Take $H \in \mathscr{L}$. By Lemma 4.4 the G-space $(G/H, \sigma/H)$ has a compact (G, σ)-extension. The latter is at the same time a (G, τ)-extension because $\sigma \subset \tau$. Since id: $(G/H, \tau/H) \to (G/H, \sigma/H)$ is continuous and $(G/H, \tau/H)$ is weakly minimal by (ii), then $\sigma/H = \tau/H$. This holds for every $H \in \mathscr{L}$, and now (i) implies $\sigma = \tau$ which is a contradiction. So every element H of $\mathscr{L}$ is not σ-closed in G. The maximality of H (see (ii) for $n=1$) shows that the σ-closure of H (which is also τ-closed in G) coincides with G, i.e. H is σ-dense in G.

Suppose $n > 1$ and the intersection of every $n-1$ elements of $\mathscr{L}$ is σ-dense in G. Let $H_1, \ldots, H_{n-1}$ be different elements of $\mathscr{L}$. Set $G' = H_1 \cap \ldots \cap H_{n-1}$. Denote by $\mathscr{L}'$ the collection of all groups $G' \cap H$, where $H \in \mathscr{L}$, $H \neq H_j$ $(j = 1, \ldots, n-1)$. If $\sigma|_{G'} = \tau|_{G'}$, then $(G', \sigma|_{G'})$ would be sup-complete and hence closed in G. Thus $\sigma|_{G'} \subsetneq \tau|_{G'}$. Applying the above arguments to G' and $\mathscr{L}'$ instead of G and $\mathscr{L}$, respectively, we see that every element of $\mathscr{L}'$ is σ-dense in G'. On the other hand G' is σ-dense in G by the inductive hypothesis, so we get that the intersection of any n elements of $\mathscr{L}$ is σ-dense in G.

The assumption $\sigma \neq \tau$ shows that $G \neq (e)$ and, as σ is Hausdorff, there is a neighbourhood $V \in \sigma$ of e in G such that $V \cdot V \neq G$. Since $\sigma \subset \tau$, $V \in \tau$, and (i) implies $H_1 \cap \ldots \cap H_n \subset V$ for some elements $H_1, \ldots, H_n$ of $\mathscr{L}$. Now it follows from above that V is σ-dense in G which is in contradiction with $V \cdot V \neq G$. □

We are going to prove the main result in the paper.

4.7. Theorem. *For every real or complex Hilbert space $\mathscr{H}$ the group $\mathrm{U}(\mathscr{H})$ of the unitary operators of $\mathscr{H}$ provided with the strong topology is a totally minimal topological group.*

Proof. The finite-dimensional case is trivial, so we assume $\mathscr{H}$ to be infinite-dimensional.

We are going to prove first that $\mathrm{U}(\mathscr{H})/Z$ is an M-group (4.5). Denote by $\mathscr{L}$ the collection of the subgroups $St(e) \cdot Z$ of $G = \mathrm{U}(\mathscr{H})/Z$, where e runs over the elements of a fixed orthonormal basis E of $\mathscr{H}$. Since $\mathrm{U}(\mathscr{H})$ is sup-complete (1.5) and Z is compact, G is also sup-complete. Clearly the condition (i) of 4.5 is fulfilled (cf. 1.5). It remains to verify (ii).

Fix a natural n and arbitrary $e_1, \ldots, e_{n-1} \in E$. Set $G' = \bigcap_{j=1}^{n-1} St(e_j) \cdot Z$. Denote by $\mathscr{H}_1$ the subspace of $\mathscr{H}$ generated by $e_1, \ldots, e_{n-1}$. There is a natural embedding

$$\varphi\colon \mathrm{U}(\mathscr{H}_1) \times \mathrm{U}(\mathscr{H}_1^\perp) \to \mathrm{U}(\mathscr{H}),$$

defined by $\varphi(T_1, T_2) = T_1 \oplus T_2$, where $T_1 \in \mathbf{U}(\mathscr{H}_1)$, $T_2 \in \mathbf{U}(\mathscr{H}_1^\perp)$, and $T_1 \oplus T_2(x, y) = T_1 x + T_2 y$ for $x \in \mathscr{H}_1$ and $y \in \mathscr{H}_1^\perp$. It is clear that

$$\varphi^{-1}(G') = K \times \mathbf{U}(\mathscr{H}_1^\perp), \tag{5}$$

where K is the group of those $T \in \mathbf{U}(\mathscr{H}_1)$ so that $Te_j \in \mathbf{S}^1 \cdot e_j$ for each $j = 1, \dots, n-1$.

Let e be an arbitrary element of $E \backslash \{e_1, \dots, e_{n-1}\}$. For $H_e = G' \cap St(e) \cdot Z$ we have

$$\varphi^{-1}(H_e) = K \times (G'_e \cdot Z'), \tag{6}$$

where G'_e is the isotropy group of e in $\mathbf{U}(\mathscr{H}_1^\perp)$ and Z' is the center of $\mathbf{U}(\mathscr{H}_1^\perp)$. Since $G'_e \cdot Z'$ is a maximal closed subgroup of $\mathbf{U}(\mathscr{H}_1^\perp)$ (cf. 1.6), then H_e is a maximal closed subgroup of G'. It is clear also that the groups of the type H_e are conjugate in G' when e runs over $E \backslash \{e_1, \dots, e_{n-1}\}$.

As $\mathbf{U}(\mathscr{H}_1^\perp)$ is isomorphic to a quotient group of G' (the isomorphism being induced by φ), it carries a natural structure of a G'-space. So $X = \mathbf{U}(\mathscr{H}_1^\perp)/G'_e \cdot Z'$ also can be naturally considered as a G'-space. By Corollary 4.3 X is a weakly minimal $\mathbf{U}(\mathscr{H}_1^\perp)$-space and therefore it is also a weakly minimal G'-space. It follows from (5) and (6) that the homomorphism φ generates a bijection

$$\psi\colon \mathbf{U}(\mathscr{H}_1^\perp)/G'_e \cdot Z' \to G'/H_e,$$

and a straightforward verification shows that ψ is G'-equivariant. Hence G'/H_e is also a weakly minimal G'-space.

Thus we have shown that $G = \mathbf{U}(\mathscr{H})/Z$ and $\mathscr{L}$ fulfil the conditions (i) and (ii) of 4.5. So G is an M-group, and by Proposition 4.6 it is minimal.

To prove that $\mathbf{U}(\mathscr{H})$ is totally minimal take a proper closed normal subgroup N of $\mathbf{U}(\mathscr{H})$. Then $N \subset Z$ (see 1.6), so there exists a natural exact sequence

$$0 \to Z/N \to \mathbf{U}(\mathscr{H})/N \to \mathbf{U}(\mathscr{H})/Z \to 0.$$

Since Z/N is compact and $\mathbf{U}(\mathscr{H})/Z$ is minimal, by [12, Corollary (7)] $\mathbf{U}(\mathscr{H})/N$ is also minimal. This completes the proof of the theorem. □

4.8. Corollary. *For every real or complex Hilbert space $\mathscr{H}$ the group $\mathbf{U}_1(\mathscr{H})$ of those $T \in \mathbf{U}(\mathscr{H})$ such that $\lambda T - I$ is a compact operator for some scalar λ with $|\lambda| = 1$ is totally minimal with respect to the strong topology.*

Proof. Since $\mathbf{U}_1(\mathscr{H})$ is dense in $\mathbf{U}(\mathscr{H})$, the assertion follows from the above theorem and the total minimality criterion ([10, Theorem 1]). □

Note that the group $\mathbf{U}_0(\mathscr{H})$ of those $T \in \mathbf{U}(\mathscr{H})$ such that $T - I$ is a compact operator is not minimal if $\dim(\mathscr{H}) = \infty$. This follows from Theorem 4.7, the minimality criterion ([4, Proposition 1]), and the fact that $\mathbf{U}_0(\mathscr{H}) \cap Z = \{I\}$.

4.9. Corollary. *Let $\mathscr{H}$ be a Hilbert space, $\mathscr{H} = \bigoplus_\alpha \mathscr{H}_\alpha$ be an orthogonal decomposition of $\mathscr{H}$, and G be the group of those $T \in \mathbf{U}(\mathscr{H})$ so that $T(\mathscr{H}_\alpha) = \mathscr{H}_\alpha$ for every α. Then G is totally minimal with respect to the strong topology.*

Proof. It is easy to see that G is topologically isomorphic to $\prod_{\alpha} \mathrm{U}(\mathscr{H}_\alpha)$. The latter group is totally minimal by Theorem 4.7 above and [11, the remark after Theorem 1]. $\square$

All known examples of minimal groups are either locally precompact (i.e. they have locally compact completions) or they are dense subgroups of groups of the type $F(X)$, X being a discrete set (see the introduction). So it is easy to see that every such a group G can be embedded in $\mathrm{U}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$ with $\dim(\mathscr{H}) = w(G)$ (the *weight* of G). Indeed, if G is locally precompact, and $\hat{G}$ is the completion of G, the regular representation $\hat{G} \to \mathrm{U}(L^2(\hat{G}))$ (cf. [16, p. 36]) induces a continuous monomorphism $G \to \mathrm{U}(L^2(\hat{G}))$ which is an embedding by the minimality of G. Clearly $w(\hat{G}) = w(G)$, and by [24, 2.3.15] $\dim(L^2(\hat{G})) = w(\hat{G})$ (for $\hat{G}$ compact the latter is proved also in [15, 24.15 and 28.2]). If $G \subset F(X)$, consider a Hilbert space $\mathscr{H}$ with an orthonormal basis E with $\mathrm{card}(E) = \mathrm{card}(X)$. Let $p: X \to E$ be a bijection. Define $T: F(X) \to \mathrm{U}(\mathscr{H})$ by $T_f(e) = pfp^{-1}(e)$ for every $e \in E$. It is easy to see that T is an embedding.

Question. Is it true that for every minimal group G there is a Hilbert space $\mathscr{H}$ (with $\dim(\mathscr{H}) = w(G)$) such that G is topologically isomorphic to a subgroup of $\mathrm{U}(\mathscr{H})$?

References

1. Adian, S.: Classifications of periodic words and their application in group theory. In: Burnside Groups, Proc. Bielefeld, Germany 1977. Workshop, edited by J.L. Mennicke, pp. 1–40. Lecture Notes in Math., vol. 806. Berlin-Heidelberg-New York: Springer 1980
2. Antonjan, S., Smirnov, Yu.: Universal objects and bicompact extensions for topological groups of transformations. Dokl. Acad. Nauk SSSR **257**, 521–526 (1981) [Russian]. Engl. Transl.: Soviet Math. Dokl. **23**, 279–284 (1981)
3. Berberian, S.: Lectures in functional analysis and operator theory. Graduate texts in mathematics, vol. 15. New York-Heidelberg-Berlin: Springer 1974
4. Banaschewski, B.: Minimal topological algebras. Math. Ann. **211**, 107–114 (1974)
5. Bourbaki, N.: Topologie Generale. Paris: Hermann 1975
6. Comfort, W., Grant, L.: Cardinal invariants, pseudocompactness and minimality: some recent advances in the topological theory of topological groups. Topology Proc. **6**, 227–265 (1981)
7. Dierolf, S., Schwanengel, U.: Un example d'un groupe topologique q-minimaux mais non precompact. Bull. Sci. Math. (2) **101**, 265–269 (1977)
8. Dierolf, S., Schwanengel, U.: Examples of locally compact non-compact minimal topological groups. Pacific J. Math. **82**, 349–355 (1979)
9. Dikranjan, D.: Minimal topologies on Abelian groups. Istituto di Algebra e Geometria de l'Università di Padova, 1983 (Preprint)
10. Dikranjan, D., Prodanov, Iv.: Totally minimal topological groups. Annuaire Univ. Sofia Fac. Math. Méc. **69**, 5–11 (1974/75)
11. Eberhardt, V., Dierolf, S.: Produits d'espaces vectoriels topologique q-minimaux. C.R. Acad. Sci. Paris **288**, 275–277 (1979)
12. Eberhardt, V., Dierolf, S., Schwanengel, U.: On the product of two (totally) minimal topological groups and the three-space-problem. Math. Ann. **251**, 123–128 (1980)
13. Harpe, P. de la: Classical Banach-Lie algebras and Banach-Lie groups of operators in Hilbert space. Lecture Notes in Math., vol. 285. Berlin-Heidelberg-New York: Springer 1972
14. Hesse, G.: Zur Topologisierbarkeit von Gruppen. Dissertation, Hannover 1979
15. Hewitt, E., Ross, K.: Abstract harmonic analysis I, Die Grundlehren der Mathematischen Wissenschaften, vol. 115. Berlin-Göttingen-Heidelberg: Springer 1963

16. Heyer, H.: Dualität lokalkompakter Gruppen. Lecture Notes in Math., vol. 150. Berlin-Heidelberg-New York: Springer 1970
17. Ludescher, H., Vries, J. de: A sufficient condition for the existence of a G-compactification. Indag. Math. (3) **42**, 263-268 (1980)
18. Prodanov, Iv.: Compact representations of continuous algebraic structures. General Topology and its Relations to Modern Analysis and Algebra II. Proc. of the Second Prague Topology Symposium, pp. 290-294 (1966) [Russian]. Prague: Academia 1967
19. Prodanov, Iv.: Some minimal group topologies are precompact. Math. Ann. **227**, 117-125 (1977)
20. Prodanov, Iv., Stojanov, L.: Every minimal Abelian group is precompact. C.R. Acad. Bulgare Sci. **37**, 23-26 (1984)
21. Schwanengel, U.: Minimale topologische Gruppen, Dissertation, München 1978
22. Shelah, S.: On a problem of Kurosh, Jonsson groups and applications. In: Word Problems II (Adian, S.I., Boone, W.W., Higman, G., eds.), pp. 373-394. Amsterdam-London: North Holland 1980
23. Takesaki, M.: Theory of operator algebras I. Berlin-Heidelberg-New York: Springer 1979
24. Vries, J. de: Topological transformation groups I. Amsterdam: Mathematical Centre 1975
25. Vries, J. de: Equivariant embeddings of G-spaces. General Topology and its Relations to Modern Analysis and Algebra IV. Proc. of the Fourth Prague Top. Symposium, pp. 485-493 (1976)

Received November 3, 1983; received in final form March 19, 1984

Math. Z. 187, 285–287 (1984)

Mathematische Zeitschrift

On the Existence of an Infinite Family of Simple 5-Designs

Tran van Trung

Institut für Mathematik, Universität Heidelberg, D-6900 Heidelberg, Federal Republic of Germany

Alltop (1972) [1] has constructed for the first time an infinite class of simple 5-$(2^n+2,\ 2^{n-1}+1,\ (2^{n-1}-3)\ (2^{n-2}-1))$ designs for $n\geqq 4$. That was the only infinite class of simple 5-designs known untill now (see [2, 3, 5, 6]). In the present paper, using the 5-designs of Alltop, we give the existence of a new infinite family of simple 5-$(2^n+3,\ 2^{n-1}+1,\ (2^n-2)\ (2^{n-1}-3)\ (2^{n-2}-1))$ designs for $n\geqq 5$.

A t-design, denoted by t-(v,k,λ), is a pair $(X,\mathscr{B})$ where $\mathscr{B}$ is a collection of k-subsets (called blocks) of a v-set X such that every t-subset of X is in exactly λ blocks of $\mathscr{B}$. A t-design is called simple if no blocks are repeated, and trivial if every k-subset of X is a block and occurs in exactly m times in $\mathscr{B}$. Throughout the paper all t-designs are assumed to be nontrivial and simple.

Let $(X,\mathscr{B})$ be a t-(v,k,λ) design, let I be an i-subset of X and let λ_i be the number of blocks in $\mathscr{B}$ which contain I. It is well known that

$$\lambda_i=\lambda\binom{v-i}{t-i}\Big/\binom{k-i}{t-i},$$

and hence the congruences

$$\lambda\binom{v-i}{t-i}\equiv 0\left(\operatorname{mod}\binom{k-i}{t-i}\right)\qquad\text{for } i=0,1,2,\dots,t,$$

are necessary for the existence of $(X,\mathscr{B})$. Note that $\lambda=\lambda_t$ and λ_0 is the total number of blocks.

Let X be a set. Then $|X|$ denotes the cardinality of X and S_X the symmetric group $S_{|X|}$ on X.

Let $(X,\mathscr{B})$ be a t-(v,k,λ) design and $(X',\mathscr{B}')$ a t-(v',k,λ') design. Then $(X,\mathscr{B})$ and $(X',\mathscr{B}')$ are called disjoint if $\mathscr{B}\cap\mathscr{B}'=\varnothing$. Of course, if $|X\cap X'|<k$, then $(X,\mathscr{B})$ and $(X',\mathscr{B}')$ are disjoint. Let $(X_i,\mathscr{B}_i)$ be a t-$(v_i,k,\lambda^{(i)})$ design for $i=1,\dots,m$. We say that $(X_i,\mathscr{B}_i)$, $i=1,\dots,m$, are mutually disjoint if $(X_h,\mathscr{B}_h)$ and $(X_j,\mathscr{B}_j)$ are disjoint for $h,j=1,\dots,m$, $h\neq j$.

If a t-(v,k,λ) design exists, then the following theorem proves the existence of a t-$(v+1,k,(v+1-t)\cdot\lambda)$ design whenever $v\cdot\lambda_0(\lambda_0-\lambda_1)<\binom{v}{k}$.

Theorem A. *If there exists a t-(v,k,λ) design with $v\cdot\lambda_0(\lambda_0-\lambda_1)<\binom{v}{k}$, then there exists a t-$(v+1,k,(v+1-t)\lambda)$ design.*

Proof. Let $X=\{x_1,x_2,\ldots,x_{(v+1)}\}$ be a $(v+1)$-set. We define $X_i=X-\{x_i\}$ for $i=1,2,\ldots,(v+1)$. Let $(X_i,\mathscr{B}_i)$ be a t-(v,k,λ) design constructed on the point set X_i. In the following we shall investigate the condition for which we can construct $(v+1)$ mutually disjoint t-designs $(X_i,\mathscr{B}_i)$, $i=1,2,\ldots,(v+1)$. Let $(X_1,\mathscr{B}_1)$ be a t-(v,k,λ) design on X_1 and $(X_2,\mathscr{B}_2^*)$ a t-(v,k,λ) design on X_2. We find an upper bound for the number of permutations $g\in S_{X_2}$ for which there is a block $B^*\in\mathscr{B}_2^*$ and a block $B\in\mathscr{B}_1$ such that $(B^*)^g=B$, if this number is less than $|S_{X_2}|=v!$, then there exists a permutation $h\in S_{X_2}$ such that $(X_2,(\mathscr{B}_2^*)^h)$ and $(X_1,\mathscr{B}_1)$ are disjoint. First note that a block $B\in\mathscr{B}_1$ containing x_2 can not be image of a block $B^*\in\mathscr{B}_2^*$ under a permutation $g\in S_{X_2}$. Therefore there are $(\lambda_0-\lambda_1)$ blocks of $\mathscr{B}_1$ which are candidates for images of blocks of $\mathscr{B}_2^*$. Given a $B^*\in\mathscr{B}_2^*$ and a $B\in\mathscr{B}_1$ with $x_2\notin B$, then there are $k!\,(v-k)!$ permutations $g\in S_{X_2}$ such that $(B^*)^g=B$. As $|\mathscr{B}_2^*|=\lambda_0$ so we get $\lambda_0(\lambda_0-\lambda_1)k!(v-k)!$ permutations $g\in S_{X_2}$ which send at least a block of $\mathscr{B}_2^*$ onto a block of $\mathscr{B}_1$. Now if $v!>\lambda_0(\lambda_0-\lambda_1)k!(v-k)!$, where $v!=|S_{X_2}|$, then there exists a permutation $h\in S_{X_2}$ such that $(\mathscr{B}_2^*)^h\cap\mathscr{B}_1=\varnothing$. We write $(X_2,\mathscr{B}_2):=(X_2,(\mathscr{B}_2^*)^h)$. Then $(X_1,\mathscr{B}_1)$ and $(X_2,\mathscr{B}_2)$ are disjoint. Let $(X_3,\mathscr{B}_3^*)$ be a t-(v,k,λ) design on the point set X_3. Similarly we find the number of $g\in S_{X_3}$ for which there is a block $B^*\in\mathscr{B}_3^*$ such that $(B^*)^g\in\mathscr{B}_1\cup\mathscr{B}_2$. Since a block $B\in\mathscr{B}_1\cup\mathscr{B}_2$ containing x_3 can not be image of a block $B^*\in\mathscr{B}_3^*$ under a $g\in S_{X_3}$, hence as before there are $2\cdot\lambda_0(\lambda_0-\lambda_1)k!(v-k)!$ permutations $g\in S_{X_3}$ which send at least a block of $\mathscr{B}_3^*$ onto a block of $\mathscr{B}_1\cup\mathscr{B}_2$. If $v!>2\cdot\lambda_0(\lambda_0-\lambda_1)k!(v-k)!$, where $v!=|S_{X_3}|$, then we have a permutation $h\in S_{X_3}$ such that $(\mathscr{B}_3^*)^h\cap(\mathscr{B}_1\cup\mathscr{B}_2)=\varnothing$. Set $(X_3,\mathscr{B}_3):=(X_3,(\mathscr{B}_3^*)^h)$, then $(X_1,\mathscr{B}_1)$, $(X_2,\mathscr{B}_2)$ and $(X_3,\mathscr{B}_3)$ are mutually disjoint. Now it is clear that if we continue this procedure for the remaining point sets $X_4,X_5,\ldots,X_{(v+1)}$, then we can construct $(v+1)$ mutually disjoint t-designs $(X_i,\mathscr{B}_i)$, $i=1,2,\ldots,(v+1)$, if the condition

$$v!>v\cdot\lambda_0(\lambda_0-\lambda_1)k!(v-k)!$$

or equivalently

$$\binom{v}{k}>v\cdot\lambda_0(\lambda_0-\lambda_1)$$

holds. Now define

$$(X,\mathscr{B}):=\dot{\bigcup_{i=1}^{v+1}}(X_i,\mathscr{B}_i).$$

Let $\{x_{j_1},x_{j_2},\ldots,x_{j_t}\}$ be a set of t points of X, then $x_{j_1},x_{j_2},\ldots,x_{j_t}$ occur together in λ blocks of $(X_i,\mathscr{B}_i)$, where $i\neq j_1,j_2,\ldots,j_t$. Hence $x_{j_1},x_{j_2},\ldots,x_{j_t}$ appear together in $(v+1-t)\cdot\lambda$ blocks in $(X,\mathscr{B})$. In other words, $(X,\mathscr{B})$ is a t-$(v+1,k,(v+1-t)\cdot\lambda)$ design. Theorem A is proved.

The next theorem is due to Alltop (1972) [1].

Theorem (Alltop). *There exists a*

$$5\text{-}(2^n+2,2^{n-1}+1,(2^{n-1}-3)(2^{n-2}-1))$$

design for every $n\geqq 4$.

Now applying Theorem A to the $5\text{-}(2^n+2, 2^{n-1}+1, (2^{n-1}-3)(2^{n-2}-1))$ designs of Alltop we see that the condition $v \cdot \lambda_0(\lambda_0-\lambda_1) < \binom{v}{k}$ is satisfied for $n \geqq 5$. Hence we get the following theorem.

Theorem B. *There exists a*

$$5\text{-}(2^n+3, 2^{n-1}+1, (2^n-2)(2^{n-1}-3)(2^{n-2}-1))$$

design for every $n \geqq 5$.

It is worth mentioning here that there are few 5-designs known with an odd number of points (see [2, 5, 6]). So the Theorem B gives the existence of infinitely many 5-designs with an odd v.

Remark. The application of Theorem A to the known infinite classes of 4-designs [2, 4] will provide other new 4-designs. For instance, using Theorem A for the

$$4\text{-}(2^n+1, 2^{n-1}-1, (2^{n-1}-3)(2^{n-2}-1)(2^{n-1}-4))$$

designs, $n \geqq 4$, constructed by Driessen (1978) [2], we get a new family of $4\text{-}(2^n+2,\ 2^{n-1}-1,\ (2^n-2)\ (2^{n-1}-3)\ (2^{n-2}-1)\ (2^{n-1}-4))$ design for $n \geqq 6$. For the other 2 classes of 4-designs constructed by Hubaut (1974) [4], namely

$$4\text{-}\left(2^n+1, 2^m, (2^m-3)\prod_{i=2}^{m-1}\frac{2^{n-i}-1}{2^{m-i}-1}\right), \quad 2<m<n$$

and

$$4\text{-}\left(2^n+1, 2^m+1, (2^m+1)\prod_{i=2}^{m-1}\frac{2^{n-i}-1}{2^{m-i}-1}\right), \quad 2 \leqq m < n, \quad m \nmid n,$$

designs, if n and m are great enough, we can see that Theorem A yields new families of 4-designs.

References

1. Alltop, W.O.: An infinite class of 5-designs. J. Comb. Theory (A) **12**, 390–395 (1972)
2. Driessen, L.H.M.E.: t-designs, $t \geqq 3$. Tech. Report, Department of Mathematics, Eindhoven University of Technology, Holland (1978)
3. Hedayat, A., Kageyama, S.: The family of t-designs – Part I. J. of Statistical Planning and Inf. **4**, 173–212 (1980)
4. Hubaut, X.: Two new families of 4-designs. Discrete Math. **9**, 247–249 (1974)
5. Kramer, E.S., Leavitt, D.W., Magliveras, S.S.: Construction procedures for t-designs and the existence of new simple 6-designs (to appear)
6. Magliveras, S.S., Leavitt, D.W.: Simple 6-(33, 8, 36) designs from $P\Gamma L_2(32)$ (to appear)

Received March 24, 1984

Math. Z. 187, 289-304 (1984)

Mathematische Zeitschrift

Compatibility Operators for Degenerate Elliptic Equations on the Ball and Heisenberg Group

C. Robin Graham*

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
Present address: Department of Mathematics, University of Washington, Seattle, WA 98195, USA

Introduction

On the unit ball $\mathbb{B}^{n+1} \subset \mathbb{C}^{n+1}$ consider the second order degenerate elliptic operators

$$\Delta^{\alpha,\beta} = (1-|w|^2)\left[\sum_{i,j=1}^{n+1} (\delta_{ij} - w_i \bar{w}_j) \frac{\partial^2}{\partial w_i \partial \bar{w}_j} + \alpha \sum_{i=1}^{n+1} w_i \frac{\partial}{\partial w_i} + \beta \sum_{i=1}^{n+1} \bar{w}_i \frac{\partial}{\partial \bar{w}_i} - \alpha\beta\right], \quad \alpha, \beta \in \mathbb{C}.$$

These operators are variants introduced by Geller [2] of the Laplace-Beltrami operator for the Bergman metric, which is $4\Delta^{0,0}$. Because of the degeneracy of $\Delta^{\alpha,\beta}$ at the boundary it is possible for solutions to $\Delta^{\alpha,\beta} u = 0$ to have boundary values which are infinitely differentiable yet not to be smooth up to the boundary themselves. This is most striking in case $k = n+1+\alpha+\beta \in \mathbb{N} = \{1,2,3,\ldots\}$. For such α, β there is a local differential condition on boundary data in order that there be a solution u to $\Delta^{\alpha,\beta} u = 0$ which is smooth up to the boundary and has the prescribed boundary values. It is the point of this paper to derive this condition and study the compatibility operators so determined.

The situation on the Siegel domain $\mathbb{D} = \{(z', z_{n+1}) \in \mathbb{C}^{n+1}\colon \operatorname{Im} z_{n+1} > |z'|^2\}$ is analogous yet simpler. In order to define the analogues of $\Delta^{\alpha,\beta}$, set $t = \operatorname{Re} z_{n+1}$, $\rho = \operatorname{Im} z_{n+1} - |z'|^2$; let $N = \frac{\partial}{\partial \rho}$, $T = \frac{\partial}{\partial t}$, and let $\mathscr{L}_0$ be the sub-laplacian of Folland-Stein [1] on the Heisenberg Group $\mathbb{H}^n$ viewed as acting on functions on $\mathbb{D}$ via the identification $\mathbb{D} = \mathbb{H}^n \times \mathbb{R}_+$. The analogues are then

$$\Delta_{\lambda,\mu} = \rho[\rho(N^2 + T^2) - \mathscr{L}_0 + (1-\lambda)N - i\mu T], \quad \lambda, \mu \in \mathbb{C}.$$

* Most of this work was carried out while the author held a NSF Postdoctoral Research Fellowship at the Courant Institute of Mathematical Sciences, New York University

Throughout this paper upper placement of the parameters will be used to denote operators on $\mathbb{B}^{n+1}$ or $\mathbb{S}^{2n+1}=b\mathbb{B}^{n+1}$ and lower placement for their analogues on $\mathbb{D}$ or $\mathbb{H}^n$. Various aspects of the Dirichlet problem for these operators were studied in [4], where in particular the existence of the aforementioned compatibility operator was established, and the operator was explicitly identified in the case $\lambda=n+1$, $\mu=0$ corresponding to the Bergman Laplacian. In §1 of this paper, it is shown that the same method can be used to identify the compatibility operator $\mathscr{C}_{k,\mu}$ for $\Delta_{k,\mu}$ in the general case $k\in\{1,2,\dots\}$, $\mu\in\mathbb{C}$. The resulting operator is a product of various of the Folland-Stein operators $\mathscr{L}_\alpha=\mathscr{L}_0+i\alpha T$, namely $\mathscr{C}_{k,\mu}=\prod_{l=1}^{k}\mathscr{L}_{\mu+k+1-2l}$.

Interestingly, products of $\mathscr{L}_\alpha$ operators of this type, i.e. in which the index increases by 2 from factor to factor, seem to be special amongst all possible products of $\mathscr{L}_\alpha$ operators. For example, in §2 it is shown that the operators $\mathscr{C}_{k,\mu}$ satisfy a simple transformation law under the Cayley transform, in which they correspond to similar products on the sphere $\mathbb{S}^{2n+1}$ of the analogues of the $\mathscr{L}_\alpha$ operators discovered by Geller [3]. The particular case $k=1$, in which $\mathscr{C}_{1,\mu}=\mathscr{L}_\mu$, thus establishes a direct connection between $\mathscr{L}_\mu$ and Geller's analogue, which itself seems to be new. The transformation law for $\mathscr{C}_{k,\mu}$ is used here to derive the explicit form of the compatibility operator for $\Delta^{\alpha,\beta}$ on the ball. It would be interesting to give a direct derivation of the resulting formula for this compatibility operator; such a derivation reduces to identifying a polynomial of two real variables determined by a particular explicit recursion relation. This reduction is carried out for $\alpha=\beta=0$ at the end of §2.

Finally, in §3 explicit fundamental solutions for the operators $\mathscr{C}_{k,\mu}$ are computed. This computation provides more evidence of the special nature of products of this type, as the fundamental solutions have a very simple form, reducing to the Folland-Stein [1] and Greiner-Kohn-Stein [5] formulae when $k=1$. There do not seem to be analogues of either the simple fundamental solutions or the transformation law under the Cayley transform for more general products of $\mathscr{L}_\alpha$ operators; for example, for $\mathscr{L}_{\alpha_1}\mathscr{L}_{\alpha_2}$ with α_1,α_2 unrelated.

1. Compatibility Operators on the Siegel Domain

We begin by recalling the degenerate elliptic operators studied in [4], which should be consulted for the details of the material sketched here. These operators are defined on the Siegel domain

$$\mathbb{D}=\{z=(z',z_{n+1}):\ \rho(z)=\operatorname{Im} z_{n+1}-|z'|^2>0\}\subset\mathbb{C}^{n+1}.$$

Via the coordinates (ζ,t,ρ), where $\zeta=z'\in\mathbb{C}^n$, $t=\operatorname{Re} z_{n+1}$, $\rho=\rho(z)$, $\mathbb{D}$ can be identified with $\mathbb{H}^n\times\mathbb{R}_+$, where $\mathbb{H}^n=\{u=(\zeta,t)\in\mathbb{C}^n\times\mathbb{R}\}$ is the Heisenberg group. Setting $\zeta_j=x_j+iy_j$, $1\leqq j\leqq n$, the standard basis for the left invariant vector fields on $\mathbb{H}^n$ is

$$X_j=\frac{\partial}{\partial x_j}+2y_j\frac{\partial}{\partial t},\qquad Y_j=\frac{\partial}{\partial y_j}-2x_j\frac{\partial}{\partial t},\qquad 1\leqq j\leqq n,$$

and $T=\dfrac{\partial}{\partial t}$; and $N=\dfrac{\partial}{\partial \rho}$ completes this to a basis for $T\mathbb{D}$. Set

$$Z_j=\tfrac{1}{2}(X_j-iY_j)=\frac{\partial}{\partial \zeta_j}+i\bar{\zeta}_j\frac{\partial}{\partial t}, \qquad 1\leqq j\leqq n,$$

and $Z_{n+1}=\frac{1}{2}(T-iN)$, and recall the Folland-Stein operators

$$\mathscr{L}_\mu=-\tfrac{1}{4}\sum_{j=1}^{n}(X_j^2+Y_j^2)+i\mu T, \qquad \mu\in\mathbb{C}.$$

The Bergman metric in $\mathbb{D}$ is the Kähler metric with Kähler form $\dfrac{i}{2}\partial\bar{\partial}\log\dfrac{1}{\rho}$. A straightforward computation (see [4], p. 446–7) shows that the Laplace-Beltrami operator in this metric is given by the formula

$$\Delta_B=4\rho[\rho(N^2+T^2)-\mathscr{L}_0-nN]. \tag{1.1}$$

We are interested in the two parameter family of operators

$$\Delta_{\lambda,\mu}=\rho[\rho(N^2+T^2)-\mathscr{L}_0+(1-\lambda)N-i\mu T], \qquad \lambda,\mu\in\mathbb{C}, \tag{1.2}$$

all of which are left invariant under Heisenberg translation and are homogeneous of degree 0 with respect to the nonisotropic dilations on $\mathbb{D}=\mathbb{H}^n\times\mathbb{R}_+$. Note that $\Delta_{n+1,0}=\frac{1}{4}\Delta_B$ and more generally $\Delta_{\lambda,\mu}=\frac{1}{4}\Delta_B+(n+1-\lambda)\rho N-i\mu\rho T$. The Dirichlet problem for these operators was studied in [4], where it was established that so long as $\operatorname{Re}\lambda>0$ and $\frac{1}{2}(n+1+\lambda\pm\mu)\notin\{0,-1,-2,\dots\}$, the problem $\Delta_{\lambda,\mu}u=0$, $u|_{b\mathbb{D}}=f$ can be solved when f is a continuous function with compact support on $\mathbb{H}^n$. Also if f is infinitely differentiable, then while u need not be smooth up to the boundary, u does have an asymptotic expansion at the boundary of the form

$$u=\begin{cases} u_1+u_2\rho^\lambda & \text{if } \lambda\notin\{1,2,3,\dots\}\\ u_1+u_2\rho^k\log\rho & \text{if } \lambda=k\in\{1,2,3,\dots\}\end{cases}$$

where $u_1,u_2\in C^\infty(\bar{\mathbb{D}})$. In the latter case, i.e. $\lambda=k\in\{1,2,3,\dots\}$, it is possible to characterize the functions f so that u is actually smooth up to the boundary: f must satisfy a "compatibility condition" on $b\mathbb{D}=\mathbb{H}^n$, a differential equation of order $2k$. The explicit form of this equation was derived in [4] only in the case $\lambda=n+1$, $\mu=0$, and although the derivation in general proceeds similarly to the special case, we include it for completeness.

The result is local, so suppose that $\Omega\subset\mathbb{C}^{n+1}$ is open with $\Omega\cap b\mathbb{D}\neq\emptyset$, and let $u\in C^\infty(\Omega\cap\bar{\mathbb{D}})$ satisfy $\Delta_{k,\mu}u=0$ in $\Omega\cap\mathbb{D}$, some $k\in\mathbb{N}=\{1,2,3,\dots\}$. Restricting the equation $\rho^{-1}\Delta_{k,\mu}u=0$ to $b\mathbb{D}=\{\rho=0\}$ results in the equation $(1-k)Nu=\mathscr{L}_\mu u$ on $b\mathbb{D}$. Since $\mathscr{L}_\mu$ is a tangential operator, for $k>1$ this equation determines Nu locally in terms of the boundary data; while if $k=1$ it is a compatibility condition on the data imposed by the existence of a smooth solution to the problem. In general, the first $k-1$ normal derivatives are determined recursively at the boundary and the k^{th} equation becomes a compatibility condition. To see this, note that $\mathscr{L}_0$, N and T commute and apply N^l

to the equation $\rho^{-1}\Delta_{k,\mu}u=0$ to obtain

$$[\rho(N^{l+2}+T^2N^l)-\mathscr{L}_\mu N^l+(l+1-k)N^{l+1}+lT^2N^{l-1}]u=0,$$

so $(k-l-1)N^{l+1}u=-\mathscr{L}_\mu N^l u+lT^2N^{l-1}u$ at $b\mathbb{D}$. Thus if we set

$$Q_l=(-1)^l(k-1)(k-2)\dots(k-l)N^l \quad \text{for } 1\leqq l\leqq k-1,$$

it follows that as differential operators on $\mathbb{H}^n$, Q_l are determined by the recursion relation $Q_0u=u$, $Q_1u=\mathscr{L}_\mu u$, and $Q_{l+1}=\mathscr{L}_\mu Q_l+l(k-l)T^2Q_{l-1}$, $1\leqq l\leqq k-2$. Also if Q_k is defined by this relation, then $Q_ku=0$ is the compatibility condition on the boundary data. It is possible to solve for Q_k: by induction one sees that Q_l is a homogeneous polynomial of degree l in the commuting variables $\mathscr{L}_\mu$, T, so that the recursion relation for Q_l reduces to the recursion for the polynomials q_l of one variable defined by $q_0=1$, $q_1=x$, $q_{l+1}=xq_l+l(k-l)q_{l-1}$, $l\geqq 1$, and Q_l is recovered from q_l by $Q_l=T^lq_l(L_\mu T^{-1})$. These polynomials are essentially the classical orthogonal polynomials known as Krawtchouk's polynomials, and a formula for q_k can be derived by introducing the generating function

$$F(x,y)=\sum_{l=0}^{\infty} q_l(x)\frac{y^l}{l!}.$$

The recursion relation for q_l implies that $(1+y^2)\dfrac{dF}{dy}=(x+(k-1)y)F$ and $F(x,0)=1$. Separating variables and integrating by partial fractions, the solution

$$F(x,y)=(1+iy)^{\frac{1}{2}(k-1-ix)}(1-iy)^{\frac{1}{2}(k-1+ix)}$$

is obtained. Note that if $x_0\in\{i(k-1), i(k-3), \dots, -i(k-3), -i(k-1)\}$, $F(x_0,y)$ is a polynomial of degree $k-1$, so $q_k(x_0)=0$. But q_k is a monic polynomial of degree k, so these are exactly its zeros and the identity

$$q_k=\prod_{l=1}^{k}(x+i(k+1-2l))$$

results. Hence

$$Q_k=\prod_{l=1}^{k}(\mathscr{L}_\mu+i(k+1-2l)T)=\prod_{l=1}^{k}\mathscr{L}_{\mu+k+1-2l}\equiv\mathscr{C}_{k,\mu}.$$

This establishes the necessity of the condition $\mathscr{C}_{k,\mu}f=0$ in order that a solution u to $\Delta_{k,\mu}u=0$ with boundary values f be smooth up to a part of the boundary, and Theorem 10.2 of [4] provides the sufficiency. Thus we have

(1.3) **Theorem.** *Let $\Omega^{\text{open}}\subset\mathbb{C}^{n+1}$ with $\Omega\cap b\mathbb{D}\neq\emptyset$, let $f\in C^\infty(\Omega\cap b\mathbb{D})$, and $k\in\mathbb{N}$, $\mu\in\mathbb{C}$.*

a) *If there is a $u\in C^\infty(\Omega\cap\overline{\mathbb{D}})$ solving $\Delta_{k,\mu}u=0$ with $u|_{\Omega\cap b\mathbb{D}}=f$, then $\mathscr{C}_{k,\mu}f=0$.*

b) *Let $\frac{1}{2}(n+1+k\pm\mu)\notin\{0,-1,-2,\dots\}$ and suppose that $u\in C(\Omega\cap\overline{\mathbb{D}})$ is any solution of $\Delta_{k,\mu}u=0$ with $u|_{b\mathbb{D}}=f$, where $\mathscr{C}_{k,\mu}f=0$. Then $u\in C^\infty(\Omega\cap\overline{\mathbb{D}})$.*

2. Compatibility Operators on the Ball

There are analogues of the operators $\Delta_{\lambda,\mu}$ on the ball $\mathbb{B}^{n+1}=\{w\in\mathbb{C}^{n+1}: |w|<1\}$ which, along with $\Delta_{\lambda,\mu}$, were first studied by Geller [2]. On the ball they are

$$\Delta^{\alpha,\beta}=(1-|w|^2)\left[\sum_{i,j=1}^{n+1}(\delta_{ij}-w_i\bar{w}_j)\frac{\partial^2}{\partial w_i\partial\bar{w}_j}+\alpha\sum_{i=1}^{n+1}w_i\frac{\partial}{\partial w_i}\right.$$
$$\left.+\beta\sum_{i=1}^{n+1}\bar{w}_i\frac{\partial}{\partial\bar{w}_i}-\alpha\beta\right],$$

$\alpha,\beta\in\mathbb{C}$. We are interested first in investigating the existence and form of compatibility operators for $\Delta^{\alpha,\beta}$. There are two ways that one could conceive of proceeding: one is to write $\Delta^{\alpha,\beta}$ in polar coordinates and successively differentiate and restrict to $b\mathbb{B}^{n+1}=\mathbb{S}^{2n+1}$ so as to recursively determine normal derivatives of a smooth solution and eventually obtain the compatibility operator, exactly as on $\mathbb{D}$. The other is to establish a relation between $\Delta^{\alpha,\beta}$ on $\mathbb{B}$ and $\Delta_{\lambda,\mu}$ on $\mathbb{D}$ and deduce the compatibility operator for $\Delta^{\alpha,\beta}$ from that for $\Delta_{\lambda,\mu}$. I tried the first procedure but was unable to solve the resulting recursion relation: it is more complicated than that on $\mathbb{D}$, reducing this time to a recursion for polynomials of two variables instead of one, the recursion involving all previously determined polynomials rather than just the last two. The second method, however, works. Thus one knows the solution of the nasty recursion on $\mathbb{B}$, but is unable to derive it directly. This matter is discussed in more detail at the end of this section.

We proceed with the derivation of the compatibility condition for $\Delta^{\alpha,\beta}$. Recall that $\mathbb{B}$ and $\mathbb{D}$ are biholomorphic via the Cayley transform $C:\mathbb{B}\to\mathbb{D}$ given by

$$C(w)=C(w',w_{n+1})=\left(\frac{w'}{1+w_{n+1}},\ i\frac{1-w_{n+1}}{1+w_{n+1}}\right).$$

Since a biholomorphism is an isometry of the Bergman metrics, it follows that $\Delta^{0,0}C^*=C^*\Delta_{n+1,0}$, where C^* denotes the operation of pulling back a function: $C^*f=f\circ C$ for $f\in C^\infty(\mathbb{D})$. Geller [2] discovered a generalization of this transformation law to the other operators in our families; in order to write it down define the multiplication operator $M^{\alpha,\beta}$ on functions on $\mathbb{B}^{n+1}$ by

$$M^{\alpha,\beta}f(w)=(1+\bar{w}_{n+1})^\alpha(1+w_{n+1})^\beta f(w).$$

Then Geller's result is:

(2.1) Let (λ,μ) and (α,β) be related by $\alpha=\frac{1}{2}(\lambda-n-1+\mu)$, $\beta=\frac{1}{2}(\lambda-n-1-\mu)$; equivalently by $\lambda=\alpha+\beta+n+1$, $\mu=\alpha-\beta$. Then $M^{\alpha,\beta}C^*\Delta_{\lambda,\mu}=\Delta^{\alpha,\beta}M^{\alpha,\beta}C^*$.

I have no further insight into the proof of this fact than Geller's admonition to perform the "laborious routine computation" leading to it.

This result immediately gives us compatibility operators on $\mathbb{B}$. In fact, if $\Delta^{\alpha,\beta}u=0$ where $\alpha+\beta+n+1=k\in\mathbb{N}$ and if u is smooth up to a portion of $b\mathbb{B}$

which does not include the point $w=(0,\dots,0,-1)$, then setting $\mu=\alpha-\beta$, it follows that $\Delta_{k,\mu}C^{-1*}M^{-\alpha,-\beta}u=0$ and $C^{-1*}M^{-\alpha,-\beta}u$ is smooth up to the corresponding part of $b\mathbb{D}$, so $\mathscr{C}_{k,\mu}C^{-1*}M^{-\alpha,-\beta}u=0$ is the compatibility condition on u. So the problem is to write this explicitly as a differential operator on $b\mathbb{B}$. The resulting operator can be expressed in terms of the analogues of the $\mathscr{L}_\mu$ operators on the sphere, due also to Geller [3]. We begin by reviewing these operators.

On the sphere $\mathbb{S}^{2n+1}$, $\mathscr{L}^\mu$ is defined as follows: first define the vector fields

$$M_{jk}=\bar{w}_j\frac{\partial}{\partial w_k}-\bar{w}_k\frac{\partial}{\partial w_j}\quad\text{and}\quad T=\frac{i}{2}\sum_{i=1}^{n+1}\left(w_i\frac{\partial}{\partial w_i}-\bar{w}_i\frac{\partial}{\partial \bar{w}_i}\right).$$

Then $\mathscr{L}^\mu=-\frac{1}{2}\sum_{j<k}(M_{jk}\overline{M}_{jk}+\overline{M}_{jk}M_{jk})+i\mu T+\frac{1}{4}(n-\mu)(n+\mu)$. Geller showed that the analogy between $\mathscr{L}^\mu$ on $\mathbb{S}^{2n+1}$ and $\mathscr{L}_\mu$ on $\mathbb{H}^n$ is very strong: for instance $\mathscr{L}^\mu$ is hypoelliptic and locally solvable unless $\mu\in\{\pm n,\pm(n+2),\dots\}$, and $\mathscr{L}^\mu$ is in an appropriate sense the diagonal part of $\Box_b$ on $(0,q)$-forms if one takes $\mu=n-2q$. He also derived in [2] a relation between $\Delta^{\alpha,\beta}$ and $\mathscr{L}^\mu$, which amounts to writing $\Delta^{\alpha,\beta}$ in polar coordinates. In order to do this, first consider $\mathscr{L}^\mu$ as a differential operator on all of $\mathbb{C}^{n+1}$ by requiring that the formulae above for M_{jk}, T, and $\mathscr{L}^\mu$ hold for all $w\in\mathbb{C}^{n+1}$. Also set $W=\sum_1^{n+1}w_i\frac{\partial}{\partial w_i}$ and $N=-\frac{1}{2}(W+\overline{W})$. Then writing out the definition of $\mathscr{L}^\mu$ in coordinates, one gets

$$\begin{aligned}\mathscr{L}^\mu=&-\sum(|w|^2\delta_{ij}-w_i\bar{w}_j)\frac{\partial^2}{\partial w_i\partial\bar{w}_j}+\tfrac{1}{2}(n+\mu)\overline{W}+\tfrac{1}{2}(n-\mu)W\\&+\tfrac{1}{4}(n+\mu)(n-\mu).\end{aligned}$$

Setting $\tilde{\mathscr{L}}^\mu=\mathscr{L}^\mu-\frac{1}{4}(n+\mu)(n-\mu)$, it follows in particular that

$$\tilde{\mathscr{L}}^0=-\sum(|w|^2\delta_{ij}-w_i\bar{w}_j)\frac{\partial^2}{\partial w_i\partial\bar{w}_j}-nN$$

so that

$$\begin{aligned}-(\tilde{\mathscr{L}}^0+nN)&=\sum(|w|^2\delta_{ij}-w_i\bar{w}_j)\frac{\partial^2}{\partial w_i\partial\bar{w}_j}\\&=|w|^2\sum(\delta_{ij}-w_i\bar{w}_j)\frac{\partial^2}{\partial w_i\partial\bar{w}_j}-(1-|w|^2)W\overline{W}\\&=\frac{|w|^2}{1-|w|^2}\Delta^{0,0}-(1-|w|^2)(N^2+T^2).\end{aligned}$$

Hence

$$\frac{|w|^2}{1-|w|^2}\Delta^{0,0}=(1-|w|^2)(N^2+T^2)-\tilde{\mathscr{L}}^0-nN. \tag{2.2}$$

This is the polar coordinate decomposition of the Bergman Laplacian on the ball which should be compared with (1.1). For $\Delta^{\alpha,\beta}$ one uses

$$\Delta^{\alpha,\beta}=\Delta^{0,0}+(1-|w|^2)[-(\alpha+\beta)N+i(\beta-\alpha)T-\alpha\beta]$$

to obtain

$$\Delta^{\alpha,\beta}=\frac{1-|w|^2}{|w|^2}\{(1-|w|^2)(N^2+T)-(\tilde{\mathscr{L}}^{\mu}+\alpha\beta)+(1-\lambda)N$$

$$+(1-|w|^2)[(\alpha+\beta)N+i\mu T+\alpha\beta]\}, \tag{2.3}$$

where (λ,μ) and (α,β) are related as in (2.1).

We immediately put this formula to use to derive the transformation law for the $\mathscr{L}_\mu$ operators under the Cayley transform.

(2.4) **Proposition.** *Let $\mu\in\mathbb{C}$ and set $\alpha=\frac{1}{2}(-n+\mu)$, $\beta=\frac{1}{2}(-n-\mu)$. Then*

$$M^{\alpha-1,\beta-1}\,C^*\,\mathscr{L}_\mu=\mathscr{L}^{\mu}\,M^{\alpha,\beta}\,C^*.$$

Proof. This follows easily from (2.1). In fact, by (1.2), if $f\in C^2(b\mathbb{D})$ and $\tilde{f}$ is any C^2 extension of f to $\mathbb{C}^{n+1}$, then at $\rho=0$, $(\rho^{-1}\Delta_{1,\mu})\tilde{f}=-\mathscr{L}_\mu f$. Similarly if $\alpha=\frac{1}{2}(-n+\mu)$, $\beta=\frac{1}{2}(-n-\mu)$, then (α,β) are related via (2.1) to $\lambda=1$, $\mu=\mu$, so by (2.3) if $g\in C^2(\mathbb{S}^{2n+1})$ and $\tilde{g}$ is any C^2 extension of g to $\mathbb{C}^{n+1}$, then on $\mathbb{S}^{2n+1}$,

$$((1-|w|^2)^{-1}\Delta^{\alpha,\beta})\tilde{g}=-(\tilde{\mathscr{L}}^{\mu}+\alpha\beta)g=-\mathscr{L}^{\mu}g.$$

Noting that $C^*\rho=\dfrac{1-|w|^2}{|1+w_{n+1}|^2}$ one obtains from (2.1) that

$$\begin{aligned}\Delta^{\alpha,\beta}M^{\alpha,\beta}C^*\tilde{f}&=M^{\alpha,\beta}C^*\Delta_{1,\mu}\tilde{f}=M^{\alpha,\beta}C^*(\rho\cdot\rho^{-1}\Delta_{1,\mu}\tilde{f})\\&=M^{\alpha,\beta}\frac{1-|w|^2}{|1+w_{n+1}|^2}C^*(\rho^{-1}\Delta_{1,\mu}\tilde{f})\\&=(1-|w|^2)M^{\alpha-1,\beta-1}C^*(\rho^{-1}\Delta_{1,\mu}\tilde{f}),\end{aligned}$$

so dividing by $1-|w|^2$ and restricting to $b\mathbb{D}$ gives $\mathscr{L}^{\mu}M^{\alpha,\beta}C^*f=M^{\alpha-1,\beta-1}C^*\mathscr{L}_\mu f$ as claimed.

Define differential operators $\mathscr{C}^{k,\mu}$ on the sphere for $k\in\mathbb{N}$, $\mu\in\mathbb{C}$ by

$$\mathscr{C}^{k,\mu}=\prod_{l=1}^{k}\mathscr{L}^{\mu+k+1-2l} \tag{2.5}$$

An easy computation shows that $\mathscr{L}^0$ and T commute, so that the order of the factors is irrelevant. The relation between these operators and their analogues on $\mathbb{H}^n$ is then given by:

(2.6) **Theorem.** Let $k\in\mathbb{N}$, $\mu\in\mathbb{C}$, and define α,β by (2.1) with $\lambda=k$. Then $M^{\alpha-k,\beta-k}\,C^*\,\mathscr{C}_{k,\mu}=\mathscr{C}^{k,\mu}\,M^{\alpha,\beta}\,C^*$.

Proof. First, by using the transformation law for the $\mathscr{L}_\mu$, this identity reduces to an identity on $\mathbb{H}^n$. Denote by $M_{\alpha,\beta}$ the multiplication operators on functions on $\mathbb{D}$ determined by $M^{\alpha,\beta}C^*=C^*M_{\alpha,\beta}$; since

$$C^*2(1-iz_{n+1})^{-1}=1+w_{n+1}$$

it follows that $M_{\alpha,\beta}f=2^{\alpha+\beta}(1+i\bar{z}_{n+1})^{-\alpha}(1-iz_{n+1})^{-\beta}f.$

Then by Proposition 2.4,

$$\mathscr{C}^{k,\mu} M^{\alpha,\beta} C^* = \prod_{l=1}^{k-1} \mathscr{L}^{\mu+k+1-2l} \mathscr{L}^{\mu-k+1} C^* M_{\alpha,\beta}$$
$$= \prod_{l=1}^{k-1} \mathscr{L}^{\mu+k+1-2l} C^* M_{\gamma-1,\delta-1} \mathscr{L}_{\mu-k+1} M_{-\gamma,-\delta} M_{\alpha,\beta},$$

where $\gamma=\frac{1}{2}(-n+\mu-k+1)$, $\delta=\frac{1}{2}(-n-\mu+k-1)$. But $\alpha=\frac{1}{2}(k-n-1+\mu)$, $\beta=\frac{1}{2}(k-n-1-\mu)$. Thus this becomes

$$\mathscr{C}^{k,\mu} M^{\alpha,\beta} C^* = \prod_{l=1}^{k-2} \mathscr{L}^{\mu+k+1-2l} \mathscr{L}^{\mu-k+3} C^* M_{\gamma-1,\delta-1} \mathscr{L}_{\mu-k+1} M_{k-1,0}.$$

Repeatedly passing C^* through each factor of $\mathscr{L}^{\mu+k+1-2l}$ leaves one with

$$\mathscr{C}^{k,\mu} M^{\alpha,\beta} C^* = C^* M_{\alpha-1,\beta-k} \mathscr{L}_{\mu+k-1} M_{-2,0} \mathscr{L}_{\mu+k-3} M_{-2,0}$$
$$\ldots M_{-2,0} \mathscr{L}_{\mu-k+1} M_{k-1,0},$$

so that Theorem 2.6 reduces to proving the curious identity

$$M_{k+2,0} \prod_{l=0}^{k} (M_{-2,0} \mathscr{L}_{\mu+k-2l}) M_{k,0} = \mathscr{C}_{k+1,\mu} \tag{2.7}$$

for $k \geqq 0$, k having been replaced by $k+1$.

For the rest of the argument we write M_α for $M_{\alpha,0}$. (2.7) will be proved by using several other identities. First set $R_\mu = \sum_{j=1}^{n} z_j Z_j + \frac{1}{2}(\mu+1)$. Then straightforward calculation shows that

$$[\mathscr{L}_\mu, M_\alpha] = \alpha M_{\alpha+1} R_\mu \quad \text{and} \quad \mathscr{L}_{\mu+2} R_\mu = R_{\mu+2} \mathscr{L}_\mu. \tag{2.8}$$

Also note that $\mathscr{C}_{k+1,\mu} = \mathscr{C}_{k,\mu+1} \mathscr{L}_{\mu-k}$, $k \geqq 1$.

Next we prove by induction on k that

$$M_1 \mathscr{C}_{k,\mu} M_{-1} = \mathscr{C}_{k,\mu} - k M_1 \mathscr{C}_{k-1,\mu+1} R_{\mu-k+1}. \tag{2.9}$$

For $k=1$ this follows from (2.8) if we take $\mathscr{C}_{0,\mu}$ = identity, since

$$M_1 \mathscr{L}_\mu M_{-1} = M_1 (M_{-1} \mathscr{L}_\mu - R_\mu) = \mathscr{L}_\mu - M_1 R_\mu.$$

And

$$M_1 \mathscr{C}_{k+1,\mu} M_{-1} = M_1 \mathscr{C}_{k,\mu+1} \mathscr{L}_{\mu-k} M_{-1} = M_1 \mathscr{C}_{k,\mu+1} (M_{-1} \mathscr{L}_{\mu-k} - R_{\mu-k})$$
$$= M_1 \mathscr{C}_{k,\mu+1} M_{-1} \mathscr{L}_{\mu-k} - M_1 \mathscr{C}_{k,\mu+1} R_{\mu-k},$$

which by the induction hypothesis is

$$= \mathscr{C}_{k,\mu+1} \mathscr{L}_{\mu-k} - k M_1 \mathscr{C}_{k-1,\mu+2} R_{\mu-k+2} \mathscr{L}_{\mu-k} - M_1 \mathscr{C}_{k,\mu+1} R_{\mu-k}$$
$$= \mathscr{C}_{k+1,\mu} - k M_1 \mathscr{C}_{k-1,\mu+2} \mathscr{L}_{\mu-k+2} R_{\mu-k} - M_1 \mathscr{C}_{k,\mu+1} R_{\mu-k}$$
$$= \mathscr{C}_{k+1,\mu} - k M_1 \mathscr{C}_{k,\mu+1} R_{\mu-k} - M_1 \mathscr{C}_{k,\mu+1} R_{\mu-k}$$
$$= \mathscr{C}_{k+1,\mu} - (k+1) M_1 \mathscr{C}_{k,\mu+1} R_{\mu-k},$$

thus establishing (2.9).

Now (2.7) can be proved by induction on k. For $k=0$ this simply states $\mathscr{C}_{1,\mu}=\mathscr{L}_\mu$. For $k>0$ write

$$\begin{aligned}
&M_{k+2}\prod_{l=0}^{k}(M_{-2}\mathscr{L}_{\mu+k-2l})M_k\\
&\quad=M_1 M_{k+1}\prod_{l=0}^{k-1}(M_{-2}\mathscr{L}_{\mu+k-2l})M_{k-1}M_{-k-1}\mathscr{L}_{\mu-k}M_k
\end{aligned}$$

and use the induction hypothesis to obtain

$$\begin{aligned}
M_1\mathscr{C}_{k,\mu+1}M_{-k-1}\mathscr{L}_{\mu-k}M_k&=M_1\mathscr{C}_{k,\mu+1}M_{-k-1}(M_k\mathscr{L}_{\mu-k}+kM_{k+1}R_{\mu-k})\\
&=M_1\mathscr{C}_{k,\mu+1}M_{-1}\mathscr{L}_{\mu-k}+kM_1\mathscr{C}_{k,\mu+1}R_{\mu-k},\quad \text{which by (2.9) is}\\
&=(\mathscr{C}_{k,\mu+1}-kM_1\mathscr{C}_{k-1,\mu+2}R_{\mu-k+2})\mathscr{L}_{\mu-k}+kM_1\mathscr{C}_{k,\mu+1}R_{\mu-k}\\
&=\mathscr{C}_{k+1,\mu}-kM_1\mathscr{C}_{k-1,\mu+2}\mathscr{L}_{\mu-k+2}R_{\mu-k}+kM_1\mathscr{C}_{k,\mu+1}R_{\mu-k}\\
&=\mathscr{C}_{k+1,\mu}-kM_1\mathscr{C}_{k,\mu+1}R_{\mu-k}+kM_1\mathscr{C}_{k,\mu+1}R_{\mu-k}\\
&=\mathscr{C}_{k+1,\mu},\ \text{proving (2.7); hence also Theorem 2.6.}
\end{aligned}$$

Combining Theorem 2.6 with our previous observation on compatibility operators for $\Delta^{\alpha,\beta}$, it follows that we have established the analogue of Theorem 1.3 for the ball.

(2.10) **Theorem.** *Let $\Omega^{\text{open}}\subset\mathbb{C}^{n+1}$ with $\Omega\cap\mathbb{S}^{2n+1}\neq\emptyset$, let $f\in C^\infty(\Omega\cap\mathbb{S}^{2n+1})$, and $\alpha,\beta\in\mathbb{C}$. Suppose that $k=n+1+\alpha+\beta\in\mathbb{N}$ and set $\mu=\alpha-\beta$.*

a) *If there is a $u\in C^\infty(\Omega\cap\overline{\mathbb{B}})$ solving $\Delta^{\alpha,\beta}u=0$ with $u|_{\Omega\cap\mathbb{S}^{2n+1}}=f$, then $\mathscr{C}^{k,\mu}f=0$.*
b) *Suppose that $\alpha,\beta\notin\{-n-1,-n-2,\dots\}$ and let $u\in C(\Omega\cap\overline{\mathbb{B}})$ be any solution of $\Delta^{\alpha,\beta}u=0$ with $u|_{\Omega\cap\mathbb{S}^{2n+1}}=f$ where $\mathscr{C}^{k,\mu}f=0$. Then $u\in C^\infty(\Omega\cap\overline{\mathbb{B}})$.*

Of course, the proof indicated via the Cayley transform only works away from $(0,\dots,0,-1)\in\mathbb{S}^{2n+1}$, but since the final result is unitarily invariant the general case follows immediately. Also note that in a) of both Theorems 1.3 and 2.10, it is not necessary that u be infinitely differentiable to get the conclusion; u need only have enough continuous derivatives at the boundary in order to allow the computations leading to the condition $\mathscr{C}_{k,\mu}u=0$. We also note that as mentioned in [4], the results of Theorem 10.2 of [4] used in Theorem 1.3 b) also hold in the real analytic case, so that in b) of both Theorems 1.3 and 2.10, if f is analytic then u is analytic up to the boundary.

Finally we discuss the possibility of deriving the compatibility operator $\mathscr{C}^{k,\mu}$ directly from the equation $\Delta^{\alpha,\beta}u=0$ in the form (2.3). It is clear that a compatibility condition arises by successively applying N and restricting to the boundary and it is possible to compute it this way for small k. But the general situation leads to a complicated recursion relation. We compute this recursion explicitly for the case $\alpha=\beta=0$, which appears to be simpler than the general case, but still contains the essential complexities. In doing the computations note that $\tilde{\mathscr{L}}^0$, N and T all commute. Also set $s=1-|w|^2$ and note that $Ns=1-s$. As the equation $s^{-1}\Delta^{0,0}u=0$ is $[s(N^2+T^2)-\tilde{\mathscr{L}}^0-nN]u=0$, the only formal difference between this case and the analogous one: $[\rho(N^2+T^2)-\mathscr{L}_0-nN]u=0$ on $\mathbb{D}$ is that $Ns=1-s$ while $N\rho=1$.

First, restricting to $s=0$ gives $Nu=-\frac{1}{n}\tilde{\mathscr{L}}^0 u$. As $N^j s=(-1)^{j+1}$ at $s=0$ for $j\geqq 1$, applying N^l to the equation $s^{-1}\Delta^{0,0}u=0$ and restricting to $s=0$ gives

$$\sum_{j=1}^{l}\binom{l}{j}(-1)^{j+1}(N^{l-j+2}+N^{l-j}T^2)u-\tilde{\mathscr{L}}^0 N^l u-nN^{l+1}u=0, \quad \text{or}$$

$$(l-n)N^{l+1}u=\tilde{\mathscr{L}}^0 N^l u-lT^2 N^{l-1}u+\sum_{j=2}^{l}(-1)^j\binom{l}{j}(N^{l-j+2}+N^{l-j}T^2)u.$$

As before introduce $Q_l u=(-1)^l\frac{n!}{(n-l)!}N^l u$ at $s=0$; then this simplifies to $Q_0=1$, $Q_1=\tilde{\mathscr{L}}^0$, and

$$Q_{l+1}=[\tilde{\mathscr{L}}^0+\tfrac{1}{2}l(l-1)]Q_l+\sum_{j=0}^{l-1}\frac{(n-j)!}{(n-l)!}\left[\binom{l}{j-2}+\binom{l}{j}T^2\right]Q_j, \qquad l\geqq 1,$$

where $\binom{l}{j-2}$ is interpreted as 0 if $j<2$.

Thus setting $x=\tilde{\mathscr{L}}^0$, $y=T$, the recursion reduces to the following for polynomials $q_l(x,y)$:

$$\text{(2.11)}\qquad \begin{aligned} &q_0=1,\ q_1=x, \quad \text{and}\\ &q_{l+1}=[x+\tfrac{1}{2}l(l-1)]q_l+\sum_{j=0}^{l-1}\frac{(n-j)!}{(n-l)!}\left[\binom{l}{j-2}+\binom{l}{j}y^2\right]q_j, \qquad l\geqq 1. \end{aligned}$$

Formula (2.5) for $\mathscr{C}^{n+1,0}$ amounts to the identity

$$q_{n+1}=\prod_{j=0}^{n}[x+i(n-2j)y+j(n-j)].$$

Is there a direct way to prove this? The proof presented here via the Cayley transform is essentially a change of variables reducing this identity to the simpler one on $\mathbb{H}^n$. Perhaps there is a direct change of variables in the recursion relation. Notice that the $\mathbb{H}^n$ recursion is contained in this one: the leading homogeneous part of q_l for $\mathbb{S}^{2n+1}$ can be identified with q_l for $\mathbb{H}^n$ and this part of the recursion and identity on the sphere coincide with the full recursion and identity on $\mathbb{H}^n$. Finally, one can ask the same questions for $\alpha,\beta\neq 0$.

3. Fundamental Solutions for $\mathscr{C}_{k,\mu}$

In this section we derive fundamental solutions for the operators $\mathscr{C}_{k,\mu}$. Folland and Stein [1] showed that $\mathscr{L}_\mu$ is hypoelliptic and locally solvable if and only if $\pm\mu\notin\{n,n+2,\dots\}$, and they found a fundamental solution for $\mathscr{L}_\mu$ in this case. Call such μ admissible. Now if k,μ are such that $\mu+k-2l+1$ is admissible for all $l\in\{1,2,\dots,k\}$, then it follows that $\mathscr{C}_{k,\mu}=\prod_{l=1}^{k}\mathscr{L}_{\mu+k-2l+1}$ is hypoelliptic and locally solvable too. For such k, μ we will compute a fundamental solution $E_{k,\mu}$ of $\mathscr{C}_{k,\mu}$. However if one of $\mu+k-2l+1$ is not admissible then one cannot expect to obtain a fundamental solution for $\mathscr{C}_{k,\mu}$. In the special cases in which the only inadmissible values of the form $\mu+k-2l+1$ are either $\pm n$, we will instead obtain a relative fundamental solution, generalizing the results of Greiner-Kohn-Stein [5]. This result in the special case $k=n+1$, $\mu=0$ was announced in [4]. The author would like to thank D. Geller for pointing out the possibility of computing explicit fundamental solutions for the other operators $\mathscr{C}_{k,\mu}$. Geller's derivation is different from the one presented here and proceeds by directly computing the Heisenberg convolutions of the fundamental solutions of the factors $\mathscr{L}_{\mu+k-2l+1}$.

In order to compute the fundamental solutions, one is led by the unitary invariance of the $\mathscr{L}_\mu$ operators in the ζ variables to look for a solution of the form $E(|\zeta|^2, t)$. Now introduce the complex variable $w=|\zeta|^2-it$, then $E=E(w,\bar{w})$ and an equation for E on the right half plane is obtained. Straightforward calculation shows that $Z_j E = 2\bar{\zeta}_j \frac{\partial E}{\partial w}$ and $TE = i\left(\frac{\partial E}{\partial \bar{w}} - \frac{\partial E}{\partial w}\right)$, so

$$\mathscr{L}_\mu E = -4\,\mathrm{Re}\, w \frac{\partial^2 E}{\partial w\, \partial \bar{w}} - (n-\mu)\frac{\partial E}{\partial w} - (n+\mu)\frac{\partial E}{\partial \bar{w}}. \tag{3.1}$$

Thus if operators $L_{r,s}$ are defined on the right half plane by

$$L_{r,s} = -\left[4\,\mathrm{Re}\, w \frac{\partial^2}{\partial w\, \partial \bar{w}} + 2r\frac{\partial}{\partial w} + 2s\frac{\partial}{\partial \bar{w}}\right],$$

then $\mathscr{L}_\mu$ on $\mathbb{H}^n$ becomes $L_{r,s}$ with $r=\frac{1}{2}(n-\mu)$, $s=\frac{1}{2}(n+\mu)$. It is rather curious that $L_{r,s}$ are essentially our degenerate operators $\Delta_{\lambda,\mu}$ in the case of one complex dimension. Recall [1] that $\Box_b$ on $(0,r)$-forms on $\mathbb{H}^n$ is $\mathscr{L}_{n-2r}$ componentwise, so taking $\mu=n-2r$ one sees that for r and s nonnegative integers $L_{r,s}$ represents the action of $\Box_b$ on $(0,r)$-forms and $s=n-r$ is the codegree of the form. This observation and the identities $\frac{\partial}{\partial w} L_{r,s} = L_{r,s+1}\frac{\partial}{\partial w}$, $\frac{\partial}{\partial \bar{w}} L_{r,s} = L_{r+1,s}\frac{\partial}{\partial \bar{w}}$ can be used to give an a-priori derivation of the fundamental solution for $\Box_b$ when $1 \leqq r \leqq n-1$ and of the Greiner-Kohn-Stein relative fundamental solution when $r=0$, by reduction to the single equation $L_{0,0}E = \frac{1}{w} + \frac{1}{\bar{w}}$, with solution $E = -\frac{1}{2}\log w \log \bar{w}$. This alternate derivation was shown to me by Reese Harvey. It, however, is limited to the case when r and s are nonnegative integers, so to analyze more general $\mathscr{L}_\mu$ and $\mathscr{C}_{k,\mu}$ we proceed differently.

First consider $\mathscr{C}_{k+1,\mu} = \prod_{l=0}^{k} \mathscr{L}_{\mu+k-2l}$ in the case in which $\mu+k-2l$ is admissible for $0 \leqq l \leqq k$. (We have switched from k to $k+1$ for notational convenience.) Note that this is equivalent to the requirement that each of the two endpoints $\mu+k$ and $\mu-k$ is admissible, which is equivalent again to $n+\mu-k$, n

$-\mu-k\notin\{0,-2,-4,\dots\}$. We begin by searching for a solution $E_{k+1,\mu}(w,\bar{w})$ to $\mathscr{C}_{k+1,\mu}E_{k+1,\mu}=0$ away from $w=0$, with a singularity of the right form at $w=0$. Using (3.1) gives $\prod_{l=0}^{k} L_{r+l,s+k-l}E=0$, with $r=\frac{1}{2}(n-\mu-k)$, $s=\frac{1}{2}(n+\mu-k)$. For $k=0$ this reads $L_{r,s}E=0$ and the solution $E=w^{-s}\bar{w}^{-r}$ can be guessed from the identity

$$L_{r,s}=-2\left[\frac{\partial}{\partial\bar{w}}\left(w\frac{\partial}{\partial w}+s\right)+\frac{\partial}{\partial w}\left(\bar{w}\frac{\partial}{\partial\bar{w}}+r\right)\right]. \tag{3.2}$$

To solve when $k\geqq 1$, one is first led by this same identity to note that

$$L_{r+l,s}w^{-s}\bar{w}^{-r}=2slw^{-s-1}\bar{w}^{-r},\qquad l\in\mathbb{C}.$$

But then this immediately yields the solution for general k inductively:

$$\begin{aligned}L_{r,s+k}\dots L_{r+k,s}w^{-s}\bar{w}^{-r}&=L_{r,s+k}\dots L_{r+k-1,s+1}2skw^{-s-1}\bar{w}^{-r}\\&=\dots=L_{r,s+k}2^k s(s+1)\dots(s+k-1)k!\,w^{-s-k}\bar{w}^{-r}=0.\end{aligned}$$

Thus we have shown that $\prod_{l=0}^{k} L_{r+l,s+k-l}w^{-s}\bar{w}^{-r}=0$ on the right half plane for any $s,r\in\mathbb{C}$, so taking $r=\frac{1}{2}(n-\mu-k)$, $s=\frac{1}{2}(n+\mu-k)$ as above shows that $\mathscr{C}_{k+1,\mu}w^{-s}\bar{w}^{-r}=0$ away from 0. It remains to analyze the singularity at the origin to prove

(3.3) **Theorem.** *Let $k\geqq 0$, $\mu\in\mathbb{C}$ be such that $n+\mu-k, n-\mu-k\notin\{0,-2,-4,\dots\}$. Then*

$$E_{k+1,\mu}=c_{k,\mu}(|\zeta|^2-it)^{-\frac{1}{2}(n+\mu-k)}(|\zeta|^2+it)^{-\frac{1}{2}(n-\mu-k)}$$

satisfies $\mathscr{C}_{k+1,\mu}E_{k+1,\mu}=\delta_0$, where

$$c_{k,\mu}=\frac{\Gamma(\frac{1}{2}(n+\mu-k))\,\Gamma(\frac{1}{2}(n-\mu-k))}{2^{2+k-n}\pi^{n+1}k!},$$

and δ_0 is the Dirac delta mass at 0.

Note. We have determined $c_{k,\mu}$ by the normalization obtained by taking $dx_1\dots dx_n dy_1\dots dy_n dt$ to be the volume form on $\mathbb{H}^n$. This explains the discrepancy in the constant with [1] in case $k=0$, where $\mathscr{C}_{1,\mu}=\mathscr{L}_\mu$.

Proof. As above set $r=\frac{1}{2}(n-\mu-k)$, $s=\frac{1}{2}(n+\mu-k)$; then the iterative computation above shows that

$$\mathscr{C}_{k+1,\mu}w^{-s}\bar{w}^{-r}=\frac{2^k\Gamma(s+k)k!}{\Gamma(s)}\mathscr{L}_{\mu+k}w^{-s-k}\bar{w}^{-r}$$

away from the origin. However as a simple limiting argument shows, at each stage of this computation the identity that was used also holds in the distribution sense across the origin, since each function $w^{-s-l}\bar{w}^{-r}$ that occurs is homogeneous and locally integrable. For the final step we are left with

$\mathscr{L}_{\mu+k}(w^{-\frac{1}{2}(n+\mu+k)}\bar{w}^{-\frac{1}{2}(n-\mu-k)})$. However it follows from the hypotheses that $\mu+k$ is admissible, so that we are in the case studied by Folland-Stein. Hence

$$\begin{aligned}\mathscr{C}_{k+1,\mu} w^{-s}\bar{w}^{-r} &= \frac{2^k\Gamma(s+k)k!}{\Gamma(s)} \mathscr{L}_{\mu+k} w^{-s-k}\bar{w}^{-r} = \frac{2^k\Gamma(s+k)k!}{\Gamma(s)} \frac{2^{2-n}\pi^{n+1}}{\Gamma(s+k)\Gamma(r)}\delta_0 \\ &= \frac{2^{2+k-n}\pi^{n+1}k!}{\Gamma(s)\Gamma(r)}\delta_0,\end{aligned}$$

and Theorem 3.3 is proved.

Next consider the case of $\mathscr{C}_{k+1,\mu} = \prod_{l=0}^{k} \mathscr{L}_{\mu+k-2l}$ in which exactly one of $\mu+k-2l$ is inadmissible and is of the form $\pm n$. Clearly either $\mu+k=n$ or $\mu-k=-n$, and as the two situations are symmetric it suffices to consider the case $\mu+k=n$. Thus the operator in question is $\mathscr{C}_{k+1,n-k} = \prod_{l=0}^{k} \mathscr{L}_{n-2l}$, and $0 \leqq k \leqq n-1$. Let

$$S = \frac{2^{n-1}n!}{\pi^{n+1}} \lim_{\rho\to 0} (|\zeta|^2+\rho-it)^{-n-1}$$

be the Cauchy-Szegö kernel on $\mathbb{H}^n$.

(3.4) **Theorem.** *Let* $0 \leqq k \leqq n-1$. *Then*

$$E_{k+1} = c_k \log\left(\frac{|\zeta|^2-it}{|\zeta|^2+it}\right)\cdot(|\zeta|^2-it)^{-n+k}$$

satisfies $\mathscr{C}_{k+1,n-k}E_{k+1} = \delta_0 - S$, *where* $c_k = \dfrac{\Gamma(n-k)}{\pi^{n+1}2^{2+k-n}k!}$.

Proof. First of all notice that $\mathscr{C}_{k+1,n-k}E_{k+1} = \mathscr{C}_{k+1,n-k}\left(-c_k\dfrac{\log\bar{w}}{w^{n-k}}\right)$. In fact, $\mathscr{L}_n$ annihilates the boundary values of holomorphic functions on $\mathbb{D}$, and since on $\mathbb{H}^n = b\mathbb{D}$, $w = |\zeta|^2 - it = -iz_{n+1}$, it follows that for $\varepsilon>0$ $\dfrac{\log(w+\varepsilon)}{(w+\varepsilon)^{n-k}}$ is the restriction to $\mathbb{H}^n$ of a function holomorphic in $\mathbb{D}$ and smooth in $\bar{\mathbb{D}}$, so $\mathscr{L}_n\dfrac{\log(w+\varepsilon)}{(w+\varepsilon)^{n-k}} = 0$. As $\varepsilon\to 0$, $\dfrac{\log(w+\varepsilon)}{(w+\varepsilon)^{n-k}} \to \dfrac{\log w}{w^{n-k}}$ as distributions, so $\mathscr{L}_n\dfrac{\log w}{w^{n-k}} = 0$, hence $\mathscr{C}_{k+1,n-k}\dfrac{\log w}{w^{n-k}} = 0$ also. Next we calculate

$$\mathscr{C}_{k+1,n-k}\frac{\log\bar{w}}{w^{n-k}} = L_{0,n}L_{1,n-1}\cdots L_{k,n-k}\frac{\log\bar{w}}{w^{n-k}}$$

by the techniques used above. Using (3.2), this is

$$\begin{aligned}&= L_{0,n}L_{1,n-1}\cdots L_{k-1,n-k+1}\left[\frac{2(n-k)}{w^{n-k+1}}(1+k\log\bar{w})\right] \\ &= L_{0,n}\cdots L_{k-1,n-k+1}\frac{2k(n-k)\log\bar{w}}{w^{n-k+1}}\end{aligned}$$

since also $\dfrac{1}{w^{n-k+1}}$ is the boundary value of a holomorphic function so is annihilated by $\mathscr{L}_n = L_{0,n}$. Iterating yields

$$\mathscr{C}_{k+1,n-k}\frac{\log\bar{w}}{w^{n-k}} = \frac{2^k k!(n-1)!}{(n-k-1)!}\mathscr{L}_n\frac{\log\bar{w}}{w^n}$$

and as before this equation holds in the distribution sense across the origin. Finally, the Greiner-Kohn-Stein result is

$$\mathscr{L}_n\frac{\log\bar{w}}{w^n} = \frac{-\pi^{n+1}}{2^{n-2}(n-1)!}[\delta_0 - S]$$

so that Theorem 3.4 follows.

Finally we investigate the case when $\pm n$ are the only inadmissible values occurring, so that the operator is $\mathscr{C}_{n+1,0} = \prod_{l=0}^{n}\mathscr{L}_{n-2l}$. By playing with identities like (3.2) for the operators $L_{r,s}$ it is possible to guess the fundamental solution away from the origin, then to carry out a limiting argument to handle the singularity at 0. Of course since both $\mathscr{L}_n$ and $\mathscr{L}_{-n}$ occur in the product, the projection operator which arises is the sum of the projections onto the mutually orthogonal L^2-kernels of $\mathscr{L}_n$ and $\mathscr{L}_{-n}$, namely $S + \bar{S}$. Rather than proceed this way, however, we instead derive the fundamental solution for $\mathscr{C}_{n+1,0}$ by differentiating that for $\mathscr{C}_{n+1,\mu}$ at $\mu = 0$, similar to one of the original derivations of Greiner-Kohn-Stein. This technique could also have been used to prove Theorem 3.4.

(3.5) **Theorem.** *Define*

$$E = (8\pi^{n+1}n!)^{-1}\left|\log\left(\frac{|\zeta|^2 + it}{|\zeta|^2 - it}\right)\right|^2.$$

Then $\mathscr{C}_{n+1,0}E = \delta_0 - (S + \bar{S})$.

Proof. Take $k = n$ and μ near 0 in Theorem 3.3 to obtain $\mathscr{C}_{n+1,\mu}\tilde{E}_\mu = \lambda_\mu\delta_0$, where $\tilde{E}_\mu = \left(\dfrac{\bar{w}}{w}\right)^{\mu/2}$ and

$$\lambda_\mu = c_{n,\mu}^{-1} = \frac{4\pi^{n+1}n!}{\Gamma(\frac{1}{2}\mu)\Gamma(-\frac{1}{2}\mu)}.$$

Now differentiate this relation twice with respect to μ and set $\mu = 0$. As $\Gamma(x) \sim \dfrac{1}{x}$ for x near 0, $\lambda''_\mu = -2\pi^{n+1}n!$. And

$$\tilde{E}''_\mu = \tfrac{1}{4}\left(\log\frac{\bar{w}}{w}\right)^2 = -\tfrac{1}{4}\left|\log\frac{\bar{w}}{w}\right|^2.$$

As $\tilde{E}_0 = 1$, the term $\mathscr{C}''_{n+1,\mu}\tilde{E}_\mu$ vanishes at $\mu = 0$, so that the only remaining term is

$$2\mathscr{C}'_{n+1,\mu}\tilde{E}'_\mu = 2\mathscr{C}'_{n+1,\mu}\left(\frac{1}{2}\log\frac{\bar{w}}{w}\right).$$

Observing as always that $\mathscr{L}_n$ annihilates boundary values of holomorphic functions and $\mathscr{L}_{-n}$ boundary values of anti-holomorphic functions and noting that $\frac{d}{d\mu}\mathscr{L}_\mu = iT$, one obtains

$$\begin{aligned}2\mathscr{C}'_{n+1,\mu}\tilde{E}'_\mu &= (\mathscr{L}_n\mathscr{L}_{n-2}\cdots\mathscr{L}_{-n+2}+\mathscr{L}_{n-2}\cdots\mathscr{L}_{-n})\,iT(\log\bar{w}-\log w)\\ &= -[\mathscr{L}_n\mathscr{L}_{n-2}\cdots\mathscr{L}_{-n+2}\bar{w}^{-1}+\mathscr{L}_{n-2}\cdots\mathscr{L}_{-n}w^{-1}].\end{aligned}$$

However

$$\mathscr{L}_{n-2}\cdots\mathscr{L}_{-n}w^{-1} = L_{1,n-1}L_{2,n-2}\cdots L_{n,0}w^{-1},$$

and repeatedly applying the relation $L_{r,s}w^{-p}=2rpw^{-p-1}$ yields $\mathscr{L}_{n-2}\cdots\mathscr{L}_{-n}w^{-1}=2^n n!^2 w^{-n-1}$. The usual limiting argument extends this across the origin so that

$$2\mathscr{C}'_{n+1,\mu}\tilde{E}'_\mu = -2^n n!^2(w^{-n-1}+\bar{w}^{-n-1}) = -2\pi^{n+1}n!(S+\bar{S}),$$

and the theorem follows upon collecting the terms.

Remarks. 1. As a relative fundamental solution for $\mathscr{C}_{n+1,0}$ is only determined up to boundary values of holomorphic and antiholomorphic functions, it can be expressed in forms that appear quite different. For example,

$$\left|\log\frac{\bar{w}}{w}\right|^2 = 2\,|\log w|^2-(\log\bar{w})^2-(\log w)^2,$$

so that $(4\pi^{n+1}n!)^{-1}\,|\log(|\zeta|^2-it)|^2$ is another relative fundamental solution.

2. In case $n=1$ the relative fundamental solution for $\mathscr{C}_{2,0}=\mathscr{L}_1\mathscr{L}_{-1}$ was derived by Laville [6]. Laville was interested in $\mathscr{L}_1\mathscr{L}_{-1}$ because it characterizes boundary values of pluriharmonic functions: a function $u\in L^2(\mathbb{H}^1)$ is the boundary value of a pluriharmonic function on $\mathbb{D}$ if and only if $\mathscr{L}_1\mathscr{L}_{-1}u=0$, as follows upon convolving both sides of the equation $\mathscr{L}_1\mathscr{L}_{-1}E=\delta_0-(S+\bar{S})$ with u and noting that $u\to u*(S+\bar{S})$ is orthogonal projection onto the boundary values of pluriharmonic functions. Analogously, on $\mathbb{S}^3$ $\mathscr{L}^1\mathscr{L}^{-1}$ characterizes boundary values of pluriharmonic functions. The natural generalization of this operator to higher dimensions is $\mathscr{L}^n\mathscr{L}^{-n}$, whose kernel can be seen to be exactly the boundary values of pluriharmonic functions by a spherical harmonics expansion. However there does not appear to be a simple formula for a relative fundamental solution for $\mathscr{L}_n\mathscr{L}_{-n}$ other than to take $\mathscr{L}_{n-2}\mathscr{L}_{n-4}\cdots\mathscr{L}_{-(n-4)}\mathscr{L}_{-(n-2)}E$, where E is the relative fundamental solution for $\mathscr{C}_{n+1,0}$ given above.

References

1. Folland, G.B., Stein, E.M.: Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group. Comm. Pure and Appl. Math. **27**, 429–522 (1974)

2. Geller, D.: Some results in H^p theory for the Heisenberg group. Duke Math. J. **47**, 365–390 (1980)
3. Geller, D.: The Laplacian and the Kohn Laplacian for the sphere. J. Diff. Geom. **15**, 417–435 (1980)
4. Graham, C.R.: The Dirichlet problem for the Bergman Laplacian, I and II. Comm. Partial Differential Equations **8**, 433–476 and 563–641 (1983)
5. Greiner, P.C., Kohn, J.J., Stein, E.M.: Necessary and sufficient conditions for solvability of the Lewy equation. Proc. Nat. Acad. Sci. **72**, 3287–3289 (1975)
6. Laville, G.: Fonctions pluriharmoniques et solution fondamentale d'un opérateur du 4ᵉ ordre. Bull. Soc. Math. France **101**, 305–317 (1977)

Received December 23, 1983

Math. Z. 187, 305–315 (1984)

Mathematische
Zeitschrift

Processus abéliens associés à un semi-groupe

Denis Feyel

Equipe d'analyse, Université Paris VI, 4 place Jussieu, 75230 Paris Cedex 05, France

Introduction

Il y a quelque temps, Akcoglu et Krengel ont introduit la notion de processus additif associé à un semi-groupe de contractions $\geqq 0$ de $\mathbb{L}^1$. L'auteur a d'autre part introduit les «familles résolues» attachées à une résolvante $\geqq 0$ à contraction dans $\mathbb{L}^1$. Il se trouve que ces deux notions se correspondent exactement par la transformation de Carson-Laplace.

Dans le présent travail, on reprend et étudie plus systématiquement ces notions (processus additifs et familles résolues nommées ici processus abéliens). Ces processus forment deux espaces complètement réticulés isomorphes par transformation de Carson-Laplace. On démontre les deux théorèmes ergodiques ponctuels: le théorème local permet de caractériser simplement la bande des processus «absolument continus» et la bande des processus «singuliers».

Pour étendre ces résultats au cas d'un semi-groupe complexe, on est amené à considérer le «module linéaire» d'un semi-groupe (notion dûe à Kipnis), et le «module linéaire» d'un processus abélien ou additif, étendant une notion de Chacon et Krengel [2]. Puis, à l'aide d'un théorème d'extension cylindrique de semi-groupes démontré dans [6], on se ramène au cas $\geqq 0$, et l'on montre qu'il y a encore décomposition unique en parties absolument continue et singulière, malgré l'absence formelle de structure réticulée: on les calcule facilement à l'aide du théorème ergodique local.

Signalons que dans la démonstration du théorème 9 a), toutes les intégrales écrites sont à prendre au sens de Bochner, bien que l'on puisse leur donner un sens «trajectoire par trajectoire», point de vue qui n'a pas été considéré.

I. Cas positif

Soit X un espace muni d'une mesure τ σ-finie. On considère une famille résolvante positive à contraction dans $\mathbb{L}^1$, c'est à dire une famille $(V_\lambda)_{\lambda>0}$ d'opérateurs positifs de $\mathbb{L}^1$ vérifiant:

a) $\|\lambda V_\lambda\| \leqq 1$ pour tout $\lambda > 0$,

b) $V_\lambda - V_\mu = -(\lambda-\mu) V_\lambda V_\mu$ pour tous $\lambda, \mu > 0$ (équation résolvante).

La théorie élémentaire des équations différentielles nous apprend que l'on peut remplacer la condition b) par l'une ou l'autre des conditions b′) ou b″):

b′) $\frac{d}{d\lambda} V_\lambda = -V_\lambda^2$ (dérivée au sens de la norme),

b″) $V_\lambda = \int_\lambda^\infty V_t^2\, dt$ (intégrale de Bochner).

Remarquons que les conditions a) et b) impliquent manifestement la relation $V_1 = \sum_{n \geqq 1} V_n V_{n+1}$ où la série converge normalement: nous utiliserons plus loin cette relation

Si P est un opérateur $\geqq 0$ de $\mathbb{L}^1$, P est prolongeable en pseudo-noyau, c'est à dire vérifiant $P(\sum_{n\geqq 1} f_n) = \sum_{n \geqq 1} Pf_n$ pour toute série $(f_n)_{n\geqq 1}$ de fonctions mesurables $\geqq 0$, intégrables ou non. Notons ∞ la fonction constante $\equiv +\infty$. La fonction $P(\infty)$ ne peut prendre que les valeurs 0 ou $+\infty$, et est donc de la forme ∞_E, où ∞_E désigne la fonction valant $+\infty$ sur E et 0 ailleurs. Si $f>0$ est intégrable, on a alors $E = \{Pf > 0\}$. On dira que P est *propre* si le complémentaire de E est τ-négligeable.

Si les V_λ ne sont pas propres, la fonction $\infty_E = V_\lambda(\infty)$ est indépendante de λ, en effet, pour $f \geqq 0$, la relation $V_\lambda f = V_\mu(f + (\mu-\lambda) V_\lambda f)$ subsiste pour $\mu > \lambda$ grâce au théorème de convergence monotone donc $V_\lambda(\infty) = V_\mu(\infty)$. De même, $V_\lambda(\infty_E)$ est de la forme ∞_A, et pour $\mu > \lambda$: $V_\lambda(\infty_E) = V_\mu(\infty_E + (\mu-\lambda) V_\lambda(\infty_E))$ $= V_\mu(\infty_E + \infty_A) = V_\mu(\infty_E)$ car $\infty_A + \infty_E = \infty_E$. Nous allons voir que $A = E$ presque sûrement: la relation $V_1 f = \sum_{n\geqq 1} V_n V_{n+1} f$ vaut aussi pour $f \geqq 0$ par le théorème de convergence monotone, donc:

$$\infty_E = V_1(\infty) = \sum_{n\geqq 1} V_n V_{n+1}(\infty) = \sum_{n \geqq 1} V_n(\infty_E) = \sum_{n\geqq 1} \infty_A = \infty_A.$$

L'ensemble E, ou «partie conservative locale» vérifie donc $V_\lambda(\infty_E) = \infty_E$ preque sûrement pour tout $\lambda > 0$.

Si l'on remplace alors la mesure τ par la mesure $\mu = 1_E \tau$ et chaque opérateur V_λ par $W_\lambda = V_\lambda I_E$, on constate que $(W_\lambda)_{\lambda>0}$ est une famille résolvante $\geqq 0$ à contraction dans $\mathbb{L}^1(\mu)$ et propre, i.e. vérifie:

c) $W_\lambda(\infty) = \infty$ μ-presque partout pour tout $\lambda > 0$.

Dans cet article, nous ferons toujours l'hypothèse de propriété pour la résolvante $(V_\lambda)_{\lambda>0}$ elle-même, relativement à la mesure τ: ce qui précède montre qu'on n'y perd pas beaucoup en généralité.

1. Definition. Un processus (u_λ) est dit abélien si:

a) $\sup_\lambda \|\lambda u_\lambda\| < \infty$ (norme de $\mathbb{L}^1$),

b) $u_\lambda - u_\mu = -(\lambda - \mu) V_\lambda u_\mu$ pour tous $\lambda, \mu > 0$.

On voit facilement comme plus haut, que $\lambda \mapsto u_\lambda$ est une fonction analytique à valeurs dans $\mathbb{L}^1$, et que l'on peut remplacer b) par b′) ou b″):

b′) $\dfrac{du_\lambda}{d\lambda} = -V_\lambda u_\lambda$,

b″) $u_\lambda = \int_0^\infty V_t u_t \, dt$.

Le processus abélien est dit *absolument continu* s'il est de la forme $u_\lambda = V_\lambda f$ avec $f \in \mathbb{L}^1$.

2. Théorème. *L'espace vectoriel $\mathscr{A}$ des processus abéliens est complètement réticulé en ordre naturel. Le sous-espace $\mathscr{A}_c$ des processus absolument continus est une bande dans $\mathscr{A}$, et l'on a les équivalences:*

a) *le processus (u_λ) est absolument continu,*
b) *la famille $(\lambda u_\lambda)_{\lambda \geqq 1}$ est faiblement relativement compacte dans $\mathbb{L}^1$,*
c) *λu_λ converge fortement quand λ tend vers $+\infty$.*

Démonstration. $\mathscr{A}$ est évidemment un espace vectoriel ordonné. On va introduire la «module linéaire» d'un élément de $\mathscr{A}$. Posons par récurrence $v_{\lambda,0} = |u_\lambda|$, $v_{\lambda,n+1} = \int_\lambda^\infty V_t v_{t,n} \, dt$.

La récurrence est possible car on a successivement

$$\|v_{\lambda,n+1}\| \leqq \int_\lambda^\infty M \, dt/t^2 \leqq M/\lambda, \qquad \text{avec } M = \sup_\lambda \|\lambda u_\lambda\|$$

et

$$v_{\lambda,0} = |u_\lambda| \leqq \int_\lambda^\infty V_t |u_t| \, dt = \int_\lambda^\infty V_t v_{t,0} \, dt = v_{\lambda,1}$$

puis par récurrence $v_{\lambda,n} \leqq v_{\lambda,n+1}$. Ainsi la suite $v_{\lambda,n}$ converge en croissant vers un processus v_λ. Pour chaque λ la convergence de $v_{\lambda,n}$ a lieu dans $\mathbb{L}^1$, donc la fonction $\lambda \mapsto v_\lambda$ est mesurable et l'on a $v_\lambda = \int_\lambda^\infty V_t v_t \, dt$ par le théorème de Lebesgue, et $\|\lambda v_\lambda\| \leqq M$. Donc (v_λ) est un processus abélien, et c'est par construction même le plus petit qui majore à la fois (u_λ) et $(-u_\lambda)$, c'est donc son module au sens de $\mathscr{A}$, et $\mathscr{A}$ est réticulé.

Si maintenent des (u_λ^i) sont $\geqq 0$ et vont en décroissant, notons u_λ la borne inférieure au sens de $\mathbb{L}^1$ de la famille u_λ^i (λ fixé) c'est aussi sa limite dans $\mathbb{L}^1$, par suite (u_λ) est un processus abélien borne inférieure de la famille (u_λ^i) au sens de $\mathscr{A}$, et $\mathscr{A}$ est complètement réticulé.

Montrons maintenant les équivalences; rappelons d'abord [5] qu'il existe un projecteur $P_0 \geqq 0$ contractant et propre de l'espace $\mathbb{L}^1$, que l'on a $P_0 f = \lim_{\lambda \to \infty} \lambda V_\lambda f$ (limite forte) pour toute $f \in \mathbb{L}^1$.

a)⇒c) En effet, $\lambda V_\lambda f$ converge fortement vers $P_0 f$.

c)⇒b) Si on la complète par sa limite pour $\lambda \to \infty$, elle est même fortement compacte comme image continue de $[1, +\infty]$.

b)⇒a) Soit f une valeur d'adhérence faible de (λu_λ) pour $\lambda \to \infty$.

Pour $\mu>0$, $V_\mu f$ est valeur d'adhérence faible de $V_\mu(\lambda u_\lambda)=\lambda V_\lambda u_\mu$, donc $V_\mu f = P_0 u_\mu$. Or on a $u_\mu - u_\lambda = (\lambda-\mu) V_\lambda u_\mu = P_0(u_\mu - u_\lambda)$ pour tout $\lambda>0$ car $P_0 V_\lambda = V_\lambda P_0 = V_\lambda$ ([5] p. 147). Faisant tendre λ vers $+\infty$, on trouve $u_\mu = P_0 u_\mu = V_\mu f$.

Montrons enfin que $\mathscr{A}_c$ est une bande: si l'on a $|u_\lambda| \leqq V_\lambda f$ pour tout $\lambda>0$, avec $f\in\mathbb{L}^1$, la famille $(\lambda u_\lambda)_{\lambda\geqq 1}$ est faiblement relativement compacte, donc $(u_\lambda)\in\mathscr{A}_c$ par le critère b).

Si des $(V_\lambda f_i)$ vont en croissant dans $\mathscr{A}_c$ et restent dominés par $(V_\lambda g)$ (g intégrable), leur borne supérieure dans $\mathscr{A}$ est majorée par $(V_\lambda g)$ et appartient donc à $\mathscr{A}_c$ par le critère b).

On notera $\mathscr{A}_s$ la bande étrangère à $\mathscr{A}_c$: les éléments de $\mathscr{A}_s$ sont dits «singuliers».

Précisant maintenant une notion de [5], on pose:

3. Definition. Soit $(f_i)_{i\in I}$ une famille de fonctions mesurables à valeurs dans $\overline{\mathbb{R}} = [-\infty, +\infty]$, filtrée par un filtre $\mathscr{F}$ sur I. On appelle limite supérieure essentielle, ou limite supérieure pour l'ordre des (classes de) fonctions mesurables, selon $\mathscr{F}$:

$$\operatorname{Lim\,ess\,sup} f_i = \operatorname*{EssInf}_{\alpha\in\mathscr{F}} \operatorname*{EssSup}_{i\in\alpha} f_i.$$

C'est une (classe de) fonction mesurable $\leqq\infty$. Si $\mathscr{F}$ est à base dénombrable, il existe un sous-ensemble dénombrable J de I, rencontrant tous les éléments de $\mathscr{F}$ et tel que la limite essentielle supérieure des f_i selon $\mathscr{F}$ soit la limite supérieure ordinaire des $(f_j)_{j\in J}$ selon la trace de $\mathscr{F}$ sur J.

On dit que les f_i convergent essentiellement pour l'ordre ou essentiellement selon $\mathscr{F}$ si la limite essentielle supérieure coincide avec la limite essentielle inférieure. Si $\mathscr{F}$ est à base dénombrable, il suffit que pour toute partie J de I, J dénombrable et appartenant à la grille de $\mathscr{F}$ (i.e. rencontrant tout élément de $\mathscr{F}$), la famille $(f_j)_{j\in J}$ converge presque partout au sens ordinaire selon le filtre trace de $\mathscr{F}$ sur J.

4. Théorème. *Soit (u_λ) un processus abélien associé à une résolvante propre. On a les propriétés suivantes:*

a) *λu_λ converge essentiellement quand $\lambda\to\infty$ vers une $f\in\mathbb{L}^1$,*

b) *le processus $(V_\lambda f)$ est la partie absolument continue de (u_λ),*

c) *si $\phi\in\mathbb{L}^1$, $\phi>0$, le rapport $u_\lambda/V_\lambda\phi$ converge essentiellement quand $\lambda\to 0$ vers une limite finie presque partout.*

Démonstration. le point a) est déjà acquis lorsque (u_λ) est un processus abélien absolument continu en vertu théorème précédent et du théorème de [5] p. 150. En général, reprenons la démonstration de [7]: si Λ est une partie finie de $]0,\infty[$, on a le lemme maximal:

$$\operatorname*{Lim\,inf}_{t\to\infty} \int_{E_\Lambda} t\,u_t\,d\tau \geqq 0 \quad \text{avec} \quad E_\Lambda = \bigcup_{\lambda\in\Lambda} \{u_\lambda>0\}.$$

Supposons donc que le processus (u_λ) soit $\geqq 0$, et soit $\psi\in\mathbb{L}^1$, $\psi < \operatorname*{Lim\,ess\,sup}_{\lambda\to\infty} \lambda u_\lambda$: le lemme maximal devient:

$$\int_{E_\Lambda} P_0\psi\,d\tau \leqq \operatorname*{Lim\,inf}_{t\to\infty} \int_{E_\Lambda} t\,u_t\,d\tau \quad \text{avec} \quad E_\Lambda = \bigcup_{\lambda\in n} \{u_\lambda > V_\lambda\psi\}.$$

On obtient à la limite lorsque Λ augmente suivant le filtre des sections de l'ensemble des parties finies de $]0,\infty[$: $X = \operatorname{Ess}_\Lambda \sup E_\Lambda$ d'après la définition de ψ, et par le théorème de convergence monotone:

$$\int P_0\,\psi\,d\tau \leqq \operatorname*{Lim\,inf}_{t\to\infty} \int t\,u_t\,d\tau.$$

Soit μ une mesure absolument continue par rapport à τ et excessive, c'est à dire vérifiant $\mu\,t\,V_t \leqq \mu$ pour tout $t>0$. La résolvante (V_λ) admet une extension en résolvante $\geqq 0$, propre et à contraction dans $\mathbb{L}^1(\mu)$, on a donc aussi:

$$\int P_0\,\psi\,d\mu \leqq \operatorname*{Lim\,inf}_{t\to\infty} \int t\,u_t\,d\mu.$$

Soit $\alpha>0$, posons $V_\lambda^\alpha = V_{\lambda+\alpha}$: (V_λ^α) est une résolvante $\geqq 0$, propre et à contraction dans $\mathbb{L}^1(\tau)$, l'opérateur terminal P_0 est inchangé. Posons $u_\lambda^\alpha = u_{\lambda+\alpha}$: (u_λ^α) est un processus abélien associé à (V_λ^α), l'on a: $\psi < \operatorname*{Lim\,ess\,sup}_{\lambda\to\infty} \lambda\,u_\lambda^\alpha$ car la différence $(\lambda+\alpha)\,u_{\lambda+\alpha} - \lambda\,u_{\lambda+\alpha} = \alpha\,u_{\lambda+\alpha}$ converge vers 0 dans $\mathbb{L}^1$ en décroissant donc aussi essentiellement. Soit $\mu \geqq 0$ une mesure de la forme $\mu(f) = \int h\,V_\alpha f\,d\tau$ où $h \in \mathbb{L}^\infty(\tau)$. On constate que μ est excessive par rapport à la résolvante (V_λ^α), donc:

$$\int h\,V_\alpha\,P_0\,\psi\,d\tau \leqq \operatorname*{Lim\,inf}_{t\to\infty} \int t\,h\,V_\alpha\,u_{t+\alpha}\,d\tau \leqq \int h\,u_\alpha\,d\tau.$$

Cela vaut pour toute $h \geqq 0$ et bornée, et tout $\alpha>0$, donc:

$$V_\alpha\,\psi \leqq u_\alpha \qquad \text{pour tout } \alpha>0.$$

Remarquons que l'on peut appliquer cela à $\psi_\varepsilon(x) = \sup(\psi(x), -\varepsilon)$ pour tout $\varepsilon>0$. Quand $\varepsilon\to 0$, ψ_ε converge vers $\psi^+(x) = \sup(\psi(x), 0)$ donc:

$$V_\alpha\,\psi^+ \leqq u_\alpha \qquad \text{pour tout } \alpha>0.$$

Supposons donc que (u_λ) soit un processus singulier, cela entraine que $V_\alpha\,\psi^+ = 0$ pour tout $\alpha>0$, donc $P_0\,\psi^+ = 0$, puis $\psi^+ = 0$ et $\psi \leqq 0$ car on a toujours $\int P_0 f\,d\tau = \int f\,d\tau$ pour toute $f \in \mathbb{L}^1(\tau)$ comme il résulte de [5] p. 147. On en déduit finalement $\lim_{\lambda\to\infty} \operatorname{ess\,sup} \lambda\,u_\lambda = 0$ ce qui achève la démonstration du a).

b) on a $u_\lambda = v_\lambda + V_\lambda\,g$ avec $g \in \mathbb{L}^1$, où (v_λ) (resp. $(V_\lambda\,g)$) est la partie singulière (resp. absolument continue) de (u_λ). On a donc $f = \operatorname*{Lim\,ess}_{\lambda\to\infty} \lambda\,u_\lambda = P_0\,g$, donc $V_\lambda f = V_\lambda\,g$ pour tout $\lambda>0$.

c) Pour $\alpha>0$, on a $u_\lambda = u_\alpha + (\alpha-\lambda)\,V_\lambda\,u_\alpha$; on est donc ramené au théorème abélien de Chacon-Ornstein, et l'on trouve

$$\operatorname*{Lim\,ess}_{\lambda\to 0} u_\lambda / V_\lambda\,\phi = u_\alpha / V\phi + Q(\alpha\,u_\alpha)$$

où $Qf = \operatorname*{Lim\,ess}_{\lambda\to 0} V_\lambda f / V_\lambda\,\phi$ est l'opérateur limite de Chacon-Ornstein et où $V\phi = \operatorname*{Ess\,sup}_{\lambda>0} V_\lambda\,\phi$.

5. Corollaire. *On suppose la résolvante dissipative* (*c'est à dire que* $V\phi<\infty$ τ-*presque sûrement pour* ϕ *intégrable* >0). *Alors tout processus abélien* $\geqq 0$ *s'écrit sous la forme* $u_\lambda = u - \lambda V_\lambda u$, *où le «potentiel»* u *est fini presque partout.*

Démonstration. On sait qu'en ce cas, on $Qf = Vf/V\phi$ pour $f \geqq 0$, donc

$$g = \operatorname*{Lim\,ess}_{\substack{\lambda\to 0\\ \lambda\to\infty}} u_\lambda / V_\lambda \phi = (u_\alpha + V(\alpha u_\alpha))/V\phi,$$

et par suite, $u_\alpha + V(\alpha u_\alpha) = gV\phi$ est indépendant de $\alpha>0$. En posant $u = \operatorname*{Ess\,sup}_{\alpha>0} u_\alpha$ on obtient $u \leqq gV\phi$, donc $u<\infty$ presque sûrement. Ensuite, pour $0<\alpha<\lambda$, on a $u_\alpha = u_\lambda + (\lambda-\alpha) V_\lambda u_\alpha = u_\lambda + (\lambda-\alpha) V_\alpha u_\lambda$, donc quand α tend vers 0 en décroissant, on obtient par le théorème de convergence monotone puisque les V_λ sont des pseudo-noyaux:

$$u = u_\lambda + \lambda V_\lambda u = u_\lambda + \lambda V u_\lambda.$$

Cela montre aussi que $g = u/V\phi$.

6. Exemple. Soit τ la mesure de Lebesgue sur $\mathbb{R}^m$ ($m \geqq 1$), soit μ une mesure bornée, et soit G^μ_λ son λ-potentiel bessélien (cf. [4]): c'est l'unique fonction intégrable u vérifiant $\lambda u - \Delta u = \mu$ au sens des distributions. Si μ est de la forme $g\tau$ avec $g \in \mathbb{L}^1$, posons $V_\lambda g = G^\mu_\lambda$. Les V_λ forment une famille résolvante vérifiant nos hypothèses et même $P_0 g = g$ pour toute $g \in \mathbb{L}^1$. Pour toute mesure μ bornée, le processus (G^μ_λ) est abélien. Donc λG^μ_λ converge essentiellement quand $\lambda \to \infty$ vers une $f \in \mathbb{L}^1$. Si $\mu \geqq 0$, on a $f \geqq 0$, et f est la plus grande fonction $\geqq 0$ vérifiant $V_\lambda f \leqq G^\mu_\lambda$ pour tout $\lambda>0$; c'est donc la densité de la partie absolument continue de μ par rapport à τ. Cela subsiste évidemment si μ n'est pas $\geqq 0$.

Dans le cas dissipatif ($m \geqq 3$), on retrouve le fait que le potentiel newtonien G^μ_0 est fini presque partout, et que $G^\mu_\lambda = G^\mu_0 - \lambda V_\lambda G^\mu_0$ (remarquer aussi que G^μ_0 n'est pas intégrable).

II. Application aux processus additifs

Soit (V_λ) une résolvante $\geqq 0$ propre et à contraction dans $\mathbb{L}^1(\tau)$. On sait qu'elle est transformée de Laplace d'un semi-groupe $(P_t)_{t\geqq 0}$ fortement continu et contractant dans $\mathbb{L}^1$, dont l'opérateur initial P_0 est précisément l'opérateur terminal de la résolvante. Les P_t sont des opérateurs $\geqq 0$, et P_0 est propre. D'ailleurs la donnée de la résolvante est équivalente à la donnée du semi-groupe (cf. par exemple [5].

Reprenant la définition d'Akcoglu et Krengel, on dit qu'un processus $(g_t)_{t\geqq 0}$ est additif s'il vérifie:

a) $g_{t+s} = g_t + P_t g_s$ pour tous $t, s \geqq 0$,
b) $\sup_{t>0} \|g_t/t\| \leqq M$ (norme dans $\mathbb{L}^1$).

Les opérateurs de Cesaro sont définis par $A_t f = \int_0^t P_s f\, ds$ (intégrale de Bochner) pour $f \in \mathbb{L}^1$ et $t \geqq 0$. On constate que pour toute $f \in \mathbb{L}^1$, le processus $(A_t f)$ est additif.

Si (g_t) est un processus additif, $t \mapsto g_t$ est lipschitzienne à valeurs dans $\mathbb{L}^1$.

On notera $\mathscr{C}$ l'espace vectoriel des processus additifs réels, et $\mathscr{C}_c$ le sous-espace des processus de la forme $(A_t f)$ qui sont dits «absolument continus».

7. Théorème. *La transformation de Carson-Laplace* $u_\lambda = \lambda \int_0^\infty e^{-\lambda t} g_t \, dt$ *(intégrale de Bochner) est un isomorphisme bi-croissant de* $\mathscr{C}$ *sur* $\mathscr{A}$. *Il s'ensuit que* $\mathscr{C}$ *est complètement réticulé. L'image de* $\mathscr{C}_c$ *est exactement* $\mathscr{A}_c$, *donc* $\mathscr{C}_c$ *est une bande dans* $\mathscr{C}$.

Démonstration. On a d'abord

$$\|u_\lambda\| \leqq M \lambda \int_0^\infty t e^{-\lambda t} dt \leqq M/\lambda \quad \text{où} \quad M = \sup_{t>0} \|g_t/t\|.$$

Notons Δ le quart de plan $t, s \geqq 0$, on a successivement par des transformations élémentaires et le théorème de Fubini:

$$V_\lambda u_\lambda = \lambda \iint_\Delta e^{-\lambda(t+s)} P_t g_s \, dt \, ds,$$

$$V_\lambda u_\lambda = \lambda \iint_\Delta e^{-\lambda(t+s)} g_{t+s} \, dt \, ds - \lambda \iint_\Delta e^{-\lambda(t+s)} g_t \, dt \, ds,$$

$$\int_\lambda^\infty V_t u_t \, dt = \int_0^\infty g_t \, dt \int_\lambda^\infty (\mu t - 1) e^{-\mu t} d\mu = \int_0^\infty e^{-\lambda t} g_t \, dt = u_\lambda.$$

Donc (u_λ) est un processus abélien, et $u_\lambda \geqq 0$ dès que les g_t le sont. Si le processus (u_λ) est nul, les g_t sont presque tous nuls donc tous nuls puisque $t \mapsto g_t$ est lipschitzienne, ainsi la transformation est injective.

Montrons qu'elle est surjective: soit (u_λ) un processus abélien, approchons le par les processus (u_λ^α) définis par $u_\lambda^\alpha = \alpha V_\alpha u_\lambda = \alpha V_\lambda u_\alpha$. Posons $g_t^\alpha = A_t(\alpha u_\alpha)$. Les processus (g_t^α) sont additifs et admettent les (u_λ^α) comme transformées de Carson-Laplace car la résolvante est transformée de Laplace du semi-groupe. On a successivement les relations:

$$\alpha u_\alpha - \beta u_\beta = (\alpha - \beta)(I - \beta V_\beta) u_\alpha,$$

$$g_t^\alpha - g_t^\beta = (\alpha - \beta) A_t (I - \beta V_\beta) u_\alpha = (\alpha - \beta)(I - P_t) V_\beta u_\alpha,$$

$$\|g_t^\alpha - g_t^\beta\| \leqq 2M |\alpha - \beta| / \alpha\beta = 2M |\alpha^{-1} - \beta^{-1}| \quad \text{avec} \quad M = \sup_\lambda \|\lambda u_\lambda\|.$$

Quand $\alpha \to \infty$, g_t^α converge donc vers une limite $g_t \in \mathbb{L}^1$. On a $\|g_t\| \leqq Mt$ et la relation d'additivité par passage à la limite, donc (g_t) est un processus additif. Le théorème de Lebesgue de convergence dominée permet de calculer la transformée de Carson-Laplace de (g_t), c'est la limite des $u_\lambda^\alpha = \alpha V_\alpha u_\lambda$ quand $\alpha \to \infty$, c'est donc $u_\lambda = P_0 u_\lambda$.

Il reste à voir que cet isomorphisme est bi-croissant, et que $\mathscr{C}_c$ correspond à $\mathscr{A}_c$. Or, il est clair que si les u_λ sont $\geqq 0$, les g_t^α le sont aussi, puis les g_t par passage à la limite. Enfin on a bien $V_\lambda f = \lambda \int_0^\infty e^{-\lambda t} A_t f \, dt$ pour toute $f \in \mathbb{L}^1$.

On note bien sûr $\mathscr{C}_s$ la bande étrangère à $\mathscr{C}_c$: c'est la bande des processus additifs singuliers, et elle correspond à $\mathscr{A}_s$.

Remarque. Akcoglu et Krengel ont montré directement que $\mathscr{C}$ était réticulé (cf. [1]).

8. Théorème. *Les conditions suivantes sont équivalentes* (P_0 *propre*):

a) (g_t) *est absolument continu,*
b) *la famille* $(g_t/t)_{0<t\leqq 1}$ *est faiblement relativement compact dans* $\mathbb{L}^1$,
c) g_t/t *converge fortement quand* $t\to 0$.

Démonstration.

a)$\Rightarrow$c) $A_t f/t$ converge fortement vers $P_0 f$ par le théorème ergodique de convergence en moyenne.

c)$\Rightarrow$b) Car $t\mapsto g_t/t$ est alors continument prolongeable à $[0,1]$.

b)$\Rightarrow$a) Soit f une valeur d'adhérence faible de g_t/t. Alors $A_t u_\lambda/t = V_\lambda(g_t/t)$ admet $V_\lambda f$ comme valeur d'adhérence, mais aussi u_λ par le théorème ergodique fort rappelé au a)$\Rightarrow$c), donc (u_λ) est un processus abélien absolument continu: le processus additif (g_t) l'est aussi par le théorème 7.

Nous retrouverons maintenant des résultats de convergence pour les processus additifs dûs à Akcoglu et Krengel:

9. Théorème. *On suppose toujours la résolvante (ou le semi-groupe) propre. Soit* (g_t) *un processus additif.*

a) *Quand* $t\to 0$, g_t/t *converge essentiellement vers une limite* $f\in\mathbb{L}^1$ *qui est la même que celle de* λu_λ *pour* $\lambda\to\infty$.
b) $(A_t f)$ *est la partie absolument continue du processus* (g_t).
c) *Quand* $t\to\infty$, *et si* $\phi>0$, $\phi\in\mathbb{L}^1$, $g_t/A_t\phi$ *converge essentiellement vers une limite finie presque partout.*

Démonstration. a) on peut se limiter au cas où les $g_t\geqq 0$. Alors le processus g_t qui est lipschitzien et croissant peut être redéfini en un processus encore noté g_t, mais dont presque toute trajectoire est une fonction croissante de t. Cela permet d'écrire pour presque tout x:

$$u_\lambda(x)=\int_0^\infty e^{-\lambda t}\,dg_t(x)$$

(intégrale de Stieltjès). Mais alors notre résultat n'est qu'un cas particulier du théorème taubérien de Karamata ([11] p. 192).

b) résulte du fait que $(V_\lambda f)$ est la partie absolument continue du processus abélien (u_λ).

c) On écrit $g_t=A_t u_\alpha+(I-P_t)u_\alpha$ pour $\alpha>0$: on est donc ramené au théorème de Chacon-Ornstein; la seule difficulté est de montrer la convergence de $P_t u_\alpha/A_t\phi$ vers 0, mais c'est précisément le lemme de Chacon-Ornstein [3]. On constate alors que la limite est la même que celle du théorème 4c).

10. Corollaire. *Si le semi-groupe est dissipatif, tout processus additif* $\geqq 0$ *s'écrit sous la forme* $g_t = u - P_t u$, *où le «potentiel»* u *est fini presque partout.*

Démonstration. On écrit $u_\lambda + \lambda V_\lambda g_t = g_t + P_t u_\lambda$ pour $\lambda, t > 0$. Le raisonnement est alors le même qu'au corollaire 5, on trouve $u = g_t + P_t u$, et u est la même fonction qu'au corollaire 5.

III. Cas complexe

Nous allons étendre ces résultats au cas d'une résolvante complexe à contraction et propre dans $\mathbb{L}^1_{\mathbb{C}}$. Il y a priori une difficulté dûe à l'absence de structure réticulée: nous verrons que ce n'est qu'une difficulté apparente.

Un opérateur $P \geqq 0$ de $\mathbb{L}^1$ majore un opérateur complexe T de $\mathbb{L}^1_{\mathbb{C}}$ si l'on a $P|f| \geqq |Tf|$ pour toute $f \in \mathbb{L}^1_{\mathbb{C}}$. Selon Chacon et Krengel, tout opérateur continu T de $\mathbb{L}^1_{\mathbb{C}}$ possède un «module linéaire», i.e. un plus petit opérateur $P \geqq 0$ majorant T, on le note $|T|$, et l'on a $\||T|\| = \|T\|$.

Si (R_λ) est une résolvante complexe à contraction dans $\mathbb{L}^1_{\mathbb{C}}$, il existe selon Kipnis [10] une plus petite résolvante $\geqq 0$ et à contraction dans $\mathbb{L}^1$ et majorant (R_λ). On peut d'ailleurs l'obtenir par récurrence (cf. [6]) en posant

$$V_{\lambda,0} = |R_\lambda|, \qquad V_{\lambda,n+1} = \int_\lambda^\infty V_{t,n}^2 \, dt,$$

alors $V_{\lambda,n}$ converge en croissant et fortement vers la résolvante V_λ que nous appellerons le module linéaire de la résolvante (R_λ).

Si les V_λ sont propres, ce qui équivaut à la même hypothèse pour les opérateurs $|R_\lambda|$, cette notion de module linéaire est conservée par la transformation de Laplace; autrement dit, (V_λ) est transformée de Laplace d'un semi-groupe propre (P_t) que est le plus petit majorant le semi-groupe complexe (T_t) dont (R_λ) est transformée de Laplace ([6], p. 75). Noter que cela implique que (R_λ) est une transformée de Laplace, ce qui n'est pas tout à fait évident.

Il y a beaucoup mieux: soit σ la mesure de Lebesgue du cercle unité Γ du plan complexe. Selon [6] p. 71 et 75, et toujours si la résolvante (V_λ) est propre, il existe un semi-groupe (Q_t) de contractions $\geqq 0$ de l'espace $\mathbb{L}^1(\sigma \otimes \tau)$ vérifiant les conditions suivantes:

a) (Q_t) est propre
b) chaque Q_t est invariant par les rotations de Γ,
c) $Q_t(1 \otimes f) = 1 \otimes P_t f$ pour toute $f \in \mathbb{L}^1$, $t \geqq 0$,
d) $Q_t(Z \otimes f) = Z \otimes T_t f$ pour toute $f \in \mathbb{L}^1_{\mathbb{C}}$, $t \geqq 0$.

On rappelle que Z désigne l'injection canonique de Γ dans $\mathbb{C}$, et la fonction $\varphi \otimes f$ est définie par $(\varphi \otimes f)(z,x) = \varphi(z) f(x)$.

La famille résolvante (W_λ) transformée de Laplace du semi-groupe (Q_t) vérifie évidemment les propriétés analogues, c'est d'ailleurs grâce à elle que l'on construit (Q_t).

On définit les notions de processus abéliens ou additifs complexes associés au semi-groupe (T_t) de la même manière que les processus réels associés à (P_t).

Cela étant, tout processus abélien (resp. additif) complexe possède lui aussi un «module linéaire», de manière précise:

11. Théorème. *Les $|R_\lambda|$ sont supposés propres. La transformation de Carson-Laplace $u_\lambda = \lambda \int_0^\infty e^{-\lambda t} g_t\, dt$ échange les processus additifs et les processus abéliens.*

Si (g_t) (resp. (u_λ)) est un processus additif (resp. abélien) associé à (T_t), il existe un plus petit processus additif $\geqq 0$ (resp. abélien $\geqq 0$) associé à (P_t), et tel que $h_t \geqq |g_t|$ pour tout t (resp. $v_\lambda \geqq |u_\lambda|$ pour tout λ). De plus

$$\sup_t \|h_t/t\| = \sup_t \|g_t/t\|, \qquad \sup_\lambda \|\lambda v_\lambda\| = \sup_\lambda \|\lambda u_\lambda\|.$$

Enfin, la transformation de Carson-Laplace échange les modules linéaires.

Démonstration. Si (u_λ) est abélien associé à (T_t), il est clair que $w_\lambda = Z \otimes u_\lambda$ est abélien complexe associé au semi-groupe (Q_t). Nous avons vu au théorème 2 la construction du module linéaire d'un processus abélien réel; la construction convient aussi dans le cas présent. On pose $s_{\lambda,0} = |w_\lambda|$, $s_{\lambda,n+1} = \int_\lambda^\infty W_t s_{t,n}\, dt$. Les $s_{\lambda,n}$ convergent en croissant vers un processus $s_\lambda \geqq 0$, le plus petit associé à (Q_t) et majorant $w_\lambda = Z \otimes u_\lambda$. Par récurrence, on voit que $s_{\lambda,n}$ est de la forme $1 \otimes v_{\lambda,n}$, donc $s_\lambda = 1 \otimes v_\lambda$ où (v_λ) est un processus abélien associé à (P_t) et c'est clairement le plus petit majorant (u_λ).

On transporte immédiatement cela aux processus additifs en utisant l'isomorphisme bi-croissant de $\mathscr{C}((Q_t))$ sur $\mathscr{A}((Q_t))$, et cela démontre entièrement le théorème.

12. Théorème. *Dans les mêmes hypothèses, soit (u_λ) un processus abélien associé à (T_t), et soit (v_λ) son module linéaire (associé à (P_t))*

a) *λu_λ converge essentiellement quand $\lambda \to \infty$, vers une $f \in \mathbb{L}^1_{\mathbb{C}}$,*

b) *λv_λ converge essentiellement quand $\lambda \to \infty$,* vers $|f|$,

c) *$(R_\lambda f)$ est la partie «absolument continue» de (u_λ),*

d) *$(u_\lambda - R_\lambda f)$ est singulier ainsi que son module linéaire qui vaut d'ailleurs $(v_\lambda - V_\lambda |f|)$,*

e) *si $\phi > 0$ et $\phi \in \mathbb{L}^1$, le rapport $u_\lambda / V_\lambda \cdot \phi$ converge essentiellement quand $\lambda \to 0$ vers une limite finie presque partout.*

Les conditions suivantes sont équivalentes:

f) *(u_λ) est absolument continu,*

g) *(v_λ) est absolument continu,*

h) *λu_λ converge fortement quand $\lambda \to \infty$,*

i) *la famille $(\lambda u_\lambda)_{\lambda \geqq 1}$ est faiblement relativement compacte dans $\mathbb{L}^1_{\mathbb{C}}$.*

Démonstration. En passant au processus $w_\lambda = Z \otimes u_\lambda$, on voit immédiatement les propriétés a) et e) et l'équivalence des propriétés f), g), h), i) car en ce cas w_λ est de la forme $w_\lambda = W_\lambda (Z \otimes f) = Z \otimes R_\lambda f$.

Pour montrer les autres propriétés, écrivons $u_\lambda = R_\lambda f + u'_\lambda$. On a donc $w_\lambda = Z \otimes u_\lambda = Z \otimes R_\lambda f + Z \otimes u'_\lambda = W_\lambda(Z \otimes f) + Z \otimes u'_\lambda$. Quand $\lambda \to \infty$, λw_λ converge essentiellement vers $Z \otimes f$, relative à la mesure $\sigma \otimes \tau$, donc la décomposition de (w_λ) ci-dessus n'est autre que celle de (w_λ) dans le complexifié de $\mathscr{A}$ $((Q_t))$: $Z \otimes R_\lambda f = W_\lambda(Z \otimes f)$ est la partie absolument continue et $Z \otimes u'_\lambda$ la partie singulière. En prenant les modules dans $\mathscr{A}$ $((Q_t))$, on obtient l'égalité car les deux termes sont étrangers, donc:

$$1 \otimes v_\lambda = 1 \otimes V_\lambda f + 1 \otimes v'_\lambda$$

où (v'_λ) est singulier relatif à (P_t) et est le module linéaire de (u'_λ). Cela prouve b), c), d).

17. *Remarque*. En posant $B_t f = \int_0^t T_s f\, ds$, on peut évidemment énoncer un théorème exactement analogue pour les processus additifs, avec une démonstration en tout point semblable.

18. *Complements*. Les résultats de ce travail sont utilisés dans les deux notes aux Comptes rendus de Paris, sur la convergence des processus sur-abéliens et sur-additifs, [8] et [9].

Bibliographie

1. Akcoglu, A., Krengel, U.: A differentiation theorem for additive processes. Math. Z. **163**, 199–210 (1978)
2. Chacon, R.V., Krengel, U.: Linear modulus of a linear operator. Proc. Amer. Math. Soc. **15**, 553–559 (1964)
3. Chacon, R.V., Ornstein, D.: A general ergodic theorem. Illinois J. Math. **4**, 153–160 (1960)
4. Donoghue, W.F.: Distributions and Fourier Transforms. New York-London: Academic Press 1969
5. Feyel, D.: Théorèmes de convergence presque-sûre, existence de semi-groupes. Advances in Math. **34**, 145–162 (1979)
6. Feyel, D.: Espaces complètement réticulés de pseudo-noyaux. Applications aux résolvantes et semi-groupes complexes. Sém. de th. du potentiel, Paris n°3. Lecture Notes in Math. **681**. Berlin-Heidelberg-New York: Springer 1978
7. Feyel, D.: Compléments sur la convergence presque-sûre des familles résolues. Sém. de th. du potentiel, Paris n°4. Lecture Notes in Math. **713**. Berlin-Heidelberg-New York: Springer 1979
8. Feyel, D.: Convergence locale des processus sur-abéliens et sur-additifs. C.R. Acad. Sci. Paris Ser. I, Math. **295**, 301 (1982)
9. Feyel, D.: Théorème ergodique ponctuel pour processus sur-abéliens. C.R. Acad. Sci. Paris Ser. I, Math. **295**, 357
10. Kipnis, C.: Majoration des semi-groupes de contractions de $\mathbb{L}^1$, applications. Ann. Inst. H. Poincaré, Sect. B 10, 369–384 (1974)
11. Widder, D.: The Laplace transform. Princeton: Princeton Univ. Press 1941 8° édition, 1972

Reçu le 26 janvier 1983; reçu en forme finale le 28 avril 1983

Math. Z. 187, 317-323 (1984)

Mathematische Zeitschrift

Analytic Functions of Topological Proper Contractions*

Mau-Hsiang Shih[1] and Kok-Keong Tan[2]

[1] Department of Mathematics, Chung Yuan University, Chung Li, Taiwan, Republic of China

[2] Department of Mathematics, Statistics and Computing Science, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 4H8

§ 1. Introduction and Notations

Throughout this paper, $\mathscr{H}$ will denote a complex Hilbert space with inner product $(\cdot,\cdot)$ and the associated norm $\|\cdot\|$. Let A be an operator (i.e. a bounded linear transformation) on $\mathscr{H}$. For a complex-valued function f on the open unit disk $\Delta := \{z \in C: |z|<1\}$, $f(A)$ will denote the operator on $\mathscr{H}$ defined by the usual Riesz-Dunford integral ([2], p. 568):

$$f(A) := \frac{1}{2\pi\sqrt{-1}} \int_{\mathscr{C}} f(z)(zI-A)^{-1}\, dz,$$

where $\mathscr{C}$ is a positively oriented simple closed rectifiable contour lying in Δ and encircling the spectrum of A. Here and elsewhere, I stands for the identity operator on $\mathscr{H}$. The limit used in defining the integral is taken in the norm topology (i.e. the uniform topology) of operators. Let $H(\Delta)$ denote the class of all complex-valued functions analytic on Δ and $B(\Delta)$ denote the class of all complex-valued functions f analytic on Δ such that $f(\Delta) \subset \Delta$. As usual, an operator A on $\mathscr{H}$ is strictly positive if $(Ax, x) > 0$ for all non-zero $x \in \mathscr{H}$. In a recent paper, Fan establishes the following theorem ([3], Theorem 1) on von Neumann's theory of spectral set ([5]) by examining the behavior at the extreme points of the Montel space of contractive holomorphic functions on Δ:

Theorem. (Ky Fan). *If A is a proper contraction on $\mathscr{H}$, i.e. $\|A\| < 1$, then $\|f(A)\| < 1$ for every $f \in B(\Delta)$.*

Let $r_\sigma(A)$ denote the spectral radius of A, $\sigma(A)$ the spectrum of A and Δ_σ the class of all operators A on $\mathscr{H}$ with $r_\sigma(A) < 1$. We denote by $SG[A_1, \ldots, A_N]$ the semigroup (under composition) generated by $\{A_1, \ldots, A_N\}$, where $A_1, \ldots, A_N$ are operators on $\mathscr{H}$.

---

* This work was partially supported by NSERC of Canada under grant No. A-8096. This work was done while the author was on sabbatical leave from Dalhousie University, Canada to the Institute of Mathematics, Academia Sinica, Taiwan, Republic of China

In this paper, a simultaneous re-Hilbert norm theorem is used to show "If $A_1, \ldots, A_N \in \Delta_\sigma$ are pairwise commutative, and suppose $\mathscr{S} := SG[A_1, \ldots, A_N]$, then there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that $\|f(A)\|_* < 1$ for all $A \in \mathscr{S}$ and for every $f \in B(\Delta)$." A theorem on maximum principle for operators and a theorem on convex univalent functions are also obtained.

§ 2. A Simultaneous re-Hilbert Norm Theorem

The proofs of our theorems rely heavily on the following result:

Theorem 2.1. *Let $A_1, \ldots, A_N$ be pairwise commutative operators on $\mathscr{H}$. Then corresponding to each $\varepsilon > 0$, there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that*

$$\|A_i\|_* \leqq r_\sigma(A_i) + \varepsilon \quad \textit{for all } i = 1, \ldots, N.$$

Proof. By Gel'fand's spectral radius formula, corresponding to each $\varepsilon > 0$, there exists a positive integer m such that

$$\|A_i^m\|^{1/m} \leqq r_\sigma(A_i) + \varepsilon \quad \text{for all } i = 1, \ldots, N.$$

For each $i = 1, \ldots, N$, let $t_i := r_\sigma(A_i) + \varepsilon$. For each $x, y \in \mathscr{H}$, define

$$(x, y)_* := \sum_{j_1=0}^{m-1} \sum_{j_2=0}^{m-1} \cdots \sum_{j_N=0}^{m-1} \frac{(A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N} x, A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N} y)}{t_1^{2j_1} t_2^{2j_2} \ldots t_N^{2j_N}}.$$

Then $(\cdot, \cdot)_*$ is an inner product on $\mathscr{H}$. Set $\|x\|_* := \sqrt{(x, x)_*}$ for all $x \in \mathscr{H}$. We have

$$\|x\| \leqq \|x\|_* \leqq \left(\sum_{j_1=0}^{m-1} \sum_{j_2=0}^{m-1} \cdots \sum_{j_N=0}^{m-1} \frac{\|A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N}\|^2}{t_1^{2j_1} t_2^{2j_2} \ldots t_N^{2j_N}} \right)^{\frac{1}{2}} \|x\|$$

for all $x \in \mathscr{H}$. Thus $\|\cdot\|_*$ is equivalent to $\|\cdot\|$. It remains to show that $\|A_i\|_* \leqq t_i$ for each $i = 1, \ldots, N$. Indeed, for each $x \in \mathscr{H}$,

$$\begin{aligned}
\|A_1 x\|_*^2 &= \sum_{j_1=0}^{m-1} \sum_{j_2=0}^{m-1} \cdots \sum_{j_N=0}^{m-1} \frac{\|A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N} A_1 x\|^2}{t_1^{2j_1} t_2^{2j_2} \ldots t_N^{2j_N}} \\
&\leqq \sum_{j_1=0}^{m-2} \sum_{j_2=0}^{m-1} \cdots \sum_{j_N=0}^{m-1} \frac{\|A_1^{j_1+1} A_2^{j_2} \ldots A_N^{j_N} x\|^2}{t_1^{2j_1} t_2^{2j_2} \ldots t_N^{2j_N}} \\
&\quad + \sum_{j_2=0}^{m-1} \cdots \sum_{j_N=0}^{m-1} \frac{\|A_1^m\|^2 \, \|A_2^{j_2} \ldots A_N^{j_N} x\|^2}{t_1^{2(m-1)} t_2^{2j_2} \ldots t_N^{2j_N}} \\
&\leqq t_1^2 \sum_{j_1=0}^{m-1} \sum_{j_2=0}^{m-1} \cdots \sum_{j_N=0}^{m-1} \frac{\|A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N} x\|^2}{t_1^{2j_1} t_2^{2j_2} \ldots t_N^{2j_N}} \\
&= t_1^2 \|x\|_*^2;
\end{aligned}$$

thus $\|A_1\|_* \leqq t_1$. By the commutativity of A_i's and the symmetry of the above proof, we have the desired conclusion. □

Example 2.2. Let $\mathscr{H}$ be a 2-dimensional Hilbert space and

$$A_n := \begin{pmatrix} 0 & n \\ 0 & 0 \end{pmatrix} \quad \text{for } n=1,2,\ldots.$$

Then $r_\sigma(A_i)=0$ and $A_i A_j = A_j A_i$ for all $i, j=1,2,\ldots$. Since $\|A_n\| = n \to \infty$, Theorem 2.1 cannot be extended to a countably infinite family of commuting operators.

Example 2.3. Let $\mathscr{H}$ be a 2-dimensional Hilbert space,

$$A_1 := \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 := \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}.$$

Then $r_\sigma(A_i)=0$ for $i=1,2$, and A_1 and A_2 are non-commutative. Let $x := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$; we see that $\|(A_2 A_1)^n x\| \to \infty$ as $n \to \infty$. Therefore the commutativity assumption in Theorem 2.1 is essential.

Remark 2.4. By Theorem 2.1, $A \in \Delta_\sigma$ iff A is a topological proper contraction, i.e., there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that $\|A\|_* < 1$.

§ 3. Commuting Topological Proper Contractions

Theorem 3.1. *Let $A_1, \ldots, A_N \in \Delta_\sigma$ be pairwise commutative, and $\mathscr{S} := SG[A_1, \ldots, A_N]$. Then there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that*

$$\|f(A)\|_* < 1 \quad \text{for all } A \in \mathscr{S} \text{ and for every } f \in B(\Delta).$$

Proof. By Theorem 2.1, there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that $\|A_i\|_* < 1$ for all $i=1, \ldots, N$. If $A \in \mathscr{S}$, then

$$A = A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N},$$

where $j_1, j_2, \ldots, j_N \in \{0, 1, 2, \ldots\}$ such that $j_1 + j_2 + \ldots + j_N \geqq 1$; thus

$$\|A\|_* = \|A_1^{j_1} A_2^{j_2} \ldots A_N^{j_N}\|_* \leqq \|A_1\|_*^{j_1} \|A_2\|_*^{j_2} \ldots \|A_N\|_*^{j_N} < 1.$$

By Fan's theorem, $\|f(A)\|_* < 1$ for every $f \in B(\Delta)$. This completes the proof. □

The following result is an infinite dimensional version of Stein's theorem ([10]). Another infinite dimensional version of Stein's theorem was obtained by Shih ([9], Theorem 8).

Lemma 3.2. *Let A be a Riesz operator* (see [6], p. 361) *on $\mathscr{H}$. Suppose that there exists a strictly positive operator B on $\mathscr{H}$ such that $B - A^* B A$ is also strictly positive. Then $r_\sigma(A) < 1$.*

Proof. As A is a Riesz operator, for each λ not in the resolvent of A, λ is an eigenvalue of A ([6], Theorem 25.5.6). Let λ be any eigenvalue of A and $x \neq 0$ a corresponding eigenvector. Since $((B-A^*BA)x, x)>0$, it follows that

$$(BAx, Ax)<(Bx, x).$$

Therefore $|\lambda|^2(Bx, x)<(Bx, x)$, so that $|\lambda|^2<1$ by strict positivity of B. As $\sigma(A)$ is compact, by Gel'fand's theorem, there exists a $\lambda_0\in\sigma(A)$ such that $r_\sigma(A) = \sup_{\lambda\in\sigma(A)} |\lambda|=|\lambda_0|<1$. This completes the proof. □

Theorem 3.3. *Let $A_1, \ldots, A_N \in \Delta_\sigma$ be pairwise commutative Riesz operators, and $\mathcal{S}:=SG[A_1, \ldots, A_N]$. Suppose that for each $i=1, \ldots, N$, there exists a strictly positive operator B_i on $\mathcal{H}$ such that $B_i - A_i^* B_i A_i$ is also strictly positive. Then there exists a Hilbert norm $\|\cdot\|_*$ on $\mathcal{H}$ equivalent to $\|\cdot\|$ such that*

$$\|f(A)\|_*<1 \quad \textit{for all } A\in\mathcal{S} \textit{ and for every } f\in B(\Delta).$$

Proof. By Lemma 3.2, $r_\sigma(A_i)<1$ for all $i=1, \ldots, N$, applying Theorem 3.1, the conclusion follows. □

Theorem 3.3 is reformulated by the Cayley transform as follows:

Theorem 3.4. *Let $A_1, \ldots, A_N \in \Delta_\sigma$ be pairwise commutative Riesz operators, and $\mathcal{S}:=SG[A_1, \ldots, A_N]$. Suppose that for each $i=1, \ldots, N$, there exists a strictly positive operator B_i on $\mathcal{H}$ such that $A_i^* B_i + B_i A_i$ is also strictly positive. Then there exists a Hilbert norm $\|\cdot\|_*$ on $\mathcal{H}$ equivalent to $\|\cdot\|$ such that for each $f\in B(\Delta)$,*

$$\|f((I-A)(I+A)^{-1})\|_*<1 \quad \textit{for all } A\in\mathcal{S}.$$

The following operator inequality may also be proved by using Theorem 3.1 and a result of Fan ([3], Example 2).

Proposition 3.5. *Let $f\in H(\Delta)$ be univalent on Δ such that $f(0)=0$ and $f'(0)=1$. Suppose that $A_1, \ldots, A_N\in\Delta_\sigma$ are pairwise commutative, and $\mathcal{S}:=SG[A_1, \ldots, A_N]$. Then there exists a Hilbert norm $\|\cdot\|_*$ on $\mathcal{H}$ equivalent to $\|\cdot\|$ such that*

(a) $$\|f(A)\|_* \leqq \frac{\|A\|_*}{(1-\|A\|_*)^2} \quad \textit{for all } A\in\mathcal{S};$$

(b) $$\|f'(A)\|_* \leqq \frac{1+\|A\|_*}{(1-\|A\|_*)^3} \quad \textit{for all } A\in\mathcal{S}.$$

§ 4. Maximum Principle for Operators

Theorem 4.1 below contains two versions of maximum principle for operators, of which part (a) was already proved in [3].

Theorem 4.1. *Let $f\in H(\Delta)$. For $0<r<1$, let*

$$M(r):=\max_{|z|=r}|f(z)|.$$

Then:

(a) $$M(r)=\max_{\|A\|\leqq r}\|f(A)\|,$$

where the maximum is taken over all operators A of norm $\leqq r$ on $\mathscr{H}$;

(b)
$$M(r) = \max_{r_\sigma(A) \leqq r} r_\sigma(f(A)),$$

where the maximum is taken over all operators A on $\mathscr{H}$ such that $r_\sigma(A) \leqq r$.

Proof. (a) See [3], Theorem 3.

(b) If f is a constant function, then the assertion follows. So we may assume that f is not a constant function. Consider a fixed r such that $0<r<1$. We have $M(r)>0$. The function g defined by

$$g(z) := f(rz)/M(r)$$

is analytic on $|z| < \frac{1}{r}$. By the maximum modulus principle, $|g(z)|<1$ for all $z \in \Delta$. If $A \in \Delta_\sigma$, by Theorem 2.1, there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that $\|A\|_* < 1$; thus by Fan's theorem, $\|g(A)\|_* < 1$ so that $r_\sigma(g(A)) < 1$. Therefore $r_\sigma(f(rA)) \leqq M(r)$ for all $A \in \Delta_\sigma$. This is equivalent to saying that $r_\sigma(f(A)) \leqq M(r)$ for all $r_\sigma(A) \leqq r$. Thus $\max\limits_{r_\sigma(A) \leqq r} r_\sigma(f(A)) \leqq M(r)$, where the maximum is taken over all operators on $\mathscr{H}$ such that $r_\sigma(A) \leqq r$. On the other hand, if $|z_0| = r$ and $|f(z_0)| = M(r)$, then $r_\sigma(z_0 I) = |z_0| = r$ and $r_\sigma(f(z_0)) = |f(z_0)| = M(r)$. This completes the proof of part (b). □

Theorem 4.1(b) can be proved by applying the spectral mapping theorem ([2], p. 569), so that Theorem 4.1(b) remains valid in an arbitrary Banach space. Note that 4.1(a) does not, in general, hold in a Banach space, e.g. the example in ([8], p. 436) can be employed to illustrate that.

Theorem 4.1 may be applied to establish certain operator inequalities, of which the following two examples serve as illustrations.

By employing Koebe's distortion theorem ([7], p. 21) and Theorem 4.1(b), we have

Example 4.2. Let $f \in H(\Delta)$ be univalent on Δ such that $f(0)=0$ and $f'(0)=1$. Then

(a)
$$r_\sigma(f(A)) \leqq \frac{r_\sigma(A)}{(1-r_\sigma(A))^2} \quad \text{for all } A \in \Delta_\sigma;$$

(b)
$$r_\sigma(f'(A)) \leqq \frac{1+r_\sigma(A)}{(1-r_\sigma(A))^3} \quad \text{for all } A \in \Delta_\sigma.$$

By employing Hadamard's three-circle theorem ([4], p. 410) and Theorem 4.1(b), we have

Example 4.3. Let $f \in H(\Delta)$. For $0<r<1$, let $M(r) := \max\limits_{r_\sigma(A) \leqq r} r_\sigma(f(A))$, where the maximum is taken over all operators on $\mathscr{H}$ such that $r_\sigma(A) \leqq r$. If $0<r_1 \leqq r \leqq r_2 < 1$, then

$$\begin{vmatrix} \log M(r) & \log r & 1 \\ \log M(r_1) & \log r_1 & 1 \\ \log M(r_2) & \log r_2 & 1 \end{vmatrix} \geqq 0.$$

§ 5. Convex Univalent Functions

We begin with the following observation:

Proposition 5.1. *Let $f \in H(\Delta)$ be univalent on Δ such that $f(0)=0$ and $f'(0)=1$. If f is starlike (i.e. the image $f(\Delta)$ is a starlike set with respect to the origin), then the set of all $f(A)$, where A runs through Δ_σ, is a starlike set of of operators.*

Proof. Observe that $\{f(A) : A \in \Delta_\sigma\} = \bigcup \{f(A) : \|A\|_* < 1\}$, where the union is taken over all Hilbert norms $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$. Applying Theorem 2.1 and a result of Fan ([3], Corollary 7), the conclusion follows. □

Let $\mathscr{F}_\sigma := \{\mathbb{F} : \mathbb{F} \subset \Delta_\sigma$ and $\mathbb{F}$ is commuting$\}$. Then $\mathscr{F}_\sigma$ is partially ordered under set inclusion such that every chain in $\mathscr{F}_\sigma$ has an upper bound in $\mathscr{F}_\sigma$. Thus by the Hausdorff maximal principle, there exists a maximal element in $\mathscr{F}_\sigma$. In fact for each operator A on $\mathscr{H}$, $\mathbb{F}_A := \{B \in \Delta_\sigma : AB = BA\}$ is maximal in $\mathscr{F}_\sigma$.

Let $K(\Delta)$ denote the class of all convex univalent functions, i.e. those functions f in $H(\Delta)$ with $f(0)=0$, $f'(0)=1$, which are univalent on Δ such that the image $f(\Delta)$ is a convex set. A theorem of Brickman-MacGregor-Wilken ([1]) states that the extreme points of the set $K(\Delta)$ (in the vector space $H(\Delta)$) are precisely the functions of the form $h_\eta(z) = z(1-\eta z)^{-1}$, where η is a complex number with $|\eta| = 1$. Corresponding to Proposition 5.1, we have the following result:

Theorem 5.2. *Let η be a complex number with $|\eta|=1$, and let $h_\eta \in K(\Delta)$ be the function $h_\eta(z) := z(1-\eta z)^{-1}$. If $\mathbb{F}$ is a maximal element in $\mathscr{F}_\sigma$, then $\{h_\eta(A) : A \in \mathbb{F}\}$ is a convex set of operators.*

Proof. When A runs through $\mathbb{F}$ and η is kept fixed, ηA also varies through $\mathbb{F}$, since $r_\sigma(\eta A) = r_\sigma(A)$. Furthermore

$$h_\eta(A) = \eta^{-1} h_1(\eta A).$$

Therefore in proving the theorem, it suffices to consider the case $\eta = 1$. Let $A, B \in \mathbb{F}$, and $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$. Let

$$D := \alpha A(I-A)^{-1} + \beta B(I-B)^{-1}.$$

Since $A, B \in \mathbb{F}$, by Theorem 2.1, there exists a Hilbert norm $\|\cdot\|_*$ on $\mathscr{H}$ equivalent to $\|\cdot\|$ such that $\|A\|_* < 1$ and $\|B\|_* < 1$. First we write

$$\begin{aligned} I + D &= (I-B)^{-1}\{(I-B)(I-A) + \alpha(I-B)A + \beta B(I-A)\}(I-A)^{-1} \\ &= (I-B)^{-1}\{I - (\beta A + \alpha B)\}(I-A)^{-1}. \end{aligned}$$

Since $\|\beta A + \alpha B\|_* \leqq \beta\|A\|_* + \alpha\|B\|_* < 1$, $I - (\beta A + \alpha B)$ is invertible, so that $I + D$ is also invertible. Following the same line of proof in ([3], pp. 287–288), if we define $C := (I+D)^{-1} D$, we have

$$\|C\|_* < 1 \quad \text{and} \quad C(I-C)^{-1} = D.$$

It remains to show that $C \in \mathbb{F}$. As $\|C\|_* < 1$, $r_\sigma(C) < 1$. Let $S \in \mathbb{F}$; then $SA = AS$ and $SB = BS$, so that $SD = DS$. Hence $SC = CS$. Since $\mathbb{F}$ is maximal, $C \in \mathbb{F}$. This completes the proof. □

Example 5.3. Let $\mathscr{H}$ be an n-dimensional Hilbert space and

$$\mathbb{F} := \left\{ (b_{ij})_{1 \leqq i, j \leqq n} : b_{ij} := \begin{cases} a_1 & \text{if } i = j; \\ 0 & \text{if } i > j; \\ a_{j-i+1} & \text{if } i < j; \end{cases} \right.$$
$$\left. i, j = 1, \ldots, n, a_1, \ldots, a_n \in \mathbb{C} \text{ and } |a_1| < 1 \right\}.$$

Then $\mathbb{F}$ is maximal in $\mathscr{F}_\sigma$. Let $h(z) := z(1-z)^{-1}$. By a direct computation,

$$h(\mathbb{F}) = \left\{ (c_{i,j})_{1 \leqq i, j \leqq n} : c_{ij} = \begin{cases} a_1 & \text{if } i = j; \\ 0 & \text{if } i > j; \\ \dfrac{1}{1-a_1} \left(\sum_{k=1}^{j-i} a_k a_{j-i+k} \right) + a_{j-i+1} & \text{if } i < j; \end{cases} \right.$$
$$\left. i, j = 1, \ldots, n, a_1, \ldots, a_n \in C \text{ and } |a_1| < 1 \right\}.$$

According to Theorem 5.2, $h(\mathbb{F})$ is convex in $\mathscr{M}_n$, the vector space of all $n \times n$ matrices over $\mathbb{C}$.

References

1. Brickman, L., MacGregor, T.H., Wilken, D.R.: Convex hull of some classical families of univalent functions. Trans. Amer. Math. Soc., **156**, 91–107 (1971)
2. Dunford, N., Schwartz, J.T.: Linear Operators, Part I: General Theory. New York: Interscience 1958
3. Fan, K.: Analytic functions of a proper contraction. Math. Z. **160**, 275–290 (1978)
4. Hille, E.: Analytic Function Theory, II. Boston: Ginn 1962
5. Neumann, J. von: Eine Spektral Theorie für allgemeine Operatoren eines unitären Raumes. Math. Nachr. **4**, 258–281 (1950/51)
6. Pietsch, A.: Operator Ideals. Berlin: VEB Deutscher Verlag der Wissenschaften 1978
7. Pommerenke, V.P.: Univalent Functions. Göttingen: Vandenhoeck-Ruprecht 1975
8. Riesz, F., SZ.-Nagy, B.: Functional Analysis. New York: Ungar 1960
9. Shih, M.-H.: Fixed points for mappings majorized by real functions. Hokkaido Math. J. **9**, 18–35 (1980)
10. Stein, P.: Some general theorems on iterants. J. Res. Nat. Bur. Standards, **48**, 82–83 (1952)

Received August 24, 1983

Math. Z. 187, 325–333 (1984)

Mathematische Zeitschrift

Minimal Immersions Into Space Forms With Two Principal Curvatures

Luquesio P. Jorge[1] and Francesco Mercuri[2]

[1] Dep. de Matemática UFC, Fortaleza - Ceará - Brazil
[2] Dep. de Matemática, UNICAMP-Campinas - SP - Brazil

§ 1. Introduction

In [2] Chern, do Carmo and Kobayashi consider minimal immersions of an n-dimensional manifold M^n into the unit sphere S^{n+N} of the $(n+N+1)$-dimensional euclidean space $\mathbb{R}^{n+N+1}$ with the property that the second fundamental form has constant length $n/(2-1/N)$. (If M is compact, this is the smallest possible value for a non-totally geodesic minimal immersion as above.) They prove that locally M is either a piece of a Veronese surface ($n=2$, $N=2$) or the product of two spaces of dimensions m and $n-m$ and of constant curvatures n/m and $n/(n-m)$. They also prove the corresponding global result under a compactness assumption for M. The main feature of the situation is that the second fundamental form in any normal direction has two distinct eigenvalues and, in general, the respective eigenspaces will determine involutive distributions whose integral manifolds will give the local product structure. This leads naturally to the study of minimal immersions into spaces of constant curvature such that the second fundamental form has, in any normal direction, at most two distinct eigenvalues. This problem was studied by Otsuki in [9] for the case of hypersurfaces. He proves the following result (see also Remark (3.7)).

(1.1) **Theorem.** *If $x\colon M^n \to \overline{M}_c^{n+1}$, $n \geqq 3$, is a minimal immersion into a simply connected space of constant curvature c, such that the second fundamental form has two distinct eigenvalues of multiplicity m and $n-m$, and $1<m<n-1$. Then:*

(i) *$c>0$ and M^n is locally the riemannian product of two spheres of dimensions m and $n-m$ with curvatures n/m and $n/(n-m)$ respectively.*

(ii) *If M^n is closed and x an embedding, then M^n is globally the product of two spheres as in* (i).

(1.2) *Remark.* In the same paper Otsuki gives a general method to produce minimal hypersurfaces of S^{n+1} with two principal curvatures, one of them being simple. This situation was generalized by do Carmo and Dajczer in [3], where they show that a hypersurface (not necessarily minimal) of a space of constant curvature with two principal curvatures λ and μ, with μ simple and $\mu=\mu(\lambda)$, is a "rotation hypersurface".

The extension of the above mentioned result to higher codimension was considered by Matsuyama in [7]. He proves the following result:

(1.3) **Theorem.** *Let x: $M^n \to \overline{M}_c^{n+N}$ be a minimal immersion into a simply connected space of constant curvature $c \geqq 0$ such that the second fundamental form in any normal direction has at most two distinct eigenvalues and, if exactly two, each with multiplicity bigger than 1. Then the second fundamental form is parallel and its norm S is either zero or $n \leqq S \leqq n^2/4$. Moreover if $c=0$, $S=0$ and if $S=n$, M is either the product of two spheres, as in* (1.1) *part* (i), *or M is the complex projective plane (if it is complete).*

(1.4) *Remark.* The proof of Theorem (1.3) does not depend on the fact that $c \geqq 0$ but only on c being constant.

The aim of this paper is to give a rather complete classification of minimal immersions into space forms of the type considered above. Since (1.1) and (1.2) take care of the codimension one case, we state our result in the following form:

(1.5) **Theorem.** *Let x: $M^n \to \overline{M}_c^{n+N}$ be a minimal immersion into a simply connected space of constant curvature c such that the second fundamental form, in any normal direction, has at most two distinct eigenvalues. If x is substantial, $N>1$ and $n \geqq 3$ we have:*

(1) *$c>0$,*

(2) *M^n is an open part of a projective plane over the complex, quaternions or Cayley numbers,*

(3) *x is a standard embedding (in the sense of* [5]).

Moreover the standard embeddings of the spaces in (2) *verify the hypothesis of the theorem.*

(1.6) *Remark.* We observe a certain analogy with the result of Chern, do Carmo, Kobayashi quoted at the beginning: A minimal immersion with our condition on the second foundamental form is either a codimension one immersion (and, in general product of spheres by (1.1)) or a Veronese type surface (by (1.5)).

We thank M. Dajczer who brought the problem to our attention and for his helpful comments and the referee for several improvements.

§ 2. Notations and Preliminary Results

Let M^n be a n-dimensional riemannian manifold and x: $M^n \to \overline{M}^{n+N}$ an isometric immersion in a $(n+N)$-dimensional riemannian manifold (superscripts will denote dimensions and will be dropped when clear from the context). We will denote by ∇, $\bar{\nabla}$ the riemannian connections of M and $\overline{M}$ respectively, by νM, $\nabla^\perp$ the normal bundle of the immersion and the normal connection, and by α: $TM \otimes TM \to \nu M$ the second fundamental form. If $p \in M$, $\xi \in \nu_p M$ we will denote by A_ξ the Weingarten operator in the ξ direction, i.e. $\langle A_\xi X, Y\rangle = \langle \alpha(X, Y), \xi\rangle$ $\forall X, Y \in T_pM$. Finally the mean curvature vector will be denoted, as usual, by H and x is minimal if $H \equiv 0$, or equivalently, trace $A_\xi = 0$ $\forall \xi \in \nu M$.

We will be interested in minimal immersions x: $M \to \overline{M}_c^{n+N}$ where $\overline{M}_c^{n+N}$ is a simply connected space of constant sectional curvature c with the following additional condition:

$2PC$: For any normal vector $\xi \in \nu M$, A_ξ has at most two distinct eigenvalues.

A minimal immersion which verify $2PC$ will be called a *$2PC$-minimal immersion.*

(2.1) *Remark.* If x is a $2PC$-minimal immersion and $A_\xi \neq 0$ then A_ξ has exactly two distinct non zero eigenvalues, since trace $A_\xi = 0$.

The space

$$\begin{aligned} N_{1,p} &= \mathrm{Span}\{\alpha(X, X) | X \in T_pM\} \\ &= \{\xi \in \nu_pM | A_\xi = 0\}^\perp \end{aligned}$$

is called the *first normal space of* x at p. If x is a $2PC$-minimal immersion and $A_\xi \neq 0$ we will denote by $\lambda(\xi)$ and $\mu(\xi)$ the two distinct eigenvalues of A_ξ, by $l(\xi)$ and $m(\xi)$ their multiplicity and by $T_{\lambda(\xi)}$ and $T_{\mu(\xi)}$ the eigenspaces relative to $\lambda(\xi)$ and $\mu(\xi)$ respectively.

(2.2) **Lemma.** *Let x: $M^n \to \overline{M}^{n+N}$ be a $2PC$-minimal immersion. If* $\dim N_{1,p} > 1$, *$\xi \in N_{1,p}$ then $l(\xi) = m(\xi)$ and $\lambda(\xi) + \mu(\xi) = 0$.*

Proof. If $\dim N_{1,p} > 1$, given $\xi \in N_{1,p}$ there exists a smooth path β: $[0,1] \to N_{1,p}$ such that $\beta(0) = \xi$, $\beta(1) = -\xi$ and $\beta(t) \neq 0$ for all $t \in [0,1]$. It is then possible to find continuous functions $\lambda_1, \ldots, \lambda_n$: $[0,1] \to \mathbb{R}$ such that the set $\{\lambda_i(t)\}$ is the set of repeated eigenvalues of $A_{\beta(t)}$ (see [6] Theorem (6.8) page 122). By (2.1), $\lambda_i(t) \neq 0$ for all $t \in [0,1]$, $i = 1, \ldots, n$, and therefore the multiplicity of the positive eigenvalue of $A_{\beta(t)}$ is constant. But $A_{\beta(1)} = -A_{\beta(0)}$ and so the multiplicity of the negative eigenvalue of $A_{\beta(0)}$ is equal to the multiplicity of the positive one and this proves the lemma.

(2.3) *Remark.* If follows easily from (2.2) that if ξ, $\eta \in N_{1,p}$, $\dim N_{1,p} > 1$, then

(a) the minimal polynomial of A_ξ is $t^2 + \lambda(\xi)\mu(\xi)$.

(b) If trace $A_\xi A_\eta = 0$ then $A_\xi A_\eta + A_\eta A_\xi = 0$. In particular, $A_\eta(T_{\lambda(\xi)}) \subset T_{\mu(\xi)}$ and $A_\eta(T_{\mu(\xi)}) \subset T_{\lambda(\xi)}$.

This allows a certain simplification of Matsuyama's proof of the parallelism of the second fundamental form (see appendix).

(2.4) **Proposition.** *Let x: $M^n \to \overline{M}_c^{n+N}$ be a $2PC$-minimal immersion, $n \geq 3$ such that $x(M^n)$ is not contained in any $(n+1)$-dimensional totally geodesic submanifold of $\overline{M}_c^{n+N}$. Then the dimension of N_1 is a constant bigger than* 1.

Proof. The proof will be based on Theorem (1.3) and on the following two facts:

(2.4.1) A minimal immersion in a space of constant curvature is real analytic.

So if M is connected and $U \subset M$ is a non empty open set we have:

(a) If $x(U)$ is contained in a totally geodesic submanifold of $\bar{M}_c^{n+N}$, so is all of $x(M^n)$.

(b) If the second fundamental form is parallel in U, it is parallel on all of M.

(2.4.2) (See [11] Lemma 2). If $x\colon M^n \to \bar{M}_c^{n+N}$ is an isometric immersion, and $U \subset M$ is a non empty connected open set such that the first normal space of x is 1-dimensional in U and for all $p \in U$ there exists a 2-plane $\sigma \subset T_p M$ with sectional curvature $K(\sigma) \neq c$, then $x(U)$ is contained in a totally geodesic $(n+1)$-dimensional submanifold of $\bar{M}_c^{n+N}$.

Let us now prove the proposition. If at some point $p \in M$ $\dim N_{1,p} > 1$, there exists an open neighborhood U of p where the same happens. By (2.2) and $n \geqq 3$, A_ξ does not have simple eigenvalues and so (1.3) implies that the second fundamental form is parallel in U and therefore, by (2.4.1), it is parallel on all of M. In particular $\dim N_1$ is a constant greater than 1. Let us suppose now $\dim N_{1,p} \leqq 1$. If there exists an open set $U \subset M$, $U \neq \phi$, such that $\dim N_1 = 0$ on U, then $x|U$ is totally geodesic and by analiticity x is a totally geodesic immersion. Suppose now that there exists an open connected non empty set $U \subseteq M$ such that $\dim N_1 = 1$ in U. Let ξ be a unit normal vector that spans N_1. Since $A_\xi \neq 0$ there exist two orthonormal eigenvectors of A_ξ, X, Y, with eigenvalues λ, μ and $\lambda \neq 0 \neq \mu$. By the Gauss equation the sectional curvature of the plane $\sigma = \operatorname{span}\{X, Y\}$ is $K(\sigma) = \lambda\mu + c \neq c$ and so, by (2.4.2), $x(U)$ is contained in an $(n+1)$-dimensional totally geodesic submanifold of $\bar{M}_c^{n+N}$ and by (2.4.1) the same is true for $x(M)$.

(2.5) **Corollary.** *Let $x\colon M^n \to \bar{M}_c^{n+N}$ be a $2PC$-minimal immersion, $n \geqq 3$, such that $x(M^n)$ is not contained in any $(n+1)$-dimensional totally geodesic submanifold of $\bar{M}_c^{n+N}$. Then $c > 0$.*

Proof. By (2.4) and (2.2) x verifies the hypothesis of (1.3). In this case Matsuyama calculated, using Simons' formula, the norm S of the second fundamental form obtaining, for a suitable choice of an orthonormal frame $\xi_1, \ldots, \xi_N \in \nu M$,

$$S\left(cn - \frac{4}{n} S\right) + \left(\frac{4}{n} - 1\right) \sum_{i=1}^{N} (\operatorname{trace} A_{\xi_i}^2)^2 = 0. \tag{2.5.1}$$

Now clearly (2.5.1) implies that if $c \leqq 0$ then $S = 0$ and therefore x is totally geodesic.

(2.6) *Remark.* Again using (2.3), we can get a simplification of the argument used to prove (2.5.1).

§ 3. $2PC$-Minimal Immersions in S^{n+N}

In view of the results of § 2 we will restrict ourselves to $2PC$-minimal immersions $x\colon M^n \to S^{n+N}$. Let $i\colon S^{n+N} \to \mathbb{R}^{n+N+1}$ be the standard inclusion. We will denote by D the riemannian connection of $\mathbb{R}^{n+N+1}$ and by $\bar{\alpha}$, $\bar{A}$, $\bar{\nu}$ the second fundamental form, the Weingarten operator and the normal bundle of the immersion $\bar{x}=i\circ x\colon M^n\to\mathbb{R}^{n+N+1}$. If X, Y are tangent vectors of M at p, since $D_X Y-\bar{\nabla}_X Y=-\langle X, Y\rangle\, x(p)$, we have:

$$(3.1)\qquad \bar{\alpha}(X, Y)=\alpha(X, Y)-\langle X, Y\rangle\, x(p).$$

If $\xi\in\bar{\nu}$ we have:

$$(3.2)\qquad \bar{A}_\xi=A_{\xi^T}-\langle x(p), \xi\rangle\, I, \quad \text{where } \xi^T=\xi-\langle x(p), \xi\rangle\, x(p).$$

The following is a well known fact:

(3.3) The second fundamental form of x is parallel if and only if the second fundamental form of $\bar{x}$ is parallel.

Isometric immersions in $\mathbb{R}^n$ with parallel second fundamental form are rather well known (see [10]). We will comment briefly some of the properties of such immersions that we will use to prove Theorem (1.5). For a large class of compact symmetric spaces which includes the classical spaces and few of the exceptional ones, it is possible to construct rigid embeddings in $\mathbb{R}^K$ such that the second fundamental form is parallel. Those embeddings, called the standard embeddings have, among others, the following properties:

(3.4) After a possible normalization of the metric, the image of a standard embedding is contained in the unit sphere and the induced embedding in that sphere is minimal.

(3.5) The standard embedding $i_M\colon M\to\mathbb{R}^K$ is tight, i.e., if ξ is a regular value of the Gauss normal map then the height function $h_\xi\colon M\to\mathbb{R}$, given by

$$h_\xi(p)=\langle i_M(p), \xi\rangle,$$

has the minimum number of critical points compatible, in the sense of Morse theory, with the topological structure of M.

In [5], Theorem 1, D. Ferus proves that if $x\colon M^n\to S^{n+N}$ is a minimal immersion, non-totally geodesic and with parallel second fundamental form, then $x(M)$ is an open part of the image of a standard embedding. So the last part of Theorem (1.5) follows easily from the following result.

(3.6) **Theorem.** *Let $i_M\colon M^n\to S^{n+N}$ be a standard embedding of a compact symmetric space such that i_M is a $2PC$-minimal immersion and $i_M(M)$ is not contained in any $(n+1)$-dimensional totally geodesic submanifold of S^{n+N}. Then i_M is the standard embedding of a projective plane over the complex, quaternions or Cayley numbers. Moreover the standard embedding of those spaces verify the hypothesis of the theorem.*

Proof. The idea of the proof is to study the homology of a compact $2PC$-minimal submanifold of S^{n+N} and using classification results for symmetric spaces, to check which symmetric spaces have a homological structure compatible with our result.

(3.6.1) *Assertion.* Let i_M be a $2PC$-minimal immersion of a compact riemannian manifold M^n into S^{n+N} such that the first normal space at each point has dimension at least two. If $\xi \in S^{n+N}$ is a regular value of the Gauss normal map the function $h_\xi\colon M^n \to \mathbb{R}$

$$h_\xi(p) = \langle i_M(p), \xi \rangle$$

has only non-degenerate critical points and the index of such a points is either 0, n or $n/2$.

Proof. It is well known (see [10]) that h_ξ has only non-degenerate critical points if ξ is a regular value of the Gauss map. Moreover the hessian of h_ξ at a critical point is given by the matrix of $\bar{A}_\xi$. If $\{X_1, \ldots, X_n\}$ is an orthonormal basis which diagonalizes A_{ξ^T} we have, from (3.2),

$$\langle \bar{A}_\xi X_i, X_j \rangle = \delta_{ij}(\lambda_i - \langle x(p), \xi \rangle).$$

From (2.2) up to a reordering of the X_i's, we have

$$\lambda_1 = \ldots = \lambda_{n/2} = -\lambda_{n/2+1} = \ldots = -\lambda_n$$

and the conclusion follows easily.

The above assertion and Morse theory applied to h_ξ imply that M has the homotopy type of a CW-complex with only cells in dimensions 0, n or $n/2$. If $n > 2$, we have the following immediate consequences:

(3.6.2) M is simply connected,

(3.6.3) $H_i(M, \mathbb{Z}) = 0$ if $i \neq 0$, n, $n/2$.

(3.6.4) *Assertion.* M is irreducible.

In fact, if $M = M_1^{n_1} \times M_2^{n_2}$, from (3.6.3) and the Künneth formula for the homology of a product, $M_i^{n_i}$ is homeomorphic to S^{n_i} and hence isometric, since M is symmetric (see [14]). It follows that $M = S^{n/2} \times S^{n/2}$ and i_M being the standard embedding has image contained in $S^{n+1} \subset \mathbb{R}^{n+N+1}$.

(3.6.5) If $H_{n/2}(M, \mathbb{Z}) = 0$, then M is homotopy equivalent to S^n and therefore i_M is a totally geodesic immersion of S^n into S^{n+N} (see [14]).

(3.6.6) If M^n is Lie group, $n \leqq 6$, then $M = S^3 \times S^3$ as follows from (3.6.2), (3.6.3) and the classification of compact Lie groups. Moreover $H^3(M, \mathbb{R}) \neq 0$ for n-dimensional Lie groups with $n \geqq 3$ and this, if $n > 6$, contradict (3.6.3) (see [13]).

After all this information we take a look at the list of the Poincaré polynomials of compact irreducible symmetric spaces which are not Lie groups, computed in [12], and we easily recognize that the only such spaces with a homology structure compatible with (3.6.2), (3.6.3), etc., are the projective planes over the complex, quaternions and Cayley numbers.

Let now $i_M\colon M^n \to S^{n+N} \subseteq \mathbb{R}^{n+N+1}$ be the standard embedding of one of the mentioned planes.

Claim. If the Weingarten operator $\bar{A}_\eta$ with respect to a unit normal vector $\eta \in \bar{\nu}_p$ is non-singular, then index $(\bar{A}_\eta) = 0$, n or $n/2$.

In fact, in this case the Gauss normal map of the immersion $\bar{i}_M$: $M^n \to \mathbb{R}^{n+N+1}$ maps an open neighbourhood of (p, η) in the unit normal bundle of $\bar{i}_M$ onto an open set in S^{n+N}. It follows that, arbitrary near (p, η), there exists (q, ξ) such that index $(\bar{A}_\eta) = \text{index}(\bar{A}_\xi)$ and that ξ is a regular value of the normal map; therefore the function h_ξ is a Morse function. By the tightness of $\bar{i}_M$ and the cohomological structure of the projective planes we get $\text{index}(\bar{A}_\xi) = 0$, n or $n/2$.

We prove now, that i_M is a $2PC$-minimal immersion. For that let $\xi \in v_p M$ be a unit normal vector with $A_\xi \neq 0$ and $\lambda_1 \leqq \lambda_2 \leqq \ldots \leqq \lambda_n$ the eigenvalues of A_ξ. Consider $\xi_\theta = (\cos\theta)\,\xi - (\sin\theta) \cdot i_M(p) \in \bar{v}_p$. Then, according to 3.2, $\bar{A}_{\xi_\theta}$ has the eigenvalues $\bar{\lambda}_i(\theta) = (\cos\theta) \cdot \lambda_i + \sin\theta$. Since trace $A_\xi = 0$, we have $\lambda_1 < 0 < \lambda_n$. If on the other hand $\lambda_j < \lambda_k$, then $\bar{\lambda}_j(\theta) < \bar{\lambda}_k(\theta)$ for $|\theta| < \pi/2$. As $\bar{\lambda}_i(\pi/2) = -\bar{\lambda}_i(-\pi/2) = 1$ for all i, there exists a $\theta \in]-\pi/2, \pi/2[$ such that $\bar{A}_{\xi_\theta}$ is non-singular and $\bar{\lambda}_j(\theta) < 0 < \bar{\lambda}_k(\theta)$. By our Claim it follows that index $(\bar{A}_{\xi_\theta}) = n/2$, i.e. $j \leqq n/2 < k$. In this way we obtain $\lambda_1 = \ldots = \lambda_{n/2} < 0 < \lambda_{1+n/2} = \ldots \lambda_n$.

(3.7) *Remark.* Theorem (1.1) may be proved using the parallelism of the second fundamental form guaranteed by (1.3), by a direct inspection of standard embeddings in S^{n+1}. Once we know that the sectional curvatures are non-negative, Theorem (1.1) can be deduced from a result of Erbacher, who classifies complete immersions in space forms with non-negative sectional curvature, parallel mean curvature and flat normal bundle (see [1]).

(3.8) *Remark.* If M is complete and x: $M^n \to S^{n+p}$ is a $2PC$-minimal immersion such that $x(M^n)$ is not contained in any $(n+1)$-dimensional totally geodesic submanifold of S^{n+p}, then x factors as composition of an isometric covering map and a standard embedding (see [4]). By (3.6) the isometric covering is actually an isometry.

Appendix

In this appendix we will give a proof of the parallelism of the second foundamental form of a $2PC$-minimal immersion since the proof given in [7] was not completely clear to us.

Theorem. *Let x: $M^n \to M_c^{n+N}$ be a $2PC$-minimal immersion and suppose that for any $\xi \in v_q M^n$, A has no simple eigenvalues. Then α is parallel.*

Proof. If $\dim N_1 \leqq 1$ it is easily seen that we can reduce the codimension to 1 and the theorem follow from Otzuki's theorem ((1.1) in our introduction).

We can assume that there exists $q \in M$ such that $\dim N_{1,q} > 1$. Since $\dim N_{1,q}$ is a semicontinuous function of q, there exists a non empty open set $\tilde{U} \subseteq M$ such that $\dim N_{1,q} = \text{const} = k > 1$ for $q \in U$. In N_1 we consider the scalar product $(\xi, \eta) = \text{trace}(A_\xi \cdot A_\eta)$ and the associate self-adjoint endomorphism B of $N_1|\tilde{U}$: $\langle B\xi, \eta\rangle = (\xi, \eta)$. If $U \subseteq \tilde{U}$ is an open set such that the eigenvalues of B have constant multiplicity in U, we can choose a smooth orthonormal frame field $\xi_1, \ldots, \xi_k$ in $N_1|U$, such that $(\xi_i, \xi_j) = 0$ if $i \neq j$. By analiticity it is sufficient to prove that α is parallel in U. Let $\xi = \xi_{i_0}$ and λ eigenvalue of A_ξ. If X, $Y \in T_\lambda$

we have, by 2.3:

$$\langle \xi_i, \alpha(X, Y)\rangle = \langle A_{\xi_i} X, Y\rangle = 0 \quad \text{if } i \neq i_0$$

and therefore

(A.1) $$\alpha(X, Y) = \lambda\langle X, Y\rangle\, \xi \quad \text{for } X, Y \in T_\lambda.$$

Now, if $Z \in T_q M$, since $A_\xi^2 = \lambda I$, $A_\xi((A_\xi - \lambda)Z) = \lambda^2 Z - \lambda A_\xi Z = -\lambda((A_\xi - \lambda)Z)$ and therefore:

(A.2) $$(A_\xi - \lambda)Z \in T_{-\lambda} \quad \forall Z \in T_q M.$$

Claim. λ is constant.

In fact let us consider the Codazzi equation projected in the ξ direction

(A.3) $$(\nabla_X A_\xi)Y - A_{\nabla_X^\perp \xi} Y = (\nabla_Y A_\xi)X - A_{\nabla_Y^\perp \xi} X.$$

If $X, Y \in T_\lambda$ we have

(A.4) $$(A_\xi - \lambda)[X, Y] + A_{\nabla_X^\perp \xi} Y - A_{\nabla_Y^\perp \xi} X = X(\lambda)\, Y - Y(\lambda)\, X.$$

Let $\{X_1, \ldots, X_n\}$ be a basis which diagonalizes A_ξ and $Z \in TM$. From (A.1) we get:

$$\langle A_{\nabla_Z^\perp \xi} X_i, A_\xi X_i\rangle = \langle \nabla_Z^\perp \xi, \alpha(X_i, A_\xi X_i)\rangle = \lambda^2 \quad \langle \nabla_Z^\perp \xi, \xi\rangle = 0$$

which implies

(A.5) $$\operatorname{trace} A_{\nabla_Z^\perp \xi} A_\xi = 0.$$

In particular, from 2.3, $A_{\nabla_Z^\perp \xi}(T_\lambda) \subseteq T_{-\lambda}$ and this, together with (A.2), implies that the left hand side of (A.4) belongs to $T_{-\lambda}$. But the right hand side of (A.4) belong to T_λ and is therefore zero. Since $\dim T_\lambda \geqq 2$, for any $X \in T_\lambda$ we can choose $Y \in T_\lambda$ independent from X and therefore $X(\lambda) = 0$.

In the same way we see that $-\lambda$ is constant in any $T_{-\lambda}$ direction and this completes the proof of our Claim.

Claim. For all $X, Y, Z \in TM$, $(\tilde{\nabla}_Z \alpha)(X, Y) \in N_1^\perp$.

It is sufficient to prove that for $i = 1, \ldots, k$

$$\langle (\tilde{\nabla}_Z \alpha)(X, \dot{Y}), \xi_i\rangle = 0.$$

As in (A.3), this is equivalent to

(A.6) $$\langle ((\nabla_Z A_{\xi_i} - A_{\nabla_Z^\perp \xi_i})\, Y, X\rangle = 0.$$

Let $\xi = \xi_i$ and $\pm\lambda$ be the eigenvalues of A_ξ. If $X \in T_{\pm\lambda}$ and $Z \in TM$, since λ is constant we get:

$$(\nabla_Z A_\xi)X = \nabla_Z(\lambda X) - A_\xi(\nabla_Z X) = -(A_\xi \mp \lambda)\, \nabla_Z X$$

and therefore by (A.2), $(\nabla_Z A_\xi)X \in T_{\mp\lambda}$ and using (A.5), we have

(A.7) $$(\nabla_Z A_\xi - A_{\nabla_Z^\perp \xi})X \in T_{\mp\lambda}, \quad \text{for } X \in T_{\mp\lambda}, \quad Z \in TM$$

and therefore (A.6) holds for X, Y in the same eigenspace. Let now $X \in T_\lambda$, $Y \in T_{-\lambda}$. From (A.7) and (A.3) we get

$$(A.8) \qquad (\nabla_X A_\xi - A_{\nabla_X^\perp \xi}) Y = 0 = (\nabla_Y A_\xi - A_{\nabla_Y^\perp \xi}) X.$$

From the above, using the Codazzi equation we get:

$$0 = \langle (\nabla_X A_\xi - A_{\nabla_X^\perp \xi}) Y, Z \rangle = \langle (\nabla_Z A_\xi - A_{\nabla_Z^\perp \xi}) Y, X \rangle,$$

hence (A.6) is proved for $X \in T_\lambda$ and $Y \in T_{-\lambda}$, and conversely. Therefore our claim is proved.

Claim. N_1 is parallel (in U).

Let $\xi = \xi_{i_0}$ be one of our frame field for N_1 choosen at the beginning of the proof and X, Y orthonormal eigenvectors of A_ξ relative to the same eigenvalue. From (A.1), we get $\alpha(X, Y) = 0$ and therefore for all $Z \in TM$, $(\tilde{\nabla}_Z \alpha)(X, Y) \in N_1$ and by our second claim, $(\tilde{\nabla}_Z \alpha)(X, Y) = 0$. In particular

$$0 = (\tilde{\nabla}_Y \alpha)(X, Y) = (\tilde{\nabla}_X \alpha)(Y, Y) = \pm \lambda \nabla_X^\perp \xi - 2\alpha(\nabla_X Y, Y)$$

and therefore $\nabla_X^\perp \xi \in N_1$ for all $X \in T_\lambda \cup T_{-\lambda}$ and therefore for all $X \in TM$.

A standard argument concludes the proof of the Theorem: From the second claim $(\tilde{\nabla}_Z \alpha)(X, Y)$ is orthogonal to N_1. If η is a unit section of $N_1^\perp$, by our last claim it follows that $\nabla_Z \eta \in N_1^\perp$, $\forall Z \in TM$ and therefore

$$\begin{aligned} \langle (\tilde{\nabla}_Z \alpha)(X, Y), \eta \rangle &= \langle \nabla_Z^\perp(\alpha(X, Y)) - \alpha(\nabla_Z X, Y) - \alpha(X, \nabla_Z Y), \eta \rangle \\ &= -\langle \alpha(X, Y), \nabla_Z^\perp \eta \rangle = 0. \end{aligned}$$

References

1. Chen, B.Y.: Geometry of Submanifolds. New York: Marcel Dekker, Inc. 1973
2. Chern, S.S., Carmo, Do, Kobayashi, S.: Minimal submanifolds of a sphere with second fundamental form of constant lenth. Functional Analysis and Related fields, pp. 59–75. Berlin-Heidelberg-New York: Springer 1970
3. Carmo, M., Do, Dajczer, M.: Rotational hypersurfaces in spaces of constant curvature. To appear in Trans. A.M.S.
4. Ferus, D.: Immersions with parallel second fundamental form. Math. Z. **140**, 87–93 (1970)
5. Ferus, D.: Symmetric submanifolds of euclidean space. Math. Ann. **274**, 81–93 (1980)
6. Kato, T.: Perturbation Theory of Linear Operators. Berlin-Heidelberg-New York: Springer 1966
7. Matsuyama, Y.: Minimal submanifolds of S^N and $\mathbb{R}^N$. Math. Z. **175**, 275–282 (1980)
8. Milnor, J.: Morse Theory. Princeton, N.J.: Princeton Academic Press 1963
9. Otsuki, T.: Minimal hypersurfaces in a Riemannian manifold of constant curvature. Amer. J. Math. XCII, (1) 1970
10. Rodríguez, L.: Geometria das Subvariedades. Monografias, IMPA, 1976
11. Rodríguez, L., Tribuzy, R.: Reduction of Codimension of regular immersion. Preprint IMPA, 1982
12. Takeuchi, M.: On Pontryagin classes of compact symmetric spaces. J. Fac. Sci. Univ. Tokyo Sect. **19**, 313–328 (1962)
13. Wolf, J.A.: Spaces of Constant Curvature. New York: McGraw-Hill 1967
14. Wolf, J.A.: Symmetric spaces which are cohomology spheres. Journal Differential Geometry **3**, 59–68 (1969)

Received March 12, 1984

Math. Z. 187, 335–344 (1984)

Mathematische Zeitschrift

Diophantine Approximation with Square-free Numbers

D.R. Heath-Brown
Magdalen College, Oxford OX1 4AU, Great Britain

1. Introduction

Let α be a fixed real irrational number. We are interested in Diophantine approximations of the form

$$|n\alpha-m|\leqq\frac{1}{n^{\theta}},$$

with m, n square-free. We shall prove:

Theorem. *Let $\alpha\in\mathbb{R}-\mathbb{Q}$ and let $\theta<2/3$. Then there are infinitely many square-free integers m, n such that*

$$|n\alpha-m|\leqq\frac{1}{n^{\theta}}. \tag{1}$$

Previously the best known result was for $\theta<1/2$. This was obtained recently by Harman [3]. A slightly weaker result was found independently, with a similar argument, by Balog and Perelli [1]. Of course if α is a quadratic irrational, then $|n\alpha-m|\gg n^{-1}$ for every $m,n\neq 0$. Thus the best range one can hope for in the theorem is $\theta\leqq 1$. Indeed it is natural to conjecture that (1) has infinitely many solutions in square-free m, n for each $\theta<1$.

The method used by Harman [3], and by Balog and Perelli [1], employs character sums. We use a more elementary method, based on a result (namely Lemma 1) from the geometry of numbers. In appropriate circumstances this provides a very effective way of counting solutions of linear congruences. We also use the "square-sieve" (see Heath-Brown [4]) to estimate the number of squares in a certain sequence.

This paper was written while the author was enjoying a visiting professorship at Oklahoma State University, Stillwater.

2. A Result from the Geometry of Numbers

The terminology of this section differs somewhat from that of standard works on the geometry of numbers.

Let $\mathbb{Z}^n \subseteq \mathbb{R}^n$ be the lattice spanned by the unit coordinate vectors of $\mathbb{R}^n$. We shall say that Λ is a "sublattice" of $\mathbb{Z}^n$ of "rank" r, if $\Lambda \subseteq \mathbb{Z}^n$ is a submodule, over $\mathbb{Z}$, of rank r. We say that Λ, or equivalently its basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$, is "primitive" if $\mathbf{b}_1, \ldots, \mathbf{b}_r$ can be extended to a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{Z}^n$. A necessary and sufficient condition for a basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$ to be primitive is that

$$\sum_1^r \lambda_i \mathbf{b}_i \in \mathbb{Z}^n, \quad \lambda_i \in \mathbb{R}, \quad (1 \leqq i \leqq r),$$

implies $\lambda_i \in \mathbb{Z}$, $(1 \leqq i \leqq r)$. (This is Corollary 3 of Cassels [2; p. 14].) In particular a single vector $\mathbf{b} \neq \mathbf{0}$ is primitive if and only if its coordinates are coprime integers. This corresponds to normal usage. For any set $S \subseteq \mathbb{R}^n$ we shall write $\langle S \rangle_{\mathbb{Z}}$ for the set of all finite linear combinations, over $\mathbb{Z}$, of elements of S. We define $\langle S \rangle_{\mathbb{Q}}$ and $\langle S \rangle_{\mathbb{R}}$ similarly.

If Λ is a sublattice of $\mathbb{Z}^n$, with basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$, we defined the "dual" lattice Λ^* as

$$\Lambda = \{\mathbf{x} \in \mathbb{Z}^n; \mathbf{x} \cdot \mathbf{b}_1 = \ldots = \mathbf{x} \cdot \mathbf{b}_r = 0\}.$$

This is independent of the choice of the basis for Λ. The dual lattice Λ^* will be a primitive sublattice of $\mathbb{Z}^n$ of rank $n-r$, and $(\Lambda^*)^* = \Lambda$. To prove these assertions we find a basis for Λ^*. Extend $\mathbf{b}_1, \ldots, \mathbf{b}_r$ to a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ for $\mathbb{Z}^n$. Define $\mathbf{b}_1^*, \ldots, \mathbf{b}_n^* \in \mathbb{Q}^n$ by

$$\mathbf{b}_i^* \cdot \mathbf{b}_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{2}$$

Thus $\langle \mathbf{b}_1, \ldots, \mathbf{b}_r \rangle_{\mathbb{Q}}$ and $\langle \mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^* \rangle_{\mathbb{Q}}$ are orthogonal complements in $\mathbb{Q}^n$, and $\langle \Lambda^* \rangle_{\mathbb{Q}} = \langle \mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^* \rangle_{\mathbb{Q}}$. If $\mathbf{e}_k$ is the k-th unit coordinate vector in $\mathbb{Z}^n$, then $\mathbf{e}_k \in \langle \mathbf{b}_1, \ldots, \mathbf{b}_n \rangle_{\mathbb{Z}} = \mathbb{Z}^n$. Consequently (2) yields $\mathbf{b}_i^* \cdot \mathbf{e}_k \in \mathbb{Z}$, for each k, whence $\mathbf{b}_i^* \in \mathbb{Z}^n$. Thus $\mathbf{b}_i^* \in \Lambda^*$ for $r < i \leqq n$. Moreover, if $\mathbf{b}^* \in \Lambda^*$, then

$$\mathbf{b}^* \in \langle \Lambda^* \rangle_{\mathbb{Q}} = \langle \mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^* \rangle_{\mathbb{Q}}.$$

Thus

$$\mathbf{b}^* = \sum_{r+1}^n \lambda_i \mathbf{b}_i^*, \quad \lambda_i \in \mathbb{Q}.$$

Since $\mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^*$ are primitive we conclude that $\lambda_i \in \mathbb{Z}$ for $r < i \leqq n$. It follows that $\mathbf{b}^* \in \langle \mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^* \rangle_{\mathbb{Z}}$. Hence $\mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^*$ is the required basis for Λ^*. It is now clear that Λ^* is primitive, with rank $n-r$, and that $(\Lambda^*)^* = \Lambda$.

We define the "determinant" of a primitive Λ to be the volume of the r-dimensional parallelopiped generated by any basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$ of Λ. Thus if $\mathbf{e}_{r+1}, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ are orthonormal vectors such that $\mathbf{b}_i \cdot \mathbf{e}_j = 0$ for each i, j, then the determinant is

$$d(\Lambda) = |\det[\mathbf{b}_1, \ldots, \mathbf{b}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n]|.$$

It is an easy exercise to check that this is independent of the choice of $\mathbf{b}_i$ and $\mathbf{e}_j$. We can now state our principal result.

Lemma 1. *If Λ is a primitive sublattice of $\mathbb{Z}^n$ then $d(\Lambda) = d(\Lambda^*)$.*

Let Λ have basis $\mathbf{b}_1, \ldots, \mathbf{b}_r$. Extend this to a basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of $\mathbb{Z}^n$, and construct $\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*$ as in (2). Let $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ be the orthonormal vectors formed by applying the Gram-Schmidt process to $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Thus

$$\langle \mathbf{b}_1, \ldots, \mathbf{b}_s \rangle_{\mathbb{R}} = \langle \mathbf{e}_1, \ldots, \mathbf{e}_s \rangle_{\mathbb{R}}, \qquad (1 \leqq s \leqq n). \tag{3}$$

We then have

$$\langle \mathbf{b}_{s+1}^*, \ldots, \mathbf{b}_n^* \rangle_{\mathbb{R}} = \langle \mathbf{e}_{s+1}, \ldots, \mathbf{e}_n \rangle_{\mathbb{R}}, \qquad (0 \leqq s < n), \tag{4}$$

since these are both orthogonal complements to $\langle \mathbf{b}_1, \ldots, \mathbf{b}_s \rangle_{\mathbb{R}}$ in $\mathbb{R}^n$. With these definitions we have

$$d(\Lambda) = |\det[\mathbf{b}_1, \ldots, \mathbf{b}_r, \mathbf{e}_{r+1}, \ldots, \mathbf{e}_n]|,$$

$$d(\Lambda^*) = |\det[\mathbf{e}_1, \ldots, \mathbf{e}_r, \mathbf{b}_{r+1}^*, \ldots, \mathbf{b}_n^*]|,$$

and we shall prove by induction on t that

$$|\det[\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{e}_{t+1}, \ldots, \mathbf{e}_n]| = |\det[\mathbf{e}_1, \ldots, \mathbf{e}_t, \mathbf{b}_{t+1}^*, \ldots, \mathbf{b}_n^*]|; \tag{5}$$

this clearly suffices for Lemma 1.

When $t=0$ we have $\det[\mathbf{e}_1, \ldots, \mathbf{e}_n] = \pm 1$, since the $\mathbf{e}_i$ are orthonormal, and $\det[\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*] = \pm 1$, since the $\mathbf{b}_i^*$ are a basis for $(\mathbb{Z}^n)^* = \mathbb{Z}^n$, and $d(\mathbb{Z}^n) = 1$. This verifies (5) when $t=0$. Now suppose (5) holds for a particular value of $t<n$. From (3), with $s=t+1$, we have

$$\mathbf{b}_{t+1} = \sum_1^{t+1} \lambda_i \mathbf{e}_i, \qquad (\lambda_i \in \mathbb{R}). \tag{6}$$

However the case $s=t$ of (3) yields

$$\sum_1^t \lambda_i \mathbf{e}_i = \sum_1^t \mu_i \mathbf{b}_i$$

for certain $\mu_i \in \mathbb{R}$. Hence

$$\mathbf{b}_{t+1} = \lambda_{t+1} \mathbf{e}_{t+1} + \sum_1^t \mu_i \mathbf{b}_i.$$

Now

$$\begin{aligned}\det[\mathbf{b}_1, \ldots, \mathbf{b}_{t+1}, \mathbf{e}_{t+2}, \ldots, \mathbf{e}_n] = \lambda_{t+1} &\det[\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{e}_{t+1}, \ldots, \mathbf{e}_n] \\ &+ \sum_1^t \mu_i \det[\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{b}_i, \mathbf{e}_{t+2}, \ldots, \mathbf{e}_n].\end{aligned}$$

The determinants in the final sum all have two equal columns, and so vanish. It follows that

$$\det[\mathbf{b}_1, \ldots, \mathbf{b}_{t+1}, \mathbf{e}_{t+2}, \ldots, \mathbf{e}_n] = \lambda_{t+1} \det[\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{e}_{t+1}, \ldots, \mathbf{e}_n]. \tag{7}$$

A similar argument based on (4) yields

$$\begin{aligned}\mathbf{b}_{t+1}^* &= \sum_{t+1}^n \nu_i \mathbf{e}_i \\ &= \nu_{t+1} \mathbf{e}_{t+1} + \sum_{t+2}^n \pi_i \mathbf{b}_i^*\end{aligned} \tag{8}$$

and

$$\det[\mathbf{e}_1, \ldots, \mathbf{e}_t, \mathbf{b}^*_{t+1}, \ldots, \mathbf{b}^*_n] = v_{t+1} \det[\mathbf{e}_1, \ldots, \mathbf{e}_{t+1}, \mathbf{b}^*_{t+2}, \ldots, \mathbf{b}^*_n]. \tag{9}$$

Since the $\mathbf{e}_i$ are orthonormal we have

$$1 = \mathbf{b}^*_{t+1} \cdot \mathbf{b}_{t+1} = \left(\sum_{t+1}^{n} v_i \mathbf{e}_i\right) \cdot \left(\sum_{1}^{t+1} \lambda_i \mathbf{e}_i\right) = v_{t+1} \lambda_{t+1},$$

by (2), (6) and (8). Hence (5), (7) and (9) yield

$$|\det[\mathbf{b}_1, \ldots, \mathbf{b}_{t+1}, \mathbf{e}_{t+2}, \ldots, \mathbf{e}_n]| = |\det[\mathbf{e}_1, \ldots, \mathbf{e}_{t+2}, \mathbf{b}^*_{t+2}, \ldots, \mathbf{b}^*_n]|.$$

This ends the induction step of our proof of (5), and thus completes our proof of Lemma 1.

We shall require an auxilliary result to count lattice points in $\mathbb{R}^2$.

Lemma 2. *Let $\Lambda \subseteq \mathbb{R}^2$ be a lattice of determinant $d(\Lambda)$. Let $E \subseteq \mathbb{R}^2$ be an ellipse, centered at the origin, together with its interior, and let A be the area of E. Then there is a positive number $\alpha = \alpha(\Lambda, E)$ and a basis $\mathbf{b}_1, \mathbf{b}_2$ of Λ, such that $g\mathbf{b}_1 + h\mathbf{b}_2 \in E$ implies $|g| \leqq \alpha$, $|h| \leqq A/(\alpha d(\Lambda))$. Moreover the number of primitive lattice points of Λ contained in E is at most $4(1 + d(\Lambda)^{-1} A)$.*

Here we use the term primitive in its conventional sense: $\mathbf{v} \in \Lambda$ is primitive if $n^{-1}\mathbf{v} \notin \Lambda$ for all integers $n > 1$. The second assertion of the lemma follows immediately from the first, since the number of available g, h with $gh \neq 0$ is at most $4A/d(\Lambda)$. Moreover the number of lattice points with $gh = 0$ is at most 4, since if $g = 0$, say, then $h = \pm 1$, on account of the condition that $g\mathbf{b}_1 + h\mathbf{b}_2$ be primitive.

To prove the main assertion of the lemma we first suppose that E is a disc of radius r. Let $\mathbf{b}_1$ be a non-zero vector in Λ of minimal length, and let $\mathbf{b}_2 \in \Lambda$ have minimal length subject to the condition that $\mathbf{b}_1$ and $\mathbf{b}_2$ should not be parallel. Then if θ is the angle between $\mathbf{b}_1$ and $\mathbf{b}_2$ we have

$$\pi/3 \leqq \theta \leqq 2\pi/3, \tag{10}$$

$$d(\Lambda) = \|\mathbf{b}_1\| \cdot \|\mathbf{b}_2\| (\sin\theta), \tag{11}$$

where $\|.\|$ denotes Euclidean length. Moreover $\mathbf{b}_1$, $\mathbf{b}_2$ form a basis for Λ. Let $\mathbf{e}_1$, $\mathbf{e}_2$ be unit vectors perpendicular to $\mathbf{b}_1$, $\mathbf{b}_2$ respectively. Then

$$|\mathbf{e}_1 \cdot \mathbf{b}_2| \cdot \|\mathbf{b}_1\| = |\mathbf{e}_2 \cdot \mathbf{b}_1| \cdot \|\mathbf{b}_2\| = d(\Lambda).$$

Now suppose that $\mathbf{v} + g\mathbf{b}_1 + h\mathbf{b}_2 \in E$. Then $|\mathbf{e}_1 \cdot \mathbf{v}| \leqq r$, since $\|\mathbf{v}\| \leqq r$. However $\mathbf{e}_1 \cdot \mathbf{v} = h\mathbf{e}_1 \cdot \mathbf{b}_2$, and so $|h|\, d(\Lambda) \leqq r\, \|\mathbf{b}_1\|$. Similarly $|g|\, d(\Lambda) \leqq r\, \|\mathbf{b}_2\|$. If we set $\alpha = r\, \|\mathbf{b}_2\|\, d(\Lambda)^{-1}$ then $|g| \leqq \alpha$ and

$$|h| \leqq \frac{r\,\|\mathbf{b}_1\|}{d(\Lambda)} = \frac{(\operatorname{cosec}\theta)}{\pi} \cdot \frac{A}{\alpha d(\Lambda)} \leqq \frac{A}{\alpha d(\Lambda)},$$

by (10) and (11), since $\operatorname{cosec}\theta \leqq (2/\sqrt{3}) \leqq \pi$.

In the general case we use a unimodular transformation M of $\mathbb{R}^2$ which takes E to a disc ME. Then ME has area A and $d(M\Lambda)=d(\Lambda)$. Thus there exist $\alpha(M\Lambda, ME)$ and a basis $\mathbf{b}'_1$, $\mathbf{b}'_2$ of $M\Lambda$ as in the lemma. If we take $\mathbf{b}_i=M^{-1}\mathbf{b}'_i$, then $\mathbf{b}_1, \mathbf{b}_2$ will be a basis for Λ. Moreover, if $g\mathbf{b}_1+h\mathbf{b}_2\in E$, then $g\mathbf{b}'_1+h\mathbf{b}'_2\in ME$, whence $|g|\leqq\alpha$, $|h|\leqq A/(\alpha d(\Lambda))$ as required. This completes the proof of Lemma 2.

We now give our application of Lemmas 1 and 2. We could have taken a shorter route, but the method used here is more illuminating, and may be of independent interest.

Lemma 3. *Let $\mathbf{v}=(v_1, v_2, v_3)\in\mathbb{Z}^3$ be a primitive vector, and let $X_i>0$ $(i=1,2,3)$ be given. Then there exists $\alpha=\alpha(\mathbf{v}, X_1, X_2, X_3)$ and primitive vectors $\mathbf{b}_1, \mathbf{b}_2\in\mathbb{Z}^3$ as follows. Every solution $\mathbf{x}=(x_1, x_2, x_3)\in\mathbb{Z}^3$ of $\mathbf{v}\cdot\mathbf{x}=0$, for which $|x_i|\leqq X_i$ $(i=1, 2, 3)$ is of the form $\mathbf{x}=g\mathbf{b}_1+h\mathbf{b}_2$ with $|g|\leqq\alpha$, and*

$$|h|\leqq 3\pi\frac{\prod X_i}{\alpha\max X_i|v_i|}.$$

Moreover the number of primitive solutions $\mathbf{x}$ is at most

$$4+12\pi\frac{\prod X_i}{\max X_i|v_i|}.$$

If Q is the 3×3 matrix representing an ellipsoid $Q(\mathbf{x})\leqq 1$, then the intersection of the ellipsoid and a plane $\{\mathbf{x}\in\mathbb{R}^3;\ \mathbf{x}\cdot\mathbf{v}=0\}$ is an ellipse E of area

$$A=\pi\,\|\mathbf{v}\|\,(Q^{-1}(\mathbf{v})\det Q)^{-1/2}.$$

We are interested in the matrix $Q=\operatorname{diag}((3X_1^2)^{-1}, (3X_2^2)^{-1}, (3X_3^2)^{-1})$, for which

$$Q^{-1}(\mathbf{v})=3X_1^2v_1^2+3X_2^2v_2^2+3X_3^2v_3^2\geqq 3(\max X_i|v_i|)^2$$

and so

$$A\leqq 3\pi\,\|\mathbf{v}\|\frac{\prod X_i}{\max X_i|v_i|}.$$

If Λ is the lattice $\langle\mathbf{v}\rangle_{\mathbb{Z}}\subseteq\mathbb{Z}^3$ then $d(\Lambda)=\|\mathbf{v}\|$. The vectors $\mathbf{x}$ in Lemma 3 lie in Λ^*, a lattice of rank 2 and determinant $d(\Lambda^*)=\|\mathbf{v}\|$, by Lemma 1. Hence Lemma 3 follows from Lemma 2.

3. Proof of the Theorem: First Steps

Let α be positive. We take a/q to be any convergent in the continued fraction for α, so that $|\alpha-a/q|\leqq q^{-2}$. Let $0<\theta<1$ and define

$$N=q^{2/(1+\theta)},\qquad L=Nq^{-1}(\log q)^{-1}, \tag{12}$$

$$S=\{(l,m,n)\in\mathbb{Z}^3;\ 1\leqq l\leqq L,\ N<n\leqq 2N,\ an-qm=l\}.$$

If $(l,m,n)\in S$ then

$$\left|\alpha-\frac{m}{n}\right|\leqq\left|\alpha-\frac{a}{q}\right|+\left|\frac{a}{q}-\frac{m}{n}\right|\leqq\frac{1}{q^2}+\frac{l}{nq}\leqq 2N^{-1-\theta}\leqq 8n^{-1-\theta}.$$

We shall estimate

$$\#\{(l,m,n)\in S;\ \mu^2(m)\,\mu^2(n)=1\}=T,$$

say, from below. Let $z=\log q$ and

$$P=\prod_{p<z}p,$$

where p runs over primes. Define

$$f(u)=\sum_{d^2|u,d|P}\mu(d),$$

so that $f(u)=1$ or 0 for all u, and

$$\mu^2(u)\geqq f(u)-\sum_{p^2|u,\,p\geqq z}1.$$

Then

$$\mu^2(m)\,\mu^2(n)\geqq f(m)\,f(n)-\sum_{p^2|m,\ p\geqq z}1-\sum_{p^2|n,\ p\geqq z}1,$$

and hence

$$T\geqq A-\sum_{p\geqq z}B_p-\sum_{p\geqq z}C_p, \tag{13}$$

with

$$A=\sum_{(l,m,n)\in S}f(m)\,f(n),$$

$$B_p=\#\{(l,m,n)\in S;p^2|m\},\qquad C_p=\#\{(l,m,n)\in S;p^2|n\}.$$

In this section we shall estimate A. On writing $m=d^2u$, $n=e^2v$ we have

$$A=\sum_{d,e|P}\mu(d)\,\mu(e)\#\{(l,v);1\leqq l\leqq L,\ Ne^{-2}<v\leqq 2Ne^{-2},$$
$$ae^2v\equiv l(\operatorname{mod} qd^2)\}.$$

If we define $(ae^2,qd^2)=D$, then the congruence condition requires $D|l$. We therefore set $l=Dk$, so that

$$A=\sum_{d,e|P}\mu(d)\,\mu(e)\#\{(k,v);\ 1\leqq k\leqq LD^{-1},\ Ne^{-2}<v\leqq 2Ne^{-2},$$
$$ae^2D^{-1}\cdot v\equiv k(\operatorname{mod} qd^2D^{-1})\}.$$

The congruence now has one solution $v(\operatorname{mod} qd^2D^{-1})$ for each value of k. It follows that

$$A=\sum_{d,e|P}\mu(d)\,\mu(e)\left\{\frac{L}{D}+O(1)\right\}\left\{\frac{ND}{qd^2e^2}+O(1)\right\}$$
$$=\frac{NL}{q}\sum_{d,e|P}\mu(d)\,\mu(e)\,d^{-2}e^{-2}+O(L4^{\omega(P)})+O(Nq^{-1}4^{\omega(P)}).$$

Here we have used the observations that there are $4^{\omega(P)}$ choices for d, e, and that $D|d^2e^2$, since $(a,q)=1$. By our choice of z we have $\omega(P)=\pi(z)=o(\log q)$, whence $4^{\omega(P)}\ll q^{\varepsilon}$ for any $\varepsilon>0$. Thus the error terms above are each $o(LN/q)$ as $q\to\infty$, by (12). Moreover

$$\sum_{d,e|P}\mu(d)\mu(e)d^{-2}e^{-2}=\prod_{p<z}\left(1-\frac{1}{p^2}\right)^2\to\zeta(2)^{-2}$$

as $q\to\infty$. It follows that

$$A\sim NLq^{-1}\zeta(2)^{-2}. \tag{14}$$

4. The Estimation of $\sum C_p$

In this section we shall estimate $\sum C_p$, the treatment of $\sum B_p$ being essentially identical. We classify the elements in C_p depending on the value of $(l,m,n)=d$, and write $C_p=\sum C_{p,d}$ accordingly. Since the number of primes p such that $p^2|n$ is $O((\log q)(\log\log q)^{-1})$, we have

$$\sum_p C_{p,d}\ll\frac{\log q}{\log\log q}\#\{(k,u,v)\in\mathbb{Z}^3;\ 1\leqq k\leqq L/d,\ N/d<v\leqq 2N/d,$$
$$av-qu=k,\ (u,v,k)=1\}.$$

The quantity above may be estimated by Lemma 3. We take $\mathbf{v}=(a,-q,-1)$, $X_1=2N/d$, $X_3=L/d$. Moreover, since $|\alpha-a/q|\leqq 1$, we have $a\ll q$. Hence

$$qu\ll k+|av|\ll(L+Nq)/d\ll Nq/d. \tag{15}$$

We may therefore take $X_2\ll N/d$. Lemma 3 now yields

$$\sum_p C_{p,d}\ll\frac{\log q}{\log\log q}\left(1+\frac{NL}{qd^2}\right).$$

Since $d|l$ we have $d\ll L$, and consequently

$$\sum_{d\geqq z}\sum_p C_{p,d}\ll\frac{\log q}{\log\log q}\left(L+\frac{NL}{qz}\right)\ll\frac{NL}{q(\log\log q)}. \tag{16}$$

In the remaining terms $C_{p,d}$ we have $d<z\leqq p$. Hence if $(l,m,n)=(dk,du,dv)\in C_{p,d}$, then $p^2|v$. On setting $v=p^2t$ it follows that $\mathbf{v}\cdot\mathbf{x}=0$, where

$$\mathbf{v}=(ap^2,-q,-1),\qquad \mathbf{x}=(t,u,k).$$

Here we have $t\ll Nd^{-1}p^{-2}$, $k\ll Ld^{-1}$ and (as in (15)) $u\ll Nd^{-1}$. Thus Lemma 3 yields

$$C_{p,d}\ll 1+\frac{LN}{qd^2p^2}.$$

We use this for $d<z$ and $p\leqq LN/(qz^2)$. Then

$$\sum_{d<z}\ \sum_{z\leqq p\leqq LN/(qz^2)}C_{p,d}\ll\frac{LN}{q(\log q)}. \tag{17}$$

For larger values of p we write

$$\sum_{R<p\leqq 2R} C_{p,d} \leqq \sum_{N/(4R^2d)<t<2N/(R^2d)} C(t), \tag{18}$$

$$C(t) = \#\{(k,u,p)\in\mathbb{Z}^3;\ 1\leqq k\leqq L/d,\ R<p\leqq 2R,\ atp^2-qu=k,\ (k,u,p)=1\}.$$

To estimate $C(t)$ we apply Lemma 3 with $\mathbf{v}=(at,-q,-1)$, $X_1\ll R^2$, $X_2\ll Nd^{-1}$, $X_3\ll Ld^{-1}$. Thus $(p^2,u,k)=g\mathbf{b}_1+h\mathbf{b}_2$ with $|g|\leqq\alpha$, $|h|\ll LR^2(\alpha dq)^{-1}$. We write these conditions as $|g|\leqq\alpha$, $|h|\leqq\beta$, where

$$\alpha\beta\ll LR^2(dq)^{-1}. \tag{19}$$

If $\beta<1$ then $(p^2,u,k)=g\mathbf{b}_1$. Since (p^2,u,k) has to be primitive we must have $g=\pm 1$, whence

$$C(t)\ll 1. \tag{20}$$

A similar argument applies if $\alpha<1$. Henceforth we shall assume that $\alpha,\beta\geqq 1$.

For each p there is at most one $(k,u,p)\in C(t)$. This is because if (k',u',p) is also in $C(t)$, then

$$|u-u'|=q^{-1}|k-k'|\leqq L(qd)^{-1}<1,$$

by (12). Hence there exist c_1, c_2 such that

$$C(t)\ll \#\{p;\ p^2=gc_1+hc_2,\ R<p\leqq 2R,\ |g|\leqq\alpha,\ |h|\leqq\beta\}.$$

By symmetry we may assume that $\alpha\geqq\beta$. If $|c_1|\geqq 4R^2$ then for each h there is at most one g for which

$$R^2<gc_1+hc_2\leqq 4R^2.$$

In this case

$$C(t)\ll\beta\ll(\alpha\beta)^{1/2}. \tag{21}$$

If $c_1=0$ then the condition $p^2=hc_2$ requires $c_2=\pm 1$ or $\pm p$ or $\pm p^2$. If $c_2=\pm p$ or $\pm p^2$ there is at most one possible value for p, so that

$$C(t)\ll 1. \tag{22}$$

If $c_2=\pm 1$ then $h=\pm p^2$. Here there are no solutions unless $\beta\geqq R^2$, in which case g and h have $O(\alpha)$ and $O(R)$ possible values respectively. Hence

$$C(t)\ll R\alpha\ll\alpha\beta R^{-1}. \tag{23}$$

There remains the range $1\leqq|c_1|\leqq 4R^2$. In this, the critical case, we shall use the "square sieve" of Heath-Brown [4; Theorem 1]. For convenience we state the result here. Let $w(j)\geqq 0$ for each $j\in\mathbb{Z}$, and suppose $\sum w(j)<\infty$. We then have:

Lemma 4. *Let $\mathscr{P}$ be a set of K primes. Suppose that $w(j)=0$ for $j=0$ or $|j|\geqq e^K$. Then*

$$\sum_{j=1}^{\infty} w(j^2) \ll K^{-1} \sum_j w(j) + K^{-2} \sum_{p_1 \neq p_2 \in \mathscr{P}} \left| \sum_j w(j) \left(\frac{j}{p_1 p_2}\right) \right|, \tag{24}$$

where $\left(\frac{j}{p_1 p_2}\right)$ *is the Jacobi symbol.*

This result has its motivation in the work of Hooley [5]. The method used there gives, when abstracted, a result of similar strength to Lemma 4, but slightly more complicated.

We take

$$w(j) = \#\{g, h; j = gc_1 + hc_2, |g| \leqq \alpha, |h| \leqq \beta\}$$

if $R^2 < j \leqq 4R^2$, and $w(j) = 0$ otherwise. Then

$$C(t) \ll \sum_{j=1}^{\infty} w(j^2). \tag{25}$$

Now let

$$\mathscr{P} = \{p;\ 3 \leqq p \leqq (\log q)^3,\ p \nmid c_1\}, \tag{26}$$

whence

$$(\log q)^2 \leqq K \leqq (\log q)^3, \tag{27}$$

since $\omega(c_1) \ll \log R \ll \log q$. It follows that $w(j) = 0$ for $j = 0$ or $|j| \geqq e^K$.

The first sum on the right of (24) is

$$\sum_j w(j) \ll \#\{g, h;\ |g| \leqq \alpha,\ |h| \leqq \beta\} \ll \alpha\beta. \tag{28}$$

Moreover

$$\sum_j w(j) \left(\frac{j}{p_1 p_2}\right) = \sum_{|h| \leqq \beta} \sum_{g \in I} \left(\frac{gc_1 + hc_2}{p_1 p_2}\right) \tag{29}$$

for a certain interval I. Since $(p_1 p_2, c_1) = 1$, the Pólya-Vinogradov inequality yields

$$\sum_{g \in I} \left(\frac{gc_1 + hc_2}{p_1 p_2}\right) \ll (p_1 p_2)^{1/2} (\log p_1 p_2) \ll (\log q)^4, \tag{30}$$

by (26). It now follows, from (25), (27)–(30) and Lemma 4, that

$$\begin{aligned} C(t) &\ll \alpha\beta(\log q)^{-2} + \beta(\log q)^4 \\ &\ll \alpha\beta(\log q)^{-2} + (\alpha\beta)^{1/2}(\log q)^4, \end{aligned}$$

since $\beta \leqq \alpha$.

We compare this with (20)–(22) and (23). On using (19) we find that

$$C(t) \ll 1 + LR(dq)^{-1} + LR^2(dq)^{-1}(\log q)^{-2} + (LR^2(dq)^{-1})^{1/2}(\log q)^4$$

in all cases. Consequently (18) yields

$$\sum_{R < p \leqq 2R} C_{p,d} \ll \frac{N}{R^2 d} + \frac{NL}{Rd^2 q} + \frac{NL}{d^2 q}(\log q)^{-2} + \frac{NL^{1/2}}{Rd^{3/2} q^{1/2}}(\log q)^4.$$

Since $C_{p,d}$ vanishes unless $p \ll N^{1/2}$ we now have

$$\sum_{d<z} \sum_{p>R_0} C_{p,d} \ll \frac{N}{R_0^2}(\log q) + \frac{NL}{R_0 q} + \frac{NL}{q}(\log q)^{-1} + \frac{NL^{1/2}}{R_0 q^{1/2}}(\log q)^4.$$

This will be $O(NLq^{-1}(\log q)^{-1})$ providing that $R_0 \geqq (q/L)^{1/2}(\log q)^5$. Comparison with (16) and (17) therefore yields

$$\sum_{p \geqq z} C_p = \sum_{d=1}^{\infty} \sum_{p \geqq z} C_{p,d} \ll \frac{NL}{q(\log\log q)}$$

under the condition

$$\frac{LN}{qz^2} \geqq \left(\frac{q}{L}\right)^{1/2} (\log q)^5. \tag{31}$$

A similar bound holds for $\sum B_p$, and so T will be positive, by (13) and (14), if q is large enough. Since (31) requires only $\theta < 2/3$, by (12), the theorem follows.

References

1. Balog, A., Perelli, A.: Diophantine approximation by square-free numbers, to appear
2. Cassels, J.W.S.: An introduction to the geometry of numbers. Berlin-Heidelberg-New York: Springer 1959
3. Harman, G.: Diophantine approximation with square-free integers, to appear
4. Heath-Brown, D.R.: The square sieve and consecutive square-free numbers. Math. Ann. **226**, 251–259 (1984)
5. Hooley, C.: On the representations of a number as the sum of four cubes: I. Proc. London Math. Soc. (3) **36**, 117–140 (1978)

Received March 12, 1984

Note Added 27th May 1984

Since this article was submitted for publication an alternative proof of Lemma 1 has appeared (P. McMullen, *Determinants of lattices induced by rational subspaces*. Bull. London Math. Soc. **16**, 275–277 (1984)). However the author has recently learned, from Professor J. Vaaler, that the result in question is an easy consequence of the so-called "Brill-Gordan Duality Theorem". (For which see P. Gordan, *Über den größten gemeinsamen Faktor*. Math. Annalen, **7**, 433–451 (1874).)

Math. Z. 187, 345–360 (1984)

Mathematische Zeitschrift

Generischer Spaltungstyp und zweite Chernklasse stabiler Vektorraumbündel vom Rang 4 auf $\mathbb{P}_4$*

Hans Jürgen Hoppe

Universität Bayreuth, Postfach 3008, D-8580 Bayreuth, Bundesrepublik Deutschland

1. Einleitung

Der generische Spaltungstyp einer semistabilen torsionsfreien Garbe vom Rang r auf $\mathbb{P}_n$

$$a_E=(a_1, \dots, a_r),$$

$a_i \in \mathbb{Z}$, $i=1, \dots, r$; $a_1 \geqq \dots \geqq a_r$, genügt nach dem Satz von Grauert-Mülich-Spindler den Bedingungen

(1.1) $$a_i - a_{i+1} \leqq 1, \quad i=1, \dots, r-1$$

(vgl. [8], p. 205, Corollary 1).

a_E ist dadurch noch nicht festgelegt, wenn $r \geqq 3$ ist. Man nennt den generischen Spaltungstyp $a_E=(a_1, \dots, a_r)$ „generisch", wenn

$$d_E = a_1 - a_r$$

minimal ist.

Die Chernklassen $c_1 = c_1 E$ und $c_2 = c_2 E$ eines stabilen r-Bündels E auf $\mathbb{P}_n$ genügen der Abschätzung

(1.2) $$(r-1)\,c_1^2 - 2r\,c_2 \leqq 0$$

(vgl. [1], p. 552, Theorem 6).

Da die Einschränkung eines stabilen 3-Bündels E auf $\mathbb{P}_3$ mit $c_1=0$ auf allgemeine Ebenen $H \subset \mathbb{P}_3$ semistabil ist (vgl. [7], p. 331, Corollary 3.1.1), erhält man:

$$c_2 E \geqq 2, \qquad \text{falls } a_E=(0, 0, 0)$$

* Diese Arbeit entstand im Rahmen eines von der Deutschen Forschungsgemeinschaft geförderten Projekts

und

$$c_2 E \leqq 3, \qquad \text{falls } a_E = (1, 0, 1).$$

Dieses und andere Beispiele auf $\mathbb{P}_3$ motivieren die Vermutung:

Ist die zweite Chernklasse eines stabilen r-Bündels auf $\mathbb{P}_n$ „minimal", so ist der generische Spaltungstyp „generisch".

Das Hauptanliegen dieser Arbeit ist zu zeigen, daß diese Vermutung für stabile 4-Bündel auf P_4 richtig ist.

Dazu werden zunächst - in Abhängigkeit vom generischen Spaltungstyp - die minimalen zweiten Chernklassen angegeben; dann werden Beispiele stabiler 4-Bündel auf $\mathbb{P}_4$ mit minimaler zweiter Chernklasse konstruiert, deren generischer Spaltungstyp „generisch" ist.

Ist E ein stabiles n-Bündel auf $\mathbb{P}_n$ mit $c_1 = -1$, dann ist $c_2 = 1$ die minimale zweite Chernklasse bezüglich 1.2.

Das Cotangentialbündel $\Omega^1_{\mathbb{P}_n}(1)$ auf $\mathbb{P}_n$ ist ein Beispiel für solche Bündel. Der generische Spaltungstyp

$$a_{\Omega^1_{\mathbb{P}_n}(1)} = (0, \ldots, 0, -1)$$

ist „generisch". Die Einschränkung des Cotangentialbündels auf allgemeine Hyperebenen $H \subset \mathbb{P}_n$ ist instabil, denn

$$\Omega^1_{\mathbb{P}_n}(1) | H \cong \mathcal{O}_H \oplus \Omega^1_H(1).$$

Es wird gezeigt, daß für stabile n-Bündel auf $\mathbb{P}_n$ mit $c_1 = -1$ und $c_2 = 1$ mit instabiler Einschränkung auf allgemeine Hyperebenen $H \subset \mathbb{P}_n$ die obige Vermutung richtig ist. Genauer gilt der folgende Satz:

Satz. *Ist E eine normierte semistabile reflexive Garbe vom Rang n auf* $\mathbb{P}_n$, *dann gilt:*

Die Einschränkung auf allgemeine Hyperebenen ist semistabil, oder

$$E \cong \Omega^1_{\mathbb{P}_n}(1) \quad bzw. \quad E \cong T_{\mathbb{P}_n}(-2).$$

Dabei heißt eine torsionsfreie Garbe E vom Rang r auf $\mathbb{P}_n$ *normiert, falls*

$$-r + 1 \leqq c_1 E \leqq 0.$$

Allgemeiner wird vermutet, daß eine stabile reflexive Garbe E vom Rang n auf $\mathbb{P}_n$ mit $c_1 = -1$ und $c_2 = 1$ isomorph zu $\Omega^1_{\mathbb{P}_n}(1)$ ist. Die Einschränkung einer solchen Garbe auf 2-dimensionale lineare Teilräume $H \subset \mathbb{P}_n$ ist instabil. Durch ein genaues Studium der Harder-Narasimhan-Filtrierung

$$0 = E_0 \subsetneqq E_1 \subsetneqq \ldots \subsetneqq E_k = E_H$$

($H \subset \mathbb{P}_n$ ist eine Ebene) kann man zeigen, daß für $n \leqq 8$ die Vermutung richtig ist (vgl. [5]).

2. Einschränkung semistabiler n-Bündel auf $\mathbb{P}_n$

Wie üblich nennen wir eine torsionsfreie Garbe E vom Rang r auf $\mathbb{P}_n$ normiert, wenn

$$-r+1 \leqq c_1 E \leqq 0.$$

Wir zeigen, daß eine semistabile normierte reflexive Garbe vom Rang n auf $\mathbb{P}_n$ semistabile Einschränkungen auf allgemeine Hyperebenen $H \subset \mathbb{P}_n$ hat, mit den evidenten Ausnahmen

$$\Omega^1_{\mathbb{P}_n}(1) \quad \text{bzw.} \quad T_{\mathbb{P}_n}(-2).$$

(2.1) **Lemma.** *Ist E eine normierte semistabile torsionsfreie Garbe vom Rang n auf $\mathbb{P}_n$, dann ist die Einschränkung E_H auf allgemeine Hyperebenen $H \subset \mathbb{P}_n$ entweder semistabil oder*

$$c_1 E \in \{-1, -n+1\}.$$

Beweis. Ist $H \subset \mathbb{P}_n$ eine allgemeine Hyperebene und ist E_H nicht semistabil, so hat E_H die Harder-Narasimhan-Filtrierung

$$0 = E_0 \subsetneqq E_1 \subsetneqq E_2 = E_H$$

und

$$rg E_1 \in \{1, n-1\}.$$

(vgl. [6], p. 86, Lemma 5.2).

Setzt man nun:

$$Q_i := E_i / E_{i-1}, \quad q_i := rg Q_i \quad \text{und} \quad d_i := c_1 Q_i$$

für $i = 1, 2$, so gilt:

$$0 < \frac{d_1}{q_1} - \frac{d_2}{q_2} \leqq \frac{1}{n-1} \tag{2.2}$$

(vgl. [7], p. 331, Theorem 3.1).

Ist $rg E_1 = 1$, dann folgt aus (2.2) mit $d_1 + d_2 = c_1$

$$\frac{c_1}{n} < d_1 \leqq \frac{1}{n} + \frac{c_1}{n}.$$

Da $-n+1 \leqq c_1 \leqq 0$ ist, erhält man

$$-1 + \frac{1}{n} < d_1 \leqq \frac{1}{n}$$

und, da $d_1 \in \mathbb{Z}$ ist, folgt $d_1 = 0$. Setzt man dies in (2.2) ein, so folgt sofort

$$c_1 = -1.$$

Ist $rg E_1 = n-1$, so ergibt sich aus (2.2)

$$1-\frac{1}{n}c_1 < d_1 \leqq \frac{1}{n} + \left(1-\frac{1}{n}\right)c_1.$$

Da $c_1 \geqq -n+1$ ist, folgt $d_1 \geqq c_1+1$ und $d_1 \leqq c_1+1$, also muß $d_1 = c_1+1$ gelten. Aus (2.2) folgt dann

$$c_1 E = -n+1.$$

Damit ist alles gezeigt.

(2.3) **Lemma.** *Ist E eine normierte semistabile reflexive Garbe vom Rang n auf $\mathbb{P}_n$ ($n \geqq 3$), dann ist die Einschränkung E_H auf allgemeine Hyperebenen $H \subset \mathbb{P}_n$ semistabil oder*

$$E \cong \Omega^1_{\mathbb{P}_n}(1) \quad bzw. \quad E \cong T_{\mathbb{P}_n}(-2).$$

Beweis. Lemma (2.1) liefert bereits die Semistabilität der Einschränkung, falls $c_1 E \notin \{-1, -n+1\}$ ist. Ist E_H instabil, und ist $c_1 E = -n+1$, dann gilt:

$$c_1 E^\vee(-1) = -1.$$

Es genügt daher zu zeigen, daß eine semistabile reflexive Garbe vom Rang n auf $\mathbb{P}_n$ mit $c_1 = -1$ und instabiler Einschränkung E_H isomorph zu $\Omega^1_{\mathbb{P}_n}(1)$ ist. Im Falle $n=3$ wurde dies von Schneider gezeigt (vgl. [10], Satz 3.1).

Aus dem Beweis von Lemma (2.1) erhält man die Harder-Narasimhan-Filtrierung von E_H:

$$O \to F \to E_H \to Q \to O. \tag{2.4}$$

Dabei gilt $c_1 F = 0$, $rg F = 1$, $c_1 Q = -1$, $rg Q = n-1$. Da Q torsionsfrei ist, und E_H reflexiv ist, ist auch F reflexiv und somit gilt

$$F \cong \mathcal{O}_H.$$

Da Q stabil ist, folgt $h^0(Q) = 0$ und somit

$$h^0(E_H) = 1.$$

Aus (2.4) folgt weiter $h^0(E_H(-1)) = 0$ und aus der Stabilität von E folgt $h^0(E) = 0$.

Mit der Methode von Gruson-Peskine (vgl. [10], pp. 181 und [2], p. 550, Proposition 3.2) zeigt man nun

$$E \cong \Omega^1_{\mathbb{P}_n}(1).$$

Man hat die exakte Sequenz

$$0 \to \mathcal{O}_H \to E_H \to Q \to 0$$

und einen nichttrivialen Morphismus

$$\psi_H\colon\ \Omega^1_H(1)\to Q$$

(vgl. [2], p. 550, Proposition 3.2). Man erhält das kommutative Diagramm

$$\text{(2.5)}\qquad \begin{tikzcd} \Omega^1_H(1) \arrow[r, "\psi_H"] \arrow[dr, "f"'] & Q_H \arrow[d] \\ & (Q_H)^{\vee\vee} \end{tikzcd}$$

f ist ein nichttrivialer Morphismus zwischen stabilen reflexiven Garben gleichen Ranges und gleicher erster Chernklasse, dann ist f ein Isomorphismus und es folgt, daß ψ_H ein Isomorphismus ist. Dann gilt $Q\cong\Omega^1_H(1)$, und man hat die spaltende exakte Sequenz

$$0\to\mathcal{O}_H\to E_H\to\Omega^1_H(1)\to 0.$$

Daraus folgt (vgl. [10], Lemma 1.1)

$$E\cong\Omega^1_{\mathbb{P}_n}(1).$$

Nützlich ist das folgende Lemma.

(2.6) **Lemma.** *Ist E ein normiertes r-Bündel auf $\mathbb{P}_n$, so ist E stabil, falls für* $1\leqq s\leqq r-1$

$$H^0((\textstyle\bigwedge^s E)_{\text{norm}})=0$$

ist.

Beweis. Es sei $F\subset E$ eine reflexive Untergarbe mit torsionsfreiem Quotienten Q. Man hat dann

$$0\to F\to E\to Q\to 0.$$

Nun gilt $\det(F)=(\bigwedge^s F)^{\vee\vee}\subset\bigwedge^s E$ und man erhält einen nichttrivialen Schnitt

$$\text{(2.7)}\qquad \mathcal{O}\to\textstyle\bigwedge^s \mathrm{E}\otimes\mathcal{O}_{\mathbb{P}_n}(-c_1F).$$

Nun gilt

$$c_1\textstyle\bigwedge^s E=\binom{r-1}{s-1}c_1E$$

und für alle $k\in\mathbb{Z}$

$$c_1(\textstyle\bigwedge^s E\otimes\mathcal{O}_{\mathbb{P}_n}(k))=\binom{r-1}{s-1}\left(c_1E+\frac{r}{s}\cdot k\right).$$

Nun gilt

$$\binom{r-1}{s-1}\left(c_1E+\frac{r}{s}k\right)\leqq 0,$$

falls $k\leqq -\frac{sc_1}{r}$ ist. Man setzt nun $k_0:=\left[-\frac{sc_1}{r}\right]$ ($[x]$ bezeichnet wie üblich die größte ganze Zahl $l\leqq x$).

Es gilt dann

$$(\bigwedge^s E)_{\text{norm}} = \bigwedge^s E \otimes \mathcal{O}_{\mathbb{P}_n}(k_0).$$

Aus (2.7) entnimmt man dann, wegen der Voraussetzungen:

$$-c_1 F > \left[-\frac{s c_1}{r}\right],$$

also

$$\mu(F) < -\frac{1}{s}\left[-\frac{s c_1}{r}\right].$$

Wäre nun

$$\mu(E) \leqq \mu(F) < -\frac{1}{s}\left[-\frac{s c_1}{r}\right],$$

so folgte

$$-\frac{s c_1}{r} \geqq -c_1 F > \left[-\frac{s c_1}{r}\right]$$

und das ist unmöglich, da $-c_1 F \in \mathbb{Z}$ ist. Also gilt für alle Untergarben F mit $1 \leqq rg F \leqq r-1$

$$\mu(F) < \mu(E),$$

und das ist die Behauptung.

3. Generischer Spaltungstyp und zweite Chernklasse bei stabilen 4-Bündeln auf $\mathbb{P}_4$ mit $c_1 = -1$ und $c_1 = -3$

Es werden stabile 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = -1$ untersucht, da man die entsprechenden Resultate für $c_1 = -3$ einfach durch den Prozeß

$$E \rightsquigarrow E^\vee(-1)$$

aus den Resultaten für $c_1 = -1$ erhält.

Wegen Lemma (2.3) genügt es, stabile 4-Bündel mit $c_1 = -1$ und stabiler Einschränkung auf Hyperebene $H \subset \mathbb{P}_4$ zu untersuchen.

(3.1) **Lemma.** *Ist E eine torsionsfreie Garbe vom Rang $r \geqq 2$ auf $\mathbb{P}_2$ mit $c_1 = -1$ und $h^0(E) = 0$ sowie $h^0(E^\vee(-1)) = 0$, dann gilt:*

$$c_2 E \geqq r - 1$$

und $c_2 E = 1$ genau dann, wenn $r = 2$ und $E \cong \Omega^1_{\mathbb{P}_2}(1)$.

Beweis. Man hat das Hilbertpolynom

$$\chi(E(k)) = r\binom{k+2}{2} - (k+1) - c_2 \tag{3.2}$$

für alle $k \in \mathbb{Z}$. Dann folgt:

$$c_2 - r + 1 = h^1(E) \geqq 0,$$

also

$$c_2 \geqq r-1 \geqq 1.$$

Ist $c_2=1$, so muß $r=2$ gelten, und man erhält das E_1^{pq}-Diagramm:

$$\begin{array}{|c|c|c|}\hline 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \quad h^p(E(q)) \quad \begin{array}{r} 0 \leqq p \leqq 2 \\ -2 \leqq q \leqq 0. \end{array}$$

Der Satz von Beilinson liefert dann:

$$E \cong \Omega^1_{\mathbb{P}_2}(1).$$

Die Umkehrung ist trivial.

Aus diesem Lemma folgt sofort:

Ist E ein stabiles 4-Bündel auf $\mathbb{P}_4$ mit stabiler Einschränkung auf allgemeine Ebenen $H \subset \mathbb{P}_4$, so gilt

(3.3) $$c_2 E \geqq 3.$$

Für das weitere benötigen wir folgendes Lemma:

(3.4) **Lemma.** *Ist E eine semistabile reflexive Garbe vom Rand r auf $\mathbb{P}_n$ mit $c_1=0$ und $h^0(E)=s$, $1 \leqq s \leqq r$, so gibt es eine exakte Sequenz*

$$0 \to \mathcal{O}_{\mathbb{P}_n}^{\oplus s} \to E \to Q \to 0.$$

Q ist semistabil und torsionsfrei.

Der Beweis verläuft ähnlich wie in [9] (vgl. [9], p. 215, Hilfssatz).

(3.5) **Corollar.** *Ist E eine semistabile reflexive Garbe vom Rang r auf $\mathbb{P}_n$ mit $c_1=0$, so folgt*

$$E \cong \mathcal{O}_{\mathbb{P}_n}^{\oplus r} \quad \textit{oder} \quad h^0(E) \leqq r-1.$$

(Für den Fall der Vektorraumbündel vgl. [9], p. 215, Hilfssatz.)

(3.6) **Corollar.** *Ist E eine nicht triviale semistabile reflexive Garbe vom Rang r auf $\mathbb{P}_2$ mit $c_1=0$, dann gilt:*

$$c_2 \geqq 1$$

und ist $c_2=1$, so gilt $a_E=(0, 0, \ldots, 0)$.

Beweis. Aus der Semistabilität folgt mit dem Satz von Riemann-Roch

$$h^0(E)-h^1(E)=r-c_2 \leqq r-1-h^1(E),$$

also $c_2 E \geqq 1$.

Ist $c_2 E=1$, so folgt $h^0(E) \geqq r-1$, also $h^0(E)=r-1$ und die Behauptung folgt aus Lemma (3.4).

(3.7) **Lemma.** *Ist E ein stabiles 4-Bündel mit $c_1=-1$ auf $\mathbb{P}_4$ und $E \neq \Omega^1_{\mathbb{P}_4}(1)$, dann gelten die Aussagen:*

(i) *Ist die Einschränkung E_H auf allgemeine Ebenen $H \subset \mathbb{P}_4$ stabil, so gilt $c_2 E \geqq 3$.*

(ii) *Ist E_H instabil für alle Ebenen $H \subset \mathbb{P}_4$, so gilt $c_2 E \geqq 2$ und $a_E=(0,0,0,-1)$.*

(3.8) **Bemerkung.** Ist $c_2 E=2$, so ist $a_E=(0,0,0,-1)$, also „generisch". Ist $a_E=(1,0,-1,-1)$, so ist E_H wegen (3.7)(ii) stabil für allgemeine Ebenen $H \subset \mathbb{P}_4$.

Beweis. Wegen Lemma (2.3) ist die Einschränkung $E_{H'}$ auf allgemeine Hyperebenen $H' \subset \mathbb{P}_4$ stabil.

zu (ii): Da E_H instabil ist, hat E_H die Harder-Narasimhan-Filtrierung

$$0=E_0 \subsetneqq E_1 \subsetneqq \ldots \subsetneqq E_k=E_H.$$

Mit Hilfe des Satzes von Grauert-Mülich-Spindler (vgl. [7], p. 331, Theorem 3.1) errechnet man die folgenden Möglichkeiten:

(3.9) $$0 \to E_1 \to E_H \to Q \to 0,$$

mit $rg E_1=2$, $c_1 E_1=0$, $rg Q=2$, $c_1 Q=-1$. E_1 ist reflexiv und semistabil, Q ist torsionsfrei und stabil.

Oder

(3.10) $$0 \to \mathcal{O}_H \to E_H \to Q \to 0,$$

wobei Q eine stabile torsionsfreie Garbe vom Rang 3 auf H mit $c_1 Q=-1$ ist.

Wir untersuchen zunächst die Filtrierung (3.9).

Aus Lemma (3.1) folgt:

$$c_2 Q \geqq 1.$$

Mit Corollar (3.5) erhält man:

$$E_1 \cong \mathcal{O}_H^{\oplus 2} \quad \text{oder} \quad h^0(E_1) \leqq 1.$$

Ist $E_1 \cong \mathcal{O}_H^{\oplus 2}$ und gilt $c_2 Q=1$, so folgt aus Lemma (3.1) $Q \cong \Omega^1_H(1)$. Dann ist aber

$$E_H \cong \mathcal{O}_H^{\oplus 2} \oplus \Omega^1_H(1).$$

Dann muß aber $E \cong \Omega^1_{\mathbb{P}_4}(1)$ sein, was der Stabilität von $E_{H'}$ widerspricht.

Also gibt es in (3.9) nur die Möglichkeiten

$$0 \to \mathcal{O}_H^{\oplus 2} \to E_H \to Q \to 0$$

mit $c_2 Q \geqq 2$ oder

$$0 \to E_1 \to E_H \to Q' \to 0$$

mit $h^0(E_1) \leqq 1$ und $c_2 Q' \geqq 1$. Da dann $c_2 E_1 \geqq 1$ ist, gilt in beiden Fällen

$$c_2 E \geqq 2.$$

Wendet man den Satz von Grauert-Mülich auf Q bzw. E_1 und Q' an, so folgt

$$a_E = (0, 0, 0, -1).$$

Zu (i). Da E_H stabil ist, folgt mit (3.3) sofort $c_2 E \geqq 3$. Das war zu zeigen.

Wir geben nun ein Beispiel für ein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = -1$ und $c_2 = 2$ an, das dann wegen (3.8) den Spaltungstyp $a_E = (0, 0, 0, -1)$ hat.

Man betrachtet das global erzeugte 6-Bündel

$$\Omega^2_{\mathbb{P}_4}(3).$$

Nach einem Satz von Serre (vgl. [8], p. 81, 4.3.1. Lemma) gibt es ein triviales Unterbündel vom Rang 2 und man erhält die exakte Sequenz:

$$(*) \qquad 0 \to \mathcal{O}_{\mathbb{P}_4}(-1)^{\oplus 2} \to \Omega^2_{\mathbb{P}_4}(2) \to E \to 0.$$

E ist ein 4-Bündel mit $c_1 E = -1$ und $c_2 E = 2$. Aus den Bottformeln entnimmt man $h^0(\Omega^2_{\mathbb{P}_4}(2)) = 0$. Dann ist $h^0(E) = 0$. Da

$$E^\vee(-1) \subset \Lambda^2(T_{\mathbb{P}_4})(-3) \cong \Omega^2_{\mathbb{P}_4}(2)$$

gilt auch $h^0(E^\vee(-1)) = 0$.

Wegen Lemma (2.6) ist E stabil, falls noch

$$H^0(\Lambda^2 E) = 0$$

gilt.

Aus der Sequenz (*) erhält man eine Filtrierung von $\Omega^2_{\mathbb{P}_4}(2)$ und daraus die exakten Sequenzen

$$0 \to \Lambda^2(T_{\mathbb{P}_4}(-3)) \to F^1 \to E \otimes T_{\mathbb{P}_4}(-3) \to 0,$$
$$0 \to F^1 \to \mathcal{O}_{\mathbb{P}_4}(-2)^{\oplus 28} \to \Lambda^2 E \to 0.$$

Da $H^1(\Lambda^2(T_{\mathbb{P}_4}(-3))) = 0$ und $H^1(E \otimes T_{\mathbb{P}_4}(-3)) = 0$ sind, gilt $H^1(F^1) = H^0(\Lambda^2 E) = 0$. Also ist E stabil.

(3.11) **Lemma.** *Ist E ein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = -3$, so gelten die Aussagen:*

(i) *Ist $a_E = (0, -1, -1, -1)$, so ist $c_2 E \geqq 5$.*

(ii) *Ist $a_E = (0, 0, -1, -2)$, so ist $c_2 E \geqq 6$.*

Beweis. Man wende Lemma (3.7) auf $E^\vee(-1)$ an.

4. Generischer Spaltungstyp und zweite Chernklasse stabiler 4-Bündel mit $c_1 = -2$

Nach Lemma (2.3) ist die Einschränkung eines stabilen 4-Bündels auf $\mathbb{P}_4$ mit $c_1 = -2$ auf allgemeine Hyperebenen $H' \subset \mathbb{P}_4$ semistabil. Spaltungstyp und $c_2 E$ hängen wie folgt zusammen:

(4.1) **Lemma.** *Ist E ein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = -2$, dann gelten die Aussagen:*

(i) *Ist $a_E = (0, 0, -1, -1)$, so ist $c_2 E \geqq 3$.*

(ii) *Ist $a_E = (1, 0, -1, -2)$, so ist $c_2 E \geqq 4$.*

(iii) *Ist $a_E = (1, 0, -1, -2)$, so ist die Einschränkung E_H auf allgemeine Ebenen $H \subset \mathbb{P}_4$ stabil.*

Beweis. Ist die Einschränkung E_H auf allgemeine Ebenen $H \subset \mathbb{P}_4$ semistabil, dann gilt bekanntlich

$$h^0(E_H) = 0 \quad \text{und} \quad h^0(E_H^\vee(-1)) = 0. \tag{4.2}$$

Da

$$\chi(E_H(-2)) = -h^1(E_H(-2)) = 3 - c_2$$

ist, folgt

$$c_2 E \geqq 3.$$

Ist $a_E = (1, 0, -1, -2)$, so gilt für allgemeine Geraden $L \subset \mathbb{P}_4$

$$h^0(E_L) = 3.$$

Es seien nun $H \subset \mathbb{P}_4$ eine Ebene, die eine allgemeine Gerade $L \subset H$ enthält, und E_H semistabil. Dann folgt aus (4.2) mit der Einschränkungssequenz

$$0 \to E_H(-1) \to E_H \to E_L \to 0$$

sofort

$$3 = h^0(E_L) \leqq h^1(E_H(-1)) = -\chi(E_H(-1)) = c_2 - 1,$$

also

$$c_2 E \geqq 4.$$

Somit sind (i) und (ii) für Bündel mit semistabiler Einschränkung E_H gezeigt.

Nun zur instabilen Einschränkung!

Mit dem Satz von Grauert-Mülich-Spindler (vgl. [7], p. 331, Theorem 3.1) errechnet man für die Harder-Narasimhan-Filtrierung:

$$0 = E_0 \subsetneqq E_1 \subsetneqq E_2 \subsetneqq E_3 = E_H$$

mit $Q_i = E_i / E_{i-1}$, $q_i = rg\, Q_i$ und $d_i = c_1 Q_i$:

$$(q_1, q_2, q_3) = (1, 2, 1) \quad \text{und} \quad (d_1, d_2, d_3) = (0, -1, -1).$$

Die Q_i sind semistabile torsionsfreie Garben und die E_i Vektorraumbündel.

Man hat dann die exakten Sequenzen:

$$0 \to E_1 \to E_2 \to Q_2 \to 0,$$
$$0 \to E_2 \to E_H \to Q_3 \to 0.$$

Dann sind $E_1 \cong \mathcal{O}_H$ und $Q_3 \cong \mathcal{I}_Z(-1)$, wobei $\mathcal{I}_Z$ das Ideal eines 0-dimensionalen Unterraumes ist.

Dann gilt $c_2\mathcal{I}_Z(-1)\geqq 0$ und aus Lemma (3.1) folgt $c_2Q_2\geqq 1$. Dann erhält man

(4.3) $$c_2E=c_2Q_2+c_2\mathcal{I}_Z(-1)+1\geqq 2.$$

Da Q_2 stabil ist, so liefert der Satz von Grauert-Mülich:

$$a_E=(0,0,-1,-1).$$

Zum Beweis von (i) müssen wir noch zeigen, daß es kein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1=-2$ und $c_2=2$ gibt.

Ist $c_2E=2$, so folgt aus (4.2) $c_2Q_2=1$ und $c_2\mathcal{I}_Z(-1)=0$. Dann folgt (vgl. Lemma (3.2))

$$Q_2\cong\Omega_H^1(1) \quad \text{und} \quad Z=\varnothing.$$

Für alle Ebenen $H\subset\mathbb{P}_4$ gilt dann

(4.4) $$E_H\cong\mathcal{O}_H\oplus\mathcal{O}_H(-1)\oplus\Omega_H^1(1).$$

Man wählt nun eine allgemeine Hyperebene H'. Es sei $H\subset H'$ eine Ebene. Man hat die exakte Sequenz:

$$\begin{aligned}H^0(E_H(k))&\to H^1(E_{H'}(k-1))\to H^1(E_{H'}(k))\to H^1(E_H(k))\\ &\to H^2(E_{H'}(k-1))\to H^2(E_{H'}(k))\to H^2(E_H(k)).\end{aligned}$$

Aus (4.4) folgt daraus mittels Theorem B

$$\begin{aligned}h^2(E_{H'}(k))&=0 \quad \text{für } k\geqq -1,\\ h^1(E_{H'}(k))&=0 \quad \text{für } k\leqq -2.\end{aligned}$$

Die Semistabilität von $E_{H'}$ liefert:

$$\begin{aligned}h^0(E_{H'}(k))&=0 \quad k\leqq 0,\\ h^3(E_{H'}(k))&=0 \quad k\geqq -3.\end{aligned}$$

Mittels des Satzes von Riemann-Roch erhält man:

$$h^2(E_{H'}(-3))=\chi(E_{H'}(-3))=\tfrac{1}{2}c_3E\geqq 0$$

und

$$-h^1(E_{H'})=\chi(E_{H'})=1+\tfrac{1}{2}c_3E\leqq 0,$$

woraus sich ein Widerspruch ergibt.

Also gilt $c_2E\geqq 3$, wenn $a_E=(0,0,-1,-1)$ ist.

Zu (iii): Ist $a_E=(1,0,-1,-2)$, so ist E_H semistabil auf allgemeine Ebenen, wie oben gezeigt.

Nimmt man nun an, daß E_H nicht stabil ist, so gibt es eine exakte Sequenz

$$0\to F\to E_H\to Q\to 0$$

mit $rgF = rgQ = 2$ und $\mu(F) = \mu(E_H) = \mu(Q) = -\frac{1}{2}$ (man beachte (4.2)). Dann sind F und Q stabil mit $c_1 F = c_1 Q = -1$. Wendet man den Satz von Grauert-Mülich auf F und Q an, dann folgt

$$a_E = (0, 0, -1, -1).$$

Das ist ein Widerspruch und somit ist alles gezeigt.

Als nächstes konstruieren wir ein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = -2$ und $c_2 = 3$.

$T_{\mathbb{P}_4}(-1)^{\oplus 2}$ ist ein global erzeugtes Vektorraumbündel vom Rang 8. Man hat ein triviales Unterbündel vom Rang 4 (vgl. [8], p. 81, 4.3.1. Lemma) und somit die exakte Sequenz

$$(**) \qquad 0 \to \mathcal{O}_{\mathbb{P}_4}(-1)^{\oplus 4} \to T_{\mathbb{P}_4}(-2)^{\oplus 2} \to E \to 0$$

mit einem 4-Bündel E, das die Chernklassen $c_1 E = -2$ $c_2 E = 3$ hat. Weiter gilt $h^0(E) = 0$ und $h^0(E^\vee(-1)) = 0$.

E ist stabil, falls auch

$$H^0(\Lambda^2 E(1)) = 0$$

ist (vgl. Lemma (2.6)). Aus der Sequenz $(**)$ erhält man bekanntlich eine Filtrierung von $\Lambda^2(T_{\mathbb{P}_4}(-2))^{\oplus 2}$ und daraus die exakten Sequenzen:

$$\begin{aligned} 0 \to \mathcal{O}_{\mathbb{P}_4}(-3) \to \quad & F^1(1) \quad \to E(-1)^{\oplus 2} \to 0, \\ 0 \to \quad F^1(1) \quad & \to \mathcal{O}_{\mathbb{P}_4}(-2)^{\oplus 15} \to \Lambda^2 E(1) \to 0. \end{aligned}$$

Da $H^1(E(k)) = 0$ für alle $k \in \mathbb{Z}$, folgt dann

$$H^0(\Lambda^2 E(1)) = H^1(F^1(1)) = H^1(E(-1))^{\oplus 2} = 0.$$

Also ist E stabil und $a_E = (0, 0, -1, -1)$.

5. Generischer Spaltungstyp und zweite Chernklasse stabiler 4-Bündel mit $c_1 = 0$

(5.1) **Lemma.** *Ist E ein stabiles 4-Bündel mit $c_1 = 0$, dann gelten die Aussagen:*

(i) *Ist $a_E = (0, 0, 0, 0)$, so gilt $c_2 E \geqq 2$.*

(ii) *Ist $a_E = (1, 0, 0, -1)$, so gilt $c_2 E \geqq 3$.*

(iii) *Ist die Einschränkung E_H auf allgemeine Ebenen $H \subset \mathbb{P}_4$ stabil, dann folgt $c_2 E \geqq 4$.*

Beweis. Nach einem Satz von Maruyama (vgl. [7], Corollary 3.1.1) ist die Einschränkung auf allgemeine Ebenen $H \subset \mathbb{P}_4$ semistabil, also gilt:

$$h^0(E_H(-1)) = 0 \quad \text{und} \quad h^0(E_H^\vee(-1)) = 0.$$

Da E als stabiles Bündel nichttrivial sein kann, muß $h^0(E_H) \leqq 3$ gelten (vgl. Corollar (3.5)).

Es sei $\varepsilon := h^0(E_H) \in \{2, 3\}$, dann hat man die exakte Sequenz

$$0 \to \mathcal{O}_H^{\oplus \varepsilon} \to E_H \to Q \to 0.$$

Q ist semistabil mit $c_1 Q = 0$ (vgl. Lemma 3.4.).

Der Satz von Grauert-Mülich erzwingt dann

$$a_E = (0, 0, 0, 0).$$

Ist $a_E = (1, 0, 0, -1)$, so muß $h^0(E_H) \leqq 1$ gelten. Mit dem Satz von Riemann-Roch folgt dann:

$$h^0(E_H) - h^1(E_H) = 4 - c_2.$$

Daraus erhält man (ii) und (iii) im Falle $a_E = (1, 0, 0, -1)$. Ist der generische Spaltungstyp trivial, so folgen analog

$$c_2 E \geqq 1$$

bzw.

$$c_2 E \geqq 4.$$

Es bleibt noch zu zeigen, daß es kein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = 0$ und $c_2 = 1$ gibt.

Spindler hat gezeigt, daß 4-Bündel auf $\mathbb{P}_4$ mit trivialem Spaltungstyp mit $c_1 = 0$, $c_2 = 1$, $c_3 = 0$ durch die Extension

$$0 \to \mathcal{O}_{\mathbb{P}_4} \to E \to E' \to 0,$$

mit

$$0 \to \mathcal{O}_{\mathbb{P}_4}(-1) \to \Omega^1_{\mathbb{P}_4}(1) \to E' \to 0$$

gegeben sind, also insbesondere nicht stabil sind (vgl. [14], p. 17, Satz 2.3.3). Es genügt also zu zeigen, daß $c_3 = 0$ sein muß. Sind H eine allgemeine Ebene und $H' \subset \mathbb{P}_4$ eine allgemeine Hyperebene mit $H \subset H'$, dann gelten:

(5.2) (i) $c_3 E \in \{-2, 0, 2\}$,

(ii) $h^0(E_H) = 3$ und $h^0(E_H(-1)) = 0$,

(iii) $h^0(E_{H'}) \geqq \frac{1}{2} c_3 + 2$.

(vgl. [13], p. 116, Lemma 7.2.).

Dann muß $h^0(E_{H'}) \leqq 2$ gelten, denn wäre $h^0(E_{H'}) = 3$, so folgte mit dem „glueing lemma“ (vgl. [2], p. 544, Proposition 1.2) $h^0(E) = 3$ im Widerspruch zur Stabilität.

Dann folgt

$$2 \geqq h^0(E_{H'}) \geqq \begin{cases} 3 & c_3 = 2 \\ 2 & c_3 = 0. \end{cases}$$

Also ein Widerspruch, falls $c_3 = 2$ ist.

Ist $c_3 = -2$, so gilt $c_3 E^\vee = 2$ und wie oben folgt ein Widerspruch. Also gilt $c_3 E = 0$, d.h. es gibt kein stabiles 4-Bündel auf $\mathbb{P}_4$ mit $c_1 = 0$ und $c_2 = 1$ (vgl.

auch [14], Bemerkung 5.2.6.). Zum Schluß wollen wir noch ein Beispiel eines stabilen 4-Bündels auf $\mathbb{P}_4$ mit $c_1=0$ und $c_2=2$ angeben.

Dazu betrachtet man die Extension

(5.3) $$0 \to \Omega^1_{\mathbb{P}_4}(1) \to E \to \mathcal{O}_{H_0}(-1) \to 0,$$

wobei $H_0 \subset \mathbb{P}_4$ eine festgewählte Hyperebene ist. Spindler hat gezeigt, daß es solche Extensionen mit einem Vektorraumbündel E gibt, das die Bedingungen

(5.4) $$h^0(E)=0 \quad \text{und} \quad h^0(E^\vee)=0$$

erfüllt (vgl. [14], p. 91).

Wir zeigen, daß diese Bündel stabil mit $a_E=(0,0,0,0)$ sind.

Wir zeigen zunächst, daß E semistabil ist. Ist $F \subset E^\vee$ eine reflexive Untergarbe mit torsionsfreiem Quotienten Q, so folgt aus

$$F \subset E^\vee \subset T_{\mathbb{P}_4}(-1)$$

und der Stabilität von $T_{\mathbb{P}_4}(-1)$:

$$\mu(F) < \tfrac{1}{4}.$$

Da $1 \leqq rg F \leqq 3$, folgt $c_1 F \leqq 0$, also $\mu(F) \leqq 0$. Dann ist $E^\vee$ und somit E semistabil.

Ist E nicht stabil, so gibt es wegen (5.4) eine reflexive Untergarbe $F \subset E^\vee$ vom Rang 2 mit $\mu(F)=\mu(E^\vee)=0$.

Man hat die exakte Sequenz

$$0 \to F \to E^\vee \to Q \to 0$$

mit einer torsionsfreien Garbe Q vom Rang 2 und $\mu(Q)=0$. Dann sind F und Q stabil mit $c_1 F = c_1 Q = 0$, denn

$$h^0(F) = h^0(Q^\vee) = 0.$$

Aus dem Satz von Grauert-Mülich folgt dann

$$a_E = (0,0,0,0).$$

Dann haben alle durch (5.3) gegebenen 4-Bündel trivialen Spaltungstyp, denn wäre $a_E=(1,0,0,-1)$, so müßte E stabil sein und aus Lemma (5.1) folgte dann $c_2 E \geqq 3$, was unmöglich ist.

Es bleibt noch zu zeigen, daß die 4-Bündel mit (5.3) und (5.4) stabil sind.

Nimmt man an, daß es eine reflexive Untergarbe $F \subset E$ vom Rang 2 (vgl. (5.4)) gibt mit $\mu(F)=\mu(E)=0$, so erhält man die exakte Sequenz

$$0 \to F \to E \to Q \to 0$$

mit einer stabilen reflexiven Garbe F mit $c_1=0$ und einer stabilen torsionsfreien Garbe Q mit $c_1=0$ und $rg Q=2$.

Nun gilt

$$c_2 Q^{\vee\vee} \leqq c_2 Q,$$

somit gilt $c_2 F \geqq 1$ und $c_2 Q \geqq 1$. Da

$$c_2 E = c_2 F + c_2 Q$$

gilt, folgt $c_2 F = c_2 Q = 1$.

Man erhält dann einen nichttrivialen Pfeil $F \xrightarrow{\varphi} \mathcal{O}_{H_0}(-1)$,

$$\begin{tikzcd} 0 \arrow[r] & \Omega^1_{\mathbb{P}_4}(1) \arrow[r] & E \arrow[r] & \mathcal{O}_{H_0}(-1) \arrow[r] & 0 \\ & & F \arrow[u, hook] \arrow[ur, "\varphi"'] & & \end{tikzcd}$$

denn wäre $F \subset \Omega^1_{\mathbb{P}_4}(1)$, so folgte aus der Stabilität von $\Omega^1_{\mathbb{P}_4}(1)$

$$\mu(F) < \mu(\Omega^1_{\mathbb{P}_4}(1)) = -\tfrac{1}{4},$$

was unmöglich ist.

Setzt man $G := \mathscr{I}m\,\varphi$, so ist $G \subset \mathcal{O}_{H_0}(-1)$ ein torsionsfreier $\mathcal{O}_{H_0}$-Modul vom Rang 1, also gilt

$$G \cong \mathscr{I}_{Z,H_0}(-k)$$

mit $k \geqq 0$ und einem 2-codimensionalen Unterraum $Z \subset H_0$.

Der Epimorphismus

$$F \to \mathscr{I}_{Z,H_0}(-k)$$

liefert einen nichttrivialen Schnitt

$$\mathcal{O}_{H_0} \to (F_{H_0})^{\vee}(-k),$$

also gilt $H^0((F_{H_0})^{\vee}(-k)) \neq 0$, und somit ist $H_0 \subset \mathbb{P}_4$ eine unstabile Hyperebene für F (vgl. [4], Proposition 9.1. und §9. Die dort angegebenen Beweise lassen sich ohne Mühe auf $\mathbb{P}_n$, $n \geqq 4$ übertragen, so daß wir hier darauf verzichten.)

Es sei $k' := \max\{l : h^0((F_{H_0})^{\vee}(-l)) \neq 0\}$; dann erhält man eine exakte Sequenz

$$0 \to F' \to F \to \mathscr{I}_{Z,H_0}(-k') \to 0.$$

F' ist eine reflexive Garbe vom Rang 2 mit den Chernklassen

$$c_1 F' = -1,$$
$$c_2 F' = 1 - k' \leqq 0,$$

da $k' \geqq k \geqq 1$.

Da $h^0(F') \leqq h^0(F) = 0$ gilt, ist F' stabil und dann folgt $c_2 F' \geqq 1$ (vgl. Lemma (3.1)), also ein Widerspruch. (Dabei ist zu beachten, daß F'_H wegen des Einschränkungssatzes von Maruyama (vgl. [6], p. 74, Theorem 3.1.) auf allgemeinen Ebenen $H \subset \mathbb{P}_4$ stabil ist.

Also ist E stabil und das war zu zeigen.

Literatur

1. Bogomolov, F.A.: Holomorphic tensors and vector bundles on projective varieties. Math. USSR Izvestija **13**, 499-555 (1979)
2. Ein, L., Hartshorne, R., Vogelaar, H.: Restriction theorems for stable rank 3 vector bundles on $\mathbb{P}^n$. Math. Ann. **259**, 541-569 (1982)
3. Hartshorne, R.: Stable vector bundles of rank 2 on $\mathbb{P}^3$. Math. Ann. **238**, 229-280 (1978)
4. Hartshorne, R.: Stable reflexive sheaves. Math. Ann. **254**, 121-176 (1980)
5. Hoppe, H.J.: Rechnungen zu $\Omega^1_{\mathbb{P}_4}(1)$, unveröffentlicht.
6. Maruyama, M.: Boundedness of semistable sheaves of small ranks. Nayoya Math. J. **78**, 65-94 (1980)
7. Maruyama, M.: The Theorem of Grauert-Mülich-Spindler. Math. Ann. **255**, 317-333 (1981)
8. Okonek, Ch., Schneider, M., Spindler, H.: Vector bundles on complex projective spaces, Progress in Math. 3. Boston: Birkhäuser 1980
9. Schneider, M.: Chernklassen semistabiler Vektorraumbündel vom Rang 3 auf dem komplex projektiven Raum. Crelles Journal **315**, 211-220 (1980)
10. Schneider, M.: Einschränkung stabiler Vektorraumbündel vom Rang 3 auf Hyperebenen des projektiven Raumes. Crelles Journal **323**, 177-192 (1981)
11. Spindler, H.: Der Satz von Grauert-Mülich für beliebige semistabile holomorphe Vektorbündel über dem n-dimensionalen komplex-projektiven Raum. Math. Ann. **243**, 131-141 (1979)
12. Spindler, H.: Die Modulräume stabiler 3-Bündel auf $\mathbb{P}_3$ mit den Chernklassen $c_1=0$, $c_3=c_2^2-c_2$. Math. Ann. **256**, 133-143 (1981)
13. Spindler, H.: Ein Satz über die Einschränkung holomorpher Vektorbündel auf $\mathbb{P}_n$ mit $c_1=0$ auf Hyperebenen. Crelles Journal **327**, 93-118 (1981)
14. Spindler, H.: Holomorphe Vektorbündel auf $\mathbb{P}_n$ mit trivialem Spaltungstyp. Habilitationsschrift, Göttingen 1981

Eingegangen in endgültiger Fassung am 4. April 1984

Math. Z. 187, 361-381 (1984)

Mathematische Zeitschrift

Orthomorphisms and Boolean Algebras of Projections

Peter G. Dodds[1] and Ben de Pagter[2]

[1] School of Mathematical Sciences, The Flinders University of South Australia, Bedford Park. S.A. 5042, Australia
[2] Department of Mathematics, Delft University of Technology, Julianalaan 132, 2628 BL Delft, The Netherlands

1. Introduction

The systematic study of operator algebras generated by Boolean algebras of projections was initiated by W.G. Bade [2]. One of the principal results given by Bade is the following reflexivity theorem. A continuous linear operator T on a Banach space X belongs to the strongly closed algebra generated by a σ-complete Boolean algebra $\mathscr{M}$ of projections if and only if T leaves invariant each $\mathscr{M}$-invariant subspace of X. The methods used by Bade depend on the existence of a linear functional with certain very specific properties and do not extend to the setting of locally convex spaces. On the other hand, it was H.H. Schaefer [15] (see also [16, 17]) who first noticed that a key role is played by partial order in the development of spectral theory in locally convex spaces. This idea was further developed by Walsh [20] who showed that the cyclic subspaces of a spectral measure carry an order structure that has very strong and useful properties. It was pointed out in [6], essentially as a consequence of earlier work of Ricker [14], that not only do the cyclic subspaces of a spectral measure have a rich lattice structure, but so also does the closed algebra generated by the range of a spectral measure.

In this paper, it is our intention to systematically investigate the role played by lattice order in the study of operator algebras generated by Boolean algebras of projections. Under some mild and natural restrictions, we show directly and simply that these algebras have the natural structure of an Archimedean f-algebra and we characterize those Boolean algebras of projections for which the algebras generated admit the structure of a Dedekind complete, locally solid complex f-algebra with Lebesgue topology. The relation between the lattice structure on the algebra and the corresponding induced lattice structure on the cyclic subspaces is exploited by using the recently developed theory of orthomorphisms in Riesz spaces as a principal tool to extend the reflexivity theorem of Bade to the locally convex setting. The key idea in our extension is to show that if T is an everywhere defined linear map leaving invariant each closed subspace which is invariant under a strongly equicontinuous Boolean algebra of projections, then the restriction of T to each cyclic subspace is an (order bounded) orthomorphism. In the more familiar setting of Banach spaces, this implies that such a map T is automatically continuous, and this appears to be new.

Our entire approach then is direct, is elementary, and is based on the theory of Riesz spaces. It is convenient to use the monograph [1] for basic information concerning the general theory of locally solid Riesz spaces. We will follow the terminology of [21], in which the reader may find an account of the theory of f-algebras and an introduction to the theory of orthomorphisms. Apart from some basic results from these sources, the present paper is self-contained. We mention finally that the intrinsic methods of the present paper are to be compared with those of [5] where similar problems have been considered from the entirely different viewpoint of the theory of spectral measures.

2. Preliminaries

In the present section we discuss some general results on Riesz spaces, orthomorphisms and f-algebras, which will be useful in this paper. For the terminology used we refer to [1, 11] and [21]. *All Riesz spaces considered will be assumed to be Archimedean.*

For any Riesz space L we denote by $L_{\mathbb{C}} = L + iL$ the complexification of L. As usual, if $h = f + ig \in L_{\mathbb{C}}$, $f, g \in L$, then we denote $f = \operatorname{Re} h$ and $g = \operatorname{Im} h$. We say that the element $f \in L_{\mathbb{C}}$ has an *absolute value*, denoted by $|f|$, if the supremum

$$|f| = \sup\{(\operatorname{Re} f)\cos\theta + (\operatorname{Im} f)\sin\theta \colon 0 \leqq \theta \leqq 2\pi\}$$

exists in L. If $|f|$ exists for all $f \in L_{\mathbb{C}}$, then $L_{\mathbb{C}}$ will be called a *complex Riesz space*. Note that if L is uniformly complete, then $L_{\mathbb{C}}$ is a complex Riesz space (see e.g. [21], Sect. 91).

Let $L_{\mathbb{C}}$ be a complex Riesz space. A subset S of $L_{\mathbb{C}}$ is called *solid* if $|f| \leqq |g|$ with $g \in S$ and $f \in L_{\mathbb{C}}$, implies that $f \in S$. A solid linear subspace of $L_{\mathbb{C}}$ is called an (*order*) *ideal* in $L_{\mathbb{C}}$. If A is an ideal in $L_{\mathbb{C}}$ then $\operatorname{Re} A = \{\operatorname{Re} f \colon f \in A\}$ is an ideal in L and $A = \operatorname{Re} A + i \operatorname{Re} A$. Conversely, if A is an ideal in L, then the complexification $A_{\mathbb{C}} = A + iA$ is an ideal in $L_{\mathbb{C}}$. Thus the ideals in $L_{\mathbb{C}}$ are precisely the complexifications of the ideals in L. Observe that the ideal $A_{\mathbb{C}}$ in $L_{\mathbb{C}}$ is uniformly closed if and only if A is uniformly closed in L. A band in $L_{\mathbb{C}}$ is, by definition, the complexification of a band in L.

Let A be a uniformly closed ideal in L, then L/A is an Archimedean Riesz space (see [11]. Theorem 60.2). Now consider the complex quotient space $L_{\mathbb{C}}/A_{\mathbb{C}}$, then it is clear that $L_{\mathbb{C}}/A_{\mathbb{C}} = L/A + iL/A$, so $L_{\mathbb{C}}/A_{\mathbb{C}}$ is the complexification of L/A. Moreover, $L_{\mathbb{C}}/A_{\mathbb{C}}$ is a complex Riesz space. For any $f \in L_{\mathbb{C}}$ the absolute value of $[f] \in L_{\mathbb{C}}/A_{\mathbb{C}}$ is given by $[|f|]$, i.e. $|[f]| = [|f|]$.

If $L_{\mathbb{C}}$ and $M_{\mathbb{C}}$ are complex Riesz spaces, then the space $\mathscr{L}^{\#}(L_{\mathbb{C}}, M_{\mathbb{C}})$ of all linear operators from $L_{\mathbb{C}}$ into $M_{\mathbb{C}}$ is the complexification of the space $\mathscr{L}^{\#}(L, M)$ of all real linear operators from L into M (see [21] Sect. 92). An operator $T\in\mathscr{L}^{\#}(L_{\mathbb{C}},M_{\mathbb{C}})$ is called *order bounded* if for every $0\leqq u\in L$ there exists $0\leqq v\in M$ such that $|Tf|\leqq v$ whenever $|f|\leqq u$ in $L_{\mathbb{C}}$. The space of all order bounded linear operators from $L_{\mathbb{C}}$ into $M_{\mathbb{C}}$ is denoted by $\mathscr{L}_b(L_{\mathbb{C}},M_{\mathbb{C}})$. It is easily seen that $\mathscr{L}_b(L_{\mathbb{C}},M_{\mathbb{C}})$ is the complexification of $\mathscr{L}_b(L,M)$.

A linear operator T from $L_{\mathbb{C}}$ into $M_{\mathbb{C}}$ is called a Riesz homomorphism if $|Tf|=T|f|$ for all $f\in L_{\mathbb{C}}$. Note that T is a Riesz homomorphism if and only if T is the canonical extension of a Riesz homomorphism from L into M.

Let L be a Riesz space. Recall that a linear operator T from L into itself is called *band preserving* if $Tf\perp g$ whenever $f\perp g$ in L (see e.g. [21], Sect. 139). Any order bounded band preserving operator is called an *orthomorphism*. The Riesz space of all orthomorphisms in L is denoted by $\mathrm{Orth}(L)$. For the general theory of orthomorphisms we refer to [21], Chap. 20.

Let $L_{\mathbb{C}}$ be a complex Riesz space. We say that $f,g\in L_{\mathbb{C}}$ are disjoint, denoted by $f\perp g$, if $|f|\wedge|g|=0$ in L. An operator $T\in\mathscr{L}^{\#}(L_{\mathbb{C}})$ is called band preserving if $Tf\perp g$ whenever $f\perp g$ in $L_{\mathbb{C}}$. It is easy to verify that $T\in\mathscr{L}^{\#}(L_{\mathbb{C}})$ is band preserving if and only if $\operatorname{Re}T$ and $\operatorname{Im}T$ are both band preserving (considered as operators in L). A band preserving operator in $L_{\mathbb{C}}$ which is order bounded is called an orthomorphism, and the collection of all orthomorphism in $L_{\mathbb{C}}$ is denoted by $\mathrm{Orth}(L_{\mathbb{C}})$. It follows from these above observations that $\mathrm{Orth}(L_{\mathbb{C}})$ is the complexification of $\mathrm{Orth}(L)$.

It is not difficult to show that for any $T\in\mathrm{Orth}(L_{\mathbb{C}})$ the absolute value $|T|$ exists and satisfies $|T|u=|Tu|$ for all $0\leqq u\in L$. Moreover for any $T\in\mathrm{Orth}(L_{\mathbb{C}})$ and $f\in L_{\mathbb{C}}$ we have $|Tf|=|T||f|=||T|f|$.

The center $Z(L_{\mathbb{C}})$ of $L_{\mathbb{C}}$ is the space of all $T\in\mathrm{Orth}(L_{\mathbb{C}})$ for which there exists $n\in\mathbb{N}$ such that $|T|\leqq nI$, i.e. $Z(L_{\mathbb{C}})$ is the order ideal in $\mathrm{Orth}(L_{\mathbb{C}})$ generated by I. Note that $Z(L_{\mathbb{C}})$ is the complexification of $Z(L)$.

Let A be a Riesz algebra, i.e. A is a Riesz space as well as an algebra with the additional property that $uv\geqq 0$ for all $0\leqq u$, $v\in A$. Recall that A is called an f-algebra if $u\wedge v=0$ in A implies that $(wu)\wedge v=(uw)\wedge v=0$ for all $0\leqq w\in A$. For general properties of f-algebras we refer to [21], Chap. 20. In particular, recall that any Archimedean f-algebra is commutative ([21], Theorem 140.10), and that any f-algebra with unit element is semi-prime ([21], Theorem 142.5). Moreover, if A is a commutative Riesz algebra with unit element e, then A is an f-algebra if and only if e is a weak order unit (see [4], 12.3.21).

Let A be an f-algebra such that the complexification $A_{\mathbb{C}}$ is a complex Riesz space. The multiplication in A extends in the obvious way to $A_{\mathbb{C}}$ such that $A_{\mathbb{C}}$ is a complex commutative algebra. In this situation $A_{\mathbb{C}}$ is called a *complex f-algebra* (see [3], Sect. 5). Note that $A_{\mathbb{C}}$ has a unit element e if and only if $e\in A$ and e is a unit element in A. For the sake of convenience we list some properties of complex f-algebras with unit element (for the proofs see [3], Sect. 5).

(i) For any $f\in A_{\mathbb{C}}$, $|f|=\sqrt{(\operatorname{Re}f)^2+(\operatorname{Im}f)^2}$.
(ii) If $f\perp g$ in $A_{\mathbb{C}}$, then $fh\perp g$ for all $h\in A_{\mathbb{C}}$.
(iii) $|fg|=|f|\cdot|g|$ for all $f,g\in A_{\mathbb{C}}$.
(iv) $A_{\mathbb{C}}$ is semi-prime, and $f\perp g$ in $A_{\mathbb{C}}$ is equivalent to $fg=0$.

As is well-known, if A is a uniformly complete f-algebra with unit element e and $u\in A$ with $u\geqq e$, then u^{-1} exists in A ([10], Theorem 3.4 or [21], Theorem 146.3). This result extends immediately to the complex case. Indeed, let $A_{\mathbb{C}}$ be a uniformly complete complex f-algebra with unit element e, and let $h=f+ig\in A_{\mathbb{C}}$, $f,g\in A$ such that $|h|\geqq e$. By property (i) above, $\sqrt{f^2+g^2}\geqq e$, so $f^2+g^2\geqq e$ and hence $(f^2+g^2)^{-1}$ exists. Now it is obvious that $h^{-1}=f(f^2+g^2)^{-1}-ig(f^2+g^2)^{-1}$.

The following lemma will be useful.

Lemma 2.1. *Let $A_{\mathbb{C}}$ be a complex f-algebra with unit element e. For an element $p\in A_{\mathbb{C}}$ the following statements are equivalent.*

(i) *p is an idempotent, i.e., $p^2=p$*

(ii) *p is a component of e, i.e., $p\in A$ and $p\wedge(e-p)=0$.*

Proof. (ii)⇒(i). $p\wedge(e-p)=0$ implies $p(e-p)=0$, i.e. $p=p^2$.
(i)⇒(ii). It follows from $p^2=p$ that $p(e-p)=0$, and since $A_{\mathbb{C}}$ is semi-prime, this implies that $p\perp e-p$. Write $p=r+is$, with $r,s\in A$, then $r+is\perp(e-r)-is$. Since $|s|\leqq|r+is|$ and $|s|\leqq|(e-r)-is|$, it follows that $s=0$, so $p=r\in A$. Therefore p is a component of e in A.

For any complex Riesz space $L_{\mathbb{C}}$ the space $\mathrm{Orth}(L_{\mathbb{C}})$ is an algebra with respect to composition as multiplication, and is the complexification of the f-algebra $\mathrm{Orth}(L)$. Hence $\mathrm{Orth}(L_{\mathbb{C}})$ is a complex f-algebra with the identity I as the unit element. Note that the components of I in $\mathrm{Orth}(L_{\mathbb{C}})$ are precisely the order projections in $L_{\mathbb{C}}$. Similarly, $Z(L_{\mathbb{C}})$ is a complex f-algebra with I as the unit element.

Let $L_{\mathbb{C}}$ be a complex Riesz space. We say that $L_{\mathbb{C}}$ is a *complex locally solid Riesz space* if there is a linear topology in $L_{\mathbb{C}}$ for which there exists a neighbourhood base at 0 consisting of solid sets. If $L_{\mathbb{C}}$ is a complex locally solid Riesz space, then it is clear that the mappings $h\mapsto|h|$, $h\mapsto\mathrm{Re}\,h$ and $h\mapsto\mathrm{Im}\,h$ are uniformly continuous. This implies in particular that L is a closed real linear subspace of $L_{\mathbb{C}}$. Furthermore, L is a locally solid Riesz space with respect to the induced topology.

For any Hausdorff locally solid Riesz space L we denote by $\hat{L}$ the topological completion of L. Then $\hat{L}$ is a locally solid Riesz space (see e.g. [1], Theorem 7.1). Now let $L_{\mathbb{C}}$ be a complex Hausdorff locally solid Riesz space with completion $\hat{L}_{\mathbb{C}}$. Observe that $\hat{L}$ is equal to the closure of L in $\hat{L}_{\mathbb{C}}$. In particular, $\hat{L}$ is a locally solid Riesz space with respect to the induced topology of $\hat{L}_{\mathbb{C}}$. The mappings $h\mapsto|h|$, $h\mapsto\mathrm{Re}\,h$ and $h\mapsto\mathrm{Im}\,h$ from $L_{\mathbb{C}}$ into L have unique uniformly continuous extensions to mappings $\hat{h}\mapsto|\hat{h}|$, $\hat{h}\mapsto\mathrm{Re}\,\hat{h}$ and $\hat{h}\mapsto\mathrm{Im}\,\hat{h}$ from $\hat{L}_{\mathbb{C}}$ into L. By continuity we get $\hat{h}=\mathrm{Re}\,\hat{h}+i\,\mathrm{Im}\,\hat{h}$ for all $\hat{h}\in\hat{L}_{\mathbb{C}}$, so $\hat{L}_{\mathbb{C}}=\hat{L}+i\hat{L}$ as a real direct sum. Hence $\hat{L}_{\mathbb{C}}$ is the complexification of $\hat{L}$. Moreover, it is not difficult to show that $\hat{L}_{\mathbb{C}}$ is a complex locally solid Riesz space.

For the sake of convenience, we summarise the above results in the next proposition.

Proposition 2.2. *For any complex Hausdorff locally solid Riesz space $L_{\mathbb{C}}$ the completion $\hat{L}_{\mathbb{C}}$ is a complex locally solid Riesz space, which is the complexification of the Riesz space $\hat{L}$.*

We conclude this section with some results which are of interest in their own right, and which will be useful in the next sections. By the above observations, we may restrict ourselves to considering real spaces only.

Proposition 2.3. *Let L be a Riesz space which is a linear subspace of the Hausdorff quasi-complete topological vector space X, such that the induced topology in L is locally solid. Then the closure $\bar{L}$ of L in X is equal to the completion of L (and hence $\bar{L}$ has the structure of a locally solid Riesz space).*

Proof. Let $\hat{L}$ be the completion of L. Observe that $\hat{L}$ is the closure of L in the completion $\hat{X}$ of X, and that $\bar{L}=\hat{L}\cap X$. As noted before, $\hat{L}$ has the structure of a locally solid Riesz space.

Let K be the order ideal generated by L in $\hat{L}$. We show first that $K\subset X$. Take any $0\leqq\hat{u}\in K$, then there exists $0\leqq v\in L$ such that $0\leqq\hat{u}\leqq v$, and there exists a net $\{f_\alpha\}$ in L such that $f_\alpha\to\hat{u}$. Define $u_\alpha=f_\alpha^+\wedge v\in L$, then $u_\alpha\to\hat{u}$ and $0\leqq u_\alpha\leqq v$ for all α. Since the topology in L is locally solid, it follows that $\{u_\alpha\}$ is a topologically bounded Cauchy net in L, and hence in X. Since X is quasi-complete, this implies that $\hat{u}\in X$. Hence $K\subset X$ i.e., $K\subset\bar{L}$.

Note that $\hat{L}$ is the completion of K and K is an ideal in $\hat{L}$. Now it follows from [1], Theorem 7.3, that for any $0\leqq\hat{w}\in\hat{L}$ there exists a net $\{w_\alpha\}$ in K with $0\leqq w_\alpha\uparrow\hat{w}$ and $w_\alpha\to\hat{w}$. Since $\{w_\alpha\}$ is a monotone Cauchy net in the locally solid Riesz space K, it follows that $\{w_\alpha\}$ is topologically bounded in K. Hence $\{w_\alpha\}$ is a topologically bounded Cauchy net in X, so $\hat{w}\in X$, as X is quasi-complete. Therefore $\hat{w}\in\bar{L}$ and we may conclude that $\bar{L}=\hat{L}$.

Let L be a Riesz space and $0<e\in L$. By $\mathscr{B}_e$ we denote the collection of all components of e in L. As is well-known, $\mathscr{B}_e$ is a Boolean algebra with respect to the ordering induced by L ([11], Sect. 30). Now assume that L is a Hausdorff locally solid Riesz space with completion $\hat{L}$. By $\hat{\mathscr{B}}_e$ we denote the Boolean algebra of components of e in $\hat{L}$. Obviously $\mathscr{B}_e\subset\hat{\mathscr{B}}_e$.

Proposition 2.4. *If L is a Hausdorff locally solid Riesz space with the principal projection property, then $\hat{\mathscr{B}}_e=\bar{\mathscr{B}}_e$ for any $0<e\in L$, where $\bar{\mathscr{B}}_e$ denotes the closure of $\mathscr{B}_e$ in $\hat{L}$.*

Proof. (compare [7], p. 2,200). First observe that any $\hat{u}\in\bar{\mathscr{B}}_e$ is a component of e in $\hat{L}$. Indeed, if $\hat{u}\in\bar{\mathscr{B}}_e$, then there exists a net $\{u_\alpha\}$ in $\mathscr{B}_e$ such that $u_\alpha\to\hat{u}$. Now $u_\alpha\wedge(e-u_\alpha)=0$ for all α implies that $\hat{u}\wedge(e-\hat{u})=0$, so $\hat{u}\in\hat{\mathscr{B}}_e$. Hence $\bar{\mathscr{B}}_e\subset\hat{\mathscr{B}}_e$.

Now take any $\hat{u}\in\hat{\mathscr{B}}_e$ and let U be a solid neighbourhood of 0 in $\hat{L}$. Let V be a solid neighbourhood of 0 such that $V+V\subseteq U$. Since L has the principal projection property, it follows from the Freudenthal spectral theorem that there exists an element $s=\sum_{i=1}^{n}\alpha_i u_i$, $u_i\in\mathscr{B}_e$, $u_i\wedge u_j=0$ $(i\neq j)$, $\sum_{i=1}^{n}u_i=e$, $0\leqq\alpha_i\leqq 1$, such that $s-\hat{u}\in V$. Now define $v=\sum_{\alpha_i>\frac{1}{2}}u_i$, then $v\in\mathscr{B}_e$, and it is not difficult to see that $|v-\hat{u}|\leqq 2|s-\hat{u}|$. Therefore $v-\hat{u}\in U$, which shows that $\hat{u}\in\bar{\mathscr{B}}_e$, and thus $\hat{\mathscr{B}}_e=\bar{\mathscr{B}}_e$.

In the next proposition we use the same notation as in the above.

Proposition 2.5. *If L is a Hausdorff locally solid Riesz space with Lebesgue topology and with the projection property, then $\hat{\mathscr{B}}_e=\mathscr{B}_e$ for all $0<e\in L$.*

Proof. Since L has Lebesgue topology, it follows from [1], Theorem 13.4 that L is order dense in $\hat{L}$. By [21], Theorem 79.2, any band $\hat{\mathscr{B}}$ in $\hat{L}$ is equal to the band generated by $\hat{\mathscr{B}}\cap L$. Using that L has the projection property, the result follows easily.

The last result in this section is concerned with the automatic order boundedness of band preserving operators. It was shown by McPolin and Wickstead (preprint) that if L is a uniformly complete Riesz space for which the order dual $L^\sim$ separates the points of L, then any band preserving operator in L is order bounded. We will use this result in Sect. 5 of this paper in the case that the space $L_n^\sim$ of normal integrals on L separates the points of L. For the sake of completeness we indicate a proof of the result for this special case.

Lemma 2.6. *Let L be a uniformly complete Riesz space with $L_n^\sim$ separating the points of L. Then any band preserving operator in L is automatically order bounded.*

Proof. Take any $0<\varphi\in L_n^\sim$ and let C_φ be the carrier of φ (see e.g. [21] Sect. 90). Since T is band preserving we can consider the restriction $T\colon C_\varphi \to C_\varphi$. Furthermore, since φ is strictly positive on C_φ, it is clear that for any disjoint sequence $0<w_n\in C_\varphi$ there exist positive real numbers $\{\lambda_n\}_{n=1}^\infty$ such that $\{\lambda_n w_n\}_{n=1}^\infty$ is not order bounded in L. Now it follows from [12], Theorem 8 that T is order bounded on some order dense ideal in C_φ. Since this holds for any $0<\varphi\in L_n^\sim$, and since $L_n^\sim$ separates the points of L, it follows that T is order bounded on some order dense ideal in L. By [12], Proposition 6, the largest ideal on which T is order bounded is a band, hence T is order bounded on L.

3. Boolean Algebras of Projections

Let X be a (complex) locally convex topological vector space and let $\mathscr{L}(X)$ be the vector space of all continuous linear operators on X. In $\mathscr{L}(X)$ we consider the strong operator topology (i.e., the topology of pointwise convergence). Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in X. Thus $\mathscr{M}$ is an equicontinuous, mutually commuting family of idempotents in $\mathscr{L}(X)$, partially ordered by range inclusion and is a Boolean algebra with respect to the lattice operations defined via the algebraic relations in $\mathscr{L}(X)$ by setting $E\vee F=E+F-EF$ and $E\wedge F=EF$ for $E,F\in\mathscr{M}$. We assume throughout that the identity operator is an element of $\mathscr{M}$. We denote by M the linear hull of $\mathscr{M}$ in $\mathscr{L}(X)$. Note that M is equal to the algebra generated by $\mathscr{M}$ in $\mathscr{L}(X)$. Each element T of M can be written as a finite sum of the form $T=\sum_{i=1}^n \lambda_i E_i$, where $E_iE_j=0$ if $i\neq j$, $\sum_{i=1}^n E_i=I$, all $E_i\neq 0$ and $\lambda_1,\dots,\lambda_n$ are mutually different scalars. This way of writing is essentially unique and will be called the standard representation of T. We denote by $\operatorname{Re} M$ the real linear hull of $\mathscr{M}$. Then $M=\operatorname{Re} M+i\operatorname{Re} M$, so M is the complexification of $\operatorname{Re} M$. As a linear space, $\operatorname{Re} M$ may be identified with the Riesz space of all real finitely-valued continuous functions on the Stone space of $\mathscr{M}$. This identification induces on $\operatorname{Re} M$ the structure of a Riesz space with the identity operator I as a strong order unit. Moreover, M is a complex Riesz space, i.e. for each $T\in M$ the absolute value $|T|$ exists. In fact, if $T=\sum_{i=1}^n \lambda_i E_i$ in standard representation, then $|T|=\sum_{i=1}^n |\lambda_i| E_i$.

Observe already that

$$\{R \in M \colon |R| \leqq I\} = \left\{ \sum_{i=1}^{n} \lambda_i E_i \colon E_i E_j = 0 \text{ if } i \neq j, \sum_{i=1}^{n} E_i = I \text{ and } 0 \leqq |\lambda_i| \leqq 1 \right\}.$$

Furthermore, for any $T \in M$ we have

$$\{S \in M \colon |S| \leqq |T|\} = \{RT \colon R \in M, |R| \leqq I\}.$$

In particular, for $0 \leqq T \in \operatorname{Re} M$ we have $[0, T] = [0, I] \cdot T$ and $[-T, T] = [-I, I] \cdot T$.

Next we will show that the topology induced by $\mathscr{L}(X)$ in M is a locally solid topology. To this end, let $\mathscr{P}$ be the family of all continuous semi-norms in X; then the topology in X is equal to the topology generated by $\mathscr{P}$. For each $x \in X$ and $p \in \mathscr{P}$ define $p_x(T) = p(Tx)$ for all $T \in \mathscr{L}(X)$. The strong operator topology in $\mathscr{L}(X)$ is the topology generated by the semi-norms $\{p_x \colon p \in \mathscr{P}, x \in X\}$.

For each $p \in \mathscr{P}$ define the semi-norm ρ_p in X by

$$\rho_p(x) = \sup\{p(Tx) \colon T \in M, |T| \leqq I\}$$

for all $x \in X$.

Lemma 3.1. *For any $p \in \mathscr{P}$ there exists $q \in \mathscr{P}$ such that $p(x) \leqq \rho_p(x) \leqq q(x)$ for all $x \in X$.*

The preceding estimate is familiar from the theory of vector valued measures. See, for example, Lemma 1 of [14] or Proposition 2.3 of [20]. However, as in Lemma 8 of [19], a direct proof in the present setting may be based on summation by parts and the fact that equicontinuity of $\mathscr{M}$ implies for any $p \in \mathscr{P}$ the existence of a $q \in \mathscr{P}$ such that $p(Ex) \leqq q(x)$ for every $E \in \mathscr{M}$ and all $x \in X$. We omit the details.

It follows from the lemma that the topology in X generated by the system $\{\rho_p \colon p \in \mathscr{P}\}$ is equal to the given topology. Moreover, the strong operator topology in $\mathscr{L}(X)$ is generated by the semi-norms $\{\rho_{p,x} \colon p \in \mathscr{P}, x \in X\}$, where $\rho_{p,x}(T) = \rho_p(Tx)$.

Lemma 3.2. *The semi-norm $\rho_{p,x}$ is a Riesz semi-norm on M for every $p \in \mathscr{P}$ and all $x \in X$. Therefore, with respect to the induced topology, M is a complex locally solid, locally convex Riesz space.*

Proof. Suppose $|S| \leqq |T|$ in M. As observed above, then $S = RT$ for some $R \in M$, $|R| \leqq I$. Now it is immediate from the definition of ρ_p that $\rho_p(Sx) \leqq \rho_p(Tx)$, i.e., $\rho_{p,x}(S) \leqq \rho_{p,x}(T)$.

Fix $x \in X$ for a moment. By $M(x)$ we denote the linear subspace of X generated by $\{Ex \colon E \in \mathscr{M}\}$. Note that $M(x) = \{Tx \colon T \in M\}$. By $\operatorname{Re} M(x)$ we denote the real linear subspace generated by $\{Ex \colon E \in M\}$, i.e. $\operatorname{Re} M(x) = \{Tx \colon T \in \operatorname{Re} M\}$. Clearly, $M(x)$ is the complexification of $\operatorname{Re} M(x)$.

The mapping φ_x from M into $M(x)$, defined by $\varphi_x(T) = Tx$, is a linear surjective mapping, and it is not difficult to see that the kernel of φ_x is a uniformly closed order ideal in the complex Riesz space M. Therefore, the complex quotient Riesz space $M/\ker\varphi_x$ is not only linearly isomorphic to $M(x)$, but induces a uniquely determined complex Riesz space structure on $M(x)$ such that $M(x)$ is the complexification of $\operatorname{Re}M(x)$, which is isomorphic to the quotient Riesz space $\operatorname{Re}M/\operatorname{Re}\ker\varphi_x$ (see Sect. 2).

Observe that φ_x is a Riesz homomorphism from M onto $M(x)$. Moreover, if $|y|\leqq|z|$ in $M(x)$, then there exist $S, T\in M$ such that $|S|\leqq|T|$ and $\varphi_x(S)=y$, $\varphi_x(T)=z$. This implies that for any $z\in M(x)$ we have

$$\{y\in M(x)\colon |y|\leqq|z|\}=\{Tz\colon T\in M, |T|\leqq I\}.$$

The next lemma follows now easily.

Lemma 3.3. *Fix $x\in X$. For every $p\in\mathscr{P}$ the semi-norm ρ_p is a Riesz semi-norm on the complex Riesz space $M(x)$. The locally convex locally solid topology in $M(x)$ generated by the collection $\{\rho_p\colon p\in\mathscr{P}\}$ of seminorms is equal to the given topology in $M(x)$.*

We introduce some notation. By $\bar{M}$ and $\overline{\operatorname{Re}M}$ we denote the respective closures of M and $\operatorname{Re}M$ in $\mathscr{L}(X)$. If $x\in X$, then we denote by $\mathscr{M}(x)$ and $\operatorname{Re}\mathscr{M}(x)$ the respective closures of $M(x)$ and $\operatorname{Re}M(x)$ in X.

Let $\hat{X}$ be the (topological) completion of X. For $x\in X$ we denote by $\hat{\mathscr{M}}(x)$ the completion of $M(x)$, i.e., $\hat{\mathscr{M}}(x)$ is the closure of $M(x)$ in $\hat{X}$. Similarly, $\operatorname{Re}\hat{\mathscr{M}}(x)$ denotes the completion of $\operatorname{Re}M(x)$. Furthermore, $\hat{M}$ denotes the completion of M and $\operatorname{Re}\hat{M}$ denotes the completion of $\operatorname{Re}M$.

Since $\operatorname{Re}M$ and $\operatorname{Re}M(x)$ are locally solid Riesz spaces, the lattice operations extend to the completions $\operatorname{Re}\hat{M}$ and $\operatorname{Re}\hat{\mathscr{M}}(x)$, so $\operatorname{Re}\hat{M}$ and $\operatorname{Re}\hat{M}(x)$ are locally solid Riesz spaces. Moreover, $\hat{M}$ and $\hat{\mathscr{M}}(x)$ are complex locally solid Riesz spaces, which are the complexifications of $\operatorname{Re}\hat{M}$ and $\operatorname{Re}\hat{\mathscr{M}}(x)$ respectively (see Proposition 2.2).

Furthermore, since the identity I is a strong order unit in the complex Riesz space M, it follows that I is a weak order unit in $\hat{M}$. Similarly, x is a weak order unit in $\hat{\mathscr{M}}(x)$. Observe that $\mathscr{M}$ is equal to the Boolean algebra of components of I in M, so $\mathscr{M}$ is isomorphic to the Boolean algebra of (principal) projection bands in M.

Next we consider the question as to when the Riesz space structures introduced above are Dedekind complete with Lebesgue topologies. To this end, we introduce the following properties of Boolean algebras of projections (see also [5]).

Definition 3.4. Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in the locally convex space X.

(i) $\mathscr{M}$ is called *strongly equicontinuous with respect to the element* $x\in X$ if $E_n x\to 0$ whenever $\{E_n\}_{n=1}^{\infty}$ is a disjoint sequence in $\mathscr{M}$.

(ii) $\mathscr{M}$ is called *strongly equicontinuous* if $\mathscr{M}$ is strongly equicontinuous with respect to every $x\in X$.

(iii) $\mathscr{M}$ is called *complete in the sense of Bade* if $\mathscr{M}$ is complete as a Boolean algebra and $\langle E_\alpha x, x^*\rangle\to\langle Ex, x^*\rangle$ for all $x\in X$ and all $x^*\in X^*$ whenever $E_\alpha\uparrow E$ in $\mathscr{M}$.

Proposition 3.5. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in X and let $x \in X$. The complex Riesz space $\hat{\mathcal{M}}(x)$ is Dedekind complete and has Lebesgue topology if and only if $\mathcal{M}$ is strongly equicontinuous with respect to x.*

Proof. First suppose that $\hat{\mathcal{M}}(x)$ has Lebesgue topology, i.e., that $\operatorname{Re}\hat{\mathcal{M}}(x)$ has Lebesgue topology, and let $\{E_n\}_{n=1}^{\infty}$ be a disjoint sequence in $\mathcal{M}$. Then $\{E_n x\}_{n=1}^{\infty}$ is a disjoint sequence in $\operatorname{Re} M(x)$ and hence in $\operatorname{Re}\hat{\mathcal{M}}(x)$. Moreover, since $0 \leqq E_n x \leqq x$ $(n=1,2,\ldots)$, the sequence is order bounded. Hence $E_n x \to 0$ in $\operatorname{Re}\hat{\mathcal{M}}(x)$, and so $\mathcal{M}$ is strongly equicontinuous with respect to x.

Now assume that $\mathcal{M}$ is strongly equicontinuous with respect to x. Since $\operatorname{Re}\hat{\mathcal{M}}(x)$ is complete, it is sufficient to show that $\operatorname{Re}\hat{\mathcal{M}}(x)$ has Lebesgue topology ([1], Theorem 10.3). By [1], Theorem 10.5, it suffices to show that the topology in $\operatorname{Re} M(x)$ is pre-Lebesgue. Hence, by [1], Theorem 10.1, it is sufficient to show that every order bounded disjoint sequence in $\operatorname{Re} M(x)$ converges to zero.

To this end, let $\{y_n\}_{n=1}^{\infty}$ be a positive order bounded disjoint sequence in $\operatorname{Re} M(x)$. Since x is a strong order unit in $\operatorname{Re} M(x)$, we may assume that $0 \leqq y_n \leqq x$ $(n=1,2,\ldots)$. It is easy to see that there exists a sequence $\{E_n\}_{n=1}^{\infty}$ in $\mathcal{M}$ such that $\{E_n x\}_{n=1}^{\infty}$ is a disjoint sequence and $0 \leqq y_n \leqq E_n x$ $(n=1,2,\ldots)$. If we can show that $E_n x \to 0$, then we are done.

Define $E_1' = E_1$ and $E_n' = E_n - (E_1 \vee \ldots \vee E_{n-1}) \wedge E_n$ for $n \geqq 2$; then $\{E_n'\}$ is a disjoint sequence in $\mathcal{M}$ and $E_n' x = E_n x$ for all n. Since $\mathcal{M}$ is strongly equicontinuous with respect to x, this implies $E_n x \to 0$.

We remark that the "if" part of the preceding proposition is contained in Proposition 3.13 of [20], for the special case that $\mathcal{M}$ is a σ-complete Boolean algebra in the sense of Bade. The preceding proposition also extends Theorem 2 of Veksler [19] and the final conclusion of Theorem 2.10 of Rall [13].

The next result is proved using analogous arguments.

Proposition 3.6. *The complex Riesz space $\hat{M}$ is Dedekind complete and has Lebesgue topology if and only if $\mathcal{M}$ is strongly equicontinuous.*

The next two results are concerned with properties of the space M.

Proposition 3.7. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the locally convex space X. The following statements are equivalent.*

(i) *$E_\alpha \uparrow E$ in $\mathcal{M}$ implies that $\langle E_\alpha x, x^* \rangle \to \langle Ex, x^* \rangle$ for all $x \in X$ and all $x^* \in X^*$.*

(ii) *$E_\alpha \uparrow E$ in $\mathcal{M}$ implies that $E_\alpha \to E$ with respect to the strong operator topology.*

(iii) *M has Lebesgue topology.*

Proof. (i)$\Rightarrow$(ii). Suppose that $E_\alpha \downarrow 0$ in $\mathcal{M}$ and take $x \in X$. It follows from the hypothesis that $E_\alpha x \to 0$ weakly in the locally solid locally convex Riesz space $M(x)$. Using that the weak closure of a convex set in $M(x)$ is equal to the closure with respect to the given topology, and that $E_\alpha x \downarrow \geqq 0$ in $M(x)$ if follows that $E_\alpha x \to 0$ in $M(x)$.

(ii)$\Rightarrow$(iii). Let $\{T_\alpha\}$ be a net in $\operatorname{Re} M$ such that $T_\alpha \downarrow 0$. We have to show that $T_\alpha \to 0$ with respect to the topology in M. Without loss of generality, we may assume that $0 \leqq T_\alpha \leqq I$ for all α. Let U be a solid neighbourhood of 0 in M and let V be a neighbourhood of 0 with $V+V \subseteq U$. Take $\varepsilon > 0$ such that $\varepsilon I \in V$.

Let E_α be the component of I in the band $\{(T_\alpha - \varepsilon I)^+\}^{dd}$ in $\operatorname{Re} M$. Then $E_\alpha \downarrow 0$ in $\mathscr{M}$, as $\{(T_\alpha - \varepsilon I)^+\}^{dd} \downarrow \{0\}$, and so, by hypothesis, $E_\alpha \to 0$. Let α_0 be such that $E_\alpha \in V$ for all $\alpha \geqq \alpha_0$. Now it follows from $0 \leqq T_\alpha \leqq E_\alpha + \varepsilon I$ that $T_\alpha \in U$ for all $\alpha \geqq \alpha_0$, and hence $T_\alpha \to 0$ in M.

(iii)$\Rightarrow$(i). Trivial.

Observe that it is an immediate consequence of the above proposition that any Boolean algebra which is complete in the sense of Bade, is strongly equicontinuous.

Corollary 3.8. *Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in the locally convex space X. Then $\mathscr{M}$ is complete in the sense of Bade if and only if M has the projection property and has Lebesgue topology.*

Proof. As observed already, $\mathscr{M}$ is isomorphic to the Boolean algebra of projection bands in M. Therefore, since M has sufficiently many projections, it follows from [11], Theorem 30.6, that $\mathscr{M}$ is complete if and only if M has the projection property. Now the result follows from the above proposition.

In the following proposition, we apply the result of Proposition 2.3.

Proposition 3.9. *Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. If $x \in X$, then the cyclic subspace $\mathscr{M}(x)$ is topologically complete. In particular, $\mathscr{M}(x)$ has the structure of a complex locally solid topologically complete Riesz space with weak order unit x, which is the complexification of the Riesz space* $\operatorname{Re}\mathscr{M}(x)$.

Moreover, $\mathscr{M}(x)$ is Dedekind complete and has Lebesgue topology if and only if $\mathscr{M}$ is strongly equicontinuous with respect to x.

The following result is similar.

Proposition 3.10. *Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in X and assume that $\mathscr{L}(X)$ is quasi-complete. Then $\bar{M}$ is complete for the strong operator topology in $\mathscr{L}(X)$. In particular, $\bar{M}$ has the structure of a complex, locally solid, topologically complete Riesz space with weak order unit I, which is the complexification of the Riesz space* $\operatorname{Re}\bar{M}$.

Moreover, $\bar{M}$ is Dedekind complete and has Lebesgue topology if and only if $\mathscr{M}$ is strongly equicontinuous.

We remark that the result of Proposition 3.9 was proved by Walsh ([20], Proposition 3.15) under the assumption that $\mathscr{M}$ is σ-complete in the sense of Bade. The direct proof, using Proposition 2.3 shows that equicontinuity of $\mathscr{M}$ suffices, provided that X is quasi-complete.

Recall that if X is quasi-complete, then the closed equicontinuous subsets of $\mathscr{L}(X)$ are complete (see e.g. [18], III 4.4). Furthermore it is easy to see that order bounded sets in M are equicontinuous. Therefore, if X is quasi-complete, then the closure in $\mathscr{L}(X)$ of an order bounded subset of M is complete. This simple observation yields the following result.

Proposition 3.11. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. Let M_I denote the order ideal generated by I in $\hat{M}$. Then M_I is contained in the closure $\bar{M}$ of M in $\mathcal{L}(X)$.*

Furthermore, the following statements are equivalent.

(i) *The Boolean algebra $\mathcal{M}$ is strongly equicontinuous.*

(ii) *The complex Riesz space M_I is Dedekind complete with Lebesgue topology.*

(iii) *The complex Riesz space $\hat{M}$ is Dedekind complete with Lebesgue topology.*

We will show next that if X is quasi-complete, then the elements of $\hat{M}$ can be identified with (not necessarily continuous) linear operators from X into itself. Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. We denote by $\mathcal{L}^{\#}(X)$ the vector space of all linear mappings from X into itself. In $\mathcal{L}^{\#}(X)$ we consider the strong operator topology. First remark that $\mathcal{L}^{\#}(X)$ is a quasi-complete space. Hence, the complex Riesz space M is a linear subspace of the quasi-complete space $\mathcal{L}^{\#}(X)$ and the topology in M is locally solid. Now it follows from Proposition 2.3 that the closure of M in $\mathcal{L}^{\#}(X)$ is topologically complete, i.e., the completion $\hat{M}$ of M is equal to the closure of M in $\mathcal{L}^{\#}(X)$. We thus have the following result.

Proposition 3.12. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. The completion $\hat{M}$ of M can be identified (algebraically and topologically) with the closure of M in $\mathcal{L}^{\#}(X)$.*

Remark 3.13. Same situation as in the above proposition. For $x \in X$, consider the Riesz homomorphism φ_x from M onto $M(x)$, defined by $\varphi_x(T) = Tx$. Clearly, φ_x is continuous. Since, by Proposition 3.9, $\hat{\mathcal{M}}(x) = \mathcal{M}(x) \subseteq X$, φ_x has a unique extension to a continuous Riesz homomorphism from $\hat{M}$ into $\mathcal{M}(x)$, which will be denoted by φ_x again. Now observe that $\hat{T}x \in \mathcal{M}(x)$ for each $\hat{T} \in \hat{M}$. Indeed, if $\hat{T} \in \hat{M}$, then there exists a net $\{T_\alpha\}$ in M such that $T_\alpha \to \hat{T}$ with respect to the strong operator topology. Hence $T_\alpha x \to \hat{T}x$. Since $T_\alpha x \in \mathcal{M}(x)$ for all α, this implies that $\hat{T}x \in \mathcal{M}(x)$. Consequently, the mapping $\hat{T} \mapsto \hat{T}x$ is a continuous linear mapping from $\hat{M}$ into $\mathcal{M}(x)$ which extends φ_x from M to $\hat{M}$. By the uniqueness, we may conclude that $\varphi_x(\hat{T}) = \hat{T}x$ for all $\hat{T} \in \hat{M}$.

4. The Algebraic Structure of $\hat{M}$

Let X be a quasi-complete locally convex space and let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in X. As observed earlier, M is a commutative algebra with respect to composition as multiplication and I is the unit element in M. Furthermore, $\operatorname{Re} M$ is a real subalgebra of M with the property that $ST \geqq 0$ whenever $S, T \geqq 0$ in $\operatorname{Re} M$, and I is a ring unit as well as a strong order unit in $\operatorname{Re} M$. This implies that $\operatorname{Re} M$ is an f-algebra, and hence M is a complex f-algebra. In particular, $|ST| = |S||T|$ for all $S, T \in M$.

Fix $x \in X$. It follows from the remarks above that any $T \in M$ leaves $M(x)$ invariant, and so by continuity, T leaves $\mathcal{M}(x)$ invariant. Let $T[x]$ denote the restriction of T to $\mathcal{M}(x)$. We assert that $T[x]\in Z(\mathcal{M}(x))$. Indeed, if $T\in M$, then there exists $n\in\mathbb{N}$ such that $|T|\leq nI$ in M. Take $y\in M(x)$, then $y=Sx$ for some $S\in M$, so $|Ty|=|TSx|=|TS|x\leq n|S|x=n|y|$. By continuity, we get $|Ty|\leq n|y|$ for all $y\in\mathcal{M}(x)$ hence, $T[x]\in Z(\mathcal{M}(x))$. In like manner, it is shown that $|T[x]|=|T|[x]$ in $Z(\mathcal{M}(x))$ for all $T\in M$. Therefore, the mapping Φ_x from M into $Z(\mathcal{M}(x))$, defined by $\Phi_x(T)=T[x]$ is a Riesz homomorphism.

Now let $\hat{T}\in\hat{M}$. There exists a net $\{T_\alpha\}$ in M such that $T_\alpha\to\hat{T}$ in $\mathcal{L}^{\#}(X)$, so in particular $T_\alpha y\to\hat{T}y$ for all $y\in\mathcal{M}(x)$. Since $T_\alpha y\in\mathcal{M}(x)$ for all α, this shows that $\hat{T}y\in\mathcal{M}(x)$ for all $y\in\mathcal{M}(x)$, i.e. $\hat{T}$ leaves $\mathcal{M}(x)$ invariant. Let $\hat{T}[x]$ denote the restriction of $\hat{T}$ to $\mathcal{M}(x)$. We claim that $\hat{T}[x]$ is an orthomorphism in $\mathcal{M}(x)$. Indeed, first observe that $\hat{T}[x]$ is order bounded, being a linear combination of positive operators. Now assume that $y\perp z$ in $\mathcal{M}(x)$ and let $T_\alpha\in M$ be such that $T_\alpha\to\hat{T}$. Since $T_\alpha[x]\in Z(\mathcal{M}(x))$ for all α, we have $T_\alpha y\perp z$ for all α, and so $T_\alpha y\to\hat{T}y$ implies that $\hat{T}y\perp z$.

Let the mapping Φ_x from $\hat{M}$ into $\mathrm{Orth}(\mathcal{M}(x))$ be defined by $\Phi_x(T)=T[x]$. It is easy to see that Φ_x is a Riesz homomorphism.

We collect the above results in the next proposition.

Proposition 4.1. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. Fix $x\in X$. Every $\hat{T}\in\hat{M}$ leaves $\mathcal{M}(x)$ invariant and, if $\hat{T}[x]$ denotes the restriction of $\hat{T}$ to $\mathcal{M}(x)$, then $\hat{T}[x]\in\mathrm{Orth}(\mathcal{M}(x))$.*

Moreover, the mapping Φ_x from $\hat{M}$ into $\mathrm{Orth}(\mathcal{M}(x))$, *defined by $\Phi_x(\hat{T})=\hat{T}[x]$ is a Riesz homomorphism.*

For any subset $H\subseteq\mathcal{L}^{\#}(X)$ with the property that every element of H leaves $\mathcal{M}(x)$ invariant, we denote by $H[x]$ the collection of all restrictions of elements of H to $\mathcal{M}(x)$. Note that $H[x]\subseteq\mathcal{L}^{\#}(\mathcal{M}(x))$. By the above proposition, $\hat{M}[x]\subseteq\mathrm{Orth}(\mathcal{M}(x))$ for all $x\in X$. Furthermore, since the mapping Φ_x is a Riesz homomorphism, it follows that $M_I[x]\subseteq Z(\mathcal{M}(x))$.

Observe that it follows from the above proposition that any two elements in $\hat{M}$ commute. Indeed, take $\hat{S},\hat{T}\in\hat{M}$ and $x\in X$, then $(\hat{S}\hat{T})(x)=\hat{S}[x]\,\hat{T}[x](x)=\hat{T}[x]\hat{S}[x](x)=(\hat{T}\hat{S})(x)$, as $\mathrm{Orth}(\mathcal{M}(x))$ is a commutative algebra.

Next we show that $\hat{M}$ is a commutative algebra with respect to composition as multiplication. First take $S\in M$ and $\hat{T}\in\hat{M}$. There exists a net $\{T_\alpha\}$ in M such that $T_\alpha\to\hat{T}$. By continuity of S we get $ST_\alpha x\to S\hat{T}x$ for all $x\in X$, i.e. $ST_\alpha\to S\hat{T}$. Since $ST_\alpha\in M$ for all α, this implies that $S\hat{T}\in\hat{M}$.

Now take $\hat{S}\in\hat{M}$ and $\hat{T}\in\hat{M}$ and let $\{S_\alpha\}$ be a net in M such that $S_\alpha\to\hat{S}$. Then $S_\alpha\hat{T}x\to\hat{S}\hat{T}x$ for all $x\in X$, i.e., $S_\alpha\hat{T}\to\hat{S}\hat{T}$. By the above $S_\alpha\hat{T}\in\hat{M}$ for all α, hence $\hat{S}\hat{T}\in\hat{M}$. Therefore, since the elements of $\hat{M}$ commute, it follows that $\hat{M}$ is a commutative algebra with the identity I as a ring unit.

Clearly, $\mathrm{Re}\,\hat{M}$ is a real subalgebra of $\hat{M}$ with the property that $\hat{S}\hat{T}\geq 0$ whenever $0\leq\hat{S},\hat{T}\in\mathrm{Re}\,\hat{M}$. Since I is a ring unit as well as a weak order unit in $\mathrm{Re}\,\hat{M}$, it follows that $\mathrm{Re}\,\hat{M}$ is an f-algebra. Consequently, $\hat{M}$ is a complex f-algebra.

We summarize the above results in the following proposition.

Proposition 4.2. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. Then the following holds.*

(i) $\hat{M}$ *is a complex f-algebra with* I *as unit element, with respect to composition as multiplication. In particular,* $\hat{M}$ *is commutative.*

(ii) *The Riesz homomorphism* Φ_x *from* $\hat{M}$ *into* $\operatorname{Orth}(\mathscr{M}(x))$*, defined by* $\Phi_x(\hat{T}) = \hat{T}[x]$*, is an algebra homomorphism as well.*

Since $\hat{M}$ is a complex f-algebra, the following inequality holds for any $0 \leqq \hat{T} \in \operatorname{Re} \hat{M}$,

$$0 \leqq \hat{T} - \hat{T} \wedge nI \leqq \frac{1}{n} \hat{T}^2 (n=1,2,\ldots)$$

(see [21], Theorem 142.7, or [10], Proposition 3.2). Since the topology in $\hat{M}$ is locally solid this implies that $\hat{T} \wedge nI \to \hat{T}$ with respect to the strong operator topology. This simple observation, combined with Proposition 3.11, yields the following interesting result.

Proposition 4.3. *Let* $\mathscr{M}$ *be an equicontinuous Boolean algebra of projections in the quasi-complete space* X*. Assume, in addition, that* $\mathscr{L}(X)$ *is sequentially complete. Then* $\hat{M} = \bar{M}$*, and so* $\bar{M}$ *is a complex, locally solid, topologically complete Riesz space with* I *as a weak order unit.*

As observed in Sect. 3 (the remarks preceding Definition 3.4), $\mathscr{M}$ is the Boolean algebra of components of I in M. Denote by $\bar{\mathscr{M}}$ the closure of $\mathscr{M}$ in $\mathscr{L}(X)$ with respect to the strong operator topology. If X is assumed to be quasi-complete, then $\bar{\mathscr{M}}$ is complete, and so $\bar{\mathscr{M}}$ is equal to the closure of $\mathscr{M}$ in $\hat{M}$. Hence we are in the situation of Proposition 2.4.

Proposition 4.4. *Let* $\mathscr{M}$ *be an equicontinuous Boolean algebra of projections in the quasi-complete space* X*. Given* $E \in \hat{M}$ *the following statements are equivalent.*

(i) $E \in \bar{M}$ *and* E *is a projection*

(ii) $E \in \bar{\mathscr{M}}$

(iii) E *is a component of* I *in the complex Riesz space* $\hat{M}$.

Proof. The equivalence of (ii) and (iii) follows from Proposition 2.4.

(i)⇒(iii). If $E \in \bar{M}$ and $E = E^2$, then E is an idempotent in the complex f-algebra $\hat{M}$, and so by Lemma 2.1 E is a component of I.

(iii)⇒(i). If E is a component of I in $\hat{M}$, then $E = E^2$ and $0 \leqq E \leqq I$. Hence E is a projection and $E \in M_I \subseteq \bar{M}$.

It is appropriate at this point to make some remarks which indicate the relation of our results to the theory of scalar type spectral operators. The connection is established via the spectral theorem of Freudenthal (see [11], Theorem 40.3). Let $\mathscr{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X. If $\mathscr{L}(X)$ is assumed sequentially complete, then $\bar{M}$ is a complex, locally solid Riesz space which is Dedekind complete, has Lebesgue topology and has I as a weak order unit, as follows by combining Propositions 4.3 and 3.6. Moreover, from Proposition 4.4, the principal components of the weak unit I are now identified with projections in $\bar{M}$. A moment's reflection now shows that the Freudenthal spectral theorem asserts that $\bar{M}$ consists entirely of scalar type spectral operators. This conclusion, which goes back to Bade [2] in the Banach space setting, is due to Ricker [14] for the case that $\mathscr{M}$ is the range of a (countably-additive) spectral measure.

Observe that it follows from a combination of Proposition 4.2, Proposition 4.4 and the fact that the idempotents in $\mathrm{Orth}(\mathscr{M}(x))$ are band projections in $\mathscr{M}(x)$, that $E[x]$ is a band projection in $\mathscr{M}(x)$ for any $E \in \bar{\mathscr{M}}$.

Now Proposition 2.5 will be used to prove our next result, which extends part of [2], Theorem 4.5.

Proposition 4.5. *Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. Then $\mathscr{M}$ is complete in the sense of Bade if and only if $\mathscr{M}$ is strongly equicontinuous and closed with respect to the strong operator topology.*

Proof. First assume that $\mathscr{M}$ is complete in the sense of Bade. By Corollary 3.8, M has the projection property and Lebesgue topology. Now it follows from Proposition 2.5 that the Boolean algebra of components of I in M is equal to the Boolean algebra of components of I in $\hat{M}$. As observed earlier, the Boolean algebra of components of I in M is $\mathscr{M}$, and it follows from Proposition 4.4 that the Boolean algebra of components of I in $\hat{M}$ is $\bar{\mathscr{M}}$. Hence $\mathscr{M} = \bar{\mathscr{M}}$, i.e., $\mathscr{M}$ is closed. As noted, it follows from Proposition 3.7 that any Boolean algebra which is complete in the sense of Bade, is strongly equicontinuous.

Now assume that $\mathscr{M}$ is strongly equicontinuous and closed. By Proposition 3.6, $\hat{M}$ is Dedekind complete and has Lebesgue topology. Therefore, $\bar{\mathscr{M}}$ is complete as a Boolean algebra, as $\bar{\mathscr{M}}$ is the Boolean algebra of components of I in $\hat{M}$. Moreover, since $\hat{M}$ has Lebesgue topology, it follows that $\bar{\mathscr{M}}$ is complete in the sense of Bade. Consequently, since $\mathscr{M} = \bar{\mathscr{M}}$, it follows that $\mathscr{M}$ is complete in the sense of Bade.

From the last part of the proof of the above proposition we immediately get the following result (cf. [5], Proposition 3.6).

Proposition 4.6. *Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. Then $\mathscr{M}$ is strongly equicontinuous if and only if $\bar{\mathscr{M}}$ is complete in the sense of Bade.*

We mention another corollary of Proposition 4.4, which extends Corollary 7, p. 2201 [7].

Proposition 4.7. *Let $\mathscr{M}$ be a Boolean algebra of projections, which is complete in the sense of Bade, in the quasi-complete space X. Then $\mathscr{M}$ contains every projection in the weakly closed algebra it generates.*

Proof. Note that the weak and strong operator closure of M coincide. Thus $\bar{M}$ is the weakly closed algebra generated by $\mathscr{M}$ in $\mathscr{L}(X)$. The result now follows from the equivalence of (i) and (ii) in Proposition 4.4 and from the fact that $\mathscr{M} = \bar{\mathscr{M}}$ (by Proposition 4.5).

If, in addition, $\mathscr{L}(X)$ is assumed to be sequentially complete, then the result of the preceding Proposition 4.7 may also be found in Corollary 4.2.2 of [5]. However, as is now clear, quasi-completeness of X alone suffices.

Let $\mathscr{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X. As before, by M_I we denote the order ideal generated by I

in $\hat{M}$. It has already been noted that $M_I \subseteq \overline{M}$. If $x \in X$, then we denote by $\mathcal{M}_x$ the order ideal generated by x in the complex Riesz space $\mathcal{M}(x)$. If $\mathcal{M}(x) = X$, then x will be called cyclic for $\mathcal{M}$.

Recall that the mapping φ_x from $\hat{M}$ into $\mathcal{M}(x)$, defined by $\varphi_x(\hat{T}) = \hat{T}x$ for all $\hat{T} \in \hat{M}$ is a Riesz homomorphism. Furthermore, it is clear that φ_x maps M_I into $\mathcal{M}_x$, so we may consider the mapping φ_x from M_I into $\mathcal{M}_x$.

Proposition 4.8. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X and assume that $x \in X$ is a cyclic vector for $\mathcal{M}$. Then φ_x is a Riesz isomorphism from M_I onto $\mathcal{M}_x$.*

Proof. It is clear that φ_x is injective. Now take $y \in \mathcal{M}_x$ with $0 \leqq y \leqq x$. Then there exists a net $\{A_\alpha\}$ in M with $A_\alpha x \to y$ and $0 \leqq A_\alpha \leqq I$ for all α. We claim that $\{A_\alpha\}$ is a Cauchy net with respect to the strong operator topology. Indeed, take $z \in X$ and let U be a solid neighbourhood of 0 in $\mathcal{M}(x) = X$. Let V be a solid neighbourhood of 0 such that $V + V + V \subseteq U$. There exists $S \in M$ such that $z - Sx \in V$. Let W be a solid neighbourhood of 0 such that $S(W) \subseteq V$. Since $\{A_\alpha x\}$ is a Cauchy net it follows that there exists α_0 such that $(A_\alpha - A_\beta)x \in W$ for all $\alpha, \beta \geqq \alpha_0$. Now it follows from $|(A_\alpha - A_\beta)z| \leqq |S(A_\alpha - A_\beta)x| + 2|z - Sx|$ that $(A_\alpha - A_\beta)z \in U$ for all $\alpha, \beta \geqq \alpha_0$. Hence $\{A_\alpha z\}$ is a Cauchy net in X. This holds for all $z \in X$, so $\{A_\alpha\}$ is a Cauchy net in M. Since $0 \leqq A_\alpha \leqq I$ for all α and since order intervals in M_I are complete, there exists $A \in M_I$, $0 \leqq A \leqq I$ with $A_\alpha \to A$. In particular $A_\alpha x \to Ax$, so $Ax = y$, i.e. $\varphi_x(A) = y$. This shows that φ_x is surjective.

Let Φ_x be the mapping from $\hat{M}$ into $\mathrm{Orth}(\mathcal{M}(x))$ as in Proposition 4.1. If $x \in X$ is cyclic for $\mathcal{M}$, then $\Phi_x(\hat{T}) = \hat{T}$ for all $\hat{T} \in \hat{M}$.

Corollary 4.9. *Let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in the quasi-complete space X, and suppose that $x \in X$ is a cyclic vector for $\mathcal{M}$. Then $\hat{M} = \mathrm{Orth}(\mathcal{M}(x))$ and $M_I = Z(\mathcal{M}(x))$.*

Proof. It follows from Proposition 4.1 that $\hat{M} \subseteq \mathrm{Orth}(\mathcal{M}(x))$. Let $T \in Z(\mathcal{M}(x))$ be given with $0 \leqq T \leqq I$. Then $0 \leqq Tx \leqq x$, so by the above proposition there exists $A \in M_I$, $0 \leqq A \leqq I$ with $Ax = Tx$. Since x is a weak order unit in $\mathcal{M}(x)$, this implies that $T = A$, hence $T \in M_I$. This shows already that $M_I = Z(\mathcal{M}(x))$.

Now take any $0 \leqq T \in \mathrm{Orth}(\mathcal{M}(x))$, then $0 \leqq T - T \wedge nI \leqq \frac{1}{n} T^2$ $(n = 1, 2, \ldots)$. Since the topology in $\mathcal{M}(x)$ is locally solid, this implies that $T \wedge nI \to T$ with respect to the strong operator topology. By the first part of the proof, $T \wedge nI \in M_I$ $(n = 1, 2, \ldots)$, and therefore $T \in \hat{M}$. We may conclude that $\hat{M} = \mathrm{Orth}(\mathcal{M}(x))$.

Our next objective is to show that Φ_x maps M_I onto $Z(\mathcal{M}(x))$ if $\mathcal{M}$ is strongly equicontinuous, without the assumption that x is a cyclic vector for $\mathcal{M}$. We denote the restriction of Φ_x to M_I by Φ_x again.

Let X be a quasi-complete space and let $\mathcal{M}$ be an equicontinuous Boolean algebra of projections in X. As before, for a subset H of $\mathcal{L}^{\#}(X)$ with the property that every element of H leaves $\mathcal{M}(x)$ invariant, we denote by $H[x]$ the collection of restrictions of elements of H to $\mathcal{M}(x)$.

Note that $\mathcal{M}[x]$ is an equicontinuous Boolean algebra of projections in $\mathcal{M}(x)$. Furthermore, it is clear that $\mathcal{M}[x](x) = \mathcal{M}(x)$, so x is a cyclic vector for

$\mathcal{M}[x]$. Hence it follows from Corollary 4.9 that $(M[x])_I = Z(\mathcal{M}(x))$. Furthermore, the components of I in $(M[x])_I$ are precisely the elements of $\overline{\mathcal{M}[x]}$ where the closure is taken with respect to the strong operator topology in $\mathcal{L}(\mathcal{M}(x))$.

Now assume that $\mathcal{M}$ is strongly equicontinuous. We claim that Φ_x is an order continuous Riesz homomorphism from M_I into $Z(\mathcal{M}(x))$. Indeed, suppose $T_\alpha \downarrow 0$ in M_I. Since $\mathcal{M}$ is strongly equicontinuous, M_I has Lebesgue topology and so $T_\alpha \to 0$. Therefore $T_\alpha[x] \to 0$ in $Z(\mathcal{M}(x))$, as Φ_x is continuous. Now it follows from $T_\alpha[x] \downarrow \geqq 0$ that $T_\alpha[x] \downarrow 0$ in $Z(\mathcal{M}(x))$, so Φ_x is order continuous. This implies in particular that the kernel N_x of Φ_x is a band in M_I. Consequently $M_I = N_x \oplus N_x^d$, and the restriction of Φ_x to N_x^d is a Riesz isomorphism. We will show next that Φ_x is surjective, i.e. that $Z(\mathcal{M}(x)) = M_I[x]$.

Proposition 4.10. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X, and take $x \in X$. Then $Z(\mathcal{M}(x)) = M_I[x]$.*

Proof. We show first that $\overline{\mathcal{M}[x]} = \bar{\mathcal{M}}[x]$, where the first closure is taken with respect to the strong operator topology in $\mathcal{L}(\mathcal{M}(x))$. Since Φ_x is continuous, it is clear that $\bar{\mathcal{M}}[x] \subseteq \overline{\mathcal{M}[x]}$. Furthermore, $\bar{\mathcal{M}}[x]$ is an equicontinuous Boolean algebra of projections in $\mathcal{M}(x)$. We assert that $\bar{\mathcal{M}}[x]$ is complete in the sense of Bade. Indeed as observed earlier, $M_I = N_x \oplus N_x^d$, where $N_x = \ker(\Phi_x)$, and Φ_x induces a Riesz isomorphism from N_x^d onto $M_I[x]$. Using that Φ_x is order continuous and that $\bar{\mathcal{M}}$ is complete in the sense of Bade, it follows easily that $\bar{\mathcal{M}}[x]$ is likewise complete in the sense of Bade. Now it follows from Proposition 4.5 that $\bar{\mathcal{M}}[x]$ is closed with respect to the strong operator topology in $\mathcal{L}(\mathcal{M}(x))$. Therefore, we may conclude that $\bar{\mathcal{M}}[x] = \overline{\mathcal{M}[x]}$.

Now take $A \in Z(\mathcal{M}(x))$ with $0 \leqq A \leqq I$. As observed above, $\overline{\mathcal{M}[x]}$ is the Boolean algebra of components of I in $Z(\mathcal{M}(x))$. Therefore, by the Freudenthal spectral theorem ([11], Theorem 40.2), there exists a sequence $\{A_n\}_{n=1}^\infty$ in $Z(\mathcal{M}(x))$ such that $0 \leqq A_n \uparrow A$ (I-uniformly) where each A_n is of the form $\sum_{i=1}^k \alpha_i F_i$, with $F_i \in \overline{\mathcal{M}[x]}$ mutually disjoint. Since $\overline{\mathcal{M}[x]} = \bar{\mathcal{M}}[x]$, there exist $T_n \in M_I$ with $T_n[x] = A_n$ for all n and $0 \leqq T_n \uparrow \leqq I$. Since M_I is Dedekind complete, there exists $T \in M_I$ such that $0 \leqq T_n \uparrow T$. By the order continuity of Φ_x, we get $A_n \uparrow T[x]$ and hence $A = T[x]$, so $A \in M_I[x]$. We may conclude, therefore, that $Z(\mathcal{M}(x)) = M_I[x]$.

Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X. Then $\bar{\mathcal{M}}$ is complete in the sense of Bade and equal to the Boolean algebra of components of I in M_I. Take $x \in X$ and let $N_x = \{T \in M_I : T[x] = 0\}$; then N_x is a band in M_I and N_x^d is isomorphic to $Z(\mathcal{M}(x))$. Since I is a strong order unit in M_I, it follows that the component of I in N_x^d is a strong order unit in N_x^d.

Definition 4.11. For $x \in X$, the *carrier of x in $\bar{\mathcal{M}}$* is defined by $E_x = \inf\{E \in \bar{\mathcal{M}} : Ex = x\}$.

Note that $E_x x = x$, and so $E_x[x] = I$. It is routine to show that E_x is the component of I in N_x^d. We thus have the following result.

Corollary 4.12. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X and let $x \in X$. Then $E_x M_I = \{T \in M_I \colon |T| \leqq n E_x$ for some $n \in \mathbb{N}\}$ is Riesz and algebra isomorphic with $Z(\mathcal{M}(x))$ (and the isomorphism is given by the restriction mapping).*

Corollary 4.13. *Same situation as above. The restriction mapping from M_I onto $Z(\mathcal{M}(x))$ is an isomorphism if and only if $E_x = I$.*

Remark. Same situation as above. Observe that $Z(\mathcal{M}(x))$ is Riesz isomorphic to $\mathcal{M}_x$, where the isomorphism is given by the mapping $T \to Tx$. Hence, the mapping φ_x from M_I into $\mathcal{M}_x$, defined by $\varphi_x(T) = Tx$, is a Riesz homomorphism onto. Moreover, φ_x is an isomorphism if and only if the carrier of x in $\bar{\mathcal{M}}$ is equal to I.

We end this section with some remarks that will be useful in the next section. Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X. Fix $x \in X$ and take $0 \leqq y \in \mathcal{M}(x)$. Since any $T \in M$ leaves $\mathcal{M}(x)$ invariant, it is clear that $M(y) \subseteq \mathcal{M}(x)$. Furthermore, it is easy to see that $M(y)$, with its canonical Riesz space structure, is a Riesz subspace of $\mathcal{M}(x)$. Therefore, $\mathcal{M}(y)$ is a Riesz subspace of $\mathcal{M}(x)$. Moreover, $\mathcal{M}(y)$ is an order ideal in $\mathcal{M}(x)$. Indeed, suppose $0 \leqq w \leqq z$ in $\mathcal{M}(x)$ with $z \in \mathcal{M}(y)$. Since $\mathcal{M}(x)$ is Dedekind complete, there exists $A \in Z(\mathcal{M}(x))$ such that $Az = w$. Using Proposition 4.10, it follows that A leaves $\mathcal{M}(y)$ invariant, so $w = Az \in \mathcal{M}(y)$. Furthermore, $\mathcal{M}(x)$ has Lebesgue topology and $\mathcal{M}(y)$ is a closed order ideal in $\mathcal{M}(x)$, hence $\mathcal{M}(y)$ is a band in $\mathcal{M}(x)$. Since any band in $\mathcal{M}(x)$ is closed, it is clear that $\mathcal{M}(y)$ is equal to the band generated by y in $\mathcal{M}(x)$. Finally, a moment's reflection shows that $E_y[x]$ is the band projection in $\mathcal{M}(x)$ onto $\mathcal{M}(y)$. We summarize the above results in the next proposition.

Proposition 4.14. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X, fix $x \in X$ and take $0 \leqq y \in \mathcal{M}(x)$. Then $\mathcal{M}(y)$ is a band in $\mathcal{M}(x)$ and $\mathcal{M}(y) = E_y \mathcal{M}(x)$.*

5. A Bade-Type Reflexivity Theorem

Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X.

Definition 5.1. An *$\mathcal{M}$-invariant subspace of* X is a closed linear subspace Y of X with $E(Y) \subseteq Y$ for all $E \in \mathcal{M}$.

In the present section, we are concerned with linear operators T from X into itself which leave invariant each $\mathcal{M}$-invariant subspace of X. First note that any such T commutes with all $E \in \mathcal{M}$.

Observe that any $T \in \hat{M} \subseteq \mathcal{L}^{\#}(X)$ leaves invariant every $\mathcal{M}$-invariant subspace of X. Indeed, suppose that Y is an $\mathcal{M}$-invariant subspace of X. Let $\{T_\alpha\}$ be a net in M with $T_\alpha \to T$, i.e. $T_\alpha x \to Tx$ for all $x \in X$. If $y \in Y$, then $T_\alpha y \in Y$ for all α, and since Y is closed, this implies that $Ty \in Y$. Hence, T leaves Y invariant.

Suppose $T \in \mathcal{L}^{\#}(X)$ leaves invariant all $\mathcal{M}$-invariant subspaces of X. Take $x \in X$, then $\mathcal{M}(x)$ is a Dedekind complete, complex, locally solid Riesz space with Lebesgue topology (Proposition 3.9). Since $\mathcal{M}(x)$ is an $\mathcal{M}$-invariant subspace, T leaves $\mathcal{M}(x)$ invariant. Let $T[x]$ be the restriction of T to $\mathcal{M}(x)$. Since the restriction of any $E\in\mathcal{M}$ to $\mathcal{M}(x)$ is a band projection in $\mathcal{M}(x)$, and since bands in $\mathcal{M}(x)$ are closed, it follows that any band in $\mathcal{M}(x)$ is an $\mathcal{M}$-invariant subspace. Hence $T[x]$ leaves all bands in $\mathcal{M}(x)$ invariant, i.e. $T[x]$ is a band preserving operator. Since $\mathcal{M}(x)$ is a locally convex, locally solid Riesz space, we can apply Proposition 2.6, and it follows that $T[x]$ is order bounded on $\mathcal{M}(x)$, i.e. $T[x]\in\mathrm{Orth}(\mathcal{M}(x))$. We thus have the following lemma.

Lemma 5.2. *Suppose $T\in\mathcal{L}^{\#}(X)$ leaves invariant all $\mathcal{M}$-invariant subspaces of X. For any $x\in X$, T leaves invariant $\mathcal{M}(x)$ and the restriction $T[x]$ of T to $\mathcal{M}(x)$ is an orthomorphism in $\mathcal{M}(x)$.*

We need the following observations concerning orthomorphisms. Let $L_{\mathbb{C}}=L+iL$ be a complex Riesz space which is Dedekind complete (i.e. L is Dedekind complete). Then $\mathrm{Orth}(L_{\mathbb{C}})$ is a complex Dedekind complete f-algebra. Take any $T\in\mathrm{Orth}(L_{\mathbb{C}})$. We claim that there exists a sequence $\{P_n\}_{n=1}^{\infty}$ of band projections in $L_{\mathbb{C}}$ such that $P_n\uparrow I$ and $|TP_n|\leqq nI$ for all n. Indeed, let P_n be the component of I in the band generated by $(nI-|T|)^+$, then $P_n\uparrow I$ and $P_n(|T|-nI)^+=0$ implies that $|P_nT|=P_n|T|\leqq nI$. Furthermore, if $T\in\mathrm{Orth}(L_{\mathbb{C}})$ and $T\neq 0$, then there exists a band projection $P\neq 0$ in $L_{\mathbb{C}}$ and real numbers $\alpha,\beta>0$ with $\alpha P\leqq|TP|\leqq\beta P$. Indeed, it follows from the above observation that there exists a band projection $Q\neq 0$ and $\beta>0$ such that $|TQ|\leqq\beta Q$ and $TQ\neq 0$. Now there exists $\alpha>0$ such that $(|TQ|-\alpha I)^+>0$. Let P be the component of I in the band generated by $(|TQ|-\alpha I)^+$. Then $0<P\leqq Q$ and $|TP|\leqq\beta P$. Since $P(\alpha I-|TQ|)^+=0$, it follows that $0<\alpha P\leqq|TP|$. Therefore α, β and P have the desired properties.

Lemma 5.3. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X. Suppose $T\in\mathcal{L}^{\#}(X)$ leaves all $\mathcal{M}$-invariant subspaces of X invariant and $Tx=0$ for some $x\in X$. Then $TE_x=0$.*

Proof. Suppose $TE_x\neq 0$. Then there exists $y\in X$ with $TE_xy\neq 0$. Consider the restriction $T[y]$ of T to $\mathcal{M}(y)$. Then $T[y]\in\mathrm{Orth}(\mathcal{M}(y))$ and $T[y]E_x[y]\neq 0$ in $\mathrm{Orth}(\mathcal{M}(y))$. By the remarks above, there exist $\alpha,\beta>0$ and $E\in\bar{\mathcal{M}}$ such that $0<\alpha E[y]E_x[y]\leqq|T[y]|E[y]E_x[y]\leqq\beta E[y]E_x[y]$. Set $EE_xy=z$, then $z\neq 0$ and $\mathcal{M}(z)=EE_x\mathcal{M}(y)$. Consider the restriction $T[z]$ of T to $\mathcal{M}(z)$, then $0<\alpha I\leqq|T[z]|\leqq\beta I$ in $Z(\mathcal{M}(z))$. Since $Z(\mathcal{M}(z))$ is a Dedekind complete, complex f-algebra, it follows that $T[z]^{-1}$ exists in $Z(\mathcal{M}(z))$ (see Sect. 2). By Corollary 4.12, the restriction mapping from E_zM_I into $Z(\mathcal{M}(z))$ is a Riesz isomorphism onto, so there exists $S\in E_zM_I$ such that $S[z]=T[z]^{-1}$.

Define $w=x+z$. Then $Tw=T(x+z)=Tz$, hence $Tw\in\mathcal{M}(z)$. Now $S(Tw)=S(T[z](z))=T[z]^{-1}T[z](z)=z$, and since S leaves $\mathcal{M}(w)$ invariant, we get $z\in\mathcal{M}(w)$. Hence $x\in\mathcal{M}(w)$.

We claim that x is a weak order unit in $\mathcal{M}(w)$. Indeed, suppose that $u\in\mathcal{M}(w)$ with $u\perp x$. Then there exists $E\in\bar{\mathcal{M}}$ such that $Ex=0$ and $Eu=u$. Now $(I-E)x=x$, so $I-E\geqq E_x$, and consequently, $(I-E)w\geqq E_xw=E_xx+E_xz=x+z=w$. Therefore $(I-E)[w]=I$, so $E[w]=0$, which implies that $Eu=0$, i.e., $u=0$. Hence, x is a weak order unit in $\mathcal{M}(w)$.

Since $T[w] \in \mathrm{Orth}(\mathcal{M}(w))$ and $T[w](x)=0$, we get $T[w]=0$. In particular $Tz = T(x+z) = Tw = T[w](w) = 0$. This contradicts the fact that $|T[z]| \geqq \alpha I > 0$ on $\mathcal{M}(z)$. We may conclude, therefore, that $TE_x y = 0$ for all $y \in x$, i.e. that $TE_x = 0$.

Lemma 5.4. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X. Suppose that $T \in \mathcal{L}^{\#}(X)$ leaves invariant all $\mathcal{M}$-invariant subspaces of X. Then $TE_x \in \hat{M}$ for all $x \in X$.*

Proof. Consider the restriction $T[x]$ of T to $\mathcal{M}(x)$, then $T[x] \in \mathrm{Orth}(\mathcal{M}(x))$. By the remarks preceding to Lemma 5.3, there exist $E_n \in \bar{\mathcal{M}}$ such that $E_n \uparrow E_x$ and $|T[x]| E_n[x] \leqq nI$. Since $|T[x]| E_n[x] \in Z(\mathcal{M}(x))$, it follows from Corollary 4.12 that there exists $S_n \in E_x M_I$ with $S_n[x] = T[x] E_n[x]$ $(n=1,2,\ldots)$. Now it follows from the lemma above that $(S_n - TE_n) E_x = 0$, i.e., $TE_n E_x = S_n$ $(n=1,2,\ldots)$. Note that T commutes with all projections in $\bar{\mathcal{M}}$, so $TE_n E_x = E_n TE_x$. Since $E_n \uparrow E_x$ in $\bar{\mathcal{M}}$, we have $E_n y \to E_x y$ for all $y \in X$, and consequently, $E_n TE_x y \to E_x TE_x y = TE_x y$ for all $y \in X$, i.e. $S_n y \to TE_x y$ for all $y \in X$. Hence $S_n \to TE_x$ with respect to the strong operator topology in $\hat{M}$, so $TE_x \in \hat{M}$.

Now we are in a position to prove the main result in this section.

Theorem 5.5 (Reflexivity Theorem). *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X, and let $T \in \mathcal{L}^{\#}(X)$ be given. Then $T \in \hat{M}$ if and only if T leaves invariant all $\mathcal{M}$-invariant subspaces of X.*

Proof. As observed already, any $T \in \hat{M}$ leaves invariant all $\mathcal{M}$-invariant subspaces of X.

Now assume that $T \in \mathcal{L}^{\#}(X)$ leaves invariant all $\mathcal{M}$-invariant subspaces of X. By Zorn's lemma, there exists a maximal disjoint system $\{E_\alpha\}_{\alpha \in A}$ of carrier projections in $\bar{\mathcal{M}}$. Note that $\sup_\alpha E_\alpha = I$ in $\bar{\mathcal{M}}$. For any finite subset $\mathcal{F}$ of A, define $E_{\mathcal{F}} = \sup_{\alpha \in \mathcal{F}} E_\alpha$, then $E_{\mathcal{F}} \uparrow I$, so $E_{\mathcal{F}} y \to y$ for all $y \in X$.

Define $T_\alpha = TE_\alpha$ and $T_{\mathcal{F}} = TE_{\mathcal{F}}$. By the Lemma above, we have $T_\alpha \in \hat{M}$ for all $\alpha \in A$ and so $T_{\mathcal{F}} \in \hat{M}$ for all finite subsets $\mathcal{F}$ of A. Now it follows from $E_{\mathcal{F}} y \to y$ for all $y \in X$ that $E_{\mathcal{F}} Ty \to Ty$ for all $y \in X$. Note that $E_{\mathcal{F}} T = TE_{\mathcal{F}} = T_{\mathcal{F}}$, hence $T_{\mathcal{F}} y \to Ty$ for all $y \in X$, i.e., $T_{\mathcal{F}} \to T$ with respect to the strong operator topology in $\mathcal{L}^{\#}(X)$. Hence $T \in \hat{M}$.

Corollary 5.6. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X. The following two statements are equivalent.*

(i) $T \in \bar{M}$

(ii) *$T \in \mathcal{L}(X)$ and T leaves invariant all $\mathcal{M}$-invariant subspaces of X.*

As observed in Proposition 4.3, if we assume in addition that $\mathcal{L}(X)$ is sequentially complete, then $\hat{M} = \bar{M}$. Combining this observation with Theorem 5.5, we get the following result.

Corollary 5.7. *Let $\mathcal{M}$ be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space X and assume that $\mathcal{L}(X)$ is sequentially complete. The following statements are equivalent.*

(i) $T \in \bar{M}$
(ii) $T \in \mathscr{L}^{\#}(X)$ *and* T *leaves invariant all* $\mathscr{M}$*-invariant subspaces of* X.

We remark that the above corollary applies in particular if X is a Banach space and in that case it extends the classical result of Bade ([1], Theorem 18, p. 2214).

We will conclude by presenting a simple application of the reflexivity theorem. For the case that X is Banach, the following result is due to Gillespie [8].

Proposition 5.8. *Let* $\mathscr{M}$ *be a strongly equicontinuous Boolean algebra of projections in the quasi-complete space* X. *Suppose that* $N \subseteq \bar{M}$ *is a strongly closed subalgebra of* $\mathscr{L}(X)$ *containing* I. *Assume that* $x \in X$ *is a cyclic vector for* $\mathscr{M}$. *The following statements are equivalent for any element* $T \in \mathscr{L}(X)$.

(i) $T \in N$
(ii) T *leaves invariant each* N*-invariant subspace.*

Proof. Since $N \subseteq \bar{M}$, each $\mathscr{M}$-invariant subspace is also N-invariant and consequently Corollary 5.6 implies that $T \in \bar{M}$. Let $x_1, \ldots x_n \in X$ and $\varepsilon > 0$ be given. In the locally convex solid Riesz space $\mathscr{M}(x)$, set $y = |x_1| + \ldots + |x_n|$ and let ρ be an arbitrary Riesz semi-norm defining the topology on $\mathscr{M}(x)$. Since Ty belongs to the smallest closed N-invariant subspace of X containing y, there exists $S \in N$ such that

$$\rho((T-S)y) = \rho(|T-S|\,y) < \varepsilon.$$

Now

$$\begin{aligned} \rho((T-S)x_i) &= \rho(|T-S|\,|x_i|) \\ &\leqq \rho(|T-S|\,y) < \varepsilon. \end{aligned}$$

It follows that $T \in N$ and the proof is complete.

Finally it seems to be an interesting question as to whether the statement of Proposition 5.8 remains valid if X does not have a cyclic vector (see [9]).

References

1. Aliprantis, C.D., Burkinshaw, O.: Locally solid Riesz spaces. New York-San Francisco-London: Academic Press 1978
2. Bade, W.G.: On Boolean algebras of projections and algebras of operators. Trans. Amer. Math. Soc. **80**, 343–359 (1955)
3. Beukers, F., Huijsmans, C.B., de Pagter, B.: Unital embedding and complexification of f-algebras. Math. Z. **183**, 131–144 (1983)
4. Bigard, A., Keimel, K., Wolfenstein, S.: Groupes et Anneaux Réticulés. Lecture Notes in Math. vol. **608**. Berlin-Heidelberg-New York: Springer 1977
5. Dodds, P.G., Ricker, W.: Spectral measures and the Bade reflexivity theorem. Preprint
6. Dodds, P.G.: Boolean algebras of projections in locally convex spaces. Proc. Mini-Conference on Operator Theory and Partial Differential Equations, Centre for Mathematical Analysis, A.N.U. 1983 (to appear)
7. Dunford, N., Schwartz, J.: Linear Operators Part III: Spectral Operators. New York: Wiley-Interscience 1971
8. Gillespie, T.A.: Cyclic Banach spaces and reflexive operator algebras. Proc. Roy. Soc. Edinburgh Sect. A **78**, 225–235 (1978)

9. Gillespie, T.A.: Boolean algebras of projections and reflexive algebras of operators. Proc. London Math. Soc. **37**, 56–74 (1978)
10. Huijsmans, C.B., de Pagter, B.: Ideal theory in f-algebras. Trans. Amer. Math. Soc. **269**, 225–245 (1982)
11. Luxemburg, W.A.J., Zaanen, A.C.: Riesz Spaces I. Amsterdam-London: North-Holland 1971
12. de Pagter, B.: Disjointness preserving operators. Proc. Amer. Math. Soc. **90**, 543–549 (1984)
13. Rall, C.: Über Boolesche Algebren von Projektionen. Math. Z. **153**, 199–217 (1977)
14. Ricker, W.: On Boolean algebras of projections and scalar-type spectral operators. Proc. Amer. Math. Soc. **87**, 73–77 (1983)
15. Schaefer, H.H.: Spectral measures in locally convex algebras. Acta Math. **107**, 125–173 (1962)
16. Schaefer, H.H.: Banach lattices and cyclic Banach spaces. Spectral Theory Symposium (Trinity College, Dublin 1974). Proc. Roy. Irish Acad. Sect. A**74**, 283–289 (1974)
17. Schaefer, H.H.: Aspects of Banach lattices. In: Studies in Functional Analysis (R.G. Bartle ed.). M.A.A. Studies in Mathematics **21**. Washington: Math. Assoc. of America 1980
18. Schaefer, H.H.: Topological Vector Spaces. Berlin-Heidelberg-New York: Springer 1971
19. Veksler, A.I.: Cyclic Banach spaces and Banach lattices. Dokl. Akad. Nauk SSSR **213**, 770–773 (1973) [Russian]; Transl.: Soviet Math. Dokl. **14**, 1773–1779 (1973)
20. Walsh, B.: Structure of spectral measures on locally convex spaces. Trans. Amer. Math. Soc. **120**, 295–326 (1965)
21. Zaanen, A.C.: Riesz Spaces II. Amsterdam-New York-Oxford: North-Holland 1983

Received January 13, 1984

Math. Z. 187, 383–392 (1984)

Mathematische Zeitschrift

Existentially Closed Locally Finite Central Extensions; Multipliers and Local Systems

Richard E. Phillips

Dept. of Mathematics, Michigan State Univ., East Lansing, Michigan 48824, USA

I. The Framework

In the recent paper [4], K. Hickin has introduced and studied a remarkable class of locally finite groups. We begin by reviewing and recasting *some* of Hickin's results specialized to the countable case.

Let A be a countable periodic Abelian group and

$$\mathscr{D}_A = \{G \mid G \text{ is countable, locally finite and } A \subseteq \zeta(G)\};$$

here, as elsewhere $\zeta(G)$ is the *center* of G. The inductive class $\mathscr{D}_A$ has existentially closed members; we let

$$E_A \text{ be any existentially closed member of } \mathscr{D}_A.$$

The three fundamental properties of E_A are the following.

(1.1) (A-universality). Let $(A)\eta \subseteq H$ be groups satisfying

i) η: $A \to (A)\eta$ is an onto isomorphism, and

ii) $(A)\eta \subseteq \zeta(H)$ and $H/(A)\eta$ is finite.

Then there is an isomorphism ρ: $H \to E_A$ such that $\eta\rho = \mathrm{id}(A)$.

(1.2) (A-homogeniety). Let H and K be subgroups of E_A with $A \subseteq H$, $A \subseteq K$ and both H/A and K/A finite. Suppose also that there is an isomorphism ψ from H onto K with $(a)\psi = a$ for all $a \in A$. Then there is a $t \in E_A$ such that $(h)\psi = h^t$ for all $h \in H$.

From (1.1) and (1.2) one easily deduces

(1.3) (A-injectivity). Suppose $(A)\eta \subseteq H \subseteq K$ are groups satisfying

i) η: $A \to (A)\eta$ is an isomorphism;

ii) $(A)\eta \subseteq \zeta(K)$ and $K/(A)\eta$ is finite;

iii) μ: $H \to E_A$ is an isomorphism with $\eta\mu = \mathrm{id}(A)$.

Then μ extends to an isomorphism ρ: $K \to E_A$.

The property (1.3) enables one to use a standard "back and forth" argument to prove the uniqueness of E_A. Thus

(1.4) E_A is the unique existentially closed member of $\mathscr{D}_A$.

In [4], Hickin's definition of the groups E_A is somewhat different, but the equivalence of our approach is noted in [4; § 1]. Our (1.3) is part of Hickin's definition of E_A (called $ULF(A)$) while (1.1) above is really a special case of (1.3). The properties (1.2) and (1.4) arise as Lemma 1 (b) and Theorem 3 (b) of [4]. As noted in [4; § 1], $E_{\{1\}}$ is P. Hall's universal countable locally finite group.

We now point out some additional relevant results of [4]; throughout, A and E_A are as above and H_A will denote the quotient group

$$H_A = E_A/A.$$

(1.5) H_A is simple; thus $A = \zeta(E_A)$ [4; Theorem 1 (a)].

(1.6) $H_A \simeq H_B$ if and only if $A \simeq B$ [4; Theorem 1 (d)].

(1.7) If the rank of A, $r(A)$, is greater than 2, then H_A is not a direct limit of finite simple groups [4; Theorem 1 (b)].

Here, as elsewhere in the sequel, $r(A)$ is the *Mal'cev rank* of A; i.e.,

$$r = r(A) = \sup\{d(B) \mid B \text{ is a finite subgroup of } A\},$$

where $d(B)$ is the smallest number of elements required to generate B. Evidently,

(1.8) if C is a section of A, then $r(C) \leqq r(A)$.

In any group G a pair of isomorphic subgroups H and K of G will be called a *homogeneous pair* if for every (onto) isomorphism ψ: $H \to K$, there is a $t \in G$ such that $(h)\psi = h^t$ for all $h \in H$.

(1.9) Let A be p-group for a prime p. Then H_A is finitely p'-homogeneous; i.e., every pair of isomorphic, finite p'-subgroups is a homogeneous pair. Also, $\mathscr{C}(H_A)$, the set of conjugacy classes of elements of order p in H_A is in $1-1$ correspondence with A/pA. Finally, H_A has two non-conjugate subgroups which are elementary Abelian of order p^2 [4; Theorem 1 (c), Theorem 5, Lemma 7].

We complete this list by noting an easy application of (1.1).

(1.10) If U is any finite group, there is an isomorphism η: $U \to H_A$. Thus, the groups H_A are universal.

II. Statement of Results

In this paper we derive additional properties of the groups E_A und H_A. Our first result is

Theorem A. *E_A is perfect and A is the Schur multiplier of H_A.*

This theorem asserts that E_A is the (unique) covering group of the perfect group H_A. Notice that (1.6) above follows readily from Theorem A. On this point we remark that Hickin's approach to (1.6) gives a technically stronger result.

The methods used in the proof of Theorem A will also be used to determine the existence of chains of finite subgroups of H_A that have certain properties. We require the notations $PSL(\ ,q)$, $PSU(\ ,q^2)$, $PSp(\ ,q)$ for the families of finite projective, special linear, unitary and symplectic *simple* groups.

Theorem B. *Let F be a finite perfect subgroup of H_A. Then there is a positive integer $d=d(F)$ and infinite sets of primes P_i, Q_i, $1 \leqq i \leqq d$ satisfying the following two properties.*

i) *For every d-tuple $\bar{p}=(p_1,\dots,p_d)\in P_1 \times \dots \times P_d$, there is a subgroup $H=H(\bar{p})$ of H_A satisfying*

a) *$F \subseteq H$, and*

b) *$H \simeq Dr\{PSL(m,p_i) \mid 1 \leqq i \leqq d\}$; here $m=m(F)$ is fixed, $m>2$, and $m \mid |F|$: m and d depend only on F and not on $\bar{p}$.*

ii) *For every d-tuple $\bar{p}=(p_1,\dots,p_d)\in Q_1 \times \dots \times Q_d$ there is a subgroup $K=K(\bar{p})$ of H_A satisfying*

a) *$F \subseteq K$, and*

b) *$K \simeq Dr\{PSU(m,p^2) \mid 1 \leqq i \leqq d\}$ (same comment on m and d as above).*

In Theorem B, the symbol Dr means, of course, direct product.

If G is any countable locally finite group, an *approximating sequence* of G is an ascending chain $\{G_n \mid n \geqq 1\}$ of finite subgroups of G with $\bigcup G_n = G$.

As a consequence of Theorem B we have the following.

Corollary. *For each A, H_A has an approximating sequence $\{H_n\}$ where each H_n is a direct product of finite non-Abelian simple groups.*

The Corollary follows easily from Theorem B when one notices

(2.1) Every countable locally finite perfect group G has an approximating sequence of finite perfect subgroups.

We will call a finite group G *completely reducible* ($=CR$) if G is a direct product of non-Abelian simple groups. While (in the $r(A)>2$ case) H_A has no approximating sequence consisting solely of simple groups, the Corollary asserts that there are an abundant number of approximating sequences consisting of CR-groups.

It seems woth mentioning at this point that all known (to the author) countable locally finite simple groups G have approximating sequences $\{G_n\}$ where $G_n/\zeta(G_n)$ is CR. It would be interesting to know whether or not this is uniformly the case (to place these remarks in context, the reader should consult [5; p. 115], [8; p. 26] and (1.7) as well as our Theorems B and C.

The proof of Theorem B (and that of Theorem A) does not use the simplicity of the groups H_A. In fact, the proof of Theorem B can easily be adapted to provide a proof of this fact. We give such a proof in Sect. (3.4).

The following proposition shows that there are severe restrictions on the types of approximating sequences in H_A (as indicated in (1.7)).

Proposition. *Let $r_0 \leqq r(A)$ with r_0 finite. Then there is a finite subgroup W of H_A such that for every perfect F (finite or infinite) with $W \leqq F \leqq H_A$ we have*

$r_0 \leqq r(M(F))$ where $M(F)$ is the Schur multiplier of F. Thus, if $\{H_n\}$ is any chain of perfect subgroups with $\bigcup H_n = H_A$, then there is an n_0 such that for all $n \geqq n_0$, $r(M(H_n)) \geqq r_0$.

In the following Theorem C, we use Theorem B and the Proposition to investigate the number of direct factors that occur in the groups H_n of an approximating sequence $\{H_n\}$ of CR subgroups of H_A.

If G is a CR-group, then $c(G)$, the *composition length* of G, is the number of direct factors of G. For any finite perfect subgroup F of H_A, we define $C \equiv C(F)$ by

$$C(F) = \min\{c(Y) \mid F \subseteq Y \subseteq H_A,\ Y \text{ is } CR\}.$$

Theorem C. i) *Let $\{H_n\}$ be an approximating sequence of perfect subgroups of H_A.* If $r = r(A)$ *is infinite, then* $\lim\limits_{n\to\infty} C(H_n) = \infty$.

ii) *Suppose r is finite and let F be a finite perfect subgroup of H_A. Then*

$$C(F) \leqq d \leqq r$$

where d is as in Theorem B. *Further, there is a finite subgroup W of H_A such that if $W \subseteq F$, then*

$$r/2 \leqq C(F) \leqq d \leqq r.$$

iii) *Let $A \simeq C_2$ (cyclic group of order 2) and F be a finite perfect subgroup of H_A. Then there are subgroups of H_A of types $PSL(\ ,p)$, $PSU(\ ,p^2)$ and $PSp(\ ,p)$ that contain F. In all cases, the primes may be chosen in an infinity of different ways.*

The lower bound aspect of (ii) of Theorem C may be viewed as a generalization of (1.7). The corresponding fact on upper bounds shows that, if $r(A) = 1$, then H_A has an approximating sequence of finite simple groups; in fact, Theorems B and C combine to show that in this case H_A has approximating sequences $\{H_n\}$ and $\{V_n\}$, where the H_n's are PSL's and the V_n's are PSU's. Part (iii) of Theorem C provides a third type of approximating sequence of simple groups for the group H_{C_2}.

The proofs of Theorem C(i) and the lower bound part of Theorem C(ii) use the classification of finite simple groups.

We close the section by indicating some homogeneity properties of the groups H_A when A is a cyclic or quâsi-cyclic p-group, p a prime. As usual, C_{p^n} is a cyclic group of order p^n and C_{p^∞} a group of type p^∞.

A locally finite group is *exactly p'-homogeneous* if every pair of finite isomorphic p'-subgroups is a homogeneous pair, and there is a pair of finite, isomorphic, non-conjugate subgroups. From (1.9) the groups H_A are exactly p'-homogeneous whenever A is a p-group. Moreover, Theorem C provides a variety of examples of simple exactly p'-homogeneous groups which are also direct limits of finite simple groups (in the case $A = C_2$ of three different types). The case $A = C_{p^\infty}$ also seems noteworthy, as here H_A is exactly p'-homogeneous and the elements of order p are all conjugate (see (1.9)).

Aside from the methods of these constructions, where homogeniety is built into the groups "a priori", it appears quite difficult to construct exactly p'-homogeneous direct limits of finite simple groups. It would be especially interesting to explicitly construct embeddings, say of a chain of PSL's, in a manner that forces exact p'-homogeniety in the direct limit. Considerations of this nature will be the topic of a subsequent paper [6].

III. Proofs

(3.1) *Preliminary Lemmas.* We first establish a somewhat stronger form of the injectivity result (1.3).

Lemma 1. i) *Let Q be a subgroup of E_A with $A \subseteq Q$ and Q/A finite. Suppose also that V is a group and that δ: $Q \to V$ is an isomorphism with $(A)\delta \subseteq \zeta(V)$ and $V/(A)\delta$ finite. Then there is an isomorphism μ: $V \to E_A$ such that $\delta\mu = \mathrm{id}(Q)$.*

ii) *Let F be a finite subgroup of E_A, S a finite group and τ: $F \to S$ an isomorphism with $(F \cap A)\tau \subseteq \zeta(S)$. Then there is an isomorphism ν: $S \to E_A$ with $\tau\nu = \mathrm{id}(F)$.*

Proof. Part (i) follows easily from (1.3) if take μ to be the extension of δ^{-1}: $(Q)\delta \to E_A$.

For the proof of (ii) we enlarge the subgroups S and F in such a way that part (i) is applicable. Note that the subgroup $Q = FA \subseteq E_A$ is a central product of F and A.

Let ψ be the isomorphism from $(F \cap A)\tau$ into A defined by $(x)\tau\psi = x$ for all $x \in A$, and let V be the central product of S and A with the identifications defined by ψ. Thus, there are isomorphisms α: $S \to (S)\alpha$ and β: $A \to (A)\beta$ such that $V = \langle (S)\alpha, (A)\beta \rangle$ and for all $x \in F \cap A$, $(x)\tau\alpha = (x)\beta$.

The relation δ: $Q \to V$ defined by $(fa)\delta = ((f)\tau\alpha)((a)\beta)$, $f \in F$, $a \in A$ defines an isomorphism from Q into V (since δ preserves defining relations); further $(A)\delta \subseteq \zeta(V)$ (since A is Abelian). Thus δ satisfies the conditions of part (i) and so there is an isomorphism μ: $V \to E_A$ such that $\delta\mu = \mathrm{id}(Q)$. Thus, if $f \in F$, we have $f = (f)\delta\mu = (f)\tau\alpha\mu$, and $\alpha\mu$ is the desired isomorphism from S into E_A.

As a first application of Lemma 1, we have

(3.1.1) If F is a finite subgroup of E_A, there is a finite subgroup W of E_A with $F \subseteq W'$. Thus, E_A is a perfect group.

To prove (3.1.1), let F be a finite subgroup of E_A and let

$$\Pi(F) = \{p \mid p \text{ is a prime and } p \mid |F|\}.$$

Now, put $R = Dr\{C_p \mid p \in \Pi(F)\}$ and let Δ: $F \to W = F\,wr\,R$ be the diagonal embedding of F into the base group of W. As is well knwon [7; p. 98] every element of $(F)\Delta$ is a commutator in W; further, $(F \cap A)\Delta \subseteq \zeta(W)$. Application of Lemma 1(b) now completes the proof.

Lemma 2. *Let R be a finite perfect group with cyclic center $B \simeq C_n$ ($n = 1$ is a possibility) and let $m = |R/B|$.*

i) *There is an infinite set of primes P, such that for each $p \in P$, there is an isomorphism α_p: $R \to SL(m, p)$ satisfying*

a) $(B)\alpha_p \subseteq \zeta(SL(m,p))$, *and*

b) $(B)\alpha_p$ *is a (possibly trivial) direct factor of* $\zeta(SL(m,p))$.

ii) *There is an infinite set of primes* Q *such that for each* $p \in Q$, *there is an isomorphism* β_p: $R \to SU(m,p^2)$ *satisfying*

a) $(B)\beta_p \subseteq \zeta(SU(m,p^2))$, *and*

b) $(B)\beta_p$ *is a direct factor of* $\zeta(SU(m,p^2))$.

Proof. Suppose first that $n \neq 1$ and let $n = p_1^{n_1} \dots p_t^{n_t}$ be the prime power factorization of n. Put

$$n^+ = p_1^{n_1+1} \dots p_t^{n_t+1};$$

then the arithmetic sequence $\{sn^+ + n + 1 \mid s \geqq 1\}$ contains an infinite set of primes P. We note

(3.1.2) if $p \in P$, then $n \mid (p-1)$, while for any prime divisor q of n, $qn \nmid (p-1)$.

For $p \in P$, let σ_p: $B \simeq C_n \to GF(p)^*$ be an isomorphism from B into the multiplicative group of the field $GF(p)$; σ_p is a 1-dimensional representation of B over $GF(p)$. Let α_p be the induced representation from R into $GL(m,p)$ (as in [2; pp. 73–75]). Recall that the matrices $(r)\alpha_p$, $r \in R$, are described as follows:

(3.1.3) Let $\{1, \dots, r_m\}$ be a transversal of B in R; if $y \in R$, the (i,j)-th entry of $(r)\alpha_p$ is given by

$$((r)\alpha_p)_{(i,j)} = \begin{cases} 0 & \text{if } r_i^{-1} r r_j \notin B \\ (r_i^{-1} r r_j)\sigma_p & \text{if } r_i^{-1} r r_j \in B. \end{cases}$$

The mapping α_p is an isomorphism [2; p. 76] and since R is perfect, we have $(R)\alpha_p \subseteq SL(m,p)$. Further, the matrices $(r)\alpha_p$, $r \in R$, are monomial, and a very easy argument shows that if $r \in B$, then $(r)\alpha_p$ is the scalar matrix $((r)\sigma_p)I_m$. Thus, $(B)\alpha_p \subseteq \zeta(SL(m,p)) = Z$.

The group Z is cyclic of order $gcd(m, p-1)$, and (3.1.2) implies that $(B)\alpha_p$ is a direct product of some of the primary components of Z. Thus, $(B)\alpha_p$ is a direct factor of Z and this completes the proof of (i) when $n \neq 1$.

If $n = 1$, take P to be the set of all primes, and proceed as above; in this case $(B)\alpha_p = 1$, and $(B)\alpha_p$ is a trivial direct factor of Z.

For the proof of (ii), we again assume $n \neq 1$; we use the same notation as in the proof of (i) and let Q be an infinite set of primes occuring in the sequence $\{sn^+ + n - 1 \mid s \geqq 1\}$. Then

(3.1.4) if $p \in Q$, then $n \mid (p+1)$ and if q is a prime dividing n, then $qn \nmid (p+1)$.

For $p \in Q$, let τ_p: $B \to U \subseteq GF(p^2)^*$ be an isomorphism from B into U, the unique subgroup of order $p+1$ in $GF(p^2)^*$. The induced representation β_p: $R \to GL(m,p^2)$ is an isomorphism and

(3.1.5) for each $r \in R$, the matrix $(r)\beta_p$ is monomial with all entries in U.

Now σ: $x \to x^p$ is the unique automorphism of order 2 of $GF(p^2)$ and for every $x \in U$, we have $xx^\sigma = x^{p+1} = 1$; thus $x \in U$ if and only if $x^{-1} = x^\sigma$. An easy check shows that

$$((r)\beta_p)^{-1} = ((r)\beta_p)^\sigma)^t,$$

where t = transpose. Thus, $(R)\beta_p \subseteq SU(m,p^2)$ and again, the scalar matrices $(B)\beta_p$ are contained in $\zeta(SU(m,p^2)) = Z$.

Since Z is cyclic of order $gcd(m,p+1)$, (3.1.4) implies that $(B)\beta_p$ is a direct factor of Z. The $n=1$ case follows as in part (i).

We adopt the following convention regarding direct decompositions of a finite Abelian group B into a direct product of cyclic groups:

$$B = Dr\{B_i \mid 1 \leqq i \leqq d,\ B_i \simeq C_{n_i}\}.$$

If $B=1$, we always assume that $d=1$ and $B_1=1$. If $B \neq 1$, all B_i's are assumed to be non-trivial.

Standard arguments can be used to prove the following lemma ... we omit the proof.

Lemma 3. *Let R be a finite perfect group with central subgroup $B = Dr\{B_i \mid 1 \leqq i \leqq d,\ B \simeq C_{n_i}\}$, and for each i, $1 \leqq i \leqq d$, let $K_i = \langle B_j \mid i \neq j\rangle$. The function*

$$\Delta\colon R \to Dr\{R/K_i \mid 1 \leqq i \leqq d\} = V \quad \textit{defined by}$$
$$(r)\Delta = (rK_1, \ldots, rK_d)$$

is an isomorphism from R into V, and $(B)\Delta \subseteq \zeta(V)$. Note that each of the direct factors R/K_i is perfect and that $C_{n_i} \simeq D_i = B/K_i \subseteq \zeta(R/K_i)$.

We will use the following combined form of Lemmas 2 and 3.

Lemma 4. *Let R, $B = Dr\{B_i \mid 1 \leqq i \leqq d,\ B_i \simeq C_{n_i}\}$, $m = |R/B|$ and $\Delta\colon R \to V = Dr\{R/K_i \mid 1 \leqq i \leqq d\}$ be as in Lemma 3. For $1 \leqq i \leqq d$, let P_i be the set of primes associated with C_{n_i}, as in Lemma 2(i). Let $\bar{p} = (p_1, \ldots, p_d) \in P_1 \times \ldots \times P_d$; we define*

$$\alpha_{\bar{p}}\colon V \to U = Dr\{SL(m,p_i) \mid 1 \leqq i \leqq d\}$$

by

$$(g_1, \ldots, g_d)\alpha_{\bar{p}} = ((g_1)\alpha_{p_1}, \ldots, (g_d)\alpha_{p_d}),$$

where the α_{p_i} are as in Lemma 2.

Then $\alpha_{\bar{p}}$ is an isomorphism, $(\zeta(V))\alpha_{\bar{p}} \subseteq \zeta(U)$ and $(B)\Delta\alpha_{\bar{p}} \subseteq \zeta(U)$.

The isomorphism

$$\beta_{\bar{p}}\colon V \to Dr\{SU(m,p_i^2) \mid 1 \leqq i \leqq d\} = Y$$

is defined in the same way and has properties identical with those of $\alpha_{\bar{p}}$; here the sets of primes are as in Lemma 2(ii).

(3.2) *Proof of Theorem A.* Here, we freely use the theory of the Schur mutlipier as presented in [1; Chap. 1 and 2].

Let R be a finite perfect subgroup of E_A and $B = R \cap A$. Then, by Lemma 4, there is an isomorphism $\delta\colon R \to U = Dr\{SL(m,p_i) \mid 1 \leqq i \leqq d\}$ such that $(B)\delta \subseteq \zeta(U)$. From Lemma 1(b) there is a subgroup $\bar{R}$ of E_A with $\bar{R} \simeq U$ and $R \subseteq \bar{R}$.

Now, by (3.1.1), E_A is perfect and according to (2.1), E_A has an approximating sequence $\{E_n\}$ with each E_n a finite perfect group. Define subgroups N_n of E_A inductively by

$$N_1 = \bar{E}_1, \ldots, N_n = \overline{\langle N_{n-1}, E_n\rangle}, \ldots.$$

Then $\{N_n\}$ is an approximating sequence of E_A. As is well knwon, the Schur multiplier $M(SL(m,p))$ is zero if $m>2$, and for any direct product U of finite groups of type $SL(m,p)$, $m \geqq 2$ we also have $M(U)=0$ [1; p. 109]. Thus, for each n, $M(N_n)=0$ and since E_A is a direct limit of the N_n's, we also have $M(E_A)=0$ [1; p. 57]. Since E_A is a central "stem" extension of A by H_A (see (3.1.1)) we have $A=M(H_A)$, as desired.

(3.3) *Proof of Theorem B.* Techniques here are basically the same as in Sect. 3.2.

Let F be a finite perfect subgroup of H_A. Then $F=W/A$ where $W \subseteq E_A$ and there is a finite $R \subseteq E_A$ such that $W=RA$. Now

(3.3.1) $F=W/A=RA/A \simeq R/R \cap A$, and replacing R by R' if necessary, we may assume that R is perfect.

Put $B=R \cap A$ and write $B=Dr\{B_i \mid 1 \leqq i \leqq d\}$ as in Lemma 3. From Lemma 4, we have the isomorphism

$$\Delta \alpha_{\bar{p}}\colon\ R \to U = Dr\{SL(m,p_i) \mid 1 \leqq i \leqq d\}.$$

For each of the coordinate subgroups $SL(m,p_i)$ of U, write $Z_i=\zeta(SL(m,p_i)) = L_i \times (D_i)\alpha_{p_i}$; this is possible by Lemma 2(i) (here $D_i \simeq B_i$ is as in Lemma 3).

Let

$$\rho\colon\ U \to Dr\{SL(m,p_i)/L_i \mid 1 \leqq i \leqq d\} = S$$

be the product of the natural quotient maps $SL(m,p_i) \to SL(m,p_i)/L_i$. Now

(3.3.2) $\Delta \alpha_{\bar{p}} \rho\colon\ R \to S$ is an isomorphism (since $\ker \rho \cap \mathrm{Im}(\Delta \alpha_{\bar{p}})=1$), and $(B)\Delta \alpha_{\bar{p}} \rho = \zeta(S)$.

From Lemma 1 (ii),

(3.3.3) there is an isomorphism $v\colon S \to E_A$ such that $(\Delta \alpha_{\bar{p}} \rho)v = \mathrm{id}(R)$. Thus, with $S_1=(S)v$, we have $R \subseteq S_1$, and $R \cap A = S_1 \cap A = \zeta(S_1)$.

Thus

(3.3.4) $F=RA/A \subseteq S_1A/A \simeq S_1/S_1 \cap A = S_1/\zeta(S_1)$, and

$$S_1/\zeta(S_1) \simeq Dr\{PSL(m,p_i) \mid 1 \leqq i \leqq d\} = Y.$$

This completes the proof of part (i) of Theorem B. The proof of part (ii) is identical; here one uses the relevant parts of Lemma 2(ii).

(3.4) *Simplicity of H_A.* Here we give an alternate proof that the groups H_A are simple. All terminology will be as in Sect. (3.3).

Let F be a finite perfect subgroup of H_A. As in (3.3.1) write $F=RA/A \simeq R/R \cap A$ where R is a finite perfect subgroup of E_A.

With $\gamma = \Delta \alpha_{\bar{p}} \rho$ as in (3.3.2) we have

(3.4.1) $$((r)\gamma)^S \zeta(S) = S \quad \text{for every } r \in R - B.$$

This follows from the fact that $(r)\gamma$ projects non-trivially onto every coordinate of S.

We let S_1 and v be as in (3.3.3); then for every $r \in R - B$,

$$(r^{S_1})B = ((r)\gamma v)^{((S)v)}(\zeta(S))v = ((r)\gamma)^S \zeta(S))v = (S)v = S$$

(here we are using (3.4.1).

If we now let $\bar{F} = S_1 A/A$, it follows from (3.4.1) and the various isomorphisms in (3.3.4) that

(3.4.2) $$x^{\bar{F}} = \bar{F} \quad \text{for all } 1 \neq x \in F.$$

Let $\{F_n\}$ be an approximating sequence of perfect subgroups of H_A. We inductively define a sequence $\{H_n\}$ by

$$H_1 = \bar{F}_1, \ldots, H_n = \overline{\langle F_n, H_{n-1} \rangle_1}, \ldots$$

Then, if $1 \neq x \in F_i \subseteq H_i$, $x^{H_t} = H_t$ for all $t \geqq i+1$; thus $x^{H_A} = H_A$ and H_A is simple.

(3.5) *Proof of the Proposition.* Let $r_0 \leqq r(A)$ be finite and A_0 be a finite subgroup of A with $r_0 = r(A_0)$. Since E_A is perfect, there is a finite perfect subgroup Y of E_A such that $A_0 \subseteq Y$. Let $W = YA/A$ and F be a perfect subgroup of H_A with $W \subseteq F$. There is a perfect subgroup D of E_A such that $F = DA/A$. Now $F \simeq D/D \cap A$ and since the group D is perfect, $D \cap A$ is an image of $M(F)$. Further, $A_0 \subseteq Y = Y' \subseteq (DA)' = D$ and we have $A_0 \subseteq D \cap A$. Thus

$$r_0 = r(A_0) \leqq r(D \cap A) \leqq r(M(F)),$$

and this completes the proof of the Proposition.

We note without proof a stronger form of the Proposition in the case where $r = r(A)$ is finite (this is not used herein).

(3.5.1) If $r = r(A)$ is finite, there is a finite subgroup W of H_A such that for *every* F with $W \subseteq F \subseteq H_A$, $r \leqq r(M(F))$.

(3.6) *Proof of Theorem C.* Parts (i) and (ii) of Theorem C depend on the classification of the finite non-Abelian simple groups and their Schur multipliers. In particular, we require

(3.6.1) [3]. If F is a finite non-Abelian simple group, then $r(M(F)) \leqq 2$.

From this it follows immediately that

(3.6.2) If F is CR, then $r(M(F)) \leqq 2(c(F))$ [1; p. 109].

(3.6.3) *Proof of Theorem* C(i). Let k be a positive integer, A_k be a subgroup of A of rank k and W_k the subgroup of the Proposition. If n_k is such that $W_k \subseteq H_{n_k}$ then the Proposition asserts that $k \leqq r(H_n)$ for all $n \geqq n_k$; this completes the proof.

(3.6.4) *Proof of Theorem* C(ii). Write $F = RA/A$ where $R' = R$ is a finite subgroup of E_A, and put $B = R \cap A$. Let $d = r(B)$; then $d \leqq r$ and B is directly decomposable as a direct product of d cyclic groups. From (3.3.4), there is a $Y \subseteq H_A$ such that $F \subseteq Y$ and $Y \simeq Dr\{PSL(m, p_i) | 1 \leqq i \leqq d\}$ (this is the d of Theorem B). Thus $C(F) \leqq d \leqq r$, as required.

For the remainder of Theorem C(ii) let W be the subgroup of the Proposition such that $W \subseteq F \subseteq H_A$ and F perfect implies $r \leqq r(M(F))$. Then if Y is a CR-subgroup of H_A with $W \subseteq F \subseteq Y$, (3.6.2) gives

$$r \leqq r(M(Y) \leqq 2(C(Y)),$$

and so

$$r/2 \leqq C(F) \leqq d \leqq r.$$

(3.6.5) *Proof of Theorem* C(iii). Let $A = C_2$ and F be a finite perfect subgroup of H_A. Then F is contained in a perfect subgroup $F_1 = R/A$ where $R' = R$ is finite and $A \subseteq R$. There is no loss in assuming that $F_1 = F$.

The assertion regarding the PSL's and PSU's follows from Theorems B and C(ii). Thus, we concern ourselves only with the PSp's.

From Lemma 2 we have isomorphisms α_p (for an infinity of p) from R into $SL(m,p)$ such that $(A)\alpha_p = (R \cap A)\alpha_p \subseteq Z = \zeta(SL(m,p))$; in fact, $(A)\alpha_p = \{I_m, -I_m\}$. Let ψ_p: $SL(m,p) \to Sp(2m,p)$ be defined by

$$(G)\psi_p = \left(\begin{array}{c|c} G & 0 \\ \hline 0 & (G^t)^{-1} \end{array}\right).$$

Then $(A)\alpha_p\psi_p = \zeta(Sp(2m,p)) \simeq C_2$. By Lemma 1(ii) there is an isomorphism v_p: $Sp(2m,p) \to E_A$ such that $\alpha_p\psi_p v_p = \mathrm{id}(R)$. Thus, with $V = (Sp(2m,p))v_p$, we have $R \subseteq V$ and $A = \zeta(V)$. Thus, $F = R/A \subset V/\zeta(V) \simeq PSp(2m,p)$ and this completes the proof.

(3.6.6) The essential feature of the preceding proof (as well as several of our other arguments) is the existence of an isomorphism τ: $R \to C =$ the covering group of $PSp(2m,p)$, with the property that $(A)\tau = (R \cap A)\tau \subseteq \zeta(C)$. One expects a similar phenomena to hold for covering groups of the alternating groups and at least some of the projective orthogonal groups. In these later cases, the covering groups are far more complicated than the cases we have handled (i.e., PSL, PSU, Psp) and herein lies the difficulty in extending Theorem C(iii) to other classes of finite simple groups.

References

1. Beyl, F., Tappe, J.: Group Extensions. Representations and the Schur Multiplier. Lecture notes in Math. Vol. **958.** Berlin-Heidelberg-New York: Springer 1982
2. Curtis, C., Reiner, I.: Representation Theory of Finite Groups and Associative Algebras. New York: Interscience 1962
3. Griess, R.L.: Schur multipliers of the known finite simple groups II, Symposia in Pure Math., AMS **37**, 279–282 (1980)
4. Hickin, K.: Universal locally finite ccentral extensions of groups. To appear
5. Kegel, O., Wehrfritz, B.: Locally Finite Groups. Amsterdam: North Holland 1973
6. Phillips, R.E.: Partial homogeniety in locally finite simple groups. To appear
7. Schenkeman. E.: Group Theory. Princeton, New Jersey: Van Nostrand 1965
8. Zalesskiĭ, A.E., Serežkin, V.N.: Linear groups generated by transvections. Math. USSR. Izvestija **10**, 25–46 (1976)

Received April 7, 1984

Math. Z. 187, 393–404 (1984)

Mathematische Zeitschrift

Eigenvalue Inequalities for Minimal Submanifolds and P-Manifolds

Antonio Ros

Departamento de Geometría y Topología, Facultad de Ciencias, Universidad, Granada, Spain

Introduction

Let M be a compact Riemannian manifold, Δ the Laplacian of M and $\operatorname{Spec}(M) = \{0<\lambda_1,\dots,\lambda_1<\lambda_2,\dots,\lambda_2<\dots\}$ the spectrum of Δ, where each eigenvalue is repeated as many times as its multiplicity indicates. We denote by $V_u = \{f \in C^\infty(M)/\Delta f = \lambda_u f\}$ *the space of the* λ_u-eigenfunctions. The first part of this paper is related to the following problem: What is the eigenvalue-behaviour of the product of eigenfunctions?

If we take as a model the λ_1-eigenfunctions of compact symmetric spaces of rank one we have that the good behaviour is given by

$$V_1 V_1 \subset V_0 + V_1 + V_2. \tag{0.0}$$

We will study the above problem only in a very special case: When M admits a minimal isometric immersion into a sphere and the eigenfunctions are the coordinate functions of the immersion. In this case we can *organize* the product of eigenfunctions as a new isometric immersion in some Euclidean space. We found that the corresponding version of (0.0) holds if and only if M verifies an intrinsic condition (M is Einstein) and an extrinsic one (certain tensor in the normal bundle is proportional to the metric). Moreover, we give inequalities involving only spectral invariants, λ_1, λ_2 and others, the equality holding if and only if (0.0) holds. So, in this special case, (0.0) is a spectral property.

In a different sense, this problem has been treated by P. Yang and S.T. Yau [12]. A parallel version for submanifolds in the complex projective space has been obtained by the author in [7, 8].

Section 4 is dedicated to a different question. We say that a compact Riemannian manifold M is a *P-manifold* if any geodesic in M is periodic (the standard reference is A. Besse [3]). The geodesics of these manifolds admit a minimum common period. By scaling we suppose that this period is 2π, and we will say that M is a $P_{2\pi}$*-manifold*. The canonical examples are compact symmetric spaces of rank one. Various conjectures assert that with certain additional conditions any $P_{2\pi}$-manifold is a canonical example. In order to obtain a possible *eigenvalue proof* of some of these, it is interesting to give good estimates for the first eigenvalue of these manifolds (for an instructive *eigenvalue proof* see the Cheng version of the sphere theorem [4]). Here, using essentially a Berger argument [1] (see also [3], p. 211), we give a sharp lower bound of λ_1, in the sense that the equality holds for canonical examples, in terms of the dimension of M and a lower bound of the Ricci curvature. More precisely, if $\dim M = n$ and k is a lower bound of the Ricci curvature of M, we obtain

$$\lambda_1 \geqq \tfrac{1}{3}(2k+n+2). \tag{0.1}$$

For the class of $P_{2\pi}$-manifolds (0.1) is an improvement of the classical Lichnerowicz's lower bound of λ_1, see [2], p. 179.

1. Preliminaries

To fix notation we give the following definitions. Let M^n be an n-dimensional isometrically immersed submanifold of $\bar{M}^m$. Let X, Y, Z (resp. ξ) be tangent (resp. normal) vector fields to M. Let $\bar{\nabla}$ and ∇ be the Riemannian connection of $\bar{M}$ and M respectively and $\nabla^\perp$ the normal connection of M in $\bar{M}$. The second fundamental form σ, and the Weingarten endomorphism A of the immersion are given by $\sigma(X,Y) = \bar{\nabla}_X Y - \nabla_X Y$ and $A_\xi X = \nabla_X^\perp \xi - \bar{\nabla}_X \xi$ respectively. Moreover the covariant derivative of σ is given by $(\nabla\sigma)_X(Y,Z) = \nabla_X^\perp \sigma(Y,Z) - \sigma(\nabla_X Y, Z) - \sigma(Y, \nabla_X Z)$, and the mean curvature vector of the immersion is $H = \frac{1}{n}\sum_i \sigma(E_i, E_i)$, where $\{E_i\}_{i=1,\dots,n}$ is an orthonormal basis in the tangent space at any point of M.

Let ϕ: $M^n \to E^m$ be an isometric immersion of a n-dimensional compact Riemannian manifold into the Euclidean space. If Δ is the Laplacian of M acting on functions, we denote also by Δ the natural extension of this operator to the space of differentiable mappings of M in E^m. We have the decomposition $\phi = \sum_{u \geqq 0} \phi_u$, $u \in \mathbb{N}$, where ϕ_u: $M \to E^m$ is a differentiable mapping, $\Delta\phi_u = \lambda_u \phi_u$, and the addition is convergent coordinate to coordinate, for the L^2-topology on $C^\infty(M)$. Moreover ϕ_0 is a constant mapping (it is the center of gravity of M) and $\{\phi_u\}_{u\in N}$ are orthogonal mappings, that is

$$\int_M \langle \phi_u, \phi_v \rangle = 0 \quad \text{for all } u, v \in \mathbb{N}, u \neq v, \tag{1.1}$$

where $\langle\ ,\ \rangle$ is the Euclidean metric on E^m.

We have the relations

$$\Delta\phi = -nH = \sum_{u \geqq 1} \lambda_u \phi_u, \tag{1.2}$$

$$\Delta^2\phi = -n\Delta H = \sum_{u \geqq 1} \lambda_u^2 \phi_u, \tag{1.3}$$

where H is the mean curvature vector of M in E^m and the addition is also taken in the above L^2-sense. Let $u_1, u_2 \in \mathbb{N}$, $1 \leqq u_1 \leqq u_2$. We say that *the immersion* ϕ *is of order* $\{u_1, u_2\}$ if

$$\phi_u = 0 \quad \text{for all } u \in \mathbb{N}, u \neq 0, u_1 \quad \text{or} \quad u_2.$$

If $u_1 = u_2$, we say simply that *the immersion is of order* u_1. From a well-known Takahashi theorem [11] we have that the immersion is of order u, for some $u \geqq 1$, if and only if M is minimal in some sphere of E^m, and in this case we have the relation $\lambda_u = n/R^2$, where R is the radius of the sphere. We say that an *order* u *immersion* ϕ *is standard* if it is of the form $\phi = K(f_1, \ldots, f_N)$, where K is a real constant and $\{f_i\}$ is an orthonormal basis in V_u.

Now we study an order 2 immersion of the sphere (it is the announced *organized* product). We consider on E^m the metric $\langle x, y \rangle = xy^t$, for all x, y of E^m, where $(\ )^t$ denote the transpose. Let $c > 0$. Then the sphere $S^m(c) = \left\{ x \in E^{m+1} / \langle x, x \rangle = \frac{1}{c} \right\}$ with the induced metric has constant sectional curvature c. The tangent space at any point x of $S^m(c)$ is given by $T_x(S^m(c)) = \{X \in E^{m+1} / \langle x, X \rangle = 0\}$.

Let D and $\bar{\nabla}$ be the Riemannian connections of E^{m+1} and $S^m(c)$ respectively. Then for any two tangent vector fields to $S^m(c)$, X, Y, we have $D_X Y = \bar{\nabla}_X Y - c\langle X, Y \rangle x$. Let $SM(m) = \{P \in gl(m, \mathbb{R}) / P = P^t\}$ the space of the symmetric matrices of order m. We define on $SM(m)$ the metric $g(P, Q) = \frac{c}{2} trPQ$, for all P, Q in $SM(m)$. We consider the mapping $f\colon S^m(c) \to SM(m+1)$ given by $f(x) = x^t x$. The differential of f is $f_*(X) = x^t X + X^t x$, for all X in $T_x(S^m(c))$, and so f is an isometric immersion. Let $X \in T_x(S^m(c))$ and $P \in SM(m+1)$. From the relation $g(f_*(X), P) = c\langle X, xP \rangle$ we have that the normal space for the immersion f at any point x of $S^m(c)$ is given by $T_x^\perp(S^m(c)) = \{P \in SM(m+1) / xP = \lambda x$, for some real $\lambda\}$. We denote also by D the Riemannian connection of $SM(m+1)$. For any two tangent vector fields to $S^m(c)$, X, Y, we have

$$\begin{aligned} D_{f_*(X)} f_*(Y) &= (D_X Y)^t x + Y^t X + X^t Y + x^t (D_X Y) \\ &= X^t Y + Y^t X - 2c\langle X, Y \rangle\, x^t x + (\bar{\nabla}_X Y)^t x + x^t (\bar{\nabla}_X Y) \\ &= X^t Y + Y^t X - 2c\langle X, Y \rangle f(x) + f_*(\bar{\nabla}_X Y). \end{aligned}$$

Hence if $\bar{\sigma}$ denotes the second fundamental form of the immersion f, we have

$$\bar{\sigma}(X, Y) = X^t Y + Y^t X - 2c\langle X, Y \rangle f(x), \tag{1.4}$$

for all X, Y in $T_x(S^m(c))$.

By similar elementary computation we obtain

$$\bar{\nabla} \bar{\sigma} = 0. \tag{1.5}$$

that is, the second fundamental form is parallel,

$$g(\bar{\sigma}(X, Y), \bar{\sigma}(V, W)) = 2c\langle X, Y \rangle \langle V, W \rangle + c\langle X, V \rangle \langle Y, W \rangle + c\langle X, W \rangle \langle Y, V \rangle, \tag{1.6}$$

for all X, Y, V, W tangent to $S^m(c)$ at x, and hence

$$\bar{A}_{\bar{\sigma}(X,Y)} V = 2c\langle X, Y\rangle V + c\langle X, V\rangle Y + c\langle Y, V\rangle X, \tag{1.7}$$

$\bar{A}$ being the Weingarten endomorphism of f. The normal connection of f is denoted by $\bar{\nabla}^{\perp}$. We also remark the relations

$$g(f(x), \bar{\sigma}(X, Y)) = -\langle X, Y\rangle, \quad g(I, \bar{\sigma}(X, Y)) = 0, \tag{1.8}$$

where I is the identity matrix in $SM(m+1)$. Finally if Δ is the Laplacian of the sphere we have

$$\Delta f(x) = 2c(m+1) f(x) - 2I. \tag{1.9}$$

Hence f is an order 2 immersion of the sphere. For more details on this immersion see K. Sakamoto [9].

2. Minimal Submanifolds in the Sphere

Let ψ: $M^n \to S^{n+p}(c)$ be a minimal isometric immersion of an n-dimensional Riemannian manifold into the sphere. We identify x and $\psi(x)$. Let $\{E_1, \ldots, E_n, \xi_{n+1}, \ldots, \xi_{n+p}\}$[1] be a local field of orthonormal frames of $S^{n+p}(c)$ such that, restricted to M, $E_1, \ldots, E_n$ are tangent to M. We denote by ∇ and $\nabla^{\perp}$ the Riemannian connection and the connection in the normal bundle of ψ, and by σ and A the second fundamental form and the Weingarten endomorphism of ψ respectively. We consider the associated isometric immersion $\phi = f \circ \psi$: $M^n \to SM(n+p+1)$.

Lemma 2.1. *We have the following relations*

$$\Delta \phi(x) = -\sum_i \bar{\sigma}(E_i, E_i)_x, \tag{2.1}$$

$$\begin{aligned} \Delta^2 \phi(x) = {} & 2(n+1)\, c \Delta \phi(x) + 2\sum_{ij} \bar{\sigma}(\sigma(E_i, E_j), \sigma(E_i, E_j))_x \\ & -2\sum_{ij} \bar{\sigma}(A_{\sigma(E_i, E_j)} E_i, E_j)_x, \end{aligned} \tag{2.2}$$

for any x in M.

Proof. Since M is minimal in $S^{n+p}(c)$ we have (2.1).

We compute the differential of the mapping $\Delta\phi$: $M^n \to SM(n+p+1)$.

$$\begin{aligned} (\Delta\phi)_*(E_j) &= -\sum_i D_{E_j} \bar{\sigma}(E_i, E_i) \\ &= \sum_i \bar{A}_{\bar{\sigma}(E_i, E_i)} E_j - 2\sum_i \bar{\sigma}(\bar{\nabla}_{E_j} E_i, E_i) \\ &= 2(n+1)\, c E_j - 2\sum_i \bar{\sigma}(\sigma(E_i, E_j), E_i) - 2\sum_i \bar{\sigma}(\nabla_{E_j} E_i, E_i), \end{aligned}$$

[1] For the range of the indices we use the following convention: $i, j, k, r, s = 1, \ldots, n \quad \alpha, \beta, \gamma, \delta = n+1, \ldots, n+p$.

where we have used (1.5) and (1.7). From (1.6) we obtain

$$g(\sum_i \bar{\sigma}(\nabla_{E_j} E_i, E_i), \bar{\sigma}(E_r, E_s)) = 0, \qquad \text{for all } j, r, s.$$

Hence $(\Delta\phi)_*(E_j) = 2(n+1)\,cE_j - 2\sum_i \bar{\sigma}(\sigma(E_i, E_j), E_i)$.

Let x be an arbitrary point in M. We may assume that $\nabla_{E_j} E_i = 0$ at x. We compute $\Delta^2\phi$ at x.

$$\begin{aligned}(\Delta^2\phi)(x) &= -\sum_j D_{E_j}(\Delta\phi)_*(E_j)\\
&= -2(n+1)\,c\sum_j D_{E_j}E_j + 2\sum_{ij} D_{E_j}\bar{\sigma}(\sigma(E_i, E_j), E_i)\\
&= 2(n+1)\,c(\Delta\phi)(x) - 2\sum_{ij} \bar{A}_{\bar{\sigma}(\sigma(E_i, E_j), E_i)}E_j\\
&\quad + 2\sum_{ij}\bar{\sigma}(\sigma(E_i, E_j), \sigma(E_i, E_j)) - 2\sum_{ij}\bar{\sigma}(A_{\sigma(E_i, E_j)}E_i, E_j)\\
&\quad + 2\sum_{ij}\bar{\sigma}((\nabla\sigma)_{E_j}(E_i, E_j), E_i).\end{aligned}$$

From Codazzi equation $\sum_i (\nabla\sigma)_{E_j}(E_i, E_j) = \sum_j (\nabla\sigma)_{E_i}(E_j, E_j) = 0$, because M is minimal. From (1.7) $\sum_{ij} \bar{A}_{\bar{\sigma}(\sigma(E_i, E_j), E_i)}E_j = 0$. Hence we have (2.2). (Q.E.D.)

We put $\sigma(E_i, E_j) = \sum_\alpha h_{ij}^\alpha \xi_\alpha$. Then we obtain

$$\begin{aligned}(\Delta^2\phi)(x) &= 2(n+1)\,c(\Delta\phi)(x) + 2\sum_{ij\alpha\beta} h_{ij}^\alpha h_{ij}^\beta \bar{\sigma}(\xi_\alpha, \xi_\beta)\\
&\quad - 2\sum_{ijk\alpha} h_{ij}^\alpha h_{ik}^\alpha \bar{\sigma}(E_j, E_k).\end{aligned} \tag{2.3}$$

We define the tensor T in the normal bundle of the immersion ψ by $T(\xi, \eta) = trA_\xi A_\eta$, for all ξ, $\eta \in T_x^\perp M$. If S is the Ricci tensor of M we have

$$\sum_{i\alpha} h_{ij}^\alpha h_{ik}^\alpha = (n-1)\,c\langle E_j, E_k\rangle - S(E_j, E_k). \tag{2.4}$$

Hence we obtain

$$\Delta^2\phi = 4nc\Delta\phi + 2\sum_{\alpha\beta} T(\xi_\alpha, \xi_\beta)\,\bar{\sigma}(\xi_\alpha, \xi_\beta) + 2\sum_{jk} S(E_j, E_k)\,\bar{\sigma}(E_j, E_k). \tag{2.5}$$

Lemma 2.2. *We have the relations*

$$g(\phi, \phi) = \frac{1}{2c}, \tag{2.6}$$

$$g(\phi, \Delta\phi) = n, \tag{2.7}$$

$$g(\Delta\phi, \Delta\phi) = 2(n+1)\,nc, \tag{2.8}$$

$$g(\Delta^2\phi, \Delta\phi) = 4(n+1)^2 nc^2 + 4c\,\|\sigma\|^2. \tag{2.9}$$

The proof can be obtained from (1.6) and (1.8). Now we characterize minimal submanifolds in the sphere for which there exist u_1, u_2 in $\mathbb{N}$ such that the product of any two coordinate functions belongs to $V_0+V_{u_1}+V_{u_2}$.

Theorem 2.3. *Let ψ: $M^n \to S^{n+p}(c)$ be an isometric immersion of a compact Riemannian manifold such that the immersion is full and minimal. Then the immersion $\phi=f\circ\psi$ is of order $\{u_1,u_2\}$ for some natural numbers $u_1,u_2\geqq 1$, if and only if M is Einstein and $T=k\langle\ ,\ \rangle$, where k is a real constant and $\langle\ ,\ \rangle$ is the metric restricted to the normal bundle of ψ.*

Proof. We suppose that ϕ is an immersion of order $\{u_1,u_2\}$, that is $\phi=\phi_0+\phi_{u_1}+\phi_{u_2}$. Then $\Delta\phi=\lambda_{u_1}\phi_{u_1}+\lambda_{u_2}\phi_{u_2}$ and $\Delta^2\phi=\lambda_{u_1}^2\phi_{u_1}+\lambda_{u_2}^2\phi_{u_2}$. Hence we obtain

$$\Delta^2\phi=(\lambda_{u_1}+\lambda_{u_2})\Delta\phi-\lambda_{u_1}\lambda_{u_2}(\phi-\phi_0). \tag{2.10}$$

We recall that ϕ_0 is a constant mapping. As ϕ, $\Delta\phi$ and $\Delta^2\phi$ are normal to $S^{n+p}(c)$ for the immersion f, we conclude that $\phi_0\in T^{\perp}_{\phi(x)}(S^{n+p}(c))$ for all x in M. So, for all x in M, there exist $\lambda\in\mathbb{R}$ such that $x\phi_0=\lambda x$. Since M is connected, $\psi(M)$ is contained in $S^{n+p}(c)\cap V$, where V is an eigenspace of the linear mapping $x\mapsto x\phi_0$ in E^{n+p+1}. As ψ is full we conclude that ϕ_0 is a scalar matrix. Let I be the identity matrix in $SM(n+p+1)$. Applying $g(I,-)$ to (2.10) we obtain $\phi_0=\{(n+p+1)c\}^{-1}I$. Hence

$$\Delta^2\phi=(\lambda_{u_1}+\lambda_{u_2})\Delta\phi-\lambda_{u_1}\lambda_{u_2}\left(\phi-\frac{1}{(n+p+1)c}I\right). \tag{2.11}$$

By a direct computation, using (1.6), (1.8), Lemma 2.1 and (2.5) we obtain

$$g\left(\bar\sigma(E_r,E_s),\phi-\frac{1}{(n+p+1)c}I\right)=-\langle E_r,E_s\rangle,$$
$$g(\bar\sigma(E_r,E_s),\Delta\phi)=-2(n+1)c\langle E_r,E_s\rangle,$$
$$g(\bar\sigma(E_r,E_s),\Delta^2\phi)=-4n(n+3)c^2\langle E_r,E_s\rangle+4cS(E_r,E_s).$$

From (2.11) and the above relations we conclude that

$$S(E_r,E_s)=\frac{1}{4c}\{\lambda_{u_1}\lambda_{u_2}-2(n+1)c(\lambda_{u_1}+\lambda_{u_2})+4n(n+3)c^2\}\langle E_r,E_s\rangle. \tag{2.12}$$

So, M is Einstein. In the same way

$$g\left(\bar\sigma(\xi_\gamma,\xi_\delta),\phi-\frac{1}{(n+p+1)c}I\right)=-\langle\xi_\gamma,\xi_\delta\rangle,$$
$$g(\bar\sigma(\xi_\gamma,\xi_\delta),\Delta\phi)=-2nc\langle\xi_\gamma,\xi_\delta\rangle,$$
$$g(\bar\sigma(\xi_\gamma,\xi_\delta),\Delta^2\phi)=-4n(n+1)c^2\langle\xi_\gamma,\xi_\delta\rangle+4cT(\xi_\gamma,\xi_\delta).$$

Hence as above we obtain

$$T(\xi_\gamma,\xi_\delta)=\frac{1}{4c}\{\lambda_{u_1}\lambda_{u_2}-2nc(\lambda_{u_1}+\lambda_{u_2})+4n(n+1)c^2\}\langle\xi_\gamma,\xi_\delta\rangle. \tag{2.13}$$

Conversely, we suppose that M^n is Einstein and that $T=k\langle\ ,\ \rangle$. Then as $S(E_i,E_j)=\frac{\rho}{n}\langle E_i,E_j\rangle$ and $T(\xi_\gamma,\xi_\delta)=\frac{\|\sigma\|^2}{p}\langle\xi_\gamma,\xi_\delta\rangle$, where ρ is the scalar curvature of M, we have from (2.5)

$$\begin{aligned}\Delta^2\phi &= 4nc\Delta\phi+\frac{2\|\sigma\|^2}{p}\sum_\alpha\bar\sigma(\xi_\alpha,\xi_\alpha)+\frac{2\rho}{n}\sum_j\bar\sigma(E_j,E_j)\\ &= 4nc\Delta\phi+\frac{2\|\sigma\|^2}{p}\Big(\sum_\alpha\bar\sigma(\xi_\alpha,\xi_\alpha)+\sum_j\bar\sigma(E_j,E_j)\Big)\\ &\quad+\left(\frac{2\rho}{n}-\frac{2\|\sigma\|^2}{p}\right)\sum_j\bar\sigma(E_j,E_j).\end{aligned}$$

Finally from (1.9) we obtain

$$\Delta^2\phi=\left(4nc-\frac{2\rho}{n}+\frac{2\|\sigma\|^2}{p}\right)\Delta\phi-\frac{4\|\sigma\|^2}{p}(n+p+1)\,c\left(\phi-\frac{1}{(n+p+1)\,c}I\right),$$

and we conclude that the immersion ϕ is of order $(u_1,u_2\}$ for some $u_1,u_2\geqq 1$. (Q.E.D)

3. Eigenvalue Inequalities

Let M^n be a compact Riemannian manifold of dimension n and ϕ: $M^n\to E^m$ and isometric immersion in the Euclidean space. Recall the decomposition $\phi=\sum_{u\geqq 0}\phi_u$. We put

$$\int_M\langle\phi_u,\phi_u\rangle=a_u\qquad\text{for all } u\in\mathbb{N}.$$

The from (1.1), (1.2) and (1.3) we have

$$\begin{aligned}\int_M\langle\Delta\phi,\phi\rangle&=\sum_{u\geqq 1}\lambda_u a_u,\\ \int_M\langle\Delta\phi,\Delta\phi\rangle&=\sum_{u\geqq 1}\lambda_u^2 a_u,\\ \int_M\langle\Delta\phi,\Delta^2\phi\rangle&=\sum_{u\geqq 1}\lambda_u^3 a_u.\end{aligned}$$

We put

$$\Xi=\int_M\langle\Delta\phi,\Delta\phi\rangle-\lambda_1\int_M\langle\Delta\phi,\phi\rangle,\quad\text{and}$$

$$\Omega=\int_M\langle\Delta\phi,\Delta^2\phi\rangle-\lambda_1\int_M\langle\Delta\phi,\Delta\phi\rangle.$$

Then from the above relations we obtain

$$\Xi=\sum_{u\geqq 2}\lambda_u(\lambda_u-\lambda_1)\,a_u\geqq 0, \tag{3.1}$$

$$\Omega = \sum_{u \geqq 2} \lambda_u^2 (\lambda_u - \lambda_1)\, a_u \geqq 0, \tag{3.2}$$

$$\Omega - \lambda_2 \Xi = \sum_{u \geqq 3} \lambda_u (\lambda_u - \lambda_1)(\lambda_u - \lambda_2)\, a_u \geqq 0. \tag{3.3}$$

The equality in (3.1) holds if and only if the immersion is of order 1. For the equality in (3.2) we have the same condition. The equality in (3.3) holds if and only if the immersion is of order $\{1,2\}$. Hence for minimal submanifolds in the sphere we have

Theorem 3.1. *Let M^n be an isometrically immersed compact minimal submanifold of $S^m(c)$. Then*

$$\{n[2(n+1)c - \lambda_1][2(n+1)c - \lambda_2] + 4n(n-1)c^2\} \operatorname{vol}(M) \geqq 4c \int_M \rho, \tag{3.4}$$

λ_1 and λ_2 being the first and the second eigenvalues of the Laplacian of M, $\operatorname{vol}(M)$ the volume of M and ρ the scalar curvature of M. If the equality holds then M is Einstein and (if the immersion is full) $T = k\langle\ ,\ \rangle$.

Proof. From Lemma 2.2, (3.3) and the relation $\rho = n(n+1) - \|\sigma\|^2$, we obtain (3.4). If the equality holds we consider the smallest totally geodesic submanifold of $S^m(c)$ which contains $\psi(M)$, and we conclude the proof from Theorem 2.3. (Q.E.D.)

Corollary 3.2. *Let M^n be a compact Riemannian manifold which admits an order 1 immersion in the Euclidean space. Then*

$$\frac{n}{4}\{2(n+5)\lambda_1 - (n+2)\lambda_2\} \operatorname{vol}(M) \geqq \int_M \rho. \tag{3.5}$$

The proof is obtained from (3.4) and the relation $c = \lambda_1/n$.

Corollary 3.3. *Let M^n be a compact Riemannian manifold which admits a standard order 1 immersion. Then*

$$\frac{n+2}{n} \cdot \frac{2m_1}{m_1+1} \lambda_1 \geqq \lambda_2, \tag{3.6}$$

where m_1 is the dimension of the λ_1-eigenspace.

Proof. Let $\psi\colon M^n \to S^{m_1-1}(\lambda_1/n)$ be the corresponding minimal immersion. If ψ is standard then the center of gravity of the associated immersion $\phi = f \circ \psi$ is a scalar matrix, hence $\phi_0 = \dfrac{n}{m_1 \lambda_1} I$. Similar to (3.3) we have

$$\begin{aligned} &\int_M \langle \Delta\phi, \Delta\phi\rangle - (\lambda_1 + \lambda_2) \int_M \left\langle \Delta\phi, \phi - \frac{n}{m_1\lambda_1} I \right\rangle + \lambda_1\lambda_2 \int_M \left\langle \phi - \frac{n}{m_1\lambda_1} I, \phi - \frac{n}{m_1\lambda_1} I \right\rangle \\ &\quad = \sum_{u \geqq 3} (\lambda_u - \lambda_1)(\lambda_u - \lambda_2)\, a_u \geqq 0, \end{aligned} \tag{3.7}$$

the equality holding if and only if ϕ is an immersion of order $\{1, 2\}$. (Q.E.D.)

Remark. The center of gravity of ϕ is essentially the matricial expression of the L^2 metric restricted to the linear functions of the immersion ψ with respect to the basis $\{x_i\}$, where x_i is the $i-th$ coordinate function of ψ.

Remark. If the equality holds in (3.5) or (3.6) we have the same conclusion as in Theorem 3.1. Two remarkable classes of Riemannian manifolds which admit a standard order 1 immersion are compact homogeneous irreducible spaces (see P. Li [5]) and strongly harmonic manifolds (see [3], p. 174 and K. Sakamoto [10]). For compact symmetric spaces of rank one equality in (3.5) and (3.6) holds.

Corollary 3.4. *Let M^n be a strongly harmonic manifold. Then*

$$\frac{n}{4}\{2(n+5)\lambda_1-(n+2)\lambda_2\}\geqq\rho, \tag{3.8}$$

$$\frac{n+2}{n}\,\frac{2m_1}{m_1+1}\lambda_1\geqq\lambda_2. \tag{3.9}$$

The equality in (3.8) *or* (3.9) *holds if and only if* M *is a compact symmetric space of rank one.*

Proof. From the above remark and because strongly harmonic manifolds are Einstein we have (3.8) and (3.9). Moreover the corresponding minimal immersion $\psi\colon M^n\to S^{m_1-1}(\lambda_1/n)$ is isotropic (see [10]). If some of the equalities holds, $T=k\langle\ ,\ \rangle$. From [6] p. 517, we conclude that M is locally symmetric. (Q.E.D.)

4. The First Eigenvalue of $P_{2\pi}$-manifolds

For the angular parametrization of S^1 the Laplacian is given by $-\dfrac{d^2}{dt^2}$ and its first eigenvalue is 1. Hence we have the following version of the Wirtinger's inequality: for any function $\psi\in C^\infty(S^1)$

$$\int_0^{2\pi}\left(\frac{d^2}{dt^2}\psi\right)^2 dt+\int_0^{2\pi}\left(\frac{d^2}{dt^2}\psi\right)\psi\,dt\geqq 0. \tag{4.1}$$

As $V_1(S^1)$ is generated by $\cos t$, $\sin t$, the equality in (4.1) holds if and only if $\psi(t)=A\cos t+B\sin t+C$, where A, B and C are real constants.

Let M be an n-dimensional $P_{2\pi}$-manifold, $\Pi\colon UM\to M$ and UM_x the unit tangent bundle and its fiber over $x\in M$ respectively. For any $u\in UM$ let γ_u be the geodesic in M given by the initial conditions $\gamma_u(0)=\Pi(u)$, $\dfrac{d}{dt}\gamma_u(0)=u$. The geodesic flow on UM is denoted by ζ^t and the canonical measures on M, UM and UM_x by dx, $d\mu$ and du respectively.

We begin by the following simple

Lemma 4.1. *Let $B\colon E^n x E^n\to\mathbb{R}$ be a symmetric bilinear form, S^{n-1} the sphere of radius 1 in E^n and du the canonical measure on S^{n-1}. Then*

$$\int_{S^{n-1}} B(u,u)\,du = \frac{\mathrm{vol}(S^{n-1})}{n} \sum_{i=1}^{n} B(e_i, e_i), \tag{4.2}$$

$$\int_{S^{n-1}} B(u,u)\,B(u,u)\,du = \frac{\mathrm{vol}(S^{n-1})}{n(n+2)} \left\{ \sum_{i,j=1}^{n} B(e_i, e_i)\,B(e_j, e_j) + 2 \sum_{i,j=1}^{n} B(e_i, e_j)\,B(e_i, e_j) \right\}, \tag{4.3}$$

where $\mathrm{vol}(S^{n-1})$ *is the volume of* S^{n-1} *and* $\{e_i\}_{i=1,\dots,n}$ *an orthonormal basis of* E^n.

Proof. Let f_1, f_2: $S^{n-1} \to \mathbb{R}$ be functions given by $f_1(u) = B(u,u)$, and $f_2(u) = B(u,u)\,B(u,u)$. Fix $u \in S^{n-1}$ and let $\{e_i\}_{i=1,\dots,n}$ be an orthonormal basis of E^n such that $e_1 = u$. By a direct calculation we have

$$(\Delta f_1)(u) = 2n f_1(u) - 2 \sum_{i=1}^{n} B(e_i, e_i),$$

$$(\Delta f_2)(u) = 4(n+2) f_2(u) - 4B(u,u) \sum_{i=1}^{n} B(e_i, e_i) - 8 \sum_{i=1}^{n} B(u, e_i)\,B(u, e_i),$$

where Δ is the Laplacian of the sphere. Integrating the above relations on S^{n-1} we conclude the proof. (Q.E.D).

Theorem 4.2. *Let* M *be a* n*-dimensional* $P_{2\pi}$*-manifold and suppose that the Ricci tensor,* S*, and the metric,* $\langle\ ,\ \rangle$*, on* M *verify the relation* $S \geqq k \langle\ ,\ \rangle$*, where* k *is a real constant. Let* λ_1 *be the first eigenvalue of the Laplacian of* M*. Then we have*

$$\lambda_1 \geqq \tfrac{1}{3}(2k + n + 2). \tag{4.4}$$

Proof. Let f: $M \to \mathbb{R}$ be any differentiable function on M. For any u in UM we have from (4.1) and because γ_u admits 2π as a period,

$$\int_0^{2\pi} \left(\frac{d^2}{dt^2} f \circ \gamma_u(t)\right)^2 dt + \int_0^{2\pi} \left(\frac{d^2}{dt^2} f \circ \gamma_u(t)\right) f \circ \gamma_u(t)\, dt \geqq 0. \tag{4.5}$$

Following the Berger idea (integrating the above inequality on the set of all the geodesics) we have

$$\int_{UM} \int_0^{2\pi} \left(\frac{d^2}{dt^2} f \circ \gamma_u(t)\right)^2 dt\, d\mu + \int_{UM} \int_0^{2\pi} \left(\frac{d^2}{dt^2} f \circ \gamma_u(t)\right) f \circ \gamma_u(t)\, dt\, d\mu \geqq 0. \tag{4.6}$$

If we denote by df and Hess f the differential and the Hessian of f we have

$$(df)(u) = \frac{d}{dt} f \circ \gamma_u(0), \qquad \mathrm{Hess}\, f(u,u) = \frac{d^2}{dt^2} f \circ \gamma_u(0).$$

In the following calculation we only use standard arguments: The Fubini theorem and the fact that $d\mu$ is invariant respect to the geodesic flow ζ^t.

$$\int_{UM}\int_0^{2\pi}\left(\frac{d^2}{dt^2}f\circ\gamma_u(t)\right)^2 dt\,d\mu=\int_{UM}\int_0^{2\pi}\operatorname{Hess}f(\zeta^t u,\zeta^t u)^2\,dt\,d\mu$$
$$=\int_0^{2\pi}\int_{UM}\operatorname{Hess}f(u,u)^2\,d\mu\,dt=2\pi\int_{x\in M}\left(\int_{UM_x}\operatorname{Hess}f_x(u,u)^2\,du\right)dx.$$

Hence, by (4.3) we obtain

$$\int_{UM}\int_0^{2\pi}\left(\frac{d^2}{dt^2}f\circ\gamma_u(t)\right)^2 dt\,d\mu=\frac{2\pi\operatorname{vol}(S^{n-1})}{n(n+2)}\int_M\{2\|\operatorname{Hess}f\|^2+(\Delta f)^2\}\,dx.$$

On the other hand

$$\int_{UM}\int_0^{2\pi}\left(\frac{d^2}{dt^2}f\circ\gamma_u(t)\right)f\circ\gamma_u(t)\,dt\,d\mu=\int_{UM}\int_0^{2\pi}\operatorname{Hess}f(\zeta^t u,\zeta^t u)f\circ\Pi(\zeta^t u)\,dt\,d\mu$$
$$=\int_0^{2\pi}\int_{UM}\operatorname{Hess}f(u,u)f\circ\Pi(u)\,d\mu\,dt=2\pi\int_{x\in M}\left(\int_{UM_x}\operatorname{Hess}f_x(u,u)\,du\right)f(x)\,dx$$
$$=-\frac{2\pi\operatorname{vol}(S^{n-1})}{n}\int_M(\Delta f)f\,dx,$$

where we have used (4.2). From (4.6) we conclude that for any differentiable f on M we have

$$\int_M\{2\|\operatorname{Hess}f\|^2+(\Delta f)^2-(n+2)(\Delta f)f\}\,dx\geqq 0,\tag{4.7}$$

and the equality holds if and only if $f\circ\gamma_u(t)=A_u\cos t+B_u\sin t+C_u$, for any u in UM.

From the Bochner-Lichnerowicz formula (see [2], p. 131) we have

$$\int_M\|\operatorname{Hess}f\|^2\,dx=\int_M\{(\Delta f)^2-S(\operatorname{grad}f,\operatorname{grad}f)\}\,dx,\tag{4.8}$$

where grad f denotes the gradient of f. So, By using (4.7), (4.8) and the hypothesis $S\geqq k\langle\ ,\ \rangle$ we obtain

$$\int_M\{3(\Delta f)^2-(2k+n+2)(\Delta f)f\}\,dx\geqq 0,\qquad\text{for all } f\in C^\infty(M).\tag{4.9}$$

Now if we take f a non zero λ_1-eigenfunction we conclude that

$$\{3\lambda_1^2-(2k+n+2)\lambda_1\}\int_M f^2\,dx\geqq 0,\tag{4.10}$$

and we obtain (4.4). (Q.E.D.)

Remark. For compact symmetric spaces of rank one equality in (4.4) holds. If the equality holds for some $P_{2\pi}$-manifold M, then for any λ_1-eigenfunction f on M and for any u in UM we have $f\circ\gamma_u(t)=A_u\cos t+B_u\sin t+C_u$. Taking into account the Obata theorem (see [2], p. 179) it is natural to expect that this holds only for a little class of manifolds.

References

1. Berger, M.: Les premières valeurs propres des variétés riemanniennes. Compositio Mathematica **26**, 129–149 (1973)
2. Berger, M., Gauduchon, P., Mazet, E.: Le spectre d'une variété riemannienne. Lecture Notes in Math. **194**. Berlin Heidelberg New York: 1971
3. Besse, A.: Manifolds all of whose geodesic are closed. Ergebnisse der Mathematik, **93**, Berlin Heidelberg New York: Springer 1978
4. Cheng, S.T.: Eigenvalue comparison theorems and its geometric applications. Math. Z. **143**, 289–297 (1975)
5. Li, P.: Minimal immersions of compact irreducible homogeneous Riemannian manifolds. J. Diff. Geom. **16**, 105–115 (1981)
6. Nakagawa, H.: A characterization of the 3rd standard immersions of spheres into a sphere. J. Diff. Geom. **16**, 511–527 (1981)
7. Ros, A.: Spectral geometry of CR-minimal submanifolds in the complex projective space. Kodai Math. J. **6**, 88–99 (1983)
8. Ros, A.: On spectral geometry of Kaehler submanifolds. J. Math. Soc. Japan (to appear)
9. Sakamoto, K.: Planar geodesic immersions. Tôhoku Math. J. **29**, 25–56 (1977)
10. Sakamoto, K.: Helical immersions into a unit sphere. Math. Ann. 63–80 (1982)
11. Takahashi, T.: Minimal immersions of Riemannian manifolds. J. Math. Soc. Japan, **18**, 380–385 (1966)
12. Yang, P., Yau, S.T.: Eigenvalues of the Laplacian of compact Riemann surfaces and minimal submanifolds. Ann. della Scuola Sup. di Pisa **7**, 55–63 (1980)

Received August 8, 1983; received in final form April 21, 1984

Math. Z. 187, 405-423 (1984)

Mathematische Zeitschrift

Ample Divisors on Fine Moduli Spaces on the Projective Plane

Stein Arild Strømme

Department of Mathematics, Bergen University, N-5014 Bergen, Norway

Contents

0. Introduction

Let $M=M(c_1,c_2)$ be the moduli space for semistable rank-2 coherent sheaves on the projective plane P with Chern classes c_1 and c_2. The existence and projectivity of M was proved by Gieseker [5, Thm. 0.3]. The open subvariety $M^0 \subseteq M$ parametrizing locally free sheaves has been studied by several authors, notably Barth, whose papers [1] and [2] provided many basic tools for dealing with rank-2 vector bundles as well as proof that M^0 is irreducible, nonsingular and rational if c_1 is even. Continuing in the same spirit, Hulek [9] verified the same assertions in the case of odd c_1. This case was also treated by Ellingsrud and the author [3] using a different method. Since that paper was never published, a souped-up version is included here, in Sect. 3 below.

Concerning properties of a more global nature of M (and M^0) we have the following results: First of all, M is irreducible and normal, and if $c_1^2-4c_2 \not\equiv 0$ (mod 8), then M is nonsingular and carries a universal family. This is due to Maruyama [13, Thm. 7.17]. Then there are two results of Le Potier: He computes $\mathrm{Pic}(M^0)$ in the case $c_1=0$ [12] and he proves that M^0 does not carry a universal family if both c_1 and $c_2-\frac{1}{4}c_1^2$ are even [11]. Finally, Ellingsrud and the author computed $\mathrm{Pic}(M^0)$ in the case $c_1=-1$ [4], using the results of [3].

The purpose of the present paper is to describe the Picard group of the projective variety M, and in particular to identify the ample linebundles. Unfortunately, we are, so far, only able to do this in the case c_1 or $c_2-\frac{1}{4}c_1^2$ odd, as the method relies heavily on the existence of a universal sheaf. Bringing together (6.1), (5.3), (5.6), (5.7), (4.6), and (7.9), we have the following

Theorem. *Assume that $c_1=-1$ (resp. $c_1=0$) and that $n=c_2\geqq 2$ (resp. $n\geqq 3$ is odd). Then there exist elements ε and δ (resp. φ and ψ) of* Pic(M) *such that:*

(i) Pic(M) *is freely generated by ε and δ (resp. φ and ψ).*

(ii) *An integral linear combination $a\varepsilon+b\delta$ (resp. $a\varphi+b\psi$) is ample if and only if $a>0$ and $b>0$.*

(iii) *$M-M^0$ is the support of a reduced and irreducible divisor in the linear system $|n\varepsilon-2\delta|$ (resp. $|(2n-1)\varphi-\psi|$).*

(iv) *The canonical class $c_1(\Omega_M)$ is -3ε (resp. -6φ).*

(v) *Consider the following subsets of M:*
$D_1=\{$sheaves with a given jumping conic (resp. line)$\}$
$D_2=\{$sheaves with a jumping line (resp. conic) passing through 1 (resp. 3) given points$\}$
Then D_1 is the support of a reduced divisor in the linear system $|\varepsilon|$ (resp. $|\varphi|$), and D_2 is the support of a reduced divisor in the linear system $|\delta|$ (resp. $|\frac{1}{2}(n-1)\psi|$).

The proof has several ingredients, of which the following stand out to me as important:

1) The idea of using a stratification for computing Pic(M^0), and the specific results of Sect. 3. This is joint work with G. Ellingsrud, and I want to express thanks for his permission to publish those results in this form.

2) The explicit description given by Gieseker [5] of M as a quotient of a "Gieseker space" by the action of a projective linear group.

3) The existence of a universal sheaf and a monad for the universal sheaf.

Notation and Conventions. For a coherent sheaf E on a proper k-scheme X, $H^i(E)$ is short for $H^i(X,E)$, $h^i(E)=\dim_k H^i(E)$, and $\chi(E)=\sum_{i\geqq 0}(-1)^i h^i(E)$. The structure sheaf $\mathcal{O}_X$ is sometimes written simply $\mathcal{O}$, when the scheme X is clear from the context. Lower-case greek letters usually denote divisor classes, or elements of some Picard group. If $\xi\in \mathrm{Pic}(X)$, then $\mathcal{O}_X(\xi)$ is the corresponding linebundle, and if E is a sheaf on X, then $E(\xi)=E\otimes\mathcal{O}(\xi)$. If $f\colon Y\to X$ is a morphism, we shall often write $f^*\mathcal{O}_X(\xi)=\mathcal{O}_Y(\xi)$, hence identifying $\xi\in\mathrm{Pic}(X)$ with its image $f^*(\xi)\in\mathrm{Pic}(Y)$.

The following symbols have a fixed meaning throughout most of the paper:

P = projective plane over an algebraically closed field k (of any characteristic).
$x=0$ or 1.
n = positive integer. Starting from Sect. 2, if $x=0$, then $n=2m+1$ for some $m\geqq 1$.
$M=M(-x,n)$ = moduli space for stable rank-2 sheaves on P with $c_1=-x$; $c_2=n$.
$M^0\subseteq M$ open subscheme corresponding to locally free sheaves
$\mathscr{E}$ = universal family on $M\times P$

$\pi\colon M\times P\to M$ the first projection
$\tau=$positive generator of $\mathrm{Pic}(P)\simeq\mathbb{Z}$
$A=R^1\pi_*(\mathscr{E}(-2\tau))$
$B=R^1\pi_*(\mathscr{E}(-\tau))$
$C=R^1\pi_*(\mathscr{E})$
$\alpha=c_1(A)\in\mathrm{Pic}(M)$
$\beta=c_1(B)\in\mathrm{Pic}(M)$
$\gamma=c_1(C)\in\mathrm{Pic}(M)$

$$\left.\begin{array}{l}\varepsilon=\gamma-\alpha\\ \delta=n\gamma-(n-1)\beta\end{array}\right\}\text{if } x=1$$

$$\left.\begin{array}{l}\varphi=\beta-\alpha\\ \psi=n\gamma-(n-2)\alpha\end{array}\right\}\text{if } x=0$$

Acknowledgement. I thank the referee for pointing out some errors in the original manuscript.

1. Properties of Stable Rank-2 Sheaves

(1.1) **Definition** (Gieseker). A rank-2 coherent sheaf E on P is *stable* (resp. *semistable*) if it is torsion free and if for all rank-1 subsheaves $F\subseteq E$, the following inequality holds for all sufficiently large integers k:

$$2\chi(F(k))<\chi(E(k))\quad(\text{resp. }\leqq).$$

(1.2) *Remarks.* (i) if $c_1(E)^2-4c_2(E)\not\equiv 0 \pmod 8$, then E is stable if (and only if) it is semistable.

(ii) If E is stable, then $h^0(\mathscr{H}om(E,E))=1$. If E is locally free, the converse also holds. If E is semistable, then $\mathrm{Hom}(E,E(-1))=0$.

(iii) If $E'\subseteq E$ is a rank-2 subsheaf such that the length of E/E' is 1, then E stable implies E' semistable, and E' stable implies E semistable. If $c_1(E)$ is odd, then E is stable if and only if $E^{\vee\vee}$ is stable. If $c_1(E)$ is even, then $E^{\vee\vee}$ stable implies E stable, and E semistable implies $E^{\vee\vee}$ "Mumford-Takemoto semistable", i.e. $H^0(E^{\vee\vee}(-1-\frac{1}{2}c_1(E)))=0$. The only bundle (locally free sheaf) which is semistable but not stable is $2\mathcal{O}_P$ and its twists.

(1.3) Fix two integers x and n, such that $x=1$ or $x=0$. Throughout the rest of this section, let E be a stable rank-2 sheaf on P such that $c_1(E)=-x$ and $c_2(E)=n$, or equivalently, with Hilbert polynomial $\chi(E(k))=(k+1)(k+2-x)-n$. Since P is regular of dimension 2, $E^{\vee\vee}$ is locally free and $pd(E)\leqq 1$, and $E^{\vee\vee}/E$ has finite length $=n-c_2(E^{\vee\vee})$. The definition of stability implies that $h^0(E)=0$, and by (1.2(iii)), $h^0(E^{\vee\vee}(x-1))=0$.

(1.4) **Lemma** (Serre duality). *For each integer k, there is a natural isomorphism*

$$H^0(E^\vee(-k-3))\xrightarrow{\ \sim\ } H^2(E(k))^\vee.$$

(1.5) **Proposition.**

(i) $h^0(E(k))=h^2(E(k))=0\quad(-2\leqq k\leqq 0)$

(ii) $h^1(E(-2))=n-x$
$h^1(E(-1))=n$
$h^1(E)=n-2+x.$

Proof. (i) follows from (1.3) and (1.4), taking into account the isomorphism $E^\vee \simeq E^{\vee\vee}(x)$. (ii) follows from the Riemann-Roch formula (1.3).

(1.6) **Proposition.** *E is the middle homology sheaf of a complex* (monad) *of the form*

$$0 \to H^1(E(-2)) \otimes \mathcal{O}(-1) \to H^1(E(-1)) \otimes \Omega(1) \to H^1(E) \otimes \mathcal{O} \to 0.$$

Proof. This is a special case of a Beilinson spectral sequence, see e.g. [15, 2.3.2].

(1.7) **Definition.** A line L (resp. a conic Γ) $\subseteq P$ is *jumping* if $h^1(E_L(x-1)) \neq 0$ (resp. $h^1(E_\Gamma) \neq 0$).

(1.8) *Remarks.* (i) This definition of a jumping conic in the case $x=0$ is not the only possible one. Although it may not be standard terminology, it is nevertheless the convenient concept for our purposes, see Sect. 5.

(ii) In characteristic zero, the Grauert-Mülich theorem [1, Thm. 1] implies that the general line (resp. conic) is not jumping. Even in positive characteristic, this is so for a general sheaf E.

(iii) In any case, the locus of jumping lines (resp. conics) has codimension $\leqq 1+x$ (resp. $\leqq 3-2x$), with equality for general E. (This is most easily seen by considering "Hulsbergen bundles" [2, Sect. 5].)

(iv) If $x=1$ (resp. $x=0$), a line (resp. conic) containing a point where E is not locally free is not necessarily jumping, in contrast with the case of conics (resp. lines).

(1.9) **Proposition.** *The numerical sequence* $\{h^1(E(k))\}$, $k \geqq -1$, *decreases strictly to zero.*

Proof. Let t be the smallest integer such that $E^\vee(t)$ has a non-zero global section, and use it to construct an exact sequence on P

$$0 \to K(-t-x) \to E \to \mathcal{O}(t) \to Q \to 0$$

where $K^{\vee\vee} \simeq \mathcal{O}$, and Q has finite support. Restricting this to a general line $L \subseteq P$, we get

$$0 \to \mathcal{O}_L(-t-x) \to E_L \to \mathcal{O}_L(t) \to 0.$$

In particular, $h^1(E_L(k))=0$ for $k \geqq t+x-1$. Assume that for some $k \geqq t+x$, $h^1(E(k-1))=h^1(E(k))$. Then the restriction map $H^0(E(k) \to H^0(E_L(k))$ is surjective. For any $m \geqq 0$, consider the commutative diagram

$$\begin{array}{ccc} H^0(\mathcal{O}(m)) \otimes H^0(E(k)) & \to & H^0(E(k+m)) \\ \downarrow & & \downarrow \\ H^0(\mathcal{O}_L(m)) \otimes H^0(E_L(k)) & \to & H^0(E_L(k+m)). \end{array}$$

The left and bottom maps are surjective, hence so is the right vertical map. It follows that $h^1(E(k+m))=h^1(E(k+m-1))$ for all $m\geqq 0$. By Serre's theorem, this must vanish. We have shown that $h^1(E(k))$ decreases strictly to zero in the range $k\geqq t+x-1$. In the remaining range $-1\leqq k\leqq t+x-1$, we have $h^0(E(k))\leqq h^0(E^{\vee\vee}(k))=h^0(E^\vee(k-x))$, which is zero by the definition of t. Hence $h^1(E(k))=n-(k+1)(k+2-x)$ in this range, obviously strictly decreasing.

(1.10) **Corollary.** *$h^1(E(k))=0$ for $k\geqq n-2+x$.*

(1.11) *Examples.* (i) Let $Y\subseteq P$ be a group of n points on a line, and construct E as any non-trivial extension of I_Y by $\mathcal{O}(-1)$. Then $h^1(E(k))=n-1-k$ for $-1\leqq k\leqq n-1$, hence (1.9–10) are "sharp" in the case $x=1$. Note that the line L containing Y is the only jumping line, and that $E_L\simeq\mathcal{O}_L(-n)\oplus\mathcal{O}_L(n-1)$ if E is locally free.

(ii) Let $Y\subseteq P$ be the union of n points on a line L and a single point outside L, and construct E as a general extension of $I_Y(1)$ by $\mathcal{O}(-1)$. Then $h^1(E(k))=n-2-k$ for $0\leqq k\leqq n-2$, so (1.9–10) are "sharp" also in the case $x=0$. Note that $E_L\simeq\mathcal{O}_L(1-n)\oplus\mathcal{O}_L(n-1)$.

(1.12) **Proposition.** *For all rank-1 subsheaves $F\subseteq E$, the following inequalities hold:*

$$2h^0(F(k))\leqq\chi(E(k))\qquad \textit{for } k\geqq n-2+x.$$

The inequality is strict if $x=0$ and $n>2$, or if $k\geqq n$.

Proof. We may assume that $h^0(F(k))\neq 0$. Then $F^{\vee\vee}\simeq\mathcal{O}(-d)$ for some d, $x\leqq d\leqq k$. If $d\geqq 1$, it follows that

$$2h^0(F(k))\leqq 2h^0(\mathcal{O}(k-d))\leqq k(k+1)\leqq(k+1)(k+2-x)-n,$$

the last inequality being strict unless $x=0$, $n=2$, $k=0$ or $x=1$, $k=n-1$. If $d=0$, the length of $F^{\vee\vee}/F$ is at most n, hence $h^1(F)\leqq n-1$. By the same method as in the proof of (1.9), one shows that $h^1(F(n-2))\leqq 1$, and $h^1(F(k))=0$ for $k\geqq n-1$. Now the polynomial $\chi(E(k))-2\chi(F(k))$ is a constant, which is positive by stability of E. Hence the proposition is proved in case $h^1(F(n-2))=0$. If not, one concludes that F is the ideal of n points, so $\chi(F(k))=\frac{1}{2}(k+2)(k+1)-n$. But then

$$2(h^0(F(k))=2\chi(F(k))+2h^1(F(k))\leqq(k+2)(k+1)-2n+2\leqq(k+2)(k+1)-n,$$

the last inequality strict unless $n=2$, and the next to last inequality strict unless $k=n-2$.

(1.13) **Definition.** A (rank-2) *Gieseker space* is a projective space of the form $G=\mathbb{P}(\mathrm{Hom}(\Lambda^2 H, W)^\vee)$, together with the natural action of $PGL(H)$ on G. Here H and W are finite vectorspaces over k. A rational point T of G will be identified with a non-zero alternating bilinear map $T\colon H\times H\to W$, modulo non-zero constants. If T is such a point and $h\in H$ is a vector, $\langle h\rangle_T\subseteq H$ denotes the subspace of all $h'\in H$ such that $T(h,h')=0$. Gieseker [5, Prop. 2.2] gave the following sufficient condition for stability or semistability of a point $T\in G$ with respect to the action of $PGL(H)$:

(1.14) **Proposition.** *T is stable (resp. semistable) if for all non-zero $h \in H$, the following inequality holds:*

$$2\dim\langle h\rangle_T < \dim V \quad (\textit{resp.} \leqq).$$

(1.15) The *determinant* $\det(E)$ of E, is by definition the double dual of $\Lambda^2(E)$. For any integer $k \geqq n-2+x$, the natural map $E(k)\times E(k)\to\det(E(k))$ induces a map

$$T_k(E)\colon H^0(E(k))\times H^0(E(k))\to H^0(\det(E(k)))$$

which is bilinear and alternating. It has the property that if s_1 and s_2 are two sections of $E(k)$, not both zero, then $T_k(E)(s_1, s_2)=0$ if and only if the subsheaf of $E(k)$ generated by s_1 and s_2 has rank 1.

(1.16) **Proposition.** *$T_k(E)$ is semistable for $k \geqq n-2+x$, and stable if $k \geqq n$ or $x=0$, $n>2$.*

Proof. The following argument has been extracted from the proof of [5, Prop. 3.1]. In order to apply (1.14), let $h \in H^0(E(k))$ be a non-zero section, and let $F'(k) \subseteq E(k)$ be the subsheaf generated by h. Let F be the kernel of the map $E \to (E/F')/\text{torsion}$. Then F is maximal in the set of rank-1 subsheaves of E containing F'. If g is any element of $\langle h\rangle_{T_k(E)}$, then the subscheaf $F''(k) \subseteq E(k)$ generated by h and g has rank 1 by (1.15) above. Hence $F'' \subseteq F$ by the maximality of F. Thus $\langle h\rangle_{T_k(E)} \subseteq H^0(F(k))$, and (1.12), together with (1.10), concludes the proof.

(1.17) *Remark.* Gieseker's formulation of [5, Prop. 3.1] implies only that $T_k(E)$ is stable for $k \gg 0$.

2. The Universal Family

(2.1) From now on we shall only consider stable sheaves with $c_1^2-4c_2 \not\equiv 0$ (mod 8). Let x and n be two integers such that $x=0$ or 1, and, if $x=0$, $n=2m+1$ is odd. Denote by $M=M(-x, n)$ the moduli space for stable rank-2 sheaves on P with Chern classes $c_1=-x$ and $c_2=n$. By [13, Thm. 7.17], M is an irreducible, nonsingular, projective variety of dimension $4n-3-x$. Furthermore, M is a fine moduli space, i.e. there exists a universal sheaf on $P\times M$. Barth ($x=0$) and Hulek ($x=1$) has shown that M is rational [2, 9]; a different proof is suggested below (Sect. 3).

(2.2) A universal sheaf $\mathscr{E}$ on $M\times P$ is not unique; however, if $\mathscr{E}'$ is another, then $\mathscr{E}'$ is isomorphic to $\mathscr{E}\otimes\pi^*\mathscr{L}$, where $\mathscr{L}$ is the linebundle $\pi_*\mathscr{H}om(\mathscr{E}, \mathscr{E}')$, see (1.2(ii)). (Here, as later on, $\pi\colon M\times P\to M$ is the first projection.) Fix a universal sheaf $\mathscr{E}$.

(2.3) Put $A=R^1\pi_*\mathscr{E}(-2\tau)$, $B=R^1\pi_*\mathscr{E}(-\tau)$, and $C=R^1\pi_*\mathscr{E}$. In view of (1.5) and standard theory of base change [8, III, 12.11] these are locally free on M and commute with base change. The Beilinson spectral sequence is easily generalized to families, so $\mathscr{E}$ is the homology sheaf of a monad on $M\times P$:

$$0 \to \pi^* A(-\tau) \to \pi^* B \otimes pr_2^* \Omega(\tau) \to \pi^* C \to 0,$$

which induces the monad of (1.6) on each fiber.

(2.4) **Proposition.** *$pd(\mathscr{E}) \leqq 1$, and $\mathscr{E}$ is reflexive.*

Proof. That $pd(\mathscr{E}) \leqq 1$ is immediate from the monad. Furthermore, $\mathscr{E}$ is locally free off a subscheme of $M \times P$ of codimension $\geqq 3$, hence $\mathscr{E}$ is reflective.

(2.5) Let α, β, and $\gamma \in \operatorname{Pic}(M)$ be the first Chern classes of A, B, and C respectively. We shall eventually prove that these generate $\operatorname{Pic}(M)$. It will be convenient to have at our disposal an alternative set of generators which is independent of the particular universal sheaf $\mathscr{E}$ chosen. This is obtained as follows:

(2.6) If $x=1$, put $\varepsilon = \gamma - \alpha$ and $\delta = n\gamma - (n-1)\beta$. If $x=0$, put $\varphi = \beta - \alpha$ and $\psi = n\gamma - (n-2)\alpha$. Then ε, δ (resp. φ, ψ) are independent of the choice of $\mathscr{E}$, because of (2.2) and (1.5 (ii)). For suppose $\lambda \in \operatorname{Pic}(M)$ and $\mathscr{E}' = \mathscr{E}(\lambda)$. Let α', β', γ' be defined as in (2.5) with respect to $\mathscr{E}'$. Then

$$\alpha' = \alpha + (n-x)\lambda$$
$$\beta' = \beta + n\lambda$$
$$\gamma' = \gamma + (n-2+x)\lambda$$

and from this the conclusion follows.

(2.7) *Remark.* An alternative approach would be to normalize $\mathscr{E}$ so as to satisfy a certain condition. For $x=1$, the equations above show that λ can be chosen so that $\beta' = \gamma'$ (take $\lambda = \gamma - \beta$). For $x=0$, λ can be chosen such that $m\beta' = (m+1)\gamma'$ (where $2m+1=n$), take $\lambda = (m+1)\gamma - m\beta$. Hence we might as well assume that $\beta = \gamma$ (resp. $m\beta = (m+1)\gamma$) from the beginning. In particular, α, β, and γ can be expressed in terms of (ε, δ) (resp. φ, ψ).

(2.8) As an illustration of what these divisor classes may mean, we work out the following example of a 2-dimensional subfamily, which will be of use later.

(2.9) *Example.* Let D be a stable, locally free sheaf on P with $c_1(D) = -x$ and $c_2(D) = n-1$, and let $L \subseteq P$ be a line such that $D_L \simeq \mathcal{O}_L(n-2) \oplus \mathcal{O}_L(2-n-x)$, cf. (1.11). Let p: $F = \mathbb{P}(D_L) \to L$ be the associated ruled surface, with tautological quotient $D_F \to \mathcal{O}_F(\lambda)$. Let r: $F \to P$ be the composition of p and the inclusion of L in P, and let j: $F \to F \times P$ be the graph of r. Define a family $\mathscr{F}$ on $F \times P$ via the exact sequence

$$0 \to \mathscr{F} \to pr_2^* D \to j_* \mathcal{O}(\lambda) \to 0.$$

Then the family $\mathscr{F}$ defines a morphism s: $F \to M$. We want to compute (s^* of) ε and δ (resp. φ and ψ) in terms of λ and μ, where $\mu \in \operatorname{Pic}(L)$ is the positive generator (hence the class of a fiber of p: $F \to L$). Put $\alpha', \beta', \gamma' = c_1$ of $R^1\pi_*\mathscr{F}(-2\tau)$, $R^1\pi_*\mathscr{F}(-\tau)$, $R^1\pi_*\mathscr{F}$ respectively. (Here π: $F \times P \to F$ denotes the projection.) From the exact sequence above and the fact that $\pi_* j_* j^* \mathcal{O}(\tau) = j^* \mathcal{O}(\tau) = \mathcal{O}(\mu)$, we get

$$\alpha' = \lambda - 2\mu,$$
$$\beta' = \lambda - \mu,$$
$$\gamma' = \lambda.$$

Hence, by (2.6), we get

$$(x=1) \quad s^*(\varepsilon) = \gamma' - \alpha' = 2\mu$$
$$s^*(\delta) = n\gamma' - (n-1)\beta' = \lambda + (n-1)\mu,$$
$$(x=0) \quad s^*(\varphi) = \beta' - \alpha' = \mu$$
$$s^*(\psi) = n\gamma' - (n-2)\alpha' = 2\lambda + 2(n-2)\mu.$$

Let us draw some conclusions from all this:

(2.10) **Proposition.** *Assume* $n \geqq 2$.

(i) *If* $x=1$, *then* ε *and* δ *are linearly independent in* $\mathrm{Pic}(M)$.

(ii) *If* $x=0$, *then* φ *and* ψ *are linearly independent in* $\mathrm{Pic}(M)$.

Furthermore, let $\xi = a\varepsilon + b\delta$ *if* $x=1$ (*resp.* $\xi = a\varphi + b\psi$ *if* $x=0$). *Then*

(iii) *If some multiple of* ξ *is generated by its global sections, then* $a \geqq 0$ *and* $b \geqq 0$.

(iv) *If* ξ *is ample, then* $a > 0$ *and* $b > 0$.

Proof. (i) and (ii) are clear from (2.9), since λ and μ are independent generators of $\mathrm{Pic}(F)$. For (iii) and (iv), we shall use (for the first time) the hypothesis on the splitting type of D_L. The quotient $D_L \to \mathcal{O}_L(2-n-x)$ gives a section $S \subseteq F$ of the ruling p with self-intersection $S^2 = -(2n-4+x)$, and its class is $[S] = \lambda - (n-2)\mu$. The intersection numbers of ξ with S and a fiber of the ruling p are as follows (note that $\lambda^2 = -x$)

$$(x=1) \quad \xi \cdot [S] = 2a$$
$$\xi \cdot \mu \;\; = b,$$
$$(x=0) \quad \xi \cdot [S] = a$$
$$\xi \cdot \mu \;\; = 2b.$$

From this the conclusion of (2.10) follows immediately.

(2.11) **Lemma.** (i) $\det(\mathscr{E}) = \mathcal{O}(-x\tau + 2\beta - \alpha - \gamma)$.

(ii) $\mathscr{E}^\vee \simeq \mathscr{E}(x\tau - 2\beta + \alpha + \gamma)$.

Proof. (i) is clear from the monad (2.3).

(ii) follows from (i) and the isomorphism $\mathscr{E}^\vee \simeq \mathscr{E} \otimes \det(\mathscr{E})^{-1}$, valid for any reflexive rank-2 sheaf on a locally factorial integral scheme [7, Prop. 1.10].

3. Classification of Bundles

(3.1) In this section we review the main results of [3]. Since that paper was never published, we give an outline of the proofs here. Throughout this section.

fix a closed point y of P. Let Q be the projective line parametrizing lines of P containing y, let $F \subseteq P \times Q$ be the incidence correspondence, with projections p and g respectively. Let $\sigma \in \mathrm{Pic}(Q)$ be the positive generator. p: $F \to P$ is the blowing up of P in y, and the class of the exceptional fiber $R = p^{-1}(y)$ is $(\tau - \sigma)$. For any rank-2 bundle (locally free sheaf) E on P, put $\tilde{E} = p^* E$. Then $\tilde{E}_R$ is the trivial bundle $2\mathcal{O}_R$. Conversely, if D is a 2-bundle on F such that $D_R \simeq 2\mathcal{O}_R$, then $E = p_* D$ is a bundle and the natural map $\tilde{E} \to D$ is an isomorphism [16, Thm. 5]. So we are reduced to the classification of rank-2 bundles $\tilde{E}$ on F such that

(i) $\tilde{E}_R \simeq 2\mathcal{O}_R$.

(ii) $c_1(\tilde{E}) = -x\tau, \quad c_2(\tilde{E}) = n\tau^2$.

(iii) $h^0(\tilde{E}) = 0$.

We do the cases $x=1$ and $x=0$ separately. Let $M^0 \subseteq M(-x, n)$ be the open subvariety corresponding to locally free sheaves.

(3.2) Assume in (3.2–3.5) that $x=1$, and let E be a bundle in M^0. The *type* of E is by definition an ordered pair (i, k) of integers uniquely determined by the condition $g_* \tilde{E}(-k\tau) \simeq \mathcal{O}_Q(-i\sigma)$. Let $M^0_{(i,k)} \subseteq M$ be the (locally closed) subscheme (with reduced structure) parametrizing bundles of a given type (i, k).

(3.3) **Theorem.** (i) *$M^0_{(i,k)}$ is nonempty if and only if the following inequalities hold:*

$$(*) \qquad \begin{aligned} &k \geqq 0 \\ &i - k > 0 \\ &n - (i-k)(k+1) - ik \geqq 0. \end{aligned}$$

(ii) *If $(*)$ holds, then $M^0_{(i,k)}$ is irreducible, nonsingular and rational of codimension $n - i + k(2 + 2i - k)$ in M^0.*

(iii) *$M^0_{(n,0)}$ is dense in M.*

(iv) *$M^0_{(n-1,0)}$ is the only stratum of codimension 1.*

(v) *If $k > 0$, then the codimension of $M^0_{(i,k)}$ is at least $n+1$.*

Proof. (i) Suppose E is of type (i, k). If the splitting of E along a general fiber of g is of the form $\mathcal{O}(j) \oplus \mathcal{O}(-1-j)$, $j \geqq 0$, then clearly $k = j \geqq 0$. As for the second inequality, if $i \leqq k$, then $\tilde{E}(i\sigma - k\tau) \subseteq \tilde{E}(k(\sigma - \tau)) \subseteq \tilde{E}$, contradicting the fact that $h^0(\tilde{E}) = h^0(E) = 0$. Any non-zero section of $\tilde{E}(i\sigma - k\tau)$ induces a short exact sequence

$$0 \to \mathcal{O}(k\tau - i\sigma) \to \tilde{E} \to I_Y(i\sigma - (k+1)\tau) \to 0$$

where Y is a finite subscheme of F of length $c_2(\tilde{E}(i\sigma - k\tau)) = n - (i-k)(k+1) - ik$. Hence this number is non-negative. Conversely, let (i, k) be given, satisfying $(*)$. Let $Y \subseteq F$ be a group of $n - (i-k)(k+1) - ik$ general points, and construct $\tilde{E}$ as a general extension as above. It is easily verified that $E = p_* \tilde{E}$ is a bundle of type (i, k).

(ii) Note that the association $E \mapsto Y$ in the proof of (i) induces a morphism $M^0_{(i,k)} \to H_{(i,k)}$, where $H_{(i,k)}$ is the open part of the Hilbert scheme of F parametrizing locally complete intersection subschemes of the relevant finite length. By the existence part of the proof of (i), this morphism is dominating. Furthermore, all the fibers are open subsets of a projective space of constant dimension, in fact, there is a locally free sheaf $\mathscr{E}xt$ on $H_{(i,k)}$ and an open embedding $M^0_{(i,k)} \to \mathbb{P}(\mathscr{E}xt^\vee)$ over $H_{(i,k)}$. The details are left to the reader.

(iii), (iv) and (v) are trivial consequences of (i) and (ii).

(3.4) **Corollary** (of proof). *$M^0_{(n,0)}$ is isomorphic to an open subspace of projective $(4n-4)$-space. In particular, $M(-1,n)$ is rational.*

(3.5) *Remark.* The type (i,k) is a measure of the splitting behaviour along lines through the point $y \in P$. Intuitively, k is the generic splitting type, and $n-(i-k)(k+1)-ik$ is the number of exceptions. In particular, type $(n,0)$ means no jumping lines through y, and type $(n-1,0)$ means one simple jumping line.

(3.6) We turn our attention to the case $x=0$, $n=2m+1$, $m \geqq 1$. Let E be a bundle in M^0. The definition of type is a little more involved in this case, but follows the same general idea. First put $k=\operatorname{rank}(R^1 g_* \tilde{E}(-\tau))$; then the restriction of $\tilde{E}$ to a general fiber of g is of the form $\mathcal{O}(k) \oplus \mathcal{O}(-k)$. If $k>0$, we proceed as before, and say that (i,k) is the type of E if $g_* \tilde{E}(-k\tau) \simeq \mathcal{O}_Q(-i\sigma)$. If $k=0$, then $g_* \tilde{E}$ is of rank two, hence of the form $g_* \tilde{E} = \mathcal{O}(-i\sigma) \oplus \mathcal{O}(-j\sigma)$ where $i \geqq j$. Then the triple $(i,j,0)$ will be called the type of E. Let $M^0_{(i,k)}$ (resp. $M^0_{(i,j,0)}) \subseteq M$ be the locally closed reduced subscheme parametrizing bundles of type (i,k) (resp. $(i,j,0)$).

(3.7) **Theorem.** (i) *$M^0_{(i,k)}$ is nonempty if and only if the following inequalities hold:*

$$(*) \qquad \begin{aligned} &k>0 \\ &i-k>0 \\ &n-k(2i-k) \geqq 0. \end{aligned}$$

(ii) *$M^0_{(i,j,0)}$ is nonempty if and only if the following inequalities hold:*

$$(**) \qquad \begin{aligned} &j>0 \\ &i-j \geqq 0 \\ &n-(i+j) \geqq 0. \end{aligned}$$

(iii) *If $(*)$ holds, then $M^0_{(i,k)}$ is irreducible, nonsingular and rational of codimension $n-2i+k(2i-k+3)-1$ in M^0.*

(iv) *If $(**)$ holds and $i-j>0$, then $M^0_{(i,j,0)}$ is irreducible, nonsingular and rational of codimension $n-(i+j)+n-(2j+1)$ in M^0.*

(v) *If $(**)$ holds and $i=j$, then $M^0_{(i,j,0)}$ is irreducible, nonsingular and unirational of codimension $2(n-2i)$ in M^0.*

(vi) *$M^0_{(m+1,m,0)}$ is dense in M.*

(vii) *The complement of $M^0_{(m+1,m,0)}$ in M^0 has codimension 2.*

(viii) *The codimension of $M^0_{(i,k)}$ is at least $n+1$.*

Proof. (i), (iii), and (viii) go exactly like the proof of (3.3), and are omitted. As for (ii), suppose E is of type $(i,j,0)$. Then $i \geqq j$ by definition, and $j>0$ for the same reason as before. Any nonzero section of $\tilde{E}(j\sigma)$ induces a short exact sequence

$$0 \to \mathcal{O}(-j\sigma) \to \tilde{E} \to I_Y(j\sigma) \to 0$$

where $Y \subseteq F$ is a closed subscheme of length $c_2(\tilde{E}(j\sigma)) = n$. Applying g_* to this sequence we see that $g_* I_Y(j\sigma) = \mathcal{O}(-i\sigma)$. In particular, $h^0(I_Y((i+j-1)\sigma)) = 0$. Since the length of Y is n, we know that $h^0(I_Y(n\sigma)) \neq 0$. Thus $i+j-1 < n$, or $n-(i+j) \geqq 0$. Conversely, suppose (**) is satisfied. Let $Y \subseteq F$ be a group of n points subject to the condition that it is contained in $(i+j)$ fibers of g, but otherwise in general position. Then construct $\tilde{E}$ as an extension as above, and verify that $E = p_* \tilde{E}$ is of type $(0, i, j)$. The proof of (iv) and (v) proceeds along the lines of (3.3). Note, however, that if $i=j$, the associated subscheme Y is not uniquely determined; rather there is a pencil of them for a given bundle. Therefore there is a $\mathbb{P}^1$-bundle over $M^0_{(i,j,0)}$ that can be described as a nice variety over the Hilbert scheme. Finally, (vi), (vii) and (viii) are again trivial consequences of (i)–(v).

(3.8) **Corollary.** *$M(0, 2m+1)$ is rational.*

(3.9) *Remark.* Barth [2] has shown that $M(0, n)$ is rational for all $n \geqq 2$. If n is even, then the method of (3.7) gives only a rational $\mathbb{P}^1$-bundle over $M^0(0, n)$.

4. Non-Bundles

(4.1) Let $S \subseteq M$ be the closed subset corresponding to sheaves which are not locally free. We shall show below (4.6) that S can be given the structure of an irreducible divisor on M. Throughout this section, assume that $n \geqq 2$ (and odd whenever $x = 0$).

(4.2) **Lemma.** *Let E belong to S. Then there exists an exact sequence on P:*

$$0 \to E \to D \to T \to 0$$

where D is a semistable sheaf with $c_1(D) = c_1(E) = -x$, $c_2(D) = c_2(E) - 1 = n-1$, and T is a torsion sheaf of length 1. Furthermore, D is locally free if and only if length $(\mathcal{E}xt^1(E, \mathcal{O})) = 1$. In this case, the bundle D and the k-point of $\mathbb{P}(D)$ given by the exact sequence above is uniquely determined by E.

Proof. Let $T_0 = E^{\vee\vee}/E$ and let $T \subseteq T_0$ be some length-1 subsheaf. The sequence above results from letting $D \subseteq E^{\vee\vee}$ be the inverse image of T. The rest of the lemma is fairly obvious from (1.2); note that D is necessarily stable and hence simple in this case.

(4.3) **Lemma.** *There exists an irreducible, nonsingular variety R and a sheaf $\mathcal{D}$ on $P \times R$, flat over R, with the following properties:*

(i) *The fibers of $\mathcal{D}$ are semistable rank-2 sheaves on P with $c_1 = -x$ and $c_2 = n-1$.*

(ii) *The induced morphism $R \to M(-x, n-1)$ is surjective.*

(iii) *$\mathcal{D}$ is reflexive of projective dimension $\leqq 1$.*

Proof. Indeed, this is the way $M(-x, n-1)$ is constructed in the first place. Let $k \gg 0$ be an integer (it suffices to take $k \geqq n$ for example), let V be a trivial vector bundle on P of rank $(k+1)(k+2-x)-(n-1)$, and let Quot be the scheme parametrizing rank-2 quotients $V(-k) \to D$ of $V(-k)$ with $c_1(D) = -x$ and $c_2(D) = n-1$. Let $R \subseteq \mathrm{Quot}$ be the open subscheme given by the two conditions

(a) D is semistable.

(b) The induced map $H^0(V) \to H^0(D(k))$ is an isomorphism.

R comes equipped with a universal exact sequence on $R \times P$

$$0 \to \mathscr{K} \to pr_2^* V(-k\tau) \to \mathscr{D} \to 0.$$

Now $M(-x, n-1)$ is a good quotient of R by the natural action of $PGL(V)$, hence R is connected. It remains only to show that R is nonsingular, and (iii). To show that R is nonsingular in the point

$$0 \to K \to V(-k) \to D \to 0$$

it suffices to show that $\mathrm{Ext}^1(K, D) = 0$, by [6, Cor. 5.2]. There is an exact sequence

$$\mathrm{Ext}^1(V(-k), D) \to \mathrm{Ext}^1(K, D) \to \mathrm{Ext}^2(D, D) \to \mathrm{Ext}^2(V(-k), D)$$

where both end terms vanish since $k \gg 0$. Thus it suffices to show that $\mathrm{Ext}^2(D, D)$ vanishes. But (essentially by Serre duality) this space is dual to $\mathrm{Hom}(D, D(-3))$, which vanishes since D is semistable. Finally, to prove (iii), note that $\mathscr{K}$ is locally free since $\mathscr{D}$ is flat and each fiber of $\mathscr{D}$ has projective dimension $\leqq 1$. Thus $pd(\mathscr{D}) \leqq 1$. Now $\mathscr{D}$ is locally free off a subscheme of $R \times P$ of codimension 3, hence reflexive.

(4.4) *Remark.* If $x = 1$, we can take $R = M(-1, n-1)$ and $\mathscr{D}$ a universal family. If $x = 0$, this does not work for two different reasons: $M(0, n-1)$ is singular, and it does not carry a universal family [11].

(4.5) **Lemma.** *Let $\mathscr{D}$ be a reflexive rank-2 sheaf on a nonsingular irreducible variety Y, such that $pd(\mathscr{D}) \leqq 1$. Then $\mathbb{P}(\mathscr{D})$ is irreducible.*

Proof. The question is local on Y, hence we may assume that Y is affine and that $\mathscr{D} \simeq \mathscr{D}^\vee$, by [7, Prop. 1.10]. Pick a presentation on Y:

$$0 \to \mathscr{D}_1 \to \mathscr{D}_0 \to \mathscr{D} \to 0$$

where the $\mathscr{D}_i$ are locally free. Dualizing, we get

$$0 \to \mathscr{D} \to \mathscr{D}_0^\vee \to \mathscr{K} \to 0$$
$$0 \to \mathscr{K} \to \mathscr{D}_1^\vee \to \mathrm{Ext}^1(\mathscr{D}, \mathscr{O}) \to 0.$$

From this we get $\mathrm{Ext}^1(\mathscr{K}, \mathscr{L}) = 0$ for any line bundle $\mathscr{L}$ on Y, and therefore the linear projection $\mathbb{P}(\mathscr{D}_0^\vee) \dashrightarrow \mathbb{P}(\mathscr{D})$ is surjective. Since $\mathbb{P}(\mathscr{D}_0^\vee)$ is irreducible, we are done.

(4.6) **Proposition.** *Assume that $x=1$ (resp. $x=0$). Then S is an irreducible and reduced divisor in the linear system $|n\varepsilon-2\delta|$ (resp. $|(2n-1)\varphi-\psi|$).*

Proof. Pick a family $\mathscr{D}$ as in (4.3), and put $r\colon T=\mathbb{P}(\mathscr{D})\to R\times P$ with tautological quotient $r^*\mathscr{D}\to\mathcal{O}_T(\lambda)$, and let $j\colon T\to T\times P$ be the graph of the composed map $pr_2\circ r\colon T\to P$. Define a sheaf $\mathscr{F}$ on $T\times P$ via the exact sequence

$$0\to\mathscr{F}\to p^*\mathscr{D}\to j_*\mathcal{O}_T(\lambda)\to 0$$

where $p\colon T\times P\to R\times P$ is the map $(pr_1\circ r)\times id_P$. Intuitively, the family $\mathscr{F}$ can be described like this: a rational point $t\in T$ consists of (i) a point $w\in R$, and the corresponding sheaf $\mathscr{D}_w$ on P, and (ii) a length-1 quotient of $\mathscr{D}_w$. Then the fiber $\mathscr{F}_t$ of $\mathscr{F}$ is the kernel of this quotient. It is clear that $\mathscr{F}$ is flat over T. Let $T'\subseteq T$ be the open subscheme the closed points of which are those $t\in T$ such that $\mathscr{F}_t$ is stable. Then the restriction of $\mathscr{F}$ to $T'\times P$ induces a morphism $T'\to M$. In view of (4.2) and (4.3), the image is exactly S. By (4.5), T' is irreducible, hence so is S.

The scheme structure on S will be defined by the zero-th Fitting ideal of $\pi_*\mathscr{E}xt^1(\mathscr{E},\mathcal{O}(-\tau))$. The local-relative Ext spectral sequence yields the following presentation of this sheaf:

$$0\to R^1\pi_*(\mathscr{E}^\vee(-\tau))\to\mathscr{E}xt^1_\pi(\mathscr{E},\mathcal{O}(-\tau)).$$

The first term is isomorphic to $R^1\pi_*\mathscr{E}((x-1)\tau)(\alpha+\gamma-2\beta)$ in view of (2.11). The second term is the dual of $R^1\pi_*(\mathscr{E}(-2\tau))$ by relative duality [10, Thm. 21]. Hence the presentation is

$$0\to C(\alpha+\gamma-2\beta)\to A^\vee$$
$$(\text{resp. } 0\to B(\alpha+\gamma-2\beta)\to A^\vee).$$

From this it follows that the class of S is $-\alpha-\gamma-(n-1)(\alpha+\gamma-2\beta)=n\varepsilon-2\delta$ (resp. $-\alpha-\beta-n(\alpha+\gamma-2\beta)=(2n-1)\varphi-\psi$). Finally, the sheaf $\pi_*\mathscr{E}xt^1(\mathscr{E},\mathcal{O}(-\tau))$ has length 1 in the generic point of S, thus the zero-th Fitting ideal is reduced.

(4.7) *Remark.* In the spirit of the first half of this proof, we may construct a double cover Z of $M^0(0,2m)$ that carries a family, in the following way: Fix a point y of P, and let $Y\subseteq M(0,2m+1)$ be the locally closed subscheme corresponding to sheaves E with the property $\mathscr{E}xt^1(\mathscr{E},\mathcal{O})\simeq k(y)$. Give Y the reduced scheme structure. The restriction $\mathscr{E}_Y$ of the universal sheaf is no longer reflexive: Indeed, $(\mathscr{E}_Y)^{\vee\vee}$ is locally free and induces $\mathscr{E}^{\vee\vee}$ on the fibers. Thus the family $(\mathscr{E}_Y)^{\vee\vee}$ induces a morphism $h\colon Y\to M^0(0,2m)$, and the fiber of h over the point in $M^0(0,2m)$ corresponding to a bundle D is the projective line $\mathbb{P}(D\otimes k(y))$. In order to construct Z, we need a linebundle $\mathscr{L}$ on Y that has degree 2 when restricted to the fibers of h. But by the same reasoning as in (2.9), one easily shows that $\mathscr{L}=\mathcal{O}_Y(\psi)$ has this property. Twisting $\mathscr{L}$ with a very ample linebundle on $M^0(0,2m)$ if necessary, we may assume that $\mathscr{L}$ is very ample. Now let $Z\subseteq Y$ be a general divisor in $|\mathscr{L}|$. By Bertini's theorem, Z is nonsingular and irreducible (since this is true for Y). Furthermore, Z carries the family $(\mathscr{E}_Y)_Z^{\vee\vee}$ on $Z\times P$, and the induced morphism $Z\to M^0(0,2m)$ is the restriction of h, hence a double cover.

5. Some Divisors on the Moduli Space

(5.1) The purpose of this section is to give the divisor classes ε and δ (resp. φ and ψ) of (2.6) a geometric interpretation, in terms of conditions on stable sheaves. We also compute the canonical divisor class. Throughout, assume $n \geqq 2$.

(5.2) Consider the divisors on M given set-theoretically as follows. Assume $x = 1$ (resp. $x=0$).

$D_1 = \{$sheaves with a given jumping conic (resp. line)$\}$,
$D_2 = \{$sheaves with a jumping line (resp. conic) passing through 1 (resp. 3) given point(s)$\}$.

The scheme structures are to be defined below. Note that each D_i is a proper subset of M, by (1.8, (iii)).

(5.3) The scheme-theoretic definition of D_1 is the following: Let $Y \subseteq P$ be a conic (resp. line), and define D_1 by the zero-th Fitting ideal of $R^1 \pi_* \mathscr{E}_Y$ (resp. $R^1 \pi_* \mathscr{E}_Y(-\tau)$). The exact sequence on $M \times P$:

$$0 \to \mathscr{E}(-2\tau) \to \mathscr{E}((x-1)\tau) \to \mathscr{E}_Y((x-1)\tau) \to 0$$

gives the following presentation on M:

$$0 \to A \to C \to R^1 \pi_* \mathscr{E}_Y \to 0$$
$$(\text{resp. } 0 \to A \to B \to R^1 \pi_* \mathscr{E}_Y(-\tau) \to 0).$$

Therefore, the divisor class of D_1 is $\gamma - \alpha = \varepsilon$ (resp. $\beta - \alpha = \varphi$).

(5.4) In order to define D_2, let Q be the projective line (resp. projective plane) parametrizing the pencil of lines passing through a point of P (resp. the net of conics passing through three general given points of P); let $\Sigma \subseteq P \times Q$ be the incidence correspondence, and $p: \Sigma \to P$, $q: \Sigma \to Q$ be the projections. Denote by σ the positive generator of $\mathrm{Pic}(Q) = \mathbb{Z}$. Then $\mathcal{O}_\Sigma$ has the following resolution in $P \times Q$

$$0 \to \mathcal{O}(-\sigma - (2-x)\tau) \to \mathcal{O} \to \mathcal{O}_\Sigma \to 0.$$

In particular, one verifies that

$$\begin{aligned}
q_* p^* \mathcal{O}(-\tau) &= 0 & & \\
R^1 q_* p^* \mathcal{O}(-\tau) &= 0 & &(\text{resp. } = \mathcal{O}(-\sigma)) \\
q_* p^* \Omega(\tau) &= \mathcal{O}(-\sigma) & &(\text{resp. } = 0) \\
R^1 q_* p^* \Omega(\tau) &= 0. & &
\end{aligned}$$

Now, using this, the monad for $\mathscr{E}$ induces an exact sequence on $M \times Q$:

$$0 \to \tilde{q}_* \tilde{p}^* \mathscr{E} \to \pi^* B(-\sigma) \to \pi^* C \to R^1 \tilde{q}_* \tilde{p}^* \mathscr{E} \to 0$$
$$(\text{resp. } 0 \to \tilde{q}_* \tilde{p}^* \mathscr{E} \to \pi^* A(-\sigma) \to \pi^* C \to R^1 \tilde{q}_* \tilde{p}^* \mathscr{E} \to 0).$$

Here $\tilde{p}=(1_M \times p)$, $\tilde{q}=(1_M \times q)$, and π: $M \times Q \to M$ is the projection. Let $Z \subseteq M \times Q$ be defined by the zero-th Fitting ideal of $R^1 \tilde{q}_* \tilde{p}^* \mathscr{E}$. Then clearly the set $D_2 \subseteq M$ of (3.2) is the image of Z under π.

(5.5) **Lemma.** *All irreducible components of Z are generically finite over $D_2 \subseteq M$. In particular, Z has pure codimension* 2 (*resp.* 3) *in $M \times Q$.*

Proof. Let $I_k \subseteq M$ be the closed locus where the fibers of Z have dimension $\geqq k$, $k=1,2$. It suffices to show that codim $I_1 \geqq 3$ (resp. codim $I_1 \geqq 3$ and codim $I_2 \geqq 4$). We divide the proof according to whether $x=1$ or $x=0$.

If $x=1$, we are to show that $\operatorname{codim}(I_1, M) \geqq 3$. In view of (3.3, (v)), the codimension of $I_1 \cap M^0$ is at least $n+1 \geqq 3$. Thus it remains to show that $I_1 \cap S$ has codimension at least 2 in S. To do so, we may throw away the set $S_y \subseteq S$ of sheaves E with the property that $\mathscr{E}xt^1(E, \mathcal{O}) \otimes \mathcal{O}_{P,y}$ has length $\geqq 2$, since S_y has codimension $\geqq 5$ in S. Now if E belongs to $I_1 \cap (S - S_y)$, then all the lines containing y are jumping for $E^{\vee\vee}$. Thus $E^{\vee\vee}$ belongs to a family of large codimension in view of (3.3, (v)).

In the case $x=0$, the proof is similar, though a little more involved. We may specialize the net Q to be the set of all line-pairs which has a fixed point $y \in P$ as a double point. If a line-pair is a jumping conic (1.7), then either each component line is jumping, or one of them is a multiple jumping line. We leave the details to the reader.

(5.6) In view of (5.5), it is reasonable to define the divisor $D_2 \subseteq M$ by the zero-th Fitting ideal of $\pi_*(\mathcal{O}_Y)$. Then the class of D_2 in $\operatorname{Pic}(M)$ can be computed as $\pi_*([Z])$, where $[Z] \in A^{3-x}(M \times P)$ is the class of Z in the Chow ring. By (5.5), $[Z]$ can be computed from the exact sequence in (5.4) using Porteous' formula [17, Sect. 6]. The result is the following:

The divisor class of D_2 is δ (resp. $\frac{1}{2}(n-1)\psi$).

(5.7) Since M is a fine moduli space, the Kodaira-Spencer map $T_M \to \mathscr{E}xt^1_\pi(\mathscr{E}, \mathscr{E})$ is an isomorphism, where T_M is the tangent bundle on M. In the Grothendieck group of M, we have the following equality, coming from the monad:

$$\begin{aligned}\sum_{i \geqq 0} (-1)^i \mathscr{E}xt^i_\pi(\mathscr{E}, \mathscr{E}) = {} & \mathscr{H}om(A, A) + \mathscr{H}om(B, B) + \mathscr{H}om(C, C) \\ & -3\,\mathscr{H}om(A, B) - 3\,\mathscr{H}om(B, C) + 3\,\mathscr{H}om(A, C).\end{aligned}$$

Since $\mathscr{H}om_\pi(\mathscr{E}, \mathscr{E}) = \mathcal{O}_M$ and $\mathscr{E}xt^i_\pi(\mathscr{E}, \mathscr{E}) = 0$ for $i \geqq 2$, we get the following expression for the canonical class $\kappa = -c_1(\mathscr{E}xt^1_\pi(\mathscr{E}, \mathscr{E}))$:

$$\begin{aligned}\kappa &= 3(c_1(A^\vee \otimes C) - c_1(A^\vee \otimes B) - c_1(B^\vee \otimes C)) \\ &= \begin{cases} -3\varepsilon & (\text{if } x=1) \\ -6\varphi & (\text{if } x=0). \end{cases}\end{aligned}$$

In the case $x=0$, this result has been obtained by Le Potier [12, Thm. 2] over the complex numbers.

Note that the anticanonical linebundle $\mathcal{O}(-\kappa)$ is not ample, by (2.10, (iv)).

6. The Picard Group

(6.1) **Theorem.** *Assume that $n \geqq 2$, $m \geqq 1$. Then*
(i) $\mathrm{Pic}(M(-1,n))$ *is freely generated by* ε *and* δ.
(ii) $\mathrm{Pic}(M(0,2m+1))$ *is freely generated by* φ *and* ψ.

Remark. ε, δ, φ, and ψ are defined in (2.6).

Proof. In view of (2.10) it suffices to show that these classes generate the Picard group of M. For this purpose, the exact sequences

$$A_\cdot(Y) \to A_\cdot(X) \to A_\cdot(X-Y) \to 0$$

where $Y \subseteq X$ is a closed subscheme, allow us to remove from M any irreducible and reduced divisor, the class of which may be expressed as an integral linear combination of ε and δ (resp. φ and ψ).

For example, in case (ii), Le Potier [12, Thm. 1] has shown that $\mathrm{Pic}(M^0(0,2m+1))$ is generated by φ (at least over the complex numbers. An independent proof, valid in any characteristic, is easily found along similar lines as the one below for $x=1$). In view of (4.6), the theorem follows in case (ii). In case (i), it also suffices to consider M^0, in view of (4.6). Fix a point y in P, and consider the divisor $D_2 \cap M^0$ of M^0 associated to y, see (5.2) and (5.6). In view of (3.3, (iv)), we are reduced to proving that ε generates the Picard group of $M^0_{(n,0)} = U$. Keeping the notation of (3.1), we see that the inclusion of U in M can be defined by a universal extension on $U \times F$:

$$0 \to \mathcal{O}(-n\sigma+\rho) \to \mathcal{D} \to \mathcal{O}(n\sigma-\tau) \to 0$$

where ρ is the linebundle on U inducing the open embedding $U \hookrightarrow \mathbb{P}(\mathrm{Ext}^1(\mathcal{O}(n\sigma-\tau), \mathcal{O}(-n\sigma))^\vee)$. In particular, $\mathrm{Pic}(U)$ is generated by ρ. A simple calculation shows that the restrictions to U of ε and δ are

$$\varepsilon = (2n-1)\rho, \qquad \delta = n(n-1)\rho.$$

Since $U \subseteq M^0 = M - S$, the divisor class $[S] = n\varepsilon - 2\delta = n\rho = 0$ in $\mathrm{Pic}(U)$. Therefore $\varepsilon = -\rho$ is a generator of $\mathrm{Pic}(U)$.

(6.2) **Corollary.** *Assume that $n \geqq 2$, $m \geqq 1$. Then*

(i) $\mathrm{Pic}(M^0(-1,n)) = \begin{cases} \mathbb{Z} & (n \text{ odd}) \\ \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} & (n \text{ even}). \end{cases}$

(ii) $\mathrm{Pic}(M^0(0,2m+1)) = \mathbb{Z}$.

(6.3) *Remark.* (i) was proved in [4] using basically similar techniques. Indeed, the proof above that $\mathrm{Pic}(U)$ is generated by ε is contained in that paper. (ii) was proved by Le Potier [12, Thm. 1] over the complex numbers. He also proved that $\mathrm{Pic}(M^0(0,2m)) = \mathbb{Z}$ for $m \geqq 2$, and $= \mathbb{Z}/3\mathbb{Z}$ for $m = 1$.

7. The Ample Cone

(7.1) We want to show that the necessary condition for ampleness (2.10, (iv)) is also sufficient. In order to do this we use the description of M as a geometric quotient and the result (1.16) on the stability of the points $T_k(E)$ in the Gieseker space. Throughout this section, assume that $x=1$, $n\geqq 2$ (resp. $x=0$, $n=2m+1\geqq 3$).

(7.2) In view of (1.10), for each $k\geqq n-2+x$, the sheaf $\mathscr{E}_k=\pi_*(\mathscr{E}(k\tau))$ is locally free on M of rank $(k+1)(k+2-x)-n$ and commutes with base change on M. Let H_k be a k-vector space of dimension $(k+1)(k+2-x)-n$, and let W_k be a vector space of dimension $h^0(\mathcal{O}_P(2k-x))=(k+x)(2k+1)$. Consider the projective bundle

$$a_k\colon \bar{X}_k=\mathbb{P}(\mathscr{H}om(\mathscr{E}_k, H_k\otimes\mathcal{O}_M)^\vee)\to M$$

with tautological linebundle $\mathcal{O}(\lambda_k)$. On $\bar{X}_k$, there is a natural homomorphism

$$f_k\colon a_k^*\,\mathscr{E}_k\to H_k\otimes\mathcal{O}(\lambda_k).$$

Let $X_k\subseteq\bar{X}_k$ be the open subscheme where f_k is an isomorphism. Then $a_k\colon X_k\to M$ is the principal $PGL(H_k)$-bundle associated to the locally free sheaf $\mathscr{E}_k$.

(7.3) Let $G_k=\mathbb{P}(\mathrm{Hom}(\Lambda^2 H_k, W_k)^\vee)$ be the associated Gieseker space, with the natural action of $PGL(H_k)$ on it. We define a $PGL(H_k)$-equivariant morphism

$$j_k\colon X_k\to G_k$$

as follows: First fix an isomorphism $b\colon \pi_*(\det(\mathscr{E}(k\tau))\to W_k\otimes\mathcal{O}_M(2\beta-\alpha-\gamma)$, cf. (2.11), and let $c\colon \Lambda^2\mathscr{E}_k\to\pi_*(\det(\mathscr{E}(k\tau)))$ be the natural map. Now take the pullback to X_k of $b\circ c$ and compose it with the inverse of $\Lambda^2 f_k$, to get the following map on X_k:

$$T_k\colon \Lambda^2(H_k\otimes\mathcal{O}(\lambda_k))\to W_k\otimes\mathcal{O}(2\beta-\alpha-\gamma).$$

This induces a map j_k as above, such that $j_k^*\,\mathcal{O}_{G_k}(1)=\mathcal{O}(2\beta-\alpha-\gamma-2\lambda_k)$.

(7.4) **Lemma.** *$j_k(X_k)\subseteq(G_k)^{ss}$ for $k\geqq n-2+x$, and $j_k(X_k)\subseteq(G_k)^s$ for $k\geqq n$ or $x=0$, where $(G_k)^s\subseteq(G_k)^{ss}\subseteq G_k$ are the open subsets of stable (resp. semistable) points.*

Proof. In more down-to-earth terms, the map j_k of (7.3) is the following: A point of X_k is a pair (E,f) where E is a stable sheaf and $f\colon H^0(E(k))\to H_k$ is an isomorphism. Then $j_k(E,f)$ is nothing but the point $T_k(E)$ of (1.15) transported to G_k via the isomorphism f. Therefore the lemma is but a rephrazing of (1.16).

(7.5) Let $g_k\colon (G_k)^{ss}\to Y_k$ be the quotient by the action of $PGL(H_k)$. By Mumford's geometric invariant theory [14], this exists and is projective; let $\eta\in\mathrm{Pic}(Y_k)$ be an ample divisor class. The morphism $g_k\circ j_k\colon X_k\to Y_k$ is $PGL(H_k)$-invariant, hence factors uniquely through M to give a commutative diagram

$$\begin{array}{ccc} X_k & \xrightarrow{j_k} & (G_k)^{ss} \\ {\scriptstyle a_k}\downarrow & & \downarrow{\scriptstyle g_k} \\ M & \xrightarrow[h_k]{} & Y_k. \end{array}$$

(7.6) **Lemma.** *The pullback* $h_k^*(\eta)$ *is a multiple of* $(k-n+1)\varepsilon+2\delta$ *(resp.* $2(k-n+2)\varphi+\psi$*).*

Proof. First note that a_k^*: $\operatorname{Pic} M \to \operatorname{Pic} X_k$ is injective; in fact, $\operatorname{Pic}(X_k)$ is generated over $\operatorname{Pic} M$ by λ_k with the single relation

$$((k+1)(k+2-x)-n)\lambda_k = c_1(\mathscr{E}_k).$$

Let N be the integer such that $g_k^*\mathcal{O}(\eta)=\mathcal{O}_{G_k}(N)$. Then by (7.3) we have

$$a_k^* h_k^*(\eta) = j_k^* g_k^* \eta = N(2\beta-\alpha-\gamma-2\lambda_k)$$

in $\operatorname{Pic}(X_k)$. Therefore, since a_k^* is injective, we have the following relation in $\operatorname{Pic}(M)$:

$$\frac{1}{N}((k+1)(k+2-x)-n)h_k^*(\eta)=((k+1)(k+2-x)-n)(2\beta-\alpha-\gamma)-2c_1(\mathscr{E}_k).$$

The latter expression is easily computed using the monad

$$0\to A\otimes H^0(\mathcal{O}_P(k-1))\to B\otimes H^0(\Omega_P(k+1))\to C\otimes H^0(\mathcal{O}_P(k))\to 0$$

for $\mathscr{E}_k$, and we get the expression of (7.6).

(7.7) **Corollary.** *For all sufficiently large integers b, the linear system* $|b\,\delta|$ *(resp.* $|b\,\psi|$*) is base-point free.*

Proof. Put $k=n-1$ (resp. $k=n-2$) in (7.6) to see that 2δ (resp. ψ) is the pullback of an ample divisor class.

(7.8) **Corollary.** *For all sufficiently large integers a, the divisor class* $a\varepsilon+2\delta$ *(resp.* $a\varphi+\psi$*) is ample.*

Proof. Indeed, Gieseker proves that h_k is a closed embedding for $k\gg 0$ [5, Sect. 4].

(7.9) **Theorem.** *Let* $\xi=a\varepsilon+b\delta$ *(resp.* $\xi=a\varphi+b\psi$*) be an element of* $\operatorname{Pic}(M)$. *Then* ξ *is ample if and only if* $a>0$ *and* $b>0$.

Proof. The necessity of the numerical condition has been noted already in (2.10). The sufficiency is immediate from (7.7) and (7.8).

(7.10) *Remark.* In characteristic zero, the Grauert-Mülich theorem implies that $|\varepsilon|$ (resp. $|\varphi|$) is base-point free. In the general case, at least ε (resp. φ) is in the closure of the ample cone ($\otimes\mathbb{R}$). I don't know whether some multiple of ε (resp. φ) is base-point free.

References

1. Barth, W.: Some properties of stable rank-2 vector bundles on $\mathbb{P}_n$. Math. Ann. **226**, 125–150 (1977)
2. Barth, W.: Moduli of vector bundles on the projective plane. Invent. Math. **42**, 63–91 (1977)
3. Ellingsrud, G., Strømme, S.A.: On the moduli space for stable rank-2 vector bundles on $\mathbb{P}^2$. Preprint, Oslo (1979)
4. Ellingsrud, G., Strømme, S.A.: The Picard group of the moduli space for stable rank-2 vector bundles on $\mathbb{P}^2$ with odd first Chern class. Preprint, Oslo (1979)
5. Gieseker, D.: On the moduli of vector bundles on an algebraic surface. Ann. of Math. **106**, 45–60 (1977)
6. Grothendieck, A.: Techniques de construction et théorèmes d'existence en géométrie algébrique IV: Les schémas de Hilbert. Sem. Bourbaki 221 (1960/61)
7. Hartshorne, R.: Stable reflexive sheaves. Math. Ann. **254**, 121–176 (1980)
8. Hartshorne, R.: Algebraic Geometry. Graduate Texts in Math. 52. Berlin, Heidelberg, New York: Springer 1977
9. Hulek, K.: Stable rank-2 vector bundles on $\mathbb{P}_2$ with c_1 odd. Math. Ann. **242**, 241–266 (1979)
10. Kleiman, S.: Relative duality for quasi-coherent sheaves. Compositio Math. **41**, 39–60 (1980)
11. Le Potier, J.: Fibrés stables de rang 2 sur $\mathbb{P}_2(\mathbb{C})$. Math. Ann. **241**, 217–256 (1979)
12. Le Potier, J.: Sur le groupe de Picard de l'espace de modules de fibrés stables sur $\mathbb{P}_2$. Ann. scient. Éc. Norm. Sup. **14**, 141–155 (1981)
13. Maruyama, M.: Moduli of stable sheaves, II. J. Math. Kyoto Univ. **18**, 557–614 (1978)
14. Mumford, D.: Geometric invariant theory. Berlin, Heidelberg, New York: Springer 1965
15. Schneider, M.: Holomorphic vector bundles on $\mathbb{P}_n$. Sem. Bourbaki 530 (1978/79)
16. Schwarzenberger, R.L.E.: Vector bundles on algebraic surfaces. Proc. London Math. Soc. **11**, 601–622 (1961)
17. Schwarzenberger, R.L.E.: The secant bundle of a projective variety. Proc. London Math. Soc. (3) **14**, 369–384 (1964)

Received March 12, 1984

Math. Z. 187, 425–432 (1984)

Mathematische Zeitschrift

On the Spectrum of Non-Compact Manifolds with Finite Volume

Robert Brooks*

Courant Inst. of Math. Sciences, New York Univ., New York, N.Y. 10012, USA**

Let M be a complete, non-compact Riemannian manifold. Then we denote by λ_0^{ess} the lower bound of the essential spectrum of the Laplacian acting on $L^2(M)$. λ_0^{ess} is defined by the variational formula

$$\lambda_0^{\mathrm{ess}} = \lim_K \inf_f \frac{\int_M \|\operatorname{grad} f\|^2}{\int_M f^2}$$

where K runs over an increasing set of compact subdomains of M such that $\bigcup K = M$, and f has compact support in $M-K$. The essential spectrum is the closure of that part of the spectrum not corresponding to discrete eigenvalues with finite multiplicity. For a compact manifold, the essential spectrum is empty, so that $\lambda_0^{\mathrm{ess}} = \infty$.

See [3, 4] for a discussion of these notions, and some results relevant to our discussion here. In particular, it was shown in [3] that if the volume of M is infinite, then

$$\tfrac{1}{4}h^2 \leqq \lambda_0^{\mathrm{ess}} \leqq \tfrac{1}{4}\mu^2$$

where h is Cheeger's isoperimetric constant (see [2]), and μ is the exponential growth of M –

$$\mu = \overline{\lim_{r\to\infty}} \frac{1}{r} \log \operatorname{vol}(B(r, x_0)),$$

where $B(r, x_0)$ is the ball of radius r about an arbitrary fixed point x_0 of M. In particular, if M has infinite volume and $\lambda_0^{\mathrm{ess}} = \infty$, i.e. M has discrete spectrum, then also $\mu = \infty$, i.e. the volume of a ball of radius r must grow extremely rapidly with r.

In [3], we left open the case where M has finite volume. Indeed, in this case the estimates of [3] look unpromising, since both h and μ are trivially 0,

---

* *Current address:* Univ. of Southern Calif., Los Angeles, California 90089-1113, USA

** Partially supported by NSF grant MCS 8102747

whereas it is known that there exist finite-volume M with discrete spectrum, $\lambda_0^{\mathrm{ess}}=\infty$.

The object of this paper is to overcome these difficulties. Indeed, let

$$\mu_f=\overline{\lim_{r\to\infty}}\,\frac{-1}{r}\log(\mathrm{vol}(M)-\mathrm{vol}(B(r,x_0)))$$

and let

$$h^{\mathrm{ess}}=\lim_K h(M-K)$$

where K runs over an exhaustion of M by compact subdomains of M. Then

Theorem 1. *If M has finite volume, then*

$$\tfrac{1}{4}(h^{\mathrm{ess}})^2\leqq\lambda_0^{\mathrm{ess}}\leqq\tfrac{1}{4}\mu_f^2.$$

In particular, if M has discrete spectrum then $\mathrm{vol}(B(r,x_0))$ must approach its maximum extremely rapidly. The left-hand inequality is essentially Cheeger's inequality, together with the Decomposition Principle that the essential spectrum is unchanged by compact perturbation [5], and is shown in [3]. The main difficulty with h^{ess} is evaluating it in specific cases, and we discuss this in §2. The right-hand inequality follows in a manner similar to the argument of [3], as we will see in §1 below.

To understand the theorem, it is perhaps useful to keep in mind the turnip (Fig. 1 below), which is the surface of revolution of a rapidly decreasing function $\gamma(x)$, $x>0$, with a bulb attached at $x=0$:

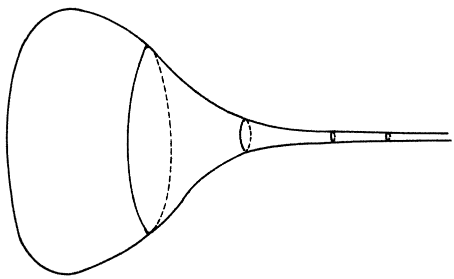


Fig. 1

This case was extensively dealt with by Baider [2]. Note that here $h=0$, since the piece of M with $x\leqq x_1$ has increasing volume as $x_1\to\infty$, but the length of the boundary tends to 0. However, if we remove from M compact pieces K_{x_2} of the form $x\leqq x_2$, then the isoperimetric constant of $M-K_{x_2}$ may indeed be positive, since there is not enough volume on the right-hand side of a curve dividing $M-K_{x_2}$ in two pieces to compensate for the length of the curve. While this is not a precise argument, we will show in §2 below how one may estimate h^{ess} from below in many cases such as this.

We also examine in §2 the surfaces of revolution M_α generated by the functions e^{-x^α}, $\alpha>0$, as did Baider in [2]. We find:

Theorem 2. (i) *For* $\alpha > 1$, M_α *has discrete spectrum*, $\lambda_0^{\mathrm{ess}}(M_\alpha) = \infty$

(ii) *For* $\alpha = 1$, $\lambda_0^{\mathrm{ess}}(M_\alpha) = \frac{1}{4}$.

(iii) *For* $0 < \alpha < 1$, $\lambda_0^{\mathrm{ess}}(M_\alpha) = 0$.

In § 3, we present some comments about the sharpness of the estimates of § 1 and § 2.

It is a pleasure to thank S. Agmon and J. Brüning for stimulating comments which were helpful in the development of this work. In particular, the invariant μ_f was suggested to me by Agmon. See also Agmon's book [1] for related results.

§ 1. Proof of Theorem 1

We will derive Theorem 1 from the following general result, which was proved in [3], see also [1]:

Theorem 3. *Let K be a compact subset of M, and $\lambda_0(M-K)$ = the greatest lower bound of the spectrum of Δ on $L^2(M-K)$, with Dirichlet boundary conditions on ∂K.*

Let $\rho(x) = \rho(x, x_0)$ denote the distance from a fixed point $x_0 \in M$. If

$$\int_{M-K} e^{-2\alpha\rho(x)}\,dx < \infty \qquad \textit{for } 0 < \alpha < \sqrt{\lambda_0}(M-K)$$

then

$$\int_{M-K} e^{2\alpha\rho(x)}\,dx < \infty.$$

Note that when M has finite volume, the first inequality is trivially satisfied, since the integral is dominated by $\mathrm{vol}(M-K)$. To establish Theorem 1, we must show that the second inequality is impossible if $2\alpha > \mu_f$. (Note that this is the opposite of the case in [3], where the second inequality was trivially false for all α, and one had to find conditions on α which made the first inequality true). But

$$(*) \qquad \int_{M-K} e^{2\alpha\rho(x)}\,dx \geqq \sum [\mathrm{vol}(B(r+1)) - \mathrm{vol}(B(r))]\, e^{2\alpha r}$$
$$= \sum [V(r) - V(r+1)]\, e^{2\alpha r}$$

where the sum begins on the first r such that $K \subset B(r)$, and where we have abbreviated $V(r) = \mathrm{Vol}(M) - \mathrm{Vol}(B(r))$. Suppose that $\dfrac{V(r+1)}{V(r)} > c$ for all r greater than some fixed r_0. Then $V(r_0+k) > c^k\, V(r_0)$ and so $-\dfrac{1}{r_0+k}\log(V(r_0+k)) < -\dfrac{k}{r_0+k}\log(c) - \dfrac{\log(V(r_0))}{r_0+k}$ and so taking $\overline{\lim}$ of both sides, we find $\mu_f \leqq -\log(c)$.

If $\mu_f > 0$, then taking c sufficiently close to 1 so that $\mu_f > -\log(c)$, we see from the above that $V(r+1) < c\,V(r)$ for infinitely many values of r. Writing $\varepsilon = 1 - c$, we may rewrite this as

$$V(r) - V(r+1) > V(r) - c\,V(r) = \varepsilon\, V(r).$$

Now assume that $2\alpha > \mu_f + \delta$, for some $\delta > 0$. Then for all r sufficiently large,

$$-\frac{1}{r}\log(V(r)) < \mu_f + \delta$$

so that

$$V(r) > e^{-r(\mu_f+\delta)}.$$

It follows that infinitely many terms in the infinite sum $(*)$ are bounded below by $\varepsilon e^{(2\alpha-\mu_f-\delta)r} > \varepsilon$, and hence this sum is not finite.

If $\mu_f = 0$, we will argue separately that $\lambda_0^{\mathrm{ess}} = 0$. Consider the test functions $f_{i,k}$ defined by

$$\begin{aligned} f_{i,k} &= 0 && \text{on } B(i) \\ &= r - i && \text{for } i \leqq r \subseteq i+k \\ &= k && \text{on } M - B(i+k). \end{aligned}$$

Then

$$\int_{M-K} \|\operatorname{grad}(f_{i,k})\|^2 \leqq \operatorname{Vol}(B(i+k)) - \operatorname{Vol}(B_i) \leqq V(i)$$

while

$$\int_{M-K} (f_{i,k})^2 \geqq k^2\, V(i+k)$$

so that $\lambda_0(M-K) \leqq \dfrac{V(i)}{k^2\, V(i+k)}$. Suppose that $\dfrac{V(i)}{k^2\, V(i+k)} \geqq \varepsilon$ for all i, k sufficiently large. Then $V(i+k) \leqq \dfrac{V(i)}{k^2\,\varepsilon}$ for all i, and hence $V(i+nk) \leqq \left(\dfrac{1}{k^2\,\varepsilon}\right)^n V(i)$, and so

$$-\frac{1}{i+nk}\log(V(i+nk)) \geqq \frac{1}{i+nk}\,[n\log(k^2\,\varepsilon) - \log(V_i)].$$

Letting n tend to ∞, we find $\mu_f \geqq \dfrac{1}{k}\log(k^2\,\varepsilon)$, which contradicts $\mu_f = 0$ as soon as $k^2 > \dfrac{1}{\varepsilon}$.

This completes the proof of Theorem 1.

Note that the inequality $\mu_f \geqq \dfrac{1}{k}\log(k^2\,\varepsilon)$ is valid also when $\mu_f > 0$, and achieves its maximum at $k = \dfrac{e}{\sqrt{\varepsilon}}$. This elementary argument gives the estimate $\lambda_0^{\mathrm{ess}} \leqq \dfrac{e^2}{4}\,\mu_f^2$, which is weaker than Theorem 1.

§ 2. Some Turnips

Consider now the surface of revolution M_γ in $\mathbb{R}^3$ generated by the graph of the positive function $z = \gamma(x)$, where $x \geqq 0$, with a bulb attached at $x = 0$. In this case, Baider [2] obtained a necessary and sufficient condition for M to have discrete spectrum, in terms of an integral involving γ.

We show here how one may estimate h^{ess} and μ_f in terms of γ. We point out in §3 that this is not enough in general to recover Baider's condition. However, we give these computations in detail in this section for $\gamma = e^{-x^\alpha}$, so that we can compute λ_0^{ess} for such manifolds. In §3 we observe that under mild restrictions on γ at infinity, these techniques give a sharp computation of $\lambda_0^{\text{ess}}(M_\gamma)$.

We begin first by estimating μ_f:

By elementary formulas, we have

$$\operatorname{vol}(M) - \operatorname{vol}(B_r) = 2\pi \int_x^\infty \gamma(1 + (\dot\gamma)^2)^{1/2}\, dt$$

where

$$r = \int_0^x (1 + (\dot\gamma)^2)^{1/2}\, dt.$$

Thus,

$$\mu_f = \overline{\lim_{x\to\infty}} \frac{-1}{\int_0^x (1+(\dot\gamma)^2)^{1/2}} \log\left(\int_x^\infty \gamma(1+(\dot\gamma)^2)^{1/2}\right).$$

We now consider the special case $\gamma = e^{-x^\alpha}$. Then

$$\dot\gamma = -a x^{\alpha-1} e^{-x^\alpha}$$

so that $(1+(\dot\gamma)^2)^{1/2}$ tends rapidly to 1 as $x \to \infty$. Thus,

$$\mu_f = \overline{\lim_{x\to\infty}} \frac{-1}{x} \log \int_x^\infty \gamma.$$

As Baider observes, the integral $\int_x^\infty e^{-x^\alpha}$ is easily estimated as

$$\int_x^\infty e^{-x^\alpha} \sim \frac{x^{1-\alpha} e^{-x^\alpha}}{\alpha},$$

so that $-\frac{1}{x}\log \int_x^\infty \gamma \sim -\frac{(1-\alpha)\log(x)}{x} + x^{\alpha-1}$, as $x \to \infty$, so that the limit is ∞ iff $\alpha > 1$, 1 if $\alpha = 1$, and 0 for $0 < \alpha < 1$. Thus, the surface of revolution generated by e^{-x^α} has $\lambda_0^{\text{ess}} = 0$ if $\alpha < 1$, and $\lambda_0^{\text{ess}} \leqq \frac{1}{4}$ if $\alpha = 1$.

We now turn to estimating h^{ess}. Here is a simple criterion for estimating it:

Lemma. *Let X be a vector-field on $M - K$. Then*

$$h(M-K) \geqq \frac{\inf|\operatorname{div}(X)|}{\sup\|X\|}$$

where $\|X\|$ denotes the Riemannian norm of X. In particular,

$$h^{\text{ess}} \geqq \lim_{r\to\infty} \frac{\inf|\operatorname{div} X|}{\sup\|X\|}.$$

Proof. Consider the $(n-1)$-form $\iota_X(\mathrm{vol})$. Then $\operatorname{div}(X)$ is defined by

$$\operatorname{div}(X)\cdot \mathrm{vol}=d(\iota_X(\mathrm{vol})).$$

Let D be a compact region in $M-K$, with smooth boundary. Then

$$\begin{aligned}\sup\|X\|\cdot \operatorname{area}(\partial D) &\geqq \Big|\int_{\partial D}\iota_x(\mathrm{vol})\Big|\\ &=\Big|\int_D d(\iota_x(\mathrm{vol}))\Big|=\Big|\int_D \operatorname{div}(X)\cdot \mathrm{vol}\Big|\\ &\geqq \inf|\operatorname{div}(X)|\cdot \mathrm{vol}(D).\end{aligned}$$

Hence $\dfrac{\operatorname{area}(\partial D)}{\mathrm{vol}(D)}\geqq\dfrac{\inf|\operatorname{div}(X)|}{\sup\|X\|}$, and taking the infimum over all D, we see that $h(M-K)\geqq\dfrac{\inf|\operatorname{div}(X)|}{\sup\|X\|}$.

We now apply these considerations to the surface of revolution generated by $\gamma(x)$, which we may represent as the solution of the equation $y^2+z^2=\gamma^2(x)$. An obvious candidate for X is then the vector-field $\dfrac{\partial}{\partial x}+y\cdot\dfrac{\dot\gamma}{\gamma}\dfrac{\partial}{\partial y}+z\cdot\dfrac{\dot\gamma}{\gamma}\dfrac{\partial}{\partial z}$, which has length $(1+(\dot\gamma)^2)^{1/2}$.

To calculate $\operatorname{div}(X)$, let us parametrize the part of M with $z>0$ by the map

$$g\colon \{(x,y)\subseteq\mathbb{R}\times[-1,1]\to(x, y(x)\cdot y, \gamma(x)\sqrt{1-y^2})$$

so that $X=g_*\left(\dfrac{\partial}{\partial x}\right)$.

Let $Y=g_*\left(\dfrac{\partial}{\partial y}\right)=\left(0,\gamma(x),\gamma(x)\dfrac{-y}{\sqrt{1-y^2}}\right)$. Then

$$g^*(\mathrm{vol})=\left\|g_*\left(\frac{\partial}{\partial x}\right)\right\|\left\|g_*\left(\frac{\partial}{\partial y}\right)\right\|dx\wedge dy$$

since $\langle X,Y\rangle=0$, and we find $g^*(\mathrm{vol})=(1+(\dot\gamma)^2)^{1/2}(\gamma)\left[\dfrac{1}{\sqrt{1-y^2}}\right]dx\wedge dy$. It follows that $\operatorname{div}(X)=\dfrac{\dot\gamma}{\gamma}+\dfrac{\dot\gamma\ddot\gamma}{(1+(\dot\gamma)^2)}$. Now we take $\gamma=e^{-x^\alpha}$, so that $\dot\gamma=-\alpha x^{\alpha-1}e^{-x^\alpha}$,

$$\ddot\gamma=\alpha^2(x^{2(\alpha-1)})e^{-x^\alpha}-(\alpha)(\alpha-1)x^{\alpha-2}e^{-x^\alpha}.$$

It is easily seen that $\dfrac{\dot\gamma}{\gamma}=-\alpha x^{\alpha-1}$, while $\dfrac{\dot\gamma\ddot\gamma}{(1+(\dot\gamma)^2)}\to 0$ as $x\to\infty$, so that $|\operatorname{div}(X)|\to\infty$ as $x\to\infty$ if and only if $\alpha>1$.

It follows that $h^{\mathrm{ess}}=\infty$ if $\alpha>1$, and $h^{\mathrm{ess}}\geqq 1$ if $\alpha=1$. Combining this with the calculation above for μ_f, we see that $\lambda_0^{\mathrm{ess}}=\infty$ for $\alpha>1$, $\lambda_0^{\mathrm{ess}}=\frac{1}{4}$ for $\alpha=1$, and $\lambda_0^{\mathrm{ess}}=0$ for $0<\alpha<1$. This establishes Theorem 2.

§ 3. Some Observations

In general, it is possible to have h^{ess} strictly less than μ_f, so that Theorem 1 is not always sharp. For instance, we may have $\gamma(x)$ containing sharp grooves, spaced far apart, so that h^{ess} is lowered, but μ_f is unaffected, as in Fig. 2.

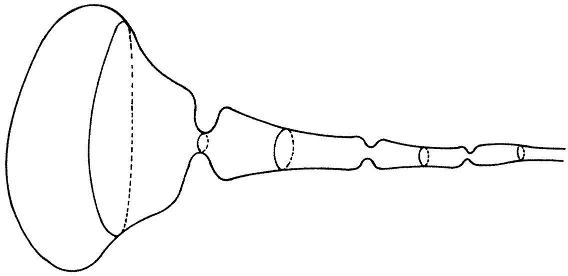


Fig. 2

However, the discussion of §2 gives the following result.

Theorem 4. *Suppose that $\dot{\gamma}$ and $\ddot{\gamma}$ tend to 0 as $x \to \infty$, while $\frac{\dot{\gamma}}{\gamma} \to -k$ as $x \to \infty$ (k may be ∞). Then $\lambda_0^{\text{ess}}(M_\gamma) = \frac{1}{4}k^2$.*

Proof. The calculations of §2 give that if $\overline{\lim}\,\frac{\dot{\gamma}}{\gamma} \leqq -k$, and $\dot{\gamma}, \ddot{\gamma}$ are as above, then

$$h^{\text{ess}} \geqq \varliminf_{x \to \infty} \frac{\inf|\operatorname{div}(x)|}{\sup\|x\|} = \varliminf \frac{\left|\frac{\dot{\gamma}}{\gamma} + \frac{\dot{\gamma}\ddot{\gamma}}{(1+(\dot{\gamma})^2)}\right|}{(1+(\dot{\gamma})^2)^{1/2}} \geqq k.$$

If $\varliminf \frac{\dot{\gamma}}{\gamma} \geqq -k$, we may integrate to get

$$\gamma \geqq (x)(\text{const})\, e^{-kx}$$

so that

$$\int_x^\infty \gamma \geqq (\text{const}) \left(\frac{1}{k}\right) e^{-kx}$$

and hence

$$\mu_f \leqq k.$$

Combining these inequalities with Theorem 1 gives Theorem 4.

References

1. Agmon, S.: Lectures on exponential decay of solutions of second-order elliptic equations: bounds on eigenfunctions of n-body Schrödinger operators. Princeton, N.J.: Princeton Univ. Press 1982

2. Baider, A.: Noncompact manifolds with discrete spectra. J. Diff. Geom. **14**, 41–57 (1979)
3. Brooks, R.: A relation between growth and the spectrum of the Laplacian. Math. Z. **178**, 501–508 (1981)
4. Donnelly, H.: On the essential spectrum of a complete riemannian manifold. Topology **20**, 1–14 (1981)
5. Donnelly, H., Li, P.: Pure point spectrum and negative curvature for non-compact manifolds. Duke Math. J. **46**, 497–503 (1979)

Received July 26, 1983

Math. Z. 187, 433–452 (1984)

Mathematische Zeitschrift

Compact Dupin Hypersurfaces with Three Principal Curvatures

Reiko Miyaoka*

Department of Mathematics, Tokyo Institute of Technology, 152 Tokyo, Japan

§ 0. Introduction

The purpose of this paper is to prove that every compact embedded Dupin hypersurface in a space form with three principal curvatures is the Lie-geometric image of an isoparametric hypersurface.

Here by a Dupin hypersurface, we mean an immersed hypersurface in a space form with the property that the multiplicities of the principal curvatures are constant and that each principal curvature is constant along its leaf of the curvature distribution.

Dupin hypersurfaces are first investigated by Cecil and Ryan [3], while the name was given later by Pinkall [6] in a slightly more general situation. Pinkall found a systematic way to treat Dupin hypersurfaces using Lie-geometry. The idea occurred to him since the class of Dupin hypersurfaces is invariant under Lie-transformations, which are transformations of the space of all oriented hyperspheres preserving contact of two hyperspheres. The group of Lie-transformations is bigger than the group of conformal transformations. Pinkall classified Dupin hypersurfaces with three different principal curvatures in E^4 up to Lie-equivalence. But his argument seems difficult to be developed into higher dimensional cases.

On the other hand, Thorbergsson got a nice global result saying that complete embedded Dupin hypersurfaces in a space form are taut [7]. He has got some more topological facts about Dupin hypersurfaces which suggest a close relation between compact Dupin hypersurfaces and isoparametric ones.

In our proof, we use Thorbergsson's results essentially and the argument is completely different from Pinkall's and Cecil and Ryan's. The latter two proved a similar result in case of two principal curvatures. In this case, and maybe in more general cases, our method will be valid.

The author thanks Professor Thorbergsson for his valuable suggestion, especially about Fact 3.2 in § 3. She thanks also Professor Pinkall for his pointing out her failure in the first preprint. She has to mention finally Professor Cecil's kind introduction to the works of Pinkall and Thorbergsson.

---

* This work was done while the author was staying at the University of Warwick, England

§ 1. Preliminaries

Let S^{n+1} be the $(n+1)$-dimensional unit sphere in E^{n+2} with center at the origin. Let D and ∇ be the riemannian connections of E^{n+2} and S^{n+1}, respectively. Consider an isometrically immersed orientable hypersurface $f\colon M \to S^{n+1}$ with a unit normal vector field ξ. Let A be the second fundamental tensor of the immersion f and $\lambda_1 \leqq \lambda_2 \leqq \dots \leqq \lambda_n$ be the principal curvaturs of M. For $\lambda \in \{\lambda_i\}$, the curvature distribution $\mathscr{D}(\lambda)$ is defined by

$$\mathscr{D}_p(\lambda) = \{X \in T_p M\colon AX = \lambda(p)X\}, \qquad p \in M.$$

The following is well-known [4]. (We use p for $f(p)$ if it is clear.)

Fact 1.1. When $\dim \mathscr{D}(\lambda)$ is constant $(=k)$ on M, $\mathscr{D}(\lambda)$ is an involutive distribution on M. Moreover if λ is constant along a leaf L of $\mathscr{D}_p(\lambda)$ (this is the case when $k>1$), then L is locally a k-sphere through p, whose center c is given by

$$c = p + \frac{H}{|H|^2}, \tag{1.1}$$

where H is the mean curvature vector of the immersion $L \to E^{n+2}$.

Definition 1.2. When the multiplicities of the principal curvatures are constant and when each principal curvature is constant along its leaf of the curvature distribution, we call M a *Dupin hypersurface.*

Remark 1.3. Dupin hypersurfaces in a space form are similarly defined. In this paper, we will mainly discuss the spherical case, since Dupin hypersurfaces in E^{n+1} and in H^{n+1} correspond to those in S^{n+1} via stereographic projections.

Examples of Dupin hypersurfaces

(1) Isoparametric hypersurfaces.

(2) Conformal images of isoparametric hypersurfaces (e.g. cyclides of Dupin).

(3) Lie-geometric images of isoparametric hypersurfaces.

(4) A part of a tube of an extrinsically symmetric submanifold in a space form.

Remark 1.4. For every n, there are Dupin hypersurfaces with any prescribed multiplicities of principal curvatures.

When M is a Dupin hypersurface, we can choose a local orthonormal frame $(e_1, \dots, e_n, p)$ of M such that each e_i is a unit principal vector for λ_i. Then we have

$$D_{e_i} e_j = \sum_{k=1}^{n} \Lambda_{ij}^k e_k + \lambda_i \delta_{ij} \xi - p,$$

where $1 \leqq i, j \leqq n$. Let $[i] = \{j\colon \lambda_i = \lambda_j\}$. By Codazzi's equation, we have

$$e_j(\lambda_i) = \Lambda_{ii}^j (\lambda_i - \lambda_j), \qquad \text{for } j \notin [i].$$

Thus we obtain:

Fact 1.5. (i) $\Lambda_{ii}^k=\Lambda_{jj}^k$ if $j\in[i]$ and $k\notin[i]$.

(ii) M is isoparametric if and only if $\Lambda_{ii}^j=0$ for all i and j such that $j\notin[i]$.

When $\lambda=\lambda_i$, (1.1) is written as

$$c_i=p+\frac{H_i}{|H_i|^2},\qquad H_i=\sum_{k\notin[i]}\Lambda_{ii}^k e_k+\lambda_i\xi-p. \tag{1.2}$$

From (i) above, we note $H_i=H_j$ for $j\in[i]$.

Now, the focal points of $p\in M$ are given by

$$f_p^i=\cos\theta_i p+\sin\theta_i\xi_p,\qquad 0<\theta_i=\cot^{-1}\lambda_i<\pi \tag{1.3}$$

and by their antipodals $\overline{f_p^i}$, $1\leqq i\leqq n$. For $x\in S^{n+1}$, let $L_x\colon M\to\mathbb{R}$ be the distance function on M, i.e. $L_x(p)=d(x,p)^2$, where $d(x,\ )$ is the distance from x on S^{n+1}. Let $p\in M$ be a non-degenerate critical point of L_x. Then the index of p is given by the sum of the multiplicities of the focal points of p on $\widehat{xp}$, where $\widehat{xp}$ denotes the shorter geodesic segment joining x and p. When $p=\bar{x}=$ the antipodal of x, it is obvious that the index of p is n. If we project S^{n+1} stereographically onto E^{n+1} from $\bar{x}$ and calculate indices of the euclidean distance function $|x-\ |^2$ (we can move x very slightly so that $\bar{x}\notin M$, if necessary), we will see that Morse theory is applicable to the spherical L_x with above indices.

§2. A New Characterization of Isoparametric Hypersurfaces

Let M be a Dupin hypersurface in S^{n+1}. Using the notation in §1, let L_p^i be the leaf of $\mathscr{D}(\lambda_i)$ at p.

Definition 2.1. For $j\notin[i]$, we say that there exists *a common normal circle γ for L_p^i and L_p^j*, if there exists a circle γ of S^{n+1} through p satisfying

(i) γ is normal to M at p.

(ii) $\gamma\cap L_p^i=\{p,p_i\}$ for $p_i\neq p$, and at p_i, γ is normal to M.

(iii) $\gamma\cap L_p^j=\{p,p_j\}$ for $p_j\neq p$, and at p_j, γ is normal to M.

When γ is a great circle, we call γ a *common normal geodesic for L_p^i and L_p^j*. When γ is common for any two leaves at p, we call γ *a common normal circle at p*.

Lemma 2.2. *For $j\notin[i]$, there exists a common normal circle γ for L_p^i and L_p^j if and only if*

$$\Lambda_{ii}^k(p)=\Lambda_{jj}^k(p)\qquad \textit{for all } k\notin[i]\cup[j]. \tag{2.1}$$

Proof. We will first show the necessity.

Choosing a suitable $e_i\in\mathscr{D}_p(\lambda_i)$, we can write $p_i\in L_p^i$ as

$$p_i=c_i+\cos\mu_1(p-c_i)+|p-c_i|\sin\mu_1 e_i,\qquad 0\leqq\mu_1<2\pi, \tag{2.2}$$

where c_i is given by (1.2). Hence we have

$$\begin{aligned}p_i=&\{1-K_i^2(1-\cos\mu_1)\}p+K_i^2\lambda_i(1-\cos\mu_1)\xi_p\\&+K_i^2(1-\cos\mu_1)\sum_{k\notin[i]}\Lambda_{ii}^k e_k+K_i\sin\mu_1 e_i,\end{aligned} \tag{2.2$'$}$$

where $K_i=\frac{1}{|H_i|}$. Let q be the center of γ in E^{n+2}. Since γ lies on the plane π_γ spanned by $q-p$ and ξ_p, we may put

$$q-p_i=a(q-p)+b\,\xi_p. \tag{2.3}$$

From

$$\langle q-p, \xi_p\rangle=0=\langle p, \xi_p\rangle,$$

we get

$$b=-\langle p_i, \xi_p\rangle=-K_i^2\,\lambda_i(1-\cos\mu_1).$$

From

$$\left\langle q-\frac{p+p_i}{2}, p-p_i\right\rangle=0 \quad \text{and} \quad |p|=|p_i|=1,$$

it follows

$$\begin{aligned}0&=\langle q, p-p_i\rangle\\&=\frac{1}{1-a}\{\langle -ap+p_i+b\,\xi_p, p-p_i\rangle\}\\&=\frac{1}{1-a}\{(1+a)\langle p, p_i\rangle-(1+a)+b^2\},\end{aligned}$$

and hence

$$a=\frac{b^2}{1-\langle p, p_i\rangle}-1=K_i^2\,\lambda_i^2(1-\cos\mu_1)-1.$$

Thus we have

$$\begin{aligned}q=(1-a)^{-1}[&\{1-a-K_i^2(1-\cos\mu_1)\}\,p\\&+K_i^2(1-\cos\mu_1)\sum_{k\notin[i]}\Lambda_{ii}^k\,e_k+K_i\sin\mu_1\,e_i].\end{aligned}$$

Note that γ is normal to M at p_i. In fact, from (1.3) and $f_p^i=f_{p_i}^i$, we get

$$\begin{aligned}\xi_{p_i}&=\xi_p+\lambda_i(p-p_i)\\&=\xi_p+\lambda_i\{K_i^2(1-\cos\mu_1)(p-\xi_p-\sum_{k\notin[i]}\Lambda_{ii}^k\,e_k)-K_i\sin\mu_1\,e_i\}\\&=\{1-K_i^2\,\lambda_i(1-\cos\mu_1)\}\,\xi_p-\lambda_i(1-a)(q-p)\\&=-a\,\xi_p-\lambda_i(1-a)(q-p)\in\pi_\gamma.\end{aligned}$$

Then from (2.3) and

$$1=|\xi_{p_i}|^2=a^2+\lambda_i^2(1-a)^2\,|q-p|^2,$$

we obtain

$$\langle q-p_i, \xi_{p_i}\rangle=-\lambda_i a(1-a)\,|q-p|^2-a\,b=-\frac{a(1-a)}{\lambda_i}-a\,b=0.$$

Thus $\dot\gamma(p_i)\in\pi_\gamma$ must be parallel to ξ_{p_i}.

Remark 2.3. Here we get a circle γ which cuts L^i_p at p and p_i orthogonally. For later use, note that γ is a great circle if and only if $\mu_1=\pi$ and $\Lambda^k_{ii}(p)=0$ for all $k\notin[i]$. Since $\sum_{k\notin[i]}(\Lambda^k_{ii})^2=|H_i|^2-\lambda_i^2-1$ is invariant on L^i_p, γ is a great circle if and only if for any $r\in L^i_p$, the normal geodesic γ_r cuts L^i_p twice orthogonally.

Now, expressing $p_j\in L^j_p$ as

$$p_j=c_j+\cos\mu_2(p-c_j)+|p-c_j|\sin\mu_2\, e_j,\qquad 0\leqq\mu_2<2\pi, \tag{2.2)''$$

for a suitable $e_j\in\mathscr{D}_p(\lambda_j)$, we get

$$\begin{aligned} q=(1-a')^{-1}[\{1-a'-K_j^2(1-\cos\mu_2)\}\,p \\ +K_j^2(1-\cos\mu_2)\sum_{k\notin[j]}\Lambda^k_{jj}e_k+K_j\sin\mu_2\, e_j], \end{aligned} \tag{2.5}$$

where $a'=K_j^2\lambda_j^2(1-\cos\mu_2)-1$. Comparing each coefficient in (2.4) and (2.5), we have

$$\frac{K_i^2(1-\cos\mu_1)}{1-a}=\frac{K_j^2(1-\cos\mu_2)}{1-a'}, \tag{2.6}$$

$$\Lambda^k_{ii}=\Lambda^k_{jj}\qquad\text{for all } k\notin[i]\cup[j], \tag{2.7}$$

$$\frac{K_i^2(1-\cos\mu_1)\Lambda^j_{ii}}{1-a}=\frac{K_j\sin\mu_2}{1-a'}, \tag{2.8}$$

$$\frac{K_j^2(1-\cos\mu_2)\Lambda^i_{jj}}{1-\mathrm{a}'}=\frac{K_i\sin\mu_1}{1-a}, \tag{2.9}$$

$$K_i^2(1-\cos\mu_1)\Lambda^{j'}_{ii}=0\qquad\text{for all } j\neq j'\in[j]; \tag{2.10}$$

$$K_j^2(1-\cos\mu_2)\Lambda^{i'}_{jj}=0\qquad\text{for all } i\neq i'\in[i]. \tag{2.11}$$

From (2.6) and (2.8), we get

$$\sin\mu_2=K_j(1-\cos\mu_2)\Lambda^j_{ii}.$$

Squaring the both sides and cancelling $(1-\cos\mu_2)$, we get

$$1+\cos\mu_2=K_j^2(1-\cos\mu_2)(\Lambda^j_{ii})^2.$$

Therefore μ_2 must satisfy

$$\cos\mu_2=\frac{K_j^2(\Lambda^j_{ii})^2-1}{K_j^2(\Lambda^j_{ii})^2+1}. \tag{2.12}$$

Obviously, a similar formula holds for μ_1. As a matter of fact, it is easy to see that

$$-1\leqq\frac{K_j^2(\Lambda^j_{ii})^2-1}{K_j^2(\Lambda^j_{ii})^2+1}<1\quad\text{and}\quad -1\leqq\frac{K_i^2(\Lambda^i_{jj})^2-1}{K_i^2(\Lambda^i_{jj})^2+1}<1. \tag{2.13}$$

Now, we will see the sufficiency of the condition (2.7) for the existence of a common normal circle for L^i_p and L^j_p.

Choose $e_i\in\mathscr{D}_p(\lambda_i)$ as a unit vector in the direction of $\mathscr{D}_p(\lambda_i)$-component of D_XX for any vector $X\in\mathscr{D}_p(\lambda_i)$. This is independent of the choice of X by

Fact 1.2. (i). Choosing e_j similarly, i.e. in the direction of $\mathscr{D}_p(\lambda_j)$-component of $D_Y Y$ for $Y \in \mathscr{D}_p(\lambda_i)$, we obtain

$$\Lambda_{ii}^{j'}=0 \quad \text{for } j \neq j' \in [j] \quad \text{and} \quad \Lambda_{jj}^{i'}=0 \quad \text{for } i \neq i' \in [i].$$

From (2.13), we can define μ_1 and μ_2 up to mod π by

$$\cos\mu_1=\frac{K_i^2(\Lambda_{jj}^i)^2-1}{K_i^2(\Lambda_{jj}^i)^2+1} \quad \text{and} \quad \cos\mu_2=\frac{K_j^2(\Lambda_{ii}^j)^2-1}{K_j^2(\Lambda_{ii}^j)^2+1}.$$

Note that $\mu_i \neq 0 \bmod 2\pi$, $i=1,2$, and hence p_i and p_j defined by (2.2) and (2.2)″ are different from p. Now, it is easy to obtain

$$\sin^2\mu_1=K_i^2(1-\cos\mu_1)^2(\Lambda_{jj}^i)^2,$$
$$\sin^2\mu_2=K_j^2(1-\cos\mu_2)^2(\Lambda_{ii}^j)^2.$$

Determine μ_1 and μ_2 up to mod 2π so that

$$\sin\mu_1=K_i(1-\cos\mu_1)\Lambda_{jj}^i \text{ and } \sin\mu_2=K_j(1-\cos\mu_2)\Lambda_{ii}^j.$$

Now, since we have

$$\begin{aligned}\frac{K_i^2(1-\cos\mu_1)}{1-a} &= \frac{K_i^2(1-\cos\mu_1)}{2-K_i^2\lambda_i^2(1-\cos\mu_1)}\\ &= \frac{2K_i^2\{K_i^2(\Lambda_{jj}^i)^2+1\}^{-1}}{2-2K_i^2\lambda_i^2\{K_i^2(\Lambda_{jj}^i)^2+1\}^{-1}}\\ &= \frac{K_i^2}{K_i^2\{(\Lambda_{jj}^i)^2-\lambda_i^2\}+1}\\ &\underset{(1.2)}{=} \{(\Lambda_{jj}^i)^2-\lambda_i^2+\sum_{k\notin[i]}(\Lambda_{ii}^k)^2+\lambda_i^2+1\}^{-1}\\ &= \{(\Lambda_{jj}^i)^2+(\Lambda_{ii}^j)^2+\sum_{k\notin[i]\cup[j]}(\Lambda_{ii}^k)^2+1\}^{-1},\end{aligned} \tag{2.14}$$

(2.6) follows from the assumption (2.7). Then it is easy to see (2.8) and (2.9). Thus the proof of the lemma is completed.

Now, since

$$\begin{aligned}|q-p|^2=\frac{b^2}{1-a^2}&=\frac{K_i^2(1-\cos\mu_1)}{1-a}\,\frac{K_i^2\lambda_i^2(1-\cos\mu_1)}{1+a}\\ &=\{(\Lambda_{jj}^i)^2+(\Lambda_{ii}^j)^2+\sum_{k\notin[i]\cup[j]}(\Lambda_{ii}^k)^2+1\}^{-1},\end{aligned}$$

by (2.14), γ is a great circle of S^{n+1} if and only if

$$\Lambda_{ii}^k=0 \text{ for all } k\notin[i] \quad \text{and} \quad \Lambda_{jj}^k=0 \text{ for all } k\notin[j].$$

An easy consequence of this fact is the following:

Proposition 2.4. *A Dupin hypersurface M in S^{n+1} is isoparametric if and only if there exists a common normal geodesic at any $p\in M$.*

Corollary 2.5. *Let M be a conformal image of an isoparametric hypersurface in S^{n+1}. Then at any $p \in M$, there exists a common normal circle, or equivalently, Λ^j_{ii} depends only on j if $j \notin [i]$.*

§3. Some Differential Topological Facts on Dupin Hypersurfaces

Recently, Thorbergsson has obtained remarkable facts on topological structures of Dupin hypersurfaces [7].

Fact 3.1. Complete embedded Dupin hypersurfaces in a simply connected space form are taut with respect to $\mathbb{Z}_2$ coefficients.

Fact 3.2. A compact embedded Dupin hypersurface M satisfies $\dim H^*(M,\mathbb{Z}_2) = 2h$, where h is the number of principal curvatures.

Fact 3.3. A compact embedded Dupin hypersurface M divides S^{n+1} into two ball bundles. In particular, $h=1,2,3,4,6$.

A proper immersion $f\colon M \to \bar{M}$ (of a manifold without boundary) into a space form is called taut if there is a field $\mathbb{F}$ such that every distance function of a point $x \in \bar{M}$, $L_x\colon M \to \mathbb{R}$, $p \to d(x,f(p))^2$, which is a Morse function has the minimal number of critical points required by the Morse inequalities with respect to $\mathbb{F}$. We refer to [1] for basic properties of taut submanifolds.

Thorbergsson proved Fact 3.1 by showing that $H_k(M^{\kappa_i}) \to H_k(M^{\kappa_i}, M^{\kappa_{i-1}})$ is surjective, where $M^\kappa = \{p \in M\colon L_x(p) \leqq \kappa\}$ and $\kappa_1 < \ldots < \kappa_l$ are the critical values of a Morse function L_x. In the proof, he constructs an iterated sphere bundle N_i as a representative of the homology class of $(M^{\kappa_i}, M^{\kappa_{i-1}})$. By tautness, we know that $[h_i(N_i)]$, $i=1,\ldots,l$ (using his notation) exhaust the homology classes of M.

The proof of Fact 3.2 is given to the author by a private letter. Let $\lambda_1 < \ldots < \lambda_h$ be the principal curvatures with multiplicities $m_1, \ldots, m_h$, respectively. Note that a point $p \in M$ is a critical point of L_x iff $x \in \gamma_p =$ the normal geodesic at p. The possible indices of p are 0, $i_\alpha = \sum_{j=1}^{\alpha} m_j$, $i'_\alpha = \sum_{j=1}^{\alpha} m_{h-j+1}$, $1 \leqq \alpha \leqq h-1$, and n. For each index, construct N_* based on p. When p_1 and p_2 are different points of M, we can join them by a curve in M and construct N_*'s at every point of the curve to obtain a homotopy between cycles based at p_1 and p_2. Thus we have $\dim H^*(M,\mathbb{Z}_2) \leqq 2h$.

Next, we insist that two cycles at the point p can not be homologous, which shows $\dim H^*(M,\mathbb{Z}_2) = 2h$. First, all the cycles corresponding to indices i_α are obviously non-homologous. The same is true for those corresponding to i'_α. Let z_1 and z_2 be cycles with respect to different types of indices i_α and i'_α. This means that for the corresponding $x_i \in \gamma_p$, $i=1,2$, $\widehat{px_1}$ and $\widehat{px_2}$ are in different normal directions at p. Suppose z_1 is homologous to z_2. Then by moving x_1 to x_2 outside of the focal submanifolds (this is possible since the codimensions of focal submanifolds $\geqq 2$ [2]), the critical point with respect to x_1 moves with the corresponding normal vector, giving a curve of normal vector connecting the positive and negative normals at p, which is impossible.

Remark 3.4. For $p_0 \in M$, let $c_0 = c'_0 = p_0$, $c_h = c'_h = M$, and for $1 \leqq \alpha \leqq h-1$,

$$c_\alpha = \bigcup_{p_1 \in L^\alpha_{p_0}} \bigcup_{p_2 \in L^{\alpha-1}_{p_1}} \cdots \bigcup_{p_{\alpha-1} \in L^2_{p_{\alpha-2}}} L^1_{p_{\alpha-1}},$$

$$c'_\alpha = \bigcup_{p_1 \in L^\alpha_{p_0}} \bigcup_{p_2 \in L^{\alpha+1}_{p_1}} \cdots \bigcup_{p_{h-\alpha} \in L^{h-1}_{p_{h-\alpha-1}}} L^h_{p_{h-\alpha-1}}.$$

Then from the way of construction of $h_i(N_i)$, we see that c_α (c'_α, resp.) is a cycle representing the homology class of $H_{i_\alpha}(M)$ ($H_{i'_\alpha}(M)$, resp.), where $0 \leqq \alpha \leqq h$.

The ball bundles in Fact 3.3 are given by $\pi_*\colon C_* \to F_*$, $* = +, -$, where F_+ (resp. F_-) is the set of first focal points in the positive (resp. negative) normal direction. Denoting $f_p^+ = f_p^{\lambda_h}$ and $f_p^- = \overline{f_p^{\lambda_1}}$ (see (1.3)), $F_+ = \{f_p^+, p \in M\}$ and $F_- = \{f_p^-, p \in M\}$. C_* is given by $C_* = \bigcup_{p \in M} \widehat{pf_p^*}$ and π_* is defined by $\pi_*(\widehat{pf_p^*}) = f_p^*$.

Applying Münzner's topological theorem on isoparametric hypersurfaces [5], Thorbergsson proved $h = 1, 2, 3, 4, 6$.

For later use we give some notation. For a principal curvature λ, we will denote $\bar{\lambda} := \cot^{-1}\lambda$, $0 < \bar{\lambda} < \pi$. Note that for p and its focal point $f_p^\lambda = \cos\bar{\lambda} p + \sin\bar{\lambda}\xi_p$, $d(p, f_p^\lambda) = \bar{\lambda}$ and $d(p, \overline{f_p^\lambda}) = \pi - \bar{\lambda}$.

For $y \in S^{n+1}$ and $r > 0$, let $B(y, r) = \{x \in S^{n+1}\colon d(x, y) \leqq r\}$ and $S(y, r) = \partial B(y, r)$. Especially, we denote $B_p^\lambda = B(f_p^\lambda, \bar{\lambda})$ and $S_p^\lambda = \partial B_p^\lambda$. Obviously, $\overline{(B_p^\lambda)^c} = B(\overline{f_p^\lambda}, \pi - \bar{\lambda})$ and $\partial\overline{(B_p^\lambda)^c} = S_p^\lambda$. Finally, let $B_p^+ = B_p^{\lambda_h}$, $B_p^- = \overline{(B_p^{\lambda_1})^c}$, $S_p^+ = \partial B_p^+$ and $S_p^- = \partial B_p^-$.

Lemma 3.5. *For any $p \in M$, we have $B_p^+ \cap M = L_p^h$ and $B_p^- \cap M = L_p^1$.*

Proof. Let $q \in B_p^+ \cap M$. From tautness, it follows $M \subset \overline{(B_p^+)^c}$ and so $q \in S_p^+$. Therefore we have $T_q M = T_q S_p^+$ or, the normal geodesic at q passes through f_p^+. The first focal point of q on γ_q in the same direction as f_p^+ is either f_q^+ or f_q^-. By tautness again, $S_q^+ \subset \overline{(B_p^+)^c}$ ($S_q^- \subset \overline{(B_p^+)^c}$, resp.) implies

$$\overline{\lambda_h(q)} \geqq \overline{\lambda_h(p)} \qquad (\pi - \overline{\lambda_1(q)} \geqq \overline{\lambda_h(p)}, \text{ resp.})$$

and $S_p^+ \subset \overline{(B_q^+)^c}$ ($S_p^+ \subset \overline{(B_q^-)^c}$, resp.) implies

$$\overline{\lambda_h(p)} \geqq \overline{\lambda_h(q)} \qquad (\overline{\lambda_h(p)} \geqq \pi - \overline{\lambda_1(q)}, \text{ resp.}).$$

Since $f_p^+ = f_q^-$ can not happen by tautness, we get $f_p^+ = f_q^+$ and $L_p^h \cup L_q^h \subset S_p^+$. From tautness and from λ_h-leaves are homologous in M, L_p^h and L_q^h should be homologous in $M \cap B_p^+$. Thus we have $L_p^h = L_q^h$.

§4. Dupin Hypersurfaces with Three Principal Curvatures

Let M be a Dupin hypersurface with three principal curvatures $\lambda < \mu < \nu$. Consider a function φ on M given by

$$\varphi = \log\frac{\mu - \lambda}{\nu - \mu}.$$

Remark 4.1. φ is conformally invariant. In particular, if M is a conformal image of an isoparametric hypersurface, then φ is constant.

Let $e_i \in \mathscr{D}(\lambda)$, $e_j \in \mathscr{D}(\mu)$ and $e_k \in \mathscr{D}(\nu)$. Then we have

$$e_i(\varphi) = \frac{\nu-\lambda}{\nu-\mu}(\Lambda^i_{jj} - \Lambda^i_{kk}),$$

$$e_j(\varphi) = \Lambda^j_{ii} - \Lambda^j_{kk},$$

$$e_k(\varphi) = \frac{\mu-\nu}{\mu-\lambda}(\Lambda^k_{jj} - \Lambda^k_{ii}).$$

Now, let M be compact. Then there exists a point $p_0 \in M$ such that $\nabla\varphi = 0$ at p_0, or equivalently,

$$\Lambda^k_{ii} = \Lambda^k_{jj}, \qquad \Lambda^i_{jj} = \Lambda^i_{kk}, \qquad \Lambda^j_{kk} = \Lambda^j_{ii} \quad \text{at } p_0. \tag{4.1}$$

Via Lemma 2.2, there exists a common normal circle γ_0 at p_0. By a conformal transformation of S^{n+1}, we may assume γ_0 to be a great circle.

When M is embedded, beginning with this common normal geodesic γ_0, we will find Lie-transformations of S^{n+1} in several steps so that the image hypersurface has a common normal geodesic at any point (Then the main theorem follows from Proposition 2.4.). In the first step, we will show that γ_0 is a common normal geodesic at all points of $\gamma_0 \cap M$. Second, we will determine the space $\mathscr{D}_q(*)$ for any $q \in \gamma_0 \cap M$. Then, after some more consideration in § 5, we will get the main theorem.

Step 1. Let p_0 and γ_0 be as above. Let L^λ_0, L^μ_0, L^ν_0 be the three leaves through p_0, and let $L^\lambda_0 \cap \gamma_0 = \{p_0, p_1\}$, $L^\mu_0 \cap \gamma_0 = \{p_0, p_2\}$ and $L^\nu_0 \cap \gamma_0 = \{p_0, p_3\}$. The situation is shown in Fig. 4.1. The location of focal points $f_i^* = f_{p_i}^*$ where $* = +, -$ is suggested by Fact 3.3. Note the position of f_2^*, which are continuously determined along L^μ_0. From Fact 3.3, there exist points $q \in M$ on $\widehat{f_2^+ f_3^-}$ (= the segment not containing p_2) and $r \in M$ on $\widehat{f_2^- f_1^+}$ ($\not\ni p_2$).

Lemma 4.2. $d(f_2^+, f_3^-) = \overline{\nu(p_2)} + \overline{\lambda(p_3)}$, $d(f_1^+, f_2^-) = \overline{\nu(p_1)} + \overline{\lambda(p_2)}$.

Proof. Since $q \in \overline{(B_3^-)^c} \cap \overline{(B_2^+)^c}$, we have immediately $d(f_2^+, f_3^-) \geqq \overline{\nu(p_2)} + \overline{\lambda(p_3)}$. Suppose $d(f_2^+, f_3^-) > \overline{\nu(p_2)} + \overline{\lambda(p_3)}$. Then since $B_2^+ \cap B_3^- = \varnothing$, we can transform S^{n+1} conformally such that $f_3^- = \overline{f_2^+}$ (see Fig. 4.2).

Let $x \in \widehat{p_2 f_2^+}$ sufficiently near f_2^+ but not equal to f_2^+. Then L_x is a Morse function on M. Since $B = B(x, \overline{\nu(p_2)} + 2\overline{\lambda(p_2)} + \varepsilon)$ contains $c_1(p_2) = L^\lambda_{p_2}$ and $c_1'(p_2) = L^\nu_{p_2}$ for small $\varepsilon > 0$, L_x has critical points of indices 0, i_1 and i_1' in B, by Remark 3.4. On the other hand, since L_x has critical points p_0, p_1 and p_3 with indices $< n$ in B^c, and the one with n in $B(\bar{x}, \overline{\lambda(p_3)} + \varepsilon) \subset B^c$, L_x has seven critical points on M, a contradiction. Similarly or more easily, $d(f_1^+, f_2^-) = \overline{\nu(p_1)} + \overline{\lambda(p_2)}$ is proved.

From this lemma, we get $q \in B_3^- \cap B_2^+$, $r \in B_1^+ \cap B_2^-$. Thus by Lemma 3.5, we obtain

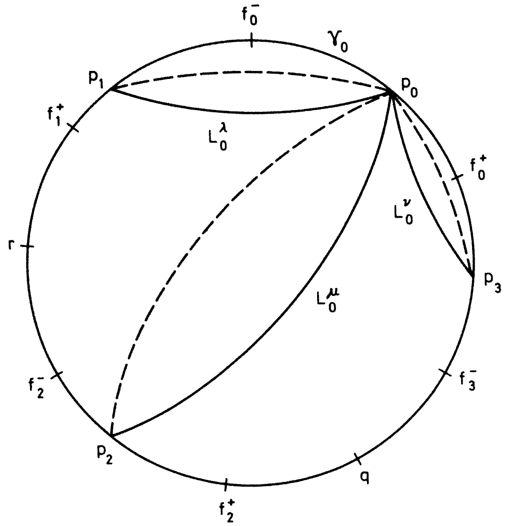


Fig. 4.1

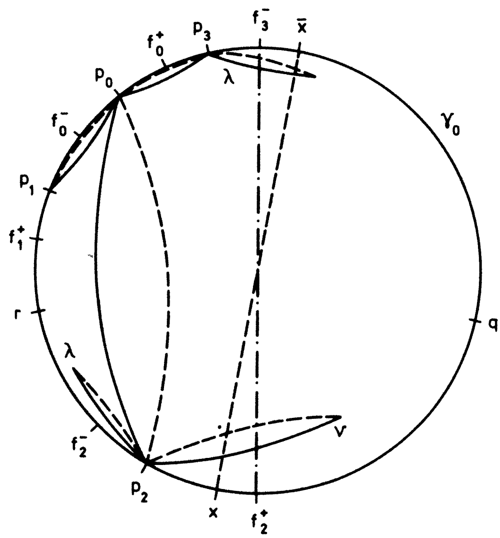


Fig. 4.2

Proposition 4.3. $L_q^\lambda = L_3^\lambda$, $L_q^\nu = L_2^\nu$, $L_r^\lambda = L_2^\lambda$, $L_r^\nu = L_1^\nu$.

To investigate the μ-leaf at each point on $\gamma_0 \cap M$, we will do some preparation.

Let $p_0 \in M$. By Remark 3.4, the representatives of $H_*(M)$ are given by $c_0 = p_0$, $c_1 = L_{p_0}^\lambda$, $c_1' = L_{p_0}^\nu$, $c_2 = \bigcup_{p \in L_{p_0}^\mu} L_p^\lambda$, $c_2' = \bigcup_{p \in L_{p_0}^\mu} L_p^\nu$ and $c_3 = M$.

Lemma 4.4. *For any $p \in M$, we have*

$$L^{\lambda}_p \cap (c_2 \cup c'_2) \neq \emptyset \quad \text{and} \quad L^{\nu}_p \cap (c_2 \cup c'_2) \neq \emptyset.$$

Proof. Suppose $L^{\lambda}_p \cap (c_2 \cup c'_2) = \emptyset$. Let $x \in S^{n+1}$ be sufficiently near f_p^{λ} such that L_x is a Morse function. Since c_2 and c'_2 are closed, we can choose (x and) $\varepsilon > 0$ so that

$$B^{\lambda}_p \subset B := B(x, \overline{\lambda(p)} + \varepsilon) \quad \text{and} \quad (c_2 \cup c'_2) \subset B^c.$$

Then, L_x has critical points of indices 0 and i_1 in B and $L_{\bar{x}}$ has critical points of indices 0, i_1, i'_1, i_2, i'_2 in B^c, since $c_2 \supset c_1$ and $c'_2 \supset c'_1$. Thus L_x has seven critical points on M, a contradiction.

More strongly, we have

Proposition 4.5. *For any $p \in M$, we have $L^{\lambda}_p \cap c'_2 \neq \emptyset$ and $L^{\nu}_p \cap c_2 \neq \emptyset$.*

Proof. Suppose $L^{\lambda}_p \cap c'_2 = \emptyset$. Then since $L^{\lambda}_p \cap L^{\mu}_{p_0} = \emptyset$, it follows $L^{\lambda}_p \cap c_2 = \emptyset$, contradicts above lemma. Q.E.D.

Corollary 4.6. *For any μ-leaf L^{μ}, we have*

$$M = \bigcup_{p \in L^{\mu}} \bigcup_{q \in L^{\lambda}_p} L^{\nu}_q = \bigcup_{p \in L^{\mu}} \bigcup_{q \in L^{\nu}_p} L^{\lambda}_q.$$

Now, we are ready to investigate μ-leaves at $\gamma_0 \cap M$.

Proposition 4.7. $L^{\mu}_{p_1} = L^{\mu}_q$, $L^{\mu}_{p_3} = L^{\mu}_r$.

Proof. Let $p_1 = \bar{q}$ by a conformal transformation of S^{n+1}. Let f be the middle point of $\widehat{p_1 q}$ containing p_2 and r (see Fig. 4.3).

Claim 1. $f^{\mu}_{p_1} = f = f^{\mu}_q$.

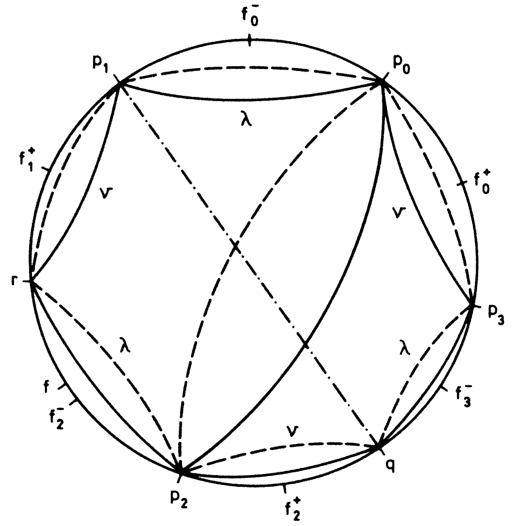
Suppose $f^{\mu}_{p_1} \in \widehat{p_1 f} - \{f\}$. Then $B^{\mu}_{p_1} \cap B^{-}_q = \emptyset$. Since $c'_2(p_1) \subset B^{\mu}_{p_1}$, we have $c'_2(p_1) \cap L^{\lambda}_q = \emptyset$, which contradicts Proposition 4.5. Now, supposing $f^{\mu}_{p_1} \in \widehat{qf} - \{f\}$, we get $c_2(p_1) \cap L^{\nu}_q = \emptyset$, a contradiction. $f_q^{\mu} = f$ will be shown similarly.

Claim 2. $L^{\mu}_{p_1} = L^{\mu}_q$.

Suppose $q \notin L^{\mu}_{p_1}$. Then $c'_2(p_1) \cap B^{-}_q = \emptyset$. This is because if $c'_2(p_1) \cap B^{-}_q \ni s$, then by Lemma 3.5, $L^{\lambda}_s = L^{\lambda}_q$ and hence $c'_2(p_1) \cap L^{\lambda}_q \neq \emptyset$. But from $B^{\mu}_{p_1} \cap B^{-}_q = \{q\}$ and $c'_2(p_1) \subset B^{\mu}_{p_1}$, it follows that $c'_2(p_1) \cap L^{\lambda}_q = \{q\}$. But this means $q \in L^{\mu}_{p_1}$, since $L^{\nu}_q \cap L^{\mu}_{p_1} \neq \emptyset$ induces $L^{\nu}_q \cap L^{\mu}_{p_1} = \{q\}$. Therefore $q \notin L^{\mu}_{p_1}$ implies $c'_2(p_1) \cap B^{-}_q = \emptyset$ or $c'_2(p_1) \cap L^{\lambda}_q = \emptyset$, which contradicts Proposition 4.5. The proof of $L^{\mu}_{p_3} = L^{\mu}_r$ is similar. Q.E.D.

Step 2. In this step, putting $p_4 = q$ and $p_5 = r$, we shall determine $\mathscr{D}_{p_i}(*)$ for $i = 0, \ldots, 5$, up to parallelism with respect to the connection of S^{n+1}. To begin with, let $\mathscr{D}_{p_0}(\lambda) = A_1$, $\mathscr{D}_{p_0}(\mu) = b$ and $\mathscr{D}_{p_0}(\nu) = C_3$.

From now on, as far as treating linear spaces, equality "=" means "be parallel to" with respect to the connection of S^{n+1}. It is obvious that $\mathscr{D}_{p_1}(\lambda) = A_1$, $\mathscr{D}_{p_2}(\mu) = b$ and $\mathscr{D}_{p_3}(\nu) = C_3$. Let $\mathscr{D}_{p_1}(\mu) = c$, $\mathscr{D}_{p_1}(\nu) = B_3$, $\mathscr{D}_{p_3}(\mu) = a$, $\mathscr{D}_{p_3}(\lambda) = B_1$, $\mathscr{D}_{p_2}(\lambda) = C_1$ and $\mathscr{D}_{p_2}(\nu) = A_3$ (see Fig. 4.4). The spaces A_1, B_1 and C_1 are of dimension m_1, a, b and c are of dimension m_2 and A_3, B_3 and C_3 are of



. Fig. 4.3

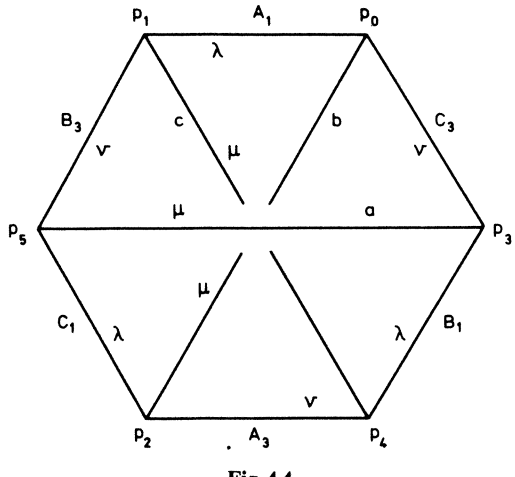


Fig. 4.4

dimension m_3. We note that at each point p_i, the three spaces are mutually orthogonal and span the tangent space $T_{p_i}M$ which is independent of i, since $\xi_{p_i}=\xi_{p_j}$ for $i\neq j$. Put $E=T_{p_i}M$.

On the other hand, the fact that any of μ-leaves at p_0 p_1 and p_3 do not intersect implies $a\cap b=b\cap c=c\cap a=\{0\}$. This is because if, say, $b\cap c\ni X\neq 0$, then considering the 2-sphere S^2 determined by X and ξ, we see that $S^2\cap L^\mu_{p_0}$ and $S^2\cap L^\mu_{p_1}$ have intersection.

Now, in this situation, via pure algebraic consideration, we have

Lemma 4.8. $m_2 \leqq m_1$, m_3. *Moreover there exist mutually orthogonal subspaces A of dimension $m_1 - m_2$ and C of dimension $m_3 - m_2$ in E such that $A_1 = a \oplus A$, $A_3 = a \oplus C$, $B_1 = b \oplus A$, $B_3 = b \oplus C$, $C_1 = c \oplus A$ and $C_3 = c \oplus C$.*

Proof. Let $\pi_1 = A_1 + b$, the space spanned by A_1 and b, and $\pi_2 = A_1 + c$, similarly. From $b \cap c = \{0\}$, it follows that $\dim(b+c) = 2m_2$. Thus $\dim(\pi_1 + \pi_2) = m_1 + 2m_2$ and since $n = m_1 + m_2 + m_3 \geqq m_1 + 2m_2$, we obtain $m_3 \geqq m_2$. Similarly, we have $m_1 \geqq m_2$.

On the other hand, from $A_1 \subset \pi_1 \cap \pi_2$ and

$$\dim(\pi_1 \cap \pi_2) = 2(m_1 + m_2) - (m_1 + 2m_2) = m_1,$$

we obtain

$$\pi_1 \cap \pi_2 = A_1.$$

Now, from $a \perp (B_3 + C_3)$, it follows that

$$a \subset \pi_1 \cap \pi_2 = A_1.$$

Therefore we decompose A_1 as

$$A_1 = a \oplus A, \quad \text{(orthogonal direct sum)}$$

and similarly

$$C_3 = c \oplus C$$

where A is an $(m_1 - m_2)$-dimensional, and C is an $(m_3 - m_2)$-dimensional subspace of E. Since $E = A_1 \oplus b \oplus C_3$, we have

$$E = a \oplus b \oplus c \oplus A \oplus C.$$

Now, the lemma follows easily.

When $n = 3$, we must have $A_1 = a = A_3$, $B_1 = b = B_3$ and $C_1 = c = C_3$. In fact, this is true for $n > 3$ as is shown below.

Proposition 4.9. *For any n, we have $m_1 = m_2 = m_3$ and $A_1 = a = A_3$, $B_1 = b = B_3$ and $C_1 = c = C_3$.*

Proof. Suppose $A \ni X \neq 0$. Note that we denote by X the parallel vector field along γ_0. Consider the great 2-sphere S^2, containing γ_0, whose tangent space along γ_0 is spanned by X and ξ. Let $l_0 = S_2 \cap L_0^\lambda$ and $l_2 = S^2 \cap L_2^\lambda$. Then there is a one-parameter family of small circles of S^2, any element of which cuts l_0 and l_2 twice for each orthogonally. To see this, transform S^{n+1} conformally so that f_0^- and f_2^- be antipodal. Then the family consists of great circles in S^2 joining f_0^- and f_2^-. Note that at each point of l_i, $i = 0, 2$, the normal vector to l_i in S^2 is the normal vector to M in S^{n+1} at that point. Now, let γ be any of these great circles and let $\{s, t\} = l_0 \cap \gamma$ and $\{u, v\} = l_2 \cap \gamma$ as in Fig. 4.5. Using tautness, we will prove:

Lemma 4.10. *Each pair of (s, v) and (t, u) is joined by either a μ-leaf or a ν-leaf.*

Proof. Transform S^{n+1} conformally so that $s = \bar{v}$ preserving γ. Let f be the middle point of $\widehat{sv}$ not containing t and u.

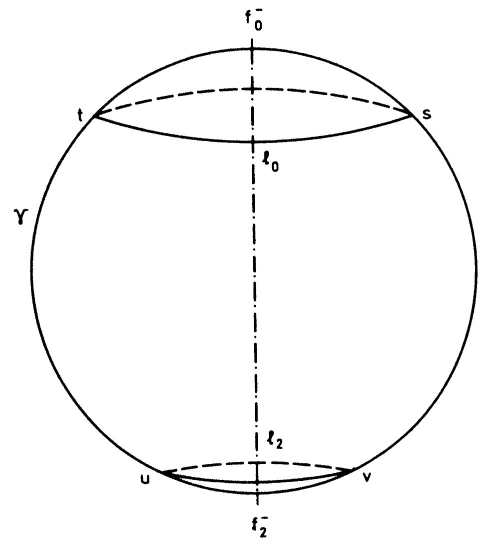


Fig. 4.5

Claim 1. $f_s^\mu = f = f_v^\mu$ if $L_s^\nu \neq L_v^\nu$.

This is shown in the same way as Claim 1 in the proof of Proposition 4.7. Note that we need $L_s^\nu \neq L_v^\nu$ to exclude $f_s^\mu \in \widehat{vf} - \{f\}$ and $f_v^\mu \in \widehat{sf} - \{f\}$.

Claim 2. $L_s^\mu = L_v^\mu$ if $L_s^\nu \neq L_v^\nu$.

See the proof of Claim 2 of Proposition 4.7.

From these claims, we know $L_s^\mu = L_v^\mu$ if $L_s^\nu \neq L_v^\nu$. Similarly we have $L_t^\mu = L_u^\mu$ if $L_t^\nu \neq L_u^\nu$.

Now, to prove Proposition 4.9, let

$$l_0^\mu = \{s \in l_0 \colon \text{joined with } v \in l_2 \text{ by a } \mu\text{-leaf}\}$$

and

$$l_0^\nu = \{s \in l_0 \colon \text{joined with } v \in l_2 \text{ by a } \nu\text{-leaf}\}.$$

Then $l_0 = l_0^\mu \cup l_0^\nu$, a disjoint union. We show that both l_0^μ and l_0^ν are closed. In fact, if $s_k \in l_0^\mu$, $k = 1, 2, \dots$, then since $\{s_k\}$ is a sequence of the compact set l_0, it converges $s_0 \in l_0 = l_0^\mu \cup l_0^\nu$. If $s_0 \in l_0^\nu$, from continuity of the principal curvature, we get $\mu(s_0) = \nu(s_0)$, a contradiction. Thus $s_0 \in l_0^\mu$ and l_0^μ is closed. Since $p_0 \in l_0^\mu$ and $p_1 \in l_0^\nu$, both l_0^μ and l_0^ν are non-empty, which contradicts the connectedness of l_0.

Thus we get $A = \{0\}$, and similarly $C = \{0\}$, which proves the proposition. Q.E.D.

§ 5. Proof of the Main Theorem

In this section, we will prove:

Theorem. *Every compact embedded Dupin hypersurface in S^{n+1} with three principal curvatures is the Lie-geometric image of an isoparametric hypersurface.*

Remark 5.1. By Remark 1.3, compact embedded Dupin hypersurfaces in E^{n+1} or in H^{n+1} with three principal curvatures are determined by this theorem.

Proof. We will use the notation in the previous sections. Let M be the Dupin hypersurface in the theorem. For $p \in M$, it is natural to define the orientation of S_p^ν to be positive and of S_p^λ to be negative. The orientation of S_p^μ is defined positive if $\mu \geqq 0$ and negative otherwise.

Let $p_0 \in M$ be a point such that $\nabla \log \dfrac{\mu-\lambda}{\nu-\mu}=0$. Transform S^{n+1} conformally so that the common normal circle at p_0 to be a great circle, which we denote by γ_0. We knew the situation about γ_0 in §4. Now, it is easy to find a conformal transformation of S^{n+1} inducing $\overline{f_0^-}=\overline{f_2^+}$ and $\overline{f_1^+}=\overline{f_3^-}$. Then by a suitable dilatation, or equivalently, considering a suitable parallel hypersurface of the original hypersurface, we may assume $\nu(p_1)=-\lambda(p_3)$, since the orientation of S_1^+ and S_3^- are opposite. Accordingly, we get $\lambda(p_0)=-\nu(p_2)$ and $\mu(p_3)=\mu(p_5)=0$. Let $\theta=\cot^{-1}\nu(p_2)$, $0<\theta<\frac{\pi}{2}$, and $\psi=\cot^{-1}\nu(p_1)$, $0<\psi<\frac{\pi}{2}$. Note that $0<\theta+\psi<\frac{\pi}{2}$ since $p_2 \neq p_5$. Obviously, M is Lie-equivalent to the present hypersurface which we denote by M_1.

Now, let S^1 be the unit circle in E^2 with center at the origin. Then any pair

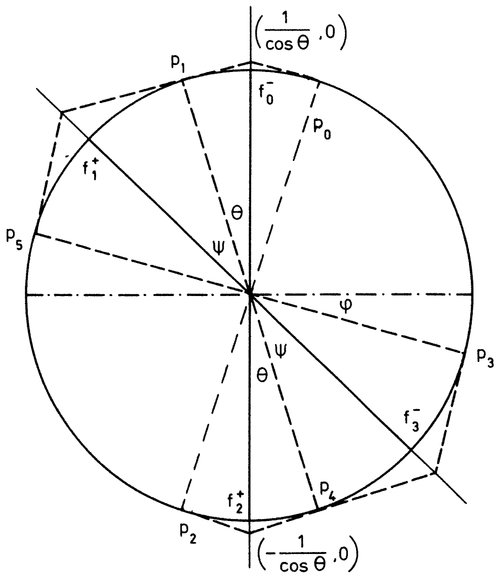


Fig. 5.1

of two points $\{p, q\}$ on S^1 is a hypersphere of S^1, and if we give it an orientation, it corresponds to a point of the quadratic Q with signature $(-, +, +, -)$ in the 3-dimensional projective space $\mathbb{P}^3$. See [6]. Identify γ_0 with this S^1. Then using a suitable homogeneous coordinates of $\mathbb{P}^3$ and considering the induced orientation of $S_i^* \cap \gamma_0$, we can express

$$k_1 = \{p_0, p_1\} = {}^t\!\left(1, \frac{1}{\cos\theta}, 0, -\tan\theta\right),$$

$$k_2 = \{p_2, p_4\} = {}^t\!\left(1, -\frac{1}{\cos\theta}, 0, \tan\theta\right),$$

$$k_3 = \{p_3, p_5\} = {}^t(0, \cos\varphi, \sin\varphi, 1)$$

where $\varphi = \frac{\pi}{2} - \theta - 2\psi$. When $\varphi = 0$, rechoose the coordinates and let

$$k_1' = \{p_3, p_4\} = {}^t\!\left(1, \frac{1}{\cos\psi}, 0, -\tan\psi\right),$$

$$k_2' = \{p_1, p_5\} = {}^t\!\left(1, -\frac{1}{\cos\psi}, 0, \tan\psi\right),$$

$$k_3' = \{p_0, p_2\} = {}^t(0, \cos\varphi', \sin\varphi', 1)$$

where $\varphi' = -\frac{\pi}{2} + 2\theta + \psi$. Let

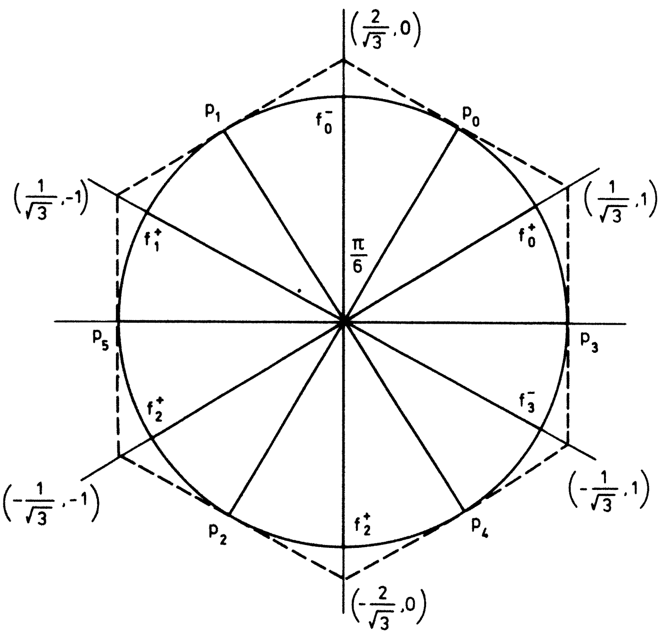


Fig. 5.2

$$h_1 = {}^t\left(1, \frac{2}{\sqrt{3}}, 0, -\frac{1}{\sqrt{3}}\right),$$
$$h_2 = {}^t\left(u, -\frac{2u}{\sqrt{3}}, 0, \frac{u}{\sqrt{3}}\right),$$
$$h_3 = {}^t(0, v, 0, v)$$

be the points of Q for $u, v \in \mathbb{R} - \{0\}$.

Claim 1. There exists a Lie-transformation $A \in O(2,2)$ of S^1 such that $Ak_i = h_i$ ($i = 1,2,3$) or $Ak_i' = h_i$ ($i = 1,2,3$).

This is shown as follows:
When $\varphi = \varphi' = 0$, we have $\theta = \psi = \frac{\pi}{6}$ and the desired Lie-transformation is given by the identity. Now, assuming $\varphi \neq 0$, we will find $A \in O(2,2)$ such that $Ak_i = h_i$. If $\varphi = 0 \neq \varphi'$, let $\theta \to \psi$ and $\varphi \to \varphi'$ in the following, then we get $A \in O(2,2)$ such that $Ak_i' = h_i$.

Let $K = (k_1, k_2, k_3, k_4)$ where $k_4 = {}^t(0,0,0,1)$ and let $Ak_4 = h_4 = {}^t(x, y, z, w)$. It is easy to see that

$$K^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{\cos\theta}{2} & -\frac{\cos\theta\cos\varphi}{2\sin\varphi} & 0 \\ \frac{1}{2} & -\frac{\cos\theta}{2} & \frac{\cos\theta\cos\varphi}{2\sin\varphi} & 0 \\ 0 & 0 & \frac{1}{\sin\varphi} & 0 \\ 0 & \sin\theta & -\frac{\sin\theta\cos\varphi + 1}{\sin\varphi} & 1 \end{pmatrix}$$

Thus we obtain $A = (h_1, h_2, h_3, h_4)K^{-1} = (a_1, a_2, a_3, a_4)$ where

$$a_1 = \tfrac{1}{2}(h_1 + h_2),$$
$$a_2 = \frac{\cos\theta}{2}(h_1 - h_2) + \sin\theta\, h_4,$$
$$a_3 = \frac{\cos\theta\cos\varphi}{2\sin\varphi}(-h_1 + h_2) + \frac{1}{\sin\varphi}h_3 - \frac{\sin\theta\cos\varphi + 1}{\sin\varphi}h_4,$$
$$a_4 = h_4.$$

Let $\langle\ ,\ \rangle$ be the symmetric bilinear form with signature $(-,+,+,-)$. Then A belongs to $O(2,2)$ if and only if

$$\langle a_1, a_1\rangle = \langle a_4, a_4\rangle = -1,$$
$$\langle a_2, a_2\rangle = \langle a_3, a_3\rangle = 1,$$
$$\langle a_i, a_j\rangle = 0 \quad \text{for all } i,j \text{ such that } i \neq j.$$

These equations are equivalent to

$$\langle h_1, h_2\rangle = -2, \tag{5.1}$$

$$\langle h_1+h_2, h_4\rangle = 0, \tag{5.2}$$

$$\langle h_1-h_2, h_3\rangle = \frac{2}{\cos\theta}(\cos\varphi+\sin\theta), \tag{5.3}$$

$$\langle h_1-h_2, h_4\rangle = \tan\theta, \tag{5.4}$$

$$\langle h_3, h_4\rangle = -1, \tag{5.5}$$

$$\langle h_4, h_4\rangle = -1. \tag{5.6}$$

We get $u=1$ from (5.1) and $x=0$ from (5.2). From $\langle h_1+h_2, h_3\rangle=0$ and (5.3), it follows $\langle h_1, h_3\rangle = a$ and $\langle h_2, h_3\rangle = -a$ where $a=\dfrac{\cos\varphi+\sin\theta}{\cos\theta}$. Thus $v=\dfrac{a}{\sqrt{3}}$. (5.4), (5.5) and (5.6) are rewritten as

$$\frac{2}{\sqrt{3}}(2y+w) = 2\tan\theta$$

$$\frac{a}{\sqrt{3}}(y-w) = -1,$$

$$y^2+z^2-w^2 = -1.$$

So we obtain

$$y=\frac{1}{\sqrt{3}}\left(\tan\theta-\frac{1}{a}\right), \quad z=\pm\frac{\sin\varphi}{\cos\varphi+\sin\theta}, \quad w=\frac{1}{\sqrt{3}}\left(\tan\theta+\frac{2}{a}\right).$$

Now, let $k_5=\{p_0, p_3\}$, $k_6=\{p_3, p_4\}$, $k_7=\{p_2, p_5\}$, $k_8=\{p_5, p_1\}$. Noting that k_i and k_j have a contact if and only if $\langle k_i, k_j\rangle=0$, and that $A\in O(2,2)$ preserves this relation, we can easily see $Ak_5={}^t\left(1, \frac{1}{\sqrt{3}}, \mp 1, \frac{1}{\sqrt{3}}\right)$ according to $z=\pm\dfrac{\sin\varphi}{\cos\varphi+\sin\theta}$. Determine A finally by $z=-\dfrac{\sin\varphi}{\cos\varphi+\sin\theta}$. Then we obtain

$$\begin{aligned} Ak_5 &= {}^t\left(1, \frac{1}{\sqrt{3}}, 1, \frac{1}{\sqrt{3}}\right), \quad . & Ak_6 &= {}^t\left(1, -\frac{1}{\sqrt{3}}, 1, -\frac{1}{\sqrt{3}}\right) \\ Ak_7 &= {}^t\left(1, -\frac{1}{\sqrt{3}}, -1, -\frac{1}{\sqrt{3}}\right), & Ak_8 &= {}^t\left(1, \frac{1}{\sqrt{3}}, -1, \frac{1}{\sqrt{3}}\right). \end{aligned} \tag{5.7}$$

When $\varphi=0\mp\varphi'$, choose $z=-\dfrac{\sin\varphi'}{\cos\varphi'+\sin\theta}$ and replace k_i by k_i' in (5.7), where $k_5'=k_2$, $k_6'=k_7$, $k_7'=k_1$, $k_8'=k_5$.

Claim 2. The image hypersurface of M_1 by the Lie-transformation A, where A is considered as an element of $O(n+2,2)$ canonically, is isoparametric.

In the following proof, all the notation shows the objects in the image hypersurface of M_1 by A. It is easy to see that $\overline{f_0^-}=f_3^\mu=f_2^+$, $f_1^+=f_0^\mu=\overline{f_3^-}$ and

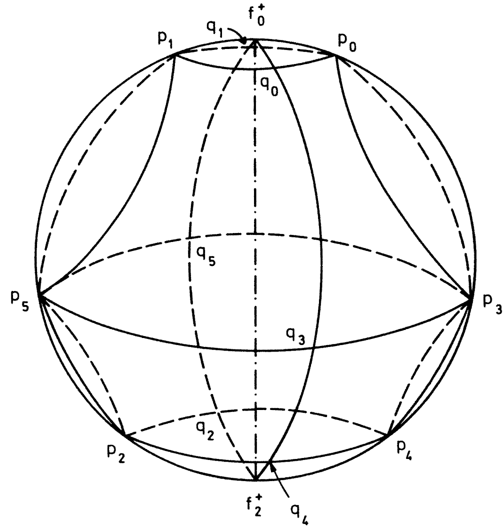


Fig. 5.3

$f_0^+ = f_1^\mu = \overline{f_2^-}$. Putting $m=m_1=m_2=m_3$, let S^{m+1} be the great $(m+1)$-sphere containing L_0^λ. From Proposition 4.9, it follows $S^{m+1} \supset L_3^\mu \cup L_2^\nu$. Using indices i, j, k such that $e_i \in \mathscr{D}(\lambda)$, $e_j \in \mathscr{D}(\mu)$ and $e_k \in \mathscr{D}(\nu)$, we note that $\sum_{l \notin [i]} (A_{ii}^l)^2 = 0$ on L_0^λ, $\sum_{l \notin [j]} (A_{jj}^l)^2 = 0$ on L_3^μ and $\sum_{l \notin [k]} (A_{kk}^l)^2 = 0$ on L_2^ν. Since $\overline{f_0^-} = f_3^\mu = f_2^+$, a great circle γ in S^{m+1} joining f_0^- and f_2^+ cuts L_0^λ, L_3^μ and L_2^ν twice for each orthogonally (see Remark 2.3). Let $\gamma \cap L_0^\lambda = \{q_0, q_1\}$, $\gamma \cap L_3^\mu = \{q_3, q_5\}$ and $\gamma \cap L_2^\nu = \{q_2, q_4\}$ as in Fig. 5.3. Then by using tautness, it is easy to see that $L_{q_1}^\nu = L_{q_5}^\nu$, $L_{q_5}^\lambda = L_{q_2}^\lambda$, $L_{q_4}^\lambda = L_{q_3}^\lambda$ and $L_{q_3}^\nu = L_{q_0}^\nu$. Moreover we see $L_{q_0}^\mu = L_{q_2}^\mu$ and $L_{q_1}^\mu = L_{q_4}^\mu$. Finally, γ with points q_i, $i=0,1,\dots,5$ on it is congruent to γ_0 with points p_i, $i \equiv 0,1,\dots,5$ on it by a rotation fixing L_0^λ. In particular, γ is the common normal geodesic at all points of $\gamma \cap M$. We can show the same thing for γ_r, $r \in L_i^*$, $i=0,1,\dots,5$ and $* = \lambda, \mu, \nu$, by using $f_1^+ = f_0^\mu = \overline{f_3^-}$, $f_0^+ = f_1^\mu = \overline{f_2^-}$ and Proposition 4.9.

Now, recall Corollary 4.6. For any $p \in M$, we can find $q \in L_0^\mu$ such that $L_p^\nu \cap L_q^\lambda \neq \emptyset$. Put $L_q^\nu \cap L_p^\lambda = r$. Then starting from γ_q which is congruent to γ_0 we see that γ_r is congruent to γ_q by a rotation preserving L_q^λ, and then γ_p is congruent to γ_r by a rotation fixing L_p^ν. Finally, for any $p \in M$, γ_p is shown to be the common normal geodesic at p. Hence the image hypersurface is isoparametric by Proposition 2.4. Q.E.D.

References

1. Carter, S., West, A.: Tight and taut immersions. Proc. London Math. Soc. **25**, 707–720 (1972)
2. Cecil, T., Ryan, P.: Focal sets of submanifolds. Pacific J. Math. **78**, 27–39 (1978)

3. Cecil, T., Ryan, P.: Focal sets, taut embeddings and the cyclides of Dupin. Math. Ann. **236**, 177–190 (1978)
4. Chen, B.Y.: Geometry of submanifolds. Pure and Applied Math. 22. New York, Marcel Dekker (1973)
5. Münzner, H.F.: Isoparametrische Hyperflächen in Sphären, II. Math. Ann. **256**, 215–232 (1981)
6. Pinkall, U.: Dupin'sche Hyperflächen, Thesis, Freiburg i.Br. (1981)
7. Thorbergsson, G.: Dupin hypersurfaces, Bull. London Math. Soc. **15**, 493–498 (1983)

Received December 30, 1983

Math. Z. 187, 453–470 (1984)

Mathematische Zeitschrift

Families of Curves on Surfaces

Augusto Nobile

Mathematics Department, Louisiana State University, Baton Rouge, Louisiana 70803, USA

Introduction

In this paper we study certain families of curves on a smooth surface X. Essentially, we are interested in families of reduced, singular curves, such that the singularities vary in an "equisingular fashion". The precise definition, which is necessarily somewhat technical, can be seen in Sect. (1.2). The case where X is the projective plane and the singularities are nodes is classical, it was treated by Severi early in this century (see [6], Anhang F; see also [5] and [7]). A generalization to the case where X is a rational or a K-3 surface (and the singularities are nodes) in presented by A. Tannenbaum in the papers [7] and [8] respectively. In the paper [13], O. Zariski studies families of curves of degree n with singularities in $\mathbb{P}^2$, parametrized by subvarieties of the $\mathbb{P}^N (N=\frac{1}{2}n(n+3))$ which parametrizes all plane curves of degree n. He shows that the parametrizing variety of such a family must have dimension $d, d \leqq 3n-g+1$, where g is the geometric genus of a general member; moreover the equality holds if and only if the family is a maximal one, consisting of curves whose only singularities are nodes.

In this paper we present a generalization of Zariski's and Tannenbaum's results to projective, smooth algebraic surfaces. Given a reduced curve C in such a surface X, we study the functor of infinitestimal deformations of C inside X, which "preserve the singularities", in the sense of the theory of Zariski and Wahl ([11, 12]); see (1.3) for the precise definition. This functor is pro-representable, and under the condition:

$$C_i \cdot K < 0, \quad \text{for all } i, \tag{1}$$

where $C_1, \ldots, C_r$ are the irreducible components of C, and K is the canonical divisor of X, the inequality $\dim T_C \leqq \dim H^0(C, N_C) - \delta$ holds, where T_C is the tangent space to the functor under consideration, N_C the normal bundle and $\delta = \sum_{P \in C} \delta(C, P)$, these being the usual "δ-invariants" (see (0.5)). This is the analogue of Zariski's inequality. Under (1), in the case where all the singularities are nodes the equality sign holds and the deformation functor is smooth. Under conditions slightly stronger than (1) (see Theorem (3.12)) the converse is also true i.e. the equality sign forces the singularities to be nodes. These results appear in Chap. 1 through 3. In § 4 we discuss some global results, namely we study certain subvarieties of the Hilbert scheme of X which are (under suitable assumptions) universal with respect to families (of the type that we are considering) parametrized by reduced schemes. In § 5, in the case of nodes, we precise these results. In this case (under suitable conditions, namely the inequalities (1)), we can represent the functor classifying our problem (see (1.2) for its precise definition). Results of this type in the case of higher singularities look more delicate, because the introduction of nilpotents is essential, as is known in the case where some singularities are cusps ([10]). We hope to treat some of these questions in the future.

The techniques of this paper are similar to those of [13]. In fact, the main results there depend on some ingenuous calculations which, although presented in [13] in a global setting, are really local (and even analytic). They can be adapted to work for surfaces too, provided the ground is suitably prepared. In particular, in § 2 we discuss the key connection between first order deformations (inducing equisingular deformations of the singularities) and adjoint sections of N_C, pointed out (in the case of $\mathbb{P}^2$) by Zariski in [13]. Of course, for nodes this was discovered by Severi many years ago.

In Chap. 0 we explain some notation and terminology. The author thanks the referee for his corrections and valuable suggestions.

§ 0. Notation and Preliminaries

(0.1) The basic terminology is taken from [3], with a few minor changes. Namely, the term "algebraic variety" denotes on reduced scheme of finite type over an algebraically closed field, but it is not necessarily irreducible. The term "generic point" will mean a closed point taken from a suitable dense open set, which will be clear from the context. Sing(Y) denotes the singular set of an algebraic variety Y. An algebraic surface is a variety of dimension 2, not necessarily projective.

(0.2) The categories of artinian local k-algebras with residue field k, of complete k-algebras with residue field k and of algebraic schemes over k will be denoted by $\mathscr{A}$, $\mathscr{A}_c$ and (Alg) respectively. If F is a functor defined on schemes, and A is a ring we often write $F(A)$ rather than $F(\operatorname{Spec} A)$.

(0.3) M_A (or $\max(A)$) will denote the maximal ideal of the local ring A. If A is a ring, $\operatorname{Fr}(A)$ denotes its total ring of quotients. If f is a power series, $\operatorname{ord}(f)$ denotes its order.

(0.4) We shall use Zariski's theory of equisingularity, as explained in [12]. Our terminology is taken from there. An element of $\{C: C$ is an algebroid plane curve$\}$ modulo the equivalence relation of "equivalence" (cf. [12], I, Def. 3 and 4) is called an "equisingular class."

We shall also use results on equisingular deformations of algebroid curves, taken from [11].

(0.5) A "curve" means an algebraic variety of dimension 1. If P is closed point of a curve C, $A=\mathcal{O}_{C,P}$, then $\delta(C,P)=\dim_k(\bar{A}/A)$. If C is projective, having irreducible components $C_1,\dots,C_r$, we have the well-known formula:

$p_a(C)=\sum_{i=1}^{r} g_i+\sum_{P\in C'}\delta_p-r+1$, where g_i is the geometric genus of C_i, $\delta_p=\delta(C,P)$.

(0.6) A branch $\mathfrak{B}$ of a curve C at P is a scheme of the form Spec(B), where $B=\hat{\mathcal{O}}_{C,P}/\mathcal{P}$, $\mathcal{P}$ being a minimal prime ideal. The irreducible component C_i of C corresponding to $\mathcal{P}\cap\mathcal{O}_{C,P}$ is, by definition, the component determined by $\mathcal{P}$. We also say in this case: C_i is the component containing $\mathfrak{B}$.

§ 1. Equisingular Families

(1.1) Let X be a smooth surface. If $T\in$(Alg) (cf. (0.2)), $D\subset X\times T=X_T$ is a relative Cartier divisor (i.e., the morphism $p\colon D\to T$ induced by the second projection is flat), $t\in T$ and $P\in D_t=p^{-1}(t)$ are closed points, then $\hat{\mathcal{O}}_{X,P}\approx k[\![x,y]\!]$, $\hat{\mathcal{O}}_{X_T,P}\approx A[\![x,y]\!]$ (where $A=\hat{\mathcal{O}}_{T,t}$) and D_t (resp. D) induces, in an obvious way, an algebroid curve $\hat{D}_{t,P}$ in Spec $K[\![x,y]\!]$ (resp., a deformation $\hat{D}_P$ of this curve over A), defined, say, by $f_0\in k[\![x,y]\!]$ (resp., $f\in A[\![x,y]\!]$).

(1.2) We shall be interested in the following functor. Fix a smooth surface X, and a finite sequence $\mathcal{I}=(\mathcal{I}_1,\dots,\mathcal{I}_r)$ of equisingular classes. Then,

(1.2.1) $$\mathcal{F}\colon(\text{Alg})\to(\text{sets})$$

(or $\mathcal{F}_{X,\mathcal{I}}$, to be more precise) is defined by: if $T\in$(Alg), $\mathcal{F}(T)=\{D\subset X\times T\colon D$ is a relative effective Cartier divisor, and condition $E(\mathcal{I})$ holds at each closed point of $T\}$, where $E(\mathcal{I})$ means the following:

(1.2.2) $E(\mathcal{I})$: the fiber D_t is a reduced curve, Sing$(D_t)=\{P_1,\dots,P_r\}$, the algebroid curve $\hat{D}_{t,P_i}$ induced by D at P_i is of class $\mathcal{I}_i$ and moreover, if $\hat{D}_{P_i}$ is the deformation of $\hat{D}_{t,P_i}$ induced by D, then there is a section $s_{t,i}$ of $\hat{D}_{P_i}\to\text{Spec}(\hat{\mathcal{O}}_{T,t})$, such that $\hat{D}_{P_i}$ is equisingular along s_i (here we use the notation of (1.1) and (0.4), and of course we assume that the points of Sing(D_t) have been suitably numbered).

Note that, according to [11], Theorem (3.2), if an equisingular section exists, it is unique, hence it is not necessary to specify the sections $s_{t,i}$ as data of our functor. We say that $\mathcal{F}$ is the functor of equisingular families of curves on X, of classes $\mathcal{I}_1,\dots,\mathcal{I}_r$.

(1.3) As usual, if $C\in\mathcal{F}(k)$ ($\mathcal{F}=\mathcal{F}_{\mathcal{I}}$, $\mathcal{I}=(\mathcal{I}_1,\dots,\mathcal{I}_r)$ where each $\mathcal{I}_i$ is an equisingular class) we may introduce the functor $\mathcal{F}_C$ (or $\mathcal{F}_{\mathcal{I},C}$):

(1.3.1) $$\mathcal{F}_C\colon A\to(\text{Sets})$$

given by: $\mathcal{F}_C(A)=\{\mathcal{D}\in\mathcal{F}(A)\colon$ under $A\to A/m_A=k$, $\mathcal{D}$ induces $C\}$.

This will be called the functor of infinitestimal equisingular deformations of C (inside X). We shall study $\mathscr{F}_C$.

Note that we may define the "Hilbert functor" $\mathscr{H}$: (Alg)$\to$(sets), $\mathscr{H}(T) = \{\mathscr{D} \subset X \times T$: $\mathscr{D}$ is a relative effective Cartier divisor$\}$, as well as the corresponding infinitesimal functor $\mathscr{H}_C$, $C \in \mathscr{F}(k)$.

Clearly, $\mathscr{F} \subset \mathscr{H}$ and, if C is reduced and has singularities $P_1, \ldots, P_r$ of types $\mathscr{I}_1, \ldots, \mathscr{I}_r$ resp., then $\mathscr{F}_{\mathscr{I},C} \subset \mathscr{H}_C$.

(1.4) We recall some well-known facts on deformation theory, mostly taken from M. Schlessinger's "Functors of Artin Rings", Trans. A.M.S. **130**, 208–222:

Given a functor $F: \mathscr{A} \to$(Sets) such that $F(k)$ is a point, and a diagram:

(1.4.1)
$$\begin{array}{ccc} B' = A' \underset{A}{\times} B & \longrightarrow & B \\ \big\downarrow & & \big\downarrow{\scriptstyle \beta} \\ A' & \xrightarrow[\alpha]{} & A \end{array}$$

where $\alpha: A' \to A$ is small (i.e., $\operatorname{Ker}(\alpha)$ is principal, say generated by η, and $\eta \operatorname{Ker}(\alpha) = 0$), consider the naturally induced map $F(B') \xrightarrow{\gamma} F(A') \underset{F(A)}{\times} F(B)$. If γ is bijective then $F(k[\varepsilon])$ ($k[\varepsilon]$ = dual numbers) is naturally a k-vector space (this can be naturally identified to the tangent space of the functor F, see, e.g., [4], Lecture 4). If γ is bijective and this vector space is finite dimensional, F is pro-representable. Also, $\mathscr{H}_C$ is pro-representable, and its tangent space is naturally isomorphic to $H^0(C, N_C)$, N_D = normal bundle of D in $X = \mathcal{O}_X(D) \underset{\mathcal{O}_X}{\otimes} \mathcal{O}_D$ (cf. [10], 1.4 and 3.2, and [4], p. 153).

(1.5) **Theorem.** *Let $X, \mathscr{I}, \mathscr{F}$ be as in* (1.1) *and* (1.2), *$D \in \mathscr{F}(k)$ a reduced curve in X. Then, $\mathscr{F}_D$* (cf. (1.3)) *is such that the map γ is bijective and $\mathscr{F}_D(k[\varepsilon]) \subset H^0(D, N_D)$. If D is projective, then $\mathscr{F}_D$ is pro-representable.*

Remark. In Sect. (2.11) we shall obtain finer information about the inclusion $\mathscr{F}_D(k[\varepsilon]) \subset H^0(D, N_D)$.

(1.6) *Proof of* (1.5). Let R be a k-algebra, $A \in \mathscr{A}$. We shall write $R_A = R \underset{k}{\otimes} A$. Recall: (i) Given the situation (1.4.1), then:

(1.5.1)
$$R_{A' \underset{A}{\times} B} = R_{A'} \underset{R_A}{\times} R_B$$

(a canonical isomorphism).

(ii) The functor $G: \mathscr{A} \to$(Sets) of infinitesimal automorphisms, i.e. $G(A) = \{A$-automorphisms of $R \otimes A$ inducing id over $k\}$ is smooth and its corresponding "map γ" (cf. 1.4) is bijective (see [10] (1.3)).

(iii) If $A' \twoheadrightarrow A$, the map $R^*_{A'} \to R^*_A$ of units is surjective. Also recall that any $\mathscr{D} \in \mathscr{F}_D(A)$ is given by an affine cover $\{U_i\}$, $U_i = \operatorname{Spec} R_i$ of X, and an "equation" $F_i \in R^{(i)}_A = \Gamma(U_i, \mathcal{O}_{X_A})$ ($X_A = X \times \operatorname{Spec} A$), this equation being determined up to multiplication times a unit $\mu \in \Gamma(U_i, \mathcal{O}^*_{X_A})$. Using these facts, one easily verifies that, given $(\mathscr{D}', \mathscr{E}) \in \mathscr{F}_D(A') \underset{\mathscr{F}_D(A)}{\times} \mathscr{F}_D(B)$, we may assume that $\mathscr{D}'$ (resp. $\mathscr{E}$) is given by an affine cover $\{U_i\}_{i\in I}$ of X, $U_i=\operatorname{Spec} R_i$, and local equations $F_i'\in R_{A'}^{(i)}$ (resp $G_i\in R_B^{(i)}$) inducing the same $F_i\in R_A^{(i)}$; then $(F_i', G_i)\in R_{B'}^{(i)}$, $B'=A'\underset{A}{\otimes}B$ (via the identification (1.5.1)), $i\in I$, defines a relative Cartier divisor $\mathscr{E}'\subset X_{B'}=X\times\operatorname{Spec}(B')$. Moreover, if $P\in X_{B'}$ is a closed point,

$$\hat{\mathcal{O}}_{X_{B'},P}\approx B'[\![x,y]\!]\approx A'[\![x,y]\!]\underset{A[\![x,y]\!]}{\times}B[\![x,y]\!]$$

(x, y a set of generators of $\max(\mathcal{O}_{X,P})$). If P is a singular point of D, by assumption there are equisingular sections s'(resp. σ) of $\hat{\mathscr{D}}'_P$ (resp. $\hat{\mathscr{E}}_P$), we may assume that the local equations are chosen so that s' and σ induce the same section of $\hat{\mathscr{D}}_P$. Then (s', σ) defines a section of $\hat{\mathscr{E}}'_P$. But by the results of [11], it follows that this will be a (unique) equisingular section of $\hat{\mathscr{E}}'_P$. It is easily checked that this construction produces a map $\mathscr{F}_D(A')\underset{\mathscr{F}_D(A)}{\times}\mathscr{F}_D(B)\to\mathscr{F}_D(B')$, which is the inverse of γ.

Concerning the tangent space, we saw: $\mathscr{F}_D\subset\mathscr{H}_D$, hence $\mathscr{F}_D(k[\varepsilon])\subset\mathscr{H}_D(k[\varepsilon])=H^0(D, N_D)$ (cf. [4], p. 153). Now, if D is projective, $H^0(D, N_D)$ is finite-dimensional, hence $\mathscr{F}_D(k[\varepsilon])$ is so.

§2. Adjoints

(2.1) Let $f\in K[\![x,y]\!]$ define an algebroid curve C, let $g\in k[\![x,y]\!]$. We shall define the notion "g is an adjoint of f", by induction on $\sigma(f)=\sigma(C)=\min\{m\,|\,a$ sequence of m quadratic transforms desingularizes $C\}$ (cf. [2], p. 799, where a more global situation is discussed).

If $\sigma(C)=0$, i.e., C is smooth, no condition is imposed. Now let $\sigma(C)=q$, $\operatorname{ord}(f)=r$. Let C have s "tangential components" (cf. [12], §2), we may assume the i-th has an equation $f_i=(y-\alpha_i x)^{r_i}+\dots$, $r_i=\operatorname{ord} f_i$. We require: (i) $\operatorname{ord} g\geqq r-1$. (ii) Consider g', defined by $g(x, xy)=x^{r-1}g'(x,y)$. For each i, $i=1,\dots,s$, g' (resp. the proper transform of f_i) can be written as a series g_i (resp. $\tilde{f}_i$) in $k[\![x,\mu_i]\!]$, $\mu_i=y-\alpha_i$ (cf. [12], §2); then g_i must be an adjoint of $\tilde{f}_i$ (this makes sense by induction, since $\sigma(\tilde{f}_i)<q$), for all i.

(2.2) In the study of adjoints, the following technique is useful (cf. [13], p. 216–218). Let $C: f=0$ be as in (2.1), $\mathcal{O}=k[\![x,y]\!]/(f)$ it local ring, K its total ring of fractions, the symbol Ω_R stands for the module of analytic differentials over k of a complete semi-local k-algebra R. A similar notation holds for the total ring of fractions of a reduced R. Assume:

(2.2.1) The line $x=0$ is not tangent to C. Given $g\in k[\![x,y]\!]$, consider:

$$\omega(g)=(g/f_y)\,dx\in\Omega_K.$$

Here, $K=\operatorname{Fr}(\bar{\mathcal{O}})$, where $\bar{\mathcal{O}}=$normalization of $\mathcal{O}$, and clearly $\Omega_{\bar{\mathcal{O}}}\subset\Omega_K$. Then

(2.3) **Lemma.** *g is an adjoint to f if and only if $\omega(g)\in\Omega_{\bar{\mathcal{O}}}$.*

The proof is easily obtained by induction on $\sigma(f)$, using the definition of an adjoint (cf. (2.1)) and the formula

(2.3.1) $$f_y(x, xy)=x^{r-1}f_y(x,y), \qquad r=\operatorname{ord} f.$$

(2.4) Recall the following facts. Given a curve $C: f=0$, $f \in k[\![x, y]\!]$ and a deformation thereof over $A \in \mathscr{A}$ given by $F \in A[\![x, y]\!]$, then a section s of $\operatorname{Spec}(A[\![x, y]\!]/F) \to \operatorname{Spec}(A)$ inducing the trivial one over k, is determined by elements α, β in $\max(A)$, satisfying $F(\alpha, \beta)=0$. We may change coordinates via $x'=x+\alpha$, $y'=y+\beta$, (which induces the identity $\bmod(m_A)$), the new corresponding section will be the trivial one (i.e., $\alpha=\beta=0$). We say that the given deformation is equimultiple along a section s if, after trivializing the section, it is defined by $F'=0$, with $\operatorname{ord} F'=\operatorname{ord} f$.

We shall need the following lemma (cf. [13], p. 218):

(2.5) **Lemma.** *Let $A=k[\varepsilon]$ (dual numbers), $f \in k[\![x, y]\!]$ and C as in* (2.1), *$F=f+\varepsilon g$ a deformation of f over A and s a section, determined by $\alpha\varepsilon$, $\beta\varepsilon$* (α, β in k). *Assume: F is equimultiple along s. Let $F'=f+\varepsilon g'$ be the quadratic transform of F along s. Then, we have* (*with the notation of* (2.2)) *the equality in Ω_K:*

$$\omega=x\omega'+[\alpha y'-\beta]\,dx,$$

where $\tilde{\omega}=\tilde{\omega}(g)$ and $\tilde{\omega}'=\tilde{\omega}(g')=g'/f'_{y'}\,dx$.

(2.6) *Note.* We explain the notation just used. The transform under consideration is described by the formulas $x=x'$, $y=y'x-\alpha y'\varepsilon+\beta\varepsilon$, this defines y'. Also, to simplify, we commit an abuse of notation: the quadratic transform of f is, in general, a curve with several "connected components", corresponding to the tangents of f (cf. [12], §2), in defining $\omega(g')$ one should apply (2.2) to each such component. To avoid overloading the notation, similar "abuses", which we believe harmless, will be committed in the proof below.

(2.7) *Proof of* (2.5). Substitute in F the quadratic transform formulas shown in (2.6). But first note that the equimultiplicity assumption implies: $g(x, xt) = x^{s-1}\tilde{g}(x, t)$, $s=\operatorname{ord} f$. In fact, $F(x+\alpha\varepsilon, y+\beta\varepsilon)=f(x, y)+\varepsilon h(x, y)$ with $h=\alpha f_x+\beta f_y+g$, since $\alpha f_x+\beta f_y$ has order $\geqq s-1$, then $\operatorname{ord} g \geqq s-1$, whence the claim. Using this and (2.3.1), we get:

$$f+\varepsilon g=x^s f(x, \tilde{y})+\varepsilon[-\alpha y' x^{s-1} f'_{y'}(x, y')+\beta x^{s-1} f'_{y'}(x, y')+x^{s-1}\tilde{g}(x, y')]$$

hence,

$$\tilde{F}=f'(x, y')-\varepsilon g'(x, y') \quad \text{where } g' \in k[\![x, y']\!]$$

is:

$$g'=\frac{1}{x}[-\alpha y' f_{y'}(x, y')+\beta f'_y(x, y')+\tilde{g}(x, y)].$$

Then,

$$\begin{aligned}
x\tilde{\omega}&=[(-\alpha y f'_y(x, y')+\beta f'_{y'}(x, y')+\tilde{g}(x, y')/f'_y(x, y')]\\
&=[(-\alpha y' f'_{y'}(x, y')+\beta f'_{y'}(x, y')+\tilde{g}(x, y'))/f_y(x, xy')/x^{s-1})]\,dx\\
&=[(-\alpha y' x^{s-1} f'_{y'}(x, y')+\beta x^{s-1} f_y(x, y')+x^{s-1}\tilde{g}(x, y'))/f_y(x, xy')]\,dx\\
&=[(-\alpha y' f_y(x, xy')+\beta f_y(x, xy)+g(x, xy'))/f_y(x, xy')]\,dx\\
&=(-\alpha y'+\beta)\,dx+(g(x, xy')/f_y(x, xy'))\,dx=(-\alpha y'+\beta)\,dx+\omega,
\end{aligned}$$

i.e., our claimed result.

(2.8) *Theorem. Let A, C, f, F be as* (2.5), *but now assume the deformation is equisingular (along a suitable section, uniquely determined). Then, g is an adjoint of f.*

Proof. Let s be the section along which F is equisingular. Then, F is equimultiple along s, and we may apply Lemma 2.5. By induction on $\sigma(f)$, we may assume (by using (2.3)) that, in (2.5.1), ω' is regular, i.e., in Ω_R, R=normalization of $(k[\![x,y]\!]/f)$. It follows that ω is in Ω_R, now use (2.3).

(2.9) **Proposition.** *Let f, C be as in* (2.1), *but now we also assume that the singularity is a node. Let $g \in k[\![x,y]\!]$ be an adjoint to C (in this case this just means $g(0,0)=0$). Then, $f+\varepsilon g$ is a first order trivial (hence equisingular) deformation of C.*

Proof. Change coordinates via $x=x'-\sigma\varepsilon$, $y=y'-\beta\varepsilon$, $\alpha, \beta \in k[\![x',y']\!]$ (to be specified later). Then,

$$f+\varepsilon g = f(x',y') + \varepsilon[g(x',y') - \alpha(x',y') f_x(x',y') - \beta(x',y') f_y(x',y')].$$

Since f has a node, f_x, f_y generate (x',y'); and since $g \in (x',y')$ (because it is an adjoint), there are $\alpha, \beta \in k[\![x,y]\!]$ such the bracket above is zero. So, in this new coordinates the deformation is given by $f=0$, and is trivial.

(2.10) We shall apply these results to the functors $\mathscr{F}_D$ of (1.3). Recall: using the notation (1.3), $\mathscr{F}_D \subset \mathscr{H}_D$, and we have an isomorphism $\mathscr{H}_D(k[\varepsilon]) \tilde{\to} H^0(D, N_D)$ (N_D=normal bundle, see [4], p. 153). Hence we get a k-linear injective map:

(2.10.1) $$\mathscr{F}_D(k[\varepsilon]) \xrightarrow{\gamma} H^0(D, N_D).$$

Given a section $s \in H^0(D, N_D)$, we call it an *adjoint* of D if the following happens: for each singular point P of D, if s induces, on $\hat{\mathcal{O}}_{X_A, P} \approx A[\![x,y]\!]$, ($A = k[\varepsilon]$), the first order deformation $f+\varepsilon g$ of $\operatorname{Spec}(\hat{\mathcal{O}}_{D,P})$ (i.e., here $f=0$ is an analytic equation of D at P), then g is an adjoint of f (cf. (2.1)). They form a vector subspace A_D of $H^0(D, N_D)$. Since, by definition of the functor $\mathscr{F}_D$, elements of $\mathscr{F}_D(k[\varepsilon])$ induce at each singular point P of D an equisingular deformation $f+\varepsilon g$ of $\hat{D}_P$, (2.8) and (2.9) imply:

(2.11) **Theorem.** (a) *the image of the homomorphism γ* (cf. (2.10.1)) *is contained in the space A_D of adjoint sections of N_D.* (b) *If all the singularities of D are nodes, then* $\operatorname{Im}(\gamma) = A_D$.

§3. Dimensions of Tangent Spaces

(3.1) We work in the situation of (1.3), but from now on we shall assume that X is a *projective* smooth surface. The symbol K_X (or just K) will denote its canonical divisor. In (2.10)–(2.11) we saw that there is a canonical injective linear map γ from the tangent space $T_C = \mathscr{F}_C(k[\varepsilon])$ of $\mathscr{F}_C$ into $H^0(C, N_C)$, which actually lands in A_C={adjoint sections of C}. Now, X being projective, $\dim H^0(C, N_C) < \infty$. We shall estimate $\dim T_C$, by using the classical method of the "characteristic linear series" (cf. [13], §3).

(3.2) Let $f: C' \to C$ be the normalization of C, then we have inclusions (up to canonical identifications).

(3.2.1) $$T_C \subset A_C \subset H^0(C, N_C) \subset H^0(C, f_* f^* N_C) = H^0(C', f^* N_C).$$

The elements of T_C, regarded as elements of $H^0(C', f^* N_C)$, "cut out" (by taking zero of these sections) a linear series Λ' (a "$g^{\rho}_{d'}$") on C'. Let us assume that C has r irreducible components $C_1, \ldots, C_r$, hence C' will have r smooth irreducible components $C'_1, \ldots, C'_r$, C'_i being the normalization of C_i. Clearly, if $N' = f^* N_C$, then

$$H^0(C', N') = \bigoplus_{i=1}^{r} H^0(C'_i, N'|_{C'_i})$$

and the divisors of Λ' are parametrized by the quotient of $H^0(C', N')$ under the action of $k^* \times \ldots \times k^*$ (r times), hence

(3.2.2) $$\dim T_C = \rho + r, \qquad \rho = \dim \Lambda'.$$

Now, Λ' always has a fixed divisor. Precisely we have:

(3.3) **Lemma.** *Let $s \in A_C$ and $D = \sum m_T T$, $T \in C'$, the divisor (of C') of zeroes of s (use (3.2.1)). Let $Q \in C'$ such that $f(Q) \in \mathrm{Sing}(C)$, $\mathfrak{B}'$ the unique branch of $\bar{C}$ at Q and $\mathfrak{B}$ the unique branch of C (at P) corresponding to $\mathfrak{B}'$ under f, where $P = f(Q)$. Then*

(3.3.1) $$m_Q \geqq 2\delta(\mathfrak{B}, P) + \sum_{\mathfrak{B} \in G} (\mathfrak{B}, \mathfrak{C})$$

where $G = \{\mathfrak{C} : \mathfrak{C}$ is a branch of C at P, $\mathfrak{C} \neq \mathfrak{B}\}$, δ is as in (0.5) and $(\mathfrak{B} \cdot \mathfrak{C})$ denotes "intersection number".

Proof. This is a consequence of the usual expressions of δ and the multiplicity of intersections in terms of multiplicities (at P and at infinitely near points, cf. [3], p. 393, (3.9.3) and 394, (3.2)), and the following observation. If R is a point infinitely near P (and lying under Q). $\mathfrak{B}_0$, C_0, S are induced (at R) by $\mathfrak{B}$, C and the adjoint s respectively, (assume C_0 has, aside from $\mathfrak{B}_0$, branches $\mathfrak{B}_1, \ldots, \mathfrak{B}_l$ at R) and m indicates multiplicity, then:

$$\begin{aligned} m(\mathfrak{B}_0)\, m(S) &\geqq m(\mathfrak{B}_0)[m(C_0) - 1) \\ &= m(\mathfrak{B}_0)\left[(m(\mathfrak{B}_0) - 1) + \sum_{j=1}^{l} \mathfrak{B}_j\right] \\ &= 2[m(\mathfrak{B}_0)(m(\mathfrak{B}_0) - 1)/2] + \sum_{j=1}^{l} m(\mathfrak{B}_0)\, m(\mathfrak{B}_j). \end{aligned}$$

(3.4) **Corollary.** *Let $s \in A_D$, D its divisor of zeroes (on C'), $E = \sum n_Q Q$, where n_Q is given by the right hand side of* (3.3.1) *if Q lies over a singular point of C and zero otherwise. Then,* (i) *$D \geqq E$ and* (ii) *$\deg E = 2\delta$, $\delta = \sum_{P \in S} \delta(C, P)$, $S = \mathrm{Sing}(C)$.*

Proof. (i) is an obvious consequence of (3.3), and (ii) of the formula $\delta(C,P) = \sum_{i=1}^{r} \delta(C_i, P) + \sum_{j \neq t} i(P, C_j \cdot C_t)$ (if C has irreducible components $C_1, \ldots, C_r$).

(3.5) **Definition.** (i) The divisor E of (3.4) is called the *standard fixed divisor* of the adjoint series. (ii) The linear series $\Lambda = \{(s) - E : s \in T_C\}$ ((s) indicates "zeroes of s") is called the *characteristic series* of $\mathscr{F}_C$.

(3.6) *Remarks.* (a) If $D' \in \Lambda'$ (3.2), then $\deg D' = \deg N_C = C^2$; if $D \in \Lambda$, then $\deg D = \deg N_C - 2\delta$, $\delta = \sum_{P \in C} \delta(C,P)$.

(b) If $s \in A_C$, $P \in C'$, $Q = f(P)$, the equation of $\hat{C}_Q$ (cf. (1.1)) is $f = 0$, $f \in \hat{\mathcal{O}}_{X,Q} \approx k[\![x,y]\!]$ and s induces the deformation $f + \varepsilon g$ of f, then the coefficient of $(s) - E$ at P is $\operatorname{ord}_P(\omega(g))$ (cf. (2.2)). Here we use, of course, the canonical bijection of maximal ideals of $\bar{\mathcal{O}}_{C,Q}$ and points of the normalization C' lying over Q). This is proved by a standard induction on $\sigma(f)$.

The following theorem estimates $\dim T_C$.

(3.7) **Theorem.** *Let C be a reduced curve on X, with irreducible components $C_1, \ldots, C_r$. Using the notation of the previous sections (in particular, a prime indicates "normalization"), let $D \in \Lambda$ (the characteristic series), $D = D_1 + \ldots + D_r$, with D_i supported at C_i'. Write $\mathcal{O}(D) = \mathcal{O}_{C'}(D)$, $\delta = \sum_{P \in C} \delta(C,P)$. Then,*

(a) $\dim T_C \leqq \dim H^0(C', \mathcal{O}(D)) = h^0(D)$.

(b) *If the following holds:*

(3.7.1) *Each D_i is a non special divisor on C_i', then:*

$$h^0(D) \leqq \deg N_C - p_a(C) + 1 - \delta = \chi(N_C) - \delta \leqq \dim H^0(C, N_C) - \delta$$

(c) *If the condition:*

$$C_i \cdot K < 0 \quad \textit{for all } i \tag{3.7.2}$$

(where K is the canonical divisor of X) holds, then each D_i is non special, i.e., (3.7.1) *holds.*

(d) *If* (3.7.2) *holds, if $|C|$ contains an integral curve Z, and if $H^1(X, \mathcal{O}_X) = 0$, then:*

$$h^0(D) \leqq \dim H^0(Z, N_Z) - \delta = \dim |Z| - \delta.$$

(This condition is equivalent to: X is rational, cf. [7], 2.3(iv)).

(3.8) *Proof of* (3.7). (a) Let D be a characteristic divisor. Let ρ be the dimension of Λ. Clearly, the linear system corresponding to $H^0(C', \mathcal{O}(D))$ has dimension $h^0(D) - r$, and this number is $\geqq \rho$. This, together with (3.2.2), shows: $\dim T_C \leqq \dim H^0(C', \mathcal{O}(D))$, i.e. (a).

(b) Under the assumption "D_i is non special" by Riemann-Roch (for a curve with r connected components, $r \geqq 1$),

$$\dim H^0(C', \mathcal{O}(D) = \deg D - g + r = \deg N_C - 2\delta - g + r = \deg N_C - p_a(C) + 1 - \delta,$$

where we used the formula $p_a(C)-\delta+r-1=g$. For the other inequalities, use Riemann-Roch applied to C:

(3.8.1) $$\dim H^0(C, N_C) \geqq \chi(N_C) = \deg N_C - p_a(C)+1$$

(cf. [3], p. 79).

(c) Let $D=D_1+\ldots+D_r$ (D_i supported at C_i') be a characteristic divisor, $g_i = p_a(C_i')$. We want to check: $\deg D_i > 2g_i-2$. Let $E_i = \sum_{P\in C_i'} n_P P$ where $E = \sum_{Q\in C'} n_Q Q$ (cf. (3.5)), let $C_i = f(C_i')$ (f is the normalization map). Then, $\deg D_i = C_i \cdot C - \sum_{P\in C_i'} n_P$, by our definitions. On the other hand, $g_i = p_a(C_i) - \sum_{P\in C_i} \delta(C_i, P)$. Hence,

$$\deg D_i - (2g_i-2) = C_i^2 - (2p_a(C_i)-2) - [\sum_{P\in C_i'} n_P - 2\sum \delta(C_i, P) - \sum_{i\neq j} C_i \cdot C_j].$$

But, by the "adjunction formula": $2p_a - 2 = C_i(C_i+K)$ (cf. [3], p. 298 and 366). Our hypothesis implies: $C_i^2-(2p_a(C_i)-2)>0$, and according to the definition of n_P (cf. (3.5)), the bracket above is zero. Hence $\deg D_i > (2g_i-2)$, as wanted.

(d) By (c), $h^0(D) \leqq \chi(N_C)-\delta$. But $\chi(N_C)=\chi(N_Z)$, because C and Z are linearly equivalent. Now, $K\cdot Z = K \cdot C < 0$ (for the same reason); from the adjunction formula, $\deg N_Z = Z^2 > 2p_a(Z)-2$, hence $\dim H^0(Z, N_Z) = \chi(N_Z) = \chi(N_C)$, and $h^0(D) \leqq \dim H^0(Z, N_Z) - \delta$. To conclude the proof, we must check: $\dim |Z| = \dim H^0(Z, N_Z)$. But, by the adjunction formula, $\dim H^0(Z, N_Z) = \chi(N_Z) = p_a(Z) - Z\cdot K - 1$. On the other hand, using the sequence $0 \to \mathcal{O} \to \mathcal{O}(Z) \to N_Z \to 0$, its cohomology sequence, and our assumptions, one gets $H^1(\mathcal{O}(Z))=0$, $H^2(\mathcal{O}) = H^2(\mathcal{O}(Z))$. Then, the Riemann-Roch theorem for surfaces gives:

$$\dim |Z| = \tfrac{1}{2} Z(Z-K) = \tfrac{1}{2} Z\cdot(Z+K) - K\cdot Z = p_a(Z) - 1 - K \cdot Z,$$

by the adjunction formula for Z (this computation appears in [7], p. 115).

(3.9) *Remark.* The conditions of (d) automatically hold if $X = \mathbb{P}^2$.

In the case of nodes, we have more precise results.

(3.10) **Theorem.** *Let C be a reduced curve on a smooth projective surface X, whose only singularities are nodes $P_1, \ldots, P_\delta$, and having irreducible components $C_1, \ldots, C_r$. Assume $C_i \cdot K < 0$, $i=1, \ldots, r$ (K = canonical divisor of X). Then,* (a) $\mathscr{F}_C$ (cf. (1.3)) *is smooth,* (b) *its tangent space has dimension* $\dim H^0(C, N_C)-\delta$, *and* (c) *it is isomorphic to* $A_C = \{s \in H^0(C, N_C)$: *s is an adjoint to* $C\}$.

Proof. The assertion (c) was proved in (2.11). Concerning (b), the space A_C of adjoints is the intersection of the δ hyperplanes $L_i = \{s\in H^0(C, N_C)\colon s(P_i)=0\}$, $i=1, \ldots, \delta$, hence $\dim T_C = \dim A_C \geqq \dim H^0(C, N_C) - \delta$. On the other hand, by (3.7), parts (a) and (c), $\dim T_C \leqq \dim H^0(C, N_C) - \dim H^1(C, N_C) - \delta$. It follows:

(3.10.1) $$H^1(C, N_C) = 0$$

and $\dim T_C = \dim H^0(C, N_C) - \delta$, i.e., (b).

Concerning (a), recall again that for nodal curves, equisingular deformations (inside X) coincide with locally trivial ones. It is well known that, given a small surjection of algebras in $\mathscr{A}$, $A' \to A$, the obstructions to lift a locally trivial deformation over A to one over A' lie in $H^1(C, N'_C)$, where $N'_C = \operatorname{Ker}(N_C \to T^1_C)$ (cf. [7], (1.3)). Hence, $N'_{C'} \subset N_{C'}$ and we have a map:

$$\lambda \colon H^1(C, N'_C) \to H^1(C, N_C).$$

If λ is injective, then (3.10.1) and the remark just made imply the smoothness of $\mathscr{F}_C$. But, in our case,

$$0 \to N'_C \to N_C \xrightarrow{\eta} T^1_C \to 0$$

is exact. In fact, η is surjective, since $T^1_{C,P}$ is a skyscraper sheaf, having k as stalk at each node and zero otherwise (use $T^1_{C,P} \approx k[\![x,y]\!]/(f, f_x, f_y)$ where f is induced by a local equation of C at P). Hence we get an exact sequence.

$$0 \to H^0(N'_C) \to H^0(N_C) \xrightarrow{\varepsilon} H^0(T^1_C) \xrightarrow{\xi} H^1(N'_C) \xrightarrow{\lambda} H^1(N_C).$$

Let $N = \dim H^0(N_C)$. Now, $H^0(N'_C)$ is the tangent space of the functor (from $\mathscr{A}$ to (Sets)) of locally trivial deformations of C (inside X). This functor is isomorphic to $\mathscr{F}_C$, hence $H^0(N'_C) \approx A_c$ (by part (c)) and its dimension will be $N - \delta$ (by (b)). Now, $\dim H^0(T^1_C) = \delta$, hence ε must be surjective and $\operatorname{Ker} \lambda = \operatorname{Im} \xi = (0)$. This proves the injectivity of λ, and hence (a).

(3.11) *Remarks.* (a) From (3.10.1), verified in the course of the proof above from the hypothesis of (3.10), it follows: the Hilbert scheme of X is smooth at C (see [4], Lecture 23).

(b) The assumption "$C_i \cdot K < 0$" cannot be dropped: Tannenbaum has shown (in [8]) that on a $K-3$ surface (were $K=0$) the formula for the dimension of T_C fails.

Theorem (3.10) has the following converse.

(3.12) **Theorem.** *Let X be as in* (3.1), *$K = K_X$, $C \subset X$ a reduced curve, with irreducible components $C_1, \ldots, C_r$; assume $C_i \cdot K < 0$, $i = 1, \ldots, r$. Assume $P \in C$ is a singular point such that, in addition, one of the following conditions holds:*

(i) *C has at P a branch $\mathfrak{B}$ of multiplicity >1, and if C_i is the component of C determined by $\mathfrak{B}$* (cf. (0.6)), *then $(C_i \cdot K) < -1$.*

(ii) *C has at P two smooth branches $\mathfrak{B}_1$, $\mathfrak{B}_2$ with a common tangent, and moreover, let C_i, C_j be the irreducible components of C containing $\mathfrak{B}_1$, $\mathfrak{B}_2$ resp. If $C_i \neq C_j$, require $C_l \cdot K < -1$, $l = i, j$. If $C_i = C_j$, require $C_i \cdot K < -2$.*

(iii) *C has at P an ordinary m-fold point, $m \geqq 3$, and there are* 3 *branches $\mathfrak{B}_i$, $i = 1, 2, 3$ of C at P such that, if C_j, $j = 1, \ldots, s$, $s \leqq 3$ are the components of C containing these branches, and C_j contains exactly m_j of them, then $(C_j \cdot K) < -m_j$. Then,* $\dim T_C < \dim H^0(C, N_C) - \delta$.

Proof. It is very similar to that of Theorem 2 of [13]. Let $s \in T_C$ induce, on $\hat{\mathcal{O}}_{X,P} \approx k[\varepsilon][\![x,y]\!]$, a deformation $F = f + \varepsilon g$ (of $f = 0$, an equation of $\hat{C}_P$). This must be equisingular, say along the section determined by $(\alpha\varepsilon, \beta\varepsilon)$. We may assume $x = 0$ is not tangent to f at P. By (2.5), if we blow up along this section, so that F gets transformed into $F'=f'+\varepsilon g'$, we get

$$\omega = x\omega' + [\alpha y - \beta]\, dx \tag{3.12.1}$$

with $\omega=\omega(g)$, $\omega'=\omega(g')$. We consider separately the three cases.

Case (i). By our assumptions, if Q is the point in C' determined by $\mathfrak{B}$, and $\operatorname{mult}(\mathfrak{B})=m>1$, then x has order m at Q, and ω order $\geqq m-1>0$. By Remark (3.6.6), this means: $(s)-E$ contains Q. Thus, the series $\Lambda_0=\Lambda-Q$ (3.5) has same dimension as Λ. By our assumption on intersections, if $D_0\in\Lambda_0$, it is non-special and $\dim H^0(C',\mathcal{O}(D_0))=\dim H^0(C,N_C)-\delta-1-r$. This proves case (i).

Case (ii). Let $s\in T_C$. Let P_1, P_2 be the points on C' corresponding to $\mathfrak{B}_1$, $\mathfrak{B}_2$. One considers inside the characteristic Λ these divisors D containing P_1, they form a sub-system Λ_1, $\dim\Lambda_1=\rho_1\geqq\rho-1$. Using (3.11.1) and the fact that $\mathfrak{B}_1$, $\mathfrak{B}_2$ have a tangent in common (hence, the first infinitely near point in common), we see: D contains $P_1\Leftrightarrow\omega(P_1)=0\Leftrightarrow[\alpha a-\beta]=0$ ($y=ax$ being the common tangent) $\Leftrightarrow\omega(P_2)=0\Leftrightarrow D$ contains P_2. Hence, any divisor D_1 of Λ_1 has 2 fixed points, and $\rho_1\leqq\dim H^0(C',\mathcal{O}(D_1))-r$. By our assumptions on intersections, $\dim H^0(C',\mathcal{O}(D_1)=\dim H^0(C,N_C)-\delta-2$ (cf. proof of (3.8.d)). Hence, $\rho_1\leqq\dim H^0(C,N_C)-\delta-2-r$, and $\rho\leqq\dim H^0(C,N_C)-\delta-1-r$. This shows (b).

Case (iii). Let $y=\alpha_i x$ be the tangent lines to $\mathfrak{B}_i$, $i=1,2,3$. Let P_i be the points of C' corresponding to $\mathfrak{B}_i$, $i=1,2,3$. Let $\Lambda_2=\{D\in\Lambda\colon D \text{ contains } P_1,P_2\}$. Then, $\dim\Lambda_2\geqq\rho-2$. We claim: $D\in\Lambda_2$ implies D contains P_3. This is a consequence of (3.11.1). In fact, at P_i, $x=0$, $\omega=0$, $dx\neq0$, then $\alpha a_i-\beta=0$, $i=1,2$. Since $a_1\neq a_2$, $\alpha=\beta=0$. Then by (3.11.1), ω also vanishes at P_3. Then, $\dim\Lambda_2=\dim\Lambda'_2=\rho_2$, where $\Lambda'_2=\Lambda_2-P_1-P_2-P_3$, and $\rho_2\leqq\dim H^0(C',\mathcal{O}(D))$, $D\in\Lambda'_2$. As in (3.8), because of the assumption on intersection, we get $\dim H^0(C',\mathcal{O}(D))\leqq\dim H^0(C,N_C)-\delta-3$, whence $\dim\Lambda-\rho\leqq\dim H^0(C,N_C)-\delta-r-1$, which proves (c).

(3.13) *Note.* The conditions of (3.12) are automatically satisfied for $X=\mathbb{P}^2$.

§ 4. Some Global Constructions

(4.1) In this chapter we study certain subvarieties of the Hilbert scheme of a surface, which have certain universal properties relative to families of curves with singularities of specified equisingular classes.

We fix once and for the rest of this chapter a system

$$\mathcal{J}=(\mathcal{J}_1,\dots,\mathcal{J}_r)$$

of equisingular classes. X will denote a fixed smooth, projective surface.

(4.2) **Definition.** A family $D\to T$ of reduced curves on X (i.e., a relative Cartier divisor $D\subset X\times T$ such that the fibers D_t, $t\in T$, are reduced for all t) is an $\mathcal{J}$-*family* if there is a dense open set $U\subset T$ such that, for all $t\in T$, D_t has exactly r singular points, of class $\mathcal{J}_1,\dots,\mathcal{J}_r$ respectively.

The following proposition is clear.

(4.3) **Proposition.** *If the family $D \to T$ is in $\mathcal{F}_{X,\mathcal{I}}(T)$ ($\mathcal{I} = (\mathcal{I}_1 \dots, \mathcal{I}_r)$ and we use the notation of* (1.2)), *then it is an $\mathcal{I}$-family.*

In the other direction, we have:

(4.4) **Proposition.** *Let $f: Z \to T$ be any family of reduced curves on X, with T integral. Then, there are equisingular classes $\mathcal{I} = (\mathcal{I}_1, \dots, \mathcal{I}_r)$ (r a suitable integer), and an open dense set $U \subset T$, such that $f^{-1}(U) \to U$ is in $\mathcal{F}_{X,\mathcal{I}}(U)$.*

In other words, any f as above is an $\mathcal{I}$-family, for a suitable $\mathcal{I}$.

Proof. First, by replacing T by its open dense subset of smooth points, we may assume T smooth. We may assume the general fiber is singular (otherwise it is obvious). By using well-known results, in particular the fact that equisingularity is an open condition (cf. [12], II), we may assume the following. Over an open dense set $U \subset T$, we have, writing $S = \mathrm{Sing}(Z) \cap f^{-1}(U)$: (a) S is smooth, (b) the induced projection f: $S \to U$ is finite, etale, (c) Z is equisingular along S, at each $s \in S$. Now it is clear that this is the U we are looking for.

(4.5) (a) If $Z \to T$ is an $\mathcal{I}$-family, $\mathcal{I} = (\mathcal{I}_1, \dots, \mathcal{I}_r)$, we shall also say that $\mathcal{I}_1, \dots \mathcal{I}_r$ are the generic equisingular classes of the family.

(b) Given a family f as in (4.4), a point $t \in T$ is called *E-generic* (for the family) if there is an open subset U of T containing t, having the property of (4.4).

(4.6) **Theorem.** *Given a polynomial P with rational coefficients, and classes $\mathcal{I} = (\mathcal{I}_1, \dots, \mathcal{I}_r)$, then there is a reduced closed subscheme $V = V_{P,\mathcal{I}}$ of the Hilbert scheme H_P of X (of divisors on X with Hilbert polynomials P) such that the family $\mathcal{D} \to V$ (induced from the universal divisor over H_P) is universal with respect to $\mathcal{I}$-families* g: *$Z \to T$, with T reduced and such that all the fibers have Hilbert polynomial P.*

Proof. Consider, inside H_P, the set $V_0 = \{C \subset X$: C is a reduced divisor having r singular points of class $\mathcal{I}_1, \dots, \mathcal{I}_r$ respectively (and no other singularity)$\}$. Let $V = \bar{V}_0$ (regarded as a reduced closed subscheme of H_p). Since V_0 is dense in V, the generic equisingular classes of $\mathcal{D} \to V$ (induced by the universal family on H_p) will be $(\mathcal{I}_1, \dots, \mathcal{I}_r)$. Now, given any $\mathcal{I}$-family g: $Z \to T$ as in the statement (4.6), there is an open dense $U \subset T$ such that $g^{-1}(t)$ is a reduced curve having r singularities of class $\mathcal{I}_1, \dots, \mathcal{I}_r$ respectively. If $h: T \to H_P$ is the morphism defined by the universal property of H_p, then clearly $h(U) \subset V_0$.

Since T is reduced it follows that $h(T) \subset V$, i.e., h factors through V, as we wanted.

(4.7) Next we shall consider the question of dimensions. Let $\mathcal{I}, V$ be as in (4.6), Σ an irreducible component of V. Let K be the canonical divisor of X. If $C \subset V$ is a reduced curve, having irreducible components $C_1, \dots, C_s$ consider the following condition:

(4.7.1) $$C_i \cdot K < 0, \quad i = 1, \dots, s.$$

(4.8) **Theorem.** *Let $C \in \Sigma$ be E-generic, assume* (4.7.1) *holds. Then:* (a)

(4.8.1) $$\dim \Sigma \leqq \dim H^0(C, N_C) - \delta.$$

(b) *If, in addition, at least one of the conditions* (i), (ii), (iii) *of* (3.11) *holds, then in* (4.8.1) *the strict inequality sign holds.*

Proof. Since, by E-genericity, in a neighborhood U of C, $D_\Sigma \to \Sigma$ induces a family in $\mathscr{F}(U)(\mathscr{F}=\mathscr{F}_{X,\mathscr{S}})$, it follows that $T_{\Sigma,C} \subset T_C$ (notation of (3.1)). Since always $\dim \Sigma \leqq \dim T_{\Sigma,C}$ the results follow from (3.7) and (3.11).

(4.9) *Remark.* The imposed conditions are automatically fulfilled if $X=\mathbb{P}^2$ (cf. [13]).

(4.10) Of course, one would like a result, sharper than (4.6), namely one stating the representability of the functor $\mathscr{F}_{X,\mathscr{S}}$. Even in the case $X=\mathbb{P}^2$, it is known that in general the use of non-reduced schemes will be necessary. This may happen for instance when we deal with plane curves of degree n, with δ nodes and κ cusps, and $\kappa > 3n$ (cf. [10], 3.G). We hope to return to this question in the future. For the case of nodes, see § 5; for higher singularities (but $X=\mathbb{P}^2$, and in a different set-up), see [5], § 3.

§ 5. Nodal Curves

(5.1) More precise results can be obtained in the case where all the singularities are nodes. In this case, under suitable numerical conditions, the functor $\mathscr{F}$ (cf. (1.2.1)) can be represented by a locally closed smooth subscheme of the Hilbert scheme of X (see Theorem (5.6)).

Some other results of the previous section can be sharpened.

To simplify the presentation, in this chapter we work over the field $\mathbb{C}$ of complex numbers.

(5.2) Choose a polynomial P with rational coefficients. Let H^P be the part of the Hilbert scheme of X parametrizing effective divisors with Hilbert polynomial P.

For the remainder of this chapter, we shall use the following convention. If C is an effective divisor on X, then the corresponding small letter c will denote the closed point of H^P corresponding to C.

Assume H^P is non-empty. Consider, inside H^P, the set $U=\{c$: the corresponding divisor C is a curve having δ nodes and no other singularities$\}$. Let V be the closure of U in H^P.

(5.3) **Theorem.** *Assume H^P is smooth at each $c \in U$. Then:* (a) *U is Zariski-open in V.* (b) *Any $c \in U$ is an E-generic point of V* (cf. (4.5.b)).

Proof. Let $\pi\colon \mathscr{D} \to H^P$ be the universal curve, $\mathscr{D} \subset H^P \times X$. If M is a subscheme (or an analytic subspace) of H^P, $\pi_M\colon \mathscr{D}_M \to M$ will denote its pull-back over M. If $c \in U$, let $P_1, \ldots, P_\delta$ be the nodes of $\mathscr{D}_c = C \times \{c\}$. Then we may find neighborhoods $\mathscr{U}$ of c (in H^P), $\mathscr{U}_i$ of P_i (in $H^P \times X$), $i=1, \ldots, \delta$ such that $p(\mathscr{U}_i) \subset \mathscr{U}$, $i=1, \ldots, \delta$ (where $p\colon H^P \times X \to H^P$ is the first projection), $\mathscr{U}_i$ is isomorphic to a polydisk in $\mathbb{C}^{N+2}$, ($N=\dim H^P$) with coordinates $u_1^{(i)}, \ldots, u_N^{(i)}$, $x^{(i)}$, $y^{(i)}$, and where $\mathscr{D}$ is defined by $f^{(i)}=0$. Let $Z_i \subset \mathscr{U}_i$ be defined by $(f^{(i)}, f^{(i)}_{x^{(i)}}, f^{(i)}_{y^{(i)}}, h^{(i)})$, where $h^{(i)}$ is the usual hessian determinant. Let $Z=\bigcup_{i=1}^{\delta} p(Z_i)$. Clearly, $C \notin Z$ {because each

P_i is a node), and $V-Z$ contains an open containing c. This shows that U is open in the metric topology. Soon we'll see it is also Zariski-open too. Now consider $\pi_V\colon \mathscr{D}_V \to V$, $\mathscr{D}_V \subset V \times X$, let $c \in U$ be as before. We may choose (metric) opens G in U, $G_1, \ldots, G_\delta$ in $\mathscr{D}_V$, containing $c, P_1, \ldots, P_\delta$ (the nodes of $\pi_V^{-1}(c)$ resp., such that for $t \in G$, $\pi_V^{-1}(t) \cap G_i$ has only one node. It follows that the induced morphism of germs $(G_i, P_i) \to (G, c)$ is a deformation of $(\pi_V^{-1}(c) \cap G_i, P_i)$ satisfying "condition E_μ" of [9] (i.e., the Milnor number remains constant). By [9] Chap. III, 2.9, for G_i, G small enough, there is a section $s_i\colon G \to G_i$, $s_i(t) =$ only node of $\pi_V^{-1}(t) \cap G_i$, along which G_i is equisingular.

Now, $\mathscr{N} = \{P \in \mathrm{Sing}(\mathscr{D}_V) = S\colon \mathscr{D}_V$ is equisingular along S at P, so that $\pi_V^{-1}(\pi(P))$ has at P the type of a node$\}$ is Zariski open, and we just saw that $P_i \in \mathscr{N}$, $i = 1, \ldots, \delta$. It easily follows that $\pi(\mathscr{N})$ is a Zariski neighborhood of c, which is contained in U. Hence U is actually Zariski open. The rest of the Theorem is now clear.

(5.4) If we impose more conditions on a curve C (of (5.2)) on X, more can be said. Namely, recall condition (3.7.2), on a curve C in X, with irreducible components $C_1, \ldots, C_r$: $C_i . K < 0$, $i = 1, \ldots, r$, $K =$ canonical divisor of X. Note that, in view of the "adjunction formula" ([3], p. 366), $\deg N_C = C^2 > 2p_a(C) - 2$; if C is irreducible this implies: $H^1(C, N_C) = 0$ (cf. [4] p. 80), which implies: the Hilbert scheme is smooth at c ([4], p. 157).

Precisely, we have:

(5.5) **Theorem.** *Assume H^p is smooth at c and C satisfies* (3.7.2). *Then:*

(a) *$\hat{\mathcal{O}}_{U,c}$ (together with the element induced by the universal curve) pro-represents the functor $\mathscr{F}_C$* (1.3.1).

(b) *U is smooth at c and the component of U containing C has dimension $\dim H^0(C, N_C) - \delta$.*

Proof. (a) We check the universal property of the pro-representing object. Given $A \in \mathscr{A}$ (cf. 0.2), and $D \in \mathscr{F}_C(A)$, by universality of H there is a unique morphism φ: $\mathrm{Spec}\, A \to H$, $\varphi(\max(A)) = c$. Now by (1.5) and (3.10), $\mathscr{F}_C$ is a smooth, pro-representable functor. By using Grothendieck's existence theorem and the uniqueness of equisingular sections, we can lift the family D (and its equisingular sections) to a family defined over Spec(B), where B is a complete regular local $\mathbb{C}$-algebra. Next, using Artin's approximations lemma we may assume that the morphism φ fits in a commutative diagram:

$$\begin{tikzcd} Y \arrow[rr] & & H \\ & \mathrm{Spec}\, A \arrow[ul] \arrow[ur] & \end{tikzcd}$$

with Y a smooth algebraic scheme, and there is $D' \in \mathscr{F}(Y)$ (cf. (1.2)) inducing $D \in \mathscr{F}_C(A)$. Now, since $U \subset V$, by (4.6) clearly the morphism φ factors through U. Hence, there is an induced unique homomorphism $\hat{\mathcal{O}}_{U,c} \to A$, such that D is induced by the canonical element. This checks (a).

(b) This is an immediate consequence of (a) and (3.10).

We also have the following:

(5.6) **Theorem.** *Assume the hypothesis of Theorem* (5.5) *hold at each* $c' \in U$. *Then, U represents the functor $\mathscr{F}$* (cf. (1.2)).

Proof. Given $Z \in (\mathrm{Alg})$ and $D \in \mathscr{F}(Z)$, by universality of the Hilbert scheme H there is a unique morphism $\alpha: Z \to H$; such that D is induced by the universal family on H. Clearly, to show (5.6) we must show that α factors through U. By (4.6), if $z \in Z$, then $\alpha(z) \in U$. To see that $\alpha(Z) \subset U$ (as schemes) we must see: if $z \in U$, $t = \alpha(z)$, $\mathcal{O}_{U,t} = \mathcal{O}_{H,t}/I$ and $\alpha_Z: \mathcal{O}_{H,t} \to \mathcal{O}_{U,t}$ is induced by α, then $\bar{\alpha}(I) = 0$. But after passing to completions, $\hat{\mathcal{O}}_{U,t} = \hat{\mathcal{O}}_{H,t}/I'$, $I' = I\hat{\mathcal{O}}_{H,t}$ and by (5.5a), $\hat{\alpha}_z(I') = 0$ (where $\hat{\alpha}_Z$ is induced by α_Z). It follows $\alpha_Z(I) = 0$, whence (5.6), holds.

(5.7) *Remark.* As pointed out in (5.4), if C is irreducible then condition (3.7.2) $(C \cdot K < 0)$ implies that the Hilbert scheme H is smooth at c.

(5.8) In Sects. (5.9)–(5.13), we explain how the theory of assigned and virtual singularities for nodal curves (cf. [6], Anhang F, [5], § 4, or [7]) can be developed in this context. Our presentation is quite close to that of [7], p. 119–122, but we can remove several restrictive hypothesis that occur there.

(5.9) So, let C be a reduced curve in X (as in (5.1)), having δ nodes as only singularities. Fix $\delta_0 < \delta$ of these, $P_1, \ldots, P_{\delta_0}$, the "assigned nodes"). The questions are:

(a) does there exist a family $f: D \to T$ of curves on X, parametrized by a smooth integral curve T, such for some $t_0 \in T$, $D_{t_0} = C$; for $t \neq t_0$, D_t is a nodal curve with δ_0 nodes, $f^{-1}(T - \{t_0\}) \to T - \{t_0\}$ satisfies condition $E(\mathscr{I})$ (1.2.2), and there are sections s_i of $\hat{D}_{P_i} \to \mathrm{Spec}(\hat{\mathcal{O}}_{T,t_0})$ (notation of (1.2)) such that $\hat{D}_{P_i}$ is equisingular along s_i; here $P_1, \ldots, P_{\delta_0}$ are the assigned nodes on C. We could have replaced the condition on the existence of equisingular sections by the requirement of "analytic local triviality" (cf. [7], Def. 2.7), for nodes both concepts agree.

(b) When can we choose such a family so that the general curve is integral?

To study (a), introduce a functor $\mathscr{F}'_C = \mathscr{F}'_{C, P_1, \ldots P_{\delta_0}}: \mathscr{A} \to (\mathrm{Sets})$, $\mathscr{F}'_C(A) = \{D \subset \mathrm{Spec}\, A \times X : D$ is a relative divisor, the special fiber is C and there are equisingular sections s_i: $\mathrm{Spec}\, A \to D$ such that $s_i(\max A) = P_i\}$. Then we have:

(5.10) **Proposition.** *The functor $\mathscr{F}'_C$ is pro-representable.*
The proof is entirely similar to that of Theorem (1.5).

(5.11) **Theorem.** *Assume $C_i \cdot K_X < 0$ for all i, where C is as in* (5.9), *and $C_1, \ldots, C_r$ are its irreducible components. Then:* (a) *$\mathscr{F}'_C$ is smooth* (b) *its tangent space T'_C has dimension* $\dim H^0(C, N_C) - \delta_0$ (c) *T'_C is isomorphic to* $\{s \in H^0(C, N_C)$: $s(P_i) = 0$, $i = 1, \ldots, \delta_0\}$.

Proof. The proofs of (b) and (c) are entirely similar to those of statements (b) and (c) of Theorem (3.10). Concerning (a), introduce a sheaf N''_C as in [7] (1.6), i.e. near P_i, N''_C is isomorphic to N'_C (cf. (3.10) or [7] (1, 3)), and near other points, to N_C. So, we have a commutative diagram, with exact rows.

(5.11.1)
$$\begin{tikzcd}
0 \arrow[r] & N''_C \arrow[r] & N_C \arrow[r] & T''_C \arrow[r] & 0 \\
0 \arrow[r] & N'_C \arrow[r] \arrow[u] & N_C \arrow[r] \arrow[u] & T^1 \arrow[r] \arrow[u, "\gamma"] & 0
\end{tikzcd}$$

The upper row induces, by taking cohomology an exact sequence.

$$H^0(T_C'')\xrightarrow{\xi'}H^1(N_C'')\xrightarrow{\lambda'}H^1(N_C).$$

As in the proof of (3.10), $H^1(N_C)=0$ (C is a nodal curve and (3.7.2) holds), and the obstructions for $\mathscr{F}_C'$ lie in $H^1(N_C'')$. Hence, if λ' is injective, i.e., ξ' is the zero map, then $\mathscr{F}_C'$ will be smooth. But this follows by considering the following diagram, induced by (5.11.1):

$$\begin{tikzcd} H^0(T_C'') \arrow[r, "\xi'"] & H^1(N_C'') \\ H^0(T_C^1) \arrow[u, "\alpha"] \arrow[r, "\xi"] & H^1(N_C') \arrow[u] \end{tikzcd}$$

where α is surjective (both are sheaves concentrated in zero dimensional subschemes of C, and γ in (5.10.1) is surjective) and ξ is zero, as shown in the proof of (3.10.a).

(5.12) By standard techniques in deformation theory (in particular, Grothendieck's existence theorem and Artin's approximation lemma), one shows the existence of a family $D\to T$, T smooth, and equisingular sections $s_i^{(t)}$, $i=1,\dots,\delta_0$, $t\in T$, having the properties described in (5.9), and inducing the universal deformation pro-representing $\mathscr{F}_C'$ (for the existence of the sections, the key point is the uniqueness of equisingular sections in the infinitesimal case, cf. [11]). Moreover,

(5.12.1) $$\dim T=\deg N_C-\delta_0=p_a(C)-\delta_0-(C\cdot K)-1$$

(the latter, by using the adjunction formula). We call such a family a locally universal family (at C or at $t_0\in T$ corresponding to C).

Clearly this solves problem (a).

Finally we have:

(5.13) **Theorem.** *Assume C satisfies the condition of* (5.11) *and is "virtually connected with respect to the assignment $P_1,\dots,P_{\delta_0}$" (i.e., $C-\{P_1,\dots,P_{\delta_0}\}$ is connected). Then, the general member of a locally universal family* (at C, cf. (5.12)) *is irreducible.*

The proof is entirely similar to that of Theorem (2.13) of [7], and we shall not present it here. In fact, all what one needs is the existence of local universal families (5.12), and the dimension formula (5.12.1). Note that we may avoid the use of the linear system $|X_i|$ of [7], (2.13), because we do not need them in the construction of the local universal families.

Bibliography

1. Artin, M.: Algebraization of formal moduli. In: Global Analysis. Princeton Univ. Press 1969
2. Brieskorn, E., Knörrer, H.: Ebene algebraische Kurven. Cambridge: Birkhäuser 1981
3. Hartshorne, R.: Algebraic Geometry. Berlin Heidelberg New York: Springer 1977

4. Mumford, D.: Lectures on curves on an algebraic surface. Ann. of Math Studies **59**. Princeton: Princeton Univ. Press 1966
5. Nobile, A.: On families of singular plane projective curves. To appear in Ann. Mat. Pura Appl. (4)
6. Severi, F.: Vorlesungen über algebraische Geometrie. Leipzig: Teubner 1921
7. Tannenbaum, A.: Families of algebraic curves with nodes. Compositio Math. **41**, 107–126 (1980)
8. Tannenbaum, A.: Families of curves with nodes on K-3 surfaces. Math. Ann. **260**, 239–253 (1982)
9. Teissier, B.: Cycles évanescents, sections planes, et conditions de Whitney. Asterisque **7–8**, 285–362 (1973)
10. Wahl, J.: Deformations of plane curves with nodes and cusps. Ann. J. Math. **96**, 529–577 (1974)
11. Wahl, J.: Equisingular deformations of plane algebroid curves. Trans. Amer. Math. Soc. **193**, 143–170 (1974)
12. Zariski, O.: Studies in equisingularity, I and II. Amer. J. Math. **87**, 507–536, 972–1006 (1965)
13. Zariski, O.: Dimension theoretic characterization of maximal irreducible algebraic systems of plane nodal curves of a given order n and with a given number d of nodes. Amer. J. Math. **104**, 209–226 (1982)

Received September 6, 1983; received in final form March 13, 1984

Math. Z. 187, 471–480 (1984)

Mathematische Zeitschrift

Strong L^p-Solutions of the Navier-Stokes Equation in R^m, with Applications to Weak Solutions

Tosio Kato*

Department of Mathematics, University of California, Berkeley, California 94720, USA

Dedicated to Ralph S. Phillips

1. Introduction

In this paper we prove some results on the existence and decay properties of local and global solutions of the Cauchy problem for the Navier-Stokes equation on space domain $\Omega=\mathbb{R}^m$, $m=2,3,\ldots$:

$$\begin{aligned} &\partial_t u-\Delta u+(u\cdot\partial)u+\partial p=f(t), \qquad t>0, x\in\mathbb{R}^m, \\ &\partial\cdot u=0, \qquad u(0,x)=a(x). \qquad (\partial=\nabla=\mathrm{grad}) \end{aligned} \tag{NS}$$

We are mainly interested in the strong solutions $u(t)$ in $PL^m=PL^m(\mathbb{R}^m;\mathbb{R}^m)$ (and its subspaces), since they exist locally in time if the initial velocity a is in PL^m (without any differentiability for a), and globally if $\|a\|_m$ is sufficiently small. Here and in what follows we denote by PL^p the subspace of $L^p(\mathbb{R}^m;\mathbb{R}^m)$ characterized by the divergence condition $\operatorname{div} u=0$, and by $\|\cdot\|_p$ the associated norm. As usual we talk about the solution u of (NS), disregarding the pressure field p, which is automatically determined by u via (NS) up to an inessential additive function of time.

There is a large literature on strong and weak solutions of (NS), but not so much has been published for the L^p-theory for unbounded space domains Ω, including $\mathbb{R}^m$. Weissler [7] gives a detailed L^p-theory in $\mathbb{R}^m_+$ (half-space) for local solutions, but not much for global solutions. In fact $\mathbb{R}^m$ is an easy case, since there is no boundary to worry about. On the other hand, there are some problems arising from the lack of compactness. For example, the strong solutions $u(t)\in PL^m$ mentioned above are not necessarily weak solutions in the sense of Leray and Hopf, since they need not have finite energy $(=\|u(t)\|_2^2/2)$. Also, the decay of global solutions as $t\to\infty$ is a nontrivial problem. In fact the decay in L^2-norm of turbulent solutions, which are ultimately strong solutions, has been left open in the famous papers by Leray [3, 4].

Our main results are summarized in the following theorems, in which we assume for simplicity that $f=0$, although it is not difficult to include nonzero

* Partially supported by NSF Grant MCS-82-00171

f under appropriate conditions. The theorems are classified into two groups; in the first group we assume that $a \in PL^m$, and in the second that $a \in PL^m \cap PL^p$. (BC denotes the class of bounded and continuous functions.)

Theorem 1. *Let $a \in PL^m$. Then there is $T>0$ and a unique solution u such that*

$$t^{(1-m/q)/2} u \in BC([0,T); PL^q) \quad \text{for } m \leqq q \leqq \infty, \tag{1.1}$$

$$t^{1-m/2q} \partial u \in BC([0,T); PL^q) \quad \text{for } m \leqq q < \infty, \tag{1.1'}$$

both with values zero at $t=0$ except for $q=m$ in (1.1), in which $u(0)=a$. Moreover, u has the additional property

$$u \in L^r(0,T_1); PL^q) \quad \text{with } 1/r=(1-m/q)/2, \quad m<q<m^2/(m-2), \tag{1.2}$$

with some $0<T_1 \leqq T$.

Theorem 2. *There is $\lambda>0$ such that if $\|a\|_m \leqq \lambda$, then the solution u in Theorem 1 is global, i.e. we may take $T=T_1=\infty$. In particular, $\|u(t)\|_q$ decays like $t^{-(1-m/q)/2}$ as $t\to\infty$, including $q=\infty$, and $\|\partial u(t)\|_q$ decays like $t^{-(1-m/2q)}$, including $q=m$.*

Theorem 2′. *In Theorem 2, we have*

$$T^{-1}\int_0^T \|u(t)\|_m dt \to 0 \quad \text{as } T\to\infty. \tag{1.3}$$

Remark 1.1. (a) As is well known, the solution u is smooth for $t>0$. Therefore, the real interest in these theorems are in the behavior of u as $t\to 0$ and, in case u is global, as $t\to\infty$. Also the spatial decay expressed by the property $u(t)\in PL^q$ is of interest. Note that q is restricted by $q \geqq m$; we are not able to give any results for $q<m$ under the assumption $a\in PL^m$.

(b) It is not known to the author whether or not $\|u(t)\|_m$ tends to zero as $t\to\infty$. Theorem 2′ shows, however, that it is the case if $\|u(t)\|_m$ is monotonically nonincreasing. Since this is true for $m=2$, where the energy inequality is available, we have answered Leray's question for $m=2$ in the affirmative. For the case $m=3$, see Remark 1.2, (b), below.

(c) The first part of Theorem 1 is valid for more general equations than (NS); only the bilinearity of the nonlinear term in u and ∂u is used in its proof. Property (1.2), on the other hand, appears to be deeper and to depend on a more special property of the nonlinear term. It was proved by Giga [1] for a bounded space domain $\Omega \subset \mathbb{R}^m$.

Theorem 3. *Let $a\in PL^m \cap PL^p$, where $1<p<m$. Then the solution u given by Theorem 1 has the following additional properties.*

$$u \text{ and } t^{1/2}\partial u \in BC([0,T_2); PL^m \cap PL^p) \quad \text{with some } 0<T_2 \leqq T. \tag{1.4}$$

Theorem 4. *There is $0<\lambda_1 \leqq \lambda$ such that if $\|a\|_m \leqq \lambda_1$, the solution u in Theorem 3 is global, and we may set $T=T_1=T_2=\infty$ in (1.1-1′), (1.2) and (1.4). (λ_1 will be small if p is close to 1.) Moreover, we have for any finite $q \geqq p$,*

$$t^{(m/p-m/q)/2} u \quad \text{and} \quad t^{(m/p-m/q+1)/2} \partial u \in BC([1,\infty); PL^q) \tag{1.5}$$

provided that the exponent of t is smaller than 1 *(separately for u or ∂u); otherwise that exponent should be replaced by an arbitrary number smaller than* 1.

Remark 1.2. (a) In Theorem 4, the assumption is that $\|a\|_m$ be small; there is no restriction on the size of $\|a\|_p$. Indeed, the only L^p-norm that has an invariant meaning is the L^m-norm; other ones are not invariant under scale change in space-time.

(b) Theorem 4 shows that all L^q norms for $u(t)$ with $q>p$ decay as $t\to\infty$ with a definite rate. For the case $q=p$, we have the following result, which contains an affirmative answer to Leray's question for $m=3$, $p=2$.

Theorem 4′. *In Theorem 4, we have $\|u(t)\|_p\to 0$ as $t\to\infty$. More precisely, we have*

$$\|u(t)-u_0(t)\|_p=O(t^{-\delta/2}) \qquad \text{as } t\to\infty, \tag{1.6}$$

where $u_0(t)=e^{-tA}a$ (in the notation introduced below) is the Stokes flow for the initial velocity a and where δ is any positive number smaller than $\min\{1, m-m/p, m/p-1\}$. *Thus the decay rate of $\|u(t)\|_p$ is at least as fast as the slower of $\|u_0(t)\|_p$ and $t^{-\delta/2}$.*

2. Proof of Theorems 1, 2, 2′

The proofs of these theorems are similar to those given in Kato-Fujita [2], which are based on the theory of analytic semigroups and fractional powers of the generator. Here we do not use the fractional powers, employing instead various L^p-norms. The semigroup is used only for the convenience of notation; its property can be obtained directly from the explicit form of the heat kernel. Since otherwise the proofs are basically the same as in [2], we shall be brief in details.

1. Thus we start by rewriting (NS) in the abstract form

$$\partial_t u+Au+F(u)=0, \tag{ABS}$$

where $A=-P\Delta=-\Delta P$ and

$$F(u)=F(u,u), \qquad F(u,v)=P(u\cdot\partial)\,v. \tag{2.1}$$

Here P is the orthogonal projection of L^2 onto the subspace PL^2; as is well known, P is extended to a bounded operator on L^p to $PL^p, 1<p<\infty$. We note that since P commutes with the Laplacian Δ, A is essentially equal to $-\Delta$ and e^{-tA} is essentially the heat operator. (ABS) is then converted into the integral equation

$$u=u_0+Gu, \tag{INT}$$

where

$$u_0(t)=e^{-tA}a, \qquad Gu(t)=-\int_0^t e^{-(t-s)A}F(u(s))\,ds. \tag{2.2}$$

The basic estimates we use are

$$\|e^{-tA}u\|_q \leqq ct^{-(m/p-m/q)/2}\|u\|_p, \qquad 1<p\leqq q<\infty. \tag{2.3}$$

$$\|\partial e^{-tA}u\|_q \leqq ct^{-(1+m/p-m/q)/2}\|u\|_p, \tag{2.3'}$$

$$\|F(u,v)\|_p \leqq c\|u\|_r\|\partial v\|_s, \qquad 1/p=1/r+1/s. \tag{2.4}$$

Here c denotes various constants that do not depend on the individual functions u, v. (2.4) is simply the Hölder inequality. (2.3) follows from the fact that the $L^p(\mathbb{R}^m)$-norm of the heat kernel is proportional to $t^{-(1-1/p)m/2}$; recall that $A=-P\Delta$ is essentially identical with $-\Delta$. Similarly (2.3′) follows from the property of the derivatives of the heat kernel.

Combining (2.3–3′) with (2.4) we obtain, with a slight change of notation,

$$\|Gu(t)\|_{m/\gamma} \leqq c\int_0^t (t-s)^{-(\alpha+\beta-\gamma)/2}\|u(s)\|_{m/\alpha}\|\partial u(s)\|_{m/\beta}ds, \tag{2.5}$$

$$\|\partial Gu(t)\|_{m/\gamma} \leqq c\int_0^t (t-s)^{-(1+\alpha+\beta-\gamma)/2}\|u(s)\|_{m/\alpha}\|\partial u(s)\|_{m/\beta}ds, \tag{2.5'}$$

where

$$\alpha, \beta, \gamma>0, \qquad \gamma\leqq\alpha+\beta<m. \tag{2.5''}$$

Of course (2.5) will be useless unless $\alpha+\beta-\gamma<2$, and (2.5′) unless $\alpha+\beta-\gamma<1$.

2. We shall now solve (INT) by successive approximation, starting with u_0 given u_0 given by (2.2) and constructing

$$u_{n+1}=u_0+Gu_n, \qquad n=0,1,2,\ldots. \tag{2.6}$$

First we show by induction that the u_n exist and satisfy the following estimates:

$$t^{(1-\delta)/2}u_n \in BC([0,\infty);\ PL^{m/\delta}) \quad \text{with norm} \leqq K_n, \tag{2.7}$$

$$t^{1/2}\partial u_n \in BC([0,\infty);\ PL^m) \quad \text{with norm} \leqq K_n'. \tag{2.8}$$

Here δ is any fixed number, $0<\delta<1$. Moreover, all the functions (2.7–8) vanish at $t=0$.

It is easy to see that (2.7–8) are true for $n=0$; we have only to apply (2.3–3′) with $q=m/\delta$ and $q=m$, respectively, and $p=m$, $u=a$. Thus we can take

$$K_0=K_0'=c\|a\|_m. \tag{2.9}$$

The continuity at $t=0$, with values zero, of the functions (2.7–8) with $n=0$ follows from the facts that the operator $t^{(1-\delta)/2}e^{-tA}$ is uniformly bounded from PL^m to $PL^{m\delta}$ and tends to zero strongly as $t\to 0$, and similarly for $t^{1/2}\partial e^{-tA}$ from PL^m to itself.

Assuming now that (2.7–8) are true for n, we proceed to prove them for $n+1$. The first term $u_0(t)$ on the right of (2.6) has been estimated. The second term $Gu_n(t)$ is estimated in $L^{m/\delta}$-norm by setting $\gamma=\alpha=\delta$, $\beta=1$ in (2.5). It follows that the required norm does not exceed

$$cK_nK_n'\int_0^t (t-s)^{-1/2}s^{-1+\delta/2}=cK_nK_n't^{-(1-\delta)/2}ds.$$

Thus we have proved (2.7) for $n+1$, with

$$K_{n+1}\leqq K_0+cK_nK_n'. \tag{2.10}$$

Similarly, differentiating (2.6) and applying (2.5') with $\alpha=\delta, \beta=\gamma=1$, we obtain (2.8) for u_{n+1}, with

$$K_{n+1}'\leqq K_0'+cK_nK_n'. \tag{2.11}$$

These computations also justify the continuity at $t=0$, with values zero, of the functions (2.7–8) with subscript $n+1$. Indeed, since this is true for subscript n by induction hypothesis, the constants K_n, K_n' can be made arbitrarily small if we restrict ourselves to a small time interval $[0,\tau)$. Since the same is true for K_0, K_0' as shown above, it follows from (2.10–11) that K_{n+1}, K_{n+1}' are also arbitrarily small when restricted to $[0,\tau)$.

The system of recurrence inequalities (2.10–11) can be solved easily, as was shown in [2]. It turns out that there is a number $\lambda>0$ such that if $K_0\leqq\lambda$, then K_n and K_n' are bounded by a fixed constant K. In view of (2.9), this is true if $\|a\|_m$ is sufficiently small. In this case the sequences (2.7–8) are uniformly bounded, and their uniform convergence on $(0,\infty)$ can be proved as in [2].

3. This does not yet give a limit function u in $BC([0,\infty); PL^m)$. To obtain u in PL^m, we have to estimate the u_n in L^m-norm. We shall do this more generally for an arbitrary L^q-norm, where $m\leqq q<\infty$. Returning to (2.6) and using (2.5) with $\gamma=m/q$, $\alpha=\delta$, $\beta=1$, we obtain

$$\begin{aligned}\|u_{n+1}(t)\|_q &\leqq \|u_0(t)\|_q+cK_nK_n'\int_0^t (t-s)^{-(1+\delta-m/q)/2}s^{-(1-\delta/2)}ds\\ &\leqq Kt^{-(1-m/q)/2}\end{aligned} \tag{2.12}$$

because K_nK_n' is bounded as shown above. (We denote by K different constants that depend on u.) Moreover, the convergence of the $u_n(t)$ in this norm can be proved as above. It follows that the limit function u is in $BC([0,T); L^m)$ (the case $q=m$) and u satisfies (1.1) for $m<q<\infty$. The continuity of functions (1.1) up to $t=0$ can be proved in the same way as above.

Similarly we can deal with ∂u to prove (1.1'); in this case we have to restrict ourselves first to $q<m/\delta$, but a bootstrap argument with a smaller β will allow us to attain any $q<\infty$. Finally, the case $q=\infty$ in (1.1) is taken care of by the Gagliardo-Nirenberg inequality $\|u(t)\|_\infty^2\leqq c\|u(t)\|_{2m}\|\partial u(t)\|_{2m}$. This completes the proof of Theorem 2 except for (1.2).

If $\|a\|_m$ is not small, the above argument fails. In this case we have to restrict the time to some finite interval $[0,T]$ such that (2.7–8) for $n=0$ hold with $K_0, K_0'\leqq\lambda$. Such a T always exists if $a\in PL^m$, since the functions (2.7–8) for $n=0$ are continuous with values zero at $t=0$ shown above. This leads to a local solution u, for which the above results hold on the interval $[0,T)$. (For details cf. [2].) This completes the proof of Theorem 1 except for (1.2). (The proof of uniqueness is similar to that in [2].)

4. The proof of (1.2) requires new ideas. One of them is the basic estimate

$$u_0 \in L^{q,r}(0,\infty) \equiv L^r((0,\infty); PL^q),$$
$$1/r = (1-m/q)/2, \qquad m<q<m^2/(m-2), \tag{2.13}$$

which has been proved by Giga [1] and in which the associated norm is arbitrarily small if $\|a\|_m$ is sufficiently small.

We can estimate the $L^{q,r}(0,\infty)$-norm of the approximate solutions u_n starting from (2.13). Or we may prefer to estimate u directly. To avoid the singularity of $u(t)$ at $t=0$, we may then estimate the L^r-norm on (ε,∞) and later go to $\varepsilon\to 0$. In any case we have to estimate the $L^{q,r}$-norm for Gu, and we shall do this formally, as if it were known that the norm in question is finite.

For this purpose it is convenient to observe that $(u\cdot\partial)u=\partial\cdot(uu)$. Using an analog of (2.5′), we thus obtain

$$\|Gu(t)\|_{m/\gamma} \leqq c\int_0^t (t-s)^{-(\alpha+\beta+1-\gamma)/2}\|u(s)\|_{m/\alpha}\|u(s)\|_{m/\beta}ds.$$

Choosing $\alpha=\beta=\gamma=m/q$ gives

$$\|Gu(t)\|_q \leqq c\int_0^t (t-s)^{-(1+m/q)/2}\|u(s)\|_q^2\,ds.$$

An application of the Hardy-Littlewood inequality thus leads to the inequality (cf. [1])

$$\|u\|_{q,r} \leqq \|u_0\|_{q,r} + c\|u\|_{q,r}^2, \tag{2.14}$$

where the time interval may be any $(0,T)$, $T>0$. It follows that $\|u\|_{q,r}$ can be estimated in terms of $\|u_0\|_{q,r}$ if the latter is sufficiently small, which is true by (2.13) if T is sufficiently small. If in paticular $\|a\|_m$ is sufficiently small, the same is true of $\|u_0\|_{q,r}$ on $(0,\infty)$ so that we can take $T=\infty$. This proves (1.2) and completes the proof of Theorems 1 and 2.

5. Finally we prove Theorem 2′. Since $u_0(t)\to 0$ in L^m as $t\to\infty$, it suffices to prove (1.3) with u replaced by Gu. Going back to (2.5), we set $\gamma=1$, $\alpha=1-2\kappa$, $\beta=4\kappa$, where κ is a small positive number, say $\kappa<1/2$. By (2.13) we have

$$\|u(\cdot)\|_{m/\alpha} \equiv \psi \in L^{1/\kappa}(0,\infty). \tag{2.15}$$

Also $\|\partial u(s)\|_{m/\beta} \leqq Ks^{-(1-2\kappa)}$ by (1.1). Thus

$$\|Gu(t)\|_m \leqq c\int_0^t (t-s)^{-\kappa}s^{-(1-2\kappa)}\psi(s)\,ds. \tag{2.16}$$

Now we integrate (2.16) on $0<t<T$, obtaining

$$\int_0^T \|Gu(t)\|_m dt \leqq c\int_0^T (T-s)^{1-\kappa}s^{-(1-2\kappa)}\psi(s)\,ds.$$

We split the integral on the right into two parts coming from the subintervals $(0,\tau)$ and (τ,T). The contribution from $(0,\tau)$ is $O(T^{1-\kappa})$ as $T\to\infty$. The contribution from (τ, T) is estimated by the Hölder inequality and is found to be majorized by T times the $L^{1/\kappa}$-norm of ψ taken over the interval (τ, T). Thus

$$\limsup T^{-1} \int_0^T \|Gu(t)\|_m dt \leqq c\|\psi\|_{1/\kappa}, \quad T\to\infty, \tag{2.17}$$

where the last norm is taken over (τ, ∞). But the latter tends to zero as $\tau\to\infty$, since $\psi\in L^{1/\kappa}(0,\infty)$. Thus the left member of (2.17) is zero, and we have proved Theorem 2′.

3. Proof of Theorems 3, 4, 4′

1. If we assume $a\in PL^m\cap PL^p$, the solution $u\in C([0,T);\ PL^m)$ exists by Theorem 1. Since we do not know yet whether or not $u(t)\in L^p$, we have to go back to the approximate solutions u_n and estimate their L^p-norms. We shall show that

$$\|u_n(t)\|_p\leqq M_n, \quad n=0,1,2,\dots,\ t\in[0,T), \tag{3.1}$$

and deduce recurrence estimates for the constants M_n.

(3.1) is obvious for n=0; we may take $M_0=c\|a\|_p$. Assuming (3.1) for n, we estimate the second term Gu_n on the right of (2.6) in L^p-norm, using (2.5) in which $\alpha=\gamma=m/p$, and $0<\beta<1$. Noting that

$$\|\partial u_n(s)\|_{m/\beta}\leqq Ks^{-(1-\beta/2)}, \tag{3.2}$$

which is the analog of (1.1′) for ∂u_n with $q=m/\beta$, we thus obtain

$$\begin{aligned}\|u_{n+1}(t)\|_p&\leqq M_0+c\int_0^t (t-s)^{-\beta/2}s^{-(1-\beta/2)}\|u_n(s)\|_p\\ &\leqq M_0+cKM_n.\end{aligned} \tag{3.3}$$

This proves (3.1) for n replaced by $n+1$, with the inequality

$$M_{n+1}\leqq M_0+cKM_n. \tag{3.4}$$

The linear recurrence inequality (3.4) has a bounded solution $M_n\leqq M$ if $cK<1$, and leads to the required estimates for u in the L^p-norm. Then the estimate for ∂u is obtained from (2.5′).

The estimates obtained above are valid for $t\in[0,T)$ on which $u(t)\in L^m$ exists, assuming that $cK<1$. Recall that K is the constant bounding the size of u_n and u in the L^m-norm. If $\|a\|_m$ is sufficiently small, we can achieve $cK<1$ with $T=\infty$ (see Sect. 2) and obtain Theorem 4 except for the estimates (1.5). If $\|a\|_m$ is not small, we have only to limit the size of T to make cK small enough (see Sect. 2). Then we obtain Theorem 3.

It remains to prove (1.5) when $T=\infty$. Since $\|u_0\|_q$ satisfies (1.5), it suffices to estimate $\|Gu(t)\|_q$ and its derivatives by using (2.5–5′). First suppose that $m/p-m/q<2$. Then we choose $\alpha=m/p$, $\gamma=m/q$, and β sufficiently small that $\alpha-\gamma+\beta<2$. Since $\|u(s)\|_p$ is bounded and $\|\partial u(s)\|_{m/\beta}\leqq Ks^{-(1-\beta/2)}$ by (1.1′), (2.5) leads to the first assertion in (1.5). If $m/p-m/q\geqq 2$, we choose p' such that $p<p'<m$ and $m/p'-m/q$ is smaller than 2 but close to 2. Then we can apply the above result to conclude that $u(t)$ has the decay rate described in Theorem 4. Note that if $a\in PL^m\cap PL^p$, then $a\in PL^{p'}$ for any p' with $p<p'<m$, and all the results proved for p remain valid for p'. The decay for ∂u can be handled in the same way.

2. Next we prove Theorem 4′. Since $a\in L^p$ implies that $u_0(t)\to 0$ in L^p as $t\to\infty$, it suffices to prove (1.6). To this end we write

$$Gu(t)=G_0u(t)+G_1u(t), \tag{3.5}$$

by splitting the integral for Gu given in (2.2) into two parts coming from the subintervals $(0,1)$ and $(1,t)$, and estimate each term by an analog of (2.5).

To estimate $G_0u(t)$, we use

$$\alpha=\delta+m/p-1, \qquad \beta=1, \qquad \gamma=m/p; \tag{3.6}$$

note that $\alpha>0$ and $\alpha+\beta=\delta+m/p<m$ by the definition of δ. Also we have

$$\|u(s)\|_{m/\alpha}=O(s^{-(1-\delta)/2}), \qquad s\to 0. \tag{3.7}$$

Indeed, (3.7) is true by (1.1) if $\alpha<1$, while the left member is bounded as $s\to 0$ if $\alpha\geqq 1$ as is seen from (1.4). Moreover, $\|\partial u(s)\|_{m/\beta}=O(s^{-1/2})$ by (1.1). Hence

$$\|G_0u(t)\|_p\leqq\int_0^1(t-s)^{-\delta/2}s^{-(1-\delta/2)}ds=O(t^{-\delta/2}). \tag{3.8}$$

To estimate $G_1u(t)$, we use

$$\alpha=\delta, \qquad \beta=\gamma=m/p. \tag{3.9}$$

Again $\alpha+\beta=\delta+m/p<m$ is satisfied, and $\|u(s)\|_{m/\alpha}=O(s^{-(m/p-\delta)/2})$, $\|\partial u(s)\|_{m/\beta}=O(s^{-1/2})$ as $s\to\infty$, by (1.5). It follows that

$$\|G_1u(t)\|_p\leqq\int_1^t(t-s)^{-\delta/2}s^{-(1+m/p-\delta)/2}ds. \tag{3.10}$$

Here $(1+m/p-\delta)/2>1$ by the definition of δ. In view of Lemma 3.1 given below, (3.10) implies that $\|G_1u(t)\|_p=O(t^{-\delta/2})$. Combined with (3.8), this proves (1.6).

Lemma 3.1. *If* $0<\alpha<1<\beta$, *then*

$$\int_1^t(t-s)^{-\alpha}s^{-\beta}ds\leqq ct^{-\alpha}, \qquad t\geqq 1.$$

The proof of this lemma is elementary and may be omitted.

4. Some Applications to Weak Solutions

A *weak solution* w of (NS) is usually supposed to have the following properties.

$$w \in C_w([0, \infty); PL^2), \tag{4.1}$$

$$\partial w \in L^2((0, \infty); L^2), \tag{4.2}$$

$$\partial_t(w, \phi) + (\partial w, \partial \phi) + ((w \cdot \partial) w, \phi) = 0 \tag{4.3}$$

for any $\phi \in \mathscr{D}(\mathbb{R}^m; \mathbb{R}^m)$ with $\partial \cdot \phi = 0$,

$$\|w(t)\|_2^2 \leqq \|w(0)\|_2^2 + \int_0^t \|\partial w(s)\|_2^2 ds, \qquad t > 0. \tag{4.4}$$

(energy inequality).

Here C_w denotes the class of weakly continuous functions; $(\cdot, \cdot)$ denotes the L^2-inner product.

Sometimes one requires the energy inequality (4.4) not only on the interval $(0, t)$ but on all intervals $(t_0, t_1) \subset (0, \infty)$ except possibly for a set of Lebesgue measure zero for $t_0 > 0$. We shall call such w a *turbulent solution*, following Leray [4].

Our results given in preceding sections have some consequences on weak and turbulent solutions, which we shall discuss briefly.

1. Uniqueness. If w is a weak solution with $w(0) \in PL^m$, then w is a strong solution on some interval $[0, T)$ with $T > 0$ (hence $w(t)$ is smooth for $0 < t < T$); moreover, we may set $T = \infty$ if $\|w(0)\|_m$ is sufficiently small. Indeed, Theorem 1 shows that there is a strong solution $u \in BC([0, T); PL^m)$ with $T > 0$ and $u(0) = w(0)$, satisfying (1.1), (1.1'), and (1.2). But (1.2), with any $q > m$, is exactly the uniqueness criterion given by Serrin [5]. It follows that $w = u$, as required. It may be noted that $u \in C([0, T); PL^m)$ already suffices for this purpose if one uses a recent result of Sohr and von Wahl [6] that generalizes Serrin's theorem.

2. Leray's theorem in $\mathbb{R}^4$. The celebrated structure theorem of Leray [4] says that if w is a turbulent solution in $\mathbb{R}^3$, $w(t)$ is smooth except for t in a compact set $\Sigma \subset [0, \infty)$ of Lebesgue measure zero. The proof of this theorem depends on the following facts. (a) (4.2) implies that $w(t_0) \in H^1(\mathbb{R}^3) \subset L^3(R^3)$ for almost all t_0 so that the uniqueness result given above applies on an interval $[t_0, t_0 + T)$, provided the energy inequality holds on that interval, which is true for almost all t_0 for a turbulent solution. (b) (4.2) also implies that $\|w(t_0)\|_3$ is sufficiently small for some t_0 that the strong solution mentioned above exists for $t_0 \leqq t < \infty$. Since $H^1 \subset L^4$ is true in $\mathbb{R}^4$, we conclude from paragraph 1 above that the structure theorem is true in $\mathbb{R}^4$ as well.

It should be noted that Leray also shows that $\Sigma |I_n|^{1/2} < \infty$, where the I_n are the finite components of the open set $(0, \infty) \setminus \Sigma$ and $|\cdot|$ denotes the length. There seems to be no analogous inequality for $\mathbb{R}^4$, however.

3. The above result naturally raises the questions: do turbulent solutions exist in $\mathbb{R}^4$? The question is not at all trivial, as will be seen by looking at Leray's proof for $\mathbb{R}^3$, although it may be trivial in a bounded domain in $\mathbb{R}^m$. In fact it turns out that Leray's proof works in $\mathbb{R}^m$ with $m = 4$, but probably

not with $m>4$. The essential part of the proof is a compactness of the set of approximate solutions w_n used in constructing the weak solution w, in the sense that they are uniformly small at infinity in L^2-norm. To prove this, Leray uses a cut-off function f such that $1-f\in\mathscr{D}$, and estimates fw_n in L^2-norm. For this purpose, it is essential that

$$\int_0^T \|p_n(t)\|_2\, dt \tag{4.5}$$

is uniformly bounded for fixed $T>0$, where p_n is the pressure accompanying w_n. Since $-\Delta p_n = \partial\cdot(J_n w_n\cdot\partial)w_n$, where J_n is a certain mollifier, we have

$$p_n = g*\partial\cdot(J_n w_n\cdot\partial)w_n = \partial g*(J_n w_n\cdot\partial)w_n, \tag{4.6}$$

where $g(x)=c|x|^{-2}$ is the fundamental solution for $-\Delta$ for $m=4$. It follows that $\|p_n\|_2 \leqq c\|(J_n w_n\cdot\partial)w_n\|_{4/3} \leqq c\|w_n\|_4\|\partial w_n\|_2 \leqq c\|\partial w_n\|_2^2$, which shows that (4.5) is uniformly bounded, since the ∂w_n are bounded in $L^2((0,\infty);L^2)$.

This computation shows that $m=4$ is a critical case and the proof would not work for $m>4$. Even if one uses a different method for constructing w, it is not likely that one can spare a corresponding compactness proof.

References

1. Giga, Y.: Regularity criteria for weak solutions of the Navier-Stokes system. Proc. AMS Summer Inst. 1983, to appear
2. Kato, T., Fujita, H.: On the non-stationary Navier-Stokes system. Rend. Sem. Mat. Univ. Padova **32**, 243–260 (1962)
3. Leray, J.: Étude de diverses équations intégrales non linéaires et de quelques problèmes que pose l'Hydrodynamique, J. Math. Pures Appl. **12**, 1–82 (1933)
4. Leray, J.: Sur le mouvement d'un liquide visqueux emplissant l'espace. Acta Math. **63**, 193–248 (1934)
5. Serrin, J.: The initial value problem for the Navier-Stokes equation. Nonlinear Problems, Proc. Symp. MRC, Univ. Wisconsin, Madison, 1963, pp. 69–98
6. Sohr, H., von Wahl, W.: On the singular set and the uniqueness of weak solutions of the Navier-Stokes equations. Sonderforschungsbereich 72 Approximation und Optimierung, Universität Bonn, Preprint series no. 635 (1984)
7. Weissler, F.: The Navier-Stokes initial value problem in L^p. Arch. Rational Mech. Anal. **74**, 219–230 (1980)

Received March 26, 1984

Noted added on May 18, 1984

Professor K. Masuda has kindly pointed out that the conclusion of Theorem 2′ can be improved to: $\lim\|u(t)\|_m\to 0$ as $t\to\infty$, provided $\|a\|_m$ is sufficiently small. The proof depends on the continuous dependence of $a\to u$ from PL^m to $BC([0,\infty);PL^m)$, which can be proved by the method used in the proof of Theorem 1. Since any $a\in PL^m$ can be approximated in L^m-norm by functions $a_n\in PL^m\cap PL^p$, where $p\in(1,m)$ is fixed, and since the corresponding solutions u_n are in $BC_0([0,\infty);PL^m)$ by Theorem 4, where BC_0 denotes the closed subspace of BC consisting of functions that vanish at infinity, it follows immediately that $u\in BC_0$ too.

Math. Z. 187, 481–490 (1984)

Mathematische Zeitschrift

Spectrum and Holonomy of the Line Bundle over the Sphere

Ruishi Kuwabara

Department of Mathematics, College of General Education, The University of Tokushima, Minami-Josanjima, Tokushima 770, Japan

1. Introduction

Let Δ be the Laplace-Beltrami operator on a compact Riemannian manifold (M, g). Weinstein [7, 8], Guillemin [3, 4], Colin de Verdière [1] etc. studied the spectrum of the Schrödinger operator $H=\Delta+V$, V being the scalar potential. In particular, when (M, g) is the unit n-sphere (S^n, g_0), the spectrum of Δ consists of

$$\lambda_k=k(k+n-1), \qquad k=0, 1, 2, \ldots,$$

and that of H consists (except for finitely many values) of clusters of eigenvalues in the interval

$$[\lambda_k+\min \bar{V}(\gamma)-\varepsilon, \lambda_k+\max \bar{V}(\gamma)+\varepsilon],$$

for any $\varepsilon>0$, where $\gamma=\gamma(t)$ $(0\leqq t\leqq 2\pi)$ is a closed geodesic of (S^n, g_0) and

$$\bar{V}(\gamma)=\frac{1}{2\pi}\int_0^{2\pi} V(\gamma(t))\,dt.$$

Moreover, they showed how eigenvalues in the k-th cluster as $k\to\infty$.

In this article we consider the spectrum of the Schrödinger operator with a "vector" potential (or "magnetic" potential). The vector potential is related with a linear connection on a complex line bundle. Let E be a C^∞ Hermitian line bundle over (M, g) and $\tilde{d}$ be a linear connection compatible with the Hermitian structure (see [5, 9]). Let e be an unitary local frame of E, i.e. a local cross-section of E such that $|e|=1$. Then $\tilde{d}e=\omega e$ holds with a purely imaginary 1-form ((co-)vector potential) $\omega=i\alpha$ (α: real) on the open set of M. For the linear connection $\tilde{d}$, there can be naturally defined a non-negative, second order, (formally) self-adjoint, elliptic differential operator L which operates on $C^\infty(E)$, the space of cross-sections of E with the natural inner product. The operator L is called the *Bochner-Laplacian* (or *Laplacian*, for short), and is locally expressed as

$$L=-\sum_{j,k=1}^{n} g^{jk}\nabla_j\nabla_k-2i\sum_{j=1}^{n} a^j\nabla_j+\sum_{j=1}^{n}(a_j a^j-i\nabla_j a^j),$$

where ∇ is the Levi-Civita connection defined by g and $\alpha=\sum a_j dx^j$, $a^j=\sum g^{jk}a_k$ (see [5]).

We will study the spectrum of the operator L on the Hermitian line bundle with linear connection $(E,\tilde{d})$ over (S^n, g_0). Particularly, we will clarify how eigenvalues of Δ might split under the influence of the vector potential. (This phenomenon is known as the Zeeman effect in a physical example.) The discussion is developed mainly based on the work by Colin de Verdière [2].

2. Result

Let $(E,\tilde{d})$ be a Hermitian line bundle with linear connection over S^n. We recall the holonomy of the linear connection. Let $c=c(t)$ $(0\leqq t\leqq T)$ be a closed curve in S^n, and $\tilde{c}(t)$ a parallel lift of $c(t)$ with respect to the connection $\tilde{d}$. Then there is $q\in\mathbb{C}$ such that $\tilde{c}(T)=q\cdot\tilde{c}(0)$. We call q the *holonomy* of $\tilde{d}$ along c, and denote it by $Q_{\tilde{d}}(c)$. Using the curvature form Ω of $\tilde{d}$, we have

$$Q_{\tilde{d}}(c)=\exp\Big(-\int_{\Sigma_c}\Omega\Big), \tag{2.1}$$

where Σ_c is a surface in S^n with $\partial\Sigma_c=c$. When $\tilde{d}$ is compatible with the Hermitian structure, $|Q_{\tilde{d}}(c)|=1$ holds. Thus we have

$$Q_{\tilde{d}}\colon \text{the set of closed curves in } S^n\to S^1=\{e^{2\pi i\theta}\colon 0\leqq\theta\leqq 1\}.$$

Let S^*S^n be the unit cosphere bundle over S^n. Each element (x,ξ) of S^*S^n corresponds to a closed geodesic $\gamma(t)$ of (S^n,g_0) by $x=\gamma(0)$, $\xi=\dot{\gamma}(0)^*$ ($*\colon TS^n\to T^*S^n$ being the bundle isomorphism defined by g_0). Hence, from $Q_{\tilde{d}}$ we have a C^∞ map

$$\bar{Q}_{\tilde{d}}\colon S^*S^n\to S^1.$$

On the manifold S^*S^n there can be defined the induced volume form $d\,\mathrm{vol}$ from the symplectic volume form $dx^1\wedge\ldots\wedge dx^n\wedge d\xi_1\wedge\ldots\wedge d\xi_n$ on T^*S^n (cf. [3, §4]).

Now let L be the Laplacian defined on $(S^n, g_0; E,\tilde{d})$ and

$$(0\leqq)\,\mu_0\leqq\mu_1\leqq\mu_2\leqq\cdots$$

be the set of eigenvalues of L (denoted by $\mathrm{Spec}(L)$). Set

$$J=[a,b]\subset[0,1],$$

and

$$J_k=[\lambda_k+aC_k,\lambda_k+bC_k],\qquad k=0,1,2,\ldots,$$

where $C_k=\lambda_{k+1}-\lambda_k=2k+n$. Then we have the following.

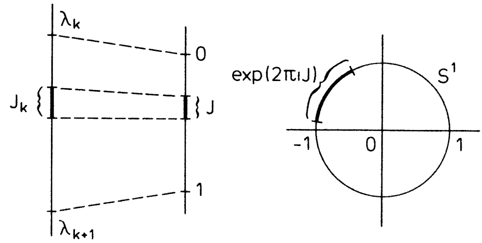


Fig. 1

Theorem. *Suppose* $\mathrm{vol}\{\bar{Q}_{\tilde{d}}^{-1}(e^{2\pi i a})\} = \mathrm{vol}\{\bar{Q}_{\tilde{d}}^{-1}(e^{2\pi i b})\} = 0$. *Then*

$$\#\{j;\ \mu_j \in J_k\} = (2\pi)^{-n}\,\mathrm{vol}\{\bar{Q}_{\tilde{d}}^{-1}(\exp(2\pi i J))\}\,k^{n-1} + o(k^{n-1}),$$

where $\#$ *denotes cardinal* (*see Fig.* 1).

Remark. The assumption for a and b in Theorem is weaker than that $e^{2\pi i a}$ and $e^{2\pi i b}$ are regular values of $\bar{Q}_{\tilde{d}}$. The assumption is satisfied for any a and b if $\bar{Q}_{\tilde{d}}$ is real analytic, and is not constant.

Roughly speaking the above theorem means that

$$\frac{\#\{j;\ \mu_j \in J_k\}}{\#\{j;\ \mu_j \in [\lambda_k, \lambda_{k+1}]\}} \sim \frac{\mathrm{vol}\{\bar{Q}_{\tilde{d}}^{-1}(\exp(2\pi i J))\}}{\mathrm{vol}\, S^* S^n} \qquad (k \to \infty).$$

Example (cf. [5]). Since $H^2(S^2, \mathbb{Z}) = \mathbb{Z}$, the set of equivalence classes of line bundles over S^2 is $\{E_m\}_{m\in\mathbb{Z}}$. On each line bundle E_m there is a unique harmonic connection $\tilde{d}_m$ whose curvature form is $\Omega_m = im\Theta/2$, Θ being the natural volume form of (S^2, g_0). The spectrum of the Laplacian L_m defined by $\tilde{d}_m$ is calculated as

$$\lambda_k^{(m)} = \left(k + \frac{|m|+1}{2}\right)^2 - \frac{m^2+1}{4}, \qquad k = 0, 1, 2, \ldots. \tag{2.2}$$

Notice that $(E_0, \tilde{d}_0)$ is the trivial bundle with the flat connection, and $L_0 = \Delta$. On the other hand, as to the holonomy of $\tilde{d}_m$ we have

$$Q_{\tilde{d}_m}(\gamma) = (-1)^m \tag{2.3}$$

for every closed geodesic γ of (S^2, g_0). From (2.2) we have

$$\lambda_k^{(\pm 2q)} = \lambda_{k+q} - q^2,$$

$$\lambda_k^{(\pm(2q+1))} = \lambda'_{k+q} - \{q(q+1) + \tfrac{1}{2}\},$$

for $k = 0, 1, 2, \ldots$ and $q = 0, 1, 2, \ldots$, where $\lambda'_k = (\lambda_{k+1} + \lambda_k)/2$. These relations correspond to (2.3).

3. Proof of the Theorem

Let L be the Laplacian on $(S^n, g_0; E, \tilde{d})$, and set $\mathrm{Spec}(L) = \{\mu_j\}_{j=0}^{\infty}$. Let E_0 be the trivial line bundle over S^n.

The proof of the theorem is performed in three steps.

3.1. We will first construct an elliptic pseudo-differential operator $A: C^\infty(E_0) \to C^\infty(E)$ of order 0 with certain properties. When E is the trivial bundle, we set $A = \text{Identity}$. (Note that the bundle is always trivial if $n \neq 2$.) Suppose $n=2$, and let $E=E_m$ ($m \in \mathbb{Z}$). Consider the principal S^1-bundle called the Hopf bundle:

$$P: SU(2) = S^3 \xrightarrow{\pi} SU(2)/S^1 = S^2,$$

over S^2. Here if we choose the coordinates (θ, φ, ψ) of $SU(2)$ as

$$h = \begin{pmatrix} \cos\frac{\theta}{2}\, e^{i(\varphi+\psi)/2} & i\sin\frac{\theta}{2}\, e^{i(\varphi-\psi)/2} \\ i\sin\frac{\theta}{2}\, e^{-i(\varphi-\psi)/2} & \cos\frac{\theta}{2}\, e^{-i(\varphi+\psi)/2} \end{pmatrix} \in SU(2),$$

$$(0 \leqq \theta \leqq \pi,\ 0 \leqq \varphi \leqq 2\pi,\ 0 \leqq \psi \leqq 4\pi),$$

then the fibre over $x \in S^2$ is the submanifold:

$$\{(\theta_0, \varphi_0, \psi) = (x, \psi) \in S^3;\ 0 \leqq \psi \leqq 4\pi\}.$$

The line bundle E_m is associated with the principal S^1-bundle P by the representation $\rho_m: S^1 \to \mathbb{C}^*$; $\rho_m(e^{it})(z) = e^{-imt} z$. We can easily see that $E_{-2} \cong T^*S^2$ as complex line bundles. For each $h \in P = S^3$ we define $\tau_m(h): \mathbb{C} \to E_m$ by

$$\tau_m(h)(z) = [(h, z)]_m \in E_m = P \times_{\rho m} \mathbb{C}.$$

Particularly,

$$\tau_{-2}(h)(z) = [(h, z)]_{-2} \stackrel{v}{=} (x, \xi) \in T^*S^2,$$

$$\tau_{-2}(h)(1) = [(h, 1)]_{-2} \stackrel{v}{=} (x, \eta) \in S^*S^2,$$

where $\pi(h) = x$, and $\stackrel{v}{=}$ means the correspondence by the bundle isomorphism $v: E_{-2} \to T^*S^2$. If $(x, \xi) \in S^*S^2$ is given, there exist just two elements $h_1 = (x, \psi_1)$ and $h_2 = (x, \psi_2)$ in $\pi^{-1}(x) (\subset P)$ such that $\tau_{-2}(h_j)(1) \stackrel{v}{=} (x, \xi)$ $(j=1,2)$, and $(x, 2\psi_1) = (x, 2\psi_2)$. We assign $h = (x, 2\psi_1) = (x, 2\psi_2) \in P$ to $(x, \xi) \in S^*S^2$, and denote this map by Φ, i.e. $\Phi(x, \xi) = h$. It is easy to see that $\Phi: S^*S^2 \to P$ is of C^∞ class. Now we define a map

$$a: T^*S^2 \backslash 0 \to \operatorname{Hom}(E_0, E),$$

$\operatorname{Hom}(E_0, E)$ being the bundle of homomorphisms of E_0 into E. For $(x, \xi) \in S^*S^2$ and $(x, z) \in E_0 = S^2 \times \mathbb{C}$, we set

$$a(x, \xi)(z) = \tau_m(\Phi(x, \xi))(z) = [(\Phi(x, \xi), z)]_m \in E_x.$$

We extend a to all of $T^*S^2 \backslash 0$ by requiring it to be homogeneous of degree 0 in ξ. We note that $a(x, \xi)$ is an unitary map of $E_{0,x}$ onto E_x for each $(x, \xi) \in T^*S^2 \backslash 0$. According to Weinstein [6, Theorem 3.1] there is an elliptic pseudo-differential operator A of order 0 such that the symbol, $\sigma(A)$, of A is a, and

$$A^*A = \text{Identity}, \quad \text{or} \quad AA^* = \text{Identity},$$

where A^* is the adjoint of A.

Set $L' = A^*LA$, and L' is a self-adjoint pseudo-differential operator of order 2 which operates on $C^\infty(E_0)$.

Lemma 3.1. (1) $L' - \Delta = Q$ *is a (formally) self-adjoint pseudo-differential operator of order* 1.

(2) *Set* $\mathrm{Spec}(L') = \{\mu'_j\}_{j=0}^\infty$. *Then there exists a constant* C_1 *not depending on* j *such that*

$$|\mu'_j - \mu_{j-l}| \leqq C_1, \tag{3.1}$$

where l *is the index of* A *(a Fredholm operator).*

Proof. (1) Since $\sigma(A)$ is an unitary map, we have $\sigma(A^*) = \sigma(A)^{-1}$ and

$$\sigma(L')(x, \xi) = \sigma(L)(x, \xi) = |\xi|^2 = \sigma(\Delta)(x, \xi),$$

which shows that Q is of order 1.

(2) We may think of A as a Fredholm operator of $L^2(E_0)$ into $L^2(E)$. Suppose A satisfies $A^*A = \text{Identity}$. Then A is injective. Let $\tilde{H} = L^2(E_0) \oplus \mathrm{Ker}(A^*)$ ($\mathrm{Ker}(A^*) \subset C^\infty(E)$), and define $\tilde{A}: \tilde{H} \to L^2(E) = \mathrm{Im}(A) \oplus \mathrm{Ker}(A^*)$ by $\tilde{A}(f \oplus u) = Af + u$ for $f \in L^2(E_0)$ and $u \in \mathrm{Ker}(A^*)$. We can see that $\tilde{A}^* = A^* \oplus P_1$, P_1 being the projection of $L^2(E)$ onto $\mathrm{Ker}(A^*)$, and that $\tilde{A}$ is an unitary operator. Next, define a self-adjoint operator $\tilde{L}'$ with domain $H_1(E_0) \oplus \mathrm{Ker}(A^*)(\subset \tilde{H})$ by $\tilde{L}'(f \oplus u) = L'f \oplus 0 = A^*LA \oplus 0$, where $H_1(E_0)$ is the set of sections s such that s and its first derivatives belong $L^2(E_0)$. Then we can show that $\tilde{L}' - \tilde{A}^*L\tilde{A}$ is a compact operator. In fact, we have

$$\tilde{A}^*L\tilde{A}(f \oplus u) = (A^*LAf + A^*Lu) \oplus (P_1LAf + P_1Lu),$$

hence,

$$\tilde{L}' - \tilde{A}^*L\tilde{A} = (-A^*LP_2) \oplus (-P_1L\tilde{A}),$$

where P_2 is the projection of $\tilde{H}$ onto $\mathrm{Ker}(A^*)$. Since P_1 and P_2 are projections onto a finite dimensional subspace, we can see that $\tilde{L}' - \tilde{A}^*L\tilde{A}$ is a compact operator. If we set $\mathrm{Spec}(\tilde{L}') = \{\tilde{\mu}'_j\}_{j=0}^\infty$, then by the definition of $\tilde{L}'$ we have

$$\tilde{\mu}'_{j+q} = \mu'_j,$$

where $q = -l = \dim \mathrm{Ker}(A^*) > 0$. On the other hand, since $\tilde{L}' = \tilde{A}^*L\tilde{A} + S$, S being a compact, and accordingly bounded operator, we can see that $|\tilde{\mu}'_j - \mu_j|$ is bounded by virtue of the min-max principle. Hence we get (3.1). We can similarly prove for the case of $AA^* = \text{Identity}$ by exchanging A for A^*. □

3.2. *Averaging the Perturbed Term.* For the perturbed term Q of L' from Δ, we consider the averaged operator Q_{av} which is defined by

$$Q_{\mathrm{av}} = \frac{1}{2\pi} \int_0^{2\pi} \exp(-itP_0)\, Q \exp(itP_0)\, dt,$$

where

$$P_0=\left(\Delta+\frac{(n-1)^2}{4}\right)^{1/2}.$$

Q_{av} is a (formally) self-adjoint pseudo-differential operator of order 1, and the following lemma is obtained.

Lemma 3.2. (1) *The principal symbol of* Q_{av} *satisfies*

$$\frac{1}{|\xi|}\,\sigma(Q_{\mathrm{av}})(x,\xi)\equiv\frac{i}{\pi}\log Q_{\tilde{d}}(\gamma(x,\xi)) \qquad (\mathrm{mod}\, 2\mathbf{Z}),$$

where $\gamma(x,\xi)$ *denotes the closed geodesic* $\gamma(t)$ $(0\leqq t\leqq 2\pi)$ *with* $\gamma(0)=x$ *and* $\dot{\gamma}(0)^*=\xi/|\xi|$.

(2) $[P_0, Q_{\mathrm{av}}]=0$.

(3) *Let* $\mathrm{Spec}(\Delta+Q_{\mathrm{av}})=\{\mu_j''\}_{j=0}^{\infty}$. *Then there exists a constant* C_2 *not depending on j such that*

$$|\mu_j''-\mu_j'|\leqq C_2. \tag{3.2}$$

Proof. (1) By Egorov's theorem

$$\sigma(Q_{\mathrm{av}})(x,\xi)=\frac{1}{2\pi}\int_0^{2\pi}\phi_t^*(x,\xi)\,\sigma(Q)\,dt,$$

where $\phi_t(x,\xi)$ denotes the integral curve in $T^*S^n\backslash 0$ of the Hamiltonian vector field defined by $\sigma(P_0)$ with the initial value (x,ξ), which is just the geodesic flow of (S^n, g_0). Since the subprincipal symbol of Δ vanishes, we have

$$\sigma(Q)=\sigma_{\mathrm{sub}}(A^*LA)=\sigma(A^*)\,\sigma_{\mathrm{sub}}(L)\,\sigma(A)+\frac{1}{i}\,\sigma(A^*)\,\mathcal{L}_H\,\sigma(A)$$

for a fixed local frame of $E=E_m$, where H is the Hamiltonian vector field associated with $\sigma(L)=\sigma(\Delta)$, and $\mathcal{L}_H$ is the Lie derivative with respect to H. Let U be a tubular neighborhood of $\gamma(x,\xi)$ in S^2, and let $s\colon U\to T^*S^2\backslash 0$ be a cross-section such that $s(\gamma(t))=\phi_t(x,\xi)$. Define an unitary local frame on U by $e(x)=\tau_m(\Phi(s(x)/|s(x)|))(1)$, $x\in U$. Then we have $\sigma(A)=\sigma(A^*)=1$ on $s(U)$ for this frame, and therefore

$$\sigma_{\mathrm{sub}}(A^*LA)=\sigma_{\mathrm{sub}}(L)=2\sum_{j=1}^{n}a^j\xi_j,$$

on the orbit $\{\phi_t(x,\xi);\,0\leqq t\leqq 2\pi\}$, where $\omega=i\alpha=i\sum a_j\,dx^j$ is the connection 1-form of $\tilde{d}$ with respect to this frame. Hence, for every n we have

$$\sigma(Q_{\mathrm{av}})(x,\xi)=\frac{|\xi|}{\pi}\int_{\gamma(x,\xi)}\alpha\equiv\frac{|\xi|}{i\pi}\int_{\Sigma_\gamma}\Omega \qquad (\mathrm{mod}\, 2\mathbf{Z}\,|\xi|).$$

Here, note that the parallel lift $\tilde{\gamma}(t)=c(t)\,e(\gamma(t))$ of $\gamma(t)$ satisfies $c(t)=c(0)\exp\left(-\int_{\gamma(t)}\omega\right)$.

(2) Same as Weinstein [8, Lemma 1.1].

(3) This is proved on the same lines as the argument in [8]. Set

$$R=\frac{i}{2\pi}\int_0^{2\pi}\left(\int_0^t \exp(-isP_0)\,Q\exp(isP_0)\,ds\right)dt,$$

and moreover set

$$S=\tfrac{1}{4}(P_0^{-1}R+RP_0^{-1}).$$

Then, R (resp. S) is a skew-symmetric pseudo-differential operator of order 1 (resp. 0). By straightforward calculations we can see that the operator

$$[\Delta, S]-(Q-Q_{\mathrm{av}})$$

is of order 0. Consider the operator $\exp S$, which is an unitary pseudo-differential operator of order 0. Then, modulo operator of order 0 we have

$$\begin{aligned}(\exp S)(\Delta+Q)(\exp(-S))&=\sum_{j=0}^{\infty}\frac{(\operatorname{Ad}S)^j}{j!}(\Delta+Q)\\ &=\Delta+Q+[S,\Delta+Q]+\frac{1}{2!}[S,[S,\Delta+Q]]+\dots\\ &\equiv\Delta+Q-(Q-Q_{\mathrm{av}})=\Delta+Q_{\mathrm{av}}.\end{aligned}$$

Thus $(\exp S)\,L'(\exp(-S))-(\Delta+Q_{\mathrm{av}})$ is a bounded operator. Hence we get (3.2). □

3.3. *Theorem of Colin de Verdière.* We first recall the theorem by Colin de Verdière. Let M be an n-dimensional compact manifold. Let $P_1,\dots,P_s$ be (formally) self-adjoint pseudo-differential operators of order 1 such that

(1) $\sum_{j=1}^{s} P_j^2$ is an elliptic operator, and

(2) $[P_j, P_k]=0$ for any j and k.

If $\psi\in C^\infty(M)$ satisfies

$$P_j\psi=\lambda^{(j)}\psi,\quad j=1,\dots,s,$$

we call ψ an eigenfunction of $(P_1,\dots,P_s)$ and $\lambda=(\lambda^{(1)},\dots,\lambda^{(s)})(\in\mathbb{R}^s)$ the associated eigenvalue. Let $\Lambda=\{\lambda_\alpha=(\lambda_\alpha^{(1)},\dots,\lambda_\alpha^{(s)})\}_{\alpha=0}^{\infty}$ be the set of all eigenvalues of $(P_1,\dots,P_s)$. Let p_j $(1\leqq j\leqq s)$ be the principal symbol of P_j, and set

$$p=(p_1,\dots,p_s)\colon T^*M\backslash 0\to\mathbb{R}^s\backslash 0.$$

Furthermore, let V be the set of all points in $T^*M\backslash 0$ where $dp_1\wedge\dots\wedge dp_s=0$, and set $W=p(V)(\subset\mathbb{R}^s\backslash 0)$.

Theorem (Y. Colin de Verdière [2]). *Assume that* $\operatorname{Spec}(P_1)=\{\lambda_\alpha^{(1)}\}\subset\mathbb{Z}+\{\mu_0\}$, *and that every integral curve of the Hamiltonian vector field defined by* p_1 *is periodic with period* 2π. *Let* C *be a conic subset of* $\mathbb{R}^s\backslash 0$ *such that* $\partial C\cap W=\varnothing$ *and* $C\cap\{x_1=1\}$ *is compact. Then*

$$\#\{\lambda\in C\cap\Lambda;\ \lambda^{(1)}=k+\mu_0\}=(2\pi)^{-n}\,\widetilde{\operatorname{vol}}\{p^{-1}(C\cap[\{k+\mu_0\}\times\mathbb{R}^{s-1}])\}+O(k^{n-2}),$$

where $\widetilde{\mathrm{vol}}$ *denotes the volume induced on the hypersurface:* $p_1=\text{const.}$ *from that on* T^*M.

Now noticing Lemma 3.2, (2), we will apply this theorem to the operators (P_0, Q_{av}). Let $\Lambda=\{(\bar{\lambda}_{k,j}, \kappa_{k,j})\}$ be the set of eigenvalues of (P_0, Q_{av}), where

$$\bar{\lambda}_{k,1}=\ldots=\bar{\lambda}_{k,N_k}=\bar{\lambda}_k=k+\frac{n-1}{2}$$

(N_k being the multiplicity of $\bar{\lambda}_k$). Then, for $\operatorname{Spec}(\Delta+Q_{av})=\{\mu''_{k,j}\}$, we have

$$\mu''_{k,j}=\lambda_{k,j}+\kappa_{k,j} \qquad (\text{where } \lambda_{k,j}=\lambda_k),$$

$\kappa_{k,j}$ being the difference between $\mu''_{k,j}$ and λ_k.

Lemma 3.3. *Let*

$$M=\max_{(x,\xi)\in S^*S^n}\sigma(Q_{av})(x,\xi)(\geqq 0).$$

Then we have

$$|\kappa_{k,j}|\leqq Mk+C_3, \tag{3.3}$$

C_3 *being a constant.*

Proof. The operator $P_0^{-1}Q_{av}$ is a self-adjoint pseudo-differential operator of order 0, whose principal symbol is equal to $\sigma(Q_{av})(x,\xi)/|\xi|$. Note that $\sigma(Q_{av})(x,-\xi)=-\sigma(Q_{av})(x,\xi)$. Hence we can write

$$MI-P_0^{-1}Q_{av}=Q_0+Q_{-1}, \qquad P_0^{-1}Q_{av}-(-MI)=Q'_0+Q'_{-1},$$

where Q_0 and Q'_0 are non-negative self-adjoint pseudo-differential operators of order 0, and Q_{-1} and Q'_{-1} are pseudo-differential operators of order -1. Let $\psi_{k,j}$ be a eigenfunction of (P_0, Q_{av}) associated with $(\bar{\lambda}_k, \kappa_{k,j})$. Then we have

$$\begin{aligned} M-\kappa_{k,j}/\bar{\lambda}_k &=(Q_0\psi_{k,j}, \psi_{k,j})+(Q_{-1}\psi_{k,j}, \psi_{k,j}) \\ &\geqq (Q_{-1}\psi_{k,j}, \psi_{k,j}) \\ &=(1/\bar{\lambda}_k)(Q_{-1}P_0\psi_{k,j}, \psi_{k,j}). \end{aligned}$$

Since $Q_{-1}P_0$ is of order 0, the last term in parentheses is bounded. Hence

$$\kappa_{k,j}\leqq\bar{\lambda}_k M+C'=Mk+C_3.$$

A similar argument shows another inequality. □

Now we proceed to prove the theorem. Note that for our case the set W of critical values of p satisfies

$$W\subset\{(x_1,x_2);\ x_2=2\nu x_1,\ e^{2\pi i\nu} \text{ is a critical value of } \bar{Q}_{\bar{a}}\}.$$

For $0\leqq a<b\leqq 1$ and sufficiently small $\varepsilon>0$ we consider the following conic subset of $\mathbb{R}^2=\{(x_1,x_2)\}$:

$$C_N^{\pm\varepsilon}=\{(x_1,x_2);\ 2(a\mp\varepsilon+N)x_1\leqq x_2\leqq 2(b\pm\varepsilon+N)x_1,\ x_1>0\},$$

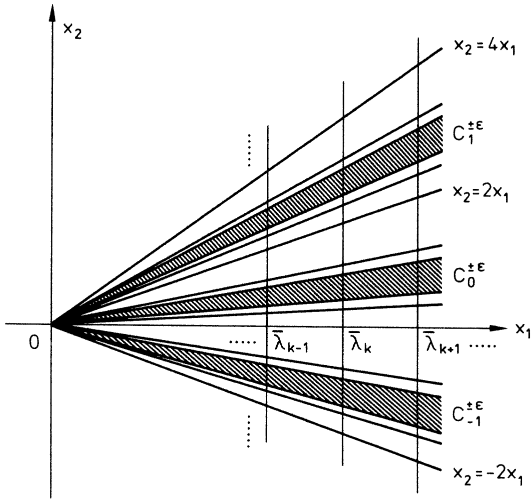


Fig. 2

for $N=0, \pm 1, \pm 2, \ldots$ (see Fig. 2). By virtue of (3.1) and (3.2) we have

$$\#\left\{\bigcup_{N=-q}^{q} (\Lambda \cap C_N^{-\varepsilon} \cap \{x_1 = \bar{\lambda}_{k-N}\})\right\} \leqq \#\{\mu_j \in J_k\} \leqq \#\left\{\bigcup_{N=-q}^{q} (\Lambda \cap C_N^{+\varepsilon} \cap \{x_1 = \bar{\lambda}_{k-N}\})\right\},$$

for sufficiently large k, where q is an integer satisfying $2q > M$ for the constant M in Lemma 3.3. If $e^{2\pi i(a\pm\varepsilon)}$ and $e^{2\pi i(b\pm\varepsilon)}$ are regular values, then by the theorem of Colin de Verdière we have

$$\begin{aligned}
&\#\left\{\bigcup_{N=-q}^{q} (\Lambda \cap C_N^{\pm\varepsilon} \cap \{x_1 = \bar{\lambda}_{k-N}\})\right\} \\
&\quad = \sum_N (2\pi)^{-n} \widetilde{\mathrm{vol}}\,\{p^{-1}(\Lambda \cap C_N^{\pm\varepsilon} \cap \{x_1 = \bar{\lambda}_{k-N}\})\} + O(k^{n-2}) \\
&\quad = (2\pi)^{-n} \sum_N \mathrm{vol}\,\{(x, \xi) \in S^* S^n;\ 2(a \mp \varepsilon + N) \leqq \sigma(Q_{\mathrm{av}})(x, \xi) \leqq 2(b \pm \varepsilon + N)\} \\
&\qquad \times \left(k - N + \frac{n-1}{2}\right)^{n-1} + O(k^{n-2}) \\
&\quad = (2\pi)^{-n} \mathrm{vol}\,\{\bar{Q}_{\bar{d}}^{-1}(\exp(2\pi i J^{\pm\varepsilon}))\}\, k^{n-1} + O(k^{n-2}),
\end{aligned}$$

where $J^{\pm\varepsilon} = [a \mp \varepsilon, b \pm \varepsilon]$. By Sard's theorem we can choose a sequence $\{\varepsilon_n (>0); n = 1, 2, \ldots\}$ such that $\varepsilon_n \downarrow 0$ as $n \to \infty$, and $e^{2\pi i(a\pm\varepsilon_n)}$ and $e^{2\pi i(b\pm\varepsilon_n)}$ are regular values. Under the assumption: $\mathrm{vol}\,\{\bar{Q}_{\bar{d}}^{-1}(e^{2\pi i a})\} = \mathrm{vol}\,\{\bar{Q}_{\bar{d}}^{-1}(e^{2\pi i b})\} = 0$, we have

$$\mathrm{vol}\,\{\bar{Q}_{\bar{d}}^{-1}(\exp(2\pi i J^{+\varepsilon_n}))\} - \mathrm{vol}\,\{\bar{Q}_{\bar{d}}^{-1}(\exp(2\pi i J^{-\varepsilon_n}))\} \to 0$$

as $n \to \infty$. Thus the proof of Theorem is completed.

Acknowledgment. The author would like to express his gratitude to the referee for pointing out that the assumption in the theorem may be improved to the present form.

References

1. Colin de Verdière, Y.: Sur le spectre des opérateurs elliptiques a bicaractéristiques toutes periodiques. Comment. Math. Helv. **54**, 508–522 (1979)
2. Colin de Verdière, Y.: Spectre conjoint d'opérateurs pseudo-differentiels qui commutent I. Le cas non integrable. Duke Math. J. **46**, 169–182 (1979)
3. Guillemin, V.: Some spectral results for the Laplace operator with potential on the n-sphere. Advances in Math. **27**, 273–286 (1978)
4. Guillemin, V.: Some spectral results on rank one symmetric spaces. Advances in Math. **28**, 129–137 (1978)
5. Kuwabara, R.: On spectra of the Laplacian on vector bundles. J. Math. Tokushima Univ. **16**, 1–23 (1982)
6. Weinstein, A.: Fourier integral operators, quantization, and the spectra of Riemannian manifolds. Colloques Internat. CNRS **237**, 289–298 (1976)
7. Weinstein, A.: Asymptotics of eigenvalue clusters for the Laplacian plus a potential. Duke Math. J. **44**, 883–892 (1977)
8. Weinstein, A.: Eigenvalues of the Laplacian plus a potential. Proceeding of the International Congress of Mathematicians Helsinki, pp. 803–805. Helsinki: Academia Scientiarum Fennica 1978
9. Wells, R.O Jr.: Differential Analysis on Complex Manifolds. 2nd ed. Berlin-Heidelberg-New York: Springer 1980

Received September 19, 1983; received in final form April 16, 1984

Math. Z. 187, 491–503 (1984)

Mathematische Zeitschrift

Generic Fixed Point Free Action of Arbitrary Finite Groups

Alexandre Turull

Department of Mathematics, University of Miami, Coral Gables, Florida 33124, USA

Introduction

Let A be a finite group of operators of the finite group G and assume $(|A|, |G|) = 1$ and $C_G(A) = 1$. A famous conjecture of Frobenius stated that if A is a cyclic group of prime order then G is nilpotent. Its proof by Thompson [15] marked the beginning of the recently completed "classification" of finite simple groups. In turn, from this classification it is easy to deduce, as remarked by Gorenstein [5, Theorem 1.48], that for any A, G will always be solvable. There remains the question, how far from nilpotent can it be? More precisely, given A, can we bound the Fitting height of G, and if so what is the smallest such bound? Gross [6] showed that if A is solvable then the bound has to be at least the number of primes that divide $|A|$ counting multiplicities.

In the last 25 years, since Thompson's proof, a great deal of effort has been put into trying to show that the number of primes that divide the order of $|A|$ counting multiplicities is in fact a bound for the Fitting height of G for any solvable A. Details on that history before 1973 can be found in [3]. In 1973, Berger [3] proved this for A nilpotent and $\mathbb{Z}_p \sim \mathbb{Z}_p$ free for all primes p. His result encompassed results by many earlier authors. Bailey [1] extended slightly Berger's results and Feldman [4] studied the case where A is a Frobenius group with cyclic kernel and complement of prime order. In [17] the case where A is supersolvable is considered and a theorem that includes all the above mentioned results is proved. In [16] the bound is proved to hold for A dihedral semihedral or generalized quaternion.

When A is not included in any of the above mentioned theorems the information is more sketchy. If A is solvable then it is shown in [16] that the Fitting height of G is at most twice the number of primes that divide the order of A counting multiplicities. This improves on some results of Kurzweil [13] which in turn improved on some results of Thompson [14].

When A is not solvable nothing was known until recently. In [12], Kurzweil studied the case where A is minimal simple and obtained some large bounds for the Fitting height of G. In particular, if A is isomorphic to A_5, the alternating group on five letters, he showed that the Fitting height of G is at most 17, and he further showed that the Fitting height of G can be at least 3 in this case. In [11], he reduced the bound on the Fitting height of G when $A \cong A_5$ from 17 to 5. In [10], he considered a more general situation and proved for many non-solvable groups A that the Fitting height of G was bounded by a large number which depends only on A.

In this paper we look at an arbitrary finite group A. Suppose it acts on a solvable group G, with $(|A|,|G|)=1$ and $C_G(A)=1$. We define this action to be *generic* if for the action of all proper sub-groups of A on sections of G there are enough regular orbits and modules (Definition 1.4). It turns out that the number of primes that divide the order of A counting multiplicities is not the correct function to look at when A is not solvable. We define, for A any finite group, $l(A)$ to be the largest integer n such that there exists a sequence of subgroups of A, $1=S_0 \subset S_1 \subset \ldots \subset S_n = A$. The main results of the paper imply the following.

Theorem A. *Suppose A is any finite group and it acts on the (solvable) group G with $(|A|,|G|)=1$ and $C_G(A)=1$. Assume further that the action of A on G is generic. Then the Fitting height of G is at most $l(A)$.*

Theorem B. *Suppose A is any finite group and p and q are any two distinct primes which do not divide the order of A. Then there is a finite (solvable) $\{p,q\}$-group G and an action of A on G such that $C_G(A)=1$ and the Fitting height of G is $l(A)$.*

We also show that "most actions" of a finite group A are generic. On the one hand we show that there exists for each A a finite set τ of primes such that if $\pi(G) \cap \tau = \varnothing$ then the action of A on G is generic. Combining this fact with Theorem A we get:

Theorem C. *Let A be any finite group. Then there exists a finite set τ of primes such that if G is any finite (solvable) group, $(|A|,|G|)=1$ and $\tau \cap \pi(G) = \varnothing$ and A acts on G with $C_G(A)=1$, then the Fitting height of G is bounded above by $l(A)$.*

On the other hand we give (Theorem 1.9 below) some sufficient conditions on A that insure, given $\pi(G)$, that the action of A on G is generic. Combining this with Theorem A we obtain a generalization (Theorem 2.6) of the result of [17] mentioned above. We give to conclude this introduction some special cases of this.

Theorem D. *Suppose A is isomorphic to $SL(2,2^p)$ (where p is any prime) or to $PSL(2,3^3)$, $PSL(2,3^5)$ or $PSL(2,3^7)$. Assume further that A acts on the finite (solvable) group G with $(|A|,|G|)=1$ and $C_G(A)=1$. Then the Fitting height of G is at most $l(A)$.*

In particular if $A \simeq A_5$, the alternating group on five letters, then the Fitting height of G is at most 4 and this is the best possible bound. This improves on some results of Kurzweil [11, 12] mentioned above.

1. Generic Action

We are interested in the following situation.

Hypothesis 1.1. *Let A be any finite group and assume it acts on the finite (solvable) group G with $(|A|, |G|)=1$ and $C_G(A)=1$.*

We define:

Definition 1.2. Let B act on G. We say that B satisfies the *regular orbit property* in its action on G if for every B-invariant irreducible section S of G (i.e. S is any section of G which is B-invariant and elementary Abelian and where B acts irreducibly) there is some $v \in S$ such that $C_B(v)=C_B(S)$.

Definition 1.3. Let B act on G. We say that B satisfies the *regular module property* in its action on G if the following holds. Suppose S a B-invariant section of G such that S contains an extraspecial subgroup P with $P \lhd BS$, $Z(P) \subseteq Z(BS)$, S/P elementary Abelian and $[S, P]=P$. Assume further that M is a $\mathbb{C}BS$-irreducible module such that for every $x \in BS$ $\det_M(x)=1$, and $M|_P$ is irreducible and faithful. Then we require that $M|_B$ contains submodules isomorphic to every irreducible $\mathbb{C}[B/C_B(P)]$-module.

Note that we do not require $M|_B$ to contain a regular $B/C_B(P)$ submodule if $B/C_B(P)$ is not Abelian.

Definition 1.4. We say that the action of A on G is *generic* if it satisfies both the regular orbit property and the regular module property in the action of every proper subgroup B of A on G.

The next proposition shows that in some sense "most" actions of any given finite group A are generic.

Proposition 1.5. *Let A be any finite group. Then there is a finite set of primes τ such that if G satisfies hypothesis 1.1 and $\pi(G) \cap \tau = \varnothing$ then the action of A on G is generic.*

Proof. For expediency, we use in this proof a rather large set τ which has the advantage of allowing a straight forward proof. Let τ be the set of primes which are smaller than or equal to $|A|$. Assume $\pi(G) \cap \tau = \varnothing$. We show that any action of A on G is generic.

Take B any proper subgroup of A and S a B-invariant irreducible section of G. Then we may view S as a $\mathbb{F}_p B$-module for p a prime, $p>|A|$. Set $C=B/C_B(S)$ and let n be the dimension of S as a $\mathbb{F}_p$-vector space. Then

$$|M \setminus \bigcup_{x \in C^\#} C_M(x)| \geqq (p^n-1)-(|C|-1)(p^{n-1}-1)$$

since for $x \in C^\#$, $|C_M(x)| \leqq p^{n-1}$. But we know that $|C|<p$, so we get

$$|M \setminus \bigcup_{x \in C^\#} C_M(x)| \geqq (p^n-1)-p(p^{n-1}-1)=p-1>0.$$

Hence we may take $v \in M$ such that for all $x \in C^{\#}$ we have $v \notin C_M(x)$. So $C_C(v) = 1$ and $C_B(v) = C_B(S)$. Hence the action of A on G satisfies the regular orbit property.

Now suppose that B, S, P and M satisfy the hypothesis of Definition 1.3. Set $C = B/C_B(P)$ and assume that $|P| = p^{2n+1}$ and that ψ is the character of M. It follows from results of Glauberman [9, Chap. 13], or more explicitly Isaacs [8] or Berger [2] that we have the following. For every $x \in B$, $\psi(x)$ is an integer and $\psi(x)^2$ is the number of fixed points in P/P' under x, and $\ker_B(M) = C_B(P)$. Hence we may consider $\psi|_B$ as a C-character and we have

$$|\psi(x)| \leqq p^{n-1} \quad \text{for } x \in C^{\#}. \tag{1.5.1}$$

Let χ be the character of an irreducible $\mathbb{C}C$-module. Then we have

$$|\chi(x)| \leqq \chi(1) \quad \text{for } x \in C. \tag{1.5.2}$$

We are now in a position to estimate the value of the inner product $(\chi, \psi|_C)_C$. We have the following using (1.5.1) and (1.5.2) and then the fact that $p > |A| > |C|$.

$$\begin{aligned}(\chi, \psi|_C)_C &= \frac{1}{|C|} \sum_{x \in C} \chi(x)\,\psi(x) \\ &\geqq \frac{1}{|C|}\left(p^n \chi(1) - (|C|-1)\,p^{n-1}\chi(1)\right) \\ &= \frac{\chi(1)\,p^{n-1}}{|C|}(p - |C| + 1) > 0.\end{aligned} \tag{1.5.3}$$

But now (1.5.3) implies that $M|_B$ has a submodule with character χ. This shows that the action of B on G satisfies the regular module property. Since we already know that the action of B on G satisfies the regular orbit property, we get that the action of A on G is generic. This completes the proof of Proposition 1.5.

Next we give a set of sufficient conditions for the action of A on G to be generic. We need to recall first some definitions from [17].

Notation 1.6. For p^n a power of à prime p and $\varepsilon = -1$ or a power of 2 such that $\varepsilon | n$ define:

$F_\varepsilon(p^n) = \mathbb{F}_{p^n}^*$ (the multiplicative group of the field of p^n elements) if $\varepsilon > 0$;
$\quad$ = subgroup of order $p^n + 1$ of $\mathbb{F}_{p^{2n}}^*$ if $\varepsilon = -1$.

$\mathrm{Gal}(1, p^n) = \mathrm{Gal}(\mathbb{F}_{p^n} : \mathbb{F}_p)$ (the Galois group of $\mathbb{F}_{p^n}$).

$\mathrm{Gal}(-1, p^n) = O_{2'}\,(\mathrm{Gal}(\mathbb{F}_{p^{2n}} : \mathbb{F}_p))$.

If $\varepsilon \neq \pm 1$ $\mathrm{Gal}(\varepsilon, p^n) = \langle \varphi \rangle \times \mathrm{Gal}(-1, p^{(n/2)})$ where φ is the automorphism of $F_\varepsilon(p^n)$: $x \to x^{-p^{(n/\varepsilon)}}$.

$G(\varepsilon, p^n) = \mathrm{Gal}(\varepsilon, p^n) \ltimes F_\varepsilon(p^n)$ (the semi-direct product).

$o(x)$ = order of x.

If $\sigma \in \mathrm{Gal}(\varepsilon, p^n)$ and $y \in F_\varepsilon(p^n)$ we denote N_σ the norm map i.e.

$$N_\sigma(y) = \prod_{\tau \in \langle \sigma \rangle} \tau(y) = y \cdot \sigma(y) \dots \sigma^{(o(\sigma)-1)}(y).$$

Suppose that q is a prime and divides $|\mathrm{Gal}(\varepsilon, p^n)|$. Define

$$GN(\varepsilon, p^n, q) = A \ltimes N \subseteq G(\varepsilon, p^n)$$

where A is the subgroup of $\mathrm{Gal}(\varepsilon, p^n)$ of order q, and

$$N = \{x \in F_\varepsilon(p^n) : \prod_{\sigma \in A} \sigma(x) = 1\}$$

the set of all elements of $F_\varepsilon(p^n)$ of norm 1 under a non-trivial element of A.

Some elementary properties of $GN(\varepsilon, p^n, q)$ are given in the following Proposition [17, Proposition 1.2].

Proposition 1.7. *For any ε, p, n, q such that $GN(\varepsilon, p^n, q)$ is defined take $N = GN(\varepsilon, p^n, q) \cap F_\varepsilon(p^n)$. We have the following:*

1) *N is a cyclic normal subgroup of $GN(\varepsilon, p^n, q)$ of index q and every element of $GN(\varepsilon, p^n, q)$ not in N has order q.*

2) *If q is odd or $s > 1$, $GN(2^s, p^n, q) = GN(1, p^n, q)$.*

3) *If $\varepsilon = \pm 1$, $|N| = (p^n - \varepsilon)/(p^{(n/q)} - \varepsilon)$.*

4) *If $\varepsilon = 2 = q$, $|N| = p^{(n/2)} - 1$.*

5) *For any prime r dividing $|N|$, we have $r \geqq q$ and either $r = q$ or $r \equiv 1\,(q)$.*

6) *If $GN(\varepsilon, p^n, q)$ is nilpotent one of the following is satisfied:*

a) *$q = 3$, $p = 2$, $\varepsilon = -1$, $n = 3$, $GN(\varepsilon, p^n, q)$ is elementary Abelian of order 9 and $F_\varepsilon(p^n)$ is a 3-group.*

b) *$q = 2$, $p = 3$, $\varepsilon = 2$, $n = 2$, $GN(\varepsilon, p^n, q)$ is elementary Abelian of order 4 and $F_\varepsilon(p^n)$ is a 2-group.*

c) *$q = 2$, p is a Mersenne prime, $\varepsilon = 1$, $n = 2$ and $GN(\varepsilon, p^n, q)$ is non-Abelian dihedral of order $2(p+1)$.*

d) *$q = 2$, p is a Fermat prime greater than 3, $\varepsilon = 2$, $n = 2$ or $p = 3$, $\varepsilon = 2$, and $n = 4$. In any case $GN(\varepsilon, p^n, q)$ is non-Abelian dihedral of order $2(p^{(n/2)} - 1)$.*

7) *If $GN(\varepsilon, p^n, q)$ is Abelian, then $\varepsilon \neq 1$ and $F_\varepsilon(p^n)$ is a q-group. Any nilpotent non-Abelian $GN(\varepsilon, p^n, q)$ contains the dihedral group of order 8.*

Definition 1.8. Let A be a finite group and π a set of primes. We say that A is *π-regular* if:

1) $\pi(A) \cap \pi = \varnothing$;

2) For any $p \in \pi$ and any elementary Abelian p-group H on which A acts and any section S of AH, if all Abelian normal sungroups of S are cyclic, S has a self-centralizing cyclic normal subgroup;

3) For any section S of A and any chief-factor X of S, $S/C(X)$ has a regular orbit on X;

4) If $\{3, 5\} \subseteq \pi$, any chief 2-factor of A is cyclic and if further $8 \,|\, |A|$, either A is supersolvable or it has a normal Sylow 2-subgroup;

5) No section of A is isomorphic to $\mathbb{Z}_r \sim \mathbb{Z}_s$ (any $r, s > 1$) or to $GN(\varepsilon, p^n, q)$ where $p \in \pi$, $n \geqq 1$ is an integer, $q \| \mathrm{Gal}(\varepsilon, p^n)|$ is a prime and if $\varepsilon \neq 1$ $\pi(F_\varepsilon(p^n)) \cap \pi \neq \varnothing$.

Note. Conditions 2), 3) and 4) are always satisfied if A is supersolvable. If A is nilpotent, satisfies 1) and is $\mathbb{Z}_p \sim \mathbb{Z}_p$-free for all p, in view of Prop. 1.7, 7), A is π-regular.

We can now state some sufficient conditions for an action to be generic.

Theorem 1.9. *Assume Hypothesis 1.1. Set $\pi = \pi(G)$ and assume that every proper subgroup of A is π-regular. Then the action of A on G is generic.*

Proof. By Definition 1.4, we need to check that for any proper subgroup B of A, the action of B on G satisfies the regular orbit property and the regular module property. Assume that the hypothesis of Definition 1.2 are satisfied. Then B is π-regular. Hence by Theorem 2.2 of [17], $B/C_B(S)$ has a regular orbit on S. So the action of B on G satisfies the regular orbit property.

Now assume B, S, P and M satisfy the hypothesis of Definition 1.3. Let r be the prime such that S/P is an r-group and let R_1 be a B-invariant Sylow r-subgroup of S. Then $BS = BR_1P$. It is well-known that $\ker(M) = C_{BR_1}(P)$. We set $C = B/C_B(P)$ and $R = R_1/C_{R_1}(P)$ and get $BS/\ker(M) = CRP$. Now it is clear that C is π-regular. Since $\{p, r\} \subseteq \pi$, it is trivial to check that we may apply Theorem 3.6 of [17] and we get that $M|_C$ contains a regular direct summand. Hence A satisfies the regular module property, by Definition 1.3. Therefore the action of A on G is generic as desired. This completes the proof of the theorem.

From the note after Definition 1.8 we see that many nilpotent and supersolvable groups will satisfy the hypothesis of Theorem 1.9. Of course there are many non-supersolvable groups A which satisfy the hypothesis of Theorem 1.9. We now give some sufficient conditions for Theorem 1.9 to be applicable to $PSL(2, q)$.

Proposition 1.10. *Suppose that A is isomorphic to $PSL(2, q)$, q a power of a prime and π is a set of primes such that $\pi \cap \pi(A) = \varnothing$. Assume further the following.*

1) *A is minimal simple;*

2) *A Sylow 2-subgroup of A is Abelian;*

3) *Assume p^n is a power of a prime $p \in \pi$, $\varepsilon_1 = \pm 1$, $\varepsilon_2 = \pm 1$ and if $\varepsilon_1 = -1$ also $\pi(p^{2n} - 1) \cap \pi \neq \varnothing$, then $p^n + \varepsilon_1 \nmid \dfrac{q + \varepsilon_2}{(q-1, 2)}$.*

Then every proper subgroup of A is π-regular.

Proof. Since A is minimal simple with Abelian Sylow 2-subgroups it follows from Dickson's Theorem [7, Hauptsatz 8.27] that every proper subgroups is either supersolvable or a Frobenius group with Abelian kernel and cyclic complement. Hence it is easy to see that every proper subgroup of A satisfies the conditions 2) and 3) of Definition 1.8. Since $3 \| A|$, we have $3 \notin \pi$ and hence 4) is also satisfied.

Since, for any r, $s\in\mathbb{Z}$ r, $s\geqq 2$, $\mathbb{Z}_r\sim\mathbb{Z}_s$ is not Abelian and is not a Frobenius group, it is easy to see that $\mathbb{Z}_r\sim\mathbb{Z}_s$ is not involved in A, using condition 2) and Dickson's Theorem. So we assume that $GN(\varepsilon,p^m,r)$ is involved in A for $p\in\pi$, $m\geqq 1$ an integer, $r\,|\,|\mathrm{Gal}(\varepsilon,p^m)|$ a prime and if $\varepsilon\neq 1$ then $\pi(F_\varepsilon(p^m))\cap\pi\neq\varnothing$. Dickson's Theorem now tells us that $r=2$ and $\frac{1}{2}|GN(\varepsilon,p^m,2)|$ divides either $\frac{q-1}{k}$ or $\frac{q+1}{k}$ where $k=(q-1,2)$. By Proposition 1.7 3) and 4) we have, with $m=2n$, $p^n+1\left|\frac{q+\varepsilon_2}{k}\right.$ for some $\varepsilon_2=\pm 1$ if $\varepsilon=1$, and $p^n-1\left|\frac{q+\varepsilon_2}{k}\right.$ for some $\varepsilon_2=\pm 1$ if $\varepsilon=2$. Remark that since $r=2$ then by 1.6 and 1.7, 2) we may assume that $\varepsilon=1$ or $\varepsilon=2$. We set $\varepsilon_1=1$ if $\varepsilon=1$ and $\varepsilon_2=-1$ if $\varepsilon=2$ and we get $p^n+\varepsilon_1\left|\frac{q+\varepsilon_2}{(q-1,2)}\right.$. If $\varepsilon_1=-1$ we have, in addition, $\varepsilon=2$ and with $|F_2(p^{2n})|=p^{2n}-1$ we have $\pi(F_2(p^{2n}))\cap\pi\neq\varnothing$. This contradicts 3) so concludes the proof of the proposition.

Examples 1.11. Assume hypothesis 1.1 and set $\pi=\pi(G)$. Then we have:

I) If r is any prime and $A\cong SL(2,2^r)$ then the action of A on G is generic. In particular this is true if $A\cong A_5$ the alternating group on five letters.

Proof. A satisfies 1) and 2) of Proposition 1.10. If $p\in\pi$ then $p^n+\varepsilon_1$ is even and $2^p+\varepsilon_2$ is odd, for any $\varepsilon_1=\pm 1$ and $\varepsilon_2=\pm 1$. So 3) is satisfied and every proper subgroup of A is π-regular. By applying Theorem 1.9 we get the result.

II) Assume r is an odd prime and 3) of 1.10 is satisfied for $q=3^r$. Then if $A\cong PSL(2,3^r)$, the action of A on G is generic. In particular this is true for A isomorphic to $PSL(2,3^3)$, $PSL(2,3^5)$ or $PSL(2,3^7)$.

Proof. It is well known that $PSL(2,3^r)$ is minimal simple if r is an odd prime so A satisfies 1) of Proposition 1.10. $3^r\equiv 3$ modulo 8 so that a Sylow 2-subgroup of A is Abelian and 2) of 1.10 is also satisfied. Hence the result.

2. Bounding the Fitting Height

Here we give a bound in terms of A for the Fitting height $h(G)$ of a group G on which A acts fixed point freely, $(|A|,|G|)=1$ and the action of A on G is generic. In the next section we prove that this is the best possible bound. We first need to recall a concept from [17].

Definition 2.1. Let G be a solvable group and A act on G. A subgroup P is called a *generating A-support subgroup* of G if:

1) $P\lhd AG$, $P\subseteq G$ and P is a p-group for some prime p.

2) There are AG-invariant subgroups P_1 and H such that

A) $P_1\subseteq Z(P)$, P/P_1 is elementary Abelian and AG-completely reducible,

B) $H\subseteq C_G(P_1)$,

C) $H/H\cap C_G(P/P_1)$ is elementary Abelian for some prime r,

D) H acts non-trivially on each H-chief factor of P/P_1.

We call the A-support of G (denoted $\mathrm{supp}_A(G)$) the subgroup generated by all subgroups $S \subseteq G$ such that $S \lhd AG$ and either S is Abelian or a generating A-support subgroup of G.

We state some elementary properties of the A-support of G which appear in Proposition 4.3 of [17]. Here $F(G)$ denotes the Fitting subgroup of G.

Proposition 2.2. *Let G be a solvable group and A act on G. Then we have the following*:

1) 1) $\bigcap C_G(X) \subseteq F(G)$ *(where X runs through the AG-chief factors of* $\mathrm{supp}_A(G)$). *In particular* $C_G(\mathrm{supp}_A(G)) \subseteq F(G)$.

2) *If* $N \subseteq G$ *and* $N \lhd AG$, $\mathrm{supp}_A(G)\, N/N \subseteq \mathrm{supp}_A(G/N)$.

3) *If* $B \subseteq A$ *and* $(|A|, |G|) = 1$, $\mathrm{supp}_B(G) \supseteq \mathrm{supp}_A(G)$.

4) $C_A(\mathrm{supp}_A(G)) \subseteq C_A(G/F(G))$.

Our main results are based on the following proposition.

Proposition 2.3. *Assume AG is a finite group where $G \lhd AG$ is solvable,* $(|A|, |G|) = 1$ *and V a $\mathbb{C}AG$-module. Assume the following*:

1) $V|_G$ *is homogeneous and faithful*;

2) $C_V(A) = 0$;

3) *Every subgroup of A satisfies both the regular orbit and the regular module property in its action on G.*

Then

$$C_V(C_A(\mathrm{supp}_A(G))) = 0.$$

Proof. This is a variation of Proposition 4.5 of [17]. For a proof just follow the proof of [17, Proposition 4.5] and use 3) in place of [17, Theorem 2.2] and [17, Theorem 3.4].

Definition 2.4. Let A be any finite group and let B be any subgroup of A. Define $l(A:B)$ to be the largest integer n such that there is a sequence of subgroups

$$B = C_0 \subset C_1 \subset \ldots \subset C_n = A,$$

each properly contained in the following one.

We are now in a position to state the main result of the paper.

Theorem 2.5. *Let A be a finite group, B a subgroup of A and let G be a solvable group on which A acts. Suppose* $(|A|, |G|) = 1$, $C_G(A) = 1$, *the action of A on G is generic and* $\bigcap_{a \in A} [B, G]^a = 1$.

Then $h(G) \leqq l(A:B)$.

Proof. Suppose false. Let A, B, G provide a counterexample with $|A:B|$ minimum. Clearly $h(G) > 1$. Set $h = h(G) - 1$. By [16, Lemma 1.4] and [16, Lemma 1.9], G contains an irreducible A-tower $(\hat{P}_i)\ i = 1, \ldots, h+1$ of height $h+1$. This means that $\hat{P}_i\ (i = 1, \ldots, h+1)$ is an A-invariant subgroup of G and the following is satisfied.

1. $\pi(\hat{P}_i)=\{p_i\}$ consists of a single prime for $i=1,\dots,h+1$.

2. $\hat{P}_i$ normalizes $\hat{P}_j$ for $i<j$.

3. We set $P_{h+1}=\hat{P}_{h+1}$ and $P_i=\hat{P}_i/C_{\hat{P}_i}(P_{i+1})$ for $i=1,\dots,h$, and $P_i\neq 1$ for $i=1,\dots,h+1$.

4. $p_i\neq p_{i+1}$ for $i=1,\dots,h$.

5. $\Phi(\Phi(P_i))=1$, $\Phi(P_i)\subseteq Z(P_i)$ and if $p_i\neq 2$ $\exp(P_i)=p_i$, for $i=1,\dots,h+1$, and $\hat{P}_{i-1}$ centralizes $\Phi(P_i)$ $i=2,\dots,h+1$; (where Φ denotes the Frattini subgroup, Z the center and exp the exponent);

6. P_1 is elementary Abelian.

7. There exists H_i an elementary Abelian subgroup of P_{i-1} normalized by A such that $[H_i,P_i]=P_i$ for $i=2,\dots,h+1$.

8. If $Q\subseteq\hat{P}_i$ for some i, Q is normalized by $A\hat{P}_1\dots\hat{P}_{i-1}$ and its image in P_i is not contained in $\Phi(P_i)$, then $Q=\hat{P}_i$.

We set $H=\hat{P}_1\dots\hat{P}_h$ and $V=P_{h+1}/\Phi(P_{h+1})$. Now by 8., setting $\mathbb{F}=\mathbb{F}_{p_{h+1}}$, V is an irreducible $\mathbb{F}AH$-module. Since $\hat{P}_{h+1}$ is normalized by A and $\bigcap_{a\in A}[B,G]^a=1$ it follows that $\hat{P}_{h+1}\nsubseteq[B,G]$. Recall that from 3. $P_{h+1}=\hat{P}_{h+1}$. Let P_0 be a maximal subgroup of P_{h+1} which contains $P_{h+1}\cap[B,G]$. Then $P_0\supseteq(P\cap[B,G])\,\Phi(P)$. Define $K=(P\cap[B,G])\,\Phi(P)/\Phi(P)$. P_0 shows that K is a proper BH-submodule of V. By Clifford's Theorem $V|_H$ is completely reducible and, since $(|B|,|V|)=1$, by Maschke's Theorem $V|_{BH}$ is completely reducible. Let S_0 be an irreducible BH-submodule of V such that $S_0\cap K=0$. Then $[B,S_0]\subseteq S_0\cap K=0$, so that $[B,H]\subseteq\ker_H(S_0)$ and $S_{0|H}$ is irreducible. Let $S\subseteq V$ be the homogeneous H-component of $S_{0|H}$. We have (from 3. it follows that $\hat{P}_h$ does not centralize V):

9. $C_S(B)\neq 0$, $\hat{P}_h\nsubseteq\ker_H(S)$ and $[B,H]\subseteq\ker_H(S)$.

We set $B_0=N_A(S)$. Since $[B,H]\subseteq\ker_H(S)$ we have $B_0\supseteq B$ and we may consider S as an irreducible B_0H-module. By Clifford and Mackey we have $V|_A\simeq S^{AH}|_A\simeq S|_{B_0}|^A$. From $C_V(A)=0$ (which follows from $C_G(A)=1$ and $(|A|,|G|)=1$) we get:

10. $C_S(B_0)=0$.

Let now k be the algebraic closure of $\mathbb{F}$. Let I be an irreducible submodule of $k\otimes_{\mathbb{F}}S$. By the Fong-Swan Theorem since B_0G is p_{h+1}-solvable we may take a $\mathbb{C}B_0G$-irreducible $\mathbb{C}B_0G$-module M_0 such that $\ker(M_0)=\ker(S)$ and M_0 "when reduced modulo p_{h+1}" gives I. Then it follows from 9. and 10. that:

11. $C_{M_0}(B)\neq 0$, $\hat{P}_h\nsubseteq\ker(M_0)$, $[B,H]\subseteq\ker(M_0)$ and $C_{M_0}(B_0)=0$.

Since $[B,H]\subseteq\ker(M_0)$, B normalizes each homogeneous component of $M_0|_H$. Let M be a homogeneous component of $M_0|_H$ such that $C_M(B)\neq 0$ and set $B_1=N_{B_0}(M)$. Then

$$M_{0|B_0}=M|_{B_1}|^{B_0},$$

so that $C_M(B_1)=0$. We get, since A normalizes $\hat{P}_h$ that $\hat{P}_h\nsubseteq\ker(M)$ from 11. So we have the following information about M:

12. M is an irreducible $\mathbb{C}B_1H$-module, $M|_H$ is homogeneous, $\hat{P}_h\nsubseteq\ker(M)$, $[B,H]\subseteq\ker(M)$, $C_M(B)\neq 0$ and $C_M(B_1)=0$.

If X is any subgroup of B_1H, we denote by X_M its image in $B_1H/\ker(M)$. Let $D=C_{B_1}(\mathrm{supp}_{B_1}(H_M))$. Set $P=(\hat{P}_h)_M$ and $R=(\hat{P}_1\dots\hat{P}_{h-1})_M$, we have $H_M=RP$. It follows from 5.–7. that

13. $P\subseteq \mathrm{supp}_{B_1}(H_M)$ and D centralizes P.

Let $C=C_R(P)$. Then $C\lhd RP$. We have $[D,P,R]=1$ and $[P,R,D]=1$ so, by the three subgroup lemma, we get $[R,D,P]=1$, i.e. $[R,D]\subseteq C$. Since $C\lhd RP$ and D centralizes P we have $[D,RP]\subseteq C$. Furthermore, $P\nsubseteq R$, since otherwise $h>1$ and by 2., P would normalize $(\hat{P}_{h-1})_M$ and hence $[\hat{P}_{h-1},\hat{P}_h]\subseteq\ker(M)$, but $[\hat{P}_{h-1},\hat{P}_h]=\hat{P}_h$ by 3. and 8. so this contradicts 12. Since $[D,RP]\subseteq C\subseteq R$ and $P\nsubseteq R$ we have $[D,H]\nsupseteq P_h$. Let $T=\bigcap_{a\in A}[D,H]^a$. Then we get $\hat{P}_h\nsubseteq T$ and therefore by 8. $\hat{P}_h\cap T$ maps into $\Phi(P_h)$, so we have by 3. and 4.:

14. $h(H/T)=h$.

If $A=D$ then $T=[A,H]$ so A centralizes H/T and hence $H=T$, against the fact that $h>0$. Therefore by 12. and the fact that the action of A on G is generic we may apply Proposition 2.3 to M and we get that $C_M(D)=0$. Since, by 12., clearly $D\supseteq B$ we obtain $B\subset D$. By the definition of T we may apply induction to A, D, H/T and deduce that $h(H/T)\leqq l(A:D)$. Since $B\subset D$ we have $l(A:D)\leqq l(A:B)-1$. From 14. we get:

15. $h\leqq l(A:B)-1$.

Now $h(G)=h+1\leqq l(A:B)$, a contradiction. This completes the proof of the theorem.

Since $l(A)=l(A:1)$, Theorem A is now an obvious corollary of Theorem 2.5. Likewise Theorem C follows immediately from Proposition 1.5 and Theorem 2.5. Furthermore from Theorem 1.9 and Theorem 2.5 we get:

Theorem 2.6. *Let A be a finite group, B a subgroup of A and let G be a solvable group on which A acts. Suppose $(|A|,|G|)=1$, $C_G(A)=1$ and $\bigcap_{a\in A}[B,G]^a=1$. Assume in addition that every proper subgroup of A is $\pi(G)$-regular (Definition 1.8). Then $h(G)\leqq l(A:B)$.*

This generalizes Theorem 4.7 of [17] to the non-supersolvable case. In addition Theorem D, which is a special case of Theorem 2.6, follows from Examples 1.11 and Theorem 2.5. It remains to show Theorem B which we will do in the next section.

3. The Bound is Sharp

Lemma 3.1. *Let G be a finite solvable group and assume that, for $i=1,\dots,h$, P_i is a p_i-subgroup of G (p_i a prime) and for $i=1,\dots,h-1$ we have $p_i\neq p_{i+1}$ and $[P_i,P_{i+1}]=P_{i+1}$. Then, if $P_h\neq 1$, we have $h(G)\geqq h$.*

Proof. Use induction on $|G|$. The lemma is clear if $h=1$. Take G satisfying the hypothesis of the lemma with $h\geqq 2$. If $F(G)\supseteq P_{h-1}$ then, since $p_{h-1}\neq p_h$, $[P_{h-1},P_h]=1$, a contradiction. So $P_{h-1}\nsubseteq F(G)$ and by induction we get $h(G/F(G))\geqq h-1$. Hence the lemma.

The following result provides the key construction for the proof that our bounds in the previous section are sharp.

Theorem 3.2. *Let A be any finite group, B_1 any subgroup of A and B a subgroup of B_1, $B_1 \neq B$. Suppose G is a finite group and A acts on G, $Q \subseteq G$ is a normal subgroup of AG such that $[B_1, G] \not\supseteq Q$. Then for any prime $p \nmid |AQ|$ there exists an $\mathbb{F}_p AG$-module M with the following properties.*

1) $[Q, M] = M$;
2) $C_M(A) = 0$;
3) $[B, GM] \not\supseteq M$.

Proof. Since $[B_1, G] \lhd G$ we may take an irreducible $\mathbb{F}_p(G/[B_1, G])$-module N where Q acts non-trivially. Since $B_1 \neq B$ we may take R a cyclic subgroup of B_1 not contained in B. Let N_0 be a non-trivial irreducible $\mathbb{F}_p(R/R \cap B)$-module. Set $N_1 = N_0^{B_1}$ and take

$$M_0 = N_1 \otimes N.$$

Since $[B_1, G]$ acts trivially on N we may consider M_0 as an $\mathbb{F}_p B_1 G$-module. It clearly satisfies.

a) $[Q, M_0] = M_0$ and $[B_1, G] \subseteq \ker(M_0)$.

Frobenius reciprocity gives:

b) $C_{M_0}(B_1) = 0$.

Furthermore, $C_{M_0}(B)$ is G-invariant since $[B, G] \subseteq \ker(M_0)$. By Mackey,

$$N_{1|B} \simeq \sum N_0^x|_{R^x \cap B}|^B$$

the sum over x, representatives of the double cosets RxB in B_1. In particular

$$N_{1|B} \supseteq N_0|_{R \cap B}|^B$$

and hence $C_{N_1}(B) \neq 0$. So

c) $C_{M_0}(B)$ is a non-zero BG-module.

Define $M = M_0^{AG}$. Since $Q \lhd AG$, 1) follows from a), 2) follows from b) by Frobenius reciprocity. By the definition of M we see that $M|_G$ is completely reducible so that by Maschke's Theorem $M|_{BG}$ is completely reducible. Set $C = C_{M_0}(B)$. By c), C is a non-zero BG-submodule of M. Take S a BG-submodule of M such that $C \cap S = 0$ and $C + S = M$. We have now $[B, G]\, S \lhd BGS$. Since $C \subseteq M_0$, by a), we get $[B, G]\, S \subseteq C_{GM}(C)$ so that $[B, G]\, S \lhd BGM$. Now $[B, M] \subseteq S$ since $C \subseteq C_M(B)$ and S is a B-submodule. Hence $[B, GM] \subseteq [B, G]\, S$ and $[B, GM] \cap M \subseteq S$, but by c) $S \neq M$. Hence we get 3). This concludes the proof of the theorem.

Theorem 3.3. *Let A be any finite group and B a proper subgroup of A. Let $n = l(A:B)$ (Definition 2.4) and let $p_1, \dots, p_n$ be a sequence of primes such that $p_i \neq p_{i+1}$ $(i = 1, \dots, n-1)$ and $p_i \nmid |A|$ $(i = 1, \dots, n)$.*

Then there exists a solvable group G, and an action of A on G, and subgroups $P_i (i = 1, \dots, n)$ of G such that:

1) *For each* $i=1,\dots,n$, P_i *is an elementary Abelian* p_i*-subgroup of* G, *is invariant under* $A\ P_1\dots P_{i-1}$, *and if* $i>1$, $[P_{i-1},P_i]=P_i$;

2) $G=P_1\dots P_n$;

3) $C_G(A)=1$;

4) $\bigcap_{a\in A}[B,G]^a=1$;

5) $P_n\neq 1$ *and* $h(G)=n$.

Proof. Use induction on $l(A:B)$. If $l(A:B)=1$ take R a cyclic subgroup of A not in B, N a non-trivial irreducible $\mathbb{F}_{p_1}(R/R\cap B)$-module and $M=N^A$. Then $C_M(A)=0$ by Frobenius reciprocity and $C_M(B)\neq 0$. Take $G=P_1=M/\bigcap_{a\in A}[B,M]^a$. Since $C_M(B)\neq 0$ we have $P_1\neq 1$ and 1)–5) are satisfied in this case.

Hence we assume $n>1$. Take $B_1\supset B$ a subgroup of A such that $l(A:B_1)=l(A:B)-1$. Then, by induction, we may take $G=P_1\dots P_{n-1}$ satisfying the conclusion of the theorem for A and B_1 and the sequence $p_1,\dots,p_{n-1}$. Since

$$\bigcap_{a\in A}[B_1,G]^a=1$$

and $1\neq P_{n-1}\lhd AG$ we have that $P_{n-1}\nsubseteq[B_1,G]$. Set $Q=P_{n-1}$ and $p=p_n$. Take an M given by Theorem 3.2. Now $C_G(A)=1$ and $C_M(A)=0$ so that we have $C_{GM}(A)=1$. Take $H=GM/\bigcap_{a\in A}[B,GM]^a$. Set P_n to be the image of M in H. Now $P_n\neq 1$ by 3) of Theorem 3.2 and by considering the images of $P_1,\dots,P_{n-1}$ in H we see that H provides a group of the form described in the statement of the theorem except possibly $h(H)=n$. But this follows from the other conditions and Lemma 3.1.

Theorem B now follows as a special case where $B=1$ and $p_i=p$ for odd i and $p_i=q$ for even i.

References

1. Bailey, J.: Nilpotent groups acting fixed point freely on solvable groups. Bull. Austral. Math. Soc. **15**, 339–346 (1976)
2. Berger, T.R.: Hall-Higman type theorems V. Pacific J. Math. **73**, 1–62 (1977)
3. Berger, T.R.: Nilpotent fixed point free automorphism groups of solvable groups. Math. Z. **131**, 305–312 (1973)
4. Feldman, A.D.: Fitting height of solvable groups admitting fixed-point-free automorphism groups. J. Algebra **53**, 268–295 (1978)
5. Gorenstein, D.: Finite simple groups. New York and London: Plenum Press, 1982
6. Gross, F.: A note on fixed point free solvable operator groups. Proc. Amer. Math. Soc. **19**, 1363–1365 (1968)
7. Huppert, B.: Endliche Gruppen I. New York: Springer 1967
8. Isaacs, I.M.: Symplectic action and the Schur index. Proc. Sympos. Pure Math. **12**, 73–75 (1971)
9. Isaacs, I.M.: Character theory of finite groups. New York: Academic Press, 1976

10. Kurzweil, H.: Auflösbare Gruppen auf denen nicht auflösbare Gruppen operieren. Manuscripta Math. **41**, 233–305 (1983)
11. Kurzweil, H.: Die Gruppe A_5 als fixpunktfreie Automorphismengruppe. Illinois J. Math. **27**, 67–76 (1983)
12. Kurzweil, H.: Minimal einfache Gruppen als fixpunktfreie Automorphismengruppen. Comm. Algebra **5**, 397–442 (1977)
13. Kurzweil, H.: p-Automorphismen von auflösbaren p'-Gruppen. Math. Z. **120**, 326–354 (1971)
14. Thompson, J.G.: Automorphisms of solvable groups. J. Algebra **1**, 259–267 (1964)
15. Thompson, J.G.: Finite groups with fixed point free automorphism group of prime order. Proc. Nat. Acad. Sci. **45**, 578–581 (1959)
16. Turull, A.: Fitting height of groups and of fixed points. J. Algebra **86**, 555–566 (1984)
17. Turull, A.: Supersolvable automorphism groups of solvable groups. Math. Z. **183**, 47–73 (1983)

Received January 3, 1984

Math. Z. 187, 505–510 (1984)

Mathematische Zeitschrift

Über λ-Ringstrukturen auf dem Burnside-Ring

Tammo tom Dieck

Mathematisches Institut der Universität, Bunsenstr. 3–5, 3400 Göttingen, Bundesrepublik Deutschland

Der Burnside-Ring $A(G)$ einer endlichen Gruppe G hat zusätzliche Struktur, nämlich λ-Operationen und daraus abgeleitete Adams-Operationen Ψ. Eine λ-Struktur, die auf äußeren oder symmetrischen Potenzen beruht, wurde von Knutson [6] eingeführt und unter anderem von Siebeneicher [10] und Gay-Morris-Morris [5] weiter untersucht. Wie Siebeneicher feststellt, macht diese λ-Ringstruktur allerdings $A(G)$ höchstens für zyklische Gruppen G zu einem speziellen λ-Ring im Sinne von Atiyah-Tall [1]. Eine im allgemeinen andere λ-Struktur wurde von Knutson [6] vorgeschlagen und in Spezialfällen von Blass [2] und Ochoa [7] untersucht.

Spezielle λ-Ringstrukturen erhält man, wenn die zugehörigen Ψ-Operationen die folgenden Eigenschaften haben – wir wollen dann von einem Ψ-Ring sprechen: $\Psi^1 = \mathrm{id}$, Ψ^k ist ein Ringhomomorphismus, $\Psi^k \Psi^l = \Psi^{kl}$ (s. Knutson [6], p. 49). Man kann sich als Leitgedanken vorstellen, daß Ψ^p für eine Primzahl p einem Frobenius-Endomorphismus „p-te Potenz" entspricht. Es ist deshalb naheliegend, p-Gruppen G zu untersuchen, bei denen die Abbildung $g \mapsto g^p$ angenehme Eigenschaften hat.

In dieser Arbeit möchte ich eine Gruppe G eine Ψ^p-Gruppe nennen, wenn sie die folgenden Eigenschaften hat: G ist eine p-Gruppe und für jede Untergruppe H von G ist $H(1) = \{g^p \mid g \in H\}$ und $H(-1) = \{g \mid g^p \in H\}$ wiederum eine Untergruppe von G. Sicherlich sind abelsche p-Gruppen Ψ^p-Gruppen, aber auch viele nicht-abelsche Gruppen, etwa Gruppen vom Exponenten p. Untergruppen und Faktorgruppen von Ψ^p-Gruppen sind wiederum Ψ^p-Gruppen. Der Begriff steht in Relation zu den P. Hallschen regulären p-Gruppen.

Der Burnside-Ring $A(G)$ wird als Teilring des Ringes $C(G)$ aller Funktionen auf der Menge $\Phi(G)$ der Konjugationsklassen (H) von Untergruppen H in die ganzen Zahlen $\mathbb{Z}$ aufgefaßt (s. tom Dieck [3], 1.3.5 für eine Beschreibung dieses Teilringes durch die Dressschen Kongruenzen). Sei G eine Ψ^p-Gruppe. Es wird $C(G)$ zu einem Ψ-Ring, wenn man Operationen Ψ^k folgendermaßen festsetzt: $\Psi^p d(H) = d(H(1))$ *für alle* $d \in C(G)$, $(H) \in \Phi(G)$, $\Psi^k d = d$ falls k zu p teilerfremd ist, und im übrigen $\Psi^k \Psi^l = \Psi^{kl}$ verlangt. Für die kanonische Basis der $[G/H]$ von $A(G) \subset C(G)$ gilt dann

$$\Psi^p[G/H] = |H(-1)/H| \, [G/H(-1)], \tag{1}$$

weshalb der Unterring $A(G)$ ebenfalls zu einem Ψ-Ring wird. Ist $IA(G)$ das Augmentationsideal von $A(G)$, nämlich der Kern von $A(G) \to \mathbb{Z}$, $d \mapsto d(1)$, so folgt aus (1) die grundlegende Eigenschaft

$$\Psi^p IA(G) \subset p IA(G), \tag{2}$$

d.h. $p^{-1} \Psi^p$ ist auf $IA(G)$ definiert. Angenommen nun, die Ψ-Operationen wären in der üblichen Weise aus λ-Operationen abgeleitet. Dann müßte wegen der Relation $\Psi^p x \equiv x^p \mod p$ auch gelten

$$x \in IA(G) \quad \text{impliziert} \quad x^p \in p IA(G), \tag{3}$$

was allerdings nicht für jede p-Gruppe G erfüllt ist. Es kann deshalb auch nicht andere Ψ-Strukturen mit der fundamentalen Eigenschaft (2) geben.

Nach diesen Vorbereitungen können wir das Hauptresultat formulieren.

Satz 1. *Sei G eine Ψ^p-Gruppe. Die oben definierte Ψ-Struktur auf $A(G)$ induziert λ^i-Operationen, die $A(G)$ zu einem speziellen λ-Ring macht.*

Zum Beweis dieses Satzes müssen wir erläutern, wie sich λ- und Ψ-Operationen ineinander umrechnen lassen. Dazu verwendet man zweckmäßig die erzeugenden formalen Potenzreihen in einer Unbestimmten t

$$\lambda_t(x) = 1 + \lambda^1(x)\, t + \lambda^2(x)\, t^2 + \dots, \tag{4}$$

$$\Psi_t(x) = 1 + \Psi^1(x)\, t + \Psi^2(x)\, t^2 + \dots. \tag{5}$$

Üblicherweise definiert man dann die Ψ-Operationen aus den λ-Operationen durch die formale Identität

$$-t \frac{d}{dt} (\log \lambda_t(x)) = \Psi_{-t}(x) \tag{6}$$

(s. etwa Atiyah-Tall [1], Knutson [6]).

Falls man in dem betrachteten Ring durch ganze Zahlen eindeutig dividieren kann, läßt sich (6) nach λ_t auflösen und man erhält

$$\lambda_t(x) = \exp\left(\sum_{i=1}^{\infty} (-1)^{i-1} \Psi^i(x) \frac{t^i}{i}\right). \tag{7}$$

Es ist ebenso gut, aber zur Gewinnung einheitlicher Formeln im folgenden nützlicher, statt (7) die Reihe

$$\lambda_{-t}(x)^{-1} = S(t,x) = 1 + S^1(x)\, t + S^2(x)\, t^2 + \dots \tag{8}$$

zu betrachten, deren Bezeichnung an symmetrische Potenzen erinnern soll. Im Fall der oben definierten Ψ-Operationen auf $A(G)$ nimmt die Formel (8) die folgende, das p-adische Herz erquickende Gestalt an (Lemma 1). Zur Formulierung verwenden wir die folgende Potenzreihe in Unbestimmten t und x

$$L(t,x)=(1-t)^x(1-t^p)^{-x/p}. \tag{9}$$

Sie ist folgendermaßen zu interpretieren: Es wird $(1-t)^x$ formal in die Binomialreihe entwickelt

$$1-\frac{x}{1}t+\frac{x(x-1)}{1\cdot 2}t^2+\dots,$$

entsprechend mit dem anderen Faktor verfahren und dann ausmultipliziert. Setzt man für x ein Element aus $A(G)$ ein, so sind die Koeffizienten von t^i natürlich zunächst nur Elemente von $A(G)\otimes\mathbb{Q}=C(G)\otimes\mathbb{Q}$.

Lemma 1. *Für $x\in A(G)$ gilt*

$$S(t,x)=L(t,x)\,S(t^p,p^{-1}\,\Psi^p p),$$

woraus durch Iteration

$$S(t,x)=L(t,x)\,L(t^p,p^{-1}\,\Psi^p x)\,L(t^{p^2},p^{-2}\,\Psi^{p^2}x)\cdot\dots$$

folgt.

Beweis. Die erste Identität in Lemma 1 bestätigt man durch eine Rechnung:

$$\sum_{i=1}^{\infty}\Psi^i(x)\frac{t^i}{i}=\sum_{(i,p)=1}x\frac{t^i}{i}+\sum_{i=1}^{\infty}\Psi^{ip}(x)\frac{t^{ip}}{ip}.$$

Der zweite Summand hat die Form

$$\sum_{i=1}^{\infty}\Psi^i\left(\frac{\Psi^p x}{p}\right)\frac{(t^p)^i}{i},$$

was durch exp in $S(t^p,p^{-1}\,\Psi^p(x))$ verwandelt wird. Der erste Summand wird als

$$\sum_{i=1}^{\infty}x\frac{t^i}{i}-\sum_{i=1}^{\infty}x\frac{t^{pi}}{pi}$$

geschrieben und durch exp bei Beachtung der formalen Identität $\exp\sum i^{-1}t^i=(1-t)^x$ in $L(t,x)$ verwandelt.

Als nächsten Schritt im Beweis von Satz 1 zeigen wir, daß die $S^i(x)$ für $x\in A(G)$ jedenfalls in $C(G)$ liegen. Wir wissen: $S(t,x+y)=S(t,x)\,S(t,y)$. Für das Einselement $1\in A(G)$ gilt $S(t,1)=1-t$. Es genügt also, $x\in IA(G)$ zu betrachten. Man erhält den Wert der Funktion $S^i(x)$ an der Stelle (H) dadurch, daß man in den Formeln aus Lemma 1 statt x die ganze Zahl $x(H)$ einsetzt. Wir wollen einsehen, daß die entstehende Potenzreihe in t ganzzahlige Koeffizienten hat. Das folgt jedenfalls dann, wenn wir wissen, daß $(\Psi^{p^i}(x)-\Psi^{p^{i-1}}(x))(H)$ jeweils durch p^i teilbar ist, weil $(1-t)^k$ für $k\in\mathbb{Z}$ ganzzahlige Koeffizienten hat. Deshalb zeigen wir

Lemma 2. *Sei G eine Ψ^p-Gruppe. Sei $H\subset G$ und $H(i)=\{g^{p^i}|g\in H\}\neq H(i-1)$. Für $d\in A(G)$ und $i\geqq 1$ gilt dann $d(H(i))\equiv d(H(i-1))\bmod p^i$.*

Beweis. $H(i)$ ist ein Normalteiler von G. Für jede endliche G-Menge X ist $|X^{H(i)}| \equiv |X^{H(i-1)}| \bmod p^i$ zu zeigen. Wir betrachten $X^{H(i)} = Y$ als $H/H(i)$-Menge. Wir zeigen $|Y| \equiv |Y^{H(i-1)/H(i)}| \bmod p^i$; mit anderen Worten: Es genügt, den Fall $G=H$, $H(i)=1$ zu betrachten, was wir auch tun wollen. Es hat dann nach Voraussetzung G Elemente der Ordnung p^i. Sei $x \in X \setminus X^{H(i-1)}$ also $H(i-1)$ nicht subkonjugiert zur Isotropiegruppe G_x. Es genügt zu zeigen: $|G/G_x| \equiv 0 \bmod p^i$. Es gibt nun eine Konjugationsklasse eines Elementes $z \in H(i-1)$, die mit G_x leeren Schnitt hat. Dann ist aber $G/G_x^z = \varnothing$ und generell gilt $|X| \equiv |X^z| \bmod p^i$, wenn z eine p^{i-1}-te Potenz eines Elementes der Ordnung p^i ist, denn es handelt sich um eine der Kongruenzen für den Burnside-Ring einer zyklischen Gruppe $\mathbb{Z}/p^i$.

Der Nutzen von Lemma 2 besteht in folgendem: Wir wollen zeigen, daß $S^i(x)$ für $x \in A(G)$ in $A(G)$ liegt. Da $S^i(x)$ in $C(G)$ liegt und $C(G)/A(G)$ eine abelsche p-Gruppe ist, genügt es zu zeigen, daß $S^i(x)$ in $A(G)_{(p)}$, in der Lokalisierung bei p, liegt. Zu diesem Zweck verwenden wir das folgende fundamentale Ganzheitslemma.

Lemma 3. *Die kommutative $\mathbb{Z}_{(p)}$-Algebra I sei als additive Gruppe torsionsfrei. Es gelte $x^p \in pI$ für alle $x \in I$. Dann hat für alle $x \in I$ die t-Reihe $L(t,x)$ Koeffizienten in I.*

Beweis. Apriori liegen die Koeffizienten in $I \otimes_{\mathbb{Z}_{(p)}} \mathbb{Q}$. Der Beweis verwendet einen Ganzheitssatz von Dwork [4], p. 634: Man betrachte die formale Potenzreihe in x und t

$$L(t,x)\, L(t^p, p^{-1} x^p)\, L(t^{p^2}, p^{-2} x^{p^2}) \cdot \ldots; \tag{10}$$

sie hat p-ganze Koeffizienten. Dies angenommen, zeigen wir induktiv nach n, daß $L(t,x)$ für $x \in I$ für $k < p^n$ Koeffizienten $q_k(x) \in I$ von t^k hat. Für $n=1$ ist das offenbar richtig, da dieser Koeffizient dann $(-1)^k \binom{x}{k}$ ist. Sei nun $p^n \leqq k < p^{n+1}$ und sei $a_j(x)$ für $j<k$ in I enthalten. Wir betrachten den Koeffizienten von t^k in (10). Nur die Faktoren $L(t^{p^i}, p^{-i} x^{p^i})$ für $i<n$ sind relevant und von dem Faktor für $i=n$ nur der Zähler $(1-t^{p^n})^y$, $y = p^{-n} x^{p^n}$. Da $p^{-i} x^{p^i} \in I$, so sind nach Induktionsvoraussetzung alle Koeffizienten dieser Faktoren, die beim Ausmultiplizieren einen Beitrag zum Koeffizienten von t^k liefern können, in I enthalten bis auf den Koeffizienten $a_k(x)$, d.h. der Koeffizient $c_k(x)$ von t^k in (10) ist modulo I gleich $a_k(x)$. Da nach dem zitierten Satz von Dwork aber $c_k(x)$ aus I ist, so folgt $a_k(x) \in I$.

Lemma 4. *Sei G eine Ψ^p-Gruppe. Dann gilt die Implikation* (3).

Wir stellen vor dem Beweis zunächst die Verbindung zu rationalen Darstellungen her. Sei $R(G;\mathbb{Q})$ der rationale Darstellungsring und sei

$$\pi_G = \pi\colon\ A(G) \to R(G;\mathbb{Q})$$

der Ringhomomorphismus, der einer endlichen G-Menge S ihre Permutationsdarstellung $Q(S)$ über Q zuordnet. Wir haben das Augmentationsideal $IR(G;\mathbb{Q})$, nämlich den Kern der Abbildung $R(G;\mathbb{Q}) \to \mathbb{Z}$, $\chi \mapsto \chi(1)$, Evalution von

Charakteren am neutralen Element. Die Abbildung π ist mit den Ψ-Operationen verträglich, wenn $R(G;\mathbb{Q})$ die üblichen Ψ-Operationen trägt, die auf Charakteren durch $\Psi^k\chi(g)=\chi(g^k)$ definiert sind. Ferner ist π mit den Augmentationen verträglich. Nach dem Satz von Segal [9] und Ritter [8] ist π für p-Gruppen surjektiv. Falls für ein Ideal I eines Ringes R gilt $x\in I\Rightarrow x^p\in pI$, so sagen wir, R erfülle die (I,p)-Bedingung. Neben den Augmentationsidealen betrachten wir auch die maximalen Ideale $MA(G)$, den Kern von $A(G)\to\mathbb{Z}/p$, $d\mapsto d(1)$, und $MR(G;\mathbb{Q})$, den Kern von $R(G;\mathbb{Q})\to\mathbb{Z}/p$, $\chi\mapsto\chi(1)$. Es erfüllt $A(G)$ die $(IA(G),p)$-Bedingung genau dann, wenn $A(G)$ die $(MA(G),p)$-Bedingung erfüllt; eine analoge Aussage gilt für $R(G;\mathbb{Q})$.

Für eine p-Gruppe G gilt nun

Lemma 5. (i) *Erfüllt $A(G)$ die $(IA(G),p)$-Bedingung, so erfüllt $R(G;\mathbb{Q})$ die $(IR(G;\mathbb{Q}),p)$-Bedingung.*

(ii) *Erfüllt $R(NH/H;\mathbb{Q})$ für alle Untergruppen H von G die $(IR(NH/H;\mathbb{Q}),p)$-Bedingung, so erfüllt $A(G)$ die $(IA(G),p)$-Bedingung.*

Beweis. (i) folgt, weil π eine surjektive Abbildung $IA(G)\to IR(G;\mathbb{Q})$ liefert.

(ii) Statt mit den I-Idealen arbeiten wir mit den oben definierten M-Idealen. Sei $x\in MA(G)$. Dann ist $p^{-1}x^p\in C(G)$.

Wir müssen zeigen, daß diese Funktion in $A(G)$ liegt, also die Dressschen Kongruenzen zur Beschreibung von $A(G)$ erfüllt. Eine solche Kongruenz wird aber folgendermaßen erhalten: Beschränkung auf die H-Fixpunktmenge liefert b_H: $A(G)\to A(NH/H)$. Dieses setze man mit π_H: $A(NH/H)\to R(NH/H;\mathbb{Q})$ zusammen und mit e_H: $R(NH/H;\mathbb{Q})\to\mathbb{Z}/|NH/H|$, $\chi\mapsto\sum_{w\in NH/H}\chi(w)$. Es liegt $d\in C(G)$ genau dann in $A(G)$, wenn für alle $H\subset G$ die Relation $e_H\pi_H d_H(d)=0$ gilt. Sei $d=p^{-1}x^p$. Dann ist $x_H=\pi_H b_H x$ aus $MR(NH/H;\mathbb{Q})$ und weil π_H und b_H Ringhomomorphismen sind, so ist wegen der vorausgesetzten $(IR(NH/H;\mathbb{Q}),p)$-Bedingung das Element d im Kern von $e_H\pi_H d_H$.

Beweis von Lemma 4. Die Gruppen NH/H f"r $H\subset G$ sind wiederum Ψ^p-Gruppen. Deshalb gilt $\Psi^p IA(NH/H)\subset pIA(NH/H)$ und wegen der Verträglichkeit von π und Ψ^p gilt $\Psi^p IR(NH/H;\mathbb{Q})\subset pIR(NH/H;\mathbb{Q})$. Da aber in dem λ-Ring $R(NH/H;\mathbb{Q})$ generell $\Psi^p x\equiv x^p \bmod p$ gilt, so erfüllt also $R(NH/H;\mathbb{Q})$ die $(IR(NH/H;\mathbb{Q}),p)$-Bedingung. Nun wende man Lemma 5 an.

Mit den durch Lemma 1–4 bereitgestellten Mitteln können wir den Beweis von Satz 1 beenden. Wie schon erläutert, genügt es, für $x\in IA(G)$ die Relation $S^i(x)\in IA(G)_{(p)}$ nachzuweisen. Wir verwenden die Identität von Lemma 1. Alle Faktoren sind nach Lemma 3 und (2) Potenzreihen mit Koeffizienten in $IA(G)_{(p)}$. Es ist übrigens $p^{-1}\Psi^p$ ein nilpotenter Operator auf $IA(G)$, so daß nur endlich viele Faktoren von 1 verschieden sind. Damit ist der Satz bewiesen.

Ich möchte die voranstehenden Überlegungen noch durch einige Bemerkungen abrunden. Für einen beliebigen torsionsfreien speziellen λ-Ring läßt sich die Reihe $S(t,x)$ durch die Formel

$$S(t,x)=\prod_{n=1}^{\infty}(1-t^n)^{f(n,x)}$$

mit $f(n,x)=\frac{1}{n}\sum_{d|n}\mu\left(\frac{n}{d}\right)\Psi^d(x)$, μ Möbius-Funktion, angeben, was man mittels Logarithmieren bestätigen kann.

Ist σ: $\Phi(G)\to\Phi(G)$ eine beliebige Abbildung, so kann man $C(G)$ durch $\Psi^p d(H)=d(\sigma H)$, $\Psi^k=\mathrm{id}$ für $(k,p)=1$, wie früher zu einem Ψ-Ring machen. Es gilt dann $\Psi^p d\equiv d^p \bmod p$ $C(G)$. Das kann man benutzen, um zu zeigen, daß $C(G)$ sogar ein spezieller λ-Ring wird. (G ist natürlich immer noch eine p-Gruppe.)

Die Struktur $\Psi^p d(H)=d(H(1))$ kann zu einer λ-Struktur auf $A(G)$ führen, selbst wenn G keine Ψ^p-Gruppe ist. Ebenso können diese Ψ-Operationen auf $A(G)$ leben, ohne daß eine λ-Struktur induziert wird. Diese Ψ- und λ-Struktur existiert auf einem Ideal von endlichem Index in $A(G)$ immer, etwa auf dem Ideal aller Funktionen $d\in C(G)$ mit durch $|G|$ teilbaren Werten.

Literatur

1. Atiyah, M.F., Tall, D.O.: Group representations, λ-rings and the J-homomorphism. Topology **8**, 253–297 (1969)
2. Blass, A.: Natural endomorphisms of Burnside rings. Trans. Amer. Math. Soc. **253**, 121–137 (1979)
3. tom Dieck, T.: Transformation groups and representation theory. Lecture Notes in Math. **766**. Berlin-Heidelberg-New York: Springer 1979
4. Dwork, B.: On the rationality of the zeta function of an algebraic variety. Amer. J. Math. **82**, 631–648 (1960)
5. Gay, C.D., Morris, G.C., Morris, I.: Computing Adams operations on the Burnside ring of a finite group. J. Reine Angew. Math. **341**, 87–97 (1983)
6. Knutson, D.: λ-rings and the representation theory of the symmetric group. Lecture Notes in Math. **308**. Berlin-Heidelberg-New York: Springer 1973
7. Ochoa, G.: λ-ring structure on the Burnside ring of Hamilton groups. To appear
8. Ritter, J.: Ein Induktionssatz für rationale Charaktere von nilpotenten Gruppen. J. Reine Angew. Math. **254**, 133–151 (1972)
9. Segal, G.: Permutation representations of finite p-groups. Quart. J. Math. Oxford (2), **23**, 375–381 (1972)
10. Siebeneicher, C.: λ-Ringstrukturen auf dem Burnsidering der Permutationsdarstellungen einer endlichen Gruppe. Math. Z. **146**, 223–238 (1976)

Eingegangen am 6. April 1984

Math. Z. 187, 511–517 (1984)

Mathematische Zeitschrift

A Global Compactness Result for Elliptic Boundary Value Problems Involving Limiting Nonlinearities

Michael Struwe*

Forschungsinstitut für Mathematik, ETH-Zentrum, CH-8092 Zürich, Switzerland

1.

Let Ω be a (smoothly) bounded domain in $\mathbb{R}^n$, $n>2$, and $2^*=\dfrac{2n}{n-2}$, $\lambda\in\mathbb{R}$.

Recently, Brezis and Nirenberg [4] presented a variational approach to the elliptic boundary value problem

$$\begin{aligned}&-\Delta u-\lambda u=u|u|^{2^*-2} \quad \text{in } \Omega\\ &u|\partial\Omega=0\end{aligned} \tag{1.1}$$

and were able to prove the following

Theorem 1.1 (Brezis-Nirenberg). *There exists $\lambda_*\in[0,\lambda_1[$ such that for any $\lambda\in]\lambda_*,\lambda_1[$ problem* (1.1) *admits a positive solution. If $n\geqq 4$, moreover, $\lambda_*=0$.*

Here λ_1 denotes the first eigenvalue of the Laplace operator acting on the Sobolev space $H_0^{1,2}(\Omega)$, the closure of $C_0^\infty(\Omega)$ in the norm

$$\|u\|_1^2=\int_\Omega |\nabla u|^2\,dx.$$

Their result provides some insight into the behaviour of the global branch of positive solutions of (1.1) emanating from $(\lambda=\lambda_1,\ u=0)$ whose existence is guaranteed by a result of Rabinowitz [13]. But more important their method had to overcome a lack of compactness for the variational problem related to (1.1) of finding stationary points of the functional

$$E(u)=\tfrac{1}{2}\int_\Omega [|\nabla u|^2-\lambda u^2]\,dx-\tfrac{1}{2^*}\int_\Omega |u|^{2^*}\,dx \tag{1.2}$$

on $H_0^{1,2}(\Omega)$.

* The author is indebted to the Eidgenössische Technische Hochschule Zürich for its kind hospitality and support

Indeed, invariance of the $H_0^{1,2}$- and L^{2^*}-norms with respect to rescaling

$$u \to u_{r,x_0} = r^{\frac{n-2}{2}}(u(r(\cdot - x_0))) \tag{1.3}$$

and the existence of non-trivial entire solutions of the "limiting problem" (cp. [9])

$$\begin{aligned} -\Delta u &= u|u|^{2^*-2} && \text{in } \mathbb{R}^n \\ u(x) &\to 0 && (|x| \to \infty), \end{aligned} \tag{1.4}$$

(the well-known [8] one-instanton solutions of Yang-Mills equations

$$u^*_{\mu,x_0}(x) = \frac{[n(n-2)\mu^{-2}]^{\frac{n-2}{4}}}{[\mu^{-2} + |x-x_0|^2]^{\frac{n-2}{2}}}, \ \mu > 0\Bigg) \tag{1.5}$$

show that the classical Palais-Smale condition (cp. [11]) cannot hold for E on $H_0^{1,2}(\Omega)$, cp. [5, Remark 2.3].

However, it has been observed that Brezis' and Nirenberg's result (and previous related existence results for Yamabe's problem by Aubin, cp. [1]) is based on a *local* Palais-Smale type condition:

Let $|u|_p^p = \int_\Omega |u|^p \, dx$ for $p \in [1, \infty[$ and let

$$S = \inf\{\|u\|_1^2 \mid u \in H_0^{1,2}(\Omega), \ |u|_{2^*} = 1\} \tag{1.6}$$

be the (universal) best constant for the embedding $H_0^{1,2}(\Omega) \to L^{2^*}(\Omega)$.

Also denote $(H_0^{1,2}(\Omega))^* = H^{-1}(\Omega)$.

Proposition 1.2. *Suppose for some sequence* $\{u_m\} \subset H_0^{1,2}(\Omega)$: $E(u_m) \leq c < \frac{1}{n}S^{n/2}$, $dE(u_m) \to 0$ *strongly in* $H^{-1}(\Omega)$[1]. *Then* $\{u_m\}$ *contains a strongly convergent subsequence.*

(Cp. [4, 5].) Using Proposition 1.2, in [5] Theorem 1.1 was generalized to obtain improved bounds for the existence of nontrivial solutions of (1.1) near arbitrary eigenvalues λ_k of $-\Delta$: $H_0^{1,2}(\Omega) \to H^{-1}(\Omega)$, extending well-known bifurcation results of Böhme [2] and Marino [10].

[5, Remark 2.3] only provides a very crude picture of how the Palais-Smale condition may fail for E. Below we investigate this phenomenon more closely. Our Proposition 2.1 essentially states that apart from "jumps" of the topological type of admissible functions compactness is preserved globally. A uniqueness result for the family (1.5) of solutions to (1.4) hence would imply various extensions of the existence results for (1.1) cited above.

In [15] a result similar to Proposition 2.1 is applied to the Plateau problem for surfaces of constant mean curvature. An analogous compactness result has independently been obtained by Brezis and Coron [3] for the corresponding Dirichlet problem.

Our proofs are mainly based on rescaling arguments. Such methods have repeatedly been used to extract convergent subsequences from families of

---

1 dE denotes the Fréchet derivative of E

(smooth) solutions or minimizing sequences to nonlinear variational problems (cp. [6, 9, 14, 16]).

Acknowledgement. I am indebted to P.-L. Lions for helpful remarks concerning a preliminary version of this note.

2. "Global Compactness"

Let $E^*: H_0^{1,2}(\mathbb{R}^n) \to \mathbb{R}$ given by

$$E^*(u) = \tfrac{1}{2} \int_{\mathbb{R}^n} |\nabla u|^2 \, dx - \tfrac{1}{2^*} \int_{\mathbb{R}^n} |u|^{2^*} dx$$

denote the energy functional corresponding to the limiting problem (1.4).

Proposition 2.1. *Let $n \geqq 3$, $\lambda \in \mathbb{R}$. Suppose $\{u_m\} \subset H_0^{1,2}(\Omega)$ satisfies:*

$$E(u_m) \leqq c, \quad dE(u_m) \to 0 \quad \text{strongly in } H^{-1}(\Omega) \ (m \to \infty).$$

Then there exists a number $k \in \mathbb{N}_0$, a solution u^0 of (1.1), solutions $u^1, \ldots, u^k$ of (1.4), sequences of points $x_m^1, \ldots, x_m^k \in \mathbb{R}^n$ and radii $r_m^1, \ldots, r_m^k > 0$ such that for some subsequence $m \to \infty$

$$\begin{aligned} &u_m^0 \equiv u_m \to u^0 && \text{weakly in } H_0^{1,2}(\Omega), \\ &u_m^j \equiv (u_m^{j-1} - u^{j-1})_{r_m^j, x_m^j} \to u^j && \text{weakly in } H_0^{1,2}(\mathbb{R}^n), \quad j = 1, \ldots, k^{2}, \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} &\|u_m\|_1^2 \to \sum_{j=0}^{k} \|u^j\|_1^2, \\ &E(u_m) \to E(u^0) + \sum_{j=1}^{k} E^*(u^j). \end{aligned} \tag{2.2}$$

Proof. It is easy to show that $\{u_m\}$ is bounded in $H_0^{1,2}(\Omega)$, cp. [4, estimates (2.18) ff.], [5, Lemma 2.1]. Hence we may assume that $u_m \to u^0$ weakly in $H_0^{1,2}(\Omega)$ as $m \to \infty$. By weak continuity of dE u^0 solves (1.1).

Let $\psi_m^1 = u_m - u^0$. Then $\psi_m^1 \to 0$ weakly and thus (again cp. [5, Lemma 2.1])

$$\begin{aligned} \|\psi_m^1\|_1^2 &= \|u_m\|_1^2 - \|u^0\|_1^2 + o(1)^{3}, \\ E^*(\psi_m^1) &= E(\psi_m^1) + o(1) = E(u_m) - E(u^0) + o(1), \\ dE^*(\psi_m^1) &= dE(\psi_m^1) + o(1) = dE(u_m) - dE(u^0) + o(1) \\ &= o(1) \in H^{-1}(\Omega), \end{aligned} \tag{2.3}$$

as Taylor's expansions show.

Suppose $\|\psi_m^1\|_1 \not\to 0$. Then we claim there exist sequences $\{r_m^1\}$ of positive radii tending to zero, $\{x_m^1\}$ of points in $\mathbb{R}^n$ such that

[2] Recall definition (1.3)

[3] Let $o(1)$ denote any quantities that tend to zero as $m \to \infty$

$$u_m^1=(\psi_m^1)_{r_m^1, x_m^1}\to u^1\not\equiv 0 \quad \text{weakly in } H_0^{1,2}(\mathbb{R}^n)^4 \tag{2.4}$$

as $m\to\infty$. By weak continuity again u^1 weakly solves the equation $-\Delta u = u|u|^{2^*-2}$ on its domain D^1.

Since $r_m^1\to 0$, either $D^1=\mathbb{R}^n$ and u^1 is a solution of (1.4), or D^1 is a half-space with $u^1|\partial D^1\equiv 0$. But in the latter case from Pohožaev's result [12] we infer that $u^1\equiv 0$. Hence u^1 solves (1.4).

Iterating, we obtain sequences $\psi_m^j=u_m^{j-1}-u^{j-1}$, $j\geqq 2$, and corresponding sequences of rescaled functions $u_m^j=(\psi_m^j)_{r_m^j, x_m^j}\to u^j$ weakly in $H_0^{1,2}(\mathbb{R}^n)$, where each u^j solves (1.4). Moreover, by induction from (2.3)

$$\begin{aligned}\|u_m^j\|_1^2 &= \|\psi_m^j\|_1^2=\|u_m^{j-1}\|_1^2-\|u^{j-1}\|_1^2+o(1)\\ &=\ldots=\|u_m\|_1^2-\sum_{i=0}^{j-1}\|u^i\|_1^2+o(1)\end{aligned} \tag{2.5a}$$

and

$$\begin{aligned}E^*(u_m^j) &= E^*(\psi_m^j)=E^*(u_m^{j-1})-E^*(u^{j-1})+o(1)\\ &=\ldots=E(u_m)-E(u^0)-\sum_{i=1}^{j-1}E^*(u^i)+o(1).\end{aligned} \tag{2.5b}$$

By the estimate[5]

$$\begin{aligned}0 &= \langle dE^*(u^j), u^j\rangle_{H_0^{1,2}}=\|u^j\|_1^2-|u^j|_{2^*}^{2^*}\\ &\geqq \|u^j\|_1^2\left(1-S^{-\frac{n}{n-2}}\|u^j\|_1^{\frac{4}{n-2}}\right)\end{aligned}$$

finally, each $\|u^j\|_1^2\geqq S^{n/2}$ and by (2.5a) the iteration must terminate at some index $k\geqq 0$.

It thus remains to establish (2.4).

First choose (tentative) radii r_m^1 and points x_m^1 such that for our (tentative) $u_m^1 =(\psi_m^1)_{r_m^1, x_m^1}$

$$\int_{\mathbb{R}^n}\frac{x}{|x|}|\nabla u_m^1|^2\,dx=0. \tag{2.6}$$

Moreover we may achieve that for some $\alpha\in\,]0, S^{n/2}[$

$$\int_{B_1(0)}|\nabla u_m^1|^2\,dx\to\alpha \quad (m\to\infty). \tag{2.7}$$

Suppose that $u_m^1\to 0$ weakly. (Otherwise we are done.) Let $A_1=\sup\|u_m^1\|_1^2<\infty$. Note that by Fubini's theorem the mapping

$$r\to\int_{\partial B_r(0)}|\nabla u_m^1|^2\,do\in L^1(\mathbb{R}_+) \quad \text{for any } m.$$

In particular for any m $u_m^1\in H^{1,2}(\partial B_r(0))$ for a.e. $r>0$. Moreover, given $\delta\in\,]0,1[$ by the estimate

$$\int_{1-\delta}^{1+\delta}\int_{\partial B_r(0)}|\nabla u_m^1|^2\,do\,dr\leqq A_1$$

4 Extending by zero outside $\tilde{\Omega}$ we may regard $H_0^{1,2}(\tilde{\Omega})\subset H_0^{1,2}(\mathbb{R}^n)$, for any $\tilde{\Omega}\subset\mathbb{R}^n$.

5 $\langle\cdot,\cdot\rangle_{H_0^{1,2}}$ denotes the dual pairing in $H_0^{1,2}(\mathbb{R}^n)$

for any m there is a set Λ_m of radii $\Lambda_m \subset [1-\delta, 1+\delta]$ of measure $|\Lambda_m| \geqq 3\delta/2$ such that

$$\int_{\partial B_r(0)} |\nabla u_m^1|^2 \, do \leqq 2\delta^{-1} A_1 = c_1$$

for all $r \in \Lambda_m$. By a selection process [7] and possibly passing to a subsequence again, we may hence choose $\underline{r}, \bar{r}$ satisfying $1-\delta \leqq \underline{r} \leqq 1 \leqq \bar{r} \leqq 1+\delta$ and such that

$$\int_{\partial B_{\underline{r}}(0)} |\nabla u_m^1|^2 \, do, \qquad \int_{\partial B_{\bar{r}}(0)} |\nabla u_m^1|^2 \, do \leqq c_1 \tag{2.8}$$

uniformly in m. Compactness of the embeddings

$$H_0^{1,2}(\mathbb{R}^n) \to L^2(\partial B_1(0)),$$
$$H^{1,2}(\partial B_1(0)) \to H^{1/2,2}(\partial B_1(0))\,[6]$$

now implies that as $m \to \infty$

$$u_m^1|_{\partial B_{\underline{r}}(0)}, \; u_m^1|_{\partial B_{\bar{r}}(0)} \to 0 \quad \text{strongly in } H^{1/2,2}. \tag{2.9}$$

Suppose $\delta \leqq 1/2$. We extend $u_m^1|_{B_{\underline{r}}(0)}$, $u_m^1|_{B_{\bar{r}}(0)}$ to functions $v_m^1, w_m^1 \in H_0^{1,2}(\mathbb{R}^n)$ as follows:

For $r \in]0,2[$, $u \in H^{1/2,2}(\partial B_r(0))$ let $v = \eta_r(u) \in H^{1,2}(\mathbb{R}^n \setminus B_r(0))$ be that function which vanishes outside $B_2(0)$, coincides with u on $\partial B_r(0)$, and which solves $-\Delta v = 0$ in the annulus $B_2(0) \setminus B_r(0)$. Note that η_r: $H^{1/2,2}(\partial B_r(0)) \to H^{1,2}(\mathbb{R}^n \setminus B_r(0))$ is continuous. Now define v_m^1 to be the extension of $u_m^1|_{B_{\underline{r}}(0)}$ by $\eta_{\underline{r}}(u_m^1|_{\partial B_{\underline{r}}(0)})$, and similarly define w_m^1 with $\bar{r}$ replacing $\underline{r}$.

By continuity of $\eta_{\underline{r}}, \eta_{\bar{r}}$ and (2.9):

$$\|v_m^1\|_1^2 = \int_{B_{\underline{r}}(0)} |\nabla u_m^1|^2 \, do + o(1),$$
$$\|w_m^1\|_1^2 = \int_{B_{\bar{r}}(0)} |\nabla u_m^1|^2 \, do + o(1). \tag{2.10}$$

Similarly we verify that e.g.

$$o(1) = \langle dE^*(u_m^1), v_m^1 \rangle_{H_0^{1,2}} = \|v_m^1\|_1^2 - |v_m|_{2^*}^{2^*} + o(1)$$
$$\geqq \|v_m^1\|_1^2 \left(1 - S^{-\frac{n}{n-2}} \|v_m^1\|_1^{\frac{4}{n-2}}\right) + o(1), \tag{2.11}$$

and an analogous estimate for w_m^1.

From (2.7), (2.10), (2.11) we thus conclude that $\|v_m^1\|_1^2 = o(1)$ while $\|w_m^1\|_1^2 \geqq S^{n/2} + o(1)$, i.e.

$$A_1 \geqq \int_{1-\delta}^{1+\delta} \int_{\partial B_r(0)} |\nabla u_m^1|^2 \, do \, dr \geqq S^{n/2} + o(1) \tag{2.12}$$

for any $\delta > 0$.

[6] $H^{1/2,2}(\partial B_1(0))$ for our purposes may be identified with the space of "traces" $\{u_{|\partial B_1(0)} | u \in H_0^{1,2}(\mathbb{R}^n)\}$

Likewise we may infer that either

$$\int_{\mathbb{R}^n\setminus B_3(0)} |\nabla u_m^1|^2\,dx = o(1)$$

or

$$\int_{\mathbb{R}^n\setminus B_2(0)} |\nabla u_m^1|^2\,dx \geqq S^{n/2} + o(1).$$

In the first case (taking account of (2.6)) for any ball B of radius $\frac{1}{2}$ centered on $\partial B_1(0)$

$$\int_B |\nabla u_m^1|^2\,dx \leqq \mu A_1 + o(1)$$

with a uniform constant $\mu < 1$ depending only on the dimension n. In the second case trivially

$$\int_B |\nabla u_m^1|^2\,dx \leqq A_1 - S^{n/2} + o(1),$$

whence in any event we obtain that

$$\limsup_{m\to\infty} \int_B |\nabla u_m^1|^2\,dx \leqq A_2 := \max\{\mu A_1, A_1 - S^{n/2}\} \tag{2.13}$$

for any ball B of radius $\frac{1}{2}$ centered on $\partial B_1(0)$.

Choosing $\delta \leqq 1/16$ we may cover the annulus $B_{1+\delta}(0)\setminus B_{1-\delta}(0)$ with a finite collection of balls $B_r(x)$ of radii $r \in [2\delta, 3\delta]$, and such that on the boundary of each of these similar to (2.8) (a subsequence: $\{u_m^1\}$ of) the sequence $\{u_m^1\}$ is equibounded in $H^{1,2}(\partial B_r(x))$ and hence accumulates at 0 with respect to the $H^{1/2,2}(\partial B_r(0))$-norm. Fix a ball $B_r(x)$ where

$$\limsup_{m\to\infty} \int_{B_r(x)} |\nabla u_m^1|^2\,dx > 0.$$

By (2.12) this is possible. Via η_r as above we may define an extension U_m^2 of $u_m^1|_{B_r(x)}$ satisfying the analogues of estimates (2.10), (2.11). Moreover, by (2.13) $\|U_m^2\|_1^2 \leqq A_2 + o(1)$. Now translate x_m^1 and rescale such that (2.6), (2.7) hold for U_m^2 with some $\alpha \in]0, S^{n/2}[$. Iterating the above argument with U_m^2 replacing $u_m^1 =: U_m^1$ and continuing in this manner in the i-th step by (2.11), (2.13) we obtain with $A_i = \max\{\mu A_{i-1}, A_{i-1} - S^{n/2}\}$ and some ball B that

$$\begin{aligned} S^{n/2} + o(1) &\leqq \int_{\mathbb{R}^n} |\nabla U_m^i|^2\,dx \leqq \int_B |\nabla U_m^{i-1}|^2\,dx + o(1) \\ &\leqq A_i + o(1). \end{aligned}$$

But $A_i \to 0$ $(i\to\infty)$. Therefore the iteration stops at a suitably rescaled sequence u_m^1 satisfying (2.4). *qed*

Remarks 2.2. i) Let

$$\sigma = \{E(u) \mid u \in H_0^{1,2}(\Omega) \text{ solves (1.1)}\},$$

$$\sigma^* = \{E^*(u) \mid u \in H_0^{1,2}(\mathbb{R}^n) \text{ solves (1.4)}\}$$

be the "spectra" of (1.1), (1.4) resp. Then in particular, Proposition 2.1 implies that any sequence $\{u_m\}$ satisfying $E(u_m)\to\beta$, $dE(u_m)\to 0$ in $H^{-1}(\Omega)$ $(m\to\infty)$ is strongly relatively compact in $H_0^{1,2}(\Omega)$, provided

$$\beta\notin\left\{\beta_0+\sum_{i=1}^{k}\beta_i^*\,\middle|\,\beta_0\in\sigma\backslash\{\beta\},\ \beta_i^*\in\sigma^*\backslash\{0\}\right\}.$$

ii) A uniqueness result (modulo translation and scaling) for the family (1.5) of non-trivial solutions of (1.4) therefore would imply various generalizations of Theorem 1.1. For example the following result could be obtained: For $n\geqq 4$, and any $l\in\mathbb{N}$ there exists a number $\lambda_l^*>0$ such that for any $\lambda\in]0,\lambda_l^*[$ problem (1.1) admits at least l pairs of solutions. We suppress the lengthy details.

References

1. Aubin, T.: Nonlinear analysis on manifolds, Monge-Ampère equations. Grundlehren **252**. Berlin Heidelberg New York: Springer 1982
2. Böhme, R.: Die Lösung der Variationsgleichungen für nichtlineare Eigenwertprobleme. Math. Z. **123**, 105–126 (1972)
3. Brezis, H., Coron, J.-M.: Convergence de solutions de H-systèmes et applications aux surfaces à courbure moyenne constante, preprint
4. Brezis, H., Nirenberg, L.: Positive solutions of nonlinear elliptic equations involving critical Sobolev exponents. Comm. Pure Appl. Math. **36** (1983)
5. Cerami, G., Fortunato, D., Struwe, M.: Bifurcation and multiplicity results for nonlinear elliptic problems involving critical Sobolev exponents, preprint
6. Donaldson, S.K.: An application of gauge theory to four dimensional topology. J. Diff. Geom. **18**, 279–315 (1983)
7. Eisen, G.: A selection lemma for measurable sets, and lower semicontinuity of multiple integrals. Manuscripta Math. **27**, 73–79 (1979)
8. Gidas, B., Ni, W.-M., Nirenberg, L.: Symmetry and related properties via the maximum principle. Comm. Math. Phys. **68**, 209–243 (1979)
9. Lions, P.-L.: Applications de la méthode de concentration-compacité á l'existence de fonctions extrémales. C.R. Acad. Sc. Paris **296**, 645–648 (1983)
10. Marino, A.: La biforcazione nel caso variazionale. Confer. Sem. Nat. Univ. Bari, **132** (1973)
11. Palais, R.S.: Morse theory on Hilbert manifolds. Topology **2**, 299–340 (1963)
12. Pohožaev, S.I.: Eigenfunctions of the equation $\Delta u+\lambda f(u)=0$. Doklady **165**, 1408–1412 (1965)
13. Rabinowitz, P.H.: Some aspects of nonlinear eigenvalue problems. Rocky Mount. J. Math. **3**, 161–202 (1973)
14. Sacks, P.-K. Uhlenbeck: On the existence of minimal immersions of 2-spheres. Ann. Math. **113**, 1–24 (1981)
15. Struwe, M.: Large H-surfaces via the mountain-pass-lemma. preprint
16. Wente, H.C.: Large solutions to the volume constrained Plateau problem. Arch. Rational Mech. Anal. **75**, 59–77 (1980)

Received January 31, 1984; in revised form May 2, 1984

Math. Z. 187, 519–525 (1984)

Mathematische Zeitschrift

Properties of Some Extremal Nonvanishing Univalent Functions

Johnny E. Brown

Department of Math., Purdue Univ., West Lafayette, Indiana 47907, USA

We let S_0 denote the class of nonvanishing univalent functions f in the unit disk D normalized so that $f(0)=1$ and we let S be those functions h univalent in D with $h(0)=0$ and $h'(0)=1$. There are simple transformations relating these classes. When investigating relationships between functions in S and their omitted values it is sometimes useful to consider the class S_0. This is an advantage since in many ways the class S_0 is much easier to handle. Principally, the existence of several transformations which preserve S_0 and the simple fact that zero is an omitted value account for this fact (see [3], for example).

Let J be a continuous complex-valued linear functional defined on the space of analytic functions in D with the topology of uniform convergence on compact subsets of D. If J is nonconstant on S_0 and $f \in S_0$ solves the extremal problem $\max\limits_{g \in S_0} \operatorname{Re} J(g)$, then $\Gamma = \mathbb{C} - f(D)$ is a single analytic arc extending from zero to infinity which satisfies a certain differential equation [3]. Such a function is called a *support point* of S_0 and has several interesting geometric properties that are close analogues of those for the support points of S. It was reported in [3, Theorem 8] that the omitted arc Γ for a function $f \in S_0$ minimizing $\operatorname{Re}\{g(\zeta)\}$, $0<\zeta<1$, over S_0, satisfies some additional geometric properties. It was claimed that $\arg\{w-f(\zeta)\}$ is monotonic for all $w \in \Gamma$ and Γ lies in a certain sector contained in the lower half-plane. Using this result, they were able to deform the path of a certain integral which was crucial in establishing the fact that for some $0<\zeta_0<1$, the function minimizing $\operatorname{Re}\{g(\zeta_0)\}$ maps D onto the complement of a certain half-ray. Unfortunately their proof of Theorem 8 contains an error. Their main results however all remain intact (as pointed out in [4]) and only the additional geometric properties of the extremal functions are in doubt. In this note we consider functions in S_0 which minimize $\operatorname{Re}\{e^{i\alpha} g(\zeta)\}$, $0 \leqq \alpha < \pi$, $0<\zeta<1$. The case $\alpha=\pi$ is omitted since the extremal function is clearly $f(z)=K_0(-z)$, where $K_0(z)=\left(\frac{1-z}{1+z}\right)^2$, for all $0<\zeta<1$. We prove that the extremal functions have omitted arcs Γ which lie in certain specific sectors and we also show that $\arg w$ is monotonic for

$w \in \Gamma \setminus \{0\}$. In the special case $\alpha=0$, we show that Γ indeed lies in a sector contained in the lower half-plane. Other results generalizing some results in [3] and [4] are also given. The results proved here may suggest properties of functions extremal for other problems in S_0. It should be pointed out that Hamilton [5] has determined the region of values $W(\zeta)=\{g(\zeta)\colon g \in S_0\}$. However no geometric properties of those functions corresponding to boundary points of $W(\zeta)$ are described.

1. Preliminaries

Let $f \in S_0$ be an extremal function for the problem

$$\min_{g \in S_0} \operatorname{Re}\{e^{i\alpha} g(\zeta)\}, \qquad 0 \leqq \alpha < \pi, \quad 0<\zeta<1. \tag{1}$$

Let $B=f(\zeta)$. Since the class S_0 is preserved under conjugation we have $\overline{f(\bar{z})} \in S_0$ and hence $\operatorname{Re}\{e^{i\alpha}B\} - \operatorname{Re}\{e^{i\alpha}\bar{B}\} \leqq 0$. Form this and the assumption $0<\alpha<\pi$ we see that $\operatorname{Im} B \geqq 0$. For the case $\alpha=0$ it is clear that $\operatorname{Im} B \geqq 0$ can be achieved again simply because $f(z)$ and $\overline{f(\bar{z})}$ belong to S_0. Moreover, by considering the function $g(z)=1-ze^{-i\alpha}$, it is clear that $f(z) \neq 1$ (cf. [4]). In [3] it is shown that $\Gamma=\mathbb{C}-f(D)$ is a single analytic arc extending from zero to infinity which satisfies the differential equation

$$\frac{e^{i\alpha}B(B-1)}{w(w-1)(w-B)}\, dw^2 > 0. \tag{2}$$

Thus Γ is a trajectory of the quadratic differential $\psi(w) \dfrac{dw^2}{w^2}$ where

$$\psi(w) = \frac{w e^{i\alpha} B(B-1)}{(w-1)(w-B)}. \tag{3}$$

Since $B(B-1) \neq 0$, the quadratic differential $\psi(w) \dfrac{dw^2}{w^2}$ has simple poles at $w=0, 1, B$ and infinity. A calculation then shows that Γ is asymptotic to the lines

$$w_0(t) = t e^{-i\alpha}(\bar{B}-1), \qquad 0<t<\infty, \tag{4}$$

and

$$w_\infty(\tau) = \tfrac{1}{3}(1+B) + \tau e^{i\alpha} B(B-1), \qquad 0<\tau<\infty, \tag{5}$$

at zero and infinity, respectively. These will prove to be crucial later.

If $w \in \Gamma$ then $f^*(z)=(1-w)f(z)/(f(z)-w)$ belongs to S_0 and if $w \neq 0$ then it is not extremal for (1) (see [3]). Also, since $f(z) \neq 1$, f^* is not extremal if $w=0$. By the extremal nature of f, we have $\operatorname{Re}\{e^{i\alpha}(f^*(\zeta)-f(\zeta))\}>0$ and hence

$$\operatorname{Re}\left\{\frac{e^{i\alpha}B(B-1)}{w-B}\right\} > 0, \qquad w \in \Gamma. \tag{6}$$

This inequality has several consequences. Putting $w=0$ in (6) gives

$$\operatorname{Re}\{e^{i\alpha}(B-1)\}<0. \tag{7}$$

Now Γ is asymptotic to the line (4) at zero and in view of (7) we can conclude that $\Gamma\setminus\{0\}$ initially lies in $\operatorname{Re} w<0$. This will be useful later.

Suppose now that $w'\in\Gamma$ and $w'=Bt'$ for some $1<t'<\infty$. Then by (6) we see that $\operatorname{Re}\{e^{i\alpha}(B-1)\}>0$, contradicting (7). Next, we observe that for any $w\in\Gamma\setminus\{0\}$ we have $f^{**}(z)=(f(z)-w)/(1-w)\in S_0$. The extremality of f then gives $\operatorname{Re}\left\{\frac{w}{w-1}e^{i\alpha}(1-B)\right\}\geqq 0$. Thus if $w\in\Gamma\cap(0,1)$ then we must have $\operatorname{Re}\{e^{i\alpha}(1-B)\}\leqq 0$, contradicting (7). Hence since $1\notin\Gamma$, we have $\Gamma\cap(0,1]=\varnothing$. Finally, since the quadratic differential $\psi(w)\frac{dw^2}{w^2}$ has simple poles at 1 and at B and no zeros, there is a unique trajectory Γ_0 joining 1 to B and hence $\Gamma\cap\Gamma_0=\varnothing$. We have thus shown that $\Gamma\cap C^*=\varnothing$, where

$$C^*=\Gamma_0\cup(0,1]\cup\{Bt\colon\ 1\leqq t<\infty\}. \tag{8}$$

In particular, the arc Γ cannot wind completely around the origin.

The following lemma will be used several times (cf. [1], [2]).

Lemma 1. *Let Ω be a simply-connected region, $0\notin\Omega$, and let Ω be bounded by a connected arc γ of a trajectory of $Q(w)\,dw^2/w^2$ and a radial segment L (L may be infinite). Let $\gamma\cap L=\{w_1,w_2\}$. Suppose $Q(w)$ is analytic and nonzero in $\bar\Omega\setminus\{w_1,w_2\}$ and suppose that w_1,w_2 are at most simple poles of $Q(w)\,dw^2/w^2$. Then there exists $w^*\in L$, not an endpoint, such that* $\operatorname{Im}\{Q(w^*)\}=0$.

Proof. Let L: $w(t)=te^{i\nu}$, $0\leqq|w_1|\leqq t\leqq|w_2|\leqq\infty$. Since Q is analytic and nonzero on $\bar\Omega\setminus\{w_1,w_2\}$, with at most simple poles at these exceptional points, we can (by continuity) find another trajectory Γ^* near γ such that $\Gamma^*\cap L=\{w_1^*,w_2^*\}$. Let γ^* be the subarc of Γ^* in Ω from w_1^* to w_2^*, L^* the corresponding segment along L and Ω^* the simply-connected subregion of Ω bounded by γ^* and L^*. See Fig. 1.

By our hypotheses Q is analytic and nonzero on $\bar\Omega^*$ and hence we have

$$\int_{\partial\Omega^*}\sqrt{Q(w)}\frac{dw}{w}=0.$$

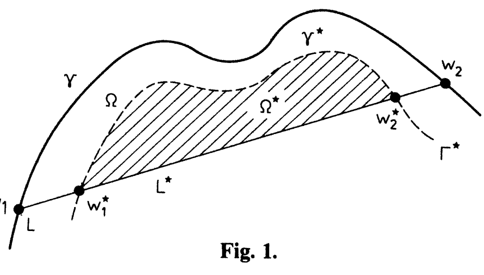


Fig. 1.

Because $\sqrt{Q(w)}\,\frac{dw}{w}$ is real along γ^*, we easily conclude that

$$0=\operatorname{Im}\int_{\partial\Omega^*}\sqrt{Q(w)}\,\frac{dw}{w}=\int_{w_1^*}^{w_2^*}\operatorname{Im}\sqrt{Q(w)}\,\frac{dw}{w}.$$

This last result implies that there must exist a point $w^*\in L$, not an endpoint, with $\operatorname{Im}Q(w^*)=0$. It is clear that the above proof is valid for Ω bounded or unbounded.

It is pointed out in [3] that if f is extremal for (1) when $\alpha=0$ then $f=K_0$ if and only if B is real and $0<B<1$. More generally, we have the following result.

Lemma 2. *If f is extremal for* (1) *and $B>0$ then $f=K_0$.*

Proof. First we assert that α must be zero. Recall that Γ_0 is the (unique) trajectory of $\psi(w)\,dw^2/w^2$ joining 1 to B. Let I denote the line segment along the real axis from 1 to B. We consider two cases. Suppose first that there is an arc $\gamma_0\subset\Gamma_0$ with $\gamma_0\subset I$. In this case we see from (2) that $\frac{e^{i\alpha}B(B-1)}{t(t-1)(t-B)}dt^2>0$ for t in some interval contained in I. In particular we see that $\alpha=0$.

Suppose now that there are no arcs $\gamma_0\subset\Gamma_0$ with $\gamma_0\subset I$. In this case Γ_0 meets I a finite number of times (since both are analytic curves). Let w' be the point of $\Gamma_0\cap I$ nearest the point 1 (possibly $w'=B$). Let Ω be the simply-connected region bounded by the arc γ of Γ_0 from w' to 1 and the segment from w' to 1. Since $B>0$ we can apply Lemma 1 to conclude that $\operatorname{Im}\psi(t')=0$ for some $0<t'<1$. It follows that $\alpha=0$. The assertion is proved.

Next, in view of (7) we see that $B<1$. As remarked above, if $\alpha=0$ and $0<B<1$ then $f(z)=K_0(z)$ and the lemma is proved.

Lemma 3. *Let $\psi(w)\,dw^2/w^2$ be the quadratic differential with ψ given by* (3). *Let L_θ be the ray $w=te^{i\theta}$, $0<t<\infty$ with $\theta\neq0$ or $\arg B$. If $\operatorname{Im}\psi(te^{i\theta})$ has at most n $(n\geqq1)$ zeros for $0<t<\infty$, then Γ meets the ray L_θ at most $(n-1)$ times.*

Proof. Assume $\Gamma\cap L_\theta=\{w_1,w_2,\dots,w_m\}$, $0<|w_1|<|w_2|<\dots<|w_m|<\infty$ with $m\geqq n$. (This set is finite since $\operatorname{Im}\psi(w)\not\equiv0$ on L_θ and ψ is rational.) From these m distinct points we can obtain $m+1$ disjoint simply-connected regions $\Omega_1,\dots,\Omega_{m+1}$ as follows. Let Ω_1 be the region bounded by the subarc γ_1 of Γ from 0 to w_1 and the segment along L_θ from 0 to w_1. Let Ω_2 be the region bounded by the subarc γ_2 of Γ from w_1 to w_2 and the segment along L_θ from w_1 to w_2. We obtain regions $\Omega_1,\dots,\Omega_m$ in this way. Finally, we let Ω_{m+1} be the unbounded region whose boundary consists of the subarc γ_{m+1} of Γ from w_m to infinity and the infinite segment along L_θ from w_m to infinity. Now because Γ does not meet the curve C^* given by (8) and $\theta\neq0$ $\arg B$, we are safely away from the poles 0, 1 and B and hence are in a position to apply Lemma 1 (with $Q=\psi$) to each of the $m+1$ regions Ω_k. Hence we conclude that $\operatorname{Im}\psi(te^{i\theta})$ vanishes at least $m+1$ times. This is a contradiction and the lemma is proved.

2. Main results

For $B=|B|e^{i\theta_0}$, $0<\theta_0\leqq\pi$, we define the sectors $\mathscr{S}_1$, $\mathscr{S}_2$ and $\mathscr{S}_3$ as follows:

$$\begin{aligned}\mathscr{S}_1&=\{\rho e^{i\theta}\colon\ 0\leqq\rho<\infty,\ \theta_0-\pi<\theta<\theta_0\}\\ \mathscr{S}_2&=\{\rho e^{i\theta}\colon\ 0\leqq\rho<\infty,\ \theta_0<\theta<\pi\}\\ \mathscr{S}_3&=\{\rho e^{i\theta}\colon\ 0\leqq\rho<\infty,\ \pi<\theta<\theta_0+\pi\}.\end{aligned} \tag{9}$$

If $\theta_0=\pi$ then $\mathscr{S}_2=\varnothing$. We can now state and prove our main results.

Theorem 1. *Let $f\in S_0$ be extremal for $\min_{g\in S_0}\operatorname{Re}\{e^{i\alpha}g(\zeta)\}$, $0\leqq\alpha<\pi$, $0<\zeta<1$ and let $\Gamma=\mathbb{C}-f(D)$. Suppose that $B=f(\zeta)$ satisfies $\operatorname{Im}B\geqq 0$. If $f\neq K_0$ then Γ lies entirely in one of the sectors $\mathscr{S}_1$, $\mathscr{S}_2$ or $\mathscr{S}_3$. In particular, Γ lies in a half-plane.*

Proof. Suppose $f\neq K_0$. Then by Lemma 2 we must have $\operatorname{Im}B>0$ or $B<0$. We have already shown that $\Gamma\cap\{Bt\colon 1\leqq t<\infty\}=\varnothing$. Hence it is enough to show that Γ does not meet any of the lines

$$w=-t,\qquad 0<t<\infty, \tag{10}$$

$$w=Bt,\qquad 0<t<1, \tag{11}$$

or

$$w=-Bt,\qquad 0<t<\infty. \tag{12}$$

Observe from (3) that

$$\operatorname{Im}\psi(-t)=\frac{-t}{(1+t)|t+B|^2}\operatorname{Im}\{e^{i\alpha}B(B-1)(t+\bar{B})\} \tag{13}$$

and

$$\operatorname{Im}\psi(-Bt)=\frac{-t}{(1+t)|Bt+1|^2}\operatorname{Im}\{e^{i\alpha}B(B-1)(t\bar{B}+1)\}. \tag{14}$$

Thus $\operatorname{Im}\psi(-t)$ vanishes at most once in $0<t<\infty$ or is identically zero. The latter case cannot arise as it implies that B is real $\alpha=0$. Thus we must have $B<0$. However, if $B<0$ then the segment from 0 to B is a trajectory of $\psi(w)\,dw^2/w^2$, which is impossible. Hence $\operatorname{Im}\psi(-t)$ vanishes at most once in $0<t<\infty$. Now since L: $w=-t$, $0<t<\infty$, is a line avoiding 1 and B we can apply Lemma 3 to conclude that Γ does not meet (10). By a similar argument using $\operatorname{Im}\psi(-Bt)$ and (14) we see that Γ does not meet (12).

Assume now that Γ meets the segment (11) at the points $\{\zeta_1,\ldots,\zeta_n\}$, $0<|\zeta_1|<\ldots<|\zeta_n|<|B|$. If $n>1$; we let Ω_1 be the simply-connected region bounded by the arc $\gamma_1\subset\Gamma$ from 0 to ζ_1 and the segment along (11) from 0 to ζ_1. Also, let Ω_2 be the simply-connected region bounded by the arc $\gamma_2\subset\Gamma$ from ζ_1 to ζ_2 and the segment from ζ_1 to ζ_2. Since $\Gamma\cap C^*=\varnothing$ we can apply Lemma 1 to Ω_1 and Ω_2 to conclude that $\operatorname{Im}\psi(Bt)=0$ for at least two distinct values of $t\in(0,1)$. This is impossible as (14) shows. Hence we must have $n=1$. Let Ω_1 be as above and let Ω_∞ be the unbounded simply-connected region bounded by Γ and the ray $w=Bt$, $-\infty<t<0$. Invoking Lemma 1 gives $\operatorname{Im}\psi(Bt)=0$ for at least two distinct values of $t\in(-\infty,1)$ $(t\neq 0)$. Again we arrive at a contradiction. Hence Γ cannot meet (11). The proof of the theorem is complete.

Theorem 1 restricts the regions in which the omitted arc Γ can lie. We should point out that (6) implies that Γ also lies in a certain half-plane through the point B defined by

$$\operatorname{Re}\left\{\frac{w-B}{e^{i\alpha}B(B-1)}\right\}>0.$$

When $\alpha=0$ we can further restrict the accessible region for Γ:

Corollary 1. *Let $f\in S_0$ be extremal for $\min_{g\in S_0} \operatorname{Re}\{g(\zeta)\}$, $0<\zeta<1$, and suppose $\operatorname{Im} B\geqq 0$. If $f\neq K_0$ then Γ lies entirely in $\mathscr{S}_3$ (given by (9)) or in $\mathscr{S}_4$ where*

$$\mathscr{S}_4=\{\rho e^{i\theta}\colon\ 0\leqq\rho<\infty,\ \theta_0+\pi<\theta<2\pi\}\quad(\theta_0=\operatorname{Arg} B).$$

Observe that both $\mathscr{S}_3$ and $\mathscr{S}_3$ lie in $\operatorname{Im} w\leqq 0$. In particular, the arc Γ does not meet the segment from 1 to B. This fact, which is proved in [4] under the restriction that $\operatorname{Re} B\geqq 0$, is used in the proofs of Theorems 7 and 9 in [3].

Proof. Clearly the assumption $\operatorname{Im} B\geqq 0$ does not restrict generality. If $f\neq K_0$ then by Lemma 2 we must have $\operatorname{Im} B>0$ or $B<0$. If $B<0$ then the segment from 0 to B is a trajectory of $\psi(w)\,dw^2/w^2$ and this is impossible. Hence $\operatorname{Im} B>0$ and from (4) and (6) we see that Γ initially lies in the third quadrant. From the theorem we can thus conclude that if $\operatorname{Re} B\leqq 0$ then Γ lies in $\mathscr{S}_3$ and if $\operatorname{Re} B>0$ then Γ lies in $\mathscr{S}_1$ or $\mathscr{S}_3$. It is enough to show that if $\operatorname{Re} B>0$ then Γ does not meet the positive real axis. Assume Γ meets the positive real axis. Let w' be the point of Γ on the positive real axis farthest from the origin and let Ω' be the simply-connected region bounded by the arc $\gamma'\subset\Gamma$ from w' to ∞ and the real axis from w' to ∞. From Lemma 1 we find that $\operatorname{Im}\psi(t')=0$ for some $1<t'<\infty$. This implies that $-t'\operatorname{Im}\{B(B-1)\}+|B|^2\operatorname{Im}\{B-1\}=0$. Now since $\operatorname{Im}\{B-1\}>0$ we see that $\operatorname{Im}\{B(B-1)\}>0$. Let $B(B-1)=|B(B-1)|\,e^{i\mu}$, $0<\mu<\pi$. From (5) we set that if $\frac{\pi}{2}\leqq\mu<\pi$ then Γ must eventually cross the line $w=Bt$, $-\infty<t<\infty$ or the negative real axis. Both are impossible. Hence we must have $0<\mu<\frac{\pi}{2}$ and since $B=|B|\,e^{i\theta_0}$ satisfies $0<\theta_0<\frac{\pi}{2}$, we see that $\operatorname{Re}\{e^{i(\mu-\theta_0)}\}>0$ which implies that $\operatorname{Re}\{B-1\}>0$. This contradicts (7). Hence Γ cannot meet the positive real axis. The proof is complete.

The second part of Theorem 8 in [3] is concerned with the monotonicity properties of the arc Γ. In this regard we have the following result:

Theorem 2. *If $f\in S_0$ is extremal for $\min_{g\in S_0} \operatorname{Re}\{e^{i\alpha}g(\zeta)\}$, $0<\zeta<1$, $0\leqq\alpha<\pi$, then $\Gamma\setminus\{0\}$ has monotonic argument.*

Proof. Consider the family of radial lines L_θ: $w=te^{i\theta}$, $0<t<\infty$, with $\theta\neq 0$ $\arg B$. It is then clear from (3) that

$$\operatorname{Im}\psi(te^{i\theta})=\frac{t\operatorname{Im}\{e^{i(\alpha+\theta)}B(B-1)(te^{-i\theta}-1)(te^{-i\theta}-\bar{B})\}}{|(te^{i\theta}-1)(te^{i\theta}-B)|^2}.$$

Since zero is a simple pole of $\psi(w)\,\frac{dw^2}{w^2}$, if $\operatorname{Im}\psi(te^{i\theta})\equiv 0$ for two distinct rays L_θ then one of them must be Γ and the other must be an orthogonal trajectory. If Γ is a radial line we are done. Hence may assume Γ is non-radial and so $\operatorname{Im}\psi(te^{i\theta_1})\equiv 0$ for at most one ray L_{θ_1}. Thus, except for $\theta=\theta_1+2\pi k$, $\operatorname{Im}\psi(te^{i\theta})$ vanishes at most twice in $0<t<\infty$. We are in a position to apply Lemma 3 to conclude that Γ meets each such L_θ at most once. This implies that $\Gamma\setminus\{0\}$ has monotonic argument and the proof of the theorem is complete.

References

1. Bombieri, E.: A geometric approach to some coefficient inequalities for univalent functions. Ann. Scuola Norm. Sup. Pisa (3) **22**, 377–397 (1968)
2. Charzynski, Z., Schiffer, M.: A geometric proof of the Bieberbach conjecture for the fourth coefficient. Scripta Math. **25**, 173–181 (1960)
3. Duren, P., Schober, G.: Nonvanishing univalent functions. Math. Z. **170**, 195–216 (1980)
4. Duren, P., Schober, G.: Nonvanishing univalent functions II. Preprint
5. Hamilton, D.: Nonzero univalent functions. Preprint.

Received October 15, 1982, in revised form February 20, 1983

Math. Z. 187, 527–547 (1984)

Mathematische Zeitschrift

The Poisson Kernel for Heisenberg Polynomials on the Disk

Charles F. Dunkl*

Department of Mathematics, University of Virginia, Charlottesville, Virginia 22903, U.S.A.

The study of harmonic polynomials on the Heisenberg groups leads to the functions $C_n^{(\alpha,\beta)}(z)$, a set of polynomials in z, $\bar{z}$, defined by $(1-r\bar{z})^{-\alpha}(1-rz)^{-\beta} = \sum_{n=0}^{\infty} r^n C_n^{(\alpha,\beta)}(z)$, for $|rz|<1$. There are existential arguments of Gaveau, [6] and Jerison [10] which can be used to show (Greiner and Koornwinder [8]) that the functions $\{\theta \mapsto C_n^{(\alpha,\beta)}(e^{i\theta}): n \geqq 0\}$ on $[0,\pi]$ span a (weighted sup-norm) dense set in the continuous functions when $\alpha+\beta$, $\alpha-\beta$ are integers and $\alpha,\beta>0$. However as yet there is no method for finding expansions of arbitrary functions. In this work we consider $\{C_n^{(\alpha,\beta)}(z)\}$ on the entire circle $T=\{e^{i\theta}: -\pi<\theta\leqq\pi\}$, being partly motivated by Gasper's [5] complex orthogonality there. To construct a basis for trigonometric polynomials on T we must double the set of functions (the prototype: $\{\cos n\theta: n\geqq 0\}$ suffices for a half circle, but $\{\sin n\theta: n\geqq 1\}$ must be added to get all of T), and so we adjoin $\{(z-\bar{z})\, C_{n-1}^{(\alpha+1,\beta+1)}(z): n\geqq 1\}$.

We obtain a biorthogonality structure for these functions in $L^2(T, |\sin\theta|^{\alpha+\beta}\, d\theta)$ and sum the resulting Poisson kernel by using the author's [2] expansion of the fundamental solution of the subelliptic Laplacian. Radial convergence of Poisson integrals holds for $C(T)$ and L^p, $1\leqq p<\infty$ when $|\alpha-\beta|\leqq 2$, and fails otherwise. Thus we are able to expand arbitrary continuous functions on T, and just as convergence theorems for cosine series follow from those for Fourier series, there is some possibility that this work will be helpful in the density problem for $\{C_n^{(\alpha,\beta)}(e^{i\theta})\}$ on $0\leqq\theta\leqq\pi$.

Here is a summary of the results:

$$\text{fix}\ \alpha,\beta>0, \quad \text{let}\ dw(\theta) := |\sin\theta|^{\alpha+\beta}\, d\theta \quad \text{on}\ (-\pi,\pi);$$

then there is an integral transform $P_r[f]$ defined for $f\in L^1(T, dw)$ such that:

i) $re^{i\theta} \to P_r[f](e^{i\theta}) = \sum_{n=0}^{\infty} \hat{f}_{+,n}\, r^n C_n^{(\alpha,\beta)}(e^{i\theta}) + 2i\sin\theta \sum_{n=1}^{\infty} \hat{f}_{-,n}\, r^n C_{n-1}^{(\alpha+1,\beta+1)}(e^{i\theta})$ is a smooth function on the open unit disk ($z=re^{i\theta}$, $|z|<1$);

---

* This research was partially supported by NSF grant MCS-8301271

ii) if f is a trigonometric polynomial, then $P_r[f](e^{i\theta})$ is a polynomial in $z, \bar{z}$ agreeing with f on T;

iii) if $|\alpha-\beta| \leqq 2$, $f \in L^p(dw)$, $1 \leqq p < \infty$ or $f \in C(T)$ and $p=\infty$, then $\|P_r[f] - f\|_p \to 0$ as $r \to 1_-$.

iv) if $|\alpha-\beta|>2$, and $\theta \neq 0, \pi$ then there exists $f \in C(T)$ such that $P_r[f](e^{i\theta})$ is divergent as $r \to 1_-$.

The author gratefully acknowledges the helpful suggestions of T.H. Koornwinder which were used to simplify the presentation of Sect. 2 and 3.

§ 1. Harmonic Polynomials on the Heisenberg Group

The motivation for the study of Heisenberg polynomials (we hereby propose this nomenclature) comes from the subelliptic Laplacian and related problems on the Heisenberg groups. These are the groups H_N, which consist of $\mathbb{C}^N \times \mathbb{R}$ with the group operation

$$(\zeta, s)\cdot(\xi, t) := (\zeta+\xi, s+t+2\operatorname{Im}\langle\zeta, \xi\rangle), (\langle\zeta, \xi\rangle := \sum_{j=1}^{N} \zeta_j \bar{\xi}_j, N=1,2,\ldots).$$

The left-invariant tangent fields are spanned by $Z_j := \frac{\partial}{\partial \xi_j} + i\bar{\xi}_j \frac{\partial}{\partial t}$, $\bar{Z}_j := \frac{\partial}{\partial \bar{\xi}_j} - i\xi_j \frac{\partial}{\partial t} (1 \leqq j \leqq N)$ and $\frac{\partial}{\partial t}$. We will be concerned with

$$L_{2\gamma} := -\frac{1}{2} \sum_{j=1}^{N} (Z_j \bar{Z}_j + \bar{Z}_j Z_j) + 2i\gamma \frac{\partial}{\partial t};$$

in particular L_0 is the sub-elliptic Laplacian (see Folland and Stein [4]). Greiner [7] initiated the study of $L_{2\gamma}$-harmonic polynomials (that is, $L_{2\gamma}\, p=0$). One stage of decomposition comes from the action of the unitary group $U(N)$ on H_N (namely, $u\cdot(\xi, t) := (u\xi, t)$) which commutes with $L_{2\gamma}$. We consider only $N \geqq 2$, since $U(1)$ is a special case, being abelian.

1.1. Definition. For $k, l \geqq 0$, let $V_{kl} = \left\{ p : p \text{ is a homogeneous polynomial on } \mathbb{C}^N \text{ of bidegree } (k,l), (p(c\xi) = c^k \bar{c}^l p(\xi), c \in \mathbb{C}), \text{ and } \sum_{j=1}^{N} \frac{\partial^2}{\partial \xi_j \partial \bar{\xi}_j} p = 0 \right\}$. This is an irreducible $U(N)$-module of dimension $\dfrac{k+l+N-1}{N-1} \dfrac{(N-1)_k (N-1)_l}{k!\,l!}$.

From the generating function stated at the beginning of the paper we have that

$$C_n^{(\alpha,\beta)}(z) = \sum_{j=0}^{n} \frac{(\alpha)_j (\beta)_{n-j}}{j!\,(n-j)!} \bar{z}^j z^{n-j}, \qquad n \geqq 0.$$

1.2. Proposition [2]. *Let $p \in V_{kl}$, then $L_{2\gamma}(p(\xi)\, C_n^{(\alpha+l, \beta+k)}(t+i|\xi|^2)) = 0$. Where $\alpha := \frac{N}{2} - \gamma$, $\beta := \frac{N}{2} + \gamma$, $n \geqq 0$; and every $L_{2\gamma}$-harmonic polynomial is a linear combination of such terms (all values of k, l, n).*

Korányi [11] also gave a proof of this.

The Dirichlet problem for the ball $B := \{(\xi, t): |\xi|^4 + t^2 \leqq 1\}$ in H_N is to find for any continuous function on ∂B a continuous extension to all of B which is $L_{2\gamma}$-harmonic in the interior.

Gaveau [6] showed that the Dirichlet problem could be solved for L_0 by a diffusion argument, and Jerison [10] later gave another proof. However these proofs give no indication how to find series $\sum_{n=0}^{\infty} a_n C_n^{(\alpha,\beta)}(z)$ which have specified boundary values on $\{z \in \mathbb{C}: |z|=1, \operatorname{Im} z \geqq 0\}$ where $z := t + i|\xi|^2$). The present work concerns a similar problem which possibly will be helpful in this one.

The basic idea comes from the fact that the two sets of functions $\{\cos n\theta: n \geqq 0\}$ and $\{\sin n\theta: n \geqq 1\}$ are each bases for $L^2(0, \pi)$ and together form a base for $L^2(T)$. This has already been generalized to Gegenbauer polynomials by Muckenhoupt and Stein [12], using the sets $\{C_n^\nu(\cos\theta): n \geqq 0\}$ and $\{\sin\theta\, C_{n-1}^{\nu+1}(\cos\theta): n \geqq 1\}$, which have similar properties for $L^2(T, |\sin\theta|^{2\nu} d\theta)$ (for $\nu > 0$, the Gegenbauer polynomial $C_n^\nu(\cos\theta) = C_n^{(\nu,\nu)}(e^{i\theta})$). Here we will adjoin the set $\{\sin\theta\, C_{n-1}^{(\alpha+1,\beta+1)}(e^{i\theta}): n \geqq 1\}$, whose relevance to the Heisenberg group will now be presented.

Consider the vector space of (0, 1)-forms $f(\xi, t) = \sum_{j=1}^{N} f_j(\xi, t)\, d\bar{\xi}_j$ on H_N with the $U(N)$ action $L(u) f(\xi, t) := \sum_{j,k=1}^{N} u_{jk} f_k(u^* \xi, t)\, d\bar{\xi}_j$, (where $u = (u_{jk}) \in U(N)$). The operator $L_{2\gamma}$ acts coefficient-wise (see Folland and Stein [4]). Further there are two $U(N)$-homomorphisms of functions to (0, 1)-forms, namely

$$\bar{\partial} f(\xi, t) := \sum_{j=1}^{N} \frac{\partial}{\partial \bar{\xi}_j} f(\xi, t)\, d\bar{\xi}_j \quad \text{and} \quad M f(\xi, t) := f(\xi, t) \sum_{j=1}^{N} \xi_j\, d\bar{\xi}_j.$$

From this we define $\bar{\partial}_b f := \sum_{j=1}^{N} (\bar{Z}_j f)\, d\bar{\xi}_j = \left(\bar{\partial} - iM \frac{\partial}{\partial t}\right) f$ (Folland and Stein [4].)

For fixed $k, l \geqq 0$, if $p \in V_{kl}$ then $\frac{\partial p}{\partial \bar{\xi}_j} \in V_{k,l-1}$ and $|\xi|^2 \frac{\partial p(\xi)}{\partial \bar{\xi}_j} - (N+k+l-1)\, \xi_j p(\xi) \in V_{k+1,l}$ for each j (for a detailed study of the $U(N)$ action on (p, q)-forms see Folland [3]). Define the $U(N)$-homomorphism T_{kl} on V_{kl} by $T_{kl} := |\xi|^2 \bar{\partial} - (N+k+l-1) M$.

Let $f(\xi, t) = g(t + i|\xi|^2)\, p(\xi)$, with $p \in V_{kl}$, then

$$\bar{\partial}_b f = \left(g(z) - \frac{z - \bar{z}}{N+k+l-1} \frac{\partial g}{\partial \bar{z}}\right) \bar{\partial} p + \frac{2i}{N+k+l-1} \frac{\partial g}{\partial \bar{z}} T_{kl} p,$$

where $z = t + i|\xi|^2$. Since $L_{2\gamma} \bar{\partial}_b = \bar{\partial}_b L_{2\gamma+2}$, we let $g(z) = C_n^{(\alpha+l-1, \beta+k+1)}(z)$ ($\alpha = N/2 - \gamma$, $\beta = N/2 + \gamma$), then the coefficients of $\bar{\partial} p$ and $T_{kl} p$ in $\bar{\partial}_b(gp)$ are $\frac{N+k+l+n-1}{N+k+l-1} C_n^{(\alpha+l-1, \beta+k)}(z)$, and $2i \frac{\alpha+l-1}{N+k+l-1} C_{n-1}^{(\alpha+l, \beta+k+1)}(z)$ respectively, by use of identities (8 ii), (10) from [2].

This leads to the definition of the conjugate series of the formal series

$\sum_{n=0}^{\infty} a_n C_n^{(\alpha,\beta)}$ to be $2i\alpha \sum_{n=1}^{\infty} \frac{a_n}{\alpha+\beta+n} C_{n-1}^{(\alpha+1,\beta+1)}$, (put $l=1$, $k=0$ in the previous discussion); for $\alpha=\beta$ this agrees with the definition of Muckenhoupt and Stein [12]. For functions $p \in V_{kl}$, a form $g_1(z)\bar{\partial}p(\xi)+g_2(z)T_{kl}p(\xi)$ is $L_{2\gamma}$-harmonic and in the range of $\bar{\partial}_b$ if and only if g_2 is the conjugate series of g_1, (if and only if it is annihilated by $\bar{\partial}_b$, where $\bar{\partial}_b \sum_{j=1}^{N} f_j d\bar{\xi}_j := \sum_{j<k} (\bar{Z}_j f_k - \bar{Z}_k f_j) d\bar{\xi}_j \wedge d\bar{\xi}_k$).

We have not lost the "$\sin\theta$" part of $\sin\theta C_{n-1}^{(\alpha+1,\beta+1)}(e^{i\theta})$ as it may appear. A reasonably general $U(N)$-invariant integral on the sphere ∂B is given by

$$I_w(f) := \int_0^{\pi} w(\theta)\, d\theta \int_S f(\xi \sin^{1/2}\theta, \cos\theta)\, d\omega(\xi)$$

where $w(\theta) \geqq 0$ on $(0,\pi)$ and $d\omega$ is the $U(N)$-invariant measure on the sphere $\{|\xi|=1\} \subset \mathbb{C}^N$. For (0, 1)-forms define the L^2-norm $N_w(f) := I_w\left(\sum_{j=1}^{N} |f_j|^2\right)$. For fixed $p \in V_{kl}$ and functions $g_1(z)$, $g_2(z)$ on $\{z \in \mathbb{C}: |z| \leqq 1, \operatorname{Im} z \geqq 0\}$,

$$N_w(g_1 \bar{\partial} p + g_2 T_{kl} p) = \int_0^{\pi} (l|g_1(e^{i\theta})|^2 + (N+k-1)|g_2(e^{i\theta})|^2 \sin^2\theta)$$
$$\cdot \sin^{k+l-1}\theta\, w(\theta)\, d\theta\, (N+k+l-1) \int_S |p(\xi)|^2\, d\omega(\xi).$$

(The derivation of this formula uses the fact that the components of $\bar{\partial}p$ and $T_{kl}p$ are in $V_{k,l-1}$ and $V_{k+1,l}$ respectively, making them orthogonal; further by Schur's lemma, $\bar{\partial}$ and T_{kl} are scalar multiples of isometries on V_{kl}, and the multipliers can be determined by computing with any convenient nonzero element of V_{kl}, e.g. $\xi_1^k \bar{\xi}_2^l$.)

Observe the factor $\sin^2\theta$ with $|g_2|^2$.

Further developments on conjugate series must wait for future work; for now this should provide adequate motivation to study the functions $C_n^{(\alpha,\beta)}(e^{i\theta})$, $\sin\theta C_{n-1}^{(\alpha+1,\beta+1)}(e^{i\theta})$ together, although the group provides a setting only for the values $0 \leqq \theta \leqq \pi$.

§ 2. Some Integrals and the Biorthogonal Set

We fix real parameters α, β with $\alpha, \beta \geqq 0$ and $\alpha+\beta>0$, and set $\nu=(\alpha+\beta)/2$, $\gamma=(\beta-\alpha)/2$. We first show that any polynomial in z, $\bar{z}$ restricted to the unit circle T has a unique expansion in terms of

$$\{C_n^{(\alpha,\beta)}(z): n \geqq 0\} \cup \{(z-\bar{z})\, C_{n-1}^{(\alpha+1,\beta+1)}(z): n \geqq 1\}.$$

It suffices to expand z^m and $\bar{z}^m$ for $m \geqq 0$ (since $z\bar{z}=1$ on T). Note that

$$C_n^{(\alpha,\beta)}(z) = \sum_{j=0}^{n} \frac{(\alpha)_j(\beta)_{n-j}}{j!(n-j)!} \bar{z}^j z^{n-j} \tag{2.1}$$

(2.2) $(z-\bar{z})\, C_{n-1}^{(\alpha+1,\beta+1)}(z) = \dfrac{1}{\alpha\beta} \sum_{j=0}^{n} \dfrac{(\alpha)_j (\beta)_{n-j}}{j!(n-j)!} \bar{z}^j z^{n-j} (\nu(n-2j)-n\gamma)$ for $\alpha\beta>0$.

(2.3) Let $D_{\alpha\beta} := (z-\bar{z}) \dfrac{\partial^2}{\partial z \partial \bar{z}} - \alpha \dfrac{\partial}{\partial z} + \beta \dfrac{\partial}{\partial \bar{z}}$, then

$$D_{\alpha\beta}\, C_n^{(\alpha,\beta)}(z) = 0 \quad \text{and} \quad D_{\alpha\beta} (z-\bar{z})\, C_{n-1}^{(\alpha+1,\beta+1)}(z)$$
$$= -(\alpha+\beta)\, C_{n-1}^{(\alpha+1,\beta+1)}(z), \quad \text{thus } D_{\alpha+1,\beta+1} \circ D_{\alpha\beta}$$

annihilates all these polynomials.

2.1. Definition. Let $h_0(\zeta) := 1$ and

$$h_n(z) := C_n^{(\alpha,\beta)}(z), \qquad h_{-n}(z) := (z-\bar{z})\, C_{n-1}^{(\alpha+1,\beta+1)}(z)$$

for $n=1,2,3\ldots$. Further define spaces of polynomial functions by $P_0 := \mathbb{C}1$ (constants) and $P_n := \operatorname{span}\{h_n, h_{-n}\}$ for $n \geqq 1$.

2.2. Proposition. *The space of function $\sum_{m=0}^{n} P_m$ restricted to T is exactly the space of trigonometric polynomials of degree $\leqq n$.*

Proof. The statement is obvious for $n=0$. Observe for $n \geqq 1$ that

$$h_n(e^{i\theta}) = ((\beta)_n e^{in\theta} + (\alpha)_n e^{-in\theta})/n! + g_1(e^{i\theta})$$

and

$$h_{-n}(e^{i\theta}) = ((\beta+1)_{n-1} e^{in\theta} - (\alpha+1)_{n-1} e^{-in\theta})/(n-1)! + g_2(e^{i\theta}),$$

where g_1 and g_2 are trigonometric polynomials of degree $\leqq n-1$. The determinant of the coefficients of $e^{in\theta}$ and $e^{-in\theta}$ in these equations is $(\alpha+\beta)(\alpha+1)_{n-1}(\beta+1)_{n-1}/(n!(n-1)!)$, which is nonzero in the specified (α,β) region. An inductive argument on n finishes the proof. □

We will use another basis for P_n. This idea is due to T.H. Koornwinder (verbal communication) and greatly simplifies the development of the rest of the section compared to its original version.

2.3. Proposition. *For* $n \geqq 1$,

(1) $$C_n^{(\alpha+1,\beta)}(z) = ((n+2\nu)\, h_n(z) - \beta h_{-n}(z))/2\nu;$$

(2) $$\bar{z}\, C_{n-1}^{(\alpha+1,\beta)}(z) = (n h_n(z) - \beta h_{-n}(z))/2\nu;$$

(3) $$h_n(z) = C_n^{(\alpha+1,\beta)}(z) - \bar{z}\, C_{n-1}^{(\alpha+1,\beta)}(z);$$

(4) $$\beta h_{-n}(z) = n\, C_n^{(\alpha+1,\beta)}(z) - (n+2\nu)\, \bar{z}\, C_{n-1}^{(\alpha+1,\beta)}(z).$$

Proof. (1) is equivalent to identity (11) in [2]. Formula (3) can be easily verified from the definition of $C_n^{(\alpha,\beta)}$. Formulas (2) and (4) follow from (1) and (3). □

2.4. Corollary. *If $\beta>0$ then*

$$P_n = \operatorname{span}\{C_n^{(\alpha+1,\beta)}(z),\ \bar{z}\, C_{n-1}^{(\alpha+1,\beta)}(z)\}.$$

By taking complex conjugates and interchanging α and β these formulas can be modified for $C_n^{(\alpha,\beta+1)}(z)$, $z\,C_{n-1}^{(\alpha,\beta+1)}(z)$. This basis would be used if $\alpha>0$ but $\beta=0$. So without loss of generality we will assume $\beta>0$.

From the work of Gasper [5], Greiner and Koornwinder [8] we know the right measure and complex weight function to use for biorthogonality results. First we introduce the (positive) normalized measure

$$dw_\nu(\theta):=(2B(\nu+\tfrac{1}{2},\tfrac{1}{2}))^{-1}|\sin\theta|^{2\nu}\,d\theta,\ -\pi<\theta<\pi;$$

and the complex weight function

$$c_{\alpha\beta}(\theta):=\frac{\Gamma(\alpha+1)\,\Gamma(\beta+1)}{\Gamma(\nu+1)^2}\,e^{(\alpha-\beta)i(\theta-\pi/2)}\qquad \text{on } 0<\theta<\pi$$

and $c_{\alpha\beta}(\theta):=c_{\alpha\beta}(\theta+\pi)$ on $-\pi<\theta<0$. (With this normalization $\int_{-\pi}^{\pi} c_{\alpha\beta}\,dw_\nu=1$.) The periodicity of $c_{\alpha\beta}$ and w_ν leads to the following simple but useful formula:

$$\int_{-\pi}^{\pi} f(\theta)\,\overline{c_{\alpha\beta}(\theta)}\,dw_\nu(\theta)=\int_0^{\pi}(f(\theta)+f(\theta-\pi))\,\overline{c_{\alpha\beta}(\theta)}\,dw_\nu(\theta) \tag{2.4}$$

for any function f integrable on $(-\pi,\pi)$. In particular, for a trigonometric polynomial

$$f(\theta)=\sum_m a_m e^{im\theta}\qquad \text{we have } \int_{-\pi}^{\pi} f\overline{c_{\alpha\beta}}\,dw_\nu$$

$$=2\int_0^{\pi}\Big(\sum_m a_{2m}e^{2mi\theta}\Big)\,\overline{c_{\alpha\beta}(\theta)}\,dw_\nu(\theta).$$

The integral underlying the biorthogonality is from Greiner and Koornwinder (Lemma 1.1, [8]), and is stated here in an equivalent form.

2.5. Lemma [8].

$$2\int_0^{\pi} C_n^{(\alpha+1,\beta)}(e^{i\theta})\,e^{mi\theta}\,\overline{c_{\alpha\beta}(\theta)}\,dw_\nu(\theta)=(2\nu+1)_n\left\{\frac{1}{(\beta+1)_n}\,\delta_{mn}-\frac{\beta}{(\alpha+1)_{n+1}}\,\delta_{m,-n-2}\right\}$$

for the values $m=-n-2,\ -n,\ -n+2,\dots,n$.

2.6. Theorem. *Let $n\geq 1$, $q\in P_n$ and let p be a trigonometric polynomial of degree $<n$, then*

$$\int_{-\pi}^{\pi} pq\overline{c_{\alpha\beta}}\,dw_\nu=0.$$

Proof. By Corollary 2.4 it suffices to consider $q(e^{i\theta})=C_n^{(\alpha+1,\beta)}(e^{i\theta})$ and $q(e^{i\theta})=e^{-i\theta}C_{n-1}^{(\alpha+1,\beta)}(e^{i\theta})$. In both cases q has the form $\sum_{m=-n}^{n} a_m e^{mi\theta}$ with $a_m=0$ for $m-n$ odd. Thus by formula (2.1) we need consider only p of the form $\sum_{m=2-n}^{n-2} b_m e^{mi\theta}$ with $b_m=0$ for $m-n$ odd, and integrate over $(0,\pi)$. The stated integrals then vanish by Lemma 2.5. $\square$

2.7. Theorem. *Let* $n \geqq 1$, *then*

(1) $$\int_{-\pi}^{\pi} (C_n^{(\alpha+1,\beta)}(e^{i\theta}))^2 \overline{c_{\alpha\beta}(\theta)}\, dw_\nu(\theta) = \frac{\beta}{\beta+n} \cdot \frac{(2\nu+1)_n}{n!};$$

(2) $$\int_{-\pi}^{\pi} (e^{-i\theta} C_{n-1}^{(\alpha+1,\beta)}(e^{i\theta}))^2 \overline{c_{\alpha\beta}(\theta)}\, dw_\nu(\theta) = -\frac{\beta}{\alpha+n} \cdot \frac{(2\nu+1)_{n-1}}{(n-1)!};$$

(3) $$\int_{-\pi}^{\pi} C_n^{(\alpha+1,\beta)}(e^{i\theta})\, e^{-i\theta} C_{n-1}^{(\alpha+1,\beta)}(e^{i\theta}) \overline{c_{\alpha\beta}(\theta)}\, dw_\nu(\theta) = 0.$$

Proof. In each case use Lemma 2.5 and pick out the appropriate coefficient. We will not explicitly state coefficients for vanishing terms.

For (1), observe $C_n^{(\alpha+1,\beta)}(e^{i\theta}) = \frac{(\beta)_n}{n!} e^{ni\theta} + \sum_{j=0}^{n-1} a_j e^{(2j-n)i\theta}$.

For (2), $e^{-2i\theta} C_{n-1}^{(\alpha+1,\beta)}(e^{i\theta}) = \frac{(\alpha+1)_{n-1}}{(n-1)!} e^{-(n+1)i\theta} + \sum_{j=0}^{n-2} b_j e^{(n-3-2j)i\theta}$, (here n is replaced by $n-1$ in Lemma 2.5).

For (3), $e^{-i\theta} C_{n-1}^{(\alpha+1,\beta)}(e^{i\theta}) = \sum_{j=0}^{n-1} b_j e^{(n-2-2j)i\theta}$. □

From these integrals it is a matter of elementary linear algebra to find the biorthogonal set for $\{h_n : n \in Z\}$, that is, a family of functions $\{g_n : n \in Z\}$ such that

$$\int_{-\pi}^{\pi} h_n \overline{g_m}\, dw_\nu = \delta_{mn} \quad \text{for } m, n \in \mathbb{Z}.$$

2.8. Definition. Let $g_0(e^{i\theta}) := c_{\alpha\beta}(\theta)$, and for $n \geqq 1$, let

$$g_n(e^{i\theta}) := \left(\frac{(n+\nu)\, n!}{\nu(2\nu)_n} C_n^{(\beta,\alpha)}(e^{i\theta}) + \frac{\gamma n!}{\nu(2\nu)_n} 2i \sin\theta\, C_{n-1}^{(\beta,\alpha)}(e^{i\theta}) \right) c_{\alpha\beta}(\theta);$$

and let

$$g_{-n}(e^{i\theta}) := \left(\frac{-\gamma n!}{\nu(2\nu)_n} C_n^{(\beta,\alpha)}(e^{i\theta}) + \frac{\alpha\beta(n+\nu)(n-1)!}{\nu(2\nu)_{n+1}} \cdot 2i \sin\theta\, C_{n-1}^{(\beta+1,\alpha+1)}(e^{i\theta}) \right) c_{\alpha\beta}(\theta).$$

2.9. Theorem. *For* $n, m \in \mathbb{Z}$

$$\int_{-\pi}^{\pi} h_m(e^{i\theta}) \overline{g_n(e^{i\theta})}\, dw_\nu(\theta) = \delta_{mn}.$$

Proof. Use the relation $(C_n^{(\beta,\alpha)}(e^{i\theta}))^- = C_n^{(\alpha,\beta)}(e^{i\theta})$ and convert the integrands to the $\{C_n^{(\alpha+1,\beta)}(e^{i\theta}),\ e^{-i\theta} C_{n-1}^{(\alpha+1,\beta)}(e^{i\theta})\}$ - basis using Proposition 2.3. The resulting integrals are done by Theorem 2.7. □

2.10. Corollary. *Each function* $f \in L^2(T, dw_\nu)$ *has the formal expansion*

$$\sum_{n \in \mathbb{Z}} \hat{f}_n h_n, \quad \text{where } \hat{f}_n := \int_{-\pi}^{\pi} f \overline{g_n}\, dw_\nu.$$

If f is a trigonometric polynomial then the series terminates and has the sum f.

§ 3. The Poisson Kernel

Having determined the biorthogonal set we are ready to define and study the Poisson kernel, which will extend functions on the circle to the disk as series of the form

$$\sum_{m\in\mathbb{Z}} \hat{f}_m h_m(\zeta), |\zeta|<1.$$

This kernel is

$$\sum_{m\in\mathbb{Z}} h_m(\zeta)\overline{g_m(e^{i\theta})}.$$

The kernel will be summed (involving just one ${}_2F_1$) by using the harmonic polynomial expansion of the fundamental solution of $L_{2\gamma}$ established in [2], (the fundamental solution has been studied by Folland and Stein [4], Greiner and Stein [9]).

3.1. Definition. Let

$$\begin{aligned} K_{\alpha\beta}(r,\theta,\phi) := &\sum_{m=0}^{\infty} \frac{m!(m+\nu)}{(2\nu)_m \nu} r^m C_m^{(\alpha,\beta)}(e^{i\theta})\, C_m^{(\alpha,\beta)}(e^{i\phi}) \\ &-\frac{2i\gamma}{\nu}\sum_{m=1}^{\infty}\frac{m!}{(2\nu)_m} r^m(\sin\theta\, C_{m-1}^{(\alpha+1,\beta+1)}(e^{i\theta})\, C_m^{(\alpha,\beta)}(e^{i\phi}) \\ &+\sin\phi\, C_{m-1}^{(\alpha+1,\beta+1)}(e^{i\phi})\, C_m^{(\alpha,\beta)}(e^{i\theta})) \\ &+4\sin\theta\sin\phi\,\alpha\beta\sum_{m=1}^{\infty}\frac{(m-1)!\,(m+\nu)}{(2\nu)_{m+1}\nu} r^m C_{m-1}^{(\alpha+1,\beta+1)}(e^{i\theta})\, C_{m-1}^{(\alpha+1,\beta+1)}(e^{i\phi}), \end{aligned}$$

for $|r|<1$, $-\pi<\theta$, $\phi\leqq\pi$. The Poisson kernel for $z=re^{i\phi}$ is exactly $K_{\alpha\beta}(r,\theta,\phi)c_{\alpha\beta}(\theta)$.

3.2. Proposition. *For fixed θ, ϕ in $[-\pi,\pi]$, $K_{\alpha\beta}(r,\theta,\phi)$ is an analytic function of r in $|r|<1$.*

Proof. The series for $K_{\alpha\beta}$ is dominated term by term by the series

$$1+\sum_{m=1}^{\infty}|r|^m\left(\frac{m+\nu}{\nu}\frac{(2\nu)_m}{m!}+\frac{4|\gamma|}{\nu}\frac{(2\nu+2)_{m-1}}{(m-1)!}+\frac{4\alpha\beta(m+\nu)}{\nu(2\nu)(2\nu+1)}\frac{(2\nu+2)_{m-1}}{(m-1)!}\right)$$

which converges for $|r|<1$. (Recall $|C_m^{(\alpha,\beta)}(e^{i\theta})|\leqq(2\nu)_m/m!$, [2]). □

The expansion of the fundamental solution of $L_{2\gamma}$ holds for $2\nu=N=2,3,4,\ldots$. It involves the disk polynomials

$$R_{kl}^{(\lambda)}(\omega) := \frac{(\lambda+1)_{k+l}}{(\lambda+1)_k(\lambda+1)_l}\,\omega^k\bar{\omega}^l\,{}_2F_1(-k,-l;-\lambda-k-l;1/(\omega\bar{\omega})).$$

$k,l\geqq 0$, $\lambda>-1$. These polynomials are orthogonal on the disk $\{\omega\in\mathbb{C}: |\omega|\leqq 1\}$ with respect to the measure $((\lambda+1)/\pi)\,(1-|\omega|^2)^\lambda\,dm_2(\omega)$ (m_2 is Lebesgue measure on $\mathbb{R}^2$), see [2].

3.3. Theorem. *For* $N=2,3,\ldots$, $\alpha=N/2-\gamma$, $\beta=N/2+\gamma$, *complex* z, w, ω *satisfying* $\operatorname{Im} z$, $\operatorname{Im} w \geqq 0$, $|zw|<1$ *and* $|\omega| \leqq 1$,

$$(3.1)\quad (1-zw-2(\operatorname{Im} z \operatorname{Im} w)^{1/2}\,\omega)^{-\beta}(1-\overline{zw}-2(\operatorname{Im} z \operatorname{Im} w)^{1/2}\,\bar{\omega})^{-\alpha}$$
$$= \sum_{k,l \geqq 0} (4 \operatorname{Im} z \operatorname{Im} w)^{(k+l)/2} R_{kl}^{(N-2)}(\omega)$$
$$\cdot \frac{(\alpha)_l (\beta)_k (N-1)_k (N-1)_l}{k!\,l!\,(N-1)_{k+l}} \sum_{m=0}^{\infty} \frac{m!}{(N+k+l)_m}\, C_m^{(\alpha+l,\beta+k)}(z)\, C_m^{(\alpha+l,\beta+k)}(w).$$

For each k, l *the* m*-sum converges absolutely in* $|zw|<1$, *and the* (k,l)*-sum converges in the Abel method applied to the partial sums* $\sum_{k+l \leqq n}$.

Proof. The expansion found in [2] can be transformed to

$$\left(1-\frac{t-i|\xi|^2+2i\langle\zeta,\xi\rangle}{s+i|\zeta|^2}\right)^{-\alpha}\left(1-\frac{t+i|\xi|^2-2i\langle\xi,\zeta\rangle}{s-i|\zeta|^2}\right)^{-\beta}$$
$$= \sum_{k,l \geqq 0} (2|\xi|\,|\zeta|)^{k+l} R_{kl}^{(N-2)}(\langle\xi,\zeta\rangle/|\xi||\zeta|)$$
$$\cdot \frac{i^{l-k}(\alpha)_l(\beta)_k(N-1)_k(N-1)_l}{k!\,l!\,(N-1)_{k+l}} \cdot \sum_{m=0}^{\infty} \frac{m!}{(N+k+l)_m}$$
$$\cdot\, C_m^{(\alpha+l,\beta+k)}(t+i|\xi|^2)\, C_m^{(\alpha+l,\beta+k)}(s+i|\zeta|^2)\,(s-i|\zeta|^2)^{-k-m}(s+i|\zeta|^2)^{-l-m},$$

where

$$(\xi,t),(\zeta,s)\in H_N \text{ (see § 1) and } t^2+|\xi|^4<s^2+|\zeta|^4.$$

The convergence of the triple sum is as stated above. Make the following substitutions: $z:=t+i|\xi|^2$, $w:=1/(s-i|\zeta|^2)$, $\omega:=(-iw/|w|)\langle\xi,\zeta\rangle/|\xi|\,|\zeta|$. This leads to the theorem (using the homogeneity $R_{kl}^{(N-2)}(c\omega)=c^k\bar{c}^l R_{kl}^{(N-2)}(\omega)$ for $c=i|w|/w$. □

In this work we need the formal expansion as an $R_{kl}^{(N-2)}(\omega)$ - series for fixed z, w, but not the Abel summability, from Theorem 3.3.

From the identity $|1-z\bar{w}|^2+4\operatorname{Im} z \operatorname{Im} w=|1-zw|^2$ we see that

$$0 \leqq \frac{4 \operatorname{Im} z \operatorname{Im} w}{|1-zw|^2} < 1 \quad \text{in the region } |zw|<1, \operatorname{Im} z \geqq 0, \operatorname{Im} w \geqq 0.$$

The following is very convenient in the sequel.

3.4. Definition. For $|r| \leqq 1$, $-\pi \leqq \theta, \phi \leqq \pi$ let

$$X(r,\theta,\phi):=4r\sin\theta\sin\phi/(1-2r\cos(\theta+\phi)+r^2).$$

3.5. Proposition. *For real* θ, ϕ *and complex* r *with* $|r|<1$, $X(r,\theta,\phi)$ *avoids* $[1,+\infty]:=\{x\colon x \geqq 1\}\cup\{\infty\}$.

Proof. We fix θ, ϕ such that $\sin\theta\sin\phi \neq 0$ (else $X \equiv 0$). Then $X(r,\theta,\phi)=F_2\circ F_1(r)$ where $F_1(r):=\frac{1}{2}(r+1/r)$ and

$$F_2(w):=2\sin\theta\sin\phi/(w-\cos(\theta+\phi)).$$

The function F_2 is one-to-one on the extended complex plane $\mathbb{C}_\infty$ and $F_2^{-1}([1, +\infty]) = [\cos(\theta \pm \phi), \cos(\theta \mp \phi)] \subset [-1, 1]$, (choosing the signs to make the left end point smaller). By a well-known result, F_1 maps $\{|r| < 1\}$ one-to-one onto $\mathbb{C}_\infty \backslash [-1, 1]$. $\square$

3.6. Theorem.

$$\sum_{m=0}^{\infty} \frac{m!}{(2v)_m} C_m^{(\alpha,\beta)}(z)\, C_m^{(\alpha,\beta)}(w) \tag{3.2}$$
$$= (1 - \overline{zw})^{-\alpha}(1 - zw)^{-\beta}\, {}_2F_1(\alpha, \beta; 2v; 4 \operatorname{Im} z \operatorname{Im} w / |1 - zw|^2)$$

holds for $\operatorname{Im} z \geqq 0$, $\operatorname{Im} w \geqq 0$, $|zw| < 1$, $v > 0$, $|\gamma| \leqq v$. Equivalently,

$$\sum_{m=0}^{\infty} \frac{m!}{(2v)_m} r^m C_m^{(\alpha,\beta)}(e^{i\theta})\, C_m^{(\alpha,\beta)}(e^{i\phi})$$
$$= (1 - re^{-i(\theta+\phi)})^{-\alpha}(1 - re^{i(\theta+\phi)})^{-\beta}\, {}_2F_1(\alpha, \beta; 2v; X(r, \theta, \phi)),$$

for $0 \leqq r < 1$; $0 \leqq \theta, \phi \leqq \pi$; $v > 0$ *and* $|\gamma| \leqq v$.

Proof. We first establish the result for $2v = N = 2, 3, \ldots$. The left side of (3.1) can be written as $F(\omega, z, w) := (1 - \overline{zw})^{-\alpha}\, (1 - zw)^{-\beta}\, (1 - \bar{u}(z, w)\,\bar{\omega})^{-\alpha}\, (1 - u(z, w)\,\omega)^{-\beta}$ where $u(z, w) := 2(\operatorname{Im} z \operatorname{Im} w)^{1/2}/(1 - zw)$. Since

$$((N-1)/\pi) \int_0^1 \int_{-\pi}^{\pi} R_{kl}(te^{i\theta})(1 - t^2)^{N-2}\, t\, d\theta\, dt = \delta_{k0}\, \delta_{l0},$$

the constant term of the expansion can be obtained as

$$((N-1)/\pi) \int_0^1 \int_{-\pi}^{\pi} F(te^{i\theta}, z, w)(1 - t^2)^{N-2}\, t\, d\theta\, dt.$$

But

$$F(te^{i\theta}, z, w) = (1 - \overline{zw})^{-\alpha}(1 - zw)^{-\beta} \sum_{j,k \geqq 0} \frac{(\alpha)_j(\beta)_k}{j!\,k!} (\bar{u}(z, w)\,t)^j (u(z, w)\,t)^k\, e^{i\theta(k-j)}$$

which converges absolutely and uniformly in $|te^{i\theta}| \leqq 1$ for fixed z, w, since $|u(z, w)| < 1$. Integrating term by term over $-\pi \leqq \theta \leqq \pi$ with respect to $(1/2\pi)\, d\theta$ yields

$$(1 - \overline{zw})^{-\alpha}(1 - zw)^{-\beta}\, {}_2F_1(\alpha, \beta; 1; |u(z, w)|^2\, t^2),$$

uniformly convergent in $0 \leqq t \leqq 1$. Next we integrate with respect to $2(N-1)\,t(1 - t^2)^{N-2}\, dt$ on $0 \leqq t \leqq 1$. The ${}_2F_1$-series is integrated term by term using $2(N-1) \int_0^1 t^{2j}(1 - t^2)^{N-2}\, t\, dt = j!/(N)_j$. This establishes the result for $v = 1, 3/2, 2, \ldots$.

To get other values of v we use Carlson's theorem (Bailey [1], p. 39: if an analytic function on a half-plane satisfies $|f(z)| \leqq Ke^{a|z|}$ with $a < \pi$ and $f(n) = 0$ for $n = n_0, n_0 + 1, n_0 + 2 \ldots$ then $f \equiv 0$; the growth condition rules out functions like $\sin \pi z$). We have established the identity (3.2), both sides of which are analytic in $2v$ for $\operatorname{Re} v > |\gamma|$ and fixed z, w, γ, for the values $2v = n_0, n_0 + 1, \ldots$ (n_0 is the least integer $\geqq |\gamma|$).

We concentrate on one value of m (as a power series in r). For fixed θ, ϕ, and any $\varepsilon>0$ there exists $r_0<1$ such that $|X(r,\theta,\phi|<\varepsilon$ and $|2r\cos(\theta+\phi)-r^2|<\varepsilon$ for $|r|\leqq r_0$, and the right side of (3.2) is analytic in a neighborhood of $\{r\in\mathbb{C}: |r|\leqq r_0\}$.

By Cauchy's integral formula,

$$\frac{m!}{(2\nu)_m} C_m^{(\alpha,\beta)}(e^{i\theta})\, C_m^{(\alpha,\beta)}(e^{i\phi}) = \frac{1}{2\pi i} \oint_{r=r_0} (1-re^{-i(\theta+\phi)})^{\gamma-\nu} \cdot (1-re^{i(\theta+\phi)})^{-\gamma-\nu}\,{}_2F_1(\nu-\gamma,\nu+\gamma;2\nu;X(r,\theta,\phi))\bar{r}^{m-1}\,dr$$

holds for $2\nu=n_0, n_0+1, \ldots$. The left side is rational in ν, thus of $O(e^{a|\nu|})$ for any $a>0$. The right side is analytic in $\operatorname{Re}\nu>|\gamma|$ and is bounded by a constant (in m) times

$$\sup_{|r|\leqq r_0}(1-2r\cos(\theta+\phi)+r^2)^{-\nu}|{}_2F_1(\nu-\gamma,\nu+\gamma;2\nu;X(r,\theta,\phi))|.$$

It suffices to consider $\gamma\geqq 0$, then $\left|\dfrac{\nu-\gamma+j}{2\nu+j}\right|\leqq 1$ for each $j=0,1,2\ldots$, when $\operatorname{Re}\nu>\gamma$.

Thus the ${}_2F_1$-series is dominated by

$$\sum_{j=0}^{\infty}\frac{(|\nu|+\gamma)_j}{j!}|X|^j=(1-|X|)^{-|\nu|-\gamma}\leqq(1-\varepsilon)^{-|\nu|-\gamma}\leqq\exp((|\nu|+\gamma)\,|\log(1-\varepsilon)|),$$

and so the right side is $O(\exp(|\nu|\,|\log(1-\varepsilon)|)$. Since ε was arbitrary >0, Carlson's theorem applies. □

(3.3) Let $F_{\alpha\beta}(r,\theta,\phi) := \sum_{m=0}^{\infty}\dfrac{m!}{(2\nu)_m} r^m C_m^{(\alpha,\beta)}(e^{i\theta})\, C_m^{(\alpha,\beta)}(e^{i\phi})$.

3.7. Corollary.

$$F_{\alpha\beta}(r,\theta,\phi)=(1-re^{-i(\theta+\phi)})^{-\alpha}\cdot(1-re^{i(\theta+\phi)})^{-\beta}\,{}_2F_1(\alpha,\beta;2\nu;X(r,\theta,\phi))$$

holds for all real θ, ϕ and complex r with $|r|<1$.

Proof. By an argument similar to that in Prop. 3.2 we see that $F_{\alpha\beta}$ is analytic in r for fixed real θ, ϕ and $|r|<1$. On the other hand $X(r,\theta,\phi)$ avoids $[1,+\infty]$ by Prop. 3.5, and the hypergeometric function is analytic on $\mathbb{C}\backslash[1,+\infty]$. Thus the two sides of the formula are analytic on $|r|<1$ and agree for $0\leqq r<1$ by the theorem (for $0\leqq\theta, \phi\leqq\pi$), hence for the entire disk. From the relations $F_{\alpha\beta}(r,\theta-\pi,\phi)=F_{\alpha\beta}(-r,\theta,\phi)$ and $X(r,\theta-\pi,\phi)=X(-r,\theta,\phi)$ we deduce the validity of the formula for all θ, ϕ in $(-\pi,\pi)$. □

This bilinear sum is similar to one found by Greiner and Koornwinder ((1.28) in [8]).

3.8. Lemma.

$$K_{\alpha\beta}(r,\theta,\phi)=\sum_{m=0}^{\infty}\frac{m!}{(2\nu+1)_m}\cdot\frac{\beta+m}{\beta}\,r^m\cdot C_m^{(\alpha+1,\beta)}(e^{i\theta})\,C_m^{(\alpha+1,\beta)}(e^{i\phi})$$
$$-e^{-i(\theta+\phi)}\sum_{m=1}^{\infty}\frac{(m-1)!}{(2\nu+1)_{m-1}}\cdot\frac{\alpha+m}{\beta}\,r^m\,C_{m-1}^{(\alpha+1,\beta)}(e^{i\theta})\,C_{m-1}^{(\alpha+1,\beta)}(e^{i\phi})$$
$$=(1/\beta)\left(\beta+r\frac{\partial}{\partial r}-re^{-i(\theta+\phi)}\left(\alpha+1+r\frac{\partial}{\partial r}\right)\right)F_{\alpha+1,\beta}(r,\theta,\phi),$$

for all real θ, ϕ and complex r, $|r|<1$.

Proof. The first equality comes from direct substitution in the definition of $K_{\alpha\beta}$ using Prop. 2.3 (or by use of the orthogonality relations in Theorem 2.7). The second equality is a standard generating function device (term-by-term manipulation). □

3.9. Theorem. *For $\alpha,\beta>0$; $-\pi\leqq\theta$, $\phi\leqq\pi$; $|r|<1$,*

$$K_{\alpha\beta}(r,\theta,\phi)=(1-r^2)(1-re^{-i(\theta+\phi)})^{-\alpha-1}\cdot(1-re^{i(\theta+\phi)})^{-\beta-1}\,{}_2F_1(\alpha+1,\beta+1;2\nu+1;X(r,\theta,\phi)). \tag{3.4}$$

Proof. Let $\xi:=1-re^{i(\theta+\phi)}$. We apply the differential operator from Lemma 3.8 to

$$F_{\alpha+1,\beta}=\bar{\xi}^{-\alpha-1}\xi^{-\beta}\,{}_2F_1(\alpha+1,\beta;2\nu+1;X).$$

In the computation we use

$$\frac{\partial}{\partial X}\,{}_2F_1(\alpha+1,\beta;2\nu+1;X)=\frac{(\alpha+1)\beta}{(2\nu+1)}\,{}_2F_1(\alpha+2;\beta+1;2\nu+2;X)$$

and $r\dfrac{\partial}{\partial r}X=(1-r^2)X/(\xi\bar{\xi})$. This leads to

$$K_{\alpha\beta}(r,\theta,\phi)=(1-r^2)\,\bar{\xi}^{-\alpha-1}\,\xi^{-\beta-1}$$
$$\left\{{}_2F_1(\alpha+1,\beta;2\nu+1;X)+\left(\frac{\alpha+1}{2\nu+1}\right)X\,{}_2F_1(\alpha+2,\beta+1;2\nu+2;X)\right\}$$

and then to (3.4) by a contiguity result for hypergeometric functions. □

We have arrived at a rather neat expression for the Poisson kernel, but it is useful to apply some transformations to the ${}_2F_1$-function. This is to make it easier to study the behaviour of the kernel as $r\to 1_-$ at the critical values $\theta=\pm\phi$.

3.10. Proposition. *For $|r|<1$ and real θ, ϕ*

$$K_{\alpha\beta}(r,\theta,\phi)=(1-r^2)(1-2r\cos(\theta-\phi)+r^2)^{-1}(1-re^{-i(\theta+\phi)})^{-\alpha}\cdot(1-re^{i(\theta+\phi)})^{-\beta}\,{}_2F_1(\alpha,\beta;2\nu+1;X(r,\theta,\phi)).$$

The ${}_2F_1$ function is uniformly bounded in the region $0\leqq r<1$ and $\sin\theta\sin\phi\geqq 0$.

Proof. Use the Pfaff transformation

$$_2F_1(a,b;c;x)=(1-x)^{c-a-b}\,{}_2F_1(c-a,c-b;c;x)$$

on the $_2F_1$ function in (3.4). Note that

$$(1-X)=(1-2r\cos(\theta-\phi)+r^2)/(1-2r\cos(\theta+\phi)+r^2).$$

The resulting $_2F_1$ function is

$$_2F_1(\alpha,\beta;2\nu+1;X)\leqq{}_2F_1(\alpha,\beta;2\nu+1;1)=\frac{\Gamma(2\nu+1)}{\Gamma(\alpha+1)\Gamma(\beta+1)}$$

by Gauss's sum (Bailey [1], p. 2), when $0\leqq X<1$, which is the case for $0\leqq r<1$ and $\sin\theta\sin\phi\geqq 0$. □

Although we require $\nu>0$ in this work, this form of $K_{\alpha\beta}$ agrees with the Poisson kernel for Fourier series when $\alpha=\beta=0$. Also for the limiting case $\nu=|\gamma|>0$, that is, $\alpha=0$ or $\beta=0$ it can be shown by direct computation (using $C_m^{(0,\beta)}(z)=((\beta)_m/m!)\,z^m$) that the above expression is valid; note the $_2F_1$ function is trivial in this case.

The following results are valid for $\nu>0$ and $|\gamma|\leqq\nu$.

3.11. Proposition. *For* $-\pi\leqq\theta,\ \phi\leqq\pi$ *and* $|r|<1$

$$K_{\alpha\beta}(r,\theta,\phi)=(1-r^2)(1-re^{i(\theta+\phi)})^{\alpha-\beta}(1-2r\cos(\theta-\phi)+r^2)^{-\alpha-1}\cdot{}_2F_1(\alpha+1,\alpha;2\nu+1;X(r,-\theta,\phi)) \tag{3.5}$$

$$=(1-r^2)(1-re^{-i(\theta+\phi)})^{\beta-\alpha}(1-2r\cos(\theta-\phi)+r^2)^{-\beta-1}\cdot{}_2F_1(\beta+1,\beta;2\nu+1;X(r,-\theta,\phi)). \tag{3.6}$$

In the region $0\leqq r<1$ *and*

$$\sin\theta\sin\phi\leqq 0,\quad 0\leqq X(r,-\theta,\phi)=X(r,\theta,-\phi)<1.$$

Proof. The transformation $_2F_1(a,b;c;x)=(1-x)^{-1}\,{}_2F_1\left(a,c-b;c;\dfrac{x}{x-1}\right)$ is used. Note that $X(r,\theta,\phi)/(X(r,\theta,\phi)-1)=X(r,-\theta,\phi)=-4r\sin\theta\ \sin\phi/(1-2r\cos(\phi-\theta)+r^2)$, which is positive when $\sin\theta\sin\phi<0$. □

This completes the description of the Poisson kernel. Form (3.5) is appropriate for $\alpha\leqq\beta$ while (3.6) is used for $\alpha>\beta$. If $\alpha<\nu$ (respectively $\beta<\nu$) the $_2F_1$ function in (3.5) is uniformly bounded in $0\leqq r<1$, $\sin\theta\sin\phi<0$ (that is, $e^{i\theta},e^{i\phi}$ on opposite sides of the real axis); in this case the only singularity in $K_{\alpha\beta}(r,\theta,\phi)$ is due to the factor $(1-r\,e^{i(\theta+\phi)})^{\alpha-\beta}$ which diverges for $\theta=-\phi$, $r\to 1_-$. Proposition 3.10 shows the nature of the singularity at $\theta=\phi$, $r\to 1_-$, which is the one expected from the known theory.

§ 4. Radial convergence of the Poisson Integral

Using the explicit forms of the Poisson kernel obtained in § 3 we can show that radial convergence holds, in the norm, in $L^p(T,dw_\nu)$, $1\leqq p<\infty$, and in $C(T)$

provided that $|\gamma| \leqq 1$. For each γ with $|\gamma|>1$ there exist continuous functions whose Poisson integrals diverge at some point on the circle. In the argument we use comparison with the known (ultraspherical) cases $\gamma=0$ wherever possible, and careful estimates elsewhere.

4.1. Definition. For $f \in L^1(T, dw_\nu)$, $0 \leqq r<1$, $-\pi<\theta \leqq \pi$, define the Poisson integral

$$P_r^{(\alpha,\beta)}[f](e^{i\theta}) := \int_{-\pi}^{\pi} f(e^{i\phi}) K_{\alpha\beta}(r,\theta,\phi) \overline{c_{\alpha\beta}(\phi)}\, dw_\nu(\phi),$$

and $P^{(\alpha,\beta)}[f](r e^{i\theta}) := P_r^{(\alpha,\beta)}[f](e^{i\theta})$.

4.2. Proposition. *For each* $f \in L^1(T, dw_\nu)$, $P^{(\alpha,\beta)}[f] \in C^\infty(\{z \in \mathbb{C}: |z|<1\})$ *and satisfies* $D_{\alpha+1,\beta+1} D_{\alpha\beta} P^{(\alpha,\beta)}[f](z)=0$ *(see (2.3)).*

Proof. From Definition 3.1 it can be seen that $K_{\alpha\beta}(r,\theta,\phi)$ as a power series in $r e^{i\theta}$, $r e^{-i\theta}$ is dominated by a series $\sum_m A_m r^m$ with $A_m = O(m^{2\nu+2})$ (uniformly in θ, ϕ). Thus for each ϕ, the function $r e^{i\theta} \to K_{\alpha\beta}(r,\theta,\phi)$ is infinitely often differentiable and satisfies the stated differential equation. Standard methods extend the result to $P^{(\alpha,\beta)}[f]$. □

4.3. Proposition. *For any trigonometric polynomial* $f(e^{i\theta}) = \sum_{m=-M}^{M} a_m e^{im\theta}$, $P_r^{(\alpha,\beta)}[f] \to f$ *uniformly as* $r \to 1-$.

Proof. $P^{(\alpha,\beta)}[f](z)$ is a polynomial in $z, \bar{z}$ which agrees with $f(z)$ on $|z|=1$ (see Corollary 2.10). □

To get any further we must consider

$$\sup\left\{\int_{-\pi}^{\pi} |K_{\alpha\beta}(r,\theta,\phi)|\, |c_{\alpha\beta}(\theta)|\, dw_\nu(\theta) \colon 0 \leqq r<1,\ -\pi<\phi \leqq \pi\right\}.$$

Since $|c_{\alpha\beta}(\theta)|$ is constant we can disregard it for now. It suffices to consider the case $0 \leqq \alpha \leqq \nu$, that is $0 \leqq \gamma \leqq \nu$. Then for $0 \leqq r<1$,

$$(4.1)\quad |K_{\alpha\beta}(r,\theta,\phi)| = (1-r^2)(1-2r\cos(\theta+\phi)+r^2)^{-\nu} \cdot (1-2r\cos(\theta-\phi)+r^2)^{-1}\, {}_2F_1(\alpha,\beta;2\nu+1;X(r,\theta,\phi))$$

for $\sin\theta \sin\phi \geqq 0$,

$$(4.2)\quad = (1-r^2)(1-2r\cos(\theta-\phi)+r^2)^{-\alpha-1} \cdot (1-2r\cos(\theta+\phi)+r^2)^{-\gamma}\, {}_2F_1(\alpha+1,\alpha;2\nu+1;X(r,-\theta,\phi))$$

for $\sin\theta \sin\phi \leqq 0$.

Since $\int_{-\pi}^{\pi} K_{\alpha\beta}(r,\theta,\phi) \overline{c_{\alpha\beta}(\theta)}\, dw_\nu(\theta)=1$ (Poisson integral of 1) and $K_{\nu\nu}(r,\theta,\phi) \geqq 0$, $c_{\nu\nu}=1$ we immediately obtain the desired L^1-norm bound for $K_{\nu\nu}$.

4.4. Lemma. *For* $0 \leqq r<1$, *and* $\sin\theta \sin\phi \geqq 0$, $|K_{\alpha\beta}(r,\theta,\phi)| \leqq K_{\nu\nu}(r,\theta,\phi)$.

Proof. It suffices to show ${}_2F_1(\alpha,\beta;2\nu+1;X)\leqq {}_2F_1(\nu,\nu;2\nu+1;X)$ when $0\leqq X<1$. But $0\leqq(\alpha+j)(\beta+j)=(\nu+j)^2-\gamma^2\leqq(\nu+j)^2$ when $j=0,1,2,\ldots$, hence $(\alpha)_j(\beta)_j\leqq(\nu)_j(\nu)_j$. □

From symmetry it suffices to consider $\int_{-\pi}^{\pi}|K_{\alpha\beta}(r,\theta,\phi)|\,dw_\nu(\theta)$ for $0\leqq\phi\leqq\pi$ Lemma 4.4 already shows that $\int_0^{\pi}|K_{\alpha\beta}(r,\theta,\phi)|\,dw_\nu(\theta)\leqq 1$, so the problem will be in $(-\pi,0)$, in particular, in a neighborhood of $-\phi$. We will show that there are constants A_γ (all $\gamma\neq 0$), B_γ (for $0<|\gamma|\leqq 1$) and functions $\phi_1(\phi)$, $\phi_2(\phi)$ on $0<\phi<\pi$ with $0<\phi_1(\phi)<\phi<\phi_2(\phi)<\pi$ such that for each ϕ in $(0,\pi)$:

i) if $0\leqq r\leqq\max(\frac{1}{2},1-\sin\phi)$, then

$$|K_{\alpha\beta}(r,\theta,\phi)|\leqq A_\gamma K_{\nu\nu}(r,\theta,\phi)\qquad\text{for all }\theta$$

(this includes the cases $\phi=0$ or π);

ii) otherwise

$$\int_{-\phi_2}^{-\phi_1}|K_{\alpha\beta}(r,\theta,\phi)|\,dw_\nu(\theta)\leqq B_\gamma$$

and

$$|K_{\alpha\beta}(r,\theta,\phi)|\leqq A_\gamma K_{\nu\nu}(r,\theta,\phi)\qquad\text{for all }\theta\text{ outside }(-\phi_2,-\phi_1).$$

We need some facts about the behavior of $X(r,\theta,\phi)$ for a fixed ϕ in $(0,\pi)$, (calculus exercises):

(4.3) For fixed θ in $(0,\pi)$, X is strictly increasing on $0\leqq r<1$

with maximum value $X(1,\theta,\phi)=(\cos(\theta-\phi)-\cos(\theta+\phi))/(1-\cos(\theta+\phi))$;

(4.4) for fixed r, $0<r\leqq 1$, X is unimodal on $0\leqq\theta\leqq\pi$ with maximum value $2/(1+\{((1-r^2)/(2r\sin\phi))^2+1\}^{\frac{1}{2}})$ at $\cos\theta=(2r/(1+r^2))\cos\phi$.

We define the intervals $[-\phi_2(\phi),-\phi_1(\phi)]$ outside which $|K_{\alpha\beta}|$ is comparable to $K_{\nu\nu}$.

4.5. Definition. For $0<\phi<\pi$. let

$$\phi_1(\phi):=\begin{cases}\phi/2 & \text{for } 0<\phi\leqq 2\pi/3\\ 2\phi-\pi & \text{for } 2\pi/3\leqq\phi<\pi;\end{cases}$$

$$\phi_2(\phi):=\begin{cases}2\phi & \text{for } 0<\phi\leqq \pi/3\\ (\pi+\phi)/2 & \text{for } \pi/3\leqq\phi<\pi;\end{cases}$$

note that $\phi_1(\pi-\phi)=\pi-\phi_2(\phi)$, a symmetry around $\phi=\pi/2$.

4.6. Proposition. *If* $0\leqq\theta\leqq\phi_1(\phi)$ *or* $\phi_2(\phi)\leqq\theta\leqq\pi$ *and* $0\leqq r<1$, *then* $X(r,\theta,\phi)\leqq 8/9$.

Proof. For given θ,ϕ, $X(r,\theta,\phi)\leqq X(1,\theta,\phi)$ (see (4.3)) and $X(1,\theta,\phi)$ has its unique maximum at $\theta=\phi$. First we show $X(1,\theta,2\theta)\leqq 8/9$ for $0<\theta<\pi/2$; indeed $X(1,\theta,2\theta)=(\cos\theta-\cos 3\theta)/(1-\cos 3\theta)=1-1/(1+2\cos\theta)^2$, a decreasing function on $0<\theta<2\pi/3$, and $\lim_{\theta\to 0}X(1,\theta,2\theta)=8/9$. Then in $0\leqq\theta\leqq\phi/2\leqq\pi/3$,

$X(1,\theta,\phi) \leqq X(1,\phi/2,\phi) \leqq 8/9$. When $0<\phi \leqq \pi/3$ and $2\phi \leqq \theta \leqq \pi$, then $X(1,\theta,\phi) \leqq X(1,2\phi,\phi) \leqq 8/9$. A symmetry argument using $X(1,\pi-\theta,\pi-\phi) = X(1,\theta,\phi)$ takes care of the other possibilities. □

4.7. Proposition. *If* $0<\phi<\pi$, $0 \leqq \theta \leqq \pi$ *and* $0 \leqq r \leqq \max(\frac{1}{2}, 1-\sin\phi)$, *then* $X(r,\theta,\phi) \leqq 8/9$.

Proof. From (4.4) $\sup\limits_{0 \leqq \theta \leqq \pi} X(r,\theta,\phi) = 2/(1+((1-r^2)/(2r\sin\phi)^2+1)^{\frac{1}{2}})$. If $0<r \leqq \frac{1}{2}$ then $1/(2r\sin\phi) \geqq 1$, so $(1-r^2)/(2r\sin\phi) \geqq 1-r^2 \geqq 3/4$ making the supremum $\leqq 8/9$. Otherwise, if $1-r \geqq \sin\phi > 0$, then

$$\frac{1-r^2}{2r\sin\phi} = \left(\frac{1}{2}+\frac{1}{2r}\right)\frac{1-r}{\sin\phi} \geqq \frac{1-r}{\sin\phi} \geqq 1, \quad \text{and} \quad \sup_\theta X \leqq 2/(1+\sqrt{2}) < 8/9. \quad \square$$

4.8. Lemma. *For* $\gamma>0$ *and* $0<\phi<\pi$, *if* $0 \leqq \theta \leqq \pi$ *or if* r,θ *satisfy* $0 \leqq X(r, -\theta,\phi) \leqq 8/9$ *then*

$$|K_{\alpha\beta}(r,\theta,\phi)| \leqq 9^\gamma \, {}_2F_1(\alpha,\alpha+1;2\nu+1;8/9)\, K_{\nu\nu}(r,\theta,\phi).$$

Proof. If $0 \leqq \theta \leqq \pi$ then $|K_{\alpha\beta}(r,\theta,\phi)| \leqq K_{\nu\nu}(r,\theta,\phi)$ by Lemma 4.4. If $-\pi<\theta<0$ then

$$|K_{\alpha\beta}(r,\theta,\phi)|/K_{\nu\nu}(r,\theta,\phi) = (1-X)^{-\gamma}\,{}_2F_1(\alpha+1,\alpha;2\nu+1;X)/({}_2F_1(\nu+1,\nu;2\nu+1;X)),$$

with $X=X(r,-\theta,\phi)$, (see Proposition 3.11). Under the hypotheses, $0 \leqq X \leqq 8/9$ implying that $(1-X)^{-\gamma} \leqq 9^\gamma$, ${}_2F_1(\alpha+1,\alpha;2\nu+1;X) \leqq {}_2F_1(\alpha+1,\alpha;2\nu+1;8/9)$ and ${}_2F_1(\nu+1,\nu;2\nu+1;X) \geqq 1$. □

Now set $A_\gamma := 9^{|\gamma|}\,{}_2F_1(\nu-|\gamma|,\nu-|\gamma|+1;2\nu+1;8/9)$. This is probably not the best possible bound, but the real problem is in $-\phi_2(\phi)<\theta<-\phi_1(\phi)$ anyway.

We establish some bounds in $\phi_1(\phi) \leqq \theta \leqq \phi_2(\phi)$ for fixed ϕ in $(0,\pi)$.

4.9. Lemma. 1) *If* $\phi_1(\phi) \leqq \theta \leqq \phi_2(\phi)$ *then* $\frac{1}{2} \leqq \dfrac{\sin\theta}{\sin\phi} \leqq 2$;

2) $\max(2\sin((\phi-\phi_1)/2), 2\sin((\phi_2-\phi)/2)) \leqq (2/\sqrt{3})\sin\phi$.

Proof. By symmetry it suffices to consider $0<\phi \leqq \pi/2$.

First we see that

$$\sup_{\phi_1 \leqq \theta \leqq \phi_2} \frac{\sin\theta}{\sin\phi} = \begin{cases} 2\cos\phi & \text{for } 0<\phi \leqq \pi/4 \\ 1/\sin\phi & \text{for } \pi/4 \leqq \phi \leqq \pi/2. \end{cases}$$

The supremum over ϕ is 2. Also $\inf\limits_{\phi_1 \leqq \theta \leqq \phi} \dfrac{\sin\theta}{\sin\phi} = \dfrac{\sin(\phi/2)}{\sin\phi} = 1/(2\cos(\phi/2)) \geqq \frac{1}{2}$.

For claim (2) the maximum of the two functions in $0<\phi \leqq \pi/2$ is

$$2\sin((\phi_2-\phi)/2) = \begin{cases} 2\sin\phi/2 & \text{in } 0<\phi \leqq \pi/3 \\ 2\sin(\pi-\phi)/4 & \text{in } \pi/3 \leqq \phi \leqq \pi/2. \end{cases}$$

Now $2\sin(\phi/2)/\sin\phi = 1/\cos(\phi/2) \leqq 1/\cos(\pi/6) = 2/\sqrt{3}$ in $0<\phi\leqq\pi/3$. Further $2\sin((\pi-\phi)/4)/\sin\phi$ decreases on $\pi/3\leqq\phi\leqq\pi/2$ and has the value $2/\sqrt{3}$ at $\phi=\pi/3$. □

4.10. Lemma. *If $\gamma>0$ there is a constant C_γ depending on α,β such that*

$$\int_{-\phi_2(\phi)}^{-\phi_1(\phi)} |K_{\alpha\beta}(r,\theta,\phi)|\,dw_v(\theta)$$

$$\leqq C_\gamma r^{-\alpha-1}(1-r^2)(\sin\phi)^{2\gamma-2}\int_0^{(2\sin\phi)/\sqrt{3}}((1-r)^2+rx^2)^{-\gamma}\,dx,$$

for each ϕ in $(0,\pi)$.

Proof. From (4.2)

$$|K_{\alpha\beta}(r,-\theta,\phi)| \leqq (1-r^2)(1-2r\cos(\theta+\phi)+r^2)^{-\alpha-1} \cdot (1-2r\cos(\theta-\phi)+r^2)^{-\gamma}\,{}_2F_1(\alpha+1,\alpha;2v+1;1);$$

also

$$(1-2r\cos(\theta+\phi)+r^2)^{-1} \leqq (4r\sin\theta\sin\phi)^{-1},$$

for $0<\theta<\pi$.
Thus the integral is bounded by a constant (involving the ${}_2F_1$-series and $B(v+\tfrac{1}{2},\tfrac{1}{2})$ from dw_v) times

$$(1-r^2)(4r)^{-\alpha-1}(\sin\phi)^{-\alpha-1}\int_{\phi_1}^{\phi_2}(1-2r\cos(\theta-\phi)+r^2)^{-\gamma}(\sin\theta)^{v+\gamma-1}\,d\theta.$$

By Lemma 4.9 (1), $(\sin\theta)^{v+\gamma}\leqq 2^{v+\gamma}(\sin\phi)^{v+\gamma}$ and $(\sin\theta)^{-1}\leqq 2(\sin\phi)^{-1}$ in $\phi_1\leqq\theta\leqq\phi_2$; thus the integral is bounded by

$$B'_\gamma r^{-\alpha-1}(1-r^2)(\sin\phi)^{2\gamma-2}\int_{\phi_1-\phi}^{\phi_2-\phi}(1-2r\cos\theta+r^2)^{-\gamma}\,d\theta$$

(having replaced θ by $\theta+\phi$; and some constant B'_γ). In the latter integral, let $x=2\sin(\theta/2)$ with $d\theta=dx/(\cos(\theta/2))$. For $\phi_1-\phi\leqq\theta\leqq\phi_2-\phi$ we have $|\theta|\leqq\pi/3$ (maximum achieved at $\phi=\pi/3$) and so $1/\cos(\theta/2)\leqq 1/\cos(\pi/6)=2/\sqrt{3}$. The limits of integration are $2\sin((\phi_1-\phi)/2)\leqq x\leqq 2\sin((\phi_2-\phi)/2)$, and since the integrand is even, the integral is bounded by twice the integral over $0\leqq x\leqq\max(2\sin((\phi-\phi_1)/2,\ 2\sin((\phi_2-\phi)/2))$, which by Lemma 4.9 (2) is included in the interval $0\leqq x\leqq(2\sin\phi)/\sqrt{3}$.

Also $1-2r\cos\theta+r^2=(1-r)^2+rx^2$. □

It remains to bound $\displaystyle\int_0^{(2\sin\phi)/\sqrt{3}}((1-r)^2+rx^2)^{-\gamma}\,dx$.

4.11. Theorem. *For $0<\gamma<1$ and $v\geqq\gamma$ there is a constant B_γ so that for $0<\phi<\pi$ and $\max(\tfrac{1}{2},1-\sin\phi)\leqq r<1$,*

$$\int_{-\phi_2(\phi)}^{-\phi_1(\phi)}|K_{\alpha\beta}(r,\theta,\phi)|\,dw_v(\theta)\leqq B_\gamma.$$

Proof. From Lemma 4.10, it remains to consider the integral

$$I(r,\phi):=r^{-\alpha-1}(1-r^2)(\sin\phi)^{2\gamma-2}\int_0^{(2\sin\phi)/\sqrt{3}}((1-r)^2+rx^2)^{-\gamma}\,dx$$
$$\leqq r^{-\alpha-3/2}(1+r)\left(\frac{1-r}{\sin\phi}\right)^{2-2\gamma}\int_0^{(2\sin\phi)/\sqrt{3}(1-r))}(1+y^2)^{-\gamma}\,dy$$

after the substitution $x=(1-r)\,y/\sqrt{r}$ and by use of

$$((2\sin\phi)/\sqrt{3})(\sqrt{r}/(1-r))\leqq(2/\sqrt{3})\sin\phi/(1-r).$$

For $s>1$,

$$\begin{aligned}\int_0^s(1+y^2)^{-\gamma}\,dy&=O(1) && \text{for } \gamma>\tfrac{1}{2},\\ &=O(1+\log s) && \text{for } \gamma=\tfrac{1}{2} \quad\text{and}\\ &=O(s^{1-2\gamma}) && \text{for } 0<\gamma<\tfrac{1}{2};\end{aligned}$$

indeed $\int_0^s(1+y^2)^{-\gamma}\,dy\leqq\int_0^1(1+y^2)^{-\gamma}\,dy+\int_1^s y^{-2\gamma}\,dy$.
By assumption $\frac{1}{2}\leqq r<1$ and thus $r^{-\alpha-3/2}(1+r)\leqq 2^{\alpha+3/2}(3/2)$. Also $\sin\phi/(1-r)\geqq 1$. Thus there is a constant B''_γ such that

$$I(r,\phi)\leqq\begin{cases}B''_\gamma((1-r)/\sin\phi) & \text{for } 0<\gamma<\tfrac{1}{2}\\ B''_\gamma((1-r)/\sin\phi)^{2-2\gamma} & \text{for } \gamma>\tfrac{1}{2}\\ B''_\gamma\left(\dfrac{1-r}{\sin\phi}\right)\left(1+\log\left(\dfrac{\sin\phi}{1-r}\right)\right) & \text{for } \gamma=\tfrac{1}{2}.\end{cases}$$

Hence there is a constant B'_γ so that $I(r,\phi)\leqq B'_\gamma$.
Also for each ϕ, $I(r,\phi)\to 0$ as $r\to 1-$ provided $0<\gamma<1$. □

4.12. Theorem. *If* $-1\leqq\gamma\leqq 1$ *then*

$$\int_{-\pi}^{\pi}|K_{\alpha\beta}(r,\theta,\phi)|\,|\overline{c_{\alpha\beta}}(\theta)|\,dw_\nu(\theta)\leqq A_{\alpha\beta},$$

for some finite constant $A_{\alpha\beta}$, $(0\leqq r<1,\ -\pi<\phi\leqq\pi)$.

Proof. Combining the bounds from Theorem 4.11 and Lemma 4.8 we obtain

$$\int_{-\pi}^{\pi}|K_{\alpha\beta}(r,\theta,\phi)|\,|\overline{c_{\alpha\beta}}(\theta)|\,dw_\nu(\theta)\leqq\frac{\Gamma(\alpha+1)\,\Gamma(\beta+1)}{\Gamma(\nu+1)^2}(A_\gamma+B_\gamma),$$

for $0<|\gamma|\leqq 1$. □

We can finally prove the desired convergence results. (The method of proof is as in Muckenhoupt and Stein [12] p. 31.)

4.13. Theorem. *Let* $\alpha,\beta>0$ *with* $|\alpha-\beta|\leqq 2$,

1) *if* $f\in C(T)$, *then* $P_r^{(\alpha,\beta)}[f]\to f$ *uniformly as* $r\to 1-$.

2) *if* $f \in L^p(T, dw_\nu)$ *with* $1 \leqq p < \infty$, *then*

$$\|P_r^{(\alpha,\beta)}[f]\|_p \leqq A_{\alpha\beta}\|f\|_p \quad \textit{and} \quad \|P_r^{(\alpha,\beta)}[f]-f\|_p \to 0 \quad \textit{as } r \to 1-.$$

Proof. We first show for $f \in L^p(dw_\nu)$, $1 \leqq p < \infty$ or $f \in C(T)$, $p=\infty$, that $\|P_r^{(\alpha,\beta)}[f]\|_p \leqq A_{\alpha\beta}\|f\|_p$. This is obvious for $C(T)$. Otherwise, let $1 \leqq p < \infty$, $1/p + 1/q = 1$, and let $f \in L^p(dw_\nu)$. Then

$$\begin{aligned}|P_r^{(\alpha,\beta)}[f](e^{i\theta})| &\leqq \int_{-\pi}^{\pi} |f(e^{i\theta})|\,|K_{\alpha\beta}(r,\theta,\phi)\overline{c_{\alpha\beta}}(\phi)|\,dw_\nu(\phi)\\ &\leqq \left(\int_{-\pi}^{\pi} |f(e^{i\phi})|^p\,|K_{\alpha\beta}(r,\theta,\phi)\overline{c_{\alpha\beta}}(\phi)|\,dw_\nu(\phi)\right)^{1/p}\cdot\\ &\quad\cdot\left(\int_{-\pi}^{\pi} |K_{\alpha\beta}(r,\theta,\phi)\overline{c_{\alpha\beta}}(\phi)|\,dw_\nu(\phi)\right)^{1/q}, \quad \text{and so}\end{aligned}$$

$$\begin{aligned}&\int_{-\pi}^{\pi} |P_r^{(\alpha,\beta)}[f](e^{i\theta})|^p\,dw_\nu(\theta)\\ &\quad\leqq A_{\alpha\beta}^{p/q}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} |f(e^{i\phi})|^p\,|K_{\alpha\beta}(r,\theta,\phi)\bar{c}_{\alpha\beta}(\theta)|\,dw_\nu(\theta)\,dw_\nu(\phi)\\ &\quad\leqq (A_{\alpha\beta})^{1+p/q}\|f\|_p^p, \quad \text{the required bound.}\end{aligned}$$

Now for $f \in L^p(dw_\nu)$ or $f \in C(T)$ and $\varepsilon > 0$ there exists a trigonometric polynomial g such that $\|f-g\|_p < \varepsilon$. Further

$$\begin{aligned}\|f-P_r^{(\alpha,\beta)}[f]\|_p &\leqq \|f-g\|_p + \|g-P_r^{(\alpha,\beta)}[g]\|_p + \|P_r^{(\alpha,\beta)}[g-f]_p\|\\ &\leqq (1+A_{\alpha\beta})\,\varepsilon + \|g-P_r^{(\alpha,\beta)}[g]\|_\infty,\end{aligned}$$

and by Proposition 4.3 the latter goes to 0 as $r \to 1-$. □

The condition $|\gamma| \leqq 1$ is necessary for convergence of the Poisson integral on $C(T)$.

4.14. Proposition. *For* $\nu \geqq \gamma > 1$, $0 < \phi < \pi$,

$$\lim_{r\to 1-} \int_{-\phi_2(\phi)}^{-\phi_1(\phi)} |K_{\alpha\beta}(r,\theta,\phi)|\,dw_\nu(\theta) = \infty.$$

Proof. In the expression (4.2) replace θ by $-\theta$ and note

$${}_2F_1(\alpha, \alpha+1; 2\nu+1; X(r,\theta,\phi)) \geqq 1.$$

Also

$$1-2r\cos(\theta+\phi)+r^2 \leqq 1+2r+r^2 \leqq 4 \quad \text{and} \quad \sin^{2\nu}\theta \geqq 2^{-2\nu}\sin^{2\nu}\phi$$

on $\phi_1 \leqq \theta \leqq \phi_2$ (Lemma 4.9(1)). Thus

$$\int_{-\phi_2(\phi)}^{-\phi_1(\phi)} |K_{\alpha\beta}(r,\theta,\phi)|\,dw_\nu(\theta) \geqq B'_\gamma \sin^{2\nu}\phi(1-r) \int_0^{\phi_2-\phi} (1-2r\cos\theta+r^2)^{-\gamma}\,d\theta$$
$$= B'_\gamma \sin^{2\nu}\phi(1-r) \int_0^{2\sin((\phi_2-\phi)/2)} ((1-r)^2+r x^2)^{-\gamma}(1-x^2/4)^{-\frac{1}{2}}\,dx$$
$$\geqq B'_\gamma \sin^{2\nu}\phi\, r^{-\frac{1}{2}}(1-r)^{2-2\gamma} \int_0^{y_0} (1+y^2)^{-\gamma}\,dy,$$

(B'_γ is some constant) where $x=2\sin(\theta/2)$, $y=\sqrt{r}\,x/(1-r)$, $y_0=2\sin((\phi_2-\phi)/2)\sqrt{r}/(1-r)$.

For each ϕ there exists $r_0>\frac{1}{2}$ so that $y_0>1$ for $r_0<r<1$, that is, the last integral $\geqq \int_0^1 (1+y^2)^{-\gamma}\,dy$ and so the last expression $\to\infty$ as $r\to 1-$ (recall $\gamma>1$). □

4.15. Corollary. *For $\nu\geqq|\gamma|>1$, $0<|\phi|<\pi$ there exists $f\in C(T)$ such that $P_r^{(\alpha,\beta)}[f](e^{i\phi})$ is divergent (unbounded) as $r\to 1-$.*

Proof. The linear functional $f\mapsto P_r^{(\alpha,\beta)}[f](e^{i\phi})$ on $C(T)$ has norm

$$\int_{-\pi}^{\pi} |K_{\alpha\beta}(r,\theta,\phi)\,\overline{c_{\alpha\beta}}(\theta)|\,dw_\nu(\theta),$$

and the conclusion follows from the Banach-Steinhaus theorem. □

§ 5. Concluding Remarks

Since trigonometric polynomials are the boundary values of their Poisson integrals for any values of α,β with $\alpha,\beta\geqq 0$ and $\alpha+\beta>0$, one suspects that there are radial limit theorems for appropriate Sobolev-type norms for values of α,β with $|\alpha-\beta|>2$.

The big question is what can be said about functions expressible as $\sum_{n=0}^{\infty} a_n C_n^{(\alpha,\beta)}(z)$ alone. Let W be the space of functions in $C(T)$ (or some $L^p(dw_\nu)$, $1\leqq p<\infty$) whose Poisson integrals are sums like $\sum_{n=0}^{\infty} a_n C_n^{(\alpha,\beta)}(z)$. Consider the operator E which for each $f\in W$ maps $f|(0,\pi)$ to $f|(-\pi,0)$. For $\gamma=0$ this situation is marvelously simple: W is exactly the space of even functions, the domain of E is $C([0,\pi])$ (or L^p) and E is an isometry. Is the situation for γ close to 0 approximately like this? On the other hand, E is unbounded in the extreme cases $\alpha=0$ or $\beta=0$ ($\nu>0$).

Another topic of interest is the problem of $L^p(dw_\nu)$-boundedness for conjugate series, defined in § 1. These problems may well be related.

References

1. Bailey, W.: Generalized Hypergeometric Series. Cambridge: Cambridge University Press 1935
2. Dunkl, C.F.: An addition theorem for Heisenberg harmonics. In: Conference on Harmonic Analysis in Honor of Antoni Zygmund, pp. 690–707. Wadsworth International 1983

3. Folland, G.B.: The tangential Cauchy-Riemann complex on spheres. Trans. Amer. Math. Soc. **171**, 83–133 (1972)
4. Folland, G.B., Stein, E.M.: Estimates for the $\bar{\partial}_b$-complex and analysis on the Heisenberg group. Comm. Pure and Appl. Math. **27**, 429–522 (1974)
5. Gasper, G.: Orthogonality of certain functions with respect to complex valued weights. Canad. J. Math. **33**, 1261–1270 (1981)
6. Gaveau, B.: Principe de moindre action, propagation de chaleur et estimées sous elliptiques sur certains groupes nilpotents. Acta Math. **139**, 95–153 (1977)
7. Greiner, P.C.: Spherical harmonics on the Heisenberg group. Canad. Math. Bull. **23**, 383–396 (1980)
8. Greiner, P.C., Koornwinder, T.H.: Variations on the Heisenberg spherical harmonics. Report ZW 186/83, Math. Centrum Amsterdam (preprint)
9. Greiner, P.C., Stein, E.M.: Estimates for the $\bar{\partial}$-Neumann problem. Mathematical Notes, Princeton: Princeton University Press 1977
10. Jerison, D.S.: The Dirichlet problem for the Kohn Laplacian on the Heisenberg group. I, II. J. Functional Anal. **43**, 97–142, 224–257 (1981)
11. Korányi, A.: Kelvin transforms and harmonic polynomials on the Heisenberg group. J. Functional Anal. **49**, 177–185 (1982)
12. Muckenhoupt, B., Stein, E.M.: Classical expansions and their relation to conjugate harmonic functions. Trans. Amer. Math. Soc. **118**, 17–92 (1965)

Received December 5, 1983; received in final form April 13, 1984

Math. Z. 187, 549–558 (1984)

Mathematische Zeitschrift

New Minimality Properties of Gaussian Quadratures

Hermann Engels and Bernd Anton Merschen
Institut für Geometrie und Praktische Mathematik der TH, Templergraben 55, D-5100 Aachen, Federal Republic of Germany

1. Introduction

Extremal properties of Gaussian quadrature formulae

$$Q_n f = \sum_{i=1}^{n} A_i f(x_i)$$

(we shortly write A_i, x_i instead of $A_i^{(n)}, x_i^{(n)}$) are well known in the theory of quadrature. The most familiar: $Q_n f$ is the only formula of highest possible degree $2n-1$ and the error coefficient τ_n of the error representation

$$E_n f := \int_{-1}^{1} W(x) f(x) dx - Q_n f = \frac{\tau_n}{(2n)!} f^{(2n)}(\xi), \qquad \xi \in (-1, 1),$$

is minimal. Since

$$\tau_n = \int_{-1}^{1} W(x) \omega_n^2(x) dx = \|\omega_n\|_2^2,$$

where $\omega_n(x) = (x - x_1) \dots (x - x_n)$ is the nodal polynomial of $Q_n f$, this is due to the fact that $\omega_n \perp \mathbb{P}_{n-1}$, and so $\|\omega_n\|_2^2$ is the unique minimum of $\|\tilde{\omega}_n\|_2^2$ among all polynomials $\tilde{\omega}_n \in \mathbb{P}_n$ (with the same leading coefficient):

$$\|\omega_n\|_2^2 \leqq \|\tilde{\omega}_n\|_2^2.$$

Here $W(x) \geqq 0$ is an integrable weight-function.

We now prove two new minimal properties of Gaussian formulae which are of considerable numerical importance, too. For this we compare the Gaussian formulae $Q_n f$ with their infinitely many dual formulae $\bar{Q}_n f$ (cf. [4, 7, 8]) of almost highest possible degree $2n-2$. Definition and some properties are in Sect. 2, the main result is in Sect. 3. Finally some comments and numerical applications are presented in the last Sect. 4.

There is no uniform notation in the literature for minimal or optimal quadratures. Here we follow the notation used by Eckhardt [1], and denote a quadrature formula as a minimal formula if the necessary conditions for an extremal value is fulfilled. If we can prove that the necessary conditions lead to a minimal value we call the corresponding quadrature formula an optimal formula.

2. Dual Quadratures

Dual quadrature formulae are closely connected with Markoff's approach to Gaussian formulae (see [6]). Markoff approximates $f(x)$ by the Hermitian interpolating polynomial $H_n(f,x) \in \mathbb{P}_{2n-1}$ which interpolates to $f(x)$ and $f'(x)$ at the same nodes $x_1, \dots, x_n$. On integration $Q_n f$ is obtained. In $Q_n f$ the weights of $f'(x_i)$ vanish and this is equivalent to $\omega_n \perp \mathbb{P}_{n-1}$ with respect to $W(x)$. Engels [3] pointed out that an alternative approach is to approximate $f(x)$ by $H_n'(F,x) \in \mathbb{P}_{2n-2}$, where

$$F(x) = \int_{-1}^{x} f(t)\,dt.$$

Now on integration we obtain

$$\bar{Q}_n f = \sum_{i=1}^{n} \bar{A}_i f(\bar{x}_i).$$

Here the weights of $F(\bar{x}_i)$ vanish thus determining $\bar{x}_1, \dots, \bar{x}_n$, not uniquely, however. So we have that

$$\bar{Q}_n f \equiv \int_{-1}^{1} W(x) H_n'(F,x)\,dx = \sum_{i=1}^{n} (\bar{A}_i f(\bar{x}_i) + \bar{B}_i F(\bar{x}_i))$$

and

$$\bar{B}_i = 0, \quad i = 1(1)n.$$

Here, by definition of $H_n(\cdot, x)$ (cf. [2]):

$$\bar{A}_i = \int_{-1}^{1} W(x)[(x - \bar{x}_i)\lambda_i^2(x)]'\,dx,$$

$$\bar{B}_i = \int_{-1}^{1} W(x)\{[1 - 2\lambda_i'(\bar{x}_i)(x - \bar{x}_i)]\lambda_i^2(x)\}'\,dx, \qquad i = 1(1)n,$$

where $\lambda_i(x)$ is a cardinal polynomial of the Lagrange interpolating polynomial

$$L_n(f,x) \equiv \sum_{i=1}^{n} f(\bar{x}_i)\lambda_i(x) \in \mathbb{P}_{n-1}. \tag{2.1}$$

Definition 1. Every quadrature formula

$$\bar{Q}_n f = \int_{-1}^{1} W(x) H_n'(F,x)\,dx \quad \text{with } \bar{B}_i = 0,\ i = 1(1)n,$$

is called a *dual quadrature formula* of $Q_n f$.

$Q_n f$ and $\bar{Q}_n f$ are called dual because they are closely related by construction, their nodes and weights are denoted as x_i, A_i and $\bar{x}_i, \bar{A}_i$ respectively. Dual quadrature formulae are first considered in [3] and discussed in detail in [7] (see also [8]). As pointed out in [7] the following theorem is valid:

Theorem 1. *The nonlinear system of equations $\bar{B}_i=0$, $i=1(1)n$, for the nodes $\bar{x}_1,\dots,\bar{x}_n$ has infinitely many real solutions.*

The corresponding dual formulae $\bar{Q}_n f$ of degree $2n-2$ have positive weights $\bar{A}_i>0$, their n real nodes are distinct, and at least $n-1$ nodes are in $(-1,1)$.

Remark 1. Since $H_n'(F,x)\in\mathbb{P}_{2n-2}$, it is to be expected that $\bar{Q}_n f$ is not uniquely determined, since the number $2n$ of unknown nodes and weights remains the same as for $Q_n f$. Hence the system $\bar{B}_i=0$ is redundant, and in fact it is equivalent to

$$\omega_n \perp \mathbb{P}_{n-2} \qquad \text{with respect to } W(x)$$

where $\omega_n(x)=(x-\bar{x}_1)\dots(x-\bar{x}_n)$. We can prove this equivalence by performing the differentiation in $\bar{B}_i$, $i=1(1)n$, taking into account that

$$\omega_n(x)=\kappa_{n,i}(x-\bar{x}_i)\,\lambda_i(x), \qquad \kappa_{n,i}=\omega_n'(\bar{x}_i)=\text{const},$$

and that the polynomial $\lambda_i'(x)-\lambda_i'(\bar{x}_i)\,\lambda_i(x)\in\mathbb{P}_{n-1}$ possesses a simple zero at $x=\bar{x}_i$ (cf. [7] or [8]).

Remark 2. Since a fortiori $\omega_n(x)=(x-x_1)\dots(x-x_n)\perp\mathbb{P}_{n-1}$ we see that the Gaussian formula $Q_n f$ is dual to itself.

Remark 3. If we impose an additional (local or global) condition, we can expect a *unique* formula $\bar{Q}_n f$. We have a *local condition* if we preassign one of the weights $\bar{A}_1,\dots,\bar{A}_n$ or one of the nodes $\bar{x}_1,\dots,\bar{x}_n$, say $\bar{x}_1=z$. The remaining nodes $\bar{x}_2,\dots,\bar{x}_n$ then are the zeros of the orthogonal polynomial $q_{n-1}(x)=(x-\bar{x}_2)\dots(x-\bar{x}_n)\perp\mathbb{P}_{n-2}$ with respect to the weight function $\tilde{W}(x):=W(x)(x-z)$ (cf. [5]). If $z\notin(-1,1)$ we still have $\tilde{W}(x)\geqq 0$ and hence no difficulty occurs. The choice $|z|=1$ corresponds to the well-known Radau-case.

If, however, $z\in(-1,1)$, the classical theory of orthogonal polynomials fails. It no longer assures the existence of $n-1$ real, distinct zeros of q_{n-1} in $(-1,1)$ which is needed. This gap is closed in [7] and intervals $I_k\subset(-1,1)$ are specified there such that q_{n-1} indeed possesses $n-1$ real, distinct zeros in $(-1,1)$, so that these are the nodes of $\bar{Q}_n f$ together with the preassigned node z.

A global condition is e.g. *symmetry of $\bar{Q}_n f$ for an even weight-function $W(x)$.* In this case we have

Lemma 1. *If $W(x)=W(-x)$, then the only symmetric solution*

$$\bar{x}_i=-\bar{x}_{n-i+1}, \qquad i=1(1)n,$$

of $\bar{B}_i=0$, $i=1(1)n$, is the set of the Gaussian nodes x_i, $i=1(1)n$, (cf. [7]).

The same solution occurs if we impose the following two *global conditions*:

I. Determine $\bar{Q}_n f$ such that the variance of $\bar{Q}_n f$ is minimal.

II. Determine $\bar{Q}_n f$ such that the diameter of $\bar{Q}_n f$ is minimal.

These two problems are treated in the following Sect. 3.

3. The New Minimal Properties

For convenience we define for any given quadrature formula $Q_n f$ with nodes x_i and weights A_i, $i=1(1)n$:

Definition 2. I. $V_n \equiv \sqrt{\sum_{j=1}^{n} A_j^2}$ is called the *variance* of $Q_n f$;

II. $d_n \equiv \max_{1 \leqq i,k \leqq n} |x_i - x_k|$ is called the *diameter* of $Q_n f$.

For simplicity we assume in the sequel that $x_1 < x_2 < \ldots < x_n$ and $\bar{x}_1 < \bar{x}_2 < \ldots < \bar{x}_n$. In this case $d_n = x_n - x_1$.

Remark. The variance has practical importance for the integration of erratic functions (cf. [2]).

Theorem 2. *Let* $W(x) = W(-x)$ *for* $x \in [-1,1]$. *Then the Gaussian formula* $Q_n f$ *possesses the two minimal properties*

I. $\min_{\{\bar{x}_1, \ldots, \bar{x}_n\}} \bar{V}_n^2 = V_n^2 = V_n^2(x_1, \ldots, x_n)$;

II. $\min_{\{\bar{x}_1, \ldots, \bar{x}_n\}} \bar{d}_n = d_n = d_n(x_1, \ldots, x_n)$.

Proof. While II. is already proved in [7] by help of an extension of orthogonal polynomial theory, we now present a uniform approach for I. and II. via the theory of Lagrangian multipliers. This method does not only provide necessary conditions for the minimisation but also the position of a minimum if there is any at all. The minimisation is constrained to the set of nodes of dual formulae $\bar{Q}_n f$, which means that the nodes $\bar{x}_i$ must satisfy

$$\bar{B}_i = \bar{B}_i(\bar{x}_1, \ldots, \bar{x}_n) = 0, \qquad i = 1(1)n-1.$$

This is equivalent to $\omega_n \perp \mathbb{P}_{n-2}$ with respect to $W(x)$, as already pointed out.

The success of the method of Lagrangian multipliers strongly depends on the judicious choice of the basis of $\mathbb{P}_{n-2}$. Otherwise we obtain a hopelessly complicated system of equations which cannot be solved analytically for arbitrary n. The space $\mathbb{P}_{n-2}$ is spanned by the $n-1$ polynomials

$$\varphi_i(x) \equiv \frac{\lambda_1(x)}{x - \bar{x}_{i+1}} \in \mathbb{P}_{n-2}, \qquad i = 1(1)n-1,$$

where

$$\lambda_1(x) = \frac{\omega_n(x)}{\omega_n'(\bar{x}_1)(x - \bar{x}_1)} \in \mathbb{P}_{n-1}$$

is the first cardinal polynomial of the Lagrange interpolating polynomial (2.1) of $f(x)$.

Hence $\omega_n \perp \mathbb{P}_{n-2}$ is equivalent to

$$\int_{-1}^{1} W(x)\, \omega_n(x)\, \varphi_i(x)\, dx = 0, \qquad i = 1(1)n-1,$$

and in turn to

$$\bar{B}_i = 0, \qquad i = 1(1)n-1,$$

so that we are allowed to replace the original constraints $\bar{B}_i = 0$ by

$$\bar{B}_i \equiv \int_{-1}^{1} W(x) \lambda_1(x) \lambda_{i+1}(x)\, dx = 0, \qquad i = 1(1)\, n-1. \tag{3.1}$$

We introduce the real Lagrangian multipliers v_i and μ_i, $i=1(1)n-1$, and thus have to minimise the two functions

$$G(\bar{x}_i, \mu_i) \equiv \bar{V}_n^2(\bar{x}_1, \ldots, \bar{x}_n) + \sum_{i=1}^{n-1} \mu_i \bar{B}_i(\bar{x}_1, \ldots, \bar{x}_n)$$

and

$$H(\bar{x}_i, v_i) \equiv \bar{d}_n(\bar{x}_1, \ldots, \bar{x}_n) + \sum_{i=1}^{n-1} v_i \bar{B}_i(\bar{x}_1, \ldots, \bar{x}_n).$$

The necessary conditions for a minimum of the functions $G(\bar{x}_i, \mu_i)$ and $H(\bar{x}_i, v_i)$ are

$$\frac{\partial}{\partial \bar{x}_k} G(\bar{x}_i, \mu_i) = 0, \qquad \frac{\partial}{\partial \mu_j} G(\bar{x}_i, \mu_i) = 0 \tag{3.2}$$

and

$$\frac{\partial}{\partial \bar{x}_k} H(\bar{x}_i, v_i) = 0, \qquad \frac{\partial}{\partial v_j} (\bar{x}_i, v_i) = 0, \tag{3.3}$$

$$k = 1(1)\, n, \qquad j = 1(1)\, n-1.$$

Here the partial derivatives with respect to the Lagrangian multipliers are identical with the constraints (3.1).

The fundamental relation for obtaining an acceptable form of the partial derivatives is

Lemma 2.

$$\frac{\partial \lambda_j(x)}{\partial \bar{x}_k} = -\lambda_j'(\bar{x}_k)\, \lambda_k(x) \equiv c_{j,k} \lambda_k(x), \qquad j, k = 1(1)\, n.$$

We can prove this relation by calculation.

As the quadrature formulae under consideration are interpolatory the weights $\bar{A}_i$ are completely determined by the nodes $\bar{x}_i$ (cf. [5]) and thus they have the representation

$$\bar{A}_i = \int_{-1}^{1} W(x) \lambda_i(x)\, dx.$$

An immediate consequence of Lemma 2 is that

$$\frac{\partial \bar{A}_i}{\partial \bar{x}_k} = c_{i,k} \int_{-1}^{1} W(x) \lambda_k(x)\, dx = c_{i,k} \bar{A}_k, \qquad i, k = 1(1)\, n.$$

Further it follows from (3.1) that

$$\left.\frac{\partial \bar{B}_i}{\partial \bar{x}_k}\right|_{\bar{B}_i = 0} = \begin{cases} c_{i+1,1} \bar{A}_1 & \text{for } k = 1 \\ 0 & \text{for } i+1 \neq k > 1 \\ c_{1,k} \bar{A}_k & \text{for } i+1 = k > 1 \end{cases}$$

where we used the fact, that if (3.1) is fulfilled, then the weights $\bar{A}_i$ may be written as (cf. [7])

$$\bar{A}_i = \int_{-1}^{1} W(x)\lambda_i^2(x)\,dx, \qquad i=1(1)n.$$

Now we are able to calculate explicitly the necessary conditions (3.2) and (3.3). The following equations occur

$$\left\{\begin{aligned} &2\sum_{i=1}^{n} c_{i,1}\bar{A}_i + \sum_{i=1}^{n-1} \mu_i c_{i+1,1} = 0 \\ &\quad 2\sum_{i=1}^{n} c_{i,k}\bar{A}_i + \mu_{k-1}c_{1,k} = 0, \qquad k=2(1)n, \end{aligned}\right\} \tag{3.4}$$

and

$$\left\{\begin{aligned} &-1+\bar{A}_1\sum_{i=1}^{n-1} \nu_i c_{i+1,1} = 0 \\ &\qquad\quad \bar{A}_k \nu_{k-1} c_{1,k} = 0, \qquad k=2(1)n-1, \\ &\quad 1+\bar{A}_n \nu_{n-1} c_{1,n} = 0 \end{aligned}\right\} \tag{3.5}$$

(from (3.2) and (3.3) respectively).

Because of the very special shape of (3.4) and (3.5) we can eliminate the Lagrangian multipliers μ_i and ν_i respectively. This elimination gives:

$$\sum_{i=1}^{n} \bar{A}_i\left(c_{i,1} - \sum_{j=2}^{n} \frac{c_{j,1}}{c_{1,j}} c_{i,j}\right) = 0,$$

and

$$-1 - \frac{c_{n,1}}{c_{1,n}} \frac{\bar{A}_1}{\bar{A}_n} = 0.$$

We replace the coefficients

$$c_{i,j} = -\lambda_i'(\bar{x}_j) = \begin{cases} -\dfrac{\omega_n'(\bar{x}_j)}{\omega_n'(\bar{x}_i)(\bar{x}_j - \bar{x}_i)} & \text{for } i \neq j, \\[2ex] -\dfrac{\omega_n''(\bar{x}_i)}{2\omega_n'(\bar{x}_i)} & \text{for } i=j, \end{cases}$$

so that we finally obtain

$$\sum_{i=1}^{n} \bar{A}_i \frac{\omega_n''(\bar{x}_i)}{[\omega_n'(\bar{x}_i)]^3} = 0, \tag{3.6}$$

and

$$1 - \left[\frac{\omega_n'(\bar{x}_1)}{\omega_n'(\bar{x}_n)}\right]^2 \frac{\bar{A}_1}{\bar{A}_n} = 0, \tag{3.7}$$

respectively.

While (3.7) is obvious, the determination of (3.6) is a bit more complicated. Since $L_n(1,x)=1$ identically we have that

$$1 = \sum_{j=1}^{n} \lambda_j(x)$$

and hence

$$\sum_{j=1}^{n} \lambda_j'(\bar{x}_i)=0, \qquad i=1(1)n,$$

or

$$\sum_{j=1}^{n} c_{j,i}=0, \qquad i=1(1)n.$$

On the other hand we can show by calculation that

$$c_{i,1}-\sum_{j=2}^{n} \frac{c_{j,1}}{c_{1,j}} c_{i,j}=-\left[\frac{\omega_n'(\bar{x}_1)}{\omega_n'(\bar{x}_i)}\right]^2 \left\{\sum_{j=1}^{n} c_{j,i}-\frac{\omega_n''(\bar{x}_i)}{\omega_n'(\bar{x}_i)}\right\}.$$

This establishes (3.6).

So finally the problem is to find a set of nodes $\bar{x}_1, \ldots, \bar{x}_n$ which simultaneously solves the system of equations $\bar{B}_i=0$, $i=1(1)n-1$, and one of the equations (3.6) or (3.7). Especially each symmetric quadrature formula satisfies conditions (3.6) and (3.7):

Symmetry means

$$\left.\begin{array}{c} \bar{A}_i=\bar{A}_{n-i+1} \\ \text{and} \qquad \bar{x}_i=-\bar{x}_{n-i+1} \end{array}\right\}, \qquad i=1(1)n,$$

so that

$$\omega_n^{(i)}(x)=(-1)^{n+i}\omega_n^{(i)}(-x), \qquad i=0(1)n.$$

Hence for even n everything is obvious.

For odd $n=2m+1$, $\bar{x}_{m+1}=0$ is a node, and it is clear by construction that $\omega_n'(0)\neq 0$ and $\omega_n''(0)=0$.

On the other hand the only symmetric solution of $\bar{B}_i=0$, $i=1(1)n-1$, is the Gaussian formula $Q_n f$ as is seen from Lemma 1. This completes the proof.

Remark. We only proved that the Gaussian formula $Q_n f$ satisfies the necessary conditions for minimality in I. and II. The method of Lagrangian multipliers is not well suited for the proof of sufficiency. For II. sufficiency has been proved in [7] using completely different methods which cannot be applied to prove the sufficiency of I.

The proof of the sufficiency for I. is an open problem; numerical calculations, however, indicate that nothing else is to be expected.

4. Comments and Examples

We comment on Theorem 2 with regard to its assumptions and its practical importance.

4.1 Symmetry. The symmetry of $W(x)$ is an essential assumption. If it is omitted there exist counterexamples such that the solution of (3.1, 3.2) and (3.1, 3.3) neither are identical with $Q_n f$ nor mutually. Such an example is $W(x) = 1+x$ and $n=3$.

Table 1

Formula	Nodes	Weights	V_n^2	d_n
Gaussian formula $Q_3 f$	-0.575319 0.181066 0.822824	0.279308 0.916964 0.803728	1.250926	1.398143
dual formula $\bar{Q}_3 f$ with minimal d_n	-0.517054 0.281082 0.863331	0.353132 0.983319 0.663549	1.237706	1.380385
dual formula $\bar{Q}_3 f$ with minimal V_n^2	-0.524015 0.267302 0.856713	0.343518 0.971232 0.685250	1.237281	1.380729

So for a not even weight-function $W(x)$ we in general obtain new dual quadrature formulae $\bar{Q}_n f$ of almost highest possible degree $2n-2$ which in view of applications have interesting properties.

4.2 Numerical Integration of Erratic Functions. Assume that the values of the integrand $f(x)$ are erratic, i.e. instead of the correct function values $f(x_i)$ we only know the erratic values

$$\tilde{f}(x_i)=f(x_i)+\varepsilon_i, \qquad i=1(1)n.$$

The errors ε_i are considered to be due to measuring or rounding. Then the total error is estimated by

$$|\tilde{E}_n f| \leqq |E_n f| + \left(\sum_{i=1}^{n} A_i^2\right)^{1/2} \left(\sum_{i=1}^{n} \varepsilon_i^2\right)^{1/2}, \quad \text{(cf. [2])}.$$

The errors ε_i usually are not at our disposal. Hence we try to minimize the variance

$$V_n^2 = \sum_{i=1}^{n} A_i^2.$$

This minimisation is performed with some constraints concerning the exactness of the required formula.

It is well known (cf. [2]) that subject to the condition degree $(Q_n f) \geqq 0$, V_n^2 is minimal iff the weights are equal,

$$A_i = \frac{m_0}{n}, \quad i=1(1)n, \qquad \text{with } m_0 \equiv \int_{-1}^{1} W(x)\,dx.$$

Theorem 2 now assures the minimisation of V_n^2 subject to the constraint degree$(\bar{Q}_n f) \geqq 2n-2$ which in this case is the highest possible degree. The statement is that for symmetric weight functions the solution is the Gaussian quadrature formula: $\bar{Q}_n f = Q_n f$. This result emphasizes the stability of Gaussian quadrature rules.

We now measure the deviation of an arbitrary quadrature formula from an equally weighted formula by

$$\sigma_*^2 := \frac{1}{n-1} \sum_{i=1}^{n} (A - A_i)^2, \qquad A := \frac{1}{n} \sum_{j=1}^{n} A_j.$$

Hence the Gaussian formula $Q_n f$ also minimises σ_n^2 among all dual quadrature formulae $\bar{Q}_n f$ for symmetric weight-functions. This is to be seen from

$$\sigma_n^2 = \frac{1}{n-1}\left\{V_n^2 - \frac{m_0}{n}\right\}.$$

4.3 Numerical Integration of Functions with Endpoint Singularities. Let us be given an even weight-function $W(x)$. Then the nodes of Gaussian quadratures are symmetric. Now the extremal property II. means that for fixed $n \in \mathbb{N}$ the Gaussian nodes are most strongly concentrated in the middle of the interval of integration (compared with all node distributions of the dual formulae $\bar{Q}_n f$).

This property impressively confirms the numerical experience that Gaussian formulae in general do not work well for functions which have integrable endpoint singularities or singularities in the immediate neighbourhood of the interval of integration.

Integrands of this kind need quadrature formulae such that the nodes cluster at the endpoints of $[-1, 1]$. Such are e.g. the Wilf-formulae (cf. [2]) or their duals (cf. [7]). We give a small but impressive example:

Table 2

$W(x) \equiv 1$ $n = 16$	Gaussian formula	dual Wilf-formula, $\varepsilon = 0.999$
nodes in $[-1; -0.9] \cup [0.9; 1]$	$x_1, x_2; x_{15}, x_{16}$	$x_1, \ldots, x_5; x_{12}, \ldots, x_{16}$
relative error for $\int_{-1}^{1} \frac{1}{1.01 - x}\,dx$	$9.92 \cdot 10^{-3}$	$2.35 \cdot 10^{-7}$

4.4 Numerical Evaluation of Cauchy Principal Value Integrals. We mention another minimal property of Gaussian quadratures which is of interest for the numerical evaluation of Cauchy principal value integrals. For stability reasons van der Sluis (cf. [9, 10]) proposed to minimise in some sense the quantity

$$s_n \equiv \sum_{j=1}^{n} \left|\frac{A_j}{x_j}\right|,$$

given a constant weight-function on $[-1, 1]$ or $[0, 1]$ alternately. A detailed discussion is beyond the scope of this note and is postponed to a forthcoming paper. We only state the following theorem subject to the same assumptions as for Theorem 2.

Theorem 3. *Let* $W(x) \equiv 1$, $n \in \mathbb{N}$ *even.*

i) *For the interval of integration* $[-1, 1]$ *the set of nodes of the Gaussian quadrature minimises* s_n.

ii) *For* $[0, 1]$ *there must not exist a real solution at all.*

The proof follows that of Theorem 2, in case of ii) we can show that the corresponding equations do not have a real solution already for $n = 2$.

References

1. Eckhardt, U.: Einige Eigenschaften Wilf'scher Quadraturformeln. Numer. Math. **12**, 1–7 (1968)
2. Engels, H.: Numerical quadrature and cubature. London: Academic Press 1980
3. Engels, H.: Variation eines Themas von Gauß-Markoff. Vortrag GAMM-Tagung 1981, Würzburg
4. Engels, H., Merschen, B.A.: Duale Quadraturen. In: Hämmerlin, G. (ed.): Numerische Integration, ISNM 57, pp. 48–54. Basel: Birkhäuser 1982
5. Krylov, V.J.: Approximate calculation of integrals, New York: ACM Monograph Series 1962
6. Markoff, A.: Sur la méthode de Gauss pour le calcul approché des intégrales. Math. Ann. **25**, 427–432 (1885)
7. Merschen, B.A.: Duale Quadraturformeln und ihre Eigenschaften. Dissertation, Aachen 1983
8. Schneider, C.: Dual quadratures. To appear in: Math. Z.
9. Sluis, A. van der: Numerical quadrature and cubature; Comments, Problems. Vortrag DMV-Tagung 1983, Köln
10. Sluis, A. van der, Zweerus, J.R.: An appraisal of some methods for computing Cauchy principal values of integrals. In: Hämmerlin, G. (ed.): Numerische Integration, ISNM 45, pp. 264–287. Basel: Birkhäuser 1979

Received December 5, 1983

Math. Z. 187, 559–562 (1984)

Mathematische Zeitschrift

Domains with Finite Dimensional Bergman Space

Jan J.O.O. Wiegerinck*

Math. Institute, Univ. of Amsterdam, Roetersstraat 15, Amsterdam, The Netherlands

1. Introduction and Result

Let Ω be a domain in $\mathbb{C}^n$ and $L^2H(\Omega)$ the Bergman space of Ω, that is:

$$L^2H(\Omega)=\{f \text{ holomorphic in } \Omega: \\ (\int_\Omega |f(z)|^2 dV(z) f \equiv \|f\|_{L^2H(\Omega)} < \infty\}.$$

It is well-known [1] that $L^2H(\Omega)$ is a closed subspace of $L^2(\Omega)$. For a bounded domain the dimension of the Bergman space is infinite, because in that case all polynomials are in L^2H. The dimension is also infinite for domains that are holomorphically equivalent to bounded domains. On the other hand $L^2H(\mathbb{C}^n)=\{0\}$ for all n. This raises the following *question: Do domains with finite dimensional Bergman space exist?*

In this note we will show that the answer is *yes*. In fact we have the following.

Theorem. *For every $k>0$ there exists a Reinhardt domain in $\mathbb{C}^2$ with k-dimensional Bergman space.*

Our construction is easily extended to $\mathbb{C}^n$ with $n\geqq 3$. However, for domains in $\mathbb{C}$ the dimension of L^2H is 0 or ∞ (Sect. 3).

M. Skwarczynski informed me that he has recently considered certain related questions, cf. [5].

2. Proof of the Theorem

The coordinates in $\mathbb{C}^2$ will be z and w. Define domains X_1 and X_2 by

$$X_1=\{(z,w)\in\mathbb{C}^2: |w|<1/(|z|\log|z|),\ |z|>e\}$$
$$X_2=\{(z,w)\in\mathbb{C}^2: |z|<1/(|w|\log|w|),\ |w|>e\}$$

* Work supported by the Netherlands' research organization Z.W.O.

and put

$$\Omega = X_1 \cup X_2 \cup \{(z,w) \in \mathbb{C}^2 : |z| < 2e,\ |w| < 2e\}.$$

Lemma 1. *The monomials contained in $L^2H(\Omega)$ are precisely*

$$a z^k w^k, \quad a \in \mathbb{C}, \quad k = 0, 1, 2, \ldots .$$

Proof. We make a small computation:

$$\int_{X_1} |z|^{2p} |w|^{2q}\, dV = (2\pi)^2 \int_{r_1=e}^{\infty} \int_{r_2=0}^{\frac{1}{r_1 \log r_1}} r_1^{2p+1} r_2^{2q+1}\, dr_1\, dr_2$$

$$= (2\pi)^2 \int_e^{\infty} \frac{r_1^{2p-2q-1}}{(\log r_1)^{2q+2}}\, dr_1 .$$

The last integral is finite if and only if $q \geqq p$. In other words

$$a z^p w^q \in L^2 H(X_1) \Leftrightarrow q \geqq p. \tag{1}$$

In the same way we obtain

$$a z^p w^q \in L^2 H(X_2) \Leftrightarrow q \leqq p. \tag{2}$$

The lemma follows from (1) and (2). □

We will enlarge Ω slightly in the directions $|z| = |w|$ in order to dispose of the monomials with high exponents. Therefore put

$$B_m = \left\{ (z,w) : |z|, |w| > 1,\ \big| |z| - |w| \big| < \frac{1}{(|z| + |w|)^m} \right\}$$

and

$$\Omega_k = \Omega \cup B_{4k} \quad k = 1, 2, \ldots .$$

We will show that the domain Ω_k has k-dimensional Bergman space. It is sketched in the Fig. 1.

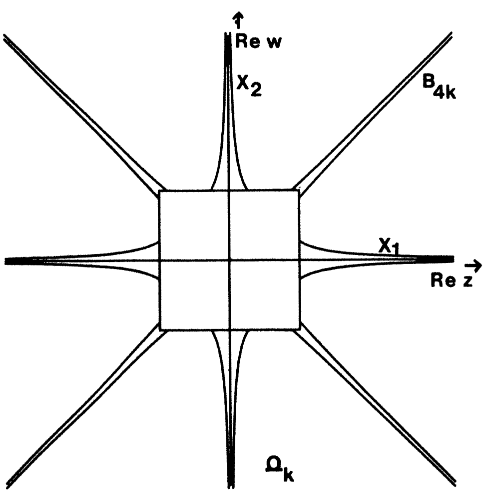


Fig. 1

Lemma 2. *The following assertion is true for* $k=0,1,2,\dots$:

$$z^p w^q \in L^2H(\Omega_k) \Leftrightarrow p=q<k.$$

Proof. We compute $\|z^p w^p\|^2_{L^2H(B_m)}$:

$$\|z^p w^p\|^2_{L^2H(B_m)} = \int_{B_m} |z^{2p} w^{2p}|\, dV = (2\pi)^2 \int_{|r_1-r_2|<\frac{1}{(r_1+r_2)^m},\, r_j>1} (r_1 r_2)^{2p+1}\, dr_1\, dr_2$$

$$\overset{r_1+r_2=t,\, r_1-r_2=s}{\approx} \int_2^\infty (2\pi)^2 \int_{-(1/t)^m}^{(1/t)^m} \frac{1}{2}\left(\frac{t^2-s^2}{4}\right)^{2p+1} ds\, dt.$$

The last integral converges if and only if $m=4k>4p+3$. Using Lemma 1 and this result we obtain Lemma 2. □

Proof of the Theorem. The logarithmically convex hull of the Reinhardt domain Ω_k is $\mathbb{C}^2$. It follows that every $f \in L^2H(\Omega_k)$ has a holomorphic extension to $\mathbb{C}^2$ (cf. [4]). Hence f has a power series expansion $\sum a_{nm} z^n w^m$ that converges to f uniformly on bounded sets. Let $B(0,R)$ be the ball in $\mathbb{C}^2$ with centre 0 and radius R. Now we estimate

$$\|f\|^2_{L^2H(\Omega_k)} \geqq \int_{\Omega_k \cap B(0,R)} |f|^2\, dV = \sum |a_{nm}|^2 \int_{\Omega_k \cap B(0,R)} |z^{2n} w^{2m}|\, dV. \tag{3}$$

The last equality follows from the fact that the monomials $z^p w^q$ form an orthogonal set on every bounded Reinhardt domain, as is easily seen using polar coordinates.

When we use Lemma 2 and let R tend to infinity in (3), we conclude that $a_{nm}=0$ unless $n=m<k$. Hence

$$L^2H(\Omega_k) = \operatorname{span}\{1, zw, \dots, (zw)^{k-1}\}. \quad \square$$

3. Domains in $\mathbb{C}$

Theorem. *For Ω a domain in $\mathbb{C}$ the dimension of $L^2H(\Omega)$ is 0 or ∞.*

Proof. After a Möbius transformation, which does not affect $\dim L^2H(\Omega)$, we can assume that Ω contains the point ∞. Suppose $f \not\equiv 0$ is in $L^2H(\Omega)$. We consider two cases:

Case 1. The function f is rational. In this case we observe that

$$\int_{\mathbb{C}} |f|^2\, dx\, dy = \infty, \qquad \int_{\Omega} |f|^2\, dx\, dy < \infty.$$

Hence the complement of Ω in $\mathbb{C}$, Ω^c, has positive 2-dimensional Lebesgue measure. It is well-known, cf. [3] that under these circumstances the Cauchy transform

$$g(z) = \int_{\Omega^c} \frac{1}{\zeta - z}\, d\xi\, d\eta, \qquad \zeta = \xi + i\eta,$$

is a non-constant bounded holomorphic function on Ω with the property that $g(\infty)=0$. It is clear that g^2, g^3, ... will be in $L^2H(\Omega)$, so $\dim L^2H(\Omega)=\infty$.

Case 2. The function f is non-rational. We expand f in a Laurent series around the point ∞:

$$f(z)=\sum_p^\infty c_k z^{-k}, \quad p\geqq 2, \quad c_p\neq 0.$$

We will construct a function $g\in L^2H(\Omega)$ such that $g\not\equiv 0$ and the Laurent series of g has no terms in $z^{-1},\dots,z^{-p}$. Let $z_1,\dots,z_{p+1}$ be different points in Ω. We take g of the form

$$g(z)=\sum_1^{p+1}\frac{b_j(f(z)-f(z_j))}{z-z_j}, \quad b_j\in\mathbb{C}.$$

Expanding around ∞ we find

$$g(z)=\sum_1^\infty a_k z^{-k},$$

with coefficients a_k depending on b_j, z_j and f. In fact, for $k=1,\dots,p$

$$a_k=\sum_{j=1}^{p+1} -b_j f(z_j)\, z_j^{k-1}.$$

Now we choose the constants b_j to be an - always existing - non-trivial solution of the homogeneous linear equations

$$a_k=0, \quad k=1,\dots,p.$$

We observe that $a_1=0$ and $f\in L^2H(\Omega)$ imply that $g\in L^2H(\Omega)$ and we also observe that g is non-rational (or else f would be rational). In particular $g\not\equiv 0$, hence for some $q>p$, $a_q\neq 0$. We can continue in this manner with g instead of f and infer that $\dim L^2H(\Omega)$ is infinite. □

For domains in $\mathbb{C}$ it is known that $L^2H(\Omega)$ contains nontrivial functions if and only if the complement of Ω has positive logarithmic capacity, cf. [2].

References

1. Bergman, S.: The kernel function and conformal mapping, 2nd ed. Providence: Amer. Math. Soc. 1970
2. Carleson, L.: Selected problems on exceptional sets. Princeton: Van Nostrand 1967
3. Garnett, J.: Analytic capacity and measure. Lecture Notes in Math. **297**. Berlin-Heidelberg-New York: Springer 1972
4. Hörmander, L.: An introduction to complex analysis in several variables. Amsterdam: North Holland 1973
5. Skwarczynski, M.: Evaluation functionals on spaces of square integrable holomorphic functions. In Prace matematyczno-fizyczne, M. Skwarczynski, W. Wasilewski editors, Wyżcza szkola inżynierska w Radomiu, Radom 1982

Received December 15, 1983; received in final form June 6, 1984

Math. Z. 187, 563–569 (1984)

Mathematische Zeitschrift

Correction to

Group Algebras of Finite Representation Type

Math. Z. **182**, 129–148 (1983)

Hagen Meltzer[1] and Andrzej Skowroński[2]

[1] Humboldt-Universität Berlin, Sektion Mathematik, DDR-1086 Berlin, German Democratic Republic
[2] Institute of Mathematics, Nicholas Copernicus University, PL-87-100 Toruń, Poland

Recently we have observed that Lemma 4 in [8] is not true and hence the characterization of group algebras of finite representation type given in [8] is not complete.

Here, we will give the correct version of [8, Theorem] and we will complete the proof given there.

We use the same notation as in [8]. In order to state the main result, we need some additional. Let F be a division algebra over a fixed field K. We will denote by $R_F(1)$ the quiver algebra FQ_1 (in the sense of [3]) of the quiver

$$Q_1:\ 1 \to 2 \leftarrow 3$$

and by $R_F(2)$ the opposite algebra to $R_F(1)$. Moreover, we will denote by $R_F(3)$ the bounden quiver algebra FQ_2/I_2 where Q_2 is the quiver

$$1 \xrightarrow{\alpha_1} 2 \xleftarrow{\alpha_2} 3 \xleftarrow{\alpha_3} 4$$

and I_2 is the ideal in the quiver algebra FQ_2 generated by $\alpha_2\alpha_3$, and by $R_F(4)$ the opposite algebra to $R_F(3)$. Finally, recall that for nonnegative integers m and n, $R_F(m, n)$ is the bounden quiver algebra $FQ_{m,n}/I_{m,n}$ where $Q_{m,n}$ is the quiver

$$\underset{-m}{\bullet} \xrightarrow{\beta_m} \underset{-m+1}{\bullet} \to \cdots \underset{-1}{\bullet} \xrightarrow{\beta_1} \underset{0}{\bullet} \xrightarrow{\alpha_1} \underset{1}{\bullet} \to \cdots \underset{n-1}{\bullet} \xrightarrow{\alpha_n} \underset{n}{\bullet}$$

and $I_{m,n}$ is the ideal in the quiver algebra $FQ_{m,n}$ generated by the composed arrows $\beta_i\beta_{i+1}$, $i=1,\ldots,m-1$, and $\alpha_{j+1}\alpha_j$, $j=1,\ldots,n-1$.

Theorem. *Let G be a finite group and let R be a connected finite-dimensional basic algebra over a field K of charcteristic p. Then RG is of finite representation type if and only if one to the following cases holds.*

(i) *p does not divide the order of G and R is of finite representation type.*

(ii) *p divides the order of G and one of the following holds:*

(1) *R is a division algebra and a p-Sylow subgroup of G is cyclic or*

(2) $p=3$, *a 3-Sylow subgroup of G is simple, and R is isomorphic to an algebra* $R_F(0,1)$ *or*

(3) $p=2$, *a 2-Sylow subgroup of G is simple, and R is isomorphic to an algebra* $R_F(m,n)$, *to an algebra* $R_F(i)$, $i=1, 2, 3, 4$, *or to a triangular matrix algebra* $\begin{pmatrix} F & {}_FM_H \\ 0 & H \end{pmatrix}$, *where F and H are division algebras,* ${}_FM_H$ *is an* $F-H$-*bimodule, K acts centrally on F, H, and* ${}_FM_H$, *and* $\dim_F({}_FM_H)\cdot\dim({}_FM_H)_H=2$.

Let R be a non-simple algebra over a fixed field K of characteristic $p>0$ and let G be a cyclic group of order $n=p^k$. From [8], for our aim, we can assume that R is isomorphic to a bounden quiver algebra FQ/I, where F is a division algebra and Q is of the form

$$Q:\ 1 \overset{\alpha_1}{—} 2 — \dots m-1 \overset{\alpha_{m-1}}{———} m$$

and $I(i,j)=0$ if i and j are neighbours in Q. Then from [8, Lemma 3] RG is isomorphic to the bounden quiver algebra FQ'/I' where Q' is the quiver

$$Q':\ \overset{\circlearrowleft\beta_1}{1} \overset{\alpha_1}{—} \overset{\circlearrowleft\beta_2}{2} \overset{\alpha_2}{—} \dots \overset{\circlearrowleft\beta_{m-1}}{m-1} \overset{\alpha_{m-1}}{———} \overset{\circlearrowleft\beta_m}{m}$$

and I' is generated by I, β_i^n, $i=1,\dots,m$, $\alpha_i\beta_i-\beta_{i+1}\alpha_i$ if $i\xrightarrow{\alpha_i}i+1$ and $\beta_i\alpha_i-\alpha_i\beta_{i+1}$, if $i\xleftarrow{\alpha_i}i+1$, $i=1,\dots,m-1$. Let us denote by $(\bar{Q},\bar{I})$ the bounden quiver constructed in the following way: vertices of $\bar{Q}$ are pairs $(i,j)\in\{1,\dots,m\}\times Z$, Z is the ring of integers, arrows of $\bar{Q}$ are formed by (i,β_j): $(i,j)\to(i,j-1)$, $i=1,\dots,m$, (α_i,j): $(i,j)\to(i+1,j)$ if α_i: $i\to i+1$ in Q and (α_i,j): $(i+1,j)\to(i,j)$ if α_i: $i+1\to i$ in Q, and $\bar{I}$ is generated by elements $(\alpha_k,j)\dots(\alpha_i,j)$ if $\alpha_k\dots\alpha_i$ belongs to I, $(i,\beta_j)(i,\beta_{j+1})\dots(i,\beta_{j+n-1})$, $(\alpha_i,j-1)(i,\beta_j)-(i+1,\beta_j)(\alpha_i,j)$ if $i\xrightarrow{\alpha_i}i+1$ and $(\alpha_i,j-1)(i+1,\beta_j)-i,\beta_j)(\alpha_i,j)$ if $i\xleftarrow{\alpha_i}i+1$, i, $k\in\{1,\dots,m\}$, $j\in Z$. Let σ be the authomorphism of $(\bar{Q},\bar{I})$ given by $\sigma((i,j))=(i,j-1)$, $\sigma((\alpha_i,j))=(\alpha_i,j-1)$, $\sigma((i,\beta_j))=(i,\beta_{j-1})$ and let (σ) be the cyclic group generated by σ.

Then $(\bar{Q},\bar{I})$ defines a locally finite dimensional K-category, the action of (σ) on $(\bar{Q},\bar{I})$ is free and locally bounded, the quotient bounden quiver $(\bar{Q},\bar{I})/(\sigma)$ is isormorphic to (Q',I'), and the canonical projection $(\bar{Q},\bar{I})\to(\bar{Q},\bar{I})/(\sigma)$ is a Galois covering. From [4, 5] we know that (Q',I') is representation-finite if and only if $(\bar{Q},\bar{I})$ is locally representation-finite.

In order to complete the proof of Theorem given in [8] we must examine carefully the case when Q contains, as a subquiver, one of the following quivers

$$\bullet\rightarrow\bullet\leftarrow\bullet \quad \text{or} \quad \bullet\leftarrow\bullet\rightarrow\bullet$$

By duality it is enough to consider only the first case.

Lemma 1. $R_F(1)$ *G is of finite representation type if and only if* $n=2$.

Proof. Assume that $n\geq 3$. Then $\mathrm{mod}_F(\bar{Q}_1,\bar{I}_1)$ contains as a full subcategory the category of all finite-dimensional representation of the following extended

Dynkin graph of type $\tilde{D}_4$

$$\begin{array}{ccccc} & & (2,2) & & \\ & & \downarrow & & \\ (1,1) & \rightarrow & (2,1) & \leftarrow & (3,1) \\ & & \downarrow & & \\ & & (2,0) & & \end{array}$$

Hence $\operatorname{mod}_F(\bar{Q}_1, \bar{I}_1)$ is not locally representation-finite and $R_F(1)G$ is not of finite representation type. Assume $n=2$. In order to prove that $(\bar{Q}_1, \bar{I}_1)$ is locally representation-finite we will prove the following: if $(V_{(i,j)})$ is an indecomposable object in $\operatorname{mod}_F(\bar{Q}_1, \bar{I}_1)$ such that $V_{(i,j)}=0$ for $j<0$ and $V_{(i,0)} \neq 0$ for some i, then $V_{(i,j)}=0$ for $j \geqq 3$. Thus our claim concerns the full subquivers (X_t, J_t) of $(\bar{Q}_1, \bar{I}_1)$ formed by the vertices (i,j), $i=1, 2, 3$, $j=0, \ldots, t$, and induced relations. The determination of the representation type of (X_t, J_t) and the computation of its Auslander-Reiten quiver can be done by calculation of dimension vectors (see [3]). The computation is recorded in Fig. 1. The double arrows denote the periodic part of the Auslander-Reiten quiver of (X_3, J_3). Repeating this periodic part infinitely many times on the left and on the right, we obtain the periodic Auslander-Reiten quiver of $(\bar{Q}_1, \bar{I}_1)$. Consequently $R_F(1)G$ has 42 nonisomorphic indecomposable right modules and this finishes the proof of the lemma.

Lemma 2. *Let $n=2$ and let R be the quiver algebra FQ where Q is one of the quivers*

$$1 \rightarrow 2 \leftarrow 3 \rightarrow 4 \quad \text{or} \quad 1 \rightarrow 2 \leftarrow 3 \leftarrow 4.$$

Then RG is of infinite representation type.

Proof. It is not hard to see that $\operatorname{mod}_F\bar{Q}$ contains as a full subcategory the category $\operatorname{mod}_F X$ where X is the following extended Dynkin graph of type $\tilde{D}_5$

$$\begin{array}{ccccc} (2,2) & & & & (4,1) \\ & \searrow & & \nearrow & \\ & (2,1) & \leftarrow & (3,1) & \\ & \nearrow & & \searrow & \\ (1,1) & & & & (3,0). \end{array}$$

From [1], FX is a hereditary algebra of finite representation type, $\operatorname{mod}_F\bar{Q}$ is not locally representation-finite, and consequently RG is not of finite representation type.

Lemma 3. *Let $n=2$ and let R be the bounden quiver algebra FQ/I where*

$$Q\colon 1 \xrightarrow{\alpha_1} 2 \xleftarrow{\alpha_2} 3 \xleftarrow{\alpha_3} 4 \xleftarrow{\alpha_4} 5$$

and I is generated by the composed arrows $\alpha_2\alpha_3$ and $\alpha_3\alpha_4$. Then RG is not of finite representation type.

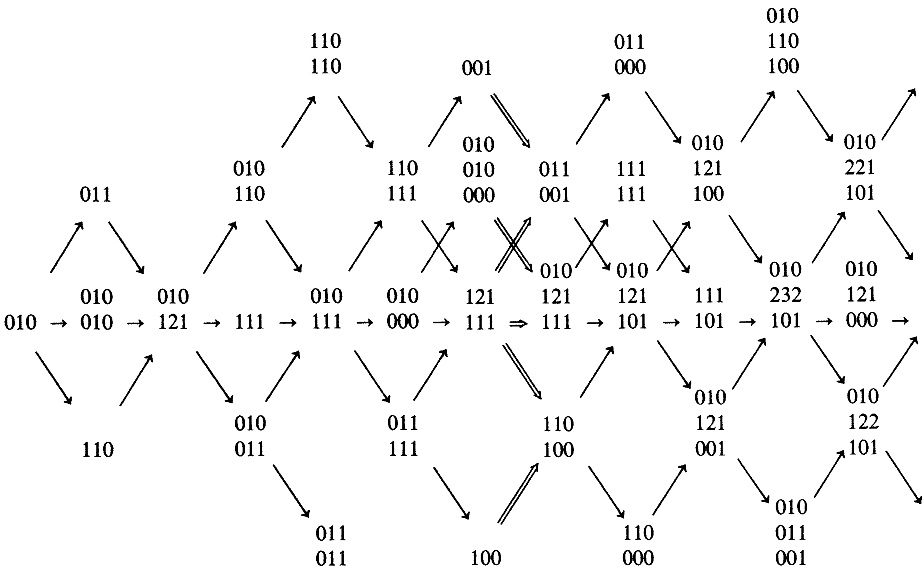


Fig. 1

Proof. It is easy to see that $\mathrm{mod}_F(\bar{Q}, \bar{I})$ contains as a full subcategory the category of all finite-dimensional representations of the following extended Dynkin graph of type $\tilde{E}_8$

$$\begin{array}{c} (2,3) \\ \downarrow \\ (1,1) \leftarrow (1,2) \rightarrow (2,2) \leftarrow (3,2) \rightarrow (3,1) \leftarrow (4,1) \rightarrow (4,0) \leftarrow (5,0). \end{array}$$

Then as above we conclude that RG is not of finite representation type.

Lemma 4. *Let $n=2$ and let R be the bounden quiver algebra FQ/I where*

$$Q\colon\ 1 \xrightarrow{\alpha_1} 2 \xrightarrow{\alpha_2} 3 \xleftarrow{\alpha_3} 4 \xleftarrow{\alpha_4} 5$$

and I is generated by the composed arrows $\alpha_2\alpha_1$ and $\alpha_3\alpha_4$. Then RG is not of finite representation type.

Proof. It is not hard to see that $\mathrm{mod}_F(\bar{Q}, \bar{I})$ contains as a full sub-category the category of all finite-dimensional representations of the following extended

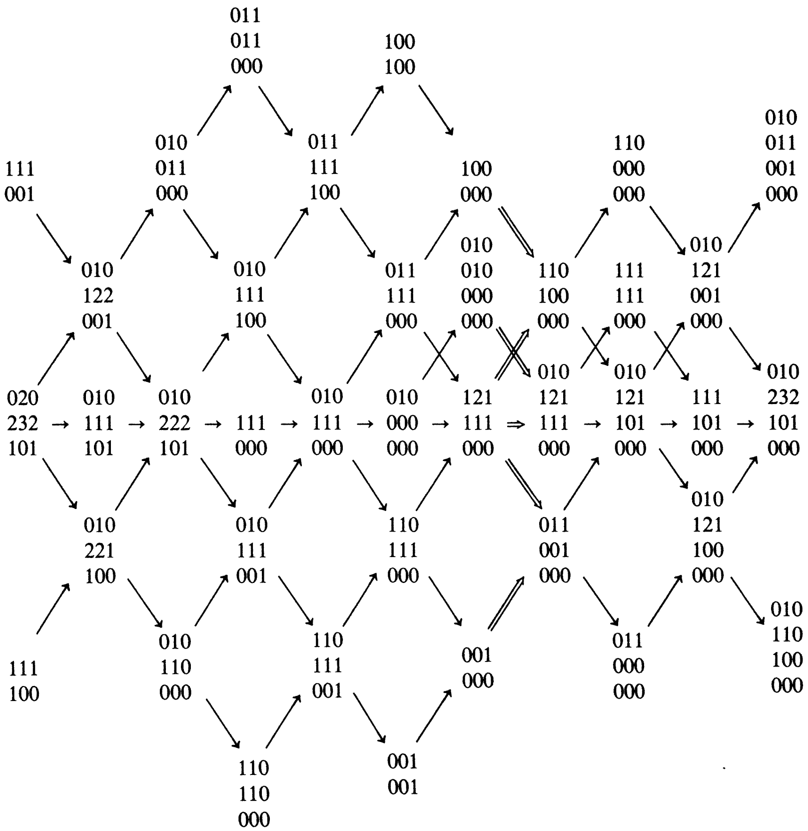


Fig. 1 (continuation)

Dynkin graph of type $\tilde{E}_7$

$$\begin{array}{c} (3,3) \\ \downarrow \\ (1,1)\to(2,1)\leftarrow(2,2)\to(3,2)\leftarrow(4,2)\to(4,1)\leftarrow(5,1). \end{array}$$

Then as above we conclude that RG is not of finite representation type.

From [8] and the above lemmas, for the proof of Theorem, it is sufficient to prove the following lemma.

Lemma 5. *If $R=R_F(3)$ and $n=2$, then RG is of finite representation type.*

Proof. From [4, Theorem 3.6] it is enough to show that $(\bar{Q}_2, \bar{I}_2)$ is locally representation-finite. For each $t \geqq 0$, let (X_t, J_t) be the full bounden subquiver of $(\bar{Q}_2, \bar{I}_2)$ formed by the vertices (i,j), $i=1, 2, 3$, $j \geqq 0$, and $(4,j)$, $j=0,\dots,t$, and induced relations. We will show that each bounden quiver (X_t, J_t) is locally representation-finite. We shall proceed by induction on t using one point extension procedure. From Lemma 1 we know that the full bounden subquiver (X, J) of $(\bar{Q}_2, \bar{I}_2)$ formed by the vertices (i,j), $i=1, 2, 3$, $j \geqq 0$, is locally representation-finite. For each point (i,j) of $(\bar{Q}_2, \bar{I}_2)$, we denote by $S_{(i,j)}$ the simple representation of $(\bar{Q}_2, \bar{I}_2)$ given by (i,j). Then, in notation of [8], each indecomposable representation of (X_0, J_0) is either an indecomposable representation of (X, J), is isomorphic to $S_{(4,0)}$, or is an indecomposable object in the subspace category $\mathfrak{U}(\mathcal{N}_F)$ where $N = S_{(3,0)}$ (see also [7]). From Fig. 1 it easily follows that $\mathfrak{U}(\mathcal{N}_F)$ is equivalent to $\mathfrak{U}(\mathcal{T}_F)$ where T is the partially ordered set given by the following family of irreducible maps in $\mathrm{mod}_F(X, J)$

$$\begin{array}{ccccccccccccccccc}
000 & & 000 & & 010 & & 010 & & 010 & & & & & & & & \\
000 & \to & 011 & \to & 121 & \to & 121 & \to & 011 & & & & & & & & \\
001 & & 001 & & 101 & & 001 & & 001 & & & & & & & & \\
 & & & & \downarrow & & \downarrow & & \downarrow & & & & & & & & \\
 & & & & 000 & & 010 & & 010 & & & & & & & & \\
 & & & & 111 & \to & 232 & \to & 122 & & & & & & & & \\
 & & & & 101 & & 101 & & 101 & & & & & & & & \\
 & & & & & & \downarrow & & \downarrow & & & & & & & & \\
 & & & & & & 010 & & 020 & & 010 & & & & & & \\
 & & & & & & 221 & \to & 232 & \to & 111 & & & & & & \\
 & & & & & & 101 & & 101 & & 101 & & & & & & \\
 & & & & & & \downarrow & & \downarrow & & \downarrow & & & & & & \\
 & & & & & & 000 & & 010 & & 010 & & 010 & & 110 & & 000 \\
 & & & & & & 111 & \to & 122 & \to & 222 & \to & 111 & \to & 111 & \to & 001 \\
 & & & & & & 001 & & 001 & & 101 & & 001 & & 001 & & 001.
\end{array}$$

Hence from [6] (see also [8, Lemma 8]). (X_0, J_0) is locally representation-finite. For $t=1$, each indecomposable representation of (X_1, J_1) is either an indecomposable representation of (X_0, J_0), is isomorphic to $S_{(4,1)}$, or is an indecomposable object in the subspace category $\mathfrak{U}(\mathcal{E}_F)$ where E is the injective envelope of $S_{(3,0)}$ in $\mathrm{mod}_F(X_0, J_0)$. Observe that $E/S_{(3,0)}$ is isomorphic to $S_{(3,1)} \oplus S_{(4,0)}$, and $S_{(4,0)}$ is a simple injective object in $\mathrm{mod}_F(X_0, J_0)$. Since there are no nonzero maps from $S_{(3,1)}$ to indecomposable objects in $\mathrm{mod}_F(X_0, J_0)$ forming the above partially ordered set T, $\mathfrak{U}(\mathcal{E}_F)$ is equivalent to $\mathfrak{U}(\mathcal{L}_F)$ where L is the partially ordered set whose elements are $S_{4,0}$, E, and the objects from $\sigma^{-1}(T)$, and with the order: $E < S_{4,0}$, $E < Y$ for any Y from $\sigma^{-1}(T)$, and $\sigma^{-1}(V) < \sigma^{-1}(W)$ if and only if $V < W$ in T. Then from [2, 6], (X_1, J_1) is locally representation-finite. Observe also $S_{(4,1)}$ is a simple injective object in $\mathrm{mod}_F(X_1, J_1)$ and there are no nonzero maps from $S_{(3,2)}$ to indecomposable objects in $\mathrm{mod}_F(X_1, J_1)$ which are not in $\mathrm{mod}_F(X, J)$. Hence repeating the above arguments we conclude that each bounden quiver (X_t, J_t) is locally representation-finite. Moreover, if $M = (M_{(i,j)})$ is an indecomposable representation of $(\bar{Q}_2, \bar{I}_2)$ such that $M_{(i,j)} = 0$ for $j < 0$ and $M_{(i,j)} \neq 0$ for some i, then M is an indecomposable representation of (X_1, J_1). Consequently $(\bar{Q}_2, \bar{I}_2)$ is locally representation-finite and $R_F(3)\ G$ is of finite representation type.

References

1. Dlab, V., Ringel, C.M.: Indecomposable representations of graphs and algebras. Mem. Amer. Math. Soc. **173** (1976)

2. Gabriel, P.: Representations indecomposables des ensembles ordonnés. In: Seminaire P. Dubreil, 26e année 1972/73. Exp. 13. Paris: Secrétariat mathématiques 1973
3. Gabriel, P.: Auslander-Reiten sequences and representation-finite algebras. In: Representation Theory I (V. Dlab, P. Gabriel, eds.). Proceedings (Ottawa 1979), pp. 1–71. Lecture Notes in Math. **831**. Berlin-Heidelberg-New York: Springer 1980
4. Gabriel, P.: The universal cover of a representation-finite algebra. In: Representations of Algebras (M. Auslander, E. Lluis, eds.). Proceedings (Puebla, Mexico 1980), pp. 68–105. Lecture Notes in Math. **903**. Berlin-Heidelberg-New York: Springer 1981
5. Martinez, R., de la Peña, J.A.: Automorphisms of representation-finite algebras. Invent. Math. **72**, 359–362 (1983)
6. Nazarova, L., Rojter, A.: Representations of partially ordered sets. Zap. Naučn. Sem. Leningrad. Otdel Mat. Inst. Steklov. (LOMI) **28**, 5–31 (1972) [Russian]
7. Ringel, C.M.: Tame algebras. In: Representation Theory I (V. Dlab, P. Gabriel, eds.). Proceedings (Ottawa 1979), pp. 137–287. Lecture Notes in Math. **831**. Berlin-Heidelberg-New York: Springer 1980
8. Meltzer, H., Skowroński, A.: Group algebras of finite representation type. Math. Z. **182**, 129–148 (1983)

Received August 11, 1983