SUB | NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

Werk

Titel: Mathematische Zeitschrift
Verlag: Springer
Jahr: 1985
Kollektion: Mathematica
Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen
Werk Id: PPN266833020_0189
PURL: http://resolver.sub.uni-goettingen.de/purl?PPN266833020_0189

Übergeordnetes Werk

Werk Id: PPN266833020
PURL: http://resolver.sub.uni-goettingen.de/purl?PPN266833020

Terms and Conditions

The Goettingen State and University Library provides access to digitized documents strictly for noncommercial educational, research and private purposes and makes no warranty with regard to their use for other purposes. Some of our collections are protected by copyright. Publication and/or broadcast in any form (including electronic) requires prior written permission from the Goettingen State- and University Library.
Each copy of any part of this document must contain there Terms and Conditions. With the usage of the library's online system to access or download a digitized document you accept the Terms and Conditions.
Reproductions of material on the web site may not be made for or donated to other repositories, nor may be further reproduced without written permission from the Goettingen State- and University Library.

For reproduction requests and permissions, please contact us. If citing materials, please give proper attribution of the source.

Contact

Niedersächsische Staats- und Universitätsbibliothek Göttingen
Georg-August-Universität Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen
Germany
Email: gdz@sub.uni-goettingen.de

Mathematische Zeitschrift

Band 189 · 1985

Herausgegeben von / Edited by
W. P. Barth, Erlangen
E. Becker, Dortmund
K. Diederich, Wuppertal
G. Faltings, Princeton
H. Heyer, Tübingen
W. Klingenberg, Bonn
C. C. Moore, Berkeley
H. H. Schaefer, Tübingen
L. Smith, Göttingen
W. v. Wahl, Bayreuth

Wissenschaftlicher Beirat / Advisory Board
T. tom Dieck, Göttingen · B. Eckmann, Zürich
E. Heinz, Göttingen · J. Herzog, Essen
K. H. Hofmann, Darmstadt · U. Krengel, Göttingen
G. Pickert, Gießen · A. Pfluger, Zürich
H. Salzmann, Tübingen · K. Zeller, Tübingen



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 aand *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:

W.P. Barth, Mathematisches Institut der Universität, Bismarckstr. 1½, D-8520 Erlangen, Federal Republic of Germany

E. Becker, FB Mathematik, Universität Dortmund, Postfach 50 05 00, D-4600 Dortmund 50, Federal Republic of Germany

K. Diederich, Gesamthochschule, Fachbereich 7, Mathematik, Gaußstr. 20, D-5600 Wuppertal 1, Federal Republic of Germany

G. Faltings, University of Princeton, Dept. of Mathematics, Fine Hall – Washington Road, Princeton, New Jersey 08544, USA

H. Heyer, Mathematisches Institut der Universität, Auf der Morgenstelle 10, D-7400 Tübingen, Federal Republic of Germany

W. Klingenberg, Mathematisches Institut der Universität, Wegelerstr. 10, D-5300 Bonn, Federal Republic of Germany

C.C. Moore, Mathematical Sciences Research Institute, 2223 Fulton Street, Room 603, Berkeley, CA 94720, USA

H.H. Schaefer, Mathematisches Institut der Universität, Auf der Morgenstelle 10, D-7400 Tübingen, Federal Republic of Germany

L. Smith, Mathematisches Institut der Universität, Bunsenstr. 3–5, D-3400 Göttingen, Federal Republic of Germany

W. v. Wahl, Mathematisches Institut der Universität, Postfach 30 08, D-8580 Bayreuth, Federal Republic of Germany

Copyright

Submission of a manuscript implies: that the work described has not been published before (except in the form of an abstract or as part of a published lecture, review, or thesis); that it is not under consideration for publication elsewhere; that its publication has been approved by all coauthors, if any, as well as by the responsible authorities at the institute where the work has been carried out; that, if and when the manuscript is accepted for publication, the authors agree to automatic transfer of the copyright to the publisher (society); and that the manuscript will not be published elsewhere in any language without the consent of the copyright holders.

All articles published in this journal are protected by copyright, which covers the exclusive rights to reproduce and distribute the article (e.g., as offprints), as well as all translation rights. No material published in this journal may be reproduced photographically or stored on microfilm, in electronic data bases, video disks, etc., without first obtaining written permission from the publisher.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if not specifically identified, does not imply that these names are not protected by the relevant laws and regulations.

While the advice and information in this journal is believed to be true and accurate at the date of its going to press, neither the authors, the editors, nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Special regulations for photocopies in the USA: Photocopies may be made for personal or in-house use beyond the limitations stipulated under Section 107 or 108 of U.S. Copyright Law, provided a fee is paid. This fee is US $0.20 per page, or a minimum of US $1.00 if an article contains fewer than five pages. All fees should be paid to the Copyright Clearance Center, Inc., 21 Congress Street, Salem, MA 01970, USA, stating the ISSN 0025-5874, the volume, and the first and last page numbers of each article copied. The copyright owner's consent does not include copying for general distribution, promotion, new works, or resale. In these cases, specific written permission must first be obtained from the publisher.

Other regulations: Authors publishing in this journal can, under certain conditions, benefit from library and photocopy fees collected by VG WORT. Authors of German nationality and those resident in the Federal Republic of Germany or Berlin (West), as well as citizens of Austria, Switzerland and member countries of the European Community, may apply to Verwertungsgesellschaft WORT, Abteilung Wissenschaft, Goethestraße 49, D-8000 München 2, for detailed information.

Springer-Verlag Berlin Heidelberg New York Tokyo
Printers: Universitätsdruckerei H. Stürtz AG, D-8700 Würzburg
© Springer-Verlag, Berlin Heidelberg 1985
Springer-Verlag GmbH & Co. KG, 1000 Berlin 33
Printed in Germany



Inhalt/Contents

Al-Sabti, G.: Framing Sphere Bundles over Spheres, the Smith Pairing, and Three-Fold Toda Brackets 457

Alt, H.W., Tomi, F.: Regularity and Finiteness of Solutions to the Free Boundary Problem for Minimal Surfaces 227

Balas, A.: Compact Hermitian Manifolds of Constant Holomorphic Sectional Curvature 193

Bantle, G.: Obere Abschätzung für die Anzahl der B-Zwillinge auf kurzen Intervallen 561

Blackadar, B., Kumjian, A.: Skew Products of Relations and the Structure of Simple C^*-Algebras 55

Bleiler, S.A.: Strongly Invertible Knots Have Property R 365

Böcherer, S.: Über die Fourier-Jacobi-Entwicklung Siegelscher Eisensteinreihen II 81

Brown, KA., Levasseur, T.: Cohomology of Bimodules Over Enveloping Algebras 393

Bruna, J., Cufí, J., Verdera, J.: Cauchy Kernels in Strictly Pseudoconvex Domains and an Application to a Mergelyan Type Approximation Problem 41

Calvert, B., s. Fitzpatrick, S. 555

Carlisle, D., Eccles, P., Hilditch, S., Ray, N., Schwartz, L., Walker, G., Wood, R.: Modular Representations of $GL(n, p)$, Splitting $\Sigma(\mathbb{C}P\infty \times \ldots \times \mathbb{C}P\infty)$, and the β-family as Framed Hypersurfaces 239

Collingwood, D.H.: A Note on Continuous Cohomology for Semisimple Lie Groups 65

Cufí, J. s. Bruna, J. et al. 41

Dade, E.C.: Generalized Clifford Correspondences for Group Characters 19

Dye, R.H.: Maximal Subgroups of the Finite Orthogonal and Unitary Groups Stabilizing Anisotropic Subspaces 111

Dziuk, G.: C^2-Regularity for Partially Free Minimal Surfaces 71

Eccles, P., s. Carlisle, D., et al. 239

Emilion, R., Hachem, B.: A Multiparameter Strongly Superadditive Ergodic Theorem 11

Fitzpatrick, S., Calvert, B.: In a Nonreflexive Space the Subdifferential is Not Onto 555
Gay, R., Sebbar, A.: Division et extension dans l'algèbre $A^{\infty}(\Omega)$ d'un ouvert pseudoconvexe à bord lisse de $\mathbb{C}^n$ 421
Ginibre, J., Velo, G.: The Global Cauchy Problem for the Non Linear Klein-Gordon Equation 487
Graham, I., Wu, H.: Characterizations of the Unit Ball B^n in Complex Euclidean Space 449
Hachem, B., s. Emilion, R. 11
Hansell, R.W., Jayne, J.E., Labuda, I., Rogers, C.A.: Boundaries of and Selectors for Upper Semi-Continuous Multi-Valued Maps 297
Hauschild, V.: Actions of Compact Lie Groups on Homogeneous Spaces 475
Haußmann, W., Jetter, K., Steinhaus, B.: Degree of Best Approximation by Triginometric Blending Functions 143
Hilditch, S., s. Carlisle, D., et al. 239
Hill, T.P.: Equipartitioning Common Domains of Non-Atomic Measures 415
Jayne, J.E., s. Hansell, R.W., et al. 297
Jetter, K., s. Haußmann, W., et al. 143
Jia-chen, Y.: Filtrations of Principal Indecomposable Modules of Frobenius Kernels of Reductive Groups 515
Kneppers, H.A.W.M.: The Hasse-Witt Matrix of a Formal Group . . . 151
Kumjian, A., s. Blackadar, B. 55
Labuda, I., s. Hansell, R.W., et al. 297
Lai, K.F.: Algebraic Cycles on Compact Shimura Surface 593
Levasseur, T., s. Brown, K.A. 393
Loos, O.: Charakterisierung symmetrischer R-Räume durch ihre Einheitsgitter 211
Makino, R.: QF-1 Algebras of Local-Colocal Type 571
Mitchell, S.A.: Splitting $B(\mathbb{Z}/p)^n$ and BT^n Via Modular Representation Theory 1
Myerson, G.: On The Values of the Dedekind Sum 337
Parrini, C., Selvaggi Primicerio, A.: Uniform Estimate for the $\bar{\partial}$-Problem in a Class of Branched Covering of Strictly Pseudoconvex Domains 465
Patrizio, G.: A Characterization of Complex Manifolds Biholomorphic to a Circular Domain 343
Petersson, H.: Über gewisse Dirichlet-Reihen mit Eulerscher Produktzerlegung 273
Ray, N., s. Carlisle, D., et al. 239
Rees, S.: Finite C_3 Geometries in Which All Lines Are Thin 263
Rogers, C.A., s. Hansell, R.W., et al. 297

Ruh, B.: Krümmungstreue Diffeomorphismen Riemannscher und pseudo-Riemannscher Mannigfaltigkeiten 371
Sauvigny, F.: Die zweite Variation von Minimalflächen im $\mathbb{R}^p$ mit polygonalem Rand 167
Schwartz, L., s. Carlisle, D., et al. 239
Sebbar, A., s. Gay, R. 421
Selvaggi Primicerio, A., s. Parrini, C. 465
Steinhaus, B., s. Haußmann, W., et al. 143
Sterk, H.: Finiteness Results for Algebraic K3 Surfaces 507
Straube, E.J.: CR-Distributions and Analytic Continuation at Generating Edges 131
Terasoma, T.: Complete Intersections with Middle Picard Number 1 Defined over Q 289
Tomi, F., s. Alt, H.W. 227
Velo, G., s. Ginibre, J. 487
Verdera, J. s. Bruna, J., et al. 41
Walker, G., s. Carlisle, D., et al. 239
Webb, G.F.: Dynamics of Populations Structured by Internal Variables 319
Wood, R., s. Carlisle, D., et al. 239
Wu, H, s. Graham, I. 449
Yamaguchi, T.: The Isometry Groups of Manifolds of Nonpositive Curvature with Finite Volume 185
Zeuner, H.: Die Existenz einer Lévy-Abbildung auf einem homogenen Raum 529

Covered by Zentralblatt für Mathematik and
Current Mathematical Publications

Mathematische Zeitschrift

1-417
94
144 ol

Band 189 Heft 1 1985 Z 4196

Herausgegeben von / Edited by
W. P. Barth, Erlangen
E. Becker, Dortmund
K. Diederich, Wuppertal
G. Faltings, Princeton
H. Heyer, Tübingen
W. Klingenberg, Bonn
C. C. Moore, Berkeley
H. H. Schaefer, Tübingen
L. Smith, Göttingen
W. v. Wahl, Bayreuth

Wissenschaftlicher Beirat / Advisory Board

T. tom Dieck, Göttingen · B. Eckmann, Zürich
E. Heinz, Göttingen · J. Herzog, Essen
K. H. Hofmann, Darmstadt · U. Krengel, Göttingen
G. Pickert, Gießen · A. Pfluger, Zürich
H. Salzmann, Tübingen · K. Zeller, Tübingen

Springer-Verlag
Berlin Heidelberg New York Tokyo

209 Math. Z. ISSN 0025-5874 MAZEAX 189 (1) 1-150 (1985)

25. März 1986
8 Z Nds 527
Niedersächsische Staats- und Universitätsbibliothek Göttingen
X, 82

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.

"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"** 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.

„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.

Von jeder Arbeit werden 100 Sonderdrucke kostenlos zur Verfügung gestellt, weitere können zum Selbstkostenpreis bezogen werden.

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

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.

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.

Geschäftsführende Herausgeber / Managing Editors

W.P. Barth — Mathematisches Institut der Universität, Bismarckstr. $1\frac{1}{2}$, D-8520 Erlangen, Federal Republic of Germany

W. v. Wahl — Mathematisches Institut der Universität, Postfach 3008, D-8580 Bayreuth, Federal Republic of Germany

Manuskripte nehmen außerdem entgegen / Manuscripts may also be sent to:

E. Becker — Institut für Mathematik der Universität, Vogelpothsweg 87, D-4600 Dortmund 50, Federal Republic of Germany

K. Diederich — Gesamthochschule, Fachbereich 7, Mathematik, Gaußstr. 20, D-5600 Wuppertal 1, Federal Republic of Germany

G. Faltings — University of Princeton, Dept. of Mathematics, Fine Hall – Washington Road, Princeton, New Jersey 08544, USA

H. Heyer — Mathematisches Institut der Universität, Auf der Morgenstelle 10, D-7400 Tübingen, Federal Republic of Germany

W. Klingenberg — Mathematisches Institut der Universität, Wegelerstr. 10, D-5300 Bonn, Federal Republic of Germany

C. C. Moore — Mathematical Sciences Research Institute, 2223 Fulton Street, Room 603, Berkeley, CA 94720, USA

H. H. Schaefer — Mathematisches Institut der Universität, Auf der Morgenstelle 10, D-7400 Tübingen, Federal Republic of Germany

L. Smith — Mathematisches Institut der Universität, Bunsenstr. 3–5, D-3400 Göttingen, Federal Republic of Germany

Math. Z. 189, 1-9 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

Splitting $B(\mathbb{Z}/p)^n$ and BT^n Via Modular Representation Theory

Stephen A. Mitchell*

Math. Dept., University of Washington, Seattle WA 98195, USA

It is by now well known that when a finite group G acts on a space X, one can use idempotents in the p-adic group ring to obtain stable splittings of the p-completion of X. Early applications of this technique tended to rely on groups of order prime to p, for the good reason that the representation theory is then relatively simple. In more recent applications [8, 6, 9] this has not been the case, but the particular representations exploited are projective and irreducible; in such cases one can, in effect, proceed as though $p \nmid |G|$. For example, in [8] the Steinberg representation of $GL_n\mathbb{F}_p$ is used obtain stable summands of $B(\mathbb{Z}_p)^n$ and BT^n (T^n is the n-torus). In this paper we show that in one sense the most interesting summand of $B(\mathbb{Z}/p)^n$ is in fact the one corresponding to the highly nonprojective trivial representation: the transfer $B(\mathbb{Z}/p)^n_+ \to S^0$ factors through this summand. Similarly, the transfer $\sum^n BT^n_+ \to S^0$ (p-completed) factors through the summand corresponding to the $\det^{-1}$ representation of $GL_n\mathbb{F}_p$ (this has been proved independently by N. Ray and the Manchester school). Hence these summands should be of considerable interest. (The case BT^2 has been extensively analyzed in [2].)

We also give (in Sect. 3) a detailed description of the stable splitting of $B(\mathbb{Z}/2)^3$ obtained by these methods. Section one contains some preliminary remarks on stable splittings of spectra. Most of this material is already implicit in the literature, but for our purposes it seems best to make it explicit. We also record (1.6) how a spectrum splits when a finite-dimensional algebra "acts" on X.

All homology and cohomology is with $\mathbb{Z}/p$ coefficients.

1. Idempotents and Stable Splittings

In this section X denotes a CW-spectrum and $[X, X]$ is the ring of self-maps of X. If $e \in [X, X]$, we let eX denote the infinite mapping telescope $\varinjlim(X \xrightarrow{e} X \xrightarrow{e} X \ldots)$. The following is standard:

* Supported by NSF Grant No. MCS 83-11676

(1.1) **Proposition.** *If e is idempotent, $X \cong eX \vee (1-e)X$. Moreover, for any spectrum Y, $[Y, eX] = e[Y, X]$ and $[eX, Y] = [X, Y]e$.* □

Remark. Of course we could use the equation $[Y, eX] = e[Y, X]$ to define eX, via Brown's representability theorem. Note also that the inverse system $[X, Y] \xleftarrow{e^*} [X, Y] \xleftarrow{e^*} \ldots$ trivially satisfies the Mittag-Leffler condition.

In practice, we are often faced with a map e that is only idempotent on mod p homology: $e_*^2 = e_*$. A convenient method of dealing with this situation is provided by the following straightforward generalization of [4], Theorem 3.4a, p. 73:

(1.2) **Proposition.** *Let R be any ring with identity, and suppose R has a complete filtration $R = F^0R \supseteq F^1R \supseteq \ldots$. Then any idempotent in R/F^1R lifts to an idempotent in R. Moreover, any splitting of $1 \in R/F^1R$ into orthogonal idempotents lifts to a splitting of $1 \in R$ into orthogonal idempotents.*

Proof. Let $R_n = R/F^{n+1}R$, and let $\pi_n : R_n \to R_{n-1}$ denote the projection. By assumption, $R = \varprojlim R_n$. Hence if $e \in R_0$ is idempotent, we need to construct, by induction on n, $e_n \in R_n$ such that $e_0 = e$, $e_n^2 = e_n$ and $\pi_n(e_n) = e_{n-1}$. We proceed exactly as in [4], p. 65: At the inductive step, choose any $a \in R_n$ with $\pi_n(a) = e_{n-1}$. Then let $e_n = a + (1-2a)(a^2 - a)$. Since $n \geqq 1$ and $a^2 - a \in F^{n+1}R$, $(a^2 - a)^2 = 0$ in R_n; one then computes $e_n^2 - e_n = 0$. For the second statement we again follow [4], p. 73. Suppose $1 = f_1 + \ldots + f_n \in R_0$, where the f_i are orthogonal idempotents. By induction we can assume there are orthogonal idempotents $e_1, \ldots, e_{n-2}$, $e \in R$ such that $\pi(e_i) = f_i$ and $\pi(e) = f_{n-1} + f_n$, where $\pi : R \to R_0$ is the projection. Now eRe is a subring of R, with identity e, and it is easily checked that the restricted filtration $(F^kR) \cap (e\,\mathrm{Re}) = eF^k\,\mathrm{Re}$ is a complete filtration on $e\,\mathrm{Re}$. By what has already been proved, we obtain $e_{n-1} \in e\,\mathrm{Re}$ with $e_{n-1}^2 = e_{n-1}$ and $\pi(e_{n-1}) = f_{n-1}$. Then e_{n-1} is orthogonal to $e_1, \ldots, e_{n-2}$, and $1 = e_1 + \ldots + e_{n-1} + (e - e_{n-1})$ is the desired splitting. □

The following is a useful corollary of the proof of (1.2):

(1.3) **Proposition.** *Suppose $f \in R_0$ is idempotent, and $g \in R$ is any (not necessarily idempotent) lift of f. Then f can be lifted to an idempotent $e \in R$ such that e_k is a polynomial in g_k, with zero constant term (e_k, g_k denote components in $\varprojlim R_k$).* □

The classical application of (1.2) is to the case when $R = KG$ is the group ring of a finite group G over a complete local field K, with F the p-adic filtration. On the other hand, when $R = [X, X]$ and F denotes the Adams filtration (with respect to $K\mathbb{Z}/p$), we have:

(1.4) **Corollary.** *Suppose X and Y are connective and p-adically complete. Then:*

(a) *Suppose $f_1, \ldots, f_n$ are orthogonal idempotents in $\mathrm{Hom}_{\mathbb{F}_p}(H_*X, H_*X)$, with $f_1 + \ldots + f_n = 1$, and suppose each f_i is the image of some $e_i \in [X, X]$. Then the e_i can be chosen to be orthogonal idempotents, with $e_1 + \ldots + e_n = 1$.*

(b) *Suppose $f \in \mathrm{Hom}_{\mathbb{F}_p}(H_*X, H_*X)$ is idempotent, $X \xrightarrow{\phi} Y$ is a map, and there exists a lift $g \in [X, X]$ of f such that $\phi g = 0$. Then there exists an idempotent lift e with $\phi e = 0$.*

(c) *Let f be as in* (b), *and suppose $g_1, g_2 \in [X, X]$ are any two lifts of f. Then $g_1 X \cong g_2 X$.*

Proof. Since the Adams filtration on $[X, X]$ is complete, (a) follows from (1.2). For (b), we use e as in (1.3). It follows that ϕe has Adams filtration n for all n, and hence is zero. For (c), lift f to any idempotent e; then

$$[eX, g_i X] = \varinjlim([X, X]e \xrightarrow{g_i} [X, X]e \xrightarrow{g_i} \ldots) \quad \text{for } i = 1, 2.$$

In particular, there is a canonical map $eX \to g_i X$ corresponding to the class of e in the limit, and it is the identity map on H_*. □

Remark. If f is as in (b), then (c) justifies the notation fX for any corresponding stable summand of X.

The structure of the ring $[X, X]$ itself is usually not accessible. Suppose, however, that B is a finite dimensional subalgebra of $\mathrm{Im}([X, X] \to \mathrm{Hom}_{\mathbb{F}_p}(H_* X, H_* X))$. Then the representation theory of such algebras can be applied to obtain stable splittings of X, as follows: There are finitely many distinct irreducible B-modules W. The projective covers P_W of the irreducibles are precisely the projective indecomposables of B. If $a_W = \dim \hom_B(W, W)$, then $a_W = 1$ if and only if W is absolutely irreducible; if $a_W = 1$ for all W we say that $\mathbb{F}_p$ is a splitting field for B. If P is projective and N is arbitrary, let $\langle P, N \rangle = \dim \hom_B(P, N)$. Then $\langle P_W, W' \rangle = a_W$ if $W = W'$ and is zero otherwise. It follows from this that ${}_B B \cong \bigoplus_W (a_W^{-1} \dim W) P_W$. Equivalently, $1_B = \sum_{W, i \leq i_W} e_{W,i}$, where $i_W = a_W^{-1} \dim W$ and the $e_{W,i}$ are primitive orthogonal idempotents, with $A e_{W,i} = P_W$. It also follows that if we let $m(N, W)$ denote the multiplicity of W in N, then $\langle P_W, N \rangle = a_W m(N, W)$. (For proofs of these assertions, see [3]).

Now the decomposition of 1_B is not unique, but we do have the following (part (a) was pointed out to the author by Reg Wood):

(1.5) **Proposition.** (a) *Any two decompositions of 1_B into primitive orthogonal idempotents are conjugate.* (b) *If N is any B-module, and e is any idempotent with $Be \cong P_W$, then* $\dim(eN) = a_W m(N, W)$.

Proof. (a) Suppose $\sum e_i, \sum e_i'$ are two such decompositions. Then without loss of generality we may assume $Be_i \cong Be_i'$. The isomorphism ${}_B B = \bigoplus Be_i \xrightarrow{\cong} \bigoplus Be_i' = {}_B B$ must be right multiplication by some unit λ, and $e_i \lambda = b_i e_i'$ with $b_i \in B$. But then $\lambda = b_i e_i'$, and hence $\lambda e_i' = b_i e_i' = e_i$ for all i, as desired. (b) There is a canonical vector space isomorphism $\mathrm{Hom}_B(Be, N) \xrightarrow{\sim} eN$, and hence $\dim(eN) = \langle P_W, N \rangle = a_W m(N, W)$. □

If N is a graded B-module of finite type, then as in [6] we let $F(N, t)$ denote the ordinary Poincare series, $F(N, B, W; t)$ the W-multiplicity series, and $F(N, B; t)$ the B-Poincare series with coefficients in the representation ring of B. We then have:

(1.6) **Proposition.** *Let X be p-complete with $H_* X$ of finite type, and let B be as above. Choose primitive idempotents e_W (one for each irreducible W) such that $A e_W \cong P_W$. Then $X \cong \bigvee_W (a_W^{-1} \dim W)(e_W X)$. Moreover $e_W X$ is independent (up to*

equivalence) of the choice of e_W, *and* $F(H_*(e_W X); t) = a_W F(H_* X, B, W; t)$. *(Note* $e_W X$ *may be trivial.)*

Proof. We have $1_B = \sum e_{W,i}$ as above, with $e_{W,1} = e_W$. Then from (1.4)(a) we have $X \cong \bigvee_{W,i} e_{W,i} X$, but $e_{W,i} X \cong e_{W,1} X$ by (1.5a). The second statement also follows from (1.5)(a), and the last is just the graded form of (1.5)(b). □

Remark. The splitting of B into "blocks" (cf. [8], § 2) or equivalently of 1_B into centrally primitive central idempotents π is in general a much coarser splitting than the above; indeed π is just the sum of all the $e_{W,i}$ such that W belongs to the block πB. However, the splitting using the block idempotents π does have some virtues: It is canonical (since the π's are unique), and it is a splitting of B-modules.

2. Splitting $B(\mathbb{Z}/p)^n$, BT^n, and the Transfer

The group $GL_n\mathbb{F}_p$ acts on the cohomology of BG when $G = (\mathbb{Z}/p)^n$ or $G = T^n$. In both cases the action is induced by the evident action of the monoid of $n \times n$-matrices over $\mathbb{Z}$ on G. Hence (1.6) describes a stable splitting of BG. Unfortunately, the modular representations of $GL_n\mathbb{F}_q$ are not as well understood as one might like. On the one hand, the irreducible representations have in a sense been completely classified by the theory of highest weight vectors (cf. [10]). In particular, there are exactly $(q-1)q^{n-1}$ distinct irreducible representations of $GL_n\mathbb{F}_q$ over $\mathbb{F}_q$, and $\mathbb{F}_q$ is a splitting field. But on the other hand, even the dimensions of the irreducibles, and of their projective covers, are in general unknown, and formulas for idempotents are hard to come by. Nevertheless, the methods of [6] yield at least a crude estimate of the sizes of the various pieces; in particular the summands arising from (1.6) are all non-trivial.

(2.1) **Proposition.** *Let* $G = (\mathbb{Z}/p)^n$ *or* T^n, *let* W *be an irreducible representation of* $GL_n\mathbb{F}_p$, *and let* e_W *be an associated primitive idempotent. Then* $F(H^*(e_W BG); t) = f_W(t) H(t)$, *where* (a) $H(t) = \prod_{i=0}^{n-1} (1 - t^{2(p^n - p^i)})^{-1}$ $\left(\text{resp. } H(t) = \prod_{0}^{n-1} (1 - t^{2^n - 2^{n-i}})^{-1}\right)$ *for* $G = T^n$ *or* p *odd (resp.* $G = (\mathbb{Z}/2)^n$*) and* (b) $f_W(t)$ *is a nonzero polynomial with nonnegative integer coefficients, such that* $f_W(1) = \dim P_W$ *if* $G = T^n$ *or* $G = (\mathbb{Z}/2)^n$, *and* $f_W(1) = 2^n \dim P_W$ *if* $G = (\mathbb{Z}/p)^n$ *and* p *odd. In particular,* $p^{\binom{n}{2}}$ *divides* $f_W(1)$.

Proof. This follows at once from (1.6) and the results of [6], Sects. 1, 2. □

It is difficult to obtain any further information about the summands. An exception is the case when W is a *projective* irreducible - i.e. $W = \det^k \cdot St$, where St is the Steinberg module and $0 \leqq k \leqq p-2$ ([8, 6]). However, the main point of this section will be to observe that in a sense the most interesting case from the viewpoint of homotopy theory lies at the opposite extreme: the case when $W = 1$, the trivial (one-dimensional) representation. (In this context, the trivial representation should be regarded as the most complicated representation!) Here we remind the reader that even the dimension of its projective

cover is unknown, and that the corresponding summand of H^*BG is much bigger than just the ring of $GL_n\mathbb{F}_p$-invariants (cf. 2.1).

We pause to recall the transfer for compact Lie groups, following [5]. Let G be a compact Lie group, with closed subgroup H. Let ξ_G, BG^{ξ_G} denote respectively the adjoint bundle over BG and its Thom spectrum. Then the transfer $t_{H,G}$ is a stable map $BG^{\xi_G} \to BH^{\xi_H}$. The definition given in [5] would appear to depend on various choices - e.g., a sequence of smooth manifolds approximating BG. However, one can easily show that the restrictions of $t_{H,G}$ to finite skeleta are independent of the choices, so that $t_{H,G}$ is at least well-defined after completion at a fixed prime p (by a standard $\lim^1$ argument). For the remainder of this section, therefore, all spectra will implicitly be completed at p. We obtain in particular transfer maps $\sum^n BT^n_+ \xrightarrow{r} B(\mathbb{Z}/p)^n_+$, $B(\mathbb{Z}/p)^n_+ \xrightarrow{s} S^0$, and $t = sr$.

Now let e_k denote a primitive idempotent associated to the representation $\det^k: GL_n\mathbb{F}_p \to GL_1\mathbb{F}_p$.

(2.2) **Theorem.** (a) *$s: B(\mathbb{Z}/p)^n_+ \to S^0$ factors through $e_0 B(\mathbb{Z}/p)^n$.*

(b) *$t: \sum^n BT^n_+ \to S^0$ factors through $e_{-1}BT^n$.*

(2.3) **Lemma.** (a) *s is $GL_n\mathbb{F}_p$-equivariant, with $GL_n\mathbb{F}_p$ acting trivially on S^0.* (b) *Let $M_n\mathbb{Z}$ denote the (multiplicative) submonoid of the $n \times n$-matrices over $\mathbb{Z}$ generated by $GL_n\mathbb{Z}$ and the diagonal matrices with nonzero determinant* mod p. *Let $M_n\mathbb{Z}$ act on $\sum^n BT^n_+ = S^n BT^n_+$ via the determinant representation on S^n and the usual action on BT^n_+, and on $B(\mathbb{Z}/p)^n$ via reduction to $GL_n\mathbb{F}_p$. Then $r: \sum^n BT^n_+ \to B(\mathbb{Z}/p)^n_+$ is $M_n\mathbb{Z}$-equivariant.*

Proof of 2.2. (a) By (2.3)(a), $sb = 0$ for any b in the kernel of the augmentation $\varepsilon: \mathbb{Z}_p GL_n\mathbb{F}_p \to \mathbb{Z}_p$. Hence it will be enough to show that if $e \in \mathbb{Z}_p GL_n\mathbb{F}_p$ is any primitive idempotent associated to a nontrivial irreducible W, then $\varepsilon(e) = 0$; or equivalently $\varepsilon(e) = 0 \bmod p$. But if $A = \mathbb{Z}/pGL_n\mathbb{F}_p$, then Ae is a projective cover of W, and hence $\mathrm{Hom}_A(Ae, \mathbb{F}_p) = 0$. (b) By (2.3) t is $M_n\mathbb{Z}$-equivariant with respect to the trivial action on S^0 and the twisted action of 2.3b. Using the twisted action, (2.2b) asserts that t factors through $e_0 BT^n_+$. Suppose $e \in \mathbb{Z}_p GL_n\mathbb{F}_p$ is associated to a nontrivial W, as above. Then we have shown that $\varepsilon(e) = 0$. Since $M_n\mathbb{Z} \to GL_n\mathbb{F}_p$ is surjective (by the Bruhat decomposition), we can then lift e to a (non-idempotent) element $\tilde{e}$ of the monoid ring $\mathbb{Z}_p M_n\mathbb{Z}$, such that $t\tilde{e} = 0$, t is trivial on eBT^n_+ by (1.4)b, and the result follows. □

Proof of 2.2. Suppose G is a compact Lie group and σ is an automorphism of G. Since the adjoint representation of G is σ-invariant, $B\sigma: BG \to BG$ extends to a bundle automorphism of ξ_G. If H is a closed subgroup with $\sigma(H) = H$, the following is then an easy consequence of the "pullback lemma" ([5], 1.16).

(2.4) **Lemma.** The following diagram commutes:

$$\begin{array}{ccc} BH^{\xi_H} & \xrightarrow{B\sigma} & BH^{\xi_H} \\ {\scriptstyle t_{H,G}}\uparrow & & \uparrow{\scriptstyle t_{H,G}} \\ BG^{\xi_G} & \xrightarrow[B\sigma]{} & BG^{\xi_G}. \end{array} \quad \square$$

Lemma 2.3(a) follows at once. For 2.3(b), first consider $\sigma \in GL_n\mathbb{Z}$. The adjoint bundle for T^n is of course trivial, and the induced map $S^n \wedge BT^n_+ \to S^n \wedge BT^n_+$ is just $\sigma \wedge B\sigma$. Here $GL_n\mathbb{Z}$ acts on S^n via inclusion in $GL_n\mathbb{R}$, and hence $\sigma: S^n \to S^n$ has degree $\det\sigma$. Now suppose $\sigma \in M_n\mathbb{Z}$ is diagonal with $\det\sigma \neq 0 \bmod p$. To prove (2.3)(b), for this σ, it is enough to consider the case when $n=1$ and $\sigma = (k)$, with $k>0$ and $k \neq 0 \bmod p$. From the pullback lemma we obtain a commutative diagram

$$\begin{array}{ccc} B\mathbb{Z}/p_+ \wedge B\mathbb{Z}/k_+ & \xrightarrow{[k]\circ\pi} & B\mathbb{Z}/p_+ \\ \uparrow{\scriptstyle t_{\mathbb{Z}/pk,S^1}} & & \uparrow{\scriptstyle t} \\ \sum \mathbb{C}P^\infty_+ & \xrightarrow[[k]]{} & \sum \mathbb{C}P^\infty_+ \end{array}$$

where $[k]$ is the kth power map of $\mathbb{C}P^\infty$ or $B\mathbb{Z}/p$, and

$$\pi: B\mathbb{Z}/pk = B\mathbb{Z}/p \times B\mathbb{Z}/k \to B\mathbb{Z}/p$$

is the projection. Since we have completed at p, the natural map $\phi: B\mathbb{Z}/p_+ \to B\mathbb{Z}/p_+ \wedge B\mathbb{Z}/k_+$ is an equivalence. Moreover, $\pi\phi = 1$ and $t_{\mathbb{Z}/p,\mathbb{Z}/pk} \circ \phi = k$, since the transfer $B\mathbb{Z}/k_+ \to S^0$ has degree k on the bottom cell. Hence $r[k] = [k]t_{\mathbb{Z}/pk,S^1} = [k]k^{-1}r$, or $[k]r = rk[k]$, as desired.

From the point of view of the transfers s, t, then, the summands of interest are respectively $e_0B(\mathbb{Z}/p)^n$ and $e_{-1}BT^n$. A crude approximation to $e_0B(\mathbb{Z}/p)^n$ is $BGL_2\mathbb{F}_q$, where $q=p^n$. The inclusion $(\mathbb{Z}/p)^n \subset GL_2\mathbb{F}_q$ as the unipotent subgroup induces an isomorphism $H^*BGL_2\mathbb{F}_q \to (H^*B(\mathbb{Z}/p)^n)^{\mathbb{F}_q^*}$ ([1]). Hence $BGL_2\mathbb{F}_q$ is a stable summand of $B(\mathbb{Z}/p)^n$, and it can be shown that $e_0B(\mathbb{Z}/p)^n$ is then a stable summand of the "smaller" space $BGL_2\mathbb{F}_q$. Unfortunately, we have been unable to progress much further in understanding $e_0B(\mathbb{Z}/p)^n$. As a meagre example, we have:

(2.5) **Proposition.** *Let $p=2$, and let $X_n = e_0B(\mathbb{Z}/2)^n$. Then* (a) *$X_n$ is precisely $(n-1)$-connected.* (b) *The transfer s induces a K-theory equivalence $X_n \to S^0$.*

Proof. (a) Note we are omitting a disjoint basepoint for $B(\mathbb{Z}/2)^n$; otherwise X_n would begin with a zero-cell. By (1.5), (a) is equivalent to the following: Let r_n = smallest positive dimension such that $H^{r_n}B(\mathbb{Z}/2)^n$ contains a trivial composition factor. Then $r_n = n$. To see that $r_n \leqq n$, let $S = H^*B(\mathbb{Z}/n)^n = \mathbb{F}_2[x_1, \ldots, x_n]$, and let $T = \mathbb{F}_2[x_1^2, \ldots, x_n^2]$. Then T is a $GL_n\mathbb{F}_2$-subalgebra of S, and hence the projection $S \to S \otimes_T \mathbb{F}_2$ is $GL_n\mathbb{F}_2$-equivariant. But in dimension n, $S \otimes_T \mathbb{F}_2$ is just the trivial module. Hence $r_n \leqq n$. On the other hand, the first positive dimensional $\mathbb{F}_{2^n}^*$-invariant of S is in dimension n ([1]), so $r_n \geqq n$ (here we are of course using the fact that $\mathbb{F}_{2^n}^*$ has odd order).

(b) It is enough to show that $s^*: K(1)^*S^0 \to K(1)^*X_n$ is an isomorphism, where $K(1) = \bmod 2$ complex K-theory (Morava K-theory). Now on ordinary K-theory, S^* is an isomorphism onto the submodule generated by the regular representation bundle (since transfer corresponds to induction on the representation rings). Hence it will be enough to show that $\dim_{K(1)^*} K(1)^*X_n = 1$. Let E be the exterior algebra as in (a), considered abstractly as just a $GL_n\mathbb{F}_2$-module.

Then by corollary (4.11) of [6], $\dim_{K(1)^*}K(1)^*X_n=\dim_{\mathbb{F}_p}Ee_0-1$ (minus one because we've omitted the base point). Hence by (1.5b), we must show the trivial representation 1 occurs exactly once in $E_1\oplus\ldots\oplus E_n$. But $E_n=1$, and more generally it is easy to show that E_k is in fact iself irreducible, of dimension $\binom{n}{k}$. (In fact, in terms of the Curtis-Richen weights (cf. [10]), the E_k $(1\leqq k\leqq n)$ are precisely the irreducibles having highest weight vectors with weights $(\mu_1,\ldots,\mu_{n-1})$ such that $\mu_a\neq 0$ for at most one a.) This completes the proof. □

3. Example: A Stable Splitting of $B(\mathbb{Z}/2)^3$

In this section we describe explicitly the splitting of $B(\mathbb{Z}/2)^3$ obtained by modular representation theory (for the splitting of $B(\mathbb{Z}/2)^2$, see [7]). The group $G=GL_3\mathbb{F}_2$ is a simple group of order 168. It has four irreducible representations over $\mathbb{F}_2$: $W_1=1$, $W_2=R$, $W_3=R^\#$, $W_4=St$. Here R is the standard representation of dimension three, and $R^\#$ is its contragredient. (Clearly these are distinct irreducibles - modular characters or highest weights can be used to distinguish R and $R^\#$. That these are all the irreducibles follows by counting 2-regular conjugacy classes, or from the weight theory.) Next, we need the dimensions of the projective covers P_i of W_i. By computing the complex character table and the decomposition matrix D for G one can easily obtain the Cartan matrix

$$C=D^TD=\begin{matrix}2&1&1&0\\1&3&2&0\\1&2&3&0\\0&0&0&1\end{matrix}.$$

Here the coefficient $C_{ij}=c_{ji}=\langle P_i,W_i\rangle$. Hence the P_i have dimensions 8, 16, 16, 8 respectively. However, this is easily seen directly as follows: Of course $P_4=W_4$, with dimension 8, and $P_3=P_2^\#$. Hence $\dim P_1+6\cdot\dim P_2=168-64=104$. On the other hand we have:

(3.1) **Proposition.** *Suppose* $\operatorname{Char}K=p$ *and* G *has a* p*-complement* H. *Then* 1_H^G *is the projective cover of the trivial representation, and* $e=|H|^{-1}\sum_{h\in H}h$ *is an associated primitive idempotent.*

Proof. By p-complement we mean a subgroup of order $[G:P]$, where P is a p-Sylow subgroup. In particular $p\nmid|H|$, so 1_H is KH-projective and hence 1_H^G is KG-projective. Since $\dim 1_H^G=|P|$, 1_H^G is indecomposable, and since $\operatorname{Hom}_{KG}(1_H^G,1_G)=\operatorname{Hom}_{KH}(1_H,1_H)\neq 0$, it must be the projective cover of 1_G. The second statement is then immediate since $1_H^G\cong e\cdot KG$.

In our situation there is a 2-complement $H=\mathbb{Z}/3\rtimes\mathbb{Z}/7$, obtained by thinking of $\mathbb{Z}/7$ as the group of units and $\mathbb{Z}/3$ as the Galois group of $\mathbb{F}_8$. Hence $\dim P_1=8$, and $\dim P_3=16$, as claimed.

(3.2) **Theorem.** *Let e_i, $1 \leqq i \leqq 4$, be primitive orthogonal idempotents associated to P_i, and let $X = B(\mathbb{Z}/2)^3$. Then there is a stable splitting $X = e_1X \vee 3e_2X \vee 3e_3X \vee 8e_4X$. The Poincaré series of $H^*(e_iX)$ is given by $f_i(t)[(1-t^4)(1-t^6)(1-t^7)]^{-1}$, where $f_i(t)$ is a polynomial whose coefficients are given in the following table:*

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
M	1	3	6	10	14	18	21	22	21	18	14	10	6	3	1	168
1	1	0	0	1	0	1	1	0	1	1	0	1	0	0	1	8
R	0	1	1	1	1	1	2	1	2	2	1	2	1	0	0	16
$R^\#$	0	0	1	2	1	2	2	1	2	1	1	1	1	1	0	16
St	0	0	0	0	1	1	1	2	1	1	1	0	0	0	0	8

Moreover $e_1X \cong B(H \tilde{\times} (\mathbb{Z}/2)^3) \cong B(\mathbb{Z}/3) \tilde{\times} GL_2\mathbb{F}_8)$, and

$$e_4X = \sum^{-4}(Sp^8S^0/Sp^4S^0) \vee \sum^{-3}(Sp^4S^0/Sp^2S^0),$$

where Sp^kS^0 is the k-fold symmetric product spectrum of S^0.

Proof. The first row of the table gives the Poincaré polynomial $F(M,t) = (1-t^4)(1-t^6)(1-t^7)(1-t)^{-3}$ of $M = S \otimes_{SG} \mathbb{F}_2$. The remaining rows give the multiplicity series $F(M, G, W_i; t)$; these agree with the f_i by (1.5b). (Recall that S^G is polynomial on generators of degree 4, 6, 7.) The total multiplicities, shown in the last column, are the dimensions of the P_i by (2.1). To obtain the rest of the table, first observe that one need only compute through degree 6: For M is a Poincare duality algebra, self-dual over G (see [6]). Thus $M_i = (M_{14-i})^\#$ as G-modules, and of course the middle module M_7 can be determined last since we know the total multiplicities. Moreover the multiplicity of St is already given in [8] (see also [6]). In any case one computes ad hoc that

$$F(S, G; t) = 1 + Rt + (R+R^\#)t^2 + (1+R+2R^\#)t^3 + (1+R+R^\#+St)t^4 + (1+2R+2R^\#+St)t^5 + (2+3R+3R^\#+St)t^6 \bmod t^7.$$

The details will be omitted, but note that the computation is greatly facilitated by making use of the A-action - for example, $Sq^2\colon S_3 \to S_5$ is injective. The table is then easily computed.

The statement about e_4 is a special case of the results of [8]. By (3.1), $H^*(e_1X) = S^H = H^*(H \tilde{\times} (\mathbb{Z}/2)^3)$; here the last equality is immediate from the Serre spectral sequence of the extension $(\mathbb{Z}/2)^3 \to H \tilde{\times} (\mathbb{Z}/2)^3 \to H$. Hence $e_1X \cong B(H \tilde{\times} (\mathbb{Z}/2)^3)$. As remarked in section two, $H^*(GL_2\mathbb{F}_8) = S^{\mathbb{F}_8^*}$. Hence if $\mathbb{Z}/3$ acts on $GL_2\mathbb{F}_8$ as the Galois group of $\mathbb{F}_8$, we have $H^*(\mathbb{Z}/3 \tilde{\times} GL_2\mathbb{F}_8) = S^H$, and $e_1X \cong B(\mathbb{Z}/3 \tilde{\times} GL_2\mathbb{F}_8)$.

Remark. The splitting can of course be refined by using the full monoid of self-maps of $(\mathbb{Z}/2)^3$. For examples, the splitting of e_4X arises in this way, and one can easily show that $B\mathbb{Z}/2$ splits off of e_2X.

References

1. Aguadé, J.: The cohomology of GL_2 of a finite field. Arch. Math. (Basel) **34**, 509–516 (1980)
2. Carlisle, D., Eccles, P., Hilditch, S., Ray, N., Schwartz, L., Walker, G., Wood, R.: Modular representations of $GL(2, p)$, splitting $\sum(\mathbb{C}p^\infty \times \mathbb{C}p^\infty)$ and the β-family as framed hypersurfaces. Preprint
3. Curtis, C., Reiner, I.: Representation Theory of Finite Groups and Associative Algebras. New York: Interscience Publishers 1962
4. Dixon, J., Puttaswamaiah, B.M.: Modular Representations of Finite Groups. New York: Academic Press 1977
5. Mann, B., Miller, E., Miller, H.: S^1-equivariant function spaces and characteristic classes, to appear
6. Mitchell, S.: Finite complexes with $A(n)$-free cohomology. Preprint
7. Mitchell, S., Priddy, S.: Symmetric product spectra and splittings of classifying spaces. To appear in Amer. J. Math.
8. Mitchell, S., Priddy, S.: Stable splittings derived from the Steinberg module. Topology **22**, 285–298 (1983)
9. Smith, J.: In preparation
10. Steinberg, R.: Lecture on Chevalley Groups. Yale University Lecture Notes 1967

Received May 2, 1984

Math. Z. 189, 11-17 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

A Multiparameter Strongly Superadditive Ergodic Theorem*

R. Emilion[1] and B. Hachem[2]

[1] Université P. & M. Curie, Laboratoire de Probabilités, Tour 56, 3ème Etage, F-75230 Paris Cedex 05, France

[2] Université de Bretagne Occidentale, Faculté des Sciences et Techniques, 6 avenue Le Gorgeu, F-29283 Brest, France

1. Introduction

The notion of strong superadditivity is due to R.T. Smythe [11]. The purpose of this article is to prove a pointwise ergodic theorem for multiparameter strongly superadditive processes with respect to a semi-group of positive operators on $L_p(X, \mathscr{F}, \mu)$. This generalizes the case of operators induced by a measure - preserving transformation (Kingman [10], Smythe [11]) and also Akcoglu-Sucheston's one-parameter theorem [4].

2. Notations and Definitions

Let $(X, \mathscr{F}, \mu)$ be a σ-finite measure space and T_1, T_2 two commuting positive operators on $L_p(X, \mathscr{F}, \mu)$, $1 \leqq p \leqq \infty$.

A sequence (or process) $F = \{F(n, m), n, m \in \mathbb{N}\}$, $F(n, m) \in L_p$, is called *superadditive* if

$$F(n+m, r) \geqq F(n, r) + T_1^n F(m, r)$$

(2.1) and

$$F(n, r+s) \geqq F(n, r) + T_2^r F(n, s)$$

for any $n, m, r, s \in \mathbb{N}$.

F will be *strongly superadditive* if the following conditions are satisfied

(2.2) $$F(n+m, r+s) \geqq F(n, r) + T_1^n F(m, r+s) + T_2^r F(n+m, s) - T_1^n T_2^r F(m, s)$$

and $F(n, 0) = F(0, n) = 0$

(2.3) $$\gamma(F) = \sup_{n, m} \frac{\|F(n, m)\|_p}{n, m} < +\infty.$$

* The present results have been announced in C.R. Acad. Sc. Paris (10 janvier 1983) Série I, p. 85-87

This notion, illustrated by interesting examples, is due to R.T. Smythe [11]. Also note that (2.2) implies (2.1).

Finally, we recall that an L_p-operator T is positive if $T(L_p^+) \subset L_p^+$ and that an L_1-*positive* operator T is called Markovian if $\int Tf d\mu = \int f d\mu$ for any $f \in L_1$.

3. Main Results

We will prove the following results.

(3.1) **Theorem.** *Let T_1 and T_2 be commuting L_1-Markovian operators which are also contractions on L_∞. Let F be a strongly superadditive process with respect to T_i $(i=1,2)$. Then* $\lim_{n \to +\infty} \frac{F(n,n)}{n^2}$ *exists a.e. on X.*

The main steps of the proof of (3.1) are a combinatorial lemma (Lemma (4.2) below) and a maximal lemma (Lemma (4.3)) obtained by using Brunel's operator [5] and the one-dimensional proof (Akcoglu-Sucheston [4] or Hachem [9]). The rest of the proof uses some arguments of Akcoglu-Krengel [2].

It is well-known that 3.1 fails even in the additive case if the operators T_i $(i=1,2)$ are only Markovian. On the other hand the trivial case $T_i = 0$ shows that the convergence does not hold on X if T_i is a positive *contraction* on L_1 and L_∞. However it is easy to deduce from (3.1) the

(3.2) **Theorem.** *Let T_1 and T_2 be commuting positive contractions on L_1 and L_∞ and F a strongly superadditive process. Then,* $\lim_{n \to +\infty} \frac{F(n,n)}{n^2}$ *exists a.e. on Y where Y is the maximal support of a T_1, T_2-invariant non negative function.*

An extra-condition on the process gives the convergence on $X \setminus Y$. It concerns the following important sequence

$$\phi(m) = \frac{1}{m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} a_{ij} \tag{3.3}$$

with

$$a_{ij} = [F(i,j) - T_1 F(i-1,j) - T_2 F(i,j-1) + T_1 T_2 F(i-1,j-1)].$$

Note that $\int \phi(m) d\mu = \frac{1}{m^2} \int F(m,m) d\mu$ if $T_i (i=1,2)$ is Markovian.

(3.4) **Theorem.** *With the same conditions as in* (3.2), $\lim_{n \to +\infty} \frac{F(n,n)}{n^2}$ *exists a.e. on $X \setminus Y$ (and then on X) if* $\liminf_{m \to +\infty} \|\phi(m)\|_1 < +\infty$.

Finally we will consider the case of L_p-processes $(1 < p < \infty)$. The following theorem generalizes the additive case ([1, 7]) and the one dimensional result [8].

(3.5) **Theorem.** *Let T_1 and T_2 be commuting positive contractions on L_p $(1<p<\infty)$ and F a strongly superadditive process with respect to T_i. If* $\liminf_{m\to+\infty} \|\phi(m)\|_p < +\infty$ *then* $\lim_{n\to+\infty} \frac{F(n,n)}{n^2}$ *exists a.e. and in L_p-norm.*

4. Proofs

Proof of Theorem (3.1). (2.1) implies the inequality

$$F(n,m) \geqq \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} T_1^i T_2^j F(1,1) = A(n,m).$$

Then $A(n,m)$ is an additive process and $F-A$ is a positive strongly superadditive process. Since A. Brunel's theorem [5] asserts the convergence of $\frac{A(n,n)}{n^2}$ a.e. on X, we may assume for the rest of the proof that F is a positive process.

(4.1) **Lemma.** $\gamma(F) = \lim_{n,m\to+\infty} \int \frac{F(n,m)}{n,m}\, d\mu.$

Proof. Indeed, since the operators T_i $(i=1,2)$ are markovian, (2.1) implies that the two-parameter family of real numbers $\int \frac{F(n,m)}{n,m}\, d\mu$ is superadditive. This yields (4.1).

In the case of transformations the following is due to Smythe [11]

(4.2) **Lemma.** *Let $\phi(m)$, $m \geqq 1$, be defined as in* (3.3). *For any n: $1 \leqq n < m$ one has*

$$\left(1-\frac{n}{m}\right)^2 F(n,n) \leqq \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} T_1^{k_1} T_2^{k_2} \phi(m).$$

Proof. Since F is positive and satisfies (2.2) one has $a_{ij} \geqq F(1,1) \geqq 0$.

Let k_1 and k_2 be such that $0<k_i<n<m$. Then

$$m^2 T_1^{k_1} T_2^{k_2} \phi(m) \geqq \sum_{i=n+1-k_1}^{m-k_1} \sum_{j=n+1-k_2}^{m-k_2} T_1^{k_1} T_2^{k_2} a_{ij} = \sum_{i=n+1}^{m} \sum_{j=n+1}^{m} T_1^{k_1} T_2^{k_2} a_{i-k_1, j-k_2}.$$

Hence

$$m^2 \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} T_1^{k_1} T_2^{k_2} \phi(m) \geqq \sum_{i=n+1}^{m} \sum_{j=n+1}^{m} \sum_{k_1=0}^{n-1} \sum_{k_2=0}^{n-1} T_1^{k_1} T_2^{k_2} a_{i-k_1, j-k_2}.$$

Now, if we observe that

$$a_{ij} = F(i,j) - T_1 F(i-1,j) - T_2[F(i,j-1) - T_1 F(i-1,j-1)]$$

we see that:

$$\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-1} T_1^{k_1}T_2^{k_2}a_{i-k_1,j-k_2}$$
$$=\sum_{k_2=0}^{n-1} T_2^{k_2}\sum_{k_1=0}^{n-1}(T_1^{k_1}F(i-k_1,j-k_2)-T_1^{k_1+1}F(i-k_1-1,j-k_2))$$
$$-\sum_{k_2=0}^{n-1} T_2^{k_2+1}\sum_{k_1=0}^{n-1}(T_1^{k_1}F(i-k_1,j-k_2-1)-T_1^{k_1+1}F(i-k_1,j-k_2-1))$$
$$=\sum_{k_2=0}^{n-1}(T_2^{k_2}F(i,j-k_2)-T_2^{k_2+1}F(i,j-k_2-1))$$
$$-T_1^n\sum_{k_2=0}^{n-1}(T_2^{k_2}F(i-n,j-k_2)-T_2^{k_2+1}F(i-n,j-k_2-1))$$
$$=F(i,j)-T_2^nF(i,j-n)-T_1^nF(i-n,j)+T_1^nT_2^nF(i-n,j-n)\geqq F(n,n)$$

because of (2.2). Therefore we obtain (4.2):

$$m^2\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-1} T_1^{k_1}T_2^{k_2}\phi(m)\geqq(m-n)^2F(n,n).$$

(4.3) **Lemma.** (Maximal ergodic lemma.) *There exists a constant $c>0$ and a function $\delta\in L_1^+$ such that $\int\delta d\mu\leqq\gamma(F)$ and $\limsup_{n\to+\infty}\frac{F(n,n)}{n^2}\leqq c\delta$. Therefore we have $\mu\left(\limsup_{n\to+\infty}\frac{F(n,n)}{n^2}\geqq\alpha\right)\leqq\frac{c\gamma(F)}{\alpha}$ for any $\alpha>0$.*

Proof. Consider the following operator defined by A. Brunel ([5], 3.7)

$$U=f(T_1)f(T_2)\quad\text{with } f(x)=\frac{1}{x}(1-\sqrt{1-x})=\sum_{j=0}^{\infty}\alpha_jx^j.$$

A. Brunel has proved that there exists a constant $c>0$ such that

$$\frac{1}{n^2}\sum_{k_1=0}^{n-1}\sum_{k_2=0}^{n-1} T_1^{k_1}T_2^{k_2}\phi(m)\leqq\frac{c}{[\sqrt{n}]+1}\sum_{j\leqq[\sqrt{n}]}U^j\phi(m),$$

for any $n\geqq1$; ($[x]$ denotes the integer part of the real x). Hence Lemma (4.2) yields

$$\left(1-\frac{n}{m}\right)^2\frac{F(n,n)}{n^2}\leqq\frac{c}{[\sqrt{n}]+1}\sum_{j\leqq[\sqrt{n}]}U^j\phi(m).$$

Now, $\alpha_j\geqq0$ and $\sum_{j=0}^{\infty}\alpha_j=1$ imply that $f(T_i)$ and U are Markovian. Therefore, since $\int\phi(m)d\mu=\frac{1}{m^2}\int F(m,m)d\mu\leqq\gamma(F)$ (3.3), the arguments in Akcoglu-Sucheston [4] show that there exists $\delta'\in L_1^+$ satisfying

$$\int\delta'd\mu\leqq\gamma(F)\quad\text{and}\quad\frac{F(n,n)}{n^2}\leqq\frac{c}{[\sqrt{n}]+1}\sum_{j\leqq[\sqrt{n}]}U^j\delta'.$$

Since U is Markovian and an L_∞-contraction, A. Brunel's ergodic theorem ([5], 4.1) shows that

$$\delta = \lim_{n\to\infty} \frac{1}{[\sqrt{n}]+1} \sum_{j\leqq[\sqrt{n}]} U^j \delta'$$

exists a.e. and in L_1-norm and Fatou's lemma implies that $\int \delta d\mu \leqq \gamma(F)$. Hence $\limsup_{n\to+\infty} \frac{F(n,n)}{n^2} \leqq c\delta$ and $\mu\left(\limsup_{n\to+\infty} \frac{F(n,n)}{n^2} > \alpha\right) \leqq \frac{c\gamma(F)}{\alpha}$ for any $\alpha>0$.

(4.4) The rest of the proof of Theorem (3.1) follows the arguments of Akcoglu-Krengel ([2], p. 58–59):

Let $\varepsilon>0$ and m_0 be such that $\int \frac{F(m_0,m_0)}{m_0^2} d\mu > \gamma(F)-\varepsilon$.

Put

$$G(i,j) = \sum_{k_1=0}^{i-1} \sum_{k_2=0}^{j-1} T_1^{k_1 m_0} T_2^{k_2 m_0} \frac{F(m_0,m_0)}{m_0^2}$$

and

$$H(i,j) = \frac{F(im_0, jm_0)}{m_0^2} - G(i,j).$$

Then G is additive and H is strongly superadditive and positive with respect to $T_1^{m_0}$ and $T_2^{m_0}$. The additive theorem of A. Brunel [5] shows that $\lim_{i\to+\infty} \frac{G(i,i)}{i^2}$ exists, hence,

$$\begin{aligned} 0 &\leqq \limsup_{i\to\infty} \frac{F(i,i)}{i^2} - \liminf_{i\to\infty} \frac{F(i,i)}{i^2} \\ &= \limsup_{i\to\infty} \frac{F(im_0, im_0)}{i^2 m_0^2} - \liminf_{i\to\infty} \frac{F(im_0, im_0)}{i^2 m_0^2} \end{aligned}$$

since F is positive and hence increasing

$$= \limsup_{i\to\infty} \frac{H(i,i)}{i^2} - \liminf_{i\to\infty} \frac{H(i,i)}{i^2} \leqq \limsup_{i\to\infty} \frac{H(i,i)}{i^2} \leqq c\cdot\delta$$

where $\int \delta d\mu \leqq \gamma(H) = \gamma(F) - \gamma(G)$ (Lemmas (4.3) and (4.1))
$\leqq \varepsilon$.

Therefore

$$\limsup_{i\to\infty} \frac{F(i,i)}{i^2} = \liminf_{i\to\infty} \frac{F(i,i)}{i^2} \quad \text{a.e. on } X.$$

The Theorem (3.1) is completely proved.

Proof of Theorem (3.2). Brunel's results ([5], (3.4) show that if Y denotes the support of the largest positive function which is invariant by T_1 and T_2, then $L_1(Y)$ and $L_1(X\setminus Y)$ are invariant by T_1 and T_2. Hence $1_Y F$ is a strongly superadditive process with respect to the restrictions of T_1 and T_2 to $L_1(Y)$. Theorem (3.2) is then a consequence of Theorem (3.1) because these restrictions are Markovian.

Proof of Theorem (3.4). The proof of Lemma (4.3) shows that

$$\left(1-\frac{n}{m}\right)^2 \frac{F(n,n)}{n^2} \leqq \frac{c}{[\sqrt{n}]+1} \sum_{j \leqq [\sqrt{n}]} U^j \phi(m).$$

Il $K = \liminf \|\phi(m)\| < +\infty$, then the one-dimensional proofs ([4, 9]) show that there exists $h \in L_1^+$ verifying $\int h d\mu \leqq K$ and

$$\frac{F(n,n)}{n^2} \leqq \frac{c}{[\sqrt{n}]+1} \sum_{j \leqq [\sqrt{n}]} U^j h.$$

Indeed these proofs only use the conditions $\sup_m \|\phi(m)\| < +\infty$ (which is satisfied for a subsequence of $\phi(m)$) and U is an L_1-positive contraction.

Now, Brunel's theorem ([5], 1.2) implies that $\frac{1}{[\sqrt{n}]+1} \sum_{j \leqq [\sqrt{n}]} U^j h$ converges a.e. and strongly to an U-invariant function $\bar{h}$. Again a result of Brunel ([5], Prop. 3.6) shows that $T_1 \bar{h} = T_2 \bar{h} = \bar{h}$. Hence $\bar{h} = 0$ on $X \setminus Y$ and $\lim_{n \to +\infty} \frac{F(n,n)}{n^2} = 0$ a.e. on $X \setminus Y$.

Proof of Theorem (3.5). As in the L_1-case, we may assume that F is a positive process because of the multiparameter additive theorem of S.A. McGrath [7].

Since $1 < p < \infty$ and since $\liminf_{m \to +\infty} \|\phi(m)\|_p < +\infty$, there exists $\phi \in L_p^+$ such that $w - \lim_{m_i \to +\infty} \phi(m_i) = \phi$ for a subsequence m_i.

Note that Brunel's operator can be constructed also for operators in L_p. We then obtain:

$$\left(1-\frac{m_i}{n}\right)^2 \leqq \frac{F(n,n)}{n^2} \leqq \frac{c}{[\sqrt{n}]+1} \sum_{j \leqq [\sqrt{n}]} U^j \phi(m_i) \quad \text{for } m_i \geqq n.$$

As $m_i \to +\infty$, we obtain $\frac{F(n,n)}{n^2} \leqq \frac{c}{[\sqrt{n}]+1} \sum_{j \leqq [\sqrt{n}]} U^j \phi$.

Let E be the maximal support of a U-invariant non negative function ([3], Proposition 3.1). $L_p(E)$ and $L_p(E^c)$ are invariant under U ([3], Proposition 3.4) and then invariant under T_1 and T_2. Hence $1_E F$ is strongly superadditive with respect to the restriction of T_1 and T_2 to $L_p(E)$. Since any U-invariant function is also T_1 and T_2-invariant ([5], 3.6), these restrictions can be considered as Markovian L_∞-contractions after a classical change of measure. Therefore $\lim_{n \to +\infty} \frac{F(n,n)}{n^2}$ exists a.e. on E because of Theorem (3.1).

On the other hand Akcoglu's theorem [1] shows that

$$\bar{\Phi} = \lim_{n \to +\infty} \frac{1}{[\sqrt{n}]+1} \sum_{j \leqq [\sqrt{n}]} U^j \phi$$

exists a.e. and verifies $U\bar{\Phi}=\bar{\Phi}$. Hence $\Phi=0$ a.e. on $X\setminus E$ and

$$\lim_{n\to+\infty}\frac{F(n,n)}{n^2}=0$$

a.e. on $X\setminus E$.

Since $\left\|\sup_n \frac{F(n,n)}{n^2}\right\|_p \leqq c\frac{p}{p-1}\|\phi\|_p$ by Akcoglu's theorem [1], we also have the L_p-norm convergence of $\frac{F(n,n)}{n^2}$ as $n\to+\infty$.

References

1. Akcoglu, M.A.: A pointwise ergodic theorem in L_p-spaces. Canad. J. Math. **27**, 1075–1082 (1985)
2. Akcoglu, M.A., Krengel, U.: Ergodic theorem for superadditive processes. J. Reine Angew. Math. **323**, 53–67 (1981)
3. Akcoglu, M., Sucheston, L.: On convergence of Iterates of positive contractions in L_p-spaces. J. Approximation Theory **13**, 348–362 (1975)
4. Akcoglu, M., Sucheston, L.: A ratio ergodic theorem for superadditive processes. Z. Wahrscheinlichkeitstheorie verw. Gebiete **44**, 269–278 (1978)
5. Brunel, A.: Théorème ergodique ponctuel pour un semi-groupe commutatif finiment engendré de contractions de L_1. Ann. Inst. H. Poincaré **XI**, n° 4, 327–343 (1973)
6. Derriennic, Y., Krengel, U.: Subadditive mean ergodic theorems. Erg. Th. and Dynam. Syst. **1**, 33–48 (1981)
7. Mc Grath, S.A.: Some ergodic theorems for commuting L_1 contractions. Studia Math. **70**, 165–172 (1980)
8. Hachem, B.: Un théorème ergodique sur-additif dans L_p $(1<p<\infty)$. C.R. Acad. Sc. Paris t. 292 (9 mars 1981), p. 531–534
9. Hachem, B.: Quelques théorèmes ergodiques sur-additifs dans L_p $(1\leqq p<\infty)$. Ann. Inst. H. Poincaré **XVIII**, n° 3 (1982)
10. Kingman, J.F.C.: Subadditive ergodic theory. Ann. Probability **1**, 883–905 (1973)
11. Smythe, R.T.: Multiparameter subadditive processes. Ann. Probability **4**, 772–782 (1976)

Received January 28, 1984; received in final form June 14, 1984

Math. Z. 189, 19-39 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Generalized Clifford Correspondences for Group Characters

Everett C. Dade*

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

Suppose that ϕ is a complex irreducible character of a normal subgroup N of a finite group G. Clifford's Theory tells us that induction is a bijection of the set $\mathrm{Irr}(G\{\phi\}|\phi)$ (of all irreducible characters χ of the stabilizer $G\{\phi\}$ of ϕ in G such that χ lies over ϕ) onto the corresponding set $\mathrm{Irr}(G|\phi)$ for G (see Hauptsatz V.17.3 in [5]). In this paper we consider the consequences of the existence of such a bijection for an irreducible character of an arbitrary subgroup of G. More specifically, we define an *inductive source* for G to be any irreducible character ϕ of a subgroup of G such that induction is a bijection of $\mathrm{Irr}(G\{\phi\}|\phi)$ onto $\mathrm{Irr}(G|\phi)$, and the associated IS-*correspondent* $\psi\{\phi\}$ of any $\psi \in \mathrm{Irr}(G|\phi)$ to be the unique character in $\mathrm{Irr}(G\{\phi\}|\phi)$ from which ψ is induced. We show that these sources and correspondents have most of the usual properties of irreducible characters of normal subgroups and the associated Clifford correspondents, including the correct 'compounding properties' in Theorem (3.1) and its corollaries below. Indeed the IS-correspondents are much easier to handle than (repeated) Clifford correspondents due to their closure under most constructions (see §4 below). In §5 and §6 we show that IS-correspondents reduce to repeated Clifford correspondents in certain interesting cases of nilpotent and odd normal subgroups.

Except in Theorem (6.5) below, we work over arbitrary fields of characteristic zero. The only problem this usually causes is that we must use isometries, instead of just bijections, of irreducible characters.

§1. Characters

Throughout this paper we shall fix a field $\mathfrak{F}$ of characteristic zero and study a finite group G. We denote by $\mathfrak{F}G$ the group algebra of G over $\mathfrak{F}$ and by $\mathrm{Ch}(\mathfrak{F}G)$ the set of all $\mathfrak{F}$-*characters* of G, i.e., of characters ϕ afforded by finite-dimensional $\mathfrak{F}G$-modules. We consider such ϕ to be functions from G (and not

* This research was supported by a grant from the National Science Foundation.

$\mathfrak{F}G$) to $\mathfrak{F}$. So any $\mathfrak{F}$-character ψ of any subgroup H of G determines H as its domain of definition:

$$H = \mathrm{Dom}(\psi).$$

Hence we may define the set $\mathrm{Sch}(\mathfrak{F}G)$ of all $\mathfrak{F}$-*subcharacters* of G to be the disjoint union:

$$\mathrm{Sch}(\mathfrak{F}G) = \dot{\bigcup}_{H \leqq G} \mathrm{Ch}(\mathfrak{F}H). \tag{1.1}$$

Here, as usual, we write '$H \leqq G$' to indicate that H is a subgroup of G.

The *inner product* $(\cdot, \cdot)$ from $\mathrm{Ch}(\mathfrak{F}G) \times \mathrm{Ch}(\mathfrak{F}G)$ to the set $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ of all natural numbers in $\mathfrak{F}$ satisfies:

$$(\phi, \psi) = |G|^{-1} \sum_{\sigma \in G} \phi(\sigma^{-1})\psi(\sigma) \tag{1.2a}$$

$$= \dim_{\mathfrak{F}}(\mathrm{Hom}_{\mathfrak{F}G}(\mathfrak{V}, \mathfrak{W})), \tag{1.2b}$$

where $\phi, \psi \in \mathrm{Ch}(\mathfrak{F}G)$ are afforded by $\mathfrak{F}G$-modules $\mathfrak{V}, \mathfrak{W}$, respectively. Here we use the usual conventions that $|G| \in \mathbb{N}$ is the order of G, that $\dim_{\mathfrak{F}}(\ldots)$ is the $\mathfrak{F}$-dimension of ... and that $\mathrm{Hom}_{\mathfrak{F}G}(\mathfrak{V}, \mathfrak{W})$ is the $\mathfrak{F}$-vector space of all $\mathfrak{F}G$-homomorphisms of $\mathfrak{V}$ into $\mathfrak{W}$.

If $H \leqq G$, then induction sends any $\phi \in \mathrm{Ch}(\mathfrak{F}H)$ afforded by some right $\mathfrak{F}H$-module $\mathfrak{V}$ into the character $\phi^G \in \mathrm{Ch}(\mathfrak{F}G)$ afforded by the induced $\mathfrak{F}G$-module $\mathfrak{V}^G = \mathfrak{V} \otimes_{\mathfrak{F}H} \mathfrak{F}G$. Evidently tensoring with the identity map Id of $\mathfrak{F}G$ onto itself is a monomorphism of the $\mathfrak{F}$-algebra $\mathrm{End}_{\mathfrak{F}H}(\mathfrak{V}) = \mathrm{Hom}_{\mathfrak{F}H}(\mathfrak{V}, \mathfrak{V})$ of all $\mathfrak{F}H$-endomorphisms of $\mathfrak{V}$ into the corresponding algebra $\mathrm{End}_{\mathfrak{F}G}(\mathfrak{V}^G)$ for $\mathfrak{V}^G$. It follows from this and (1.2b) that:

$$(\phi, \phi) \leqq (\phi^G, \phi^G), \quad \textit{for all } \phi \in \mathrm{Sch}(\mathfrak{F}G), \tag{1.3a}$$

(1.3b) *Equality holds in* (1.3a) *if and only if the* $\mathfrak{F}$*-algebras* $\mathrm{End}_{\mathfrak{F}\mathrm{Dom}(\phi)}(\mathfrak{V})$ *and* $\mathrm{End}_{\mathfrak{F}G}(\mathfrak{V}^G)$ *are isomorphic for any* $\mathfrak{F}\,\mathrm{Dom}(\phi)$*-module* $\mathfrak{V}$ *affording* ϕ.

An *irreducible* $\mathfrak{F}$-character ϕ of G is, of course, the $\mathfrak{F}$-character afforded by some simple $\mathfrak{F}G$-module. We denote by $\mathrm{Irr}(\mathfrak{F}G)$ the set of all irreducible $\mathfrak{F}$-characters of G. Following (1.1) we define the set $\mathrm{Irs}(\mathfrak{F}G)$ of all *irreducible* $\mathfrak{F}$-*subcharacters* of G to be the disjoint union:

$$\mathrm{Irs}(\mathfrak{F}G) = \dot{\bigcup}_{H \leqq K} \mathrm{Irr}(\mathfrak{F}H). \tag{1.4}$$

Because $\mathfrak{F}$ has characteristic zero, the algebra $\mathfrak{F}H$ is semi-simple for any $H \leqq G$. Hence an $\mathfrak{F}H$-module $\mathfrak{V}$ is simple if and only if $\mathrm{End}_{\mathfrak{F}H}(\mathfrak{V})$ is a division algebra. This and (1.3b) tell us that:

(1.5) *If* $\phi \in \mathrm{Irs}(\mathfrak{F}G)$ *and* $(\phi, \phi) = (\phi^G, \phi^G)$, *then* $\phi^G \in \mathrm{Irr}(\mathfrak{F}G)$.

Of course it is quite possible to have $\phi^G \in \mathrm{Irr}(\mathfrak{F}G)$ without the equality $(\phi, \phi) = (\phi^G, \phi^G)$, since $\mathfrak{F}$ need not be a splitting field for G.

A similar discussion can be made for the restriction ψ_H to $H \leqq G$ of any $\psi \in \mathrm{Ch}(\mathfrak{F}G)$. If $\mathfrak{W}$ is an $\mathfrak{F}G$-module affording ψ, then its restriction $\mathfrak{W}_H$ to an $\mathfrak{F}H$-module affords $\psi_H \in \mathrm{Ch}(\mathfrak{F}H)$. Clearly $\mathrm{End}_{\mathfrak{F}G}(\mathfrak{W})$ is a subalgebra of $\mathrm{End}_{\mathfrak{F}H}(\mathfrak{W}_H)$. As in the case of (1.3) and (1.5), this implies that:

(1.6a) $(\psi, \psi) \leqq (\psi_H, \psi_H)$, *for all* $\psi \in \mathrm{Ch}(\mathfrak{F}G)$ *and subgroups* H *of* G,

(1.6b) *Equality holds in* (1.6a) *if and only if the* $\mathfrak{F}$*-algebras* $\mathrm{End}_{\mathfrak{F}G}(\mathfrak{W})$ *and* $\mathrm{End}_{\mathfrak{F}H}(\mathfrak{W}_H)$ *are isomorphic for any* $\mathfrak{F}G$*-module* $\mathfrak{W}$ *affording* ψ,

(1.6c) *If* $\psi \in \mathrm{Irr}(\mathfrak{F}G)$ *and* $(\psi, \psi) = (\psi_H, \psi_H)$ *for some* $H \leqq G$, *then* $\psi_H \in \mathrm{Irr}(\mathfrak{F}H)$.

We shall make little use of this remark.

There is a natural partial ordering $\leqq$ on $\mathrm{Irs}(\mathfrak{F}G)$ given by:

(1.7) *If* $\phi, \psi \in \mathrm{Irs}(\mathfrak{F}G)$, *then* $\phi \leqq \psi$ *if and only if* $\mathrm{Dom}(\phi) \leqq \mathrm{Dom}(\psi)$ *and* $(\phi^{\mathrm{Dom}(\psi)}, \psi) \neq 0$ (*i.e.,* ψ *is an irreducible* $\mathfrak{F}$*-constituent of* $\phi^{\mathrm{Dom}(\psi)}$).

Of course the condition $(\phi^{\mathrm{Dom}(\psi)}, \psi) \neq 0$ in (1.7) is equivalent to $(\phi, \psi_{\mathrm{Dom}(\phi)}) \neq 0$, i.e., to ϕ's being an irreducible $\mathfrak{F}$-constituent of $\psi_{\mathrm{Dom}(\phi)}$, by the Frobenius Reciprocity Law. If X is any 'functor', such as Irr or Irs, mapping any group algebra $\mathfrak{F}H$ into a subset of $\mathrm{Irs}(\mathfrak{F}H)$, then we define:

$$X(\mathfrak{F}H \mid \phi) = \{\psi \in X(\mathfrak{F}H) \mid \psi \geqq \phi\}, \tag{1.8a}$$

$$X(\chi \mid \mathfrak{F}H) = \{\psi \in X(\mathfrak{F}H) \mid \chi \geqq \psi\}, \tag{1.8b}$$

$$X(\chi \mid \mathfrak{F}H \mid \phi) = \{\psi \in X(\mathfrak{F}H) \mid \chi \geqq \psi \geqq \phi\}, \tag{1.8c}$$

for any $H \leqq G$ and any $\chi, \phi \in \mathrm{Irs}(\mathfrak{F}G)$. For example, we have:

(1.9a) $\mathrm{Irr}(\mathfrak{F}G \mid \phi)$ *is the set of all irreducible* $\mathfrak{F}$*-constituents of* ϕ^G,

(1.9b) $\mathrm{Irs}(\mathfrak{F}G \mid \phi) = \dot{\bigcup}_{\mathrm{Dom}(\phi) \leqq K \leqq G} \mathrm{Irr}(\mathfrak{F}K \mid \phi)$,

(1.9c) $\mathrm{Irr}(\psi \mid \mathfrak{F}H)$ *is the set of all irreducible* $\mathfrak{F}$*-constituents of* ψ_H,

whenever $\phi \in \mathrm{Irs}(\mathfrak{F}G)$, $\psi \in \mathrm{Irr}(\mathfrak{F}G)$ and $H \leqq G$.

The group G acts by conjugation on $\mathrm{Sch}(\mathfrak{F}G)$. Conjugation by any $\sigma \in G$ sends an element $\tau \in G$ into $\tau^\sigma = \sigma^{-1}\tau\sigma \in G$, a subgroup $H \leqq G$ to $H^\sigma = \sigma^{-1}H\sigma \leqq G$ and a character $\phi \in \mathrm{Ch}(\mathfrak{F}H)$ to $\phi^\sigma \in \mathrm{Ch}(\mathfrak{F}H^\sigma)$ given by:

$$\phi^\sigma(\tau^\sigma) = \phi(\tau), \quad \textit{for all } \tau \in H.$$

This action of G leaves invariant the subset $\mathrm{Irs}(\mathfrak{F}G)$, its partial ordering $\leqq$, induction, restriction and inner products. If $H \leqq G$ and $\phi \in \mathrm{Sch}(\mathfrak{F}G)$, then we denote by $H\{\phi\}$ the stabilizer:

$$H\{\phi\} = \{\sigma \in H \mid \phi^\sigma = \phi\} \leqq H \tag{1.10}$$

of ϕ in H. Notice that:

$$\mathrm{Dom}(\phi) \trianglelefteq G\{\phi\}, \quad \textit{for all } \phi \in \mathrm{Sch}(\mathfrak{F}G), \tag{1.11}$$

where, as usual, '$\trianglelefteq$' means 'is a normal subgroup of'.

§ 2. Inductive Sources

We say that *induction to G is an isometry on* $\mathrm{Irr}(\mathfrak{F}H|\phi)$, for some $H \leqq G$ and $\phi \in \mathrm{Irs}(\mathfrak{F}H)$, if:

$$(\psi, \chi) = (\psi^G, \chi^G), \quad \textit{for all } \psi, \chi \in \mathrm{Irr}(\mathfrak{F}H|\phi).$$

Such inductive isometries relate $\mathrm{Irr}(\mathfrak{F}H|\phi)$ to $\mathrm{Irr}(\mathfrak{F}G|\phi)$.

(2.1) **Proposition.** *Suppose that induction to G is an isometry on* $\mathrm{Irr}(\mathfrak{F}H|\phi)$, *for some* $H \leqq G$ *and* $\phi \in \mathrm{Irs}(\mathfrak{F}H)$. *Then induction is an isometric bijection of* $\mathrm{Irr}(\mathfrak{F}H|\phi)$ *onto* $\mathrm{Irr}(\mathfrak{F}G|\phi)$.

Proof. Statement (1.5) implies that $\psi^G \in \mathrm{Irr}(\mathfrak{F}G)$, for any $\psi \in \mathrm{Irr}(\mathfrak{F}H|\phi)$. Since distinct $\psi, \chi \in \mathrm{Irr}(\mathfrak{F}G|\phi)$ satisfy:

$$0 = (\psi, \chi) < (\psi, \psi)$$

by (1.2b), their images ψ^G, χ^G are distinct. From (1.9a) it follows that:

$$\mathrm{Irr}(\mathfrak{F}G|\phi) = \bigcup_{\psi \in \mathrm{Irr}(\mathfrak{F}H|\phi)} \mathrm{Irr}(\mathfrak{F}G|\psi) = \bigcup_{\psi \in \mathrm{Irr}(\mathfrak{F}H|\phi)} \{\psi^G\}.$$

Thus the proposition holds.

The following consequence of Mackey's Formula is crucial.

(2.2) **Theorem.** *Fix* $H \leqq G$ *and* $\phi \in \mathrm{Irs}(\mathfrak{F}H)$. *Then induction to G is an isometry on* $\mathrm{Irr}(\mathfrak{F}H|\phi)$ *if and only if:*

$$((\phi^\sigma)_{\mathrm{Dom}(\phi^\sigma) \cap \mathrm{Dom}(\phi)}, \phi_{\mathrm{Dom}(\phi^\sigma) \cap \mathrm{Dom}(\phi)}) = 0 \tag{2.3}$$

for all $\sigma \in G - H$.

Proof. Let D be $\mathrm{Dom}(\phi)$. Mackey's Formula (Satz V.16.9 in [5]) says that:

$$(\phi^G, \phi^G) = \sum_\sigma ((\phi^\sigma)_{D^\sigma \cap D}, \phi_{D^\sigma \cap D}), \tag{2.4}$$

where σ runs over representatives for the D, D-double cosets $D\sigma D$ in G. The summands in (2.4) for which $D\sigma D \subseteq H$ add up to (ϕ^H, ϕ^H) by Mackey's Formula. Equality (2.3) for all $\sigma \in G - H$ just says that the remaining summands in (2.4) are zero. Since all summands in (2.4) are nonnegative, we conclude that (2.3) holds for all $\sigma \in G - H$ if and only if:

$$(\phi^H, \phi^H) = (\phi^G, \phi^G). \tag{2.5}$$

Let $\chi_1, \ldots, \chi_n$ be the distinct elements of $\mathrm{Irr}(\mathfrak{F}H|\phi)$. By (1.9a) there are integers $m_i > 0$, for $i = 1, \ldots, n$, such that:

$$\phi^H = \sum_{i=1}^n m_i \chi_i.$$

It follows that:

$$\phi^G = \sum_{i=1}^n m_i \chi_i^G$$

and hence that:

$$(\phi^G, \phi^G) = \sum_{i,j=1}^{n} m_i m_j (\chi_i^G, \chi_j^G) \quad \text{and} \quad (\phi, \phi) = \sum_{i,j=1}^{n} m_i m_j (\chi_i, \chi_j).$$

If $i \neq j$, then:

$$m_i m_j (\chi_i^G, \chi_j^G) \geqq 0 = (\chi_i, \chi_j),$$

with equality if and only if $(\chi_i^G, \chi_j^G) = 0$. If $i = j$, then:

$$m_i^2 (\chi_i^G, \chi_i^G) \geqq m_i^2 (\chi_i, \chi_i)$$

by (1.3a), with equality if and only if $(\chi_i^G, \chi_i^G) = (\chi_i, \chi_i)$. It follows that (2.5) holds if and only if induction is an isometry on $\mathrm{Irr}(\mathfrak{F}H|\phi)$. So the theorem is proved.

We shall apply Theorem (2.2) to the special case in which H is the stabilizer $G\{\phi\}$ of ϕ in G. We define an *inductive source* of $\mathfrak{F}G$ to be any $\phi \in \mathrm{Irr}(\mathfrak{F}G)$ such that induction to G is an isometry on $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\phi)$. We denote by $\mathrm{IS}(\mathfrak{F}G)$ the set of all inductive sources of $\mathfrak{F}G$.

(2.6) **Corollary.** *The set* $\mathrm{IS}(\mathfrak{F}G)$ *consists of all* $\phi \in \mathrm{Irr}(\mathfrak{F}G)$ *satisfying* (2.3) *for all* $\sigma \in G - G\{\phi\}$.

Proof. This is an immediate consequence of Theorem 2.2.

(2.7) **Corollary** (Clifford). *If* $N \trianglelefteq G$, *then* $\mathrm{Irr}(\mathfrak{F}N) \subseteq \mathrm{IS}(\mathfrak{F}G)$.

Proof. If $\phi \in \mathrm{Irr}(\mathfrak{F}N)$, then $\mathrm{Dom}(\phi^\sigma) = \mathrm{Dom}(\phi) = N$ for any $\sigma \in G - G\{\phi\}$. Since ϕ^σ and ϕ are then distinct members of $\mathrm{Irr}(\mathfrak{F}N)$, it follows that (2.3) holds. Thus $\phi \in \mathrm{IS}(\mathfrak{F}G)$ by Corollary (2.6).

The following corollary of Theorem 2.2 is useful enough to be called a theorem in its own right.

(2.8) **Theorem.** *Suppose that* $\phi \in \mathrm{Irr}(\mathfrak{F}G)$, *that* $\mathrm{Dom}(\phi) \leqq H \leqq G\{\phi\}$ *and that induction to* G *is an isometry on* $\mathrm{Irr}(\mathfrak{F}H|\phi)$. *Then* $\phi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $H = G\{\phi\}$.

Proof. We know from Theorem (2.2) that (2.3) holds for all $\sigma \in G - H$. Clearly no $\sigma \in G\{\phi\}$ can satisfy (2.3). Therefore $G\{\phi\} \leqq H$. Since we have assumed that $H \leqq G\{\phi\}$, this implies the theorem.

(2.9) **Corollary** (Mackey). *If* $\phi \in \mathrm{Irr}(\mathfrak{F}G)$ *and* $(\phi, \phi) = (\phi^G, \phi^G)$, *then* $\phi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $G\{\phi\} = \mathrm{Dom}(\phi)$.

Proof. Apply the theorem to $H = \mathrm{Dom}(\phi)$.

If $\phi \in \mathrm{IS}(\mathfrak{F}G)$, then Proposition (2.1) says that induction is an isometric bijection of $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\phi)$ onto $\mathrm{Irr}(\mathfrak{F}G|\phi)$. We denote the inverse to this bijection by $\cdot\{\phi\}: \psi \mapsto \psi\{\phi\}$. Thus:

(2.10) *If* $\phi \in \mathrm{IS}(\mathfrak{F}G)$, *then* $\cdot\{\phi\}$ *is the isometric bijection of* $\mathrm{Irr}(\mathfrak{F}G|\phi)$ *onto* $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\phi)$ *sending any* $\psi \in \mathrm{Irr}(\mathfrak{F}G|\phi)$ *to the unique character* $\psi\{\phi\} \in \mathrm{Irr}(\mathfrak{F}G\{\phi\}|\phi)$ *such that* $\psi\{\phi\}^G = \psi$.

To extend $\cdot\{\phi\}$ to a map of $\mathrm{Irs}(\mathfrak{F}G|\phi)$ onto $\mathrm{Irs}(\mathfrak{F}G\{\phi\}|\phi)$ we use:

Proposition 2.11. *If* $\phi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $\mathrm{Dom}(\phi) \leqq H \leqq G$, *then* $\phi \in \mathrm{IS}(\mathfrak{F}H)$.

Proof. Because (2.3) holds for all $\sigma \in G - G\{\phi\}$ it holds for all $\sigma \in H - H\{\phi\} \subseteq G - G\{\phi\}$. So Corollary (2.6) implies this proposition.

Proposition (2.11) implies that the map $\cdot\{\phi\}$ of (2.10) also sends $\mathrm{Irr}(\mathfrak{F}H|\phi)$ onto $\mathrm{Irr}(\mathfrak{F}H\{\phi\}|\phi)$, for each H such that $\mathrm{Dom}(\phi) \leqq H \leqq G$. It follows from this and (1.9b) that $\cdot\{\phi\}$ is a map of $\mathrm{Irs}(\mathfrak{F}G|\phi)$ onto $Irs(\mathfrak{F}G\{\phi\}|\phi)$. While this extended $\cdot\{\phi\}$ is not a bijection (unless $G\{\phi\} = G$), it is an isometry in the following sense:

(2.12) **Proposition.** *Suppose that* $\phi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $\psi, \chi \in \mathrm{Irs}(\mathfrak{F}G|\phi)$ *with* $\mathrm{Dom}(\psi) \leqq \mathrm{Dom}(\chi)$. *Then:*

$$(\psi^{\mathrm{Dom}(\chi)}, \chi) = (\psi\{\phi\}^{\mathrm{Dom}(\chi\{\phi\})}, \chi\{\phi\}), \tag{2.13a}$$

$$(\psi, \chi_{\mathrm{Dom}(\psi)}) = (\psi\{\phi\}, \chi\{\phi\}_{\mathrm{Dom}(\psi\{\phi\})}). \tag{2.13b}$$

Proof. Let D be $\mathrm{Dom}(\chi)$, so that $D\{\phi\}$ is $\mathrm{Dom}(\chi\{\phi\})$ by (2.10). In view of (1.9a), the distinct elements $\theta_1, \ldots, \theta_n$ of $\mathrm{Irr}(\mathfrak{F}D\{\phi\}|\psi\{\phi\})$ satisfy:

$$\psi\{\phi\}^{D\{\phi\}} = m_1\theta_1 + \ldots + m_n\theta_n,$$

for some integers $m_i > 0$. By (2.10) this implies that:

$$\psi^D = (\psi\{\phi\}^{\mathrm{Dom}(\psi)})^D = m_1\theta_1^D + \ldots + m_n\theta_n^D.$$

Because $\chi\{\phi\}$ and $\theta_1, \ldots, \theta_n$ are all members of $\mathrm{Irr}(\mathfrak{F}D\{\phi\}|\phi)$, which is mapped isometrically onto $\mathrm{Irr}(\mathfrak{F}D|\phi)$ by induction, we have:

$$(\psi^D, \chi) = \sum_{i=1}^{n} m_i(\theta_i^D, \chi\{\phi\}^D) = \sum_{i=1}^{n} m_i(\theta_i, \chi\{\phi\}) = (\psi\{\phi\}^{D\{\phi\}}, \chi\{\phi\}).$$

So (2.13a) holds. This is the same as (2.13b) by the Frobenius Reciprocity Law.

(2.14) **Corollary.** *If* $\phi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $\psi \in \mathrm{Irs}(\mathfrak{F}G|\phi)$, *then* $\cdot\{\phi\}$ *is an isometric bijection of* $\mathrm{Irr}(\mathfrak{F}G|\psi)$ *onto* $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$. *Hence induction is an isometry of* $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$ *onto* $\mathrm{Irr}(\mathfrak{F}G|\psi)$.

Proof. The set $\mathrm{Irr}(\mathfrak{F}G|\psi)$ consists of all $\chi \in \mathrm{Irr}(\mathfrak{F}G)$ such that $(\psi^G, \chi) > 0$. Any such χ satisfies $\chi \geqq \psi \geqq \phi$, and hence lies in $\mathrm{Irr}(\mathfrak{F}G|\phi)$. In view of (2.13a) and (2.10), the map $\cdot\{\phi\}$ is an isometric bijection of $\mathrm{Irr}(\mathfrak{F}G|\psi)$ onto the set of all $\chi' \in \mathrm{Irr}(\mathfrak{F}G\{\phi\}|\phi)$ such that $(\psi\{\phi\}^{G\{\phi\}}, \chi') > 0$, i.e., onto $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$. This proves the first statement of the corollary. The rest follows by (2.10).

While we shall only need the definition (2.10) of $\cdot(\phi)$ via characters, it is worth remarking that this map can also be defined in terms of modules. Since $\mathfrak{F}\mathrm{Dom}(\phi)$ is a semi-simple algebra, the restriction to it of any $\mathfrak{F}G$-module $\mathfrak{V}$ has a unique direct sum decomposition:

$$\mathfrak{V}_{\mathrm{Dom}(\sigma)} = \dot{\sum_{\chi \in \mathrm{Irr}(\mathfrak{F}\mathrm{Dom}(\phi))}} \mathfrak{V}\{\chi\},$$

where the *χ-primary component* $\mathfrak{V}\{\chi\}$ is the sum of all the simple $\mathfrak{F}\,\mathrm{Dom}(\phi)$-submodules $\mathfrak{W}$ of $\mathfrak{V}$ affording $\chi \in \mathrm{Irr}(\mathfrak{F}\,\mathrm{Dom}(\phi))$. Since $\mathrm{Dom}(\phi) \trianglelefteq G\{\phi\}$, we have:

$$\mathfrak{V}\{\chi\}\sigma = \mathfrak{V}\{\chi^\sigma\}, \quad \textit{for all } \chi \in \mathrm{Irr}(\mathfrak{F}\,\mathrm{Dom}(\phi)) \quad \textit{and all } \sigma \in G\{\phi\}.$$

In particular, the subspace $\mathfrak{V}\{\phi\}$ is an $\mathfrak{F}G\{\phi\}$-submodule of $\mathfrak{V}$.

(2.15) **Proposition.** *If $\phi\in\mathrm{IS}(\mathfrak{F}G)$ and $\mathfrak{V}$ is a simple $\mathfrak{F}G$-module whose character ψ lies in $\mathrm{Irr}(\mathfrak{F}G|\phi)$, then $\mathfrak{V}\{\phi\}$ is a simple $\mathfrak{F}G\{\phi\}$-module affording $\psi\{\phi\}$.*

Proof. Since $\psi=\psi\{\phi\}^G$, there is some $\mathfrak{F}G\{\phi\}$-submodule $\mathfrak{U}$ of $\mathfrak{V}$ affording $\psi\{\phi\}$ such that $\mathfrak{V}$ is the direct sum:

$$\mathfrak{V}=\dot{\sum_{\sigma}}\,\mathfrak{U}\sigma \qquad (\textit{as } \mathfrak{F}\textit{-vector spaces}), \tag{2.16}$$

where σ runs over representatives for the left cosets $G\{\phi\}\sigma$ of $G\{\phi\}$ in G. To prove the proposition it suffices to show that $\mathfrak{U}=\mathfrak{V}\{\phi\}$.

Because $\psi\{\phi\}\in\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\phi)$ and $\mathrm{Dom}(\phi)\trianglelefteq G\{\phi\}$. the restriction $\psi\{\phi\}_{\mathrm{Dom}(\phi)}$ is a multiple $m\phi$ of ϕ by Clifford's Theory (Hauptsatz V.17.3 of [5]). It follows that $\mathfrak{U}\subseteq\mathfrak{V}\{\phi\}$.

Suppose that $\mathfrak{U}\subset\mathfrak{V}\{\phi\}$. Then there must be some $\sigma\in G-G\{\phi\}$ such that the projection π_σ of $\mathfrak{V}$ onto $\mathfrak{U}\sigma$ in (2.16) sends $\mathfrak{V}\{\phi\}$ onto a non-zero subspace $\pi_\sigma(\mathfrak{V}\{\phi\})$ of $\mathfrak{U}\sigma$. Evidently π_σ is an $\mathfrak{F}(D^\sigma\cap D)$-homomorphism, where $D=\mathrm{Dom}(\phi)$. So $\pi_\sigma(\mathfrak{V}\{\phi\})$ is a non-zero $\mathfrak{F}(D^\sigma\cap D)$-factor module of $\mathfrak{V}\{\phi\}$. It follows that $\pi_\sigma(\mathfrak{V}\{\phi\})$ has a simple $\mathfrak{F}(D^\sigma\cap D)$-submodule $\mathfrak{W}$ whose character χ is an irreducible $\mathfrak{F}$-constituent of $\phi_{D^\sigma\cap D}$. But $\mathfrak{W}$ is also an $\mathfrak{F}(D^\sigma\cap D)$-submodule of the $\mathfrak{F}G\{\phi\}^\sigma$-module $\mathfrak{U}\sigma$, which affords $\psi\{\phi\}^\sigma$. So χ is an irreducible $\mathfrak{F}$-constituent of $(\psi\{\phi\}^\sigma)_{D^\sigma}=m\phi^\sigma$. Therefore χ is an irreducible $\mathfrak{F}$-constituent of both $(\phi^\sigma)_{D^\sigma\cap D}$ and $\phi_{D^\sigma\cap D}$. This contradicts (2.3), which holds for this $\sigma\in G-G\{\phi\}$ by Corollary (2.6). Hence $\mathfrak{U}=\mathfrak{V}\{\phi\}$ and the proposition is proved.

§ 3. Composite Maps

The map $\cdot\{\phi\}$ of (2.10) treats inductive sources $\psi\in\mathrm{IS}(\mathfrak{F}G|\phi)$ properly. It even preserves the maps $\cdot\{\psi\}$ attached to them.

(3.1) **Theorem.** *Suppose that $\phi\in\mathrm{IS}(\mathfrak{F}G)$ and $\psi\in\mathrm{Irs}(\mathfrak{F}G|\phi)$. Then $\psi\in\mathrm{IS}(\mathfrak{F}G)$ if and only if:*

(3.2a) $$\psi\{\phi\}\in\mathrm{IS}(\mathfrak{F}G\{\phi\}) \quad \textit{and:}$$

(3.2b) $$G\{\phi\}\{\psi\{\phi\}\} \quad \textit{normalizes } \mathrm{Dom}(\psi).$$

In that case we have:

(3.3a) $$\psi\{\phi\}\in\mathrm{IS}(\mathfrak{F}G),$$

(3.3b) $$G\{\psi\}\{\phi\}=G\{\phi\}\{\psi\{\phi\}\}=G\{\psi\{\phi\}\},$$

(3.3c) $$\chi\{\psi\}\{\phi\}=\chi\{\phi\}\{\psi\{\phi\}\}=\chi\{\psi\{\phi\}\}, \quad \textit{for all } \chi\in\mathrm{Irs}(\mathfrak{F}G|\psi),$$

(3.3d) $$G\{\psi\}=G\{\psi\{\phi\}\}\,\mathrm{Dom}(\psi),$$

(3.3e) $$G\{\psi\{\phi\}\}\cap\mathrm{Dom}(\psi)=\mathrm{Dom}(\psi\{\phi\}).$$

Before starting to prove this, let's make sure we know what the notation means. The group $G\{\phi\}\{\psi\{\phi\}\}$ is the stabilizer in $G\{\phi\}$ of $\psi\{\phi\}\in\mathrm{Irs}(\mathfrak{F}G\{\phi\}|\phi)$, while $G\{\psi\}\{\phi\}$ is the stabilizer in $G\{\psi\}$ of $\phi\in\mathrm{Irs}(\mathfrak{F}G\{\psi\})$ (see (1.10)). The character $\chi\{\psi\}\{\phi\}$ in (3.3c) is the image under $\cdot\{\phi\}$ of $\chi\{\psi\}\in\mathrm{Irs}(\mathfrak{F}G\{\psi\}|\psi)\subseteq\mathrm{Irs}(\mathfrak{F}G\{\psi\}|\phi)$. The character $\chi\{\phi\}\{\psi\{\phi\}\}$ there is the image under $\cdot\{\psi\{\phi\}\}$ of $\chi\{\phi\}$, which lies in $\mathrm{Irs}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$ by Corollary (2.14).

Proof of Theorem (3.1). If $\psi\in\mathrm{IS}(\mathfrak{F}G)$, then induction to G is an isometry on $\mathrm{Irr}(\mathfrak{F}G\{\psi\}|\psi)$. We know that $\phi\in\mathrm{IS}(\mathfrak{F}G\{\psi\})$ by Proposition (2.11) and hence that induction is an isometry of $\mathrm{Irr}(\mathfrak{F}G\{\psi\}\{\phi\}|\psi\{\phi\})$ onto $\mathrm{Irr}(\mathfrak{F}G\{\psi\}|\psi)$ by Corollary (2.14). Thus induction to G is an isometry on $\mathrm{Irr}(\mathfrak{F}G\{\psi\}\{\phi\}|\psi\{\phi\})$. Since $G\{\psi\}\{\phi\}$ stabilizes the unique function $\psi\{\phi\}$ of ψ and ϕ, this and Theorem (2.8) tell us that (3.3a) holds and that the first and last groups in (3.3b) are equal. Evidently (3.3a) implies (3.2a) by Proposition (2.11). The inclusions:

$$G\{\phi\}\{\psi\{\phi\}\}\leqq G\{\psi\{\phi\}\}=G\{\psi\}\{\phi\}\leqq G\{\psi\}$$

imply (3.2b) since $\mathrm{Dom}(\psi)\trianglelefteq G(\psi)$.

Now suppose that (3.2a, b) hold. Then induction is an isometry of $\mathrm{Irr}(\mathfrak{F}G\{\phi\}\{\psi\{\phi\}\}|\psi\{\phi\})$ onto $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$. It is also an isometry of $\mathrm{Irr}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$ onto $\mathrm{Irr}(\mathfrak{F}G|\psi)$ by Corollary (2.14). Hence induction to G is an isometry on $\mathrm{Irr}(\mathfrak{F}G\{\phi\}\{\psi\{\phi\}\}|\psi\{\phi\})$. Since $G\{\phi\}\{\psi\{\phi\}\}\leqq G\{\psi\{\phi\}\}$, this and Theorem (2.8) imply that (3.3a) and the second equality in (3.3b) hold.

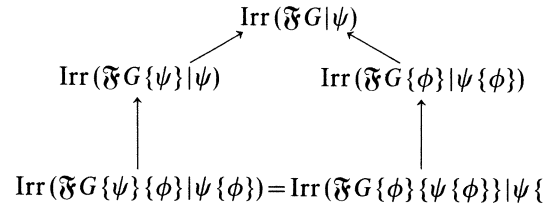
Since $G\{\psi\{\phi\}\}=G\{\phi\}\{\psi\{\phi\}\}$ normalizes $\mathrm{Dom}(\psi)$ and stabilizes $\psi\{\phi\}$, it stabilizes $\psi=\psi\{\phi\}^{\mathrm{Dom}(\psi)}$. Therefore $H=G\{\psi\{\phi\}\}\,\mathrm{Dom}(\psi)$ is a subgroup of $G\{\psi\}$. We know from (3.3a) and Proposition (2.11) that $\psi\{\phi\}\in\mathrm{IS}(\mathfrak{F}H)$. Hence induction is an isometry of $\mathrm{Irr}(\mathfrak{F}G\{\psi\{\phi\}\}|\psi\{\phi\})$ onto both $\mathrm{Irr}(\mathfrak{F}H|\psi\{\phi\})$ and $\mathrm{Irr}(\mathfrak{F}G|\psi\{\phi\})$. These last two sets are precisely $\mathrm{Irr}(\mathfrak{F}H|\psi)$ and $\mathrm{Irr}(\mathfrak{F}G|\psi)$, since $\mathrm{Dom}(\psi)\leqq H\leqq G$ and $\psi=\psi\{\phi\}^{\mathrm{Dom}(\psi)}$. Therefore induction is an isometry of $\mathrm{Irr}(\mathfrak{F}H|\psi)$ onto $\mathrm{Irr}(\mathfrak{F}G|\psi)$. In view of Theorem (2.8), this implies that $\psi\in\mathrm{IS}(\mathfrak{F}G)$ and that (3.3d) holds.

We have now proven that $\psi\in\mathrm{IS}(\mathfrak{F}G)$ if and only if (3.2a, b) hold, and have shown incidently that (3.3a, b, d) are true when these conditions are satisfied. Since $\mathrm{Dom}(\psi\{\phi\})=G\{\phi\}\cap\mathrm{Dom}(\psi)$, Eq. (3.3e) follows easily from (3.3b). Any $\chi\in\mathrm{Irs}(\mathfrak{F}G|\psi)$ is induced from $\chi\{\psi\}\in\mathrm{Irs}(\mathfrak{F}G\{\psi\}|\psi)$, which, in turn, is induced from $\chi\{\psi\}\{\phi\}\in\mathrm{Irs}(\mathfrak{F}G\{\psi\}\{\phi\}|\psi\{\phi\})$ by Corollary (2.14). This, (3.3b) and the definition (2.10) of $\chi\{\psi\{\phi\}\}$ imply that $\chi\{\psi\}\{\phi\}=\chi\{\psi\{\phi\}\}$. Similarly χ is induced from $\chi\{\phi\}\in\mathrm{Irs}(\mathfrak{F}G\{\phi\}|\psi\{\phi\})$, which, in turn, is induced from

$$\chi\{\phi\}\{\psi\{\phi\}\}\in\mathrm{Irs}(\mathfrak{F}G\{\phi\}\{\psi\{\phi\}\}|\psi\{\phi\}).$$

This implies that $\chi\{\phi\}\{\psi\{\phi\}\}=\chi\{\psi\{\phi\}\}$ by (3.3b). So the proofs of (3.3c) and the theorem are complete.

(3.4) **Corollary.** *If* $\phi\in\mathrm{IS}(\mathfrak{F}G)$ *and* $\psi\in\mathrm{IS}(\mathfrak{F}G|\phi)$, *then the following diagram, in which each arrow represents induction of characters, consists entirely of isometric bijections and commutes:*

(3.5)


Proof. It follows from (2.10) and Corollary (2.14) that each arrow in (3.5) is an isometric bijection. The equality in the bottom row comes from (3.3b), and the commutivity comes from (3.3c) in view of (2.10).

(3.6) **Corollary.** *Suppose that M, N are two normal subgroups of G with $M \leqq N$, that $\phi \in \mathrm{Irr}(\mathfrak{F}M)$ and that $\psi \in \mathrm{Irr}(\mathfrak{F}N|\phi)$. Then $\phi \in \mathrm{IS}(\mathfrak{F}G)$ and $\psi \in \mathrm{IS}(\mathfrak{F}G|\phi)$. Hence* (3.2), (3.3) *and the conclusions of Corollary* (3.4) *all hold for ϕ and ψ. Here $M = \mathrm{Dom}(\phi) \trianglelefteq G\{\psi\}$ and $N\{\phi\} = \mathrm{Dom}(\psi\{\phi\}) \trianglelefteq G\{\phi\}$. So the arrows in* (3.5) *represent the usual inductive bijections of Clifford's Theory (Hauptsatz* V.17.3e) *of* [5]).

Proof. This is an immediate consequence of the theorem and Corollaries (3.4) and (2.7).

It should be remarked that the commutativity of (3.5) in the situation of Corollary (3.6) is fairly well known. For example, when $\mathfrak{F}$ is algebraically closed the diagram (3.5) appears as part of the commutative diagram in Corollary 17.10 in [1].

(3.7) **Corollary.** *Fix $\phi \in \mathrm{IS}(\mathfrak{F}G)$ and $\psi \in \mathrm{Irs}(\mathfrak{F}G\{\phi\}|\phi)$. Then $\psi \in \mathrm{IS}(\mathfrak{F}G)$ if and only if $\psi \in \mathrm{IS}(\mathfrak{F}G\{\phi\})$. In that case $G\{\psi\}$ is $G\{\phi\}\{\psi\}$.*

Proof. Since $\mathrm{Dom}(\psi) \leqq G\{\phi\}$, it follows from (2.10) that $\psi\{\phi\} = \psi$. Hence (3.2b) holds and the theorem implies this corollary.

§ 4. IS-Correspondents

If $\psi \in \mathrm{Irs}(\mathfrak{F}G)$, then $\mathrm{IS}(\psi|\mathfrak{F}G)$ is precisely the set of all $\phi \in \mathrm{IS}(\mathfrak{F}G)$ for which $\psi\{\phi\}$ is defined (see (1.8) and (2.10)). We call the characters $\psi\{\phi\}$ obtained in this way the $\mathrm{IS}(\mathfrak{F}G)$-*correspondents* of ψ. In the spirit of (1.8b) we denote the set of all these correspondents by:

$$\mathrm{ISC}(\psi|\mathfrak{F}G) = \{\psi\{\phi\} \mid \phi \in \mathrm{IS}(\psi|\mathfrak{F}G)\}. \tag{4.1}$$

We shall only be interested in $\mathrm{ISC}(\psi|\mathfrak{F}G)$ when ψ is itself an inductive source of $\mathfrak{F}G$. In that case, the study of this set is closely related to the partial ordering $\ll$ on $\mathrm{Irs}(\mathfrak{F}G)$ defined by:

(4.2) *If $\phi, \chi \in \mathrm{Irs}(\mathfrak{F}G)$, then $\phi \ll \chi$ if and only if $\mathrm{Dom}(\phi) \leqq \mathrm{Dom}(\chi)$ and $\chi = \phi^{\mathrm{Dom}(\chi)}$ with $(\phi.\,\phi) = (\chi,\chi)$.*

Evidently $\ll$ is stronger than the partial ordering $\leqq$ of (1.7).

(4.3) **Proposition.** *If* $\psi \in \mathrm{IS}(\mathfrak{F}G)$, *then:*

$$\mathrm{ISC}(\psi|\mathfrak{F}G) = \{\phi \in \mathrm{IS}(\mathfrak{F}G) | \phi \ll \psi\} = \{\phi \in \mathrm{IS}(\psi|\mathfrak{F}G) | \psi\{\phi\} = \phi\}.$$

Proof. Any element of $\mathrm{ISC}(\psi|\mathfrak{F}G)$ has the form $\psi\{\phi\}$ for some $\phi \in \mathrm{IS}(\psi|\mathfrak{F}G)$. Then $\psi\{\phi\} \in \mathrm{IS}(\mathfrak{F}G)$ by (3.3a) and $\psi\{\phi\} \ll \psi$ by (2.10).

Now suppose that $\phi \in \mathrm{IS}(\mathfrak{F}G)$ and $\phi \ll \psi$. Then $\phi \leqq \psi$, so that $\phi \in \mathrm{IS}(\psi|\mathfrak{F}G)$. Furthermore $\mathrm{Dom}(\psi) \cap G\{\phi\}$ is $\mathrm{Dom}(\phi)$ by Corollary (2.9). Hence $\psi\{\phi\} = \phi$ by (2.10).

Finally, assume that $\phi \in \mathrm{IS}(\psi|\mathfrak{F}G)$ with $\psi\{\phi\} = \phi$. Then $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$ by (4.1). So the proposition is proved.

(4.4) **Corollary.** *If* $\psi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$, *then* $G\{\phi\} = G\{\psi\}\{\phi\} \leqq G\{\psi\}$. *Indeed, we have* $G\{\psi\} = G\{\phi\}\mathrm{Dom}(\psi)$ *with* $G\{\phi\} \cap \mathrm{Dom}(\psi) = \mathrm{Dom}(\phi)$.

Proof. By the proposition ϕ belongs to $\mathrm{IS}(\psi|\mathfrak{F}G)$ and $\psi\{\phi\} = \phi$. So this corollary follows from (3.3b, d, e).

(4.5) **Corollary.** *The two partial orderings* $\ll$ *and* $\leqq$ *coincide on* $\mathrm{ISC}(\psi|\mathfrak{F}G)$ *when* $\psi \in \mathrm{IS}(\mathfrak{F}G)$.

Proof. If $\phi, \chi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$, then $\phi \ll \psi$ and $\chi \ll \psi$ by the proposition. It follows immediately from this and (1.3) that $\phi \leqq \chi$ if and only if $\phi \ll \chi$.

The set $\mathrm{ISC}(\psi|\mathfrak{F}G)$ only depends on what happens inside the subgroup $G\{\psi\}$ when ψ is an inductive source of $\mathfrak{F}G$.

(4.6) **Lemma.** *If* $\psi \in \mathrm{IS}(\mathfrak{F}G)$, *then* $\mathrm{ISC}(\psi|\mathfrak{F}G) = \mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$.

Proof. If $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$, then $\phi \in \mathrm{IS}(\mathfrak{F}G)$ with $\phi \ll \psi$ by Proposition (4.3). Proposition (2.11) now says that $\phi \in \mathrm{IS}(\mathfrak{F}G\{\psi\})$ with $\phi \ll \psi$. So $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$ by Proposition (4.3).

Now suppose that $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$. Then $\phi \in \mathrm{IS}(\psi|\mathfrak{F}G\{\psi\})$ with $\psi\{\phi\} = \phi$ by Proposition (4.3). So Corollary (2.14) tells us that induction is an isometry of $\mathrm{Irr}(\mathfrak{F}G\{\psi\}\{\phi\}|\phi)$ onto $\mathrm{Irr}(\mathfrak{F}G\{\psi\}|\psi)$. But induction to G is an isometry on $\mathrm{Irr}(\mathfrak{F}G\{\psi\}|\psi)$ since $\psi \in \mathrm{IS}(\mathfrak{F}G)$. Hence it is an isometry on $\mathrm{Irr}(\mathfrak{F}G\{\psi\}\{\phi\}|\phi)$. Because $G\{\psi\}\{\phi\} \leqq G\{\phi\}$, this and Theorem (2.8) imply that $\phi \in \mathrm{IS}(\psi|\mathfrak{F}G)$ with $\psi\{\phi\} = \phi$. Therefore $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$ by Proposition (4.3) and the lemma is proved.

Evidently Lemma (4.6) and Corollary (4.4) reduce the study of $\mathrm{ISC}(\psi|\mathfrak{F}G)$, for $\psi \in \mathrm{IS}(\mathfrak{F}G)$, to the special case in which $G = G\{\psi\}$. In view of (1.11) and Corollary (2.7), this is precisely the situation where:

(4.7) $$G \text{ fixes } \psi \in \mathrm{Irr}(\mathfrak{F}N), \quad \text{for some } N \trianglelefteq G.$$

In Corollary 4.3 of [6] Isaacs proved the difficult half of the following characterization of $\mathrm{ISC}(\psi|\mathfrak{F}G)$ for the case of algebraically closed $\mathfrak{F}$.

(4.8) **Theorem.** *If* (4.7) *holds, then:*

$$\mathrm{ISC}(\psi|\mathfrak{F}G) = \{\phi \in \mathrm{Irs}(\mathfrak{F}N) | \phi \ll \psi \text{ and } G = G\{\phi\}N\}.$$

Proof. If $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$, then $\phi \ll \psi$ by Proposition (4.3). Hence $\phi \in \mathrm{Irs}(\mathfrak{F}\,\mathrm{Dom}(\psi)) = \mathrm{Irs}(\mathfrak{F}N)$. Furthermore, Corollary (4.4) gives:

$$G = G\{\psi\} = G\{\phi\}\,\mathrm{Dom}(\psi) = G\{\phi\}N.$$

Now suppose that $\phi \in \mathrm{Irs}(\mathfrak{F}N)$ satisfies $\phi \ll \psi$ and $G = G\{\phi\}N$. Then $\phi \in \mathrm{IS}(\mathfrak{F}N)$ by Corollary (2.9). So Corollary (2.6) tells us that (2.3) holds for all $\sigma \in N - N\{\phi\}$. Because $G = G\{\phi\}N$, this implies (2.3) for all $\sigma \in G - G\{\phi\}$. Therefore $\phi \in \mathrm{IS}(\mathfrak{F}G)$ by Corollary (2.6). Since $\psi \in \mathrm{IS}(\mathfrak{F}G)$ and $\phi \ll \psi$, Proposition (4.3) now says that $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$. So the theorem is proved.

(4.9) **Corollary.** *If ψ is an arbitrary element of* $\mathrm{IS}(\mathfrak{F}G)$, *then:*

$$\mathrm{ISC}(\psi|\mathfrak{F}G) = \{\phi \in \mathrm{Irs}(\mathfrak{F}\,\mathrm{Dom}(\psi)) \mid \phi \ll \psi \text{ and } G\{\psi\} = G\{\psi\}\{\phi\}\,\mathrm{Dom}(\psi)\}.$$

Proof. The theorem says that the right side of the above equality is $\mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$. So it is equal to the left side by Lemma (4.6).

(4.10) **Corollary.** *Suppose that* (4.7) *holds, that* $\phi \in \mathrm{ISC}(\psi|\mathfrak{F}G)$ *and that* $\mathrm{Dom}(\phi) \leqq H \leqq N$. *Then* $\phi^H \in \mathrm{ISC}(\psi|\mathfrak{F}G)$ *if and only if* $G\{\phi\}$ *normalizes* H. *In that case we have:*

$$G\{\phi^H\} = G\{\phi\}H, \tag{4.11a}$$

$$G\{\phi\} \cap H = \mathrm{Dom}(\phi), \tag{4.11b}$$

$$\phi \in \mathrm{ISC}(\phi^H|\mathfrak{F}G) = \mathrm{ISC}(\phi^H|\mathfrak{F}G\{\phi^H\}), \tag{4.11c}$$

$$\phi \ll \phi^H \ll \psi. \tag{4.11d}$$

Proof. Since $\phi \ll \psi$, it follows easily from (1.3) and (1.5) that $\phi^H \in \mathrm{Irs}(\mathfrak{F}N)$ satisfies (4.11d).

If $\phi^H \in \mathrm{ISC}(\psi|\mathfrak{F}G)$, then both ϕ and ϕ^H lie in $\mathrm{IS}(\mathfrak{F}G)$ by Proposition (4.3). That proposition and (4.11d) imply that $\phi \in \mathrm{ISC}(\phi^H|\mathfrak{F}G)$. Now Corollary (4.4) gives (4.11a, b) and Lemma (4.6) gives the equality in (4.11c). Evidently (4.11a) implies that $G\{\phi\}$ normalizes $H = \mathrm{Dom}(\phi^H)$.

Conversely, if $G\{\phi\}$ normalizes H it stabilizes ϕ^H. So $G\{\phi\}H$ is a subgroup of $G\{\phi^H\}$. Since $G\{\phi\}N = G$ by Theorem (4.8), we must have $G\{\phi^H\}N = G$. This and (4.11d) imply that $\phi^H \in \mathrm{ISC}(\psi|\mathfrak{F}G)$ by the theorem. So the corollary is proved.

§ 5. Clifford Correspondents

Suppose that $\psi \in \mathrm{Irs}(\mathfrak{F}G)$. If $N \leqq \mathrm{Dom}(\psi)$ is a normal subgroup of $G\{\psi\}$, then any $\phi \in \mathrm{Irr}(\psi|\mathfrak{F}N)$ lies in $\mathrm{IS}(\psi|\mathfrak{F}G\{\psi\})$ by Corollary (2.7). Furthermore, (2.10) tells us that $\psi\{\phi\} \in \mathrm{Irr}(\mathfrak{F}(\mathrm{Dom}(\psi) \cap G\{\phi\})|\phi)$ is the usual character corresponding to $\psi \in \mathrm{Irr}(\mathfrak{F}\,\mathrm{Dom}(\psi)|\phi)$ in Clifford's theory (Hauptsatz V.17.3 in [5]). We call the characters $\psi\{\phi\}$ obtained in this way the *direct $\mathfrak{F}G$-Clifford correspondents* of ψ. In view of (4.1), the set $\mathrm{DCC}(\psi|\mathfrak{F}G)$ of all direct $\mathfrak{F}G$-Clifford correspondents of ψ satisfies:

$$\begin{aligned}\mathrm{DCC}(\psi|\mathfrak{F}G) &= \{\psi\{\phi\} \mid \phi \in \mathrm{Irs}(\psi|\mathfrak{F}G) \text{ and } \mathrm{Dom}(\phi) \trianglelefteq G\{\psi\}\} \\ &\subseteq \mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\}).\end{aligned} \tag{5.1}$$

Although the inclusion in (5.1) is normally strict, certain elements of $\mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$ automatically lie in $\mathrm{DCC}(\psi|\mathfrak{F}G)$.

(5.2) **Proposition.** *If* $\psi\in\mathrm{Irs}(\mathfrak{F}G)$ *and* $\phi\in\mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$ *with* $\mathrm{Dom}(\phi)\trianglelefteq G\{\psi\}$, *then* $\phi\in\mathrm{DCC}(\psi|\mathfrak{F}G)$.

Proof. It follows from (1.11) and Corollary (2.7) that $\psi\in\mathrm{IS}(\mathfrak{F}G\{\psi\})$. So Proposition (4.3) tells us that $\phi=\psi\{\phi\}\in\mathrm{DCC}(\psi|\mathfrak{F}G)$.

As usual, things are more interesting when $\psi\in\mathrm{IS}(\mathfrak{F}G)$. In that case $\mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$ is $\mathrm{ISC}(\psi|\mathfrak{F}G)$ by Lemma (4.6). Hence (5.1) and Proposition (4.3) imply:

(5.3) *If* $\psi\in\mathrm{IS}(\mathfrak{F}G)$, *then* $\mathrm{DCC}(\psi|\mathfrak{F}G)\subseteq\mathrm{ISC}(\psi|\mathfrak{F}G)\subseteq\mathrm{IS}(\psi|\mathfrak{F}G)$.

If $\phi,\psi\in\mathrm{Irs}(\mathfrak{F}G)$, we write '$\phi\ dcc-\mathfrak{F}G\ \psi$' to indicate that ϕ is a direct $\mathfrak{F}G$-Clifford correspondent of ψ, i.e., that $\phi\in\mathrm{DCC}(\psi|\mathfrak{F}G)$. The relation $dcc-\mathfrak{F}G$ is reflexive but not usually transitive. If we extend it to a transitive relation we obtain the partial ordering $cc-\mathfrak{F}G$ on $\mathrm{Irs}(\mathfrak{F}G)$ defined by:

(5.4) *If* $\phi,\psi\in\mathrm{Irs}(\mathfrak{F}G)$, *then* $\phi\ cc-\mathfrak{F}G\ \psi$ *if and only if there exist* $\chi_0,\chi_1,\ldots,\chi_n\in\mathrm{Irs}(\mathfrak{F}G)$, *for some* $n\in\mathbb{N}$, *such that:*

$$\phi=\chi_0\ dcc-\mathfrak{F}G\ \chi_1\ dcc-\mathfrak{F}G\ldots dcc-\mathfrak{F}G\ \chi_n=\psi. \tag{5.5}$$

In that case we say that "ϕ *is an* $\mathfrak{F}G$*-Clifford correspondent of* ψ*".*

Evidently (2.10) and (5.1) imply that the partial ordering $cc-\mathfrak{F}G$ on $\mathrm{Irs}(\mathfrak{F}G)$ is stronger than the partial ordering $\ll$ of (4.2), and hence stronger than the partial ordering $\leqq$ of (1.7).

We denote by $\mathrm{CC}(\psi|\mathfrak{F}G)$ the set $\{\phi\in\mathrm{Irs}(\mathfrak{F}G)|\phi\ cc-\mathfrak{F}G\ \psi\}$ of all $\mathfrak{F}G$-Clifford correspondents of any $\psi\in\mathrm{Irs}(\mathfrak{F}G)$.

(5.6) **Proposition.** *If* $\psi\in\mathrm{IS}(\mathfrak{F}G)$, *then* $\mathrm{CC}(\psi|\mathfrak{F}G)\subseteq\mathrm{ISC}(\psi|\mathfrak{F}G)\subseteq\mathrm{IS}(\psi|\mathfrak{F}G)$.

Proof. If $\phi\in\mathrm{CC}(\psi|\mathfrak{F}G)$, then there exist $n\in\mathbb{N}$ and $\chi_0,\chi_1,\ldots,\chi_n\in\mathrm{Irs}(\mathfrak{F}G)$ satisfying (5.5). If $n=0$, then $\phi=\psi\in\mathrm{ISC}(\psi|\mathfrak{F}G)$. If $n>0$, then $\chi_{n-1}\in\mathrm{DCC}(\psi|\mathfrak{F}G)$ implies that $\chi_{n-1}\in\mathrm{ISC}(\psi|\mathfrak{F}G)$ by (5.3). Hence Proposition (4.3) gives $\chi_{n-1}\in\mathrm{IS}(\mathfrak{F}G)$ with $\chi_{n-1}\ll\psi$. Using induction on n, we may suppose that $\phi\in\mathrm{ISC}(\chi_{n-1}|\mathfrak{F}G)$. Then $\phi\in\mathrm{IS}(\mathfrak{F}G)$ with $\phi\ll\chi_{n-1}\ll\psi$, so that $\phi\in\mathrm{ISC}(\psi|\mathfrak{F}G)$ by Proposition (4.3). Since we have $\mathrm{ISC}(\psi|\mathfrak{F}G)\subseteq\mathrm{IS}(\psi|\mathfrak{F}G)$ by Proposition (4.3), this proves the proposition.

(5.7) **Corollary.** *If* $\psi\in\mathrm{IS}(\mathfrak{F}G)$, *then* $\mathrm{CC}(\psi|\mathfrak{F}G)=\mathrm{CC}(\psi|\mathfrak{F}G\{\psi\})$.

Proof. If (5.5) holds with $G\{\psi\}$ in place of G, then the proposition says that each χ_i lies in $\mathrm{ISC}(\psi|\mathfrak{F}G\{\psi\})$, which is $\mathrm{ISC}(\psi|\mathfrak{F}G)$ by Lemma (4.6). So $G\{\chi_i\}=G\{\psi\}\{\chi_i\}$ by Corollary (4.4). Hence (5.5) holds for G. This implies the corollary.

The first inclusion in Proposition (5.6) is normally strict (see Example (5.12) below), as is the second. However, there is one important case in which the former is equality. We shall obtain this case as a corollary to:

(5.8) **Theorem.** *If* (4.7) *holds and if* $\phi \in \mathrm{ISC}(\psi | \mathfrak{F}G)$ *with* $\mathrm{Dom}(\phi)$ *subnormal in* N, *then* $\phi \in \mathrm{CC}(\psi | \mathfrak{F}G)$.

Proof. Because $\mathrm{Dom}(\phi)$ is subnormal in N there exist $n \in \mathbb{N}$ and subgroups $N(0), N(1), \ldots, N(n)$ satisfying:

$$\mathrm{Dom}(\phi) = N(0) \trianglelefteq N(1) \trianglelefteq \ldots \trianglelefteq N(n) = N. \tag{5.9}$$

We claim that the $N(i)$ in (5.9) can be chosen to be $G\{\phi\}$-invariant. Clearly $G\{\phi\}$ normalizes $N(n) = N \trianglelefteq G$. Suppose that $G\{\phi\}$ normalizes $N(i+1), N(i+2), \ldots, N(n)$, for some $i = 0, 1, \ldots, n-1$. If we replace $N(i)$ by:

$$N'(i) = \bigcap_{\sigma \in G\{\phi\}} N(i)^{\sigma} \trianglelefteq N(i+1),$$

and $N(j)$ by:

$$N'(j) = N(j) \cap N'(i),$$

for $j = 0, 1, \ldots, i-1$, then we obtain a new sequence $N'(0), \ldots, N'(i), N(i+1), \ldots, N(n)$, also satisfying (5.9), in which $G\{\phi\}$ normalizes all elements from $N'(i)$ on. By induction we eventually reach a $G\{\phi\}$-invariant sequence satisfying (5.9).

Once $G\{\phi\}$ normalizes each $N(i)$ in (5.9), Corollary (4.10) tells us that $\chi_i = \phi^{N(i)}$ lies in $\mathrm{ISC}(\psi | \mathfrak{F}G)$ with $G\{\chi_i\} = G\{\phi\} N(i)$, for each $i = 0, 1, \ldots, n$. Since $N(i)$ is $G\{\phi\}$-invariant and normal in $N(i+1)$, we have:

$$\mathrm{Dom}(\chi_i) = N(i) \trianglelefteq G\{\phi\} N(i+1) = G\{\chi_{i+1}\}, \quad \textit{for } i = 0, 1, \ldots, n-1.$$

Because $\chi_i \in \mathrm{ISC}(\chi_{i+1} | \mathfrak{F}G\{\chi_{i+1}\})$ by (4.11c) (with χ_i and $N(i+1)$ in place of ϕ and H), this and Proposition (5.2) imply that $\chi_i \in \mathrm{DCC}(\chi_{i+1} | \mathfrak{F}G)$, for $i = 0, 1, \ldots, n-1$. Therefore (5.5) holds and the theorem is proved.

(5.10) **Corollary.** *If* (4.7) *holds with nilpotent* N, *then* $\mathrm{CC}(\psi | \mathfrak{F}G) = \mathrm{ISC}(\psi | \mathfrak{F}G)$.

Proof. In this case $\mathrm{Dom}(\phi)$ is subnormal in N for any $\phi \in \mathrm{Irs}(\mathfrak{F}N)$. So the theorem implies this corollary.

(5.11) **Corollary.** *If* $\psi \in \mathrm{IS}(\mathfrak{F}G)$ *and* $\mathrm{Dom}(\psi)$ *is nilpotent, then* $\mathrm{CC}(\psi | \mathfrak{F}G) = \mathrm{ISC}(\psi | \mathfrak{F}G)$.

Proof. From Corollaries (5.7) and (5.10), and Lemma (4.6), we have:

$$\mathrm{CC}(\psi | \mathfrak{F}G) = \mathrm{CC}(\psi | \mathfrak{F}G\{\psi\}) = \mathrm{ISC}(\psi | \mathfrak{F}G\{\psi\}) = \mathrm{ISC}(\psi | \mathfrak{F}G),$$

since (4.7) holds with $G\{\psi\}$ and $\mathrm{Dom}(\psi)$ in place of G and N, respectively. Thus the corollary is proved.

The following example shows that 'nilpotent' in Corollaries (5.10) and (5.11) cannot be replaced by 'solvable of nilpotent length two'.

(5.12) *Example.* Let p, q be odd primes such that $p \equiv 1 \pmod{q}$. Then there is some $w \in \mathbb{N}$ such that $w \not\equiv 1 \pmod{p}$ but $w^q \equiv 1 \pmod{p}$. Let Q be the cyclic group $\langle \sigma \rangle$ generated by some element σ of order q, and E be the extra-special group of order p^3 and exponent p. Then Q can act as automorphisms of E so that:

(5.13a) $\tau^\sigma = \tau^{w^2}$, *for all* $\tau \in Z$

(5.13b) $\rho^\sigma = \rho^w$, *for all* $\rho \in E/Z$,

where Z is the center $Z(E)$ of E. For G we take the resulting semidirect product QE of Q with E. For N we take G itself. For $\mathfrak{F}$ we take the field $\mathbb{C}$ of all complex numbers.

Fix any $\lambda \neq 1$ in $\mathrm{Irr}(\mathbb{C}Z)$. Since q is odd we have $w^2 \not\equiv 1 \pmod{p}$. So (5.13a) implies that $\lambda^\sigma \neq \lambda$. It follows that:

(5.14) $$G\{\lambda\} = E.$$

By (5.13b) any subgroup A satisfying $Z < A < E$ is normal in $G = QE$. Such an A is elementary abelian of order p^2. So λ can be extended to some $\alpha \in \mathrm{Irr}(\mathbb{C}A \mid \lambda)$. In view of (5.14), α satisfies:

$$G\{\alpha\} = E\{\alpha\} = A.$$

Therefore $\psi = \alpha^G$ lies in $\mathrm{Irr}(\mathbb{C}G)$ by Corollary (2.7). Furthermore $\phi = \alpha^{QA} \in \mathrm{Irr}(\mathbb{C}QA)$ satisfies:

$$\alpha \ll \phi \ll \psi.$$

Since $N = G$, Theorem 4.8 tells us that $\phi \in \mathrm{ISC}(\psi \mid \mathbb{C}G)$.

The construction of G forces any proper normal subgroup M of G to satisfy:

$$Z \leqq M \leqq E.$$

Then any $\theta \in \mathrm{Irr}(\psi \mid \mathbb{C}M)$ must lie over some G-conjugate λ^δ of λ. Since $Z = \mathrm{Dom}(\lambda^\delta)$ is central in M, this implies that:

$$\mathrm{Dom}(\psi\{\theta\}) = G\{\theta\} \leqq G\{\lambda^\delta\} = E^\delta = E.$$

It follows from this, (5.1) and (5.4) that any $\chi \neq \psi$ in $\mathrm{CC}(\psi \mid \mathbb{C}G)$ must satisfy $\mathrm{Dom}(\chi) \leqq E$. Therefore $\phi \notin \mathrm{CC}(\psi \mid \mathbb{C}G)$ even though $\phi \in \mathrm{ISC}(\psi \mid \mathbb{C}G)$.

§ 6. Compound Clifford Correspondents

In this section we always assume that (4.7) holds, i.e., that $N \trianglelefteq G$ and that $\psi \in \mathrm{Irr}(\mathfrak{F}N)$ is G-invariant. We define the set $\mathrm{CCC}(\psi \mid \mathfrak{F}G)$ of all *compound* $\mathfrak{F}G$-*Clifford correspondents* of ψ by:

(6.1) $\mathrm{CCC}(\psi \mid \mathfrak{F}G)$ *consists of all* $\phi \in \mathrm{Irs}(\mathfrak{F}G)$ *for which there exist* $n \in \mathbb{N}$ *and* $\chi_0, \chi_1, \ldots, \chi_{2n} \in \mathrm{Irs}(\mathfrak{F}N)$ *satisfying:*

(6.2a) $$\chi_0 = \psi \quad \text{and} \quad \chi_{2n} = \phi,$$

(6.2b) $$\chi_{2i-1} \in \mathrm{CC}(\chi_{2i-2} \mid \mathfrak{F}G) \cap \mathrm{CC}(\chi_{2i} \mid \mathfrak{F}G), \quad \text{for all } i = 1, \ldots, n,$$

(6.2c) $$G\{\chi_{2i-1}\} \text{ normalizes } \mathrm{Dom}(\chi_{2i}), \quad \text{for all } i = 1, \ldots, n.$$

In essence $\mathrm{CCC}(\psi \mid \mathfrak{F}G)$ consists of all members of $\mathrm{ISC}(\psi \mid \mathfrak{F}G)$ which can be obtained from ψ by repeated applications of Clifford's correspondence, both

upwards and downwards, such that all intermediate characters also lie in $\mathrm{ISC}(\psi|\mathfrak{F}G)$.

(6.3) **Proposition.** $\mathrm{CC}(\psi|\mathfrak{F}G)\subseteq\mathrm{CCC}(\psi|\mathfrak{F}G)\subseteq\mathrm{ISC}(\psi|\mathfrak{F}G)$.

Proof. If $\phi\in\mathrm{CC}(\psi|\mathfrak{F}G)$, then (6.2) holds with $n=1$ and $\chi_1=\phi$. So $\phi\in\mathrm{CCC}(\psi|\mathfrak{F}G)$.

Now assume that $\phi\in\mathrm{CCC}(\psi|\mathfrak{F}G)$. Then there exist $n\in\mathbb{N}$ and $\chi_0,\chi_1,\ldots,\chi_{2n}\in\mathrm{Irs}(\mathfrak{F}|N)$ such that (6.2) holds. So $\chi_0=\psi\in\mathrm{ISC}(\psi|\mathfrak{F}G)$.

Suppose that $\chi_i\in\mathrm{ISC}(\psi|\mathfrak{F}G)$, for some $i=0,1,\ldots,2n-1$. Since $\psi\in\mathrm{IS}(\mathfrak{F}G)$, Proposition (4.3) tells us that:

$$\chi_i\in\mathrm{IS}(\mathfrak{F}G)\quad\textit{and}\quad\chi_i\ll\psi.$$

If i is even, then $\chi_{i+1}\in\mathrm{CC}(\chi_i|\mathfrak{F}G)\subseteq\mathrm{ISC}(\chi_i|\mathfrak{F}G)$ by (6.2b) and Proposition (5.6). So Proposition (4.3) says that:

$$\chi_{i+1}\in\mathrm{IS}(\mathfrak{F}G)\quad\textit{and}\quad\chi_{i+1}\ll\chi_i\ll\psi,$$

and then that $\chi_{i+1}\in\mathrm{ISC}(\psi|\mathfrak{F}G)$.

If i is odd, then $\chi_i\in\mathrm{CC}(\chi_{i+1}|\mathfrak{F}G)$ by (6.2b). It follows that $\chi_i\ll\chi_{i+1}$. Since $G\{\chi_i\}$ normalizes $H=\mathrm{Dom}(\chi_{i+1})$ by (6.2c), Corollary (4.10) says that $\chi_{i+1}=\chi_i^H\in\mathrm{ISC}(\psi|\mathfrak{F}G)$.

We have shown by induction that each χ_i, $i=0,1,\ldots,n$, lies in $\mathrm{ISC}(\psi|\mathfrak{F}G)$. Hence $\phi=\chi_{2n}\in\mathrm{ISC}(\psi|\mathfrak{F}G)$ by (6.2a), and the proposition is proved.

The character ϕ in Example (5.12) lies in $\mathrm{CCC}(\psi|\mathfrak{F}G)$ but not in $\mathrm{CC}(\psi|\mathfrak{F}G)$. So the first inclusion in Proposition (7.3) can be strict. The following example shows that the second inclusion can also be strict, even for solvable groups.

(6.4) *Example.* Let G be the group $\mathrm{GL}(2,3)$ of all non-singular 2×2 matrices with entries in the field of 3 elements. For N we take G itself, for $\mathfrak{F}$ the field $\mathbb{C}$ of complex numbers and for ψ the unique faithful character of degree 4 in $\mathrm{Irr}(\mathbb{C}G)$. Then (4.7) holds, and Theorem (4.8) just says that:

$$\mathrm{ISC}(\psi|\mathbb{C}G)=\{\phi\in\mathrm{Irs}(\mathbb{C}G)|\,\phi^G=\psi\}.$$

The character ψ is monomial, being induced from a suitable linear character ϕ of the normalizer H of a 3-Sylow subgroup of G. However, $\mathrm{CCC}(\psi|\mathbb{C}G)$ consists only of ψ and the two conjugate characters of degree 2 lying in $\mathrm{Irr}(\psi|\mathbb{C}\mathrm{SL}(2,3))$, where $\mathrm{SL}(2,3)$ is the normal subgroup of all matrices of determinant 1 in $G=\mathrm{GL}(2,3)$. So $\phi\notin\mathrm{CCC}(\psi|\mathbb{C}G)$ even though $\phi\in\mathrm{ISC}(\psi|\mathbb{C}G)$.

There is no example similar to this when $|N|$ is odd and $\mathfrak{F}$ is $\mathbb{C}$.

(6.5) **Theorem.** *If* (4.7) *holds with* $\mathfrak{F}=\mathbb{C}$ *and* $|N|$ *odd, then* $\mathrm{CCC}(\psi|\mathbb{C}G)=\mathrm{ISC}(\psi|\mathbb{C}G)$.

Proof. We shall prove the theorem in a series of lemmas, all based on the hypothesis that it is true for all strictly smaller values of $|G|$ and that ϕ has been chosen in $\mathrm{ISC}(\psi|\mathbb{C}G)-\mathrm{CCC}(\psi|\mathbb{C}G)$ so as to maximize $|G\{\phi\}|$.

(6.6) **Lemma.** *ψ is a faithful character of N.*

Proof. Because ψ is G-invariant, its kernel $K=\mathrm{Ker}(\psi)$ is a normal subgroup of G contained in N. If $K>1$, then the factor group $G'=G/K$, its normal subgroup $N'=N/K$ and the unique character $\psi'\in\mathrm{Irr}(\mathbb{C}N')$ from which ψ is inflated satisfy (4.7) with $|N'|$ odd and $|G'|<|G|$. Since $\psi=\phi^N$ (see (4.2) and Proposition (4.3)) we have $\mathrm{Ker}(\phi)\geqq\mathrm{Ker}(\psi)=K$. Hence ϕ is inflated from some unique $\phi'\in\mathrm{Irr}(\mathbb{C}(\mathrm{Dom}(\phi)/K))$. From Proposition (4.3) it is clear that $\phi\in\mathrm{ISC}(\psi|\mathbb{C}G)$ implies that $\phi'\in\mathrm{ISC}(\psi'|\mathbb{C}G')$. So $\phi'\in\mathrm{CCC}(\psi'|\mathbb{C}G')$ by the minimality of $|G|$. It is evident that this implies that $\phi\in\mathrm{CCC}(\psi|\mathbb{C}G)$ (just inflate all the subcharacters and subgroups of G' used in (5.1), (5.4) and (6.1) to define $\phi'\in\mathrm{CCC}(\psi'|\mathbb{C}G')$ to subcharacters and subgroups of G showing that $\phi\in\mathrm{CCC}(\psi|\mathbb{C}G)$). This contradicts the choice of ϕ. So the lemma is proved.

(6.7) **Lemma.** *$G\{\phi\}$ is a maximal subgroup of G and $G\{\phi\}N=G$.*

Proof. We know from Theorem (4.8) that $G\{\phi\}N=G$. If $G\{\phi\}=G$, then Corollary (4.4) implies that $\mathrm{Dom}(\phi)=G\{\phi\}\cap N=N$. So $\phi=\phi^N=\psi\in\mathrm{CCC}(\psi|\mathbb{C}G)$, contradicting the choice of ϕ. Therefore $G\{\phi\}<G$.

Suppose that $G\{\phi\}<K<G$. Since $G\{\phi\}N=G$, we must have $K=G\{\phi\}H$, where $H=K\cap N$ is normalized by $G\{\phi\}$ and satisfies $\mathrm{Dom}(\phi)<H<N$. Corollary (4.10) says that $\phi^H\in\mathrm{ISC}(\psi|\mathbb{C}G)$ with $G\{\phi^H\}=K$. The maximality of $|G\{\phi\}|$ implies that $\phi^H\in\mathrm{CCC}(\psi|\mathbb{C}G)$. So there exist $n\in\mathbb{N}$ and $\chi_0,\chi_1,\ldots,\chi_{2n}\in\mathrm{Irs}(\mathbb{C}N)$ such that (6.2) holds with ϕ^H in place of ϕ.

Evidently (4.7) is satisfied by $G\{\phi^H\}$, its normal subgroup H and the latter's character ϕ^H. Furthermore $H<N$ has odd order and $|G\{\phi^H\}|<|G|$. Since $\phi\in\mathrm{ISC}(\phi^H|\mathbb{C}G\{\phi^H\})$ by (4.11c), the minimality of $|G|$ implies that $\phi\in\mathrm{CCC}(\phi^H|\mathbb{C}G\{\phi^H\})$. Hence there exist $m\in\mathbb{N}$ with $m\geqq n$ and characters $\chi_{2n+1},\ldots,\chi_{2m}\in\mathrm{Irs}(\mathbb{C}H)$ such that:

(6.8a) $$\chi_{2m}=\phi,$$

(6.8b) $$\chi_{2i-1}\in\mathrm{CC}(\chi_{2i-2}|\mathbb{C}G\{\phi^H\})\cap\mathrm{CC}(\chi_{2i}|\mathbb{C}G\{\phi^H\}),\quad \textit{for all } i=n+1,\ldots,m,$$

(6.8c) $$G\{\phi^H\}\{\chi_{2i-1}\}\ \textit{normalizes}\ \mathrm{Dom}(\chi_{2i}),\quad \textit{for all } i=n+1,\ldots,m.$$

(In comparing this with (6.2), remember that χ_{2n} has already been chosen to be ϕ^H.) We know that $\phi^H\in\mathrm{ISC}(\psi|\mathbb{C}G)\subseteq\mathrm{IS}(\mathbb{C}G)$ (see Proposition (4.3)) and that $\chi_i\in\mathrm{CCC}(\phi^H|\mathbb{C}G\{\phi^H\})\subseteq\mathrm{ISC}(\phi^H|\mathbb{C}G\{\phi^H\})$, for $i=2n,\ldots,2m$ (see Proposition (6.3)). It follows from Lemma (4.6) that $\chi_i\in\mathrm{ISC}(\phi^H|\mathbb{C}G)\subseteq\mathrm{IS}(\mathbb{C}G)$ for $i=2n,\ldots,2m$. So Corollary (4.4) implies that $G\{\phi^H\}\{\chi_i\}=G\{\chi_i\}$, for $i=2n,\ldots,2m$. Hence (6.8c) gives (6.2c), with m in place of n. Furthermore, Corollary (5.7) gives:

$$\begin{aligned}\mathrm{CC}(\chi_i|\mathbb{C}G\{\phi^H\})&=\mathrm{CC}(\chi_i|\mathbb{C}G\{\phi^H\}\{\chi_i\})=\mathrm{CC}(\chi_i|\mathbb{C}G\{\chi_i\})\\&=\mathrm{CC}(\chi_i|\mathbb{C}G)\end{aligned}$$

for $i=2n,2n+1,\ldots,2m$. Hence (6.8b) implies (6.2b) with m in place of n. Therefore $\phi\in\mathrm{CCC}(\psi|\mathbb{C}G)$. This contradiction with the choice of ϕ proves the lemma.

The G-core:

$$C = \bigcap_{\sigma \in G} \mathrm{Dom}(\phi)^{\sigma} \tag{6.9}$$

of $\mathrm{Dom}(\phi)$ is a normal subgroup of G contained in N. The next two lemmas pin down its structure.

(6.10) **Lemma.** *Suppose that $A \leqq C$ is abelian with $A \trianglelefteq G$. Then there is only one character α in $\mathrm{Irr}(\psi|\mathbb{C}A)$. This α is G-invariant and faithful. Hence A is cyclic and central in G.*

Proof. Since $\psi = \phi^N$ and $\mathrm{Dom}(\phi) \geqq C \geqq A$, there is some $\alpha \in \mathrm{Irr}(\psi|\mathbb{C}A)$ with $\alpha \leqq \phi$. In view of Lemma (6.6), it will suffice to show that α is the only element of $\mathrm{Irr}(\psi|\mathbb{C}A)$. If it isn't, then Clifford's Theory tells us that $N\{\alpha\} < N$. Hence $G\{\alpha\}$, its normal subgroup $N\{\alpha\}$ and the $G\{\alpha\}$-invariant character $\psi\{\alpha\} \in \mathrm{Irr}(\mathbb{C}N\{\alpha\})$ satisfy (4.7) with $|N\{\alpha\}|$ odd and $|G\{\alpha\}| < |G|$.

We know from Corollary (2.7) that $\alpha \in \mathrm{IS}(\mathbb{C}G)$ and from Proposition (4.3) that $\phi \in \mathrm{IS}(\psi|\mathbb{C}G|\alpha)$ with $\psi\{\phi\} = \phi$. So Corollary (2.14) and Theorem (3.1) imply that $\phi\{\alpha\} \in \mathrm{IS}(\psi\{\alpha\}|\mathbb{C}G\{\alpha\})$ with:

$$\psi\{\alpha\}\{\phi\{\alpha\}\} = \psi\{\phi\}\{\alpha\} = \phi\{\alpha\}.$$

Therefore $\phi\{\alpha\} \in \mathrm{ISC}(\psi\{\alpha\}|\mathbb{C}G\{\alpha\})$. The minimality of $|G|$ forces $\phi\{\alpha\}$ to lie in $\mathrm{CCC}(\psi\{\alpha\}|\mathbb{C}G\{\alpha\})$. So there exist $n \in \mathbb{N}$ with $n \geqq 2$ and $\chi_2, \chi_3, \ldots, \chi_{2n-2} \in \mathrm{Irs}(\mathbb{C}N\{\alpha\})$ such that:

(6.11a) $$\chi_2 = \psi\{\alpha\} \quad \textit{and} \quad \chi_{2n-2} = \phi\{\alpha\},$$

(6.11b) $$\chi_{2i-1} \in \mathrm{CC}(\chi_{2i-2}|\mathbb{C}G\{\alpha\}) \cap \mathrm{CC}(\chi_{2i}|\mathbb{C}G\{\alpha\}), \quad \textit{for } i = 2, \ldots, n-1,$$

(6.11c) $$G\{\alpha\}\{\chi_{2i-1}\} \textit{ normalizes } \mathrm{Dom}(\chi_{2i}), \quad \textit{for } i = 2, \ldots, n-1.$$

Fix $i = 2, \ldots, 2n-2$. Since $\alpha \in \mathrm{IS}(\mathbb{C}G)$ and $\psi \in \mathrm{IS}(\mathbb{C}G|\alpha)$, Theorem (3.1) tells us that $\psi\{\alpha\} \in \mathrm{IS}(\mathbb{C}G) \cap \mathrm{IS}(\mathbb{C}G\{\alpha\})$ with:

$$G\{\alpha\}\{\psi\{\alpha\}\} = G\{\psi\{\alpha\}\}.$$

So Proposition (6.3) and Lemma (4.6) give:

$$\begin{aligned}\chi_i \in \mathrm{CCC}(\psi\{\alpha\}|\mathbb{C}G\{\alpha\}) \subseteq \mathrm{ISC}(\psi\{\alpha\}|\mathbb{C}G\{\alpha\}) &= \mathrm{ISC}(\psi\{\alpha\}|\mathbb{C}G\{\alpha\}\{\psi\{\alpha\}\}) \\ &= \mathrm{ISC}(\psi\{\alpha\}|\mathbb{C}G\{\psi\{\alpha\}\}) = \mathrm{ISC}(\psi\{\alpha\}|\mathbb{C}G).\end{aligned}$$

In particular, $\chi_i \in \mathrm{IS}(\mathbb{C}G)$ by Proposition (4.3). Because A is abelian, no positive multiple of α can be induced from a character of any $B < A$. Since $(\chi_i^{N\{\alpha\}})_A = \psi\{\alpha\}_A$ is such a multiple of α, this forces χ_i to lie in $\mathrm{IS}(\mathbb{C}G|\alpha)$, with $\chi_i\{\alpha\} = \chi_i$. Now Theorem (3.1) tells us that:

$$G\{\alpha\}\{\chi_i\} = G\{\alpha\}\{\chi_i\{\alpha\}\} = G\{\chi_i\{\alpha\}\} = G\{\chi_i\}.$$

Corollary (5.7) and the fact that $\chi_i \in \mathrm{IS}(\mathbb{C}G\{\alpha\}) \cap \mathrm{IS}(\mathbb{C}G)$ give:

$$\mathrm{CC}(\chi_i|\mathbb{C}G\{\alpha\}) = \mathrm{CC}(\chi_i|\mathbb{C}G\{\alpha\}\{\chi_i\}) = \mathrm{CC}(\chi_i|\mathbb{C}G\{\chi_i\}) = \mathrm{CC}(\chi_i|\mathbb{C}G).$$

The preceding two equations imply that (6.11b, c) are just (6.2b, c) for $i=2, \ldots, n-1$. Evidently $\psi\{\alpha\} \in \mathrm{DCC}(\psi|\mathbb{C}G) \subseteq \mathrm{CC}(\psi|\mathbb{C}G)$ (see (5.1) and Proposition (5.6)). So (6.2b, c) also hold for $i=1$ if we set $\chi_0=\psi$ and $\chi_1=\psi\{\alpha\}=\chi_2$. Finally, $\phi\{\alpha\}$ lies in $\mathrm{CC}(\phi|\mathbb{C}G)$ and $G\{\phi\{\alpha\}\}=G\{\phi\}\{\alpha\}$ by (3.3b). So (6.2b, c) also hold for $i=n$ if we set $\chi_{2n-1}=\phi\{\alpha\}=\chi_{2n-2}$ and $\chi_{2n}=\phi$. Thus (6.2) holds and $\phi \in \mathrm{CCC}(\psi|\mathbb{C}G)$. This contradicts the choice of ϕ and proves the lemma.

(6.12) **Lemma.** *C is cyclic and central in G.*

Proof. Since N has odd order it is solvable [4]. Hence so is its subgroup C. If the Fitting subgroup (the largest normal nilpotent subgroup) F of C is abelian, then F is cyclic and central in G by Lemma (6.10). Since F contains its centralizer $\mathrm{C}(F \text{ in } C)$ in the solvable group C (see Satz III.4.2b) in [5]), the lemma holds in this case. So we may assume that F is not abelian. Then there is some odd prime p such that the p-Sylow subgroup F_p of F is not abelian.

Evidently F_p is a characteristic subgroup of C and hence is normal in G. So Lemma (6.10) implies that any characteristic abelian subgroup of F_p is cyclic. Since p is odd, a theorem of P. Hall (Satz III.13.10 in [5]) tells us that F_p is the central product of its cyclic center and some extra-special subgroup E of exponent p. This subgroup E is characteristic in F_p (it consists of all elements of F_p whose order divides p) and hence is normal in G. So its center $Z=\mathrm{Z}(E)$ is central in G by Lemma (6.10).

Conjugation in G defines a homomorphism γ of G into the group $\mathrm{Aut}(E|Z)$ of all automorphisms of E centralizing Z. It is known (see the proof of Theorem (1.4) in [2], that $\mathrm{Aut}(E|Z)$ is the semi-direct product of the group $\mathrm{Inn}(E)$ of all inner automorphism of E and the centralizer $\mathrm{Aut}(E|Z)\{\iota\}$ of some involution ι inverting E/Z. Since $\gamma(E)=\mathrm{Inn}(E)$, it follows that the inverse image $G\{\iota\}$ of $\mathrm{Aut}(E|Z)\{\iota\}$ in G satisfies:

$$G=G\{\iota\}E, \tag{6.13a}$$

$$G\{\iota\} \cap E=Z, \tag{6.13b}$$

$$\mathrm{C}(E \text{ in } G) \leqq G\{\iota\}. \tag{6.13c}$$

By Lemma (6.10) there is a unique character $\lambda \in \mathrm{Irr}(\psi|\mathbb{C}Z)$, and this λ is faithful and G-invariant. Let Λ be the unique character in $\mathrm{Irr}(\mathbb{C}E|\lambda)$ (see Satz V.16.14 of [5]). It is known (see Proposition 2.3 in [7]) that Λ can be extended to a character Λ^e of the semi-direct product $\mathrm{Aut}(E|Z)\{\iota\} \ltimes E$ of $\mathrm{Aut}(E|Z)\{\iota\}$ and E. Pulling back Λ^e via γ, we obtain a character Ψ of the semi-direct product $G\{\iota\} \ltimes E$ such that $\psi_E=\Lambda$ (here we identify E with its image in $G\{\iota\} \ltimes E$, but do not so identify $G\{\iota\}$ because of the problems raised by $Z=G\{\iota\} \cap E$).

It is evident from (6.13) that the map $H \mapsto H\{\iota\}=H \cap G\{\iota\}$ sends $\{H|E \leqq H \leqq G\}$ one-to-one onto the set $\{K|Z \leqq K \leqq G\{\iota\}\}$, while preserving inclusion and normalizers. As in Theorem (5.10) of [2], the existence of Ψ gives us an isometric bijection $\cdot\{\iota\}$ of $\mathrm{Irs}(\mathbb{C}G|\Lambda)$ onto $\mathrm{Irs}(\mathbb{C}G\{\iota\}|\lambda)$ satisfying:

(6.14) *If* $E \leqq H \leqq G$ *and* $\theta \in \mathrm{Irr}(\mathbb{C}H|\iota)$, *then* $\theta\{\iota\}$ *is the unique character in* $\mathrm{Irr}(\mathbb{C}H\{\iota\}|\lambda)$ *such that:*

$$\theta(\sigma\tau) = \theta\{\iota\}(\sigma)\Psi(\sigma \ltimes \tau), \quad \text{for all } \sigma \in H\{\iota\} \text{ and } \tau \in E.$$

One easily computes that $\cdot\{\iota\}$ preserves $\leqq$ and stabilizers:

(6.15a) *If* $\theta, \chi \in \mathrm{Irs}(\mathbb{C}G|\Lambda)$ *and* $H = \mathrm{Dom}(\theta) \leqq K = \mathrm{Dom}(\chi)$, *then:*

$$(\theta^K, \chi) = (\theta\{\iota\}^{K\{\iota\}}, \chi\{\iota\}).$$

Hence $\theta \leqq \chi$ $(\theta \ll \chi)$ *if and only if* $\theta\{\iota\} \leqq \chi\{\iota\}$ *(resp.,* $\theta\{\iota\} \ll \chi\{\iota\}$*).*

(6.15b) *If* $\theta \in \mathrm{Irs}(\mathbb{C}G|\Lambda)$, *then* $G\{\theta\}\{\iota\} = G\{\iota\}\{\theta\{\iota\}\}$.

Since $N \trianglelefteq G$ we have $N\{\iota\} \trianglelefteq G\{\iota\}$. From $\psi \in \mathrm{Irr}(\mathbb{C}N|\lambda)$ and $N \geqq E$ it follows that ψ lies over the unique character $\Lambda \in \mathrm{Irr}(\mathbb{C}E|\lambda)$. Hence $\psi\{\iota\} \in \mathrm{Irr}(\mathbb{C}N\{\iota\}|\lambda)$ is defined. By (6.15b) the G-invariance of ψ implies the $G\{\iota\}$-invariance of $\psi\{\iota\}$. Thus $G\{\iota\}$, $N\{\iota\}$ and $\psi\{\iota\}$ satisfy (4.7) with $|N\{\iota\}|$ odd and $|G\{\iota\}| < |G|$ (see (6.13)).

Proposition (4.3) says that $\phi \in \mathrm{IS}(\mathbb{C}G)$ and $\phi \ll \psi$. Since Z is central, it follows from $\phi^N = \psi \geqq \lambda$ and $\mathrm{Dom}(\phi) \geqq C \geqq E$ that $\phi \geqq \lambda$ and $\phi \geqq \Lambda$. Hence $\phi\{\iota\} \in \mathrm{Irs}(\mathbb{C}N\{\iota\}|\lambda)$ is defined. By (6.15a) we have $\phi\{\iota\} \ll \psi\{\iota\}$. Since induction to G is an isometry on $\mathrm{Irr}(\mathbb{C}G\{\phi\}|\phi)$, it follows from (6.15) that induction to $G\{\iota\}$ is an isometry on $\mathrm{Irr}(\mathbb{C}G\{\iota\}\{\phi\{\iota\}\}|\phi\{\iota\})$. Therefore $\phi\{\iota\} \in \mathrm{IS}(\mathbb{C}G\{\iota\})$ and $\phi\{\iota\} \in \mathrm{ISC}(\psi\{\iota\}|\mathbb{C}G\{\iota\})$ by Proposition (4.3).

The minimality of $|G|$ now tells us that $\phi\{\iota\} \in \mathrm{CCC}(\psi\{\iota\}|\mathbb{C}G\{\iota\})$. So there exist $n \in \mathbb{N}$ and characters $\chi'_0, \chi'_1, \ldots, \chi'_{2n} \in \mathrm{Irs}(\mathbb{C}N\{\iota\})$ such that:

(6.16a) $$\chi'_0 = \psi\{\iota\} \quad \text{and} \quad \chi'_{2n} = \phi\{\iota\},$$

(6.16b) $\chi'_{2i-1} \in \mathrm{CC}(\chi'_{2i-2}|\mathbb{C}G\{\iota\}) \cap \mathrm{CC}(\chi'_{2i}|\mathbb{C}G\{\iota\})$, *for all* $i = 1, \ldots, n$,

(6.16c) $G\{\iota\}\{\chi'_{2i-1}\}$ *normalizes* $\mathrm{Dom}(\chi'_{2i})$, *for all* $i = 1, \ldots, n$.

Since each χ'_i, for $i = 0, 1, \ldots, 2n$, satisfies $(\chi'_i)^{N\{\iota\}} = \psi\{\iota\} \geqq \lambda$ and Z is central, we have $\chi'_i \geqq \lambda$. Hence there exist unique characters $\chi_0, \chi_1, \ldots, \chi_{2n} \in \mathrm{Irs}(\mathbb{C}N|\Lambda)$ such that $\chi'_i = \chi_i\{\iota\}$ for $i = 0, 1, \ldots, 2n$. It follows from (6.16a) that (6.2a) holds, and from (6.15b) and (6.16c) that (6.2c) holds. We must show that (6.16b) implies (6.2b). It suffices to prove:

(6.17) *If* $\theta, \chi \in \mathrm{Irs}(\mathbb{C}G|\Lambda)$ *and* $\theta\{\iota\}$ $cc-\mathbb{C}G\{\iota\}$ $\chi\{\iota\}$, *then* θ $cc-\mathbb{C}G$ χ.

By (5.4) there exist $m \in \mathbb{N}$ and $\theta'_0, \theta'_1, \ldots, \theta'_m \in \mathrm{Irs}(\mathbb{C}G\{\iota\})$ such that:

$$\theta\{\iota\} = \theta'_0 \; dcc-\mathbb{C}G\{\iota\} \; \theta'_1 \; dcc-\mathbb{C}G\{\iota\} \ldots dcc-\mathbb{C}G\{\iota\} \; \theta'_m = \chi\{\iota\}.$$

Each θ'_i lies in $\mathrm{CC}(\chi\{\iota\}|\mathbb{C}G\{\iota\})$, and hence satisfies $(\theta'_i)^{\mathrm{Dom}(\chi\{\iota\})} = \chi\{\iota\} \geqq \lambda$. Since Z is central, this implies that $\theta'_i \geqq \lambda$. So there are unique characters $\theta_0, \theta_1, \ldots, \theta_m \in \mathrm{Irs}(\mathbb{C}G|\Lambda)$ such that $\theta'_i = \theta_i\{\iota\}$, for $i = 0, 1, \ldots, n$. From this and (5.4) we conclude that (6.17) will be implied by:

(6.18) *If* $\theta, \chi \in \mathrm{Irs}(\mathbb{C}G|\Lambda)$ *and* $\theta\{\iota\}$ $dcc-\mathbb{C}G\{\iota\}$ $\chi\{\iota\}$, *then* θ $dcc-\mathbb{C}G$ χ.

Now (5.1) gives us a character $\delta' \in \mathrm{Irs}(\chi\{\iota\}|\mathbb{C}G\{\iota\})$ such that $D = \mathrm{Dom}(\delta') \trianglelefteq G\{\iota\}\{\chi\{\iota\}\}$ and $\theta\{\iota\} = \chi\{\iota\}\{\delta'\}$. If $D \not\geqq Z$, then $D \cap Z = 1$, since $|Z| = p$ is a prime. So the product $DZ = D \times Z$ is direct. Evidently the character $\delta' \times \lambda \in \mathrm{Irr}(\mathbb{C}(D \times Z))$ has the same stabilizer $G\{\iota\}\{\delta' \times \lambda\} = G\{\iota\}\{\delta'\}$ as δ' since Z is central. It follows (see (2.10)) that $D \times Z \trianglelefteq G\{\iota\}\{\chi\{\iota\}\}$, that $\delta \times \lambda \leqq \chi\{\iota\}$ and that $\chi\{\iota\}\{\delta' \times \lambda\} = \chi\{\iota\}\{\delta'\}$. So we may replace δ' by $\delta' \times \lambda$ and assume that:

(6.19) $$D \geqq Z.$$

Since Z is central, it follows from (6.19) and $\delta' \leqq \chi\{\iota\} \geqq \lambda$ that $\delta' \geqq \lambda$. Hence there is a unique character $\delta \in \mathrm{Irs}(\mathbb{C}G|\Lambda)$ such that $\delta' = \delta\{\iota\}$. It follows from (6.15b) that $D \trianglelefteq G\{\chi\}\{\iota\}$ and hence that $\mathrm{Dom}(\delta) = DE \trianglelefteq G\{\chi\}\{\iota\}E = G\{\chi\}$. Using (6.15a) and (2.10), we see that $\chi\{\delta\}\{\iota\} = \chi\{\iota\}\{\delta\{\iota\}\} = \theta\{\iota\}$. Hence $\chi\{\delta\} = \theta$. This completes the proofs of (6.18), of (6.17) and of the lemma.

The group N of odd order is solvable [4]. So Lemma (6.7) implies the existence of a subgroup B satisfying:

(6.20a) $$C < B \leqq N \quad \textit{and} \quad B \trianglelefteq G,$$

(6.20b) $$G\{\phi\}B = G \quad \textit{and} \quad G\{\phi\} \cap B = C,$$

(6.20c) *B/C is an elementary abelian p-group, for some odd prime p,*

(6.20d) *$G\{\phi\}/C$ acts irreducibly on B/C by conjugation.*

(6.20e) *$N\{\phi\}/C = \mathrm{Dom}(\phi)/C$ acts faithfully on B/C.*

It is easy to show that the action (6.20d) is non-trivial.

(6.21). **Lemma.** *There is some normal p'-subgroup K of $G\{\phi\}$ such that $[B/C, K] = B/C$ and $K \leqq \mathrm{Dom}(\phi)$.*

Proof. As usual, $[B/C, K]$ is the subgroup of B/C generated by all commutators $[\sigma, \tau] = \sigma^{-1}\sigma^{\tau}$, for $\sigma \in B/C$ and $\tau \in K$.

Suppose that $\mathrm{Dom}(\phi)$ centralizes B/C. Then (6.20e) implies that $\mathrm{Dom}(\phi) = C \trianglelefteq G$. So $\phi \in \mathrm{DCC}(\psi|\mathbb{C}G)$ by Proposition (5.2). Since $\mathrm{DCC}(\psi|\mathbb{C}G) \subseteq \mathrm{CCC}(\psi|\mathbb{C}G)$ (see (5.4) and (6.1)), this contradicts the choice of ϕ. So $\mathrm{Dom}(\phi)$ does not centralize B/C. Since $\mathrm{Dom}(\phi) \trianglelefteq G\{\phi\}$, we may choose a subgroup Q of $\mathrm{Dom}(\phi)$ such that $C < Q$ and Q/C is a minimal normal subgroup of $G\{\phi\}/C$. Because $\mathrm{Dom}(\phi) \leqq N$ is solvable, Q/C is an elementary abelian q-group, for some prime q. This and the centrality of C (see Lemma 6.12) force Q to be nilpotent group whose q-Sylow subgroup K is normal in $G\{\phi\}$ and satisfies $KC = Q$. From (6.20c, d, e) it follows that $q \neq p$ and that $[B/C, K] = [B/C, Q/C] = B/C$. So the lemma is proved.

(6.22) **Lemma.** *B is abelian.*

Proof. If B is not abelian, then (6.20) and Lemma (6.12) imply that B is the central product of C and some characteristic extra-special subgroup E of exponent p, with the center $Z = Z(E)$ amalgamated. By Lemma (6.10) there is a unique faithful character $\lambda \in \mathrm{Irr}(\psi|\mathbb{C}Z)$. The centrality of $Z \leqq C$ and the equation $\phi^N = \psi \in \mathrm{Irr}(\mathbb{C}N|\lambda)$ imply that $\phi \geqq \lambda$. In view of Lemma (6.21) we may apply Lemma (4.4) of [3] with the present $\mathrm{Dom}(\phi)$ and N in place of the H

and G used there. Since $\mathrm{Dom}(\phi)\cap E = C\cap E = Z$ by (6.20b), we obtain a contradiction with Corollary (4.8) of [3], which says that $\mathrm{Dom}(\phi)\cap E$ must be a maximal abelian subgroup of E. Thus the lemma holds.

(6.23) **Lemma.** *B is not abelian.*

Proof. If B is abelian, then (6.20a, b) and Lemmas (6.12) and (6.21) imply that $B = C\times[B,K]$, with both C and $[B,K]$ normal in G. Since $\mathrm{Dom}(\phi)\cap B = C$, the restriction $\psi_{[B,K]}$ of $\psi=\phi^N$ must be a multiple of the regular character of $[B,K]$. This is impossible for the restriction to the non-trivial normal subgroup $[B,K]$ of the irreducible character ψ. Therefore the lemma holds.

The evident contradiction between Lemmas (6.22) and (6.23) proves Theorem (6.5).

§ 7. Open Questions

The following questions are suggested by the above material.

(7.1) *Question. Is Theorem* (6.5) *valid with arbitrary fields $\mathfrak{F}$ of characteristic zero in place of $\mathbb{C}$?*

The answer to the above question is probably yes, but any proof would involve delicate arguments with Clifford extensions going far beyond the scope of this article.

Corollary (4.10) says, in part, that if (4.7) holds and $\phi,\chi\in\mathrm{ISC}(\psi|\mathfrak{F}G)$ satisfy $\phi\leqq\chi$, then $\phi\in\mathrm{ISC}(\chi|\mathfrak{F}G\{\chi\})$ (this is (4.11c) since $\chi=\phi^{\mathrm{Dom}(\chi)}$ by Corollary 4.5). One can pose the corresponding questions for the subsets $\mathrm{CC}(\psi|\mathfrak{F}G)$ and $\mathrm{CCC}(\psi|\mathfrak{F}G)$ of $\mathrm{ISC}(\psi|\mathfrak{F}G)$.

(7.2a) *Question. If* (4.7) *holds and $\phi,\chi\in\mathrm{CC}(\psi|\mathfrak{F}G)$ satisfy $\phi\leqq\chi$, is ϕ an element of* $\mathrm{CC}(\chi|\mathfrak{F}G\{\chi\})$?

(7.2b) *Question. If* (4.7) *holds and $\phi,\chi\in\mathrm{CCC}(\psi|\mathfrak{F}G)$ satisfy $\phi\leqq\chi$, is ϕ an element of* $\mathrm{CCC}(\chi|\mathfrak{F}G\{\chi\})$?

The answer to Question (7.2a) is most likely no, but we have no example showing this. It is possible that the answer to Question (7.2b) is yes, due to the much greater freedom we have in constructing members of "CCC" than in constructing those of "CC".

References

1. Dade, E.: Compounding Clifford's Theory. Ann. of Math. (2) **91**, 236–290 (1970)
2. Dade, E.: Characters of Groups with Normal Extra-Special Subgroups. Math Z. **152,** 1–31 (1976)
3. Dade, E.: Monomial Characters and Normal Subgroups. Math. Z. **178**, 401–420 (1981)
4. Feit, W., Thompson, J.: Solvability of Groups of Odd Order. Pacific J. Math. **13**, 755–1029 (1963)
5. Huppert, B.: Endliche Gruppen I. Berlin-Heidelberg-New York: Springer 1967
6. Isaacs, I.M.: Characters of π-Separable Groups. Preprint.
7. Ward, H.N.: Representations of Symplectic Groups. J. Algebra **20**, 182–195 (1972)

Received July 4, 1984

Math. Z. 189, 41-53 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Cauchy Kernels in Strictly Pseudoconvex Domains and an Application to a Mergelyan Type Approximation Problem

Joaquim Bruna[1], Julià Cufí[1] and Joan Verdera[2]

[1] Secció de Matemàtiques, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain
[2] Facultat de Matemàtiques, Universitat de Barcelona, Gran Via 585, Barcelona 7, Spain

Introduction

Let D be a bounded, strictly pseudoconvex domain in $\mathbb{C}^n$ and let $H(\zeta, z)$, $\zeta \in bD$, $z \in D$, denote the Henkin-Ramírez kernel for D ([5]). For an holomorphic function h in D, continuous on $\bar{D}$, one has

$$h(z) = c_n \int_{bD} h(\xi)\, H(\zeta, z), \qquad z \in D.$$

Starting from Henkin's work Øvrelid defined a kernel $C(\zeta, z)$ for $(\zeta, z) \in \bar{D} \times \bar{D}$, $\zeta \neq z$, such that $C(\zeta, z) = H(\zeta, z)$ for $\zeta \in bD$, $z \in D$, and the following "generalized Cauchy formula" holds for arbitrary $u \in C^1(D)$

$$u(z) = c_n \int_{bD} u(\zeta)\, C(\zeta, z) + c_n \int_D \bar{\partial} u(\zeta) \wedge C(\zeta, z), \qquad z \in D.$$

One of the main motivations of this construction was that the kernel $C(\zeta, z)$ serves then to define an explicit solution integral operator for the $\bar{\partial}$-equation: for a regular $(0,1)$ form f in $\bar{D}$, $\bar{\partial} f = 0$, the function

$$Tf(z) = c_n \int_D f(\zeta) \wedge C(\zeta, z), \qquad z \in D$$

satisfies $\bar{\partial} Tf = f$ in D.

Other expressions for Tf (also for forms of arbitrary bedegree) have been given by several authors, using them to obtain optimal estimates in different norms for solutions of the $\bar{\partial}$-equation (see the survey [5] and the references there).

However, Øvrelid's construction is technically involved, and the resulting kernels depend on many choices. In this paper we construct a kernel $C(\zeta, z)$ for which the generalized Cauchy formula holds (and $C(\zeta, z) = H(\zeta, z)$, $\zeta \in bD$) and which, in our opinion, is much easier to handle in applications. The main reason is that near the diagonal of bD it is given by a simple formula just depending on the defining function for D. In case D is strictly convex, the

kernel $C(\zeta, z)$ admits a particularly transparent writing, and for the case of the unit ball it is the same as the one obtained by Charpentier in [4]. We show that $C(\zeta, z)$ enjoys a well known property of Cauchy's kernel in one variable:

$$\int_{B_r(z)\cap D} |C(\zeta, z)|\, dm(\zeta) \leqq \text{const.}\, r.$$

We point out that this can be used to provide very simple proofs of the L^p and Hölder estimates for the $\bar{\partial}$-equation.

In Sect. 2 we apply the results to the following approximation problem. We consider functions f satisfying the equations

$$\frac{\partial^2 f}{\partial \bar{z}_i \partial \bar{z}_j} = 0, \qquad i, j = 1, \ldots, n$$

in D (or, which is the same, $f = h_0 + \sum_{j=1}^{n} h_j \bar{z}_j$ with h_j holomorphic in D, $0 \leqq j \leqq n$). We prove that every continuous function f on $\bar{D}$ satisfying these equations in D is the uniform limit on $\bar{D}$ of functions satisfying them in a (fixed) neighborhood of $\bar{D}$. Recently, problems of this kind, i.e., approximation by modules over the holomorphic functions have been considered in some detail by several authors (see [3, 10, 11] and the references given there).

Section 1

1.1. First we briefly recall Koppelman's formulas (see [7] for proofs). We denote by Δ the diagonal in $\mathbb{C}^n \times \mathbb{C}^n$ and by E the open set $E = \{(\xi, \eta)/\langle \xi, \eta \rangle \neq 0\}$, where $\langle \xi, \eta \rangle = \sum_{j=1}^{n} \xi_j \eta_j$. Let $\omega'(\xi) = \sum_{j=1}^{n} (-1)^{j-1} \xi_j \bigwedge_{i \neq j} d\xi_i$ and $\omega(\eta) = d\eta_1 \wedge \ldots \wedge d\eta_n$. The differential form (the Cauchy-Leray form)

$$\mu = \langle \xi, \eta \rangle^{-n} \mu = \langle \xi, \eta \rangle^{-n} \omega'(\xi) \wedge \omega(\eta)$$

is well defined on E and $d\mu = 0$.

Now let D be a domain with smooth boundary in $\mathbb{C}^n$ and let $s: \bar{D} \times \bar{D} \to \mathbb{C}^n$ be of class C^1 and such that

(1) $|s(\zeta, z)| \leqq c|\zeta - z|, \qquad \zeta, z \in \bar{D}$

(2) $\langle s, \zeta - z \rangle \neq 0$, $(\zeta, z) \in \bar{D} \times \bar{D} \setminus \Delta$, and $|\langle s, \zeta - z \rangle| \geqq c_L |\zeta - z|^2$

for $\zeta \in \bar{D}$ and z in a compact set L of D. Consider now $\psi: \bar{D} \times \bar{D} \setminus \Delta \to E$, $\psi(\zeta, z) = (s(\zeta, z), \zeta - z)$ and define $k(s) = \psi^*(\mu)$, $k_{p,q}(s) =$ component of $k(s)$ of bedegree (p, q) in z and $(n-p, n-q-1)$ in ζ. Then $k_{p,q} = k_{p,q}(s)$ is integrable in D, for each $z \in D$ fixed, and the Koppelman formula reads as follows:

Theorem (Koppelman's formula). *Let $f \in C^1_{(p,q)}(\bar{D})$. Then for $q \geqq 1$*

$$f = C_{p,q,n}\{\int_{bD} f \wedge k_{p,q} + (-1)^{p+q+1} (\int_D \bar{\partial} f \wedge k_{p,q} - \bar{\partial}_z \int_D f \wedge k_{p,q-1})\}$$

and for $q=0$

$$f=C_{p,n}\{\int_{bD} f\wedge k_{p,0}+(-1)^{p+1}\int_D \bar\partial f\wedge k_{p,0}\}.$$

Let S denote the form $S=\sum_{j=1}^{n} s_j d(\zeta_j-z_j)$ if $s=(s_1,\dots,s_n)$. In terms of S the kernel $k(s)$ is written

$$k(s)=(-1)^{n(n-1)/2}/(n-1)!\,\langle s,\zeta-z\rangle^{-n} S\wedge(dS)^{n-1}.$$

This expression shows that $k(s)$ is invariant under dilations of s. It can be also used to show that if s also satisfies

(3) For $\zeta\in bD$, $s(\zeta,\cdot)$ is a multiple of an holomorphic mapping in D,

then $k_{p,q}=0$ for $\zeta\in bD$ and so the following holds:

Theorem. *Let* $f\in C^1_{(p,q)}(\bar D)$ *such that* $\bar\partial f=0$, $q\geqq 1$. *Then*

$$u(z)=C_{p,q,n}\int_D f(\zeta)\wedge k_{p,q-1}(\zeta,z)$$

satisfies $\bar\partial u=f$.

1.2. Now let D be a strictly pseudoconvex domain in $\mathbb{C}^n$ with boundary bD of class C^2. Our aim is to construct a particular map s such that (1) and (2) hold, its restriction to $bD\times D$ is a multiple of Henkin's map ([5]), and for which the kernels $k_{p,q}$ are easy to estimate. Such extension was previously done in [8] but the resulting kernels there are not easy to estimate for our purposes. Our construction will closely follow the one in the preliminary version [2] (see also [4] for the unit ball).

We first recall Henkin's construction. Let ρ be a function of class C^2 in a neighborhood of $\bar D$ which defines D and which is strictly plurisubharmonic in a neighborhood of bD. Let $V_\delta=\{z/|\rho(z)|<\delta\}$, $D_\delta=\{z/\rho(z)<\delta\}$ and $U_{\varepsilon,\delta}=\{(\zeta,z)\in V_\delta\times D_\delta/|\zeta-z|<\varepsilon\}$.

Lemma (Henkin). *There exist positive constants* ε,δ *and* c, *and there exist functions* $\Phi\in C^1(V_\delta\times D_\delta)$, $F,G\in C^1(U_{\varepsilon,\delta})$, *holomorphic in* z *such that*

(a) $\Phi=FG$ *in* $U_{\varepsilon,\delta}$, $F(\zeta,\zeta)=0$, $|G|>c$ *in* $U_{\varepsilon,\delta}$, $|\Phi|>c$ *outside* $U_{\varepsilon,\delta}$.

(b) $2\,\mathrm{Re}\,F(\zeta,z)\geqq\rho(\zeta)-\rho(z)+c|\zeta-z|^2$, $(\zeta,z)\in U_{\varepsilon,\delta}$.

(c) $d_\zeta F(\zeta,z)_{|\zeta=z}=-d_z F(\zeta,z)_{|z=\zeta}=\partial\rho(z)$, $z\in V_\delta$.

Then one writes $\Phi(\zeta,z)=\sum_{j=1}^{n} P_j(\zeta,z)(\zeta_j-z_j)$, with $P_j\in C^1(V_\delta\times D_\delta)$, P_j holomorphic in z, and Henkin's reproducing kernel is the one associated to the map $(\zeta,z)\to P(\zeta,z)=(P_1,\dots,P_n)$ in bD.

1.3. We begin our construction by considering a first extension of the above map to $\bar D\times\bar D$ contained in fact in [1]. Let φ be a C^∞ function, $0\leqq\varphi\leqq 1$, which equals 1 in a neighborhood of bD and vanishes outside V_δ and set

$$u(\zeta,z)=-\rho(\zeta)(\bar\zeta-\bar z)+P(\zeta,z)\,\overline{\Phi(\zeta,z)}\,\varphi(\zeta),\qquad(\zeta,z)\in D_\delta\times D_\delta.$$

Then (1), (2) and (3) are trivially verified.

To obtain more symmetric kernels we will modify u near the diagonal of $bD \times bD$, which is the region of interest concerning the estimate of the kernels. Let

$$V_{\varepsilon,\delta} = \{(\zeta, z)/|\rho(\zeta)|, |\rho(z)| < \delta, |\zeta - z| < \varepsilon\} \quad \text{and} \quad W_{\varepsilon,\delta} = V_{\varepsilon,\delta} \cap (\bar{D} \times \bar{D});$$

we introduce

$$a(\zeta, z) = -\rho(\zeta) + F(\zeta, z), \qquad b(\zeta, z) = a(z, \zeta)$$

and define[1] v in $V_{\varepsilon,\delta}$ by

$$v(\zeta, z) = \rho(\zeta)\frac{P(z,\zeta)}{G(z,\zeta)} + b(\zeta, z)\frac{P(\zeta, z)}{G(\zeta, z)}.$$

Observe that v is a multiple of the mapping considered by Skoda in [9, p. 251].

Note that $v(\zeta, \zeta) = 0$ and that

$$2\,\mathrm{Re}\, a(\zeta, z) \geqq -\rho(\zeta) - \rho(z) + c|\zeta - z|^2 \tag{4}$$

and hence

$$\langle v, \zeta - z\rangle = -\rho(\zeta) F(z, \zeta) + b(\zeta, z) F(\zeta, z) = a(\zeta, z) b(\zeta, z) - \rho(\zeta)\rho(z)$$

satisfies in $W_{\varepsilon,\delta}$

$$\begin{aligned} |\langle v, \zeta - z\rangle| &\geqq |a|\,|b| - \rho(\zeta)\rho(z) \geqq |\mathrm{Re}\, a|\,|\mathrm{Re}\, b| + |\mathrm{Im}\, a|\,|\mathrm{Im}\, b| - \rho(\zeta)\rho(z) \\ &\geqq c\{(\rho(\zeta) - \rho(z))^2 + |\zeta - z|^4 + (-\rho(\zeta) - \rho(z)|\zeta - z|^2 \\ &\quad + |\mathrm{Im}\, F(\zeta, z)|\,|\mathrm{Im}\, F(z, \zeta)|\}. \end{aligned} \tag{5}$$

Multiplying by a scalar function (which will not change the kernels) we may assume that $\langle v, \zeta - z\rangle > 0$ in $V_{\varepsilon,\delta} \setminus \Delta$, and then we can patch together u, v

$$s = \varphi v + (1 - \varphi) u$$

with φ now supported in $V_{\varepsilon,\delta}$, $0 \leqq \varphi \leqq 1$, equal to 1 in $V_{\varepsilon',\delta'}$, $\varepsilon' < \varepsilon$, $\delta' < \delta$. Writing ε, δ instead of ε', δ', this gives a map $s\colon \bar{D}_\delta \times \bar{D}_\delta \to \mathbb{C}^n$ of class C^1 for which (1) and (2) hold, its restriction to $bD \times D$ is a multiple of Henkin's map and $k(s) = k(v)$ near $\Delta \cap (bD \times bD)$, say in $V_{\varepsilon,\delta}$.

The invariance by dilations remarked before also implies that $k_{0,0}$ restricted to $bD \times D$ is Henkin's reproducing kernel. Hence, Koppelman's formula for $p = q = 0$ and the map s can be viewed as the analogue of the Cauchy-Pompeiu formula in one variable. In the following we will write k for $k(s)$ and T for the integral solution operator for the $\bar{\partial}$-equation given by

$$Tf(z) = c_{p,q,n} \int_D f(\zeta) \wedge k_{p,q-1}(\zeta, z), \qquad f \in C^1_{(p,q)}(\bar{D}),\ \bar{\partial} f = 0.$$

---

[1] It has been brought to our attention that a similar construction appears in "Extension dans de classes de Hardy de fonctions holomorphes et estimations de type "mesures de Carleson" pour l'equation $\bar{\partial}''$ by A. Cumenge. Ann. Inst. Fourier **XXXIII**, 59–97 (1983)

1.4. Now we will write down $k(v)$ more explicitely. Set $F_i = P_i/G$, $d(\zeta, z) = a(\zeta, z)\, b(\zeta, z) - \rho(\zeta)\, \rho(z)$ and

$$\alpha(\zeta, z) = \sum_{j=1}^{n} F_j(\zeta, z)(d\zeta_j - dz_j), \qquad \beta(\zeta, z) = \sum_{j=1}^{n} F_j(z, \zeta)(d\zeta_j - dz_j).$$

Then $V = \sum_{j=1}^{n} v_j(d\zeta_j - dz_j) = \rho(\zeta)\,\beta + b\alpha$, $dV = d\rho \wedge \beta + \rho d\beta + db \wedge \alpha + b d\alpha$. Using Newton's binomial formula (which is legitim for 2-forms) we obtain

$$k(v) = c_n d(\zeta, z)^{-n}\{(\rho(\zeta)\,\beta + b\alpha) \wedge (\rho d\beta + b d\alpha)^{n-1} + (n-1)(\rho d\beta + b d\alpha)^{n-2} \wedge (-\rho db + b d\rho) \wedge \beta \wedge \alpha\}. \tag{6}$$

The components $k_{p,q}$ can be easily written using this expression. We will do in case $p = q = 0$. Setting

$$\alpha_1(\zeta, z) = \sum_{j=1}^{n} F_j(\zeta, z)\, d\zeta_j, \qquad \beta_1(\zeta, z) = \sum_{j=1}^{n} F_j(z, \zeta)\, d\zeta_j$$

and using that $F_j(z, \zeta)$ is holomorphic in ζ it turns out that

$$k_{0,0}(v) = c_n \frac{b^{n-1}}{d^n}\{\rho(\zeta)\,\beta_1 \wedge (\bar\partial_\zeta \alpha_1)^{n-1} + b\alpha_1 \wedge (\bar\partial_\zeta \alpha_1)^{n-1} + (n-1)\,\beta_1 \wedge \alpha_1 \wedge \bar\partial \rho(\zeta) \wedge (\bar\partial_\zeta \alpha_1)^{n-2}\}. \tag{7}$$

This is the "Cauchy kernel" $C(\zeta, z)$ we referred to in the introduction. Analogous simple expressions can be derived for the kernels $k_{0,q}(v)$.

1.5. By (4), $-\rho(\zeta) \leqq 2|b(\zeta, z)|$ for $(\zeta, z) \in W_{\varepsilon,\delta}$. It follows from (6) and the fact that all components of $\beta \wedge \alpha$ and all coefficients of $\rho(\zeta)\,\beta + b\alpha$ vanish when $\zeta = z$ that

$$|k_{p,q}(v)(\zeta, z)| \leqq \text{const.} \frac{|b(\zeta, z)|^{n-1}|\zeta - z|}{|d(\zeta, z)|^n}, \qquad (\zeta, z) \in W_{\varepsilon,\delta}. \tag{8}$$

To make easier the exposition we state an auxiliary lemma collecting all estimates we will need on the quantities a, b, d. For $(\zeta, z) \in V_{\varepsilon,\delta}$ we set

$$\psi(\zeta, z) = |F(\zeta, z)| + |F(z, \zeta)|.$$

Lemma. (a) $a(\zeta, z) \approx b(\zeta, z) \approx -\rho(\zeta) - \rho(z) + |\zeta - z|^2 + |\operatorname{Im} F(\zeta, z)|$.

(b) $\psi(\zeta, z) \approx |\rho(\zeta) - \rho(z)| + |\zeta - z|^2 + |\operatorname{Im} F(\zeta, z)|$

(c) $|d(\zeta, z)| \gtrsim \psi^2(\zeta, z) + (-\rho(\zeta) - \rho(z))|\zeta - z|^2$

(d) $|\zeta - z|^2 |b(\zeta, z)| \lesssim |d(\zeta, z)| \leqq |b(\zeta, z)|^2$

(e) $|d(\zeta, z)| \lesssim |\zeta - z|\,|b(\zeta, z)|$ *for* $(\zeta, z) \in W_{\varepsilon,\delta}$.

Proof. That $a(\zeta, z) \approx -\rho(\zeta) - \rho(z) + |\zeta - z|^2 + |\operatorname{Im} F(\zeta, z)|$ follows from (4) and from the fact that in part (b) of Lemma 1.2 also holds the opposite inequality. Since $|F(\zeta, z) + F(z, \zeta)| = O(|\zeta - z|^2)$ one has $|a + b| \lesssim -\rho(\zeta) - \rho(z) + c|\zeta - z|^2$ and

so (a) holds. For (b), just distinguish the cases $\rho(\zeta)\geqq\rho(z)$ and $\rho(\zeta)\leqq\rho(z)$ and use again that $|F(\zeta,z)+F(z,\zeta)|=O(|\zeta-z|^2)$. The estimate (c) is then a restatement of (5). The definition of $d(\zeta,z)$ and (a) trivially imply that $|d|\lesssim|b|^2$ and $|d|\lesssim|\zeta-z|\,|b|$. Finally, $|\zeta-z|^2|b|\lesssim|d|$ follows from the trivial estimate $|b|\leqq-\rho(z)+\psi(\zeta,z)$, from (b) and (c).

Part (c) of the lemma and (7) give

$$(9)\qquad |k_{p,q}(v)(\zeta,z)|\leqq\text{const.}\left\{\frac{|\zeta-z|}{\psi^{n+1}(\zeta,z)}+\frac{|\rho(z)|^{n-1}|\zeta-z|}{[\psi^2+|\rho(z)|\,|\zeta-z|^2]^n}\right\},\qquad (\zeta,z)\in W_{\varepsilon,\delta}.$$

1.6. In this paragraph we will prove the following theorem, which generalizes a familiar property of Cauchy's kernel in one variable.

Theorem. *There exists $c>0$ such that*

$$\int_{B(z,r)\cap\bar{D}}|k_{p,q}(\zeta,z)|\,dm(\zeta)\leqq Cr,\qquad z\in\bar{D},\ r>0.$$

Remark. A (weaker) result of this type has been obtained by Henkin-Leiterer [6] with the Euclidean balls replaced by the smaller Hörmander balls.

Proof of the Theorem. By (1) and (2), $|k_{p,q}(\zeta,z)|\leqq c_L|\zeta-z|^{1-2n}$ for $\zeta\in\bar{D}$ and z in a compact set L of D, and so we may assume that z is close to bD. Since it is enough to prove the estimate for r small, we can replace k by $k(v)$ and use (9).

There exists (see [5]) positive numbers δ_0,r_0, such that for $|\rho(z)|<\delta_0$ it is possible to make a change of coordinates $(\zeta_1,\ldots,\zeta_n)\to(\eta_1(\zeta),\ldots,\eta_n(\zeta))$ in $B(z,r_0)$ such that

$$\eta_1(\zeta)=\rho(\zeta)-\rho(z)+i\,\mathrm{Im}\,F(\zeta,z),\qquad \eta(z)=0,\ |\eta(\zeta)|\approx|\zeta-z|$$

and the jacobian of the transformation is bounded above and below by constants independent of z. We will assume in the following that $|\rho(z)|<\delta_0$ and $r<r_0$.

We express everything in terms of the coordinates $\eta_j=t_{2j-1}+it_{2j}$. Then according to lemma 1.6.(b), $|\psi|\approx|t_1|+|t_2|+|t|^2$ and, also, $|\zeta-z|\approx|t|$. The domain of integration, modulo constants, is contained in the cube $|t_i|\leqq r$. Thus we are led to the following estimates

Estimate 1. $$\int_{|t_i|\leqq r}\frac{|t|}{[|t_1|+|t_2|+|t|^2]^{n+1}}\,dt_1\ldots dt_{2n}\leqq\text{const.}\,r$$

Estimate 2. $$\int_{|t_i|\leqq r}\frac{|t|}{[t_1^2+t_2^2+m|t|^2]^n}\,dt_1\ldots dt_{2n}\leqq const.\,m^{1-n}r\ \text{(here } m=|\rho(z)|).$$

Proof of estimates 1, 2. Changing to polar coordinates we are led to

$$c\int_0^r x^{n-1}\,dx\int_0^\pi\int_0^\pi\frac{\sin^{2n-2}\theta\sin^{2n-3}\varphi}{[|\cos\theta|+\sin\theta|\cos\varphi|+x]^{n+1}}\,d\theta\,d\varphi.$$

When integrating in φ, the change $u=\cos\varphi$ converts it into

$$\int_{-1}^{1} \frac{\sin^{2n-2}\theta \sin^{2n}\varphi}{[|\cos\theta|+u\sin\theta+x]^{n+1}}\,du$$
$$\leqq \sin^{2n-3}\theta \int_{-1}^{1} \frac{\sin\theta}{[|\cos\theta|+u\sin\theta+x]^{n+1}}\,du$$
$$\leqq \frac{\sin^{2n-3}\theta}{[|\cos\theta|+x]^{n}}.$$

The integral of this last expression in θ in $[0,\pi]$ is evaluated in a similar way and is bounded by const. x^{1-n}. The estimate 2 follows in a similar way. This ends the proof of the theorem.

Remark. From the above theorem a simple proof of the L^p-estimates for T can be derived. Just note that the simmetry of the kernels and the theorem imply that

$$\int_D |k_{p,q}(\zeta,z)|\,dm(\zeta)\leqq C, \qquad z\in D$$
$$\int_D |k_{p,q}(\zeta,z)|\,dm(z)\leqq C, \qquad \zeta\in D$$

and use the generalized Young's inequality.

Also the Lip 1/2 estimate can be obtained along the same lines. If $z_1, z_2\in D$ and $\delta=|z_1-z_2|$ is small enough, $\eta=\delta^{1/2}$ and $w=(z_1+z_2)/2$, then

$$|Tf(z_1)-Tf(z_2)|\leqq \|f\|_\infty\{\int_{D\cap B_\eta(w)} (|k(\zeta,z_1)|+|k(\zeta,z_2)|)\,dm(\zeta)$$
$$\times \int_{D\setminus B_\eta(w)} k(\zeta,z_1)-k(\zeta,z_2)|\,dm(\zeta)\}.$$

The first integral is bounded by η, according to the theorem. In the second integral we use the mean value theorem and the easily established inequality

$$|\nabla_z k_{p,q}(\zeta,z)|\leqq \frac{|b(\zeta,z)|^n|\zeta-z|}{|d(\zeta,z)|^{n+1}}, \qquad (\zeta,z)\in W_{\varepsilon,\delta}$$

to bound it by

$$\delta \int_{D\setminus B_\eta(w)} \frac{|b(\zeta,w)|^n|\zeta-w|}{|d(\zeta,w)|^{n+1}}\,dm(\zeta)$$

which in turn is estimated as before by $\delta\cdot\eta^{-1}=\delta^{1/2}$.

Section 2

2.1. In this section we are going to prove the following Mergelyan type result.

Theorem. *Let $D\Subset\mathbb{C}^n$ be a strictly pseudoconvex domain with a C^3 boundary. Then each function $f\in C(\bar{D})$ satisfying*

(10) $$\bar\partial_i \bar\partial_j f = 0, \quad 1 \leqq i, j \leqq n,$$

in D, can be uniformly approximated on $\bar D$ by functions satisfying (10) *in a neighborhood of $\bar D$.*

Remark. As it is easily seen, a function f satisfies (10) in an open set D if and only if there are holomorphic functions in D, $h_0, \ldots, h_n$, with $f = h_0 + h_1 \bar z_1 + \ldots + h_n \bar z_n$ in D. Since $h_j = \bar\partial_j f$, $1 \leqq j \leqq n$, the h_j are determined by f. It can happen that f extends continuously to $\bar D$ and the h_j are even unbounded functions. For example, take D to be the unit ball, $h_j = 0$, $2 \leqq j \leqq n$, $h_1 = -h_0$, and $h_0(z) = (1 - z_1)^{-\frac{1}{2}}$. Therefore the approximation problem we deal with cannot be reduced trivially to an approximation problem for holomorphic functions.

2.2. First we give an outline of the proof of the theorem. In the following, $A_2(D)$ stands for the space of continuous functions on $\bar D$ which satisfy (10) in D. A well known argument ([5], p. 663–664) shows that the approximation result will follow at once, if we can separate the singularities of each function in $A_2(D)$. To do this we find, for $f \in A_2(D) \cap C^1(\bar D)$, an integral formula expressing f as the sum of two boundary integrals. We are then able to write explicitely, for such f, the operators of separation of singularities, which unfortunatly contain a term of the form

$$\int_{bD} \bar\partial f \wedge L(\zeta, z)$$

which makes no sense when f is a general function in $A_2(D)$. The idea is then to move the derivative on f to a derivative on the kernel by performing an integration by parts via Stokes theorem. The final step consists in showing that the new operators, which can be now defined on $A_2(D)$, indeed separate the singularities of each $f \in A_2(D)$.

2.3. To get concise expressions for our formulas we introduce an operator, denoted by θ, acting on forms in the following way

$$\sum f_{IJ} dz_I \wedge d\bar z_J \xrightarrow{\theta} \sum f_{IJ} d\bar z_I \wedge dz_J.$$

This operator satisfies the following properties.

(a) $\theta(\alpha \wedge \beta) = \theta\alpha \wedge \theta\beta$ for all α, β

(b) If α is a form of degree $2n$, then $\theta\alpha = (-1)^n \alpha$.

(c) A function f in $C^2(D)$ satisfies (10) in D if and only if $\bar\partial\theta\,\bar\partial f = 0$ in D.

From now on we will write c_n for the constant $c_{0,n}$, and $C(\zeta, z)$ for $k_{0,0}(\zeta, z)$, s being the map constructed in 1.3. Then, for $(\zeta, z) \in D \times D$,

$$C(\zeta, z) = \langle s, \zeta - z\rangle^{-n} \sum_{i=1}^{n} (-1)^{i-1} s_i(\zeta, z) \bigwedge_{j \neq i} \bar\partial_\zeta s_j(\zeta, z) \wedge \omega(\zeta)$$

and, for $\zeta \in bD$, $C(\zeta, z)$ is Henkin's kernel. We also put

$$L(\zeta, z) = \langle s, \zeta - z\rangle^{-n} \sum_{i,j=1}^{n} (-1)^{i+j} (\bar\zeta_j - \bar z_j)\, s_i \cdot \bar\partial_\zeta s_1 \wedge \ldots \wedge \widehat{\bar\partial_\zeta s_i} \wedge \ldots \wedge \bar\partial_\zeta s_n \wedge d\zeta_1 \wedge \ldots \wedge \widehat{d\zeta_j} \wedge \ldots \wedge d\zeta_n.$$

Lemma. *For each* $f\in A_2(D)\cap C^1(\bar{D})$, *the following integral formula holds*

$$f(z)=c_n\Big(\int_{bD} f(\zeta)\,C(\zeta,z)+(-1)^n\int_{bD}(\theta\bar{\partial}f)(\zeta)\wedge L(\zeta,z)\Big),\qquad z\in D.$$

Proof. Write

$$f(z)=h_0(z)+h_1(z)\,\bar{z}_1+\ldots+h_n(z)\,\bar{z}_n,\qquad z\in\bar{D},$$

with $h_j\in H(D)\cap C(\bar{D})$, $0\leqq j\leqq n$. Then

$$c_n\int_{bD} f(\zeta)\,C(\zeta,z)=f(z)+c_n\int_{bD}\sum_{i=1}^{n} h_i(\zeta)(\bar{\zeta}_i-\bar{z}_i)\,C(\zeta,z).$$

An easy computation shows that the form in the integral of the right-hand side is

$$(-1)^{n-1}(\theta\bar{\partial})f(\zeta)\wedge L(\zeta,z).$$

2.4. To perform the separation of singularities of $f\in A_2(D)\cap C^1(\bar{D})$ we proceed in the following standard way. Given $\chi\in C_0^\infty(\mathbb{C}^n)$, let $U_\chi=D\cup\{z\in D_\delta/d(z,\operatorname{spt}\chi)>\varepsilon\}$ and set

$$(S_\chi f)(z)=c_n\int_{bD} f(\zeta)\,\chi(\zeta)\,C(\zeta,z)$$
$$+(-1)^n c_n\int_{bD}\chi(\zeta)(\theta\bar{\partial}f)(\zeta)\wedge L(\zeta,z),\qquad z\in U_\chi.$$

Let $(B_i)_{1\leqq i\leqq N}$ be a covering of bD by open balls (of small radii), and consider functions $(\chi_i)_{1\leqq i\leqq N}$ in $C_0^\infty(\mathbb{C}^n)$, $0\leqq\chi_i\leqq 1$ $\operatorname{spt}\chi_i\subset B_i$, $\sum_i\chi_i=1$ in a neighborhood of bD. Then, writing S_i in place of S_{χ_i} and U_i in place of U_{χ_i}, we have the following

Lemma. *Let* $f\in A_2(D)\cap C^1(\bar{D})$. *Then*

(a) $f(z)=\sum_i(S_i f)(z)$, $z\in D$.

(b) $\bar{\partial}\theta\bar{\partial}(S_i f)(z)=0$, $z\in U_i$.

(c) $S_i f$ *extends continuously to* $\bar{D}$, $1\leqq i\leqq N$.

Proof. (a) is a direct consequence of Lemma 2.3., and (b) follows from the fact that, for fixed $\zeta\in bD$, $C(\zeta,z)$ and $L(\zeta,z)$ are defined and satisfy (10) in U_i. To prove (c) simply write, using lemma 2.3.,

$$(S_i f)(z)=\chi_i(z)f(z)+c_n\int_{bD} f(\zeta)(\chi_i(\zeta)-\chi_i(z))\,C(\zeta,z)$$
$$+(-1)^n c_n\int_{bD}(\chi_i(\zeta)-\chi_i(z))(\theta\bar{\partial}f)(\zeta)\wedge L(\zeta,z),\qquad z\in D.$$

The middle term extends continuously up to the boundary according to the logarithmic singularity of Henkin's kernel (see [5, p. 655]), and then so does the third one, its kernel being less singular.

2.5. We need the following computational lemma, the proof of which is omitted.

Lemma 1. $\bar{\partial}_\zeta L(\zeta, z) = \theta C(\zeta, z)$, $\zeta \neq z$.

We come now to the integration by parts mentioned above.

Lemma 2. *If* $f \in A_2(D) \cap C^1(\bar{D})$ *and* $\chi \in C_0^\infty(\mathbb{C}^n)$, *then we have*

$$(S_\chi f)(z) = \chi(z) f(z) + \sum_{j=1}^{4} I_j(z), \qquad z \in U_\chi \tag{11}$$

where

$$I_1(z) = -c_n \int_{bD} f(\zeta)(\theta \bar{\partial} \chi)(\zeta) \wedge \theta L(\zeta, z)$$

$$I_2(z) = c_n \int_D f(\zeta)(\bar{\partial}\theta\, \bar{\partial}\chi)(\zeta) \wedge \theta L(\zeta, z)$$

$$I_3(z) = c_n \int_D f(\zeta)\, \bar{\partial}\chi(\zeta) \wedge C(\zeta, z)$$

$$I_4(z) = (-1)^{n-1} c_n \int_D f(\zeta)\, \bar{\partial}\chi(\zeta) \wedge \theta \bar{\partial}_\zeta \theta L(\zeta, z).$$

Proof. Applying Stokes theorem, Lemma 1 and the properties of the operator θ we find

$$\begin{aligned} &\int_{bD} \chi(\zeta)\, \theta \bar{\partial} f(\zeta) \wedge L(\zeta, z) \\ &\quad = \int_D \bar{\partial}\chi(\zeta) \wedge \theta \bar{\partial} f(\zeta) \wedge L(\zeta, z) + (-1)^{n-1} \int_D \chi(\zeta)\, \bar{\partial} f(\zeta) \wedge C(\zeta, z) \\ &\quad = \mathrm{I} + \mathrm{II} \end{aligned}$$

To deal with I, write

$$\bar{\partial}\chi(\zeta) \wedge \theta \bar{\partial} f(\zeta) \wedge L(\zeta, z) = (-1)^{n-1} \bar{\partial} f(\zeta) \wedge \theta \bar{\partial}\chi(\zeta) \wedge \theta L(\zeta, z),$$

and then apply Stoke's theorem to the $(n, n-1)$ form

$$f(\zeta)\, \theta \bar{\partial}\chi(\zeta) \wedge \theta L(\zeta, z)$$

in $D \setminus \bar{B}_r(z)$. Taking into account that

$$|\theta L(\zeta, z)| \leqq c_z |\zeta - z|^{2-2n} \tag{12}$$

$$|\bar{\partial}\theta L(\zeta, z)| \leqq c_z |\zeta - z|^{1-2n} \tag{13}$$

we obtain, letting $r \to 0$,

$$\begin{aligned} (-1)^{n-1} \mathrm{I} = &\int_{bD} f(\zeta)(\theta \bar{\partial}\chi)(\zeta) \wedge \theta L(\zeta, z) \\ &+ \int_D f(\zeta)(\theta \bar{\partial}\chi)(\zeta) \wedge \bar{\partial}_\zeta \theta L(\zeta, z) \\ &- \int_D f(\zeta)\, \bar{\partial}\theta \bar{\partial}\chi(\zeta) \wedge \theta L(\zeta, z). \end{aligned}$$

For II, using Koppelman's formula for the function $f\chi$ we get

$$(-1)^{n-1}\,\mathrm{II}=\int_{bD} f(\zeta)\chi(\zeta)\,C(\zeta,z)-c_n^{-1}f(z)\chi(z)$$
$$-\int_D f(\zeta)\,\bar\partial\chi(\zeta)\wedge C(\zeta,z).$$

This completes the proof of the Lemma.

2.6. In view of the above lemma it is natural to define $S_\chi f$ in case $f\in A_2(D)$ by (11). What remains to show now is that the conclusions of lemma 2.4 still hold for each $f\in A_2(D)$.

Proof of (a) *and* (b). Set $D^\eta=\{z/\rho_\eta(z)=\rho(z)+\eta<0\}$, $\eta>0$. For η small enough D^η is strictly pseudoconvex. Observe that the mappings $s^\eta(\zeta,z)=\varphi(\zeta)\,v^\eta(\zeta,z)+(1-\varphi(\zeta))\,u^\eta(\zeta,z)$ corresponding to D^η can be defined on $D_\delta\times D_\delta$ from the same ε,δ and patching functions φ which were used to construct $s(\zeta,z)$. Some easy computations lead to

$$\langle s^\eta(\zeta,z),\zeta-z\rangle=\langle s(\zeta,z),\zeta-z\rangle+\eta O(|\zeta-z|^2),$$

so that the kernels C^η and L^η converge uniformly to C and L when $z\in U_i$, $\zeta\in\mathrm{spt}\,\chi_i\cap D$ and $|z-\zeta|\geqq\varepsilon$. Moreover for z in a compact subset K of D and $\zeta\in\mathrm{spt}\,\chi_i\cap D$ one has the estimate

$$\langle s^\eta(\zeta,z),\zeta-z\rangle|\geqq C_K|\zeta-z|^2,\qquad \eta>0,$$

and thus, inequalities (12) and (13) hold uniformly in η.

Define $S_i^\eta f=S_{\chi_i}^\eta f$ by formula (11) with D^η in place of D and the kernels C^η and L^η in place of C and L. It is clear that the domains $U_i^\eta=D^\eta\cup\{z\in D_\delta/d(z,\mathrm{sp}\,\chi_i)>\varepsilon\}$ of $S_i^\eta f$ increase to U_i. Since $f\in A_2(D^\eta)\cap C^1(\bar D^\eta)$, we get from lemma 2.4

$$f(z)=\sum_i (S_i^\eta f)(z),\qquad z\in D^\eta.$$

From what has been said above we can let $\eta\to 0$ in the preceding identity to get (a). For (b) just note that $S_i^\eta f\to S_i f$ as $\eta\to 0$, uniformly on compact subsets of U_i.

Proof of (c). We are going to prove that, for $\chi\in C_0^\infty(\mathbb{C}^n)$ and $f\in A_2(\bar D)$, $S_\chi f$ has a continuous extension to $\bar D$. For $I_1(z)$ this is essentially shown in [5, Th. 2.1., p. 655]. The result for $I_3(z)$ follows easily from Theorem 1.6. Since the kernel in $I_2(z)$ is less singular than that in $I_3(z)$, it suffices to pay attention to $I_4(z)$.

It will be enough to prove

$$\int_{B(z,r)\cap D}|\theta\bar\partial_\zeta\theta L(\zeta,z)|\,dm(\zeta)\lesssim r,\qquad z\in D,\ r>0,$$

which, in view of the proof of Theorem 1.6, is a consequence of the following inequality

$$|\theta\bar\partial_\zeta\theta L(\zeta,z)|\lesssim\frac{|\zeta-z|\,|b(\zeta,z)|^{n-1}}{|d(\zeta,z)|^n},\tag{14}$$

for ζ close to bD and z close to ζ.

According to (7) we can write

$$C(\zeta,z)=\frac{b^{n-1}(\zeta,z)}{d^n(\zeta,z)}A(\zeta,z)\wedge\omega(\zeta)$$

where $A(\zeta,z)$ is a $(0,n-1)$ form such that $|A|=O(|\zeta-z|)$. From the definition of $L(\zeta,z)$ one has then

$$L(\zeta,z)=\sum_{j=1}^{n}(-1)^{j-1}(\bar\zeta_j-\bar z_j)\frac{b^{n-1}(\zeta,z)}{d^n(\zeta,z)}A(\zeta,z)\wedge d\zeta_1\wedge\ldots\wedge\widehat{d\zeta_j}\wedge\ldots\wedge d\zeta_n.$$

A short computation leads to

$$\begin{aligned}\theta\bar\partial_\zeta\theta L(\zeta,z)=&\sum_{j=1}^{n}(-1)^{j-1}(\bar\zeta_j-\bar z_j)\frac{b^{n-1}}{d^n}(\theta\bar\partial_\zeta\theta A)\wedge d\zeta_1\wedge\ldots\wedge\widehat{d\zeta_j}\wedge\ldots\wedge d\zeta_n\\&+(-1)^{n-1}n\,C(\zeta,z)-n\sum_{j=1}^{n}(-1)^{j-1}(\bar\zeta_j-\bar z_j)\frac{b^{n-1}}{d^{n+1}}(\theta\bar\partial_\zeta d)\\&\wedge A\wedge d\zeta_1\wedge\ldots\wedge\widehat{d\zeta_j}\wedge\ldots\wedge d\zeta_n=\mathrm{I}+\mathrm{II}+\mathrm{III},\end{aligned}$$

and clearly

$$|\mathrm{I}|+|\mathrm{II}|\lesssim\frac{|\zeta-z|\,|b(\zeta,z)|^{n-1}}{|d(\zeta,z)|^n},$$

so we must worry only about III. Since

$$\bar\partial d(\zeta,z)=-F(z,\zeta)\,\bar\partial\rho(\zeta)+b(\zeta,z)\,B(\zeta,z),$$

where

$$B(\zeta,z)=\sum_{i=1}^{n}\bar\partial_\zeta F_i(\zeta,z)(\zeta_i-z_i),$$

we get

$$\begin{aligned}\mathrm{III}=&n\sum_{j=1}^{n}(-1)^{j-1}(\bar\zeta_j-\bar z_j)\frac{b^{n-1}}{d^{n+1}}F(z,\zeta)\,\theta\bar\partial\rho(\zeta)\\&\wedge A\wedge d\zeta_1\wedge\ldots\wedge\widehat{d\zeta_j}\wedge\ldots\wedge d\zeta_n\\&-n\sum_{j=1}^{n}(-1)^{j-1}(\bar\zeta_j-\bar z_j)\frac{b^n}{d^{n+1}}\theta B\wedge A\wedge d\zeta_1\wedge\ldots\wedge\widehat{d\zeta_j}\wedge\ldots\wedge d\zeta_n\\=&\mathrm{IV}+\mathrm{V}.\end{aligned}$$

The term IV is equal to

$$(-1)^{n-1}n\frac{b^{n-1}(\zeta,z)}{d^{n+1}(\zeta,z)}F(z,\zeta)\left(\sum_{j=1}^{n}\frac{\partial\rho}{\partial\bar\zeta_j}(\zeta)(\bar\zeta_j-\bar z_j)\right)A\wedge\omega(\zeta)$$

and thus, from

$$\sum_{j=1}^{n}\frac{\partial\rho}{\partial\zeta_j}(\zeta)(\zeta_j-z_j)=F(\zeta,z)+O(|\zeta-z|^2)$$

we obtain

$$|\mathrm{IV}| \lesssim \frac{|b(\zeta,z)|^{n-1}}{|d(\zeta,z)|^{n+1}}(|\zeta-z|\,|F(z,\zeta)|\,|F(\zeta,z)|+|\zeta-z|^3|F(z,\zeta)|)$$
$$\lesssim \frac{|\zeta-z|\,|b(\zeta,z)|^{n-1}}{|d(\zeta,z)|^n}+\frac{|\zeta-z|^3\,|b(\zeta,z)|^n}{|d(\zeta,z)|^{n+1}}.$$

On the other hand, since $|B|=O(|\zeta-z|)$ we see that

$$|\mathrm{V}| \lesssim \frac{|\zeta-z|^3\,|b(\zeta,z)|^n}{|d(\zeta,z)|^{n+1}},$$

and therefore, appealing to part (d) of Lemma 1.5, we get (14). This completes the proof of the theorem.

References

1. Berndtsson, B., Andersson, M.: Henkin-Ramírez formulas with weight factors. Ann. Inst. Fourier (Grenoble) **XXXII**, 91–110 (1982)
2. Bruna, J., Cufí, J.: Some explicit kernels for solving the $\bar{\partial}$-equation in strictly convex domains. Preprint
3. Carmona, J.J.: A necessary and sufficient condition for uniform approximation by certain rational modules. Proc. Amer. Math. Soc. **86**, 487–490 (1982)
4. Charpentier, P.: Formules explicites pur les solutions minimales de l'equation $\bar{\partial}u=f$ dans la boule et dans le polydisque. Ann. Inst. Fourier (Grenoble) **XXX**, 121–154 (1980)
5. Henkin, G.M., Cirka, E.M.: Boundary properties of holomorphic functions of several complex variables. J. Soviet Math. **5**, 612–687 (1976)
6. Henkin, G.M., Leiterer, J.: Theory of functions on strictly pseudoconvex sets with non-smooth boundary. Akad. Wiss. DDR. Inst. Math. Rep. Math. 02/81 (1981)
7. Koppelman, W.: The Cauchy integral for differential forms. Bull. Amer. Math. Soc. **73**, 554–556 (1967)
8. Øvrelid, N.: Integral representation formulas and L^p-estimates for the $\bar{\partial}$-equation. Math. Scand. **29**, 137–160 (1971)
9. Skoka, H.: Valeurs au bord pour les solutions de L'opérateur d'' et caracterisation des zéros des fonctions de la classe de Nevanlinna. Bull. Soc. Math. France **104**, 225–299 (1976)
10. Trent, T., Wang, J.L.: The uniform closure of rational modules. Bull. London Math. Soc. **13**, 415–420 (1981)
11. Verdera, J.: Approximation by rational modules in Sobolev and Lipschitz norms. J. Funct. Anal. **58**, 267–290 (1984)

Received July 15, 1983; received in final form January 20, 1984

Math. Z. 189, 55–63 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Skew Products of Relations and the Structure of Simple C^*-Algebras

Bruce Blackadar[1] and Alexander Kumjian[2]

[1] Department of Mathematics, University of Nevada, Reno, Nevada 89557, USA
[2] School of Mathematics, University of New South Wales, P.O. Box 1, Kensington, NSW 2033, Australia

In this paper, we give a general construction which can be applied to give some interesting examples of simple C^*-algebras. The construction is described in two ways: as an inductive limit, reminiscent of the construction in [1] and [2], in Sect. 1; and as the C^*-algebra of a skew product of relations in Sect. 2. The second approach presents the C^*-algebras as groupoid C^*-algebras [21] with an easily described diagonal [16].

The specific examples constructed are simple C^*-algebras with a shortage of projections. The first algebra A_1 is a stable simple C^*-algebra with no nonzero projections. Earlier examples of such C^*-algebras were constructed in [1], [8], and [11; p. 110]; but the construction given here is perhaps the most elementary, requiring neither K-theory nor the theory of AF algebras – it really only uses the Dixmier-Douady classification of continuous trace C^*-algebras [13], along with a related result of Serre and Grothendieck [15]. The skew product description of this algebra owes much to Raeburn and Taylor [20].

The second example A_2 (and its variant A_3) has a more subtle lack of projections: although every hereditary C^*-subalgebra of $A_3 \otimes \mathscr{K}$ contains a nonzero projection, there are nonetheless distinct traces on $A_2 \otimes \mathscr{K}$ and $A_3 \otimes \mathscr{K}$ which agree on all projections. Thus, although A_3 has no minimal projections, it is in a sense "more projectionless" than the examples constructed in [2], which at least have enough projections in matrix algebras to distinguish traces.

As a related result in the spirit of Sect. 2, we show in Sect. 3 that every hereditary C^*-subalgebra of a Bunce-Deddens weighted shift algebra [7] has an approximate identity of projections.

Most of the work presented in this paper was actually done by one or the other author individually. However, the whole has turned out to be more than the sum of its parts, and it is clear to us that the overall results of the paper never would have been obtained without the benefit of the numerous scintillating conversations we have had together over the last several months. This work was done while we were both visitors at the Mathematics Institute,

Universität Tübingen, supported by Forschungsstipendiums from the Alexander von Humboldt-Stiftung. We would like to thank the people of both institutions for their generous hospitality and support.

1. Inductive Limit Construction of the Examples

We first describe a functorial construction of a "twisted double embedding" θ_ρ: $B \to M_2(B)$ for any C^*-algebra B and $\rho \in \mathrm{Aut}(B)$, which can be iterated to give a (functorial) inductive limit construction.

Definition 1.1. Given a C^*-algebra B and $\rho \in \mathrm{Aut}(B)$, let $\theta_\rho(b) = \begin{bmatrix} b & 0 \\ 0 & \rho(b) \end{bmatrix} \in M_2(B)$. Let $M_\infty^\rho(B) = \varinjlim (M_{2^n}(B), \theta_\rho \otimes \mathrm{id})$.

There are of course many possible variations of this construction – the embedding multiplicities and automorphisms can be varied from step to step, for example.

Remark 1.2. We may alternately write $M_\infty^\rho(B)$ as a crossed product as in [22; 2.5(4)] as follows. Let $G = \bigoplus_{\mathbb{N}} \mathbb{Z}_2$ act on $A = B \otimes C(\Omega)$, where Ω is the Cantor set $\{0,1\}^{\mathbb{N}}$, by letting $g_n = (0, \ldots, 0, 1, 0, \ldots)$ act on $A = A_0^{(n)} \oplus A_1^{(n)}$ ($A_i^{(n)} = B \otimes \chi_i^{(n)} C(\Omega) \chi_i^{(n)}$, where $\chi_i^{(n)}$ is the characteristic function of $\{(e_k) : e_n = i\} \subseteq \Omega$) by $\lambda(g_n)(b_0 \otimes f_0, b_1 \otimes f_1) = (\rho(b_1) \otimes f_1, \rho^{-1}(b_0) \otimes f_0)$. Then $M_\infty^\rho(B) \cong A \times_\lambda G$.

Proposition 1.3. *If $X = \mathrm{Prim}(B)$ is a compact Hausdorff space and the homeomorphism of X induced by ρ is minimal, then $M_\infty^\rho(B)$ is simple.*

Proof. The argument is quite similar to the proof of [1; 3.2]. If $y \in X = \mathrm{Prim}(B)$, write $y^{(n)}$ for the primitive ideal of $M_{2^n}(B)$ consisting of matrices with all entries in y. Then

$$y^{(n+k)} \cap \theta_\rho^k(M_{2^n}(B)) = \theta_\rho^k(\{b : b \in y^{(n)}, \rho(b) \in y^{(n)}, \ldots, \rho^k(b) \in y^{(n)}\})$$
$$= \theta_\rho^k(y^{(n)} \cap \rho^{-1}(y^{(n)}) \cap \ldots \cap \rho^{-k}(y^{(n)})).$$

If I is a (closed two-sided) ideal of $M_{2^{n+k}}(B)$ and $I_0 = (\theta_\rho^k)^{-1}(I \cap \theta_\rho^k(M_{2^n}(B))$ is the corresponding ideal of $M_{2^n}(B)$, then I and I_0 correspond to closed subsets E and E_0 of X respectively. $I = \bigcap_{y \in E} y^{(n+k)}$,

$$I_0 = (\theta_\rho^k)^{-1}\left(\bigcap_{y \in E} [y^{(n+k)} \cap \theta_\rho^k(M_{2^n}(B))]\right) = \bigcap_{y \in E} [y^{(n)} \cap \ldots \cap \rho^{-k}(y^{(n)})]$$
$$= \bigcap \{y^{(n)} : y \in E \cup \rho^{-1}(E) \cup \ldots \cup \rho^{-k}(E)\},$$

so $E_0 = E \cup \rho^{-1}(E) \cup \ldots \cup \rho^{-k}(E)$. Now if J is a proper (closed two-sided) ideal of $M(B)$, let $J_n = (\theta_\rho^\infty)^{-1}[J \cap \theta_\rho^\infty(M_{2^n}(B)]$, and F_n the corresponding closed subset of X. For n sufficiently large, J_n is a proper ideal, so F_n is nonempty; and

$$F_n = F_{n+k} \cup \rho^{-1}(F_{n+k}) \cup \ldots \cup \rho^{-k}(F_{n+k})$$

for each k, since $J_n=(\theta_\rho^k)^{-1}[J_{n+k}\cap\theta_\rho^k(M_{2^n}(B))]$, so if $x\in F_n$, for any k we must have $\rho^r(x)\in F_{n+k}$ for some r, $0\leqq r\leqq k$, and hence $\{\rho^{-r}(x),\rho^{-r+1}(x),\ldots,\rho^{-r+k}(x)\}\subseteq F_n$. Thus F_n contains an interval of the orbit of x, including x itself, of length k. Since k is arbitrary, we have either $\{\rho^k(x):k\geqq 0\}$ or $\{\rho^{-k}(x):k\geqq 0\}$ contained in F_n. Each of these sets is dense by the following lemma, so $F_n=X$, $J_n=0$. This is true for all n, so $J=0$. □

Lemma 1.4. *Let X be a compact Hausdorff space, ρ a minimal homeomorphism of X. Then, for any $x\in X$, $\Sigma_+=\{\rho^n(x):n\geqq 0\}$ and $\Sigma_-=\{\rho^{-n}(x):n\geqq 0\}$ are dense in X.*

Proof. Let $E_k=\{\rho^n(x):n\geqq k\}^-$. Then $\bigcap E_k$ is a nonempty closed subset of X which is invariant under ρ, so $\bigcap E_k=X$ by minimality; thus $(\Sigma_+)^-=E_0\supseteq\bigcap E_k=X$. The argument for Σ_- is similar. □

Example 1.5. (A_1) Let B be a stable continuous trace C^*-algebra with $\hat{B}=\mathbb{T}^3$, the 3-torus, with nontrivial Dixmier-Douady invariant ([13], [12; 10.7.14]). (Recall that $H^3(\mathbb{T}^3,\mathbb{Z})\cong\mathbb{Z}$.) Let ρ be the minimal homeomorphism of $\mathbb{T}^3$ given by coordinatewise rotation by three irrational angles θ_1, θ_2, θ_3 which are linearly independent over $\mathbb{Q}$. Since ρ is homotopic to the identity, the induced automorphism ρ^* of $H^3(\mathbb{T}^3,\mathbb{Z})$ is the identity. The stable continuous trace C^*-algebras with fixed base space X are parametrized up to isomorphism exactly by $H^3(X,\mathbb{Z})$ [12; 10.8.4], and so there is an automorphism of B (also denoted ρ) which induces the homeomorphism ρ on $\hat{B}$. We may assume $\rho\cong\gamma\otimes\mathrm{id}$ on $B\cong B\otimes\mathscr{K}$. A specific ρ with these properties is given in 2.5. By [15; 1.7], B contains no nonzero projections since the Dixmier-Douady obstruction is of infinite order. (If B contained a projection p, then pBp would be an Azumaya algebra with the same H^3 invariant as B.) The same is true of $M_{2^n}(B)\cong B$ and hence of $A_1=M_\infty^\rho(B)$. A_1 is simple by 1.3, and stable because a tensor product factor of $\mathscr{K}$ can be split off at each stage: $M_\infty^\rho(B)\cong M_\infty^\gamma(B)\otimes\mathscr{K}$.

The above construction would appear to depend strongly on the choice of the angles θ_1, θ_2, θ_3 (as well as ρ), but it is not clear how to find invariants which will distinguish between the algebras corresponding to different angles. For example, any two such algebras will have identical K-groups. In fact, any such algebra will have the same shape [5] as $B\otimes M_{2^\infty}$, where M_{2^∞} is the UHF algebra of type 2^∞.

Example 1.6. (A_2,A_3) Let X be a connected compact Hausdorff space, and ρ a minimal homeomorphism of X for which there is more than one invariant probability measure. For example, we can take $X=\mathbb{T}^2$, ρ the homeomorphism constructed in [14; 1.17]; there are uncountably many extremal invariant measures in this case. Let $B=C(X)$; we use ρ also to denote the corresponding automorphism of B. Then $A_2=M_\infty^\rho(B)$ is simple. A_2 has many tracial states (any invariant measure on X induces one), but all tracial states of A_2 agree on projections (it is enough to show this for $M_{2^n}(B)$, where it is obvious.) The same is true for any matrix algebra over A_2, and hence also for $A_3=A_2\otimes M_{2^\infty}$.

Recall the following properties ([19, 4]) for a C^*-algebra B.

(LP): The linear span of the projections in B is dense.

(FS): Every self-adjoint element of B can be approximated arbitrarily closely by a self-adjoint element of finite spectrum.

(HP): Every hereditary C^*-subalgebra of B has an approximate identity of projections.

(SP): Every nonzero hereditary C^*-subalgebra of B contains a nonzero projection.

We have (LP) $\Leftarrow$ (FS) $\Leftrightarrow$ (HP) $\Rightarrow$ (SP) and (LP) $\not\Leftarrow$ (FS) by the results of [19] and [4; 2.7]. It is clear that A_2 and A_3 do not have (LP) (all traces on A_2 or A_3 agree on linear combinations of projections), and hence (HP) also fails. However, A_3 does satisfy (SP). In fact, we have a more general result.

If $x\in B$, we let B_x denote the hereditary C^*-subalgebra $(x^*Bx)^-$ of B. For $\delta>0$, we let f_δ be the continuous function on $\mathbb{R}$ which is 0 on $(-\infty, \delta/2]$, 1 on $[\delta, \infty)$, and linear on $[\delta/2, \delta]$.

The proof of the next lemma essentially uses G. Elliott's argument from [19; Prop. 14].

Lemma 1.7. *Let B be a C^*-algebra, x_n, $x\in B_+$ with $x_n\to x$. Suppose for some fixed $\delta>0$ there is a projection $p_n\in B_{f_\delta(x_n)}$ for all n. Let $\varepsilon>0$. Then for all sufficiently large n, there is a projection $q_n\in B_{f_{\delta/2}(x)}$ with $\|p_n-q_n\|<\varepsilon$.*

Proof. Let $0<\eta<1/2$ be such that $2\eta+4\eta(1+\sqrt{1-2\eta})^{-1}<\varepsilon$. We have $f_{\delta/2}(x_n)\to f_{\delta/2}(x)$; choose N so that $\|f_{\delta/2}(x_n)-f_{\delta/2}(x)\|<\eta$ for all $n\geqq N$. Since $f_{\delta/2}(x_n)$ is a unit for p_n, we have (for $n\geqq N$) $\|p_n f_{\delta/2}(x)^2 p_n-p_n\|<2\eta<1$; so $p_n f_{\delta/2}(x)^2 p_n$ and hence also $[p_n f_{\delta/2}(x)^2 p_n]^{1/2}=z_n$ is invertible in B_{p_n}, i.e. there is an $r_n\in(B_{p_n})_+$ with $r_n z_n=z_n r_n=p_n$. We have $\|r_n-p_n\|<2\eta(1+\sqrt{1-2\eta})^{-1}$. Let $u=f_{\delta/2}(x)p_n r_n$. Then $u^*u=p_n$, and $q_n=uu^*\in B_{f_{\delta/2}(x)}$.

$$\begin{aligned}\|u-p_n\| &\leqq \|(f_{\delta/2}(x)p_n-p_n)p_n\| \\ &\quad + \|f_{\delta/2}(x)(r_n-p_n)\|<\eta+2\eta(1+\sqrt{1-2\eta})^{-1},\end{aligned}$$

so

$$\begin{aligned}\|q_n-p_n\| &\leqq \|q_n-p_n u^*\|+\|p_n u^*-p_n\| \\ &\leqq 2\|u-p_n\|<2\eta+4\eta(1+\sqrt{1-2\eta})^{-1}<\varepsilon. \quad \square\end{aligned}$$

Lemma 1.8. *Let B be a C^*-algebra, S a dense subset of B_+.*

(a) *If, for every $x\in S$ and $\delta>0$ with $f_\delta(x)\neq 0$, $B_{f_\delta(x)}$ contains a nonzero projection, then B has (SP).*

(b) *If each $B_{f_\delta(x)}$ has an approximate identity of projections, then B has (HP).*

Proof. We prove only (b); the other is similar but easier. Fix $\varepsilon>0$ and $x\in B_+$, $\|x\|=1$. Let $x_n\in S$ with $x_n\to x$; set $\delta=\varepsilon/5$, and let p_n be a projection in $B_{f_\delta(x_n)}$ with $\|f_\delta(x_n)-f_\delta(x_n)p_n\|<\varepsilon/5$. Choose n large enough that $\|f_\delta(x_n)-f_\delta(x)\|<\varepsilon/5$, and let $q_n\in B_{f_{\delta/2}(x)}\subseteq B_x$ with $\|p_n-q_n\|<\varepsilon/5$. Then we have

$\|f_\delta(x)-f_\delta(x)q_n\|<3\varepsilon/5$. Since $\|x-xf_\delta(x)\|<\delta<\varepsilon/5$, we have

$$\|x-xq_n\|<2\|x-xf_\delta(x)\|+\|xf_\delta(x)-xf_\delta(x)q_n\|<\varepsilon.$$

The general result now follows from [6; I.1.10]. □

Theorem 1.9. *Let B be a simple C^*-algebra and D a UHF algebra. If B contains a nonzero projection, then $A=B\otimes D$ has (SP).*

Proof. The argument is similar to that of [3; A2]. Write $D=\bigotimes_{k=1}^{\infty} D_k$, with D_k a matrix algebra. Let $A_n=B\otimes\left(\bigotimes_{k=1}^{n} D_k\right)\otimes 1\subseteq A$; then $A=[\bigcup A_n]^-$. Let $\delta>0$ and $x\in\left[B\otimes\left(\bigotimes_{k=1}^{n} D_k\right)\right]_+$, with $f_\delta(x)\neq 0$. If p is a nonzero projection in B, then $(p\otimes 1)\otimes 1$ is in the 2-sided ideal of A_n generated algebraically by $f_\delta(x)\otimes 1$. So by the argument of [10; 1.1 IV and 1.7], for sufficiently large m, we have $p\otimes 1\otimes q\otimes 1$ equivalent to a projection in the hereditary C^*-subalgebra of A_{n+m} generated by $f_\delta(x)\otimes 1\otimes 1$, where q is a one-dimensional projection in $\bigotimes_{k=n+1}^{n+m} D_k$. Thus A satisfies the hypotheses of 1.8(a). □

Corollary 1.10. *A_3 satisfies (SP).*

It follows that (SP) $\not\Rightarrow$ (LP), and *a fortiori* (SP) $\not\Rightarrow$ (HP). This is perhaps somewhat surprising, since one might hope to use some kind of exhaustion argument to deduce (HP) from (SP).

2. Skew Products of Relations

The twisted double embedding construction of Sect. 1 may be viewed as a C^*-cross-product where the group is replaced by a groupoid (the action is given by a 1-cocycle with values in the automorphism group of the C^*-algebra.) In this section we confine ourselves to introducing a notion of skew product of relations; here the action is given by a 1-cocycle defined on one relation taking values in the automorphism group of another (two kindred notions may be found in [21] under the names skew-product and semi-direct product.) The examples of Sect. 1 can then be exhibited as the C^*-algebras of skew products with appropriate 2-cocycles.

Recall that a principal discrete groupoid is a locally compact topological groupoid G with unit space X and source and range maps s, r: $G\to X$ which are local homeomorphisms, such that the map $\alpha\mapsto(r(\alpha),s(\alpha))$ from G to $X\times X$ is one-to-one. G may then be regarded as an equivalence relation on X. In the following, the phrase "R is a relation on X" means that R is a principal discrete groupoid with unit space X. The inverse is denoted by an asterisk (*).

Let R be a relation on X and Q a relation on Y; we consider continuous 1-cocycles on R with values in the automorphism group of Q.

Definition 2.1. *A* continuous function $\phi: R \times Q \to Q$ is an $R-Q$ cocycle (written $(\alpha, \beta) \mapsto \phi(\alpha)\beta$) if:

(i) $\phi(\alpha_1\alpha_2)\beta = \phi(\alpha_1)\phi(\alpha_2)\beta \quad \alpha_1, \alpha_2 \in R, \beta \in Q$

(ii) $\phi(\alpha)(\beta_1\beta_2) = (\phi(\alpha)\beta_1)(\phi(\alpha)\beta_2) \quad \alpha \in R, \beta_1, \beta_2 \in Q$

(iii) $\phi(s(\alpha))\beta = \beta \quad \alpha \in R, \beta \in Q$.

Note. In order for (ii) to make sense $\phi(\alpha)$ must map composable pairs to composable pairs. It follows that each $\phi(\alpha)$ is a structure-preserving homeomorphism on Q (with inverse $\phi(\alpha^*)$), hence an automorphism.

Definition 2.2. Given an $R-Q$ cocycle ϕ, the skew product $R \underset{\phi}{\times} Q$ is the relation on $X \times Y$ which, as a topological space, is $R \times Q$, with groupoid structure given by the formulae:

i) $s(\alpha, \beta) = (s(\alpha), s(\beta))$

ii) $(\alpha, \beta)^* = (\alpha^*, \phi(\alpha)\beta^*)$

iii) $(\alpha_1, \phi(\alpha_2)\beta_1)(\alpha_2, \beta_2) = (\alpha_1\alpha_2, \beta_1\beta_2)$.

Proposition 2.3. *With the above operations,* $R \underset{\phi}{\times} Q$ *is a relation on* $X \times Y$.

Proof. Clearly, all the above operations are continuous. The range map is given by $r(\alpha, \beta) = (r(\alpha), r(\phi(\alpha)\beta))$. Then the image of $R \underset{\phi}{\times} Q$ under the map $r \times s$ constitutes an equivalence relation on $X \times Y$. □

Note that if $\phi: R \times Q \to Q$ is trivial (i.e. $\phi(\alpha)\beta = \beta$), then the skew product is simply the direct product relation $R \times Q$. This is true more generally for suitably defined coboundaries. An $R-Q$ cocycle ϕ is said to be a coboundary if there is a continuous map $h: X \times Q \to Q$ with $h(x) \in \mathrm{Aut}(Q)$ for each $x \in X$, such that $\phi(\alpha) = h(r(\alpha))^{-1} h(s(\alpha))$. In this case, $R \underset{\phi}{\times} Q \cong R \times Q$ under the correspondence $(\alpha, \beta) \mapsto (\alpha, h(s(\alpha))\beta)$.

Example 2.4. Let R be the *CAR* relation on the Cantor set Ω and $m: R \to \mathbb{Z}$ be the number cocycle (i.e. $\Omega = \{(e_i): e_i = 0 \text{ or } 1, i \geqq 0\}$ and $(e_i) \sim (f_i)$ if there is $n \geqq 0$ so that $e_i = f_i$ for all $i \geqq n$; $m(e, f) = \sum_i (e_i - f_i)$, cf. [21; pp. 129–130].) Let ρ be a homeomorphism on a locally compact space X. Then, viewing X as the trivial relation on itself, we define an $R-X$ cocycle ϕ by $\phi(\alpha)x = \rho^{m(\alpha)}(x)$. Note that if ρ is a minimal homeomorphism, then the skew product $R \underset{\phi}{\times} X$ is a minimal relation, i.e. there are no nontrivial invariant open subsets of $\Omega \times X$. Further, the associated C^*-algebra $C^*(R \underset{\phi}{\times} X)$ is isomorphic to $M_\infty^\rho(C_0(X))$. This may be seen as follows.

Topologically, $R \underset{\phi}{\times} X$ is a countable disjoint union of spaces homeomorphic to $\Omega \times X$, indexed by $G = \bigoplus_{k=1}^{\infty} \mathbb{Z}_2$, and $C^*(R \underset{\phi}{\times} X)$ is the completion of $C_c(R \underset{\phi}{\times} X)$ as described in [21]. If A_n is the set of functions which vanish outside $G_n = \{0, 1\}^n \times \{0\}$ and whose values in each copy of $\Omega \times X$ depend only on the point in X and the first n coordinates in Ω, then A_n is a C^*-subalgebra of $C^*(R \underset{\phi}{\times} X)$ and $C^*(R \underset{\phi}{\times} X) = \varinjlim A_n$. Note that $A_n \cong C^*(R_n \underset{\phi_n}{\times} X)$, where R_n is the transitive relation on the discrete space $\Omega_n = \{0, 1\}^n$, and ϕ_n the $R_n - X$ cocycle

defined in the same way as ϕ. We have $C^*(R_n \underset{\phi_n}{\times} X) \cong M_{2^n}(C_0(X))$, and the corresponding embedding of A_n into A_{n+1}, when identified as an embedding of $M_{2^n}(C_0(X))$ into $M_{2^{n+1}}(C_0(X))$, is easily seen to be θ_ρ. The algebra A_2 of Example 1.6 can thus be described as the C^*-algebra of a relation; A_3 can be described as the C^*-algebra of the product relation $R \underset{\phi}{\times} (R \underset{\phi}{\times} X)$.

Example 2.5. Using a technique of Raeburn and Taylor [20] for constructing a continuous trace algebra with prescribed Dixmier-Douady invariant, we describe the algebra A_1 of Example 1.4 as the C^*-algebra of a skew product relation with 2-cocycle (cf. [18]).

If T is a locally compact space, it is known that $H^3(T, \mathbb{Z})$ is canonically isomorphic to second cohomology with coefficients in the sheaf of germs of continuous circle-valued functions; the construction in [20] uses cocycles from the sheaf cohomology. Let ρ be the homeomorphism of $T = \mathbb{T}^3$ given in 1.5. Choose an open cover $\mathcal{U} = \{U_i : i \in I\}$ of T so that $\rho^n(U_i) \in \mathcal{U}$ for each n, i (write $U_{i'} = \rho(U_i)$), together with a nontrivial Čech 2-cocycle λ relative to $\mathcal{U}$ such that:

$$\lambda_{i'j'k'}(\rho(t)) = \lambda_{ijk}(t) \quad \text{for } t \in U_i \cap U_j \cap U_k.$$

This may be done by choosing a finite cover $\{V_1, \ldots, V_m\}$ such that $V_i \cap \rho^n(V_j)$ is contractible for all i, j, n; then a nontrivial 2-cocycle relative to $\{V_i\}$ may be extended to the cover consisting of all translates under ρ.

Put $Y = \coprod_i U_i$ (the disjoint union) and let $\psi : Y \to T$ be the "reinclusion" map (i.e. $t \in U_i \mapsto t \in T$). Let $R(\psi) = \{(x, y) \in Y \times Y : \psi(x) = \psi(y)\}$ (cf. [17]) and note that $R(\psi)$ is homeomorphic to $\coprod_{i,j} U_i \cap U_j$. The groupoid of composable pairs in $R(\psi)$ is homeomorphic to $\coprod_{i,j,k} U_i \cap U_j \cap U_k$, so the Čech 2-cocycle λ defines a 2-cocycle $\tilde{\lambda}$ on $R(\psi)$. The resulting C^*-algebra, $C^*(R(\psi), \tilde{\lambda})$, is continuous trace with Dixmier-Douady invariant $[\lambda]$ [20].

Let $\hat{\rho} \in \mathrm{Aut}(R(\psi))$ be such that $\hat{\rho}(U_i \cap U_j) = U_{i'} \cap U_{j'}$. Observe that $\hat{\rho}$ leaves the 2-cocycle $\tilde{\lambda}$ invariant; thus $\hat{\rho}$ induces an automorphism ρ of $C^*(R(\psi), \tilde{\lambda})$.

We define an $R - R(\psi)$ cocycle (R is again the CAR relation):

$$\phi(\alpha)\beta = \hat{\rho}^{m(\alpha)}(\beta) \quad \text{(where } m \text{ is the number cocycle).}$$

We define a 2-cocycle on the skew product by

$$\sigma((\alpha_1, \phi(\alpha_2)\beta_1), (\alpha_2, \beta_2)) = \tilde{\lambda}(\beta_1, \beta_2).$$

Then, as in 2.4, $C^*(R \underset{\phi}{\times} R(\psi), \sigma) \cong A_1$ (if the ρ of 1.5 is chosen to be the one constructed here).

3. The Bunce-Deddens Algebras

In this section, we will show that if A is one of the simple C^*-algebras constructed in [7], then $A \otimes \mathcal{K}$ has the (HP) property. The two constructions which will be relevant for our purposes are:

(1) $A \otimes \mathcal{K}$ is isomorphic to the crossed product $C_0(\mathbb{R}) \times_\tau G$, where G is a

dense subgroup of $\mathbb{Q}$ containing $\mathbb{Z}$, acting by translation on $\mathbb{R}$. A itself is $C(\mathbb{T}) \times_\tau H$, where H is a dense torsion subgroup of $\mathbb{T}$; in fact, $H = G/\mathbb{Z}$ [23].

(2) $A = \varinjlim A_n$, where $A_n = M_{k_n}(C(\mathbb{T}))$, and the embedding of A_n into A_{n+1} is an "r_n-times-around" embedding. We may take $k_1 = 1$.

Viewed in either way, the canonical copy of $C(\mathbb{T})$ in A is a diagonal, and it is shown in [16; §8, 13] that if $g \in C(\mathbb{T})$, then A_g has an approximate identity of projections.

Our approach will be to use this fact, along with 1.8(b), to show that $A \otimes \mathscr{K}$ has (HP).

First note that if n is fixed, then $A \cong M_{k_n} \otimes B$ for some Bunce-Deddens algebra B, with $A_n \cong M_{k_n} \otimes C(\mathbb{T})$, where $C(\mathbb{T})$ is a diagonal in B as above. Thus, if g is any element of A_n which is diagonal under this decomposition, then A_g has an approximate identity of projections. (Such a g can be written as a sum of orthogonal elements each of which is of the form considered in [16].)

It is not quite true that every positive element g of $C(\mathbb{T}, M_k)$ is conjugate to a diagonal element. But a slightly weaker result holds, which is still strong enough to satisfy the hypotheses of 1.8:

Lemma 3.1. *Let $g \in C(\mathbb{T}, M_k)_+$, $\delta > 0$. Then there are elements $u_1, \ldots, u_k \in C(\mathbb{T}, M_k)$ such that:*

(1) $u_i^* u_i \perp u_j^* u_j$ *for* $i \neq j$;

(2) $\sum u_i^* u_i$ *is a unit for* $f_\delta(g)$;

(3) $u_i u_i^* \leqq e_{ii} \in C(\mathbb{T}, M_k)$, *where* $e_{ii} \in M_k$ *is the* ii*'th matrix unit, interpreted as a constant function.*

Proof. We construct the u_i inductively. Let $\mathscr{U}_1 = \{z \in \mathbb{T}: g(z) \neq 0\}$. For each $z \in \mathscr{U}_1$, choose a one-dimensional projection p_z in M_k orthogonal to the projection onto the null space of $g(z)$. Because $[g\,C(\mathbb{T}, M_k)\,g]^-$ is a continuous trace C^*-algebra, the corresponding continuous field satisfies Fell's condition [12; 10.5.7], and thus p_z can be continuously extended over a neighborhood of z to a one-dimensional projection orthogonal to the null space of g. Choosing a cover of $\mathscr{U}_1$ with such intervals so that only successive intervals overlap, the sections can be fitted together to get a global field $\{q_z: z \in \mathscr{U}_1\}$ of one-dimensional projections orthogonal to null(g). By [9; Lemma 7], we can then find a continuous field $\{v_z: z \in \mathscr{U}_1\}$ with $v_z^* v_z = q_z$, $v_z v_z^* = e_{11}$. Set $v_z = 0$ for $z \notin \mathscr{U}_1$. Then v is discontinuous, but $u_1 = v f_{\delta/2}(g)$ is continuous and has the right properties. (v is a multiplier of $C_0(\mathscr{U}_1, M_k)$.) Now replace g by

$$g_2 = (1 - v^* v)\, g (1 - v^* v) \in C(\mathbb{T}, M_k), \qquad \mathscr{U}_2 = \{z \in \mathbb{T}: g_2(z) = 0\}. \quad \square$$

Return to the case of $A \cong M_{k_n} \otimes B$, $A_n \cong M_{k_n} \otimes C(\mathbb{T})$. If g is an arbitrary positive element of A_n and $\delta > 0$, let $u_1, \ldots, u_{k_n}$ be as in Lemma 3.1. Set $u = \sum u_i$, and $h = \sum (u_i^* u_i) f_\delta(g) (u_i^* u_i)$. Then $A_h = A_{f_\delta(g)}$ (it is enough to check that $h(z)$ and $[f_\delta(g)](z)$ have the same null space for each $z \in \mathbb{T}$). Then $x \mapsto uxu^*$ is an isomorphism of A_h onto A_{uhu^*}, and uhu^* is diagonal; hence these hereditary algebras have approximate identities of projections. Therefore, by 1.8(b) A has (HP). The same argument shows that $M_r(A)$ has (HP) for all r, and so another application of 1.8(b) shows that $A \otimes \mathscr{K}$ has (HP) too.

This example shows that even a simple C^*-algebra with (HP) can contain a projectionless masa, in fact a projectionless diagonal. We conjecture that even UHF algebras contain projectionless masas, although it seems unlikely that such a masa could be a diagonal.

It would be of great interest to know whether the irrational rotation algebras also have (HP). The argument given here does not seem to work, since there is no known reasonable way of writing the irrational rotation algebras as inductive limits.

References

1. Blackadar, B.: A simple C^*-algebra with no nontrivial projections. Proc. Amer. Math. Soc. **78**, 504–508 (1980)
2. Blackadar, B.: A simple unital projectionless C^*-algebra. J. Operator Theory **5**, 63–71 (1981)
3. Blackadar, B.: A stable cancellation theorem for simple C^*-algebras. Proc. London Math. Soc. **47**, 303–305 (1983)
4. Blackadar, B.: Notes on the structure of projections in simple C^*-algebras. Semesterbericht Funktionalanalysis, Tübingen, W83
5. Blackadar, B.: Shape theory for C^*-algebras. To appear in Math. Scand.
6. Blackadar, B., Handelman, D.: Dimension functions and traces on C^*-algebras. J. Functional Anal. **45**, 297–340 (1982)
7. Bunce, J., Deddens, J.: A family of simple C^*-algebras related to weighted shift operators. J. Functional Analysis **19**, 13–24 (1975)
8. Connes, A.: An analogue of the Thom isomorphism for crossed products of a C^*-algebra by an action of $\mathbb{R}$. Advances in Math. **39**, 31–55 (1981)
9. Conway, J.: On the Calkin algebra and the covering homotopy property. Trans. Amer. Math. Soc. **211**, 135–142 (1975)
10. Cuntz, J.: The structure of addition and multiplication in simple C^*-algebras. Math. Scand. **40**, 215–233 (1977)
11. Cuntz, J.: The internal structure of simple C^*-algebras. Proc. Symp. Pure Math. **38** (ed. R.V. Kadison), AMS 1982
12. Dixmier, J.: C^*-Algebras. Amsterdam: North-Holland 1977
13. Dixmier, J., Douady, A.: Champs continus d'espaces Hilbertiens et de C^*-algèbres. Bull. Soc. Math. France **91**, 227–284 (1963)
14. Effros, E., Hahn, F.: Locally compact transformation groups and C^*-algebras. Mem. Amer. Math. Soc. No. **75** (1967)
15. Grothendieck, A.: Le group de Brauer, I: algèbres d'Azumaya et interpretations diverses. Séminaire Bourbaki 1964/65, Exposé 290
16. Kumjian, A.: On localizations and simple C^*-algebras. Pacific J. Math. **112**, 141–192 (1984)
17. Kumjian, A.: Preliminary algebras arising from local homeomorphisms. Math. Scand. **52**, 269–278 (1983)
18. Kumjian, A.: On C^*-diagonals and twisted relations. Semesterbericht Funktionalanalysis, Tübingen, W83. Complete version in preparation
19. Pedersen, G.: The linear span of projections in simple C^*-algebras. J. Operator Theory **4**, 289–296 (1980)
20. Raeburn, I., Taylor, J.: Continuous trace C^*-algebras with given Dixmier-Douady class. To appear in J. Australian Math. Soc.
21. Renault, J.: A groupoid approach to C^*-algebras. Lecture Notes in Math. **793**. Berlin Heidelberg New York: Springer 1980
22. Masuda, T.: Groupoid dynamical systems and crossed products II – the case of C^*-systems. To appear
23. Ghatage, P.: C^*-algebras generated by weighted shifts. Indiana Univ. Math. J. **30**, 539–545 (1981)

Received January 1, 1984; received in final form May 17, 1984

Math. Z. 189, 65–70 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

A Note on Continuous Cohomology for Semisimple Lie Groups

David H. Collingwood

Dept. of Mathematics, Rutgers University, New Brunswick, New Jersey 08903, USA

Arguing by "asymptotic induction", we compute the continuous cohomology of any irreducible admissible representation (not necessarily unitary), in the case of a real rank one matrix group. This completes the program in [2, VI. 3]. In the general case, we remark on Zuckerman's approach to the problem, using Vogan's work on the Kazhdan-Lusztig conjectures.

1. Asymptotic Induction

Let G be a connected semisimple real matrix group, K a maximal compact subgroup and V the space of K-finite vectors of an irreducible admissible representation of G. Denote real Lie algebras by $\mathfrak{g}_0, \mathfrak{k}_0$, etc. and the corresponding complexifications without subscript "o". We are interested in the *relative* $(\mathfrak{g}, K)$*-cohomology* with coefficients in V

$$H^*(\mathfrak{g}, K, V). \tag{1.1}$$

The zero cohomology group consists of $\mathfrak{g}$-invariant vectors of V and the higher cohomology may be obtained using left derived functors in an appropriate "relative category"; cf. [2]. The problem we consider, and completely solve in the real rank one case, is the explicit computation of (1.1).

By a well-known vanishing theorem, we may as well assume V has trivial infinitesimal character. In the monograph [2], Borel and Wallach outline a program for computing (1.1), which we now recall. For simplicity, assume $rkG = rkK$; i.e., G contains a discrete series. Enumerate the discrete series, having trivial infinitesimal character, as $D_1, D_2, \ldots, D_r$. Next, let (P, σ, ν) be a collection of *trivial Langlands data:* (i) $P = MAN$ is the Langlands decomposition of a proper cuspidal parabolic subgroup of G; (ii) σ is a discrete series representation for M; (iii) $\nu \in \mathfrak{a}^*$ (algebraic dual) satisfies $\mathrm{Re}(\nu|\alpha) \geqq 0$ for the positive $\mathfrak{a}$-roots determined by P; and (iv) the (normalized) generalized principal series representation $\pi_P^G(\sigma \otimes e^{\nu} \otimes 1)$ has trivial infinitesimal character. We will denote

the K-finite vectors of $\pi_P^G(\sigma \otimes e^{\nu} \otimes 1)$ by $I(P, \sigma, \nu)$. Let

$$\hat{G}_\rho = \left\{ \begin{array}{l} \text{equivalence classes of irreducible admissible} \\ (\mathfrak{g}, K) \text{ modules with trivial infinitesimal character} \end{array} \right\};$$

cf. [2] for appropriate definitions. According to *Langlands classification* [2, IV]:

$$\hat{G}_\rho = \{D_1, D_2, \ldots, D_r\} \cup \{J(P, \sigma, \nu)\}_{(P, \sigma, \nu)}, \tag{1.2}$$

where (P, σ, ν) runs over all equivalence classes of trivial Langlands data and $J(P, \sigma, \nu)$ denotes *the* unique irreducible quotient of $I(P, \sigma, \nu)$.

We have an exact sequence

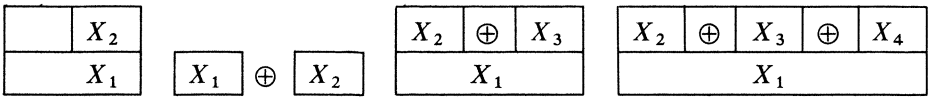
$$0 \to U(P, \sigma, \nu) \to I(P, \sigma, \nu) \to J(P, \sigma, \nu) \to 0. \tag{1.3}$$

A further consequence of Langlands classification is the fact: all irreducible subquotients of $U(P, \sigma, \nu)$ have asymptotic behavior which is *strictly* better than that of $J(P, \sigma, \nu)$; this can be made precise in terms of roots, etc. By [2, III], $H^*(\mathfrak{g}, K, I(P, \sigma, \nu))$ lives in a "small" explicitly computable interval. Elementary K-type considerations [2, II] lead to $H^*(\mathfrak{g}, K, D_i)$, $1 \leqq i \leqq r$. Thus, in view of (1.2), our problem is to compute $H^*(\mathfrak{g}, K, J(P, \sigma, \nu))$. Apply $H^*(\mathfrak{g}, K, -)$ to (1.3), obtaining the usual long exact sequence of cohomology. The above remarks, and on induction argument on the asymptotic behavior of irreducible constituents of $I(P, \sigma, \nu)$, leads to information on $H(\mathfrak{g}, K, J(P, \sigma, \nu))$. Our philosophy is: this algorithm should actually compute the cohomology of $J(P, \sigma, \nu)$.

2. The Case of Real Rank One

A major obstruction to the success of the algorithm in § 1, is knowing the "structure" of $U(P, \sigma, \nu)$. Of course, using the Kazhdan-Lusztig conjectures (which are theorems in our context [8]), we know the structure of $U(P, \sigma, \nu)$ in the Grothendieck group. For further connections with the Kazhdan-Lusztig conjectures, see § 6.

We now assume, for the remainder of this section, G has real rank one. In [3], the explicit structure of $U(P, \sigma, \nu)$ is determined. More precisely, either $U(P, \sigma, \nu)$ is irreducible, or decomposes as follows:



(a) (b) (c) (d)

Here, our notation means the following: (i) the X_i are the irreducible constituents of $U(P, \sigma, \nu)$; (ii) X_1 (resp. $X_1 \oplus X_2$) is the largest semisimple submodule of $U(P, \sigma, \nu)$ in cases (a), (c), (d) (resp. (b)); and (iii) the irreducible constituents of $U(P, \sigma, \nu)/X_1$ have trivial extentions with one another. This is all made precise in [3], keeping track of the various $J(P, \sigma, \nu)$'s and D_i's. The point is

(2.1) **Theorem.** *In the real rank one case, this information is sufficient for the algorithm of Sect.* 1 *to explicitly compute* $H^*(\mathfrak{g}, K, J(P, \sigma, \nu))$.

The verification of (2.1) is an extremely straightforward verification, though somewhat tedious. We leave the details to the reader, who is urged to first consider the cases of $\mathfrak{g}_0 = \mathfrak{so}(n,1)$, $\mathfrak{su}(n,1)$; these cases are worked out in [2, VI.4]. In this paper, we state the new cases of $\mathfrak{g}_0 = \mathfrak{sp}(n,1)$, $\mathfrak{f}_{4(-20)}$.

3. The Case of $Sp(n,1)$

Fix $G = Sp(n,1)$, $P = MAN$ a minimal parabolic subgroup of G, $H = TA$ a (compatible) maximally split θ-stable Cartan subgroup of G; here, θ = Cartan involution. Form the roots of $\mathfrak{h}$ in $\mathfrak{g}$, denoted Φ; the restricted roots determined by P, denoted Σ; the roots of $\mathfrak{h}$ in $\mathfrak{m} \oplus \mathfrak{a}$, denoted $\Phi_{\mathfrak{m}}$. Choose positive systems Φ^+, Σ^+ satisfying the compatibility: $\alpha \in \Phi^+$, $\alpha|_{\mathfrak{a}} \neq 0 \Rightarrow \alpha|_{\mathfrak{a}} \in \Sigma^+$. Define W and $W_{\mathfrak{m}}$ to be the Weyl groups of Φ and $\Phi_{\mathfrak{m}}$, respectively. The irreducible finite dimensional representations of $\mathfrak{m} \oplus \mathfrak{a}$ are parametrized by $W^{\mathfrak{m}} \overset{\text{def}}{=} W/W_{\mathfrak{m}}$. We index the elements of $W^{\mathfrak{m}}_+ = \{\omega_{ij}\}_{(i,j)}$ as in [3, 4], where (i,j) runs over the index set

$$\{(i,j): 0 \leqq i \leqq n-1,\ 0 \leqq i < j \leqq 2n-i\}.$$

More precisely, if ρ is the half-sum of Φ^+, then the $(P, \omega_{ij}\rho)$ correspond to the various sets of Langlands data appearing in (1.2). For notational simplicity, we write J_{ij} for the attached Langlands representation. This indexing scheme has the following properties:

(3.1) (a) J_{01} = trivial representation;

(b) ω_{ij} has length $i+j-1$, with respect to Φ^+;

(c) Suppose $i+j < k+l$ (i.e. ω_{ij} and ω_{kl} lie in different "rows" of $W^{\mathfrak{m}}$, when filtered by the Bruhat order), then the Langlands data $\mathfrak{a}_0$-parameter of J_{ij} has a projection onto $\mathfrak{a}_0$ which is greater than or equal to the Langlands data $\mathfrak{a}_0$-parameter of J_{kl}.

In the sense of (3.1c), we have a nice filtration on $\hat{G}_\rho$, by asymptotic growth. It is this filtration upon which one inducts to compute $H^*(\mathfrak{g}, K, V)$. If $k+l = 2n$, then J_{kl} is *minimal* in the sense of [2, VI.1], where $H^*(\mathfrak{g}, K, J_{kl})$ was explicitly computed. (Any such J_{kl} is a "first wall crossing" of some discrete series D_i.)

(3.2) **Theorem.** *Adhering to the above notation,*

(i) $$H^q(\mathfrak{g}, K, D_i) = \begin{cases} \mathbb{C}, & q = 2n \\ 0, & \text{otherwise.} \end{cases}$$

(ii) *If* $1 \leqq i \leqq n$,

$$H^q(\mathfrak{g}, K, J_{n-i, n-i+1}) = \begin{cases} \mathbb{C}, & q = 2(n-i)+4l, \quad 0 \leqq l \leqq i \\ 0, & \text{otherwise.} \end{cases}$$

(iii) *In all remaining cases,*

$$H^q(\mathfrak{g}, K, J_{ij}) = \begin{cases} \mathbb{C}, & q=(i+j)-1+2l, \quad 0 \leqq l \leqq 2n-(i+j)+1 \\ 0, & \text{otherwise.} \end{cases}$$

4. The Case of F_4

Consider the exceptional real rank one form $\mathfrak{f}_{4(-20)}$ and the corresponding group F_4. $\hat{G}_\rho$ is parametrized as follows [3, 4]:

(4.1) (a) D_1, D_2, D_3, discrete series;

(b) $J_{01}, J_{02}, J_{03}, J_{04}, J_{05}, J_{14}, J_{24}, J_{15}$
$J_{25}, J_{16}, J_{26}, J_{17}$, non-tempered.

The indexing satisfies the same sort of rules as in § 3. We have summarized the non-zero cohomology in Table 1.

Table 1. $H^*(\mathfrak{g}, K, -)$ for $\mathfrak{g}_0 = \mathfrak{f}_{4(-20)}$

$\hat{G}_\rho$	q for $H^q(\mathfrak{g}, K, -)$																	Behavior
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
D_1									$\mathbb{C}$									discrete series
D_2									$\mathbb{C}$									discrete series
D_3									$\mathbb{C}$									discrete series
J_{17}								$\mathbb{C}$		$\mathbb{C}$								unitary
J_{26}								$\mathbb{C}$		$\mathbb{C}$								unitary
J_{16}							$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$							non-unitary
J_{25}							$\mathbb{C}$				$\mathbb{C}$							unitary
J_{15}						$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$						non-unitary
J_{24}						$\mathbb{C}$						$\mathbb{C}$						unitary
J_{05}					$\mathbb{C}$				$\mathbb{C}$				$\mathbb{C}$					isolated unitary
J_{14}					$\mathbb{C}$		$\mathbb{C}$				$\mathbb{C}$		$\mathbb{C}$					non-unitary
J_{04}				$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$				non-unitary
J_{03}			$\mathbb{C}$				$\mathbb{C}$		$\mathbb{C}$		$\mathbb{C}$				$\mathbb{C}$			non-unitary
J_{02}		$\mathbb{C}$						$\mathbb{C}$		$\mathbb{C}$						$\mathbb{C}$		non-unitary
J_{01}	$\mathbb{C}$								$\mathbb{C}$								$\mathbb{C}$	isolated unitary

5. Remarks

Assume the hypothesis of § 2. We make three remarks.

(i) Let F be an irreducible finite dimensional representation of G. Form the *twisted* cohomology groups $H^*(\mathfrak{g}, K, V \otimes F)$. These groups are non-zero, only if the infinitesimal character of V coincides with that of F^*. The results of § 3, 4 easily translate into this context.

(ii) Consideration of continuous cohomology with coefficients in a *non-unitary* module, arises in Harder's work on Eisenstein cohomology.

(iii) Suppose V is an irreducible unitary representation of G and $H^*(\mathfrak{g}, K, V) \neq 0$. By the work of Vogan and Zuckerman [9], we conclude V is isomorphic to some A_q. (We shall not explicitly define A_q, but only remark that $q = \ell \oplus u$ is a θ-stable parabolic subalgebra of $\mathfrak{g}$ and A_q is "cohomologically induced" from the trivial representation of ℓ, up to $\mathfrak{g}$; cf. [9].) Furthermore, it is shown

$$H^*(\mathfrak{g}, K, V) = H^{*'}(\ell, L \cap K, \mathbb{C}); \tag{5.1}$$

$*'$ being a shifted index set. In other words, the cohomology of V looks like that of the trivial representation for some smaller group. If $\mathfrak{g}_0 = \mathfrak{so}(n, 1)$ or $\mathfrak{su}(n, 1)$, then all representations in $\hat{G}_\rho$ are unitary and can thus be inductively computed using (5.1). For $\mathfrak{g}_0 = \mathfrak{sp}(n, 1)$ or $\mathfrak{f}_{4(-20)}$, $\hat{G}_\rho$ contains non-unitary representations. Indeed, using the results of Silva [5] and Barbasch-Silva [1], one can explicitly determine which J_{ij} are unitary. These unitary parameters are indicated for F_4 in Table 1. The structures of $[\ell_0, \ell_0]$ for the q attached to non-tempered unitary Aq are as follows: (a) J_{17} or J_{26}, $[\ell_0, \ell_0] = \mathfrak{sl}(2, \mathbb{R})$; (b) J_{25}, $[\ell_0, \ell_0] = \mathfrak{so}(4, 1)$; (c) J_{24}, $[\ell_0, \ell_0] = \mathfrak{so}(6, 1)$; (d) J_{05}, $[\ell_0, \ell_0] = \mathfrak{sp}(2, 1)$; and (e) J_{01}, $[\ell_0, \ell_0] = \mathfrak{g}$.

6. The General Problem

Assume the hypothesis of § 1; in particular, $rkG = rkK$. Recall, the key step in our proof of (2.1) involved explicit knowledge of the socle filtration for $U(P, \sigma, \nu)$. As discussed in [3], these results, in turn, are a consequence of Vogan's work on the Kazhdan-Lusztig conjectures. In short, Theorem (2.1) is a consequence of these conjectures (which are theorems, as we previously noted). The referee has pointed out that (when $rkG = rkK$), the Kazhdan-Lusztig conjectures actually solve the general problem; up to combinatorics involving the Kazhdan-Lusztig polynomials of [8]. Precisely, we have

(6.1) **Theorem** (Zuckerman). *Assume $rkG = rkK$ and let V be an irreducible Harish-Chandra module with non-trivial $(\mathfrak{g}, K)$-cohomology. Then the dimension of $H^i(\mathfrak{g}, K, V)$ is given as a sum of certain coefficients of the Kazhdan-Lusztig polynomials.*

We will only sketch the proof of (6.1), which is an unpublished result of Zuckerman. (The following remarks are a consequence of the author's discussion with Gregg Zuckerman.) We assume full familiarity with Vogan's work on computing characters, the U_α-construction, Kazhdan-Lusztig polynomials, etc.; cf. [6–8]. The idea is to start with the known cohomology of the discrete series representations and compute the cohomology of any other representation $X \in \hat{G}_\rho$, by induction. The vehicle for such an induction argument is the U_α-construction. The key ingredients are:

(i) the long exact sequences arising when $H^*(\mathfrak{g}, K, -)$ is applied to [7, (2.4)];

(ii) the "Ext" interpretation of $(\mathfrak{g}, K)$-cohomology [2];

(iii) the fact $H^*(\mathfrak{g}, K, \varphi_\alpha \psi_\alpha X) = 0$, (a consequence of (ii) and adjointness of translation functors).

Built into the induction, one simultaneously proves the following: If $\mathscr{E} = \{i : H^i(\mathfrak{g}, K, X) \neq 0\}$, then all elements of $\mathscr{E}$ are congruent, modulo 2. This is the crucial use of our equal rank hypothesis.

As a final note, when $rkG \neq rkK$, it is an open problem to compute $H^*(\mathfrak{g}, K, X)$, for a general irreducible X.

Acknowledgements. The author wishes to acknowledge several people for their helpful comments regarding this paper. Nolan Wallach suggested that [3] should lead to the results herein. Dan Barbasch's comments on unitary representations of F_4 were appreciated. The author is grateful to the referee for pointing out Zuckerman's result in §6 and also wishes to thank Gregg Zuckerman for ellaborating on (6.1). This work was carried out during the author's stay at the Institute for Advanced Study and he wishes to acknowledge gratefully their support.

References

1. Barbasch, D., Silva, W.: The unitary spectrum for real rank one groups. Invent Math. **72**, 27–55 (1983)
2. Borel, A., Wallach, N.: Continuous cohomology, discrete subgroups and representations of reductive groups. Princeton University Press, Annals of Math. Study **94**, 1980
3. Collingwood, D.: Embeddings of Harish-Chandra modules, π-homology and the composition series problem: the case of real rank one. Trans. Amer. Math. Soc. **285**, 565–579 (1984)
4. Silva, W.: The embeddings of the discrete series in the principal series for semisimple groups of real rank one. Trans. Amer. Math. Soc. **261**, 303–368 (1980)
5. Silva, W.: The unitary dual of $Sp(n, 1)$ $(n \geqq 2)$. Duke Math. J. **48**, 549–583 (1981)
6. Vogan, D.: Irreducible characters of semisimple Lie groups I. Duke Math. J. **46**, 61–108 (1979)
7. Vogan, D.: Irreducible characters of semisimple Lie groups II: the Kazhdan-Lusztig conjectures. Duke Math. J. **46**, 805–859 (1979)
8. Vogan, D.: Irreducible characters of semisimple Lie groups III: Proof of the Kazhdan-Lusztig conjectures in the integral case. Invent. Math. **71**, 381–417 (1983)
9. Vogan, D., Zuckerman, G.: Unitary representations with non-zero cohomology. Compositio Math. **53**, 51-90 (1984)

Received December 13, 1983

Math. Z. 189, 71–79 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

C^2-Regularity for Partially Free Minimal Surfaces*

Gerhard Dziuk

Institut für Mathematik, Rheinisch-Westfälische Technische Hochschule, Templergraben 55, D-5100 Aachen, Federal Republic of Germany

1. Introduction

Consider a configuration in Euclidean N-space consisting of a smooth Jordan arc C having its end points P_1 and P_2 on a smooth hypersurface S, but no other points in common with S. Let $B := \{(u,v) | u^2+v^2<1\}$ and denote by $\partial^+ B$ ($\partial^- B$) its boundary portions in $v>0$ ($v<0$). Let $Z(C,S)$ be the set of all surfaces $x = x(u,v) = (x_1(u,v), \ldots, x_N(u,v)) \in C^0(B) \cap H^{1,2}(B)$ which are bounded by C and S in the following sense: x maps $\partial^+ B$ continuously and in weakly monotonic manner onto C such that $x(-1,0)=P_1$, $x(1,0)=P_2$ and $x(0,1)=P_3$ for some fixed third point on C, different from P_1 and P_2. The free boundary condition is expressed as

$$\lim_{(u,v)\to(u_0,v_0)} \inf_{y\in S} |x(u,v)-y| = 0$$

for $(u_0,v_0) \in \partial^- B$. This implies the continuity of the distance function $\operatorname{dist}(x,S)$ but not the continuity of x itself. It is well known (see VI in [2]) that the variational problem

$$\int_B |y_u|^2 + |y_v|^2 \, du\, dv \to \text{Min} \quad (y \in Z(C,S)) \tag{1}$$

has at least one solution $x \in Z(C,S)$ which is a minimal surface, i.e. it solves the system

$$\Delta x = 0, \qquad x_u \cdot x_v = 0, \qquad |x_u| = |x_v| \tag{2}$$

on B. Note that in general a minimal surface does not represent a minimum of the corresponding variational problem.

There is satisfactory information concerning the behaviour of a solution x of (1) at the "fixed" boundary portion; see [13], pp. 281–325. Roughly spoken the result is that, up to $\partial^+ B$, x is as smooth as the curve C itself is. The crucial point in proving regularity at the free boundary portion $\partial^- B$ is to show continuity of x on $B \cup \partial^- B$. For solutions of (1), i.e. for area minimizing

* Die Arbeit ist Teil der Habilitationsschrift des Autors

minimal surfaces, regularity at the free boundary has been proved by J.C.C. Nitsche in [12], K.H. Goldhorn and S. Hildebrandt in [6] and by W. Jäger in [10]. These results are to the effect that $x \in C^{k,\mu}(B \cup \partial^- B)$ if the supporting surface S is of that class and is admissable in some sense. For $k \geqq 3$ this has been proved by W. Jäger in [10]. The case $S \in C^{2,\mu}$ has been investigated by J.C.C. Nitsche in [14]. M. Grüter, S. Hildebrandt and J.C.C. Nitsche in [8] and the author in [4] proved regularity of the free boundary of stationary points x of the variational problem (1) under slightly different assumptions. In both papers the supporting surface S was assumed to be of class C^3.

The aim of this paper is to manage the cases $S \in C^{2,\mu}$ and $S \in C^2(C^{1,1})$ which means that we prove: $x \in C^{2,\mu}(B \cup \partial^- B)$ in the first case and $x \in C^{1,\mu}(B \cup \partial^- B)$ for every $\mu \in (0,1)$ in the second case. The proof consists in reflecting a given stationary solution x at the supporting surface S. W. Jäger proved in [10] weak transversality for such stationary solutions and this we use to prove that the reflected solution solves a quasilinear elliptic system of second order. In general there is no continuity for solutions of such systems. But using the conformality relations in (2) we may apply a method developed by M. Grüter in [7] to prove regularity.

2. Results

First of all we have to specify what an admissable supporting surface S is. We adopt W. Jäger's notions since they are easy to verify. For example every compact hypersurface S in $\mathbb{R}^N$ given by $f(x)=0$, $f \in C^{m,\mu}(\mathbb{R}^N, \mathbb{R})$, $\nabla f \neq 0$ on S, represents such an admissable supporting surface.

Definition 1. Let S be an $(N-1)$-dimensional $C^{m,\mu}$-manifold in $\mathbb{R}^N$ with the following properties: For every point $x_0 \in S$ there is some neighbourhood U_0 of x_0 in $\mathbb{R}^N$ and some real-valued function $f_0 \in C^{m,\mu}(U_0)$ $(m \geqq 2,\ 0 \leqq \mu \leqq 1)$ with $\nabla f_0 \neq 0$ in U_0 and $f(x)=0$ iff $x \in S \cap U_0$. There is some positive number d and there are functions ξ, a, $n \in C^{m-1,\mu}(U_d)$ in the strip $U_d = \{x \in \mathbb{R}^N \mid \operatorname{dist}(x,S) < d\}$ such that every $x \in U_d$ can be written as

$$x = a(x) + \xi(x)\, n(x)$$

where $a(x) \in S$, $n(x)$ is normal to S in $a(x)$ $(|n(x)|=1)$ and $|\xi(x)| = \operatorname{dist}(x,S)$. In addition to that we assume:

$$\sup_{U_d} |\nabla n| \leqq C_1.$$

Under these conditions we call S an admissable supporting surface of class $C^{m,\mu}$.

Let us look now at the basic variational problem (1) presented in the introduction. Since we are concerned with the free boundary portion only we may localize the situation by cutting out a neighbourhood of a point in $\partial^- B$ and we may achieve the following situation by conformal mapping. At this point it is obvious that our restriction to the special boundary configuration (C, S) is only for the sake of simplicity.

Definition 2. Let $x \in C^0(\overline{Q^+} \setminus \bar{I})$ be a solution of (2) on $Q^+ = \{(u,v) | |u| < 1, 0 < v < 1\}$ where $I = (-1,1) \times \{0\}$. We call $y \in C^0(\overline{Q^+} \setminus \bar{I}) \cap H^{1,2}(Q^+)$ admissable relative to x, if

$$\lim_{(u,v) \to (u_0,0)} \operatorname{dist}(y(u,v), S) = 0 \tag{3}$$

for every $u_0 \in [-1,1]$ and if $y = x$ on $\partial Q^+ \setminus \bar{I}$. x is called stationary if

$$\frac{d}{d\varepsilon} \int_{Q^+} |y_u^{(\varepsilon)}|^2 + |y_v^{(\varepsilon)}|^2 \, du\, dv \Big|_{\varepsilon = 0} = 0 \tag{4}$$

exists for admissable $y^{(\varepsilon)} = x + \varepsilon z^{(\varepsilon)}$ ($|\varepsilon| < \varepsilon_0$).

We now are able to formulate our results:

Theorem. *Let S be an admissable supporting surface of class $C^{m,\mu}$ ($m \in \mathbb{N}$, $m \geqq 2$, $0 \leqq \mu \leqq 1$) and let x be a stationary minimal surface on Q^+. Then*

$$x \in C^{m,\mu}(Q^+ \cup I), \quad \text{if } 0 < \mu < 1,$$

$$x \in C^{m-1,\nu}(Q^+ \cup I) \quad \text{for every } \nu \in (0,1), \text{ if } \mu = 0$$

and

$$x \in C^{m,\nu}(Q^+ \cup I) \quad \text{for every } \nu \in (0,1), \text{ if } \mu = 1.$$

As the reader will see, the classes $C^{m-1,1}$ for S generate the same regularity properties for x as $S \in C^{m,0}$ does. From now on we will assume the assumptions of the theorem to be satisfied. Because of (3) we find some $\delta > 0$ such that $x(\bar{Q}_\delta^+ \setminus I) \subset U_d$, where $Q_\delta^+ = \{(u,v) | |u| < 1, 0 < v < \delta\}$.

3. Preliminary Lemmata

In [10], Lemma 1 W. Jäger proves weak transversality of stationary free minimal surfaces on the free boundary:

1. Lemma. *For every $\Phi \in C_0^1(Q^+ \cup I)$*

$$\lim_{\substack{v \to 0 \\ v > 0}} \int_{-1}^{1} (x_v - n(x) \cdot x_v n(x)) \cdot \Phi \, du = 0. \tag{5}$$

Let us reflect the solution surface x at the supporting surface S

$$y(u,v) = \begin{cases} x(u,v) & \text{in } (u,v) \in \bar{Q}_\delta^+ \\ x(u,-v) - 2\xi(x(u,-v))\, n(x(u,-v)) & \text{if } (u,-v) \in Q_\delta^+ \end{cases} \tag{6}$$

and claim that y solves some quasilinear elliptic system on $Q_\delta = \{(u,v) | |u| < 1, |v| < \delta\}$. We introduce some abbreviations related to the reflection mapping $y = F(x) = x - 2\xi(x)\, n(x)$. Define the matrices H, B and T by $H_{ik} = \xi_{x_i x_k}$, $T_{ik} = \delta_{ik} - 2n_i n_k$ ($i, k = 1, \ldots, N$) and $B = (I - 2\xi H)^{-1}$. Then the following relations are easily verified: $Hn = 0$, $HT = TH = H$, $\dfrac{\partial F}{\partial x} = TB^{-1}$, $\det \dfrac{\partial F}{\partial x} = -1 + O(d)$ for small d, $F = F^{-1} \in C^{m-1,\mu}(U_d)$, $\xi \circ F = -\xi$, $n \circ F = n$, $H \circ F = BH$, $B \circ F = B^{-1}$.

2. Lemma. *There is some $\delta>0$ such that y is in $H^{1,2}(Q_\delta)\cap C^1(Q_\delta\setminus I)$ and for every test function $\Phi\in\mathring{H}^{1,2}(Q_\delta)\cap L^\infty(Q_\delta)$ we have:*

$$\int_{Q_\delta} Ay_u\cdot\Phi_u+Ay_v\cdot\Phi_v\,du\,dv=\int_{Q_\delta} f\cdot\Phi\,du\,dv \tag{7}$$

and

$$Ay_u\cdot Ay_v=0,\qquad |Ay_u|=|Ay_v| \tag{8}$$

on $Q_\delta\setminus I$, where we have set $A=I$, $f=0$ on $\bar{Q}_\delta^+$ and $A=B(y)^{-1}$,

$$\begin{aligned} f=&2(H(y)y_u\cdot B(y)^{-1}y_u+H(y)\,y_v\cdot B(y)^{-1}y_v)\,n(y)\\ &-2(n(y)\cdot y_u H(y)y_u+n(y)\cdot y_v H(y)y_v)\end{aligned}$$

on $Q_\delta^-=\{(u,v)\,|\,|u|<1,\ -\delta<v<0\}$.

Proof: Since $\xi\circ x=0$ on I continuously we may chose $\delta>0$ small enough for our calculations. For $\Phi\in C_0^1(Q_\delta)$ and small $\varepsilon>0$ we derive from the harmonicity of $x=y$ on Q^+:

$$\int_{Q_\delta^+\setminus Q_\varepsilon^+} y_u\cdot\Phi_u+y_v\cdot\Phi_v\,du\,dv=-\int_{-1}^{1} x_v\cdot\Phi\,du|_{v=\varepsilon}. \tag{9}$$

On the lower half plane the situation is more complicated. Let us agree for the following lines to take x, x_u and x_v in $(u,-v)$. On Q_δ^-:

$$y_u=T(x)\,B(x)^{-1}x_u,\qquad y_v=-T(x)\,B(x)^{-1}x_v.$$

From this and $x_{uu}+x_{vv}=0$ we infer:

$$\begin{aligned}\int_{Q_\delta^-\setminus Q_\varepsilon^-} y_u\cdot\Phi_u+y_v\cdot\Phi_v\,du\,dv=&2\int_{Q_\delta^-\setminus Q_\varepsilon^-}((H(x)x_u\cdot x_u+H(x)x_v\cdot x_v)\,n(x)\\ &+n(x)\cdot x_u H(x)x_u+n(x)\cdot x_v H(x)x_v)\cdot\Phi-\xi(x)\,H(x)x_u\cdot\Phi_u\\ &+\xi(x)\,H(x)x_v\cdot\Phi_v\,du\,dv-\int_{-1}^{1}(x_v-2n(x)\cdot x_v n(x))(u,\varepsilon)\cdot\Phi(u,-\varepsilon)\,du.\end{aligned}$$

We combine this with (9) and get:

$$\begin{aligned}\int_{Q_\delta\setminus Q_\varepsilon} Ay_u\cdot\Phi_u+Ay_v\cdot\Phi_v\,du\,dv=&\int_{Q_\delta\setminus Q_\varepsilon} f\cdot\Phi\,du\,dv\\ &-\int_{-1}^{1} x_v(u,\varepsilon)\cdot(\Phi(u,\varepsilon)-\Phi(u,-\varepsilon))\,du\\ &-2\int_{-1}^{1}(x_v-n(x)\cdot x_v\,n(x))(u,\varepsilon)\cdot\Phi(u,-\varepsilon)\,du.\end{aligned}$$

Transversality (5) completes the proof for $\varepsilon\to 0$.

M. Grüter proved in [7], (2.5) Proposition a fundamental inequality for surfaces given in conformal parameters. In our situation y is a smooth function on $Q_\delta\setminus I$ closely related to x. Here this inequality is derived easily by using that dx/dw is holomorphic and $\sum_{k=1}^{3}(dx_k/dw)^2\equiv 0$ on Q^+, if x is not a constant.

3. Lemma. *Assume that $x(\overline{Q_\delta^+}\setminus I)\subset U_{d/2}$. Then for every ball $B(w_1,r)=\{w=(u,v)\,|\,|(u,v)-(u_1,0)|<r\}\Subset Q_\delta$ $(w_1=(u_1,0)\in I)$ and for every w_0*

$=(u_0, v_0)\in B(w_1, r)\setminus I$ *we have:*

$$\limsup_{\rho\to 0,\rho>0} \rho^{-2} \int\limits_{B(w_1,r)\cap\{w\,|\,|y(w)-y(w_0)|<\rho\}} |y_u|^2+|y_v|^2\,du\,dv \geqq \left(\frac{1-2dC_1}{1+2dC_1}\right)^2 2\pi. \quad (10)$$

Now we are going to prove the main Lemma of this paper. It will enable us to prove continuity of y in $w_1\in I$. The most important step is to control the integral in (10) from above uniformly in w_0.

4. Lemma. *Let* w_1, $B(w_1, r)$ *and* w_0 *be as in* 3. *Lemma and assume that* $x(\overline{Q_\delta^+}\setminus I)\subset U_{d/6}$. *Then there is some* $R_0>0$ *such that for every* $R\in[0, R_0]$ *from*

$$\inf_{\partial B(w_1,r)} |y-y(w_0)|>R \quad (11)$$

it follows that

$$R\leqq C_2\Big(\int\limits_{B(w_1,r)} |y_u|^2+|y_v|^2\,du\,dv\Big)^{1/2}.$$

The constants R_0 *and* C_2 *depend on* d *and* C_1 *only.*

Proof. Chose $\lambda\in C^1(\mathbb{R})$ such that $0\leqq\lambda\leqq 1$, $\lambda'\geqq 0$, $\lambda=0$ on $(-\infty, 0]$, $\lambda=1$ on $[\varepsilon, \infty)$, $\lambda>0$ on $(0, \infty)$ where ε is some positive number to be specified later. Let us define

$$z=\begin{cases} y-y(w_0) & \text{on } \overline{Q_\delta^+} \\ y-y(w_0)-2\xi(y)(n(y)-n(y(w_0))) & \text{on } Q_\delta^- \end{cases}$$

and introduce the test function

$$\Phi=\begin{cases} z\lambda(\rho-|z|) & \text{on } B(w_1, r) \\ 0 & \text{elsewhere,} \end{cases}$$

where $\rho\in(0, (1-dC_1/3)R)$. Φ *is a test function for* (7) *since* $\Phi=0$ for $|z|\geqq\rho$ and $|z|\geqq|y-y(w_0)|-d/3|n(y)-n(y(w_0))|\geqq(1-dC_1/3)|y-y(w_0)|$ for $|z|<\rho$, so that (11) gives us $\Phi=0$ on $\partial B(w_1, r)$. Thus we may insert Φ into (7) and estimate the integrands separately on Q_δ^+ and Q_δ^-. On $Q_\delta^+\cap B(w_1, r)$ we have:

$$\begin{aligned} A\,y_u\cdot\Phi_u+A\,y_v\cdot\Phi_v &= (|y_u|^2+|y_v|^2)\,\lambda(\rho-|z|)-\Big(\Big(\frac{z}{|z|}\cdot y_u\Big)^2+\Big(\frac{z}{|z|}\cdot y_v\Big)^2\Big)|z|\,\lambda'(\rho-|z|) \\ &\geqq (|y_u|^2+|y_v|^2)(\lambda(\rho-|z|)-\tfrac{1}{2}\rho\lambda'(\rho-|z|)). \end{aligned} \quad (12)$$

Here we used the conformality relations (8) with $A=I$. On $Q_\delta^-\cap B(w_1, r)$ we get:

$$\begin{aligned} &A\,y_u\cdot\Phi_u+A\,y_v\cdot\Phi_v \\ &\quad=\{|B(y)^{-1}y_u|^2+|B(y)^{-1}y_v|^2-2(n(y)\cdot y_u B(y)^{-1}y_u+n(y)\cdot y_v B(y)^{-1}y_v) \\ &\qquad\cdot(n(y)-n(y(w_0)))\}\,\lambda(\rho-|z|)-\Big\{\Big(\frac{z}{|z|}\cdot B(y)^{-1}y_u\Big)^2+\Big(\frac{z}{|z|}\cdot B(y)^{-1}y_v\Big)^2 \\ &\qquad-2\Big(n(y)\cdot y_u\frac{z}{|z|}\cdot B(y)^{-1}y_u+n(y)\cdot y_v\frac{z}{|z|}\cdot B(y)^{-1}y_v\Big)\frac{z}{|z|} \\ &\qquad\cdot(n(y)-n(y(w_0)))\}\,|z|\,\lambda'(\rho-|z|) \\ &\quad\geqq(|B(y)^{-1}y_u|^2+|B(y)^{-1}y_v|^2)((1-C_3\rho)\,\lambda(\rho-|z|)-\tfrac{1}{2}\rho(1+C_3\rho)\,\lambda'(\rho-|z|)), \end{aligned} \quad (13)$$

$C_3 = 2C_1/(1 - dC_1/3)$. In this estimate we have used (8) with $A = B(y)^{-1}$ and $|n(y) - n(y(w_0))| \leqq C_1 |y - y(w_0)| \leqq (C_3/2)\rho$ since $|z| \leqq \rho$.

The right hand side of system (7) can be estimated from above as follows:

$$f \cdot \Phi \leqq C_4(|B(y)^{-1} y_u|^2 + |B(y)^{-1} y_v|^2)\, \rho\, \lambda(\rho - |z|). \tag{14}$$

If we write

$$\Psi(\rho) := \int_{B(w_1, r)} (|A y_u|^2 + |A y_v|^2)\, \lambda(\rho - |z|)\, du\, dv,$$

estimates (12), (13) and (14) give us:

$$(1 - C_3\rho)\, \Psi(\rho) - \tfrac{1}{2}\rho(1 + C_3\rho)\, \Psi'(\rho) \leqq C_4 \rho\, \Psi(\rho),$$

whence

$$-\frac{d}{d\rho}(\rho^{-2}(1 + C_3\rho)\, \Psi(\rho)) \leqq C_5 \rho^{-2}(1 + C_3\rho)\, \Psi(\rho)$$

and integration from ρ_1 to ρ_2, $0 < \rho_1 < \rho_2 < (1 - dC_1/3)\, R$ yields

$$\rho_1^{-2}(1 + C_3\rho_1)\, \Psi(\rho_1) \leqq e^{C_5(\rho_2 - \rho_1)} \rho_2^{-2}(1 + C_3\rho_2)\, \Psi(\rho_2).$$

The right hand side of this inequality can be estimated from above by

$$C_6 \rho_2^{-2} \int_{B(w_1, r)} |y_u|^2 + |y_v|^2\, du\, dv$$

and the left hand side from below by

$$C_7 \rho^{-2} \int_{B(w_1, r)} \lambda(\rho - |y - y(w_0)|)(|y_u|^2 + |y_v|^2)\, du\, dv,$$

where we have written $\rho = \rho_1/(1 + dC_1/3)$.

We take $\varepsilon \in (0, \rho)$ and employ the properties of λ to arrive at

$$\rho^{-2} \int_{B(w_1, r) \cap \{w \,|\, |y(w) - y(w_0)| < \rho - \varepsilon\}} |y_u|^2 + |y_v|^2\, du\, dv \leqq C_8 \rho_2^{-2} \int_{B(w_1, r)} |y_u|^2 + |y_v|^2\, du\, dv.$$

For $\varepsilon \to 0$, $\rho \to 0$ and $\rho_2 \to (1 - dC_1/3)R$ we infer from 3. Lemma for some constant C_2 the inequality

$$C_2^{-2} \leqq R^{-2} \int_{B(w_1, r)} |y_u|^2 + |y_v|^2\, du\, dv$$

which proves the Lemma.

4. Proof of the Theorem

First of all let us recall that 4. Lemma states the smallness of $\inf |y - y(w_0)|$ taken over $\partial B(w_1, r)$ for small Dirichlet integral uniformly in $w_0 \in B(w_1, r) \setminus I$. As we shall show now this and the Courant-Lebesgue Lemma imply the continuity of y in w_1.

5. Lemma. *$y \in C^{0,\nu}(Q_\delta)$ for every $\nu \in (0,1)$.*

Proof. Let $w_1 \in I$ and let $\delta > 0$ be small enough. Assume that $B(w_1, r_1) \Subset Q_\delta$. For given $R \in (0, R_0]$ we chose r_2 from $(0, r_1]$ such that

$$C_2 \Big(\int\limits_{B(w_1, r_2)} |y_u|^2 + |y_v|^2 \, du\, dv \Big)^{\frac{1}{2}} < R.$$

According to the Lemma of Courant-Lebesgue (see [9]) there is some r_3 from $[r_2/2, r_2]$ with

$$\operatorname*{osc}_{\partial B(w_1, r_3)} y \leqq (\pi/\log 2)^{\frac{1}{2}} \Big(\int\limits_{B(w_1, r_2)} |y_u|^2 + |y_v|^2 \, du\, dv \Big)^{\frac{1}{2}}.$$

5. Lemma gives us for $r = r_3$ and $w_0 \in B(w_1, r_3) \setminus I$:

$$\inf_{\partial B(w_1, r_3)} |y - y(w_0)| \leqq R.$$

On account of

$$|y(w_{01}) - y(w_{02})| \leqq |y(w_{01}) - y(w')| + |y(w') - y(w'')| + |y(w'') - y(w_{02})|$$

$(w_{01}, w_{02} \in B(w_1, r_3)$, w', $w'' \in \partial B(w_1, r_3))$, we obtain that

$$|y(w_{01}) - y(w_{02})| \leqq C_9 R$$

or

$$\operatorname*{osc}_{B(w_1, r_3)} y \leqq C_9 R$$

and y is continuous in w_1.

The system (7) can be written as

$$\int\limits_{Q_\delta} y_u \cdot \Phi_u + y_v \cdot \Phi_v \, du\, dv = \int\limits_{Q_\delta} f \cdot \Phi + g_1 \cdot \Phi_u + g_2 \cdot \Phi_v \, du\, dv$$

with suitable $|g_1| + |g_2| \leqq C_{10}(|y_u| + |y_v|) \sup\limits_{Q_\delta} |\xi(y)|$. Since the oscillation of y on small balls is small we may apply Theorem 3 from [3] to get Hölder continuity for every Hölder exponent.

Thus we have proved the crucial starting regularity of the free boundary. For $S \in C^{2,\mu}$ $(\mu > 0)$ higher regularity follows from well known regularity results. But since we assume S only to be of class $C^2(C^{1,1})$ we have to prove Hölder continuity of the first derivatives of x. This will be done by reflection of the complex derivative of x.

6. Lemma. *$x \in C^{1,\nu}(Q^+ \cup I)$ for every $\nu \in (0,1)$.*

Proof. First of all we observe that $x \in H^{1,p}(Q_{(\delta)} \cap Q^+)$ for every $p \in [1, \infty)$, where $Q_{(\delta)} = \{(u,v) \,|\, |u| < 1 - \delta,\ |v| < 1 - \delta\}$. This follows easily from the fact that x is harmonic on Q^+ and Hölder continuous on $Q^+ \cup I$ for every Hölder exponent:

$$|x_u(w_0)| + |x_v(w_0)| \leqq \frac{2}{v_0} \sup_{\{w \,||w - w_0| < v_0\}} |x(w) - x(w_0)|.$$

But then $\xi(x)\in H^{1,p}(Q_{(\delta)}\cap Q^+)$ and $\Delta\xi(x)\in L^p(Q_{(\delta)}\cap Q^+)$ because of (15). $\xi(x)$ vanishes on I and from Theorem 15.1 in [11] we obtain:

$$\xi(x)\in C^{1,\nu}(\bar{Q}_{(\delta')}\cap\overline{Q^+}) \quad \text{for every } \nu\in(0,1) \text{ and } \delta'>\delta.$$

In order to prove the same regularity for x itself we employ a reflection trick for the complex derivative $dx/dw=(x_u-ix_v)/2$, $(w=u+iv)$.

We define

$$F(w)=\begin{cases}\dfrac{dx}{dw}(w) & \text{on } Q^+\\[2ex] T(x(\bar{w}))\dfrac{dx}{d\bar{w}}(\bar{w}) & \text{on } Q^-.\end{cases}$$

and show that for every complex valued function Φ from $C_0^1(Q_{(\delta')})$

$$\int_{Q_{(\delta')}} F\Phi_{\bar{w}}\,du\,dv=\int_{Q_{(\delta')}} G\Phi\,du\,dv$$

where $G=0$ on Q^+ and

$$G(w)=-\sum_{j=1}^{N}\frac{\partial T}{\partial x_j}(x(\bar{w}))\frac{dx_j}{dw}(\bar{w})\frac{dx}{d\bar{w}}(\bar{w})$$

on Q^-. This is done in the same way as in the proof of 2. Lemma, if we make use of transversality (5) and of $\xi(x)_u=n(x)\cdot x_u=0$ on the real axis. Theorem 1.17 in [15] says that

$$F(w)=F_0(w)-\frac{1}{\pi}\int_{Q_{(\delta'')}}\frac{G(z)}{z-w}\,d\alpha\,d\beta$$

$(z=\alpha+i\beta)$ almost everywhere on $Q_{(\delta'')}$ and F_0 is some holomorphic function. Since $G\in L^p$ for every $p<\infty$, Theorem 1.20 in [15] gives us: $F\in C^{0,\nu}(Q_{(\delta''')})$ $(\delta'''>\delta'')$ for every $\nu\in(0,1)$. Thus 6. Lemma is proved. The *proof of* our *theorem* follows directly from 6. Lemma if we use well known results on boundary regularity of Dirichlet and Neumann problems (see Theorem 12.1 in [11], Theorem 6.26 in [5] and Theorem 6.2 in [1]. We only have to formulate the boundary value problems. x is a solution of the problem

$$\Delta x=0 \text{ on } Q^+, \quad x_v=\xi(x)_v\,n(x) \text{ on } I\ .$$

and $\xi(x)$ solves

$$\Delta\xi(x)=H(x)x_u\cdot x_u+H(x)x_v\cdot x_v \text{ on } Q^+, \quad \xi(x)=0 \text{ on } I. \tag{15}$$

References

1. Agmon, S., Douglis, A., Nirenberg, L.: Estimates near the boundary for solutions of elliptic differential equations satisfying general boundary conditions I. Comm. Pure Appl. Math. **12**, 623-727 (1959)
2. Courant, R.: Dirichlet's principle, conformal mapping, and minimal surfaces. Berlin-Heidelberg-New York: Springer 1977 (Reprint)

3. Dziuk, G.: Das Verhalten von Flächen beschränkter mittlerer Krümmung an C^1-Randkurven. Nachr. Akad. Wiss. Göttingen, II. Math.-Phys. Kl. 1979, 21–28
4. Dziuk, G.: Über die Stetigkeit teilweise freier Minimalflächen. Manuscripta Math. **36**, 241–251 (1981)
5. Gilbarg, D., Trudinger, N.S.: Elliptic partial differential equations of second order. Grundlehren der mathematischen Wissenschaften 224, Berlin-Heidelberg-New York: Springer 1977
6. Goldhorn, K.H., Hildebrandt, S.: Zum Randverhalten der Lösungen gewisser zweidimensionaler Variationsprobleme mit freien Randbedingungen. Math. Z. **118**, 241–253 (1970)
7. Grüter, M.: Regularity of weak H-surfaces. J. Reine Angew. Math. **329**, 1–15 (1981)
8. Grüter, M., Hildebrandt, S., Nitsche, J.C.C.: On the boundary behavior of minimal surfaces with a free boundary which are not minima of the area. Manuscripta Math. **35**, 387–410 (1981)
9. Heinz, E.: Ein Regularitätssatz für schwache Lösungen nichtlinearer elliptischer Systeme. Nachr. Akad. Wiss. Göttingen, II. Math.-Phys. Kl. 1975, 1–13
10. Jäger, W.: Behavior of minimal surfaces with free boundaries. Comm. Pure Appl. Math. **23**, 803–818 (1970)
11. Ladyzhenskaya, O.A.; Ural'tseva, N.N.: Linear and quasilinear elliptic equations. New York-London: Academic Press 1968
12. Nitsche, J.C.C.: Minimal surfaces with partially free boundary. Least area property and Hölder continuity for boundaries satisfying a chord-arc condition. Arch. Rational Mech. Anal. **39**, 131–145 (1970)
13. Nitsche, J.C.C.: Vorlesungen über Minimalflächen. Grundlehren der mathematischen Wissenschaften 199, Berlin-Heidelberg-New York: Springer 1975
14. Nitsche, J.C.C.: The regularity of the trace for minimal surfaces. Ann. Scuola Norm. Sup Pisa, Ser. IV, **3**, 139–155 (1976)
15. Vekua, I.N.: Verallgemeinerte analytische Funktionen. Berlin: Akademie-Verlag 1963

Received July 30, 1984

Math. Z. 189, 81-110 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Über die Fourier-Jacobi-Entwicklung Siegelscher Eisensteinreihen II

Siegfried Böcherer

Mathematisches Institut der Universität, Hebelstr. 29, D-7800 Freiburg, Bundesrepublik Deutschland

1. Einleitung

In [4] hatte ich die Siegelsche Eisensteinreihe $(n+m)$-ten Grades

$$\Psi_{n+m}^k(Z)=\sum_{\{C,D\}} \det(CZ+D)^{-k} \tag{1}$$

vom geraden Gewicht $k>n+m+1$ hinsichtlich der Aufspaltung

$$Z=\begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}, \quad z_1\in\mathfrak{H}_n,\ z_4\in\mathfrak{H}_m \tag{2}$$

untersucht. Dabei bezeichnet $\mathfrak{H}_n$ die Siegelsche Halbebene n-ten Grades, die aus allen komplexen symmetrischen n-reihigen Matrizen mit positiv definitem Imaginärteil besteht; in (1) ist über alle nicht assoziierten teilerfremden symmetrischen Matrizenpaare $\{C,D\}$ zu summieren.

In Verbindung mit einem schönen Resultat von Garrett [12] hatten sich in [4] unter anderem Aussagen über die Darstellbarkeit von Modulformen durch Thetareihen sowie Formeln für Fourierkoeffizienten von verallgemeinerten Eisensteinreihen im Sinne von Klingen [17] ergeben.

In der vorliegenden Arbeit, die als Fortsetzung von [4] anzusehen ist, greife ich die Aufspaltung (2) im Spezialfall $n=m$ (dies war schon in [4] der eigentlich interessante Fall) nochmals auf.

Es wird zunächst ein Differentialoperator $\mathscr{D}_k$ eingeführt, der bezüglich der Aufspaltung (2) ein übersichtliches Transformationsverhalten zeigt:

Man bettet die symplektische Gruppe $\mathrm{Sp}(n,\mathbb{R})$ vermöge

$$M=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\mapsto M^{\uparrow}:=\begin{pmatrix} a & 0 & b & 0 \\ 0 & 1 & 0 & 0 \\ c & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

beziehungsweise

$$M^{\downarrow} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & b \\ 0 & 0 & 1 & 0 \\ 0 & c & 0 & d \end{pmatrix}$$

auf zwei Weisen in die Gruppe $\mathrm{Sp}(2n, \mathbb{R})$ ein.

Der Operator $\mathscr{D}_k$ erfüllt dann für alle holomorphen f: $\mathfrak{H}_{2n} \to \mathbb{C}$ und alle $M \in \mathrm{Sp}(n, \mathbb{R})$ die Vertauschungsrelationen

$$\mathscr{D}_k(f|_k M^{\uparrow}) = (\mathscr{D}_k f)|_{k+1} M^{\uparrow} \tag{3a}$$

$$\mathscr{D}_k(f|_k M^{\downarrow}) = (\mathscr{D}_k f)|_{k+1} M^{\downarrow}. \tag{3b}$$

Dabei wird die übliche Peterssonsche Schreibweise

$$f|_k \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(CZ+D)^{-k} f((AZ+B)(CZ+D)^{-1})$$

benutzt.

Dieser Differentialoperator kann auch in anderem Zusammenhang von Interesse sein (im Falle $n=1$ ist ein ähnlicher Operator von Eichler und Zagier im Rahmen ihrer Untersuchungen über Jacobische Modulformen [8] eingeführt worden; die Ausführungen in [8, I, § 3] haben mir in mancher Hinsicht als Vorbild gedient). Eine Verallgemeinerung auf eine beliebige Aufspaltung (2) mit $m \neq n$ ist möglich; dabei wird man aber zwangsläufig auf vektorwertige Funktionen geführt, worauf ich in dieser Arbeit nicht eingehen will.

Die Untersuchung des Differentialoperators $\mathscr{D}_k$ – insbesondere der Nachweis der Transformationseigenschaften (3a), (3b) – bilden den mathematischen Kern der vorliegenden Arbeit; die Anwendungen, um deretwillen ich diesen Operator einführe, ergeben sich durch vergleichsweise einfache Rechnungen.

Der iterierte Operator

$$\mathscr{D}_k^{\nu} := \mathscr{D}_{k+\nu-1} \circ \ldots \circ \mathscr{D}_k$$

führt für $\nu > 0$ in Verbindung mit der Restriktion $z_2 = 0$ zu einer Abbildung

$$\mathscr{D}_{k,0}^{\nu} : \mathfrak{M}_{2n}^k \to \mathfrak{S}_n^{k+\nu} \otimes \mathfrak{S}_n^{k+\nu},$$

wobei mit $\mathfrak{M}_n^k$ (bzw. $\mathfrak{S}_n^k$) der Raum der Siegelschen Modulformen (bzw. Spitzenformen) n-ten Grades vom Gewicht k bezeichnet sei.

Es wird dann gezeigt, daß die Ergebnisse aus [4] und [12] auch noch für $\mathscr{D}_{k,0}^{\nu} \Psi_{2n}^k$ sinngemäß weiterhin gültig bleiben. Daraus erhält man eine Reihe interessanter Anwendungen:

In Abschnitt 5 werden die von Cohen [6] und Zagier [29] für $n=1$ eingeführten Modulformen verallgemeinert. Man erhält Modulformen mit rationalen Fourierkoeffizienten, deren Peterssonsches Skalarprodukt mit einer Spitzenform f durch einen speziellen Wert einer f zugeordneten Dirichletreihe ausgedrückt werden kann. Abschnitt 6 bringt dann Rationalitätsaussagen für spezielle Werte der Andrianovschen (Standard-)Zetafunktion einer Hecke-Ei-

genform. Solche Aussagen findet man (unter schärferen Bedingungen an n und k) auch bei Harris [14] und Sturm [25]. Dabei wird – in Anlehnung an eine Idee von Garrett [12] – entscheidend davon Gebrauch gemacht, daß die Siegelsche Eisensteinreihe Ψ_{2n}^k selbst rationale Fourierkoeffizienten hat.

Der Rest der Arbeit ist dem Basisproblem gewidmet.

Es wird zunächst gezeigt, daß man vermöge der Abbildung $\mathscr{D}_{k,0}^{\nu}$ aus gewöhnlichen Thetareihen $2n$-ten Grades Thetareihen n-ten Grades mit harmonischen Koeffizienten gewinnen kann. Es handelt sich dabei um Reihen des Typs

$$\vartheta_{S,P}^n(z) = \sum_{G \in \mathbb{Z}^{(m,n)}} P(S^{\frac{1}{2}}G)\, e^{\pi i \operatorname{Spur}(S[G]z)}.$$

Dabei ist S die Matrix einer positiv definiten quadratischen Form in m Variablen, mit $S^{\frac{1}{2}}$ wird wie üblich die positiv definite Wurzel aus S bezeichnet und P ist eine harmonische Form vom Gewicht ν. Solche Thetareihen sind für $n=1$ von Hecke [15] und für beliebiges n von Maaß [20] eingeführt worden. Wir betrachten hier nur Thetareihen $\vartheta_{S,P}^n$ zu geraden unimodularen quadratischen Formen S (diese existieren nur für $m \equiv 0 \bmod 8$). Bei festem m, n und ν spannen alle solchen Thetareihen einen Teilraum $B_{n,\frac{m}{2}+\nu}(m)$ von $\mathfrak{M}^{\frac{m}{2}+\nu}$ auf, für $\nu > 0$ handelt es sich um Spitzenformen.

Wichtige Resultate über diese Räume stammen von Freitag: Er hat sie funktionentheoretisch charakterisiert [9] und gezeigt, daß sie invariant sind unter den Heckeoperatoren [10]. Im letzten Abschnitt dieser Arbeit wird – auf ähnlichem Wege wie in [4], wo der Fall $\nu=0$ behandelt wurde – das Hauptergebnis dieser Arbeit bewiesen: Es gilt

$$B_{n,l}(m) = \mathfrak{S}_n^l \quad \text{für alle Gewichte } l > \frac{m}{2},$$

sofern nur $m \equiv 0 \bmod 8$ und $m > 4n$; darüber hinaus kann man unter den genannten Bedingungen jede Eigenform aller Heckeoperatoren in einfacher Weise explizit als endliche Linearkombination von Thetareihen aus $B_{n,l}(m)$ darstellen. Sätze dieses Typs gibt es bisher nur für $n=1$ [26, 8].

Herr Dr. R. Weissauer (Heidelberg) hat mir freundlicherweise in einem Brief mitgeteilt, daß er – auf völlig anderem Wege – ebenfalls zu einer Lösung des Basisproblems gelangt ist.

Ich danke Herrn Prof. Zagier für die Übersendung der Arbeit [8] sowie Herrn Dr. Weissauer, der mich auf einen Fehler in der Formel (63) aufmerksam gemacht hat.

Bezeichnungen. Einige Standardbezeichnungen werden aus [4] übernommen. Satz 7 und die Gleichung (21) aus [4] werden in der Form Satz I.7 und (I.21) zitiert.

Mit $\mathrm{Hol}(\mathfrak{H}_n)$ wird die Menge der holomorphen Funktionen auf $\mathfrak{H}_n$ bezeichnet.

2. Der Differentialoperator $\mathscr{D}_k$

2.1. Die $\sqcap$-Multiplikation

Es wird der benötigte formale Apparat der multilinearen Algebra eingeführt. Genaueres findet man in [11, III, § 6].

Es sei k ein kommutativer Ring mit Eins. Man versehe den freien Modul $V=k^n$ mit der kanonischen Basis $e_1, \ldots, e_n$. Dann bilden die Elemente $e_a = e_{a_1} \wedge \ldots \wedge e_{a_p}$ eine Basis von $\Lambda^p V$, der p-ten äußeren Potenz von V $(0 \leqq p \leqq n)$, wenn $a=\{a_1, \ldots, a_p\}$ mit $a_1 < \ldots < a_p$ alle p-elementigen Teilmengen von $N=\{1, \ldots, n\}$ durchläuft $\left(\text{man schreibe dafür kurz: } a \in \binom{N}{p}\right)$.

Jeder linearen Abbildung $A\colon \Lambda^p V \to \Lambda^p V$ entspricht dann vermöge

$$A e_a = \sum_x A_x^a e_x$$

eine $\binom{n}{p}$-reihige Matrix $(A_b^a)_{a,\, b \in \binom{N}{p}}$. Zwischen der Abbildung A und der zugehörigen Matrix wollen wir nicht mehr unterscheiden. Jedem Paar linearer Abbildungen $A\colon \Lambda^p V \to \Lambda^p V$ und $B\colon \Lambda^q V \to \Lambda^q V$ wird eine lineare Abbildung $A \sqcap B\colon \Lambda^{p+q} V \to \Lambda^{p+q} V$ zugeordnet gemäß

$$(A \sqcap B)_b^a = \frac{1}{\binom{p+q}{p}} \sum_{\substack{a=a' \cup a'' \\ b=b' \cup b''}} \varepsilon(a', a'')\, \varepsilon(b', b'')\, A_{b'}^{a'} B_{b''}^{a''}.$$

Für $a'=\{a_1', \ldots, a_p'\}$, $a''=\{a_1'', \ldots, a_q''\}$ mit $a_1' < \ldots < a_p'$ und $a_1'' < \ldots < a_q''$ sei dabei $\varepsilon(a', a'')$ das Vorzeichen derjenigen Permutation, die benötigt wird, um das $(p+q)$-Tupel $(a_1', \ldots, a_p', a_1'', \ldots, a_q'')$ in die natürliche Reihenfolge zu bringen. Die Verknüpfung

$$\sqcap\colon \operatorname{End}_k(\Lambda^p V) \times \operatorname{End}_k(\Lambda^q V) \to \operatorname{End}_k(\Lambda^{p+q} V)$$

ist bilinear, assoziativ und kommutativ.

Einer linearen Abbildung $A\colon V \to V$ ordnet man vermöge

$$(A^{[p]})_b^a = |A|_b^a$$
$$(\mathrm{Ad}^{[p]} A)_b^a = \varepsilon(a, N \setminus a)\, \varepsilon(b, N \setminus b)\, |A|_{N \setminus a}^{N \setminus b}$$

zwei Abbildungen $A^{[p]}$ und $\mathrm{Ad}^{[p]} A$ zu, die beide $\Lambda^p V$ in sich überführen. Mit $|A|_b^a$ wird dabei die p-reihige Unterdeterminante

$$|A|_b^a = \det\left((A_{ij})_{\substack{i \in a \\ j \in b}}\right)$$

bezeichnet.

Für $A\colon V \to V$, $B\colon V \to V$, $C\colon \Lambda^p V \to \Lambda^p V$ und $D\colon \Lambda^q V \to \Lambda^q V$ gelten die Gesetze

$$A^{[p]} = A \sqcap \ldots \sqcap A \tag{4a}$$

$$(A+B)^{[p]}=\sum_{\alpha+\beta=p}\binom{p}{\alpha}A^{[\alpha]}\sqcap B^{[\beta]} \tag{4b}$$

$$A^{[p+q]}(C\sqcap D)=(A^{[p]}C)\sqcap(A^{[q]}D) \tag{4c}$$

$$(C\sqcap D)A^{[p+q]}=(CA^{[p]})\sqcap(DA^{[q]}) \tag{4d}$$

$$(\mathrm{Ad}^{[p]}A)A^{[p]}=\det(A)1_{\binom{n}{p}} \tag{4e}$$

2.2. *Differentialoperatoren auf* $\mathfrak{H}_{2n}$

Es werden Differentialoperatoren betrachtet, die auf holomorphe Funktionen $f\colon \mathfrak{H}_{2n}\to\mathbb{C}$ bzw. $g\colon \mathfrak{H}_{2n}\to\mathrm{End}_{\mathbb{C}}(\Lambda^p\mathbb{C}^n)$ wirken. Der in [11] entwickelte Kalkül wird der Aufspaltung

$$Z=\begin{pmatrix}z_1 & z_2\\ z_3 & z_4\end{pmatrix},\qquad Z\in\mathfrak{H}_{2n},\ z_1,z_4\in\mathfrak{H}_n,\ z_2=z_3'\in\mathbb{C}^{(n,n)} \tag{5}$$

angepaßt. Die elementaren Bausteine sind die Operatoren

$$\partial_{ij}=\tfrac{1}{2}(1+\delta_{ij})\frac{\partial}{\partial z_{ij}}\qquad(1\leqq i,j\leqq 2n),$$

($\delta=$ Kroneckersymbol), die man in der Operatorenmatrix

$$\partial=\begin{pmatrix}\partial_1 & \partial_2\\ \partial_3 & \partial_4\end{pmatrix},\qquad \partial_3'=\partial_2$$

zusammenfaßt. Die Operatoren ∂_{ij} erzeugen über $\mathbb{C}$ einen kommutativen Ring von Operatoren. Die äußeren Pozenzen $\partial_i^{[p]}$ $(1\leqq i\leqq 4)$ werden im folgenden eine wichtige Rolle spielen; diese werden auf Funktionen $f\colon \mathfrak{H}_{2n}\to\mathbb{C}$ und $g\colon \mathfrak{H}_{2n}\to\mathrm{End}_{\mathbb{C}}(\Lambda^q\mathbb{C}^n)$ angewandt. Die Anwendung ist gemäß

$$(\partial_i^{[p]}f)_b^a=|\partial_i|_b^a f$$

$$(\partial_i^{[p]}\sqcap g)_b^a=\frac{1}{\binom{p+q}{p}}\sum_{\substack{a=a'\cup a''\\ b=b'\cup b''}}\varepsilon(a',a'')\,\varepsilon(b',b'')\,|\partial_i|_{b'}^{a'}g_{b''}^{a''}$$

komponentenweise zu verstehen.

Als Testfunktionen für Operatoridentitäten werden Funktionen vom Typ

$$f_{\mathfrak{T}}(Z)=e^{\mathrm{Spur}(\mathfrak{T}Z)}=e^{\mathrm{Spur}(t_1z_1+2t_2'z_2+t_4z_4)} \tag{6}$$

von Bedeutung sein. Dabei ist $\mathfrak{T}$ stets eine (i.allg. komplexe) $2n$-reihige symmetrische Matrix. die in der Form $\mathfrak{T}=\begin{pmatrix}t_1 & t_2\\ t_3 & t_4\end{pmatrix}$, $t_3'=t_2$ in n-reihige Bestandteile aufgespalten wird.

Es gilt dann

$$\partial_i^{[p]}f_{\mathfrak{T}}=t_i^{[p]}f_{\mathfrak{T}}\qquad(1\leqq i\leqq 4) \tag{7}$$

und

$$\partial_2^{[p]} e^{\mathrm{Spur}(tz_2)} = 2^{-p}(t')^{[p]} e^{\mathrm{Spur}(tz_2)} \tag{8}$$

für $t \in \mathbb{C}^{(n,n)}$.

Wichtig sind weiterhin die für $f, g \in \mathrm{Hol}(\mathfrak{H}_{2n})$ gültige Produktregel

$$\partial_1^{[p]}(fg) = \sum_{\alpha+\beta=p} \binom{p}{\alpha} (\partial_1^{[\alpha]} f) \sqcap (\partial_1^{[\beta]} g) \tag{9}$$

sowie die Formel

$$\partial_4^{[p]} \det(z_4)^s = C_p(s) \det(z_4)^s (z_4)^{-[p]} \tag{10}$$

mit

$$C_p(s) = s(s+\tfrac{1}{2}) \ldots \left(s + \frac{p-1}{2}\right). \tag{11}$$

2.3. *Konstruktion des Operators* $\mathscr{D}_k$

In diesem Abschnitt wird in drei Schritten aus jeweils einfacheren Bestandteilen der in der Einleitung genannte Differentialoperator $\mathscr{D}_k$ aufgebaut.

1. Schritt: Es seien α, β, p nichtnegative ganze Zahlen mit $\alpha+\beta+p=n$ und $\delta(p, \alpha, \beta)$ der durch

$$\delta(p, \alpha, \beta) = z_2^{[\alpha]} \partial_4^{[\alpha]} \sqcap ((1_n^{[p]} \sqcap z_2^{[\beta]} \partial_3^{[\beta]})(\mathrm{Ad}^{[p+\beta]} \partial_1)\, \partial_2^{[p+\beta]}) \tag{12}$$

gegebene Operator.

Für Funktionen $f \in \mathrm{Hol}(\mathfrak{H}_{2n})$ der speziellen Gestalt

$$f(Z) = g(z_4, z_2)\, e^{\mathrm{Spur}(tz_1)},\ t = t',\ \det(t) \neq 0 \tag{13}$$

wird $\delta(p, \alpha, \beta)$ beschrieben durch

$$\delta(p, \alpha, \beta) f = t^{[\alpha+\beta]} z_2^{[\alpha+\beta]} (\partial_4^{[\alpha]} \sqcap \partial_3^{[\beta]} t^{-[\beta]} \partial_2^{[\beta]}) \sqcap \partial_2^{[p]} f. \tag{14}$$

Zur Begründung hat man die Identität

$$\begin{aligned} &z_2^{[\alpha]} \partial_4^{[\alpha]} \sqcap (1_n^{[p]} \sqcap z_2^{[\beta]} \partial_3^{[\beta]})(\mathrm{Ad}^{[p+\beta]} t)\, \partial_2^{[p+\beta]} \\ &\quad = t^{[\alpha+\beta]} z_2^{[\alpha+\beta]} (\partial_4^{[\alpha]} \sqcap \partial_3^{[\beta]} t^{-[\beta]} \partial_2^{[\beta]}) \sqcap \partial_2^{[p]} \end{aligned}$$

nachzuweisen; diese erhält man leicht durch (mehrfache) Anwendung der Regeln (4c), (4d) sowie (4e).

2. Schritt: Für $p+q=n$ $(p, q \geqq 0)$ setze man

$$\Delta(p, q) = \sum_{\alpha+\beta=q} (-1)^\beta\, \delta(p, \alpha, \beta).$$

Für $f(Z) = g(z_4, z_2)\, e^{\mathrm{Spur}(tz_1)}$, $t = t'$, $\det(t) \neq 0$ folgt aus (14) und (4b):

$$\Delta(p, q) f = t^{[q]} z_2^{[q]} (\partial_4 - \partial_3 t^{-1} \partial_2)^{[q]} \sqcap \partial_2^{[p]} f. \tag{15}$$

Da die $\sqcap$-Multiplikation kommutativ ist, kann man (15) auch in der (häufig günstigeren) Form

$$\Delta(p,q)f=\partial_2^{[p]}\sqcap t^{[q]}\underline{z}_2^{[q]}(\partial_4-\partial_3 t^{-1}\partial_2)^{[q]}f \tag{16}$$

schreiben, wobei das Unterstreichen von z_2 andeute, daß z_2 hinsichtlich der nachfolgenden durch $\partial_2^{[p]}$ gegebenen Differentiation als Konstante angesehen werden soll.

Der Operator $\Delta(p,q)$ besitzt eine zusätzliche Symmetrieeigenschaft bezüglich

$$Z\mapsto V\langle Z\rangle:=Z\left[\begin{pmatrix}0&1\\1&0\end{pmatrix}\right]=\begin{pmatrix}z_4&z_3\\z_2&z_1\end{pmatrix}.$$

Es gilt nämlich für alle $f\in\mathrm{Hol}(\mathfrak{H}_{2n})$

$$\Delta(p,q)(f|V)=(\Delta(p,q)f)|V. \tag{17}$$

Nach einer bekannten Schlußweise [11, III, §6] genügt es, (17) für Testfunktionen vom Typ $f_{\mathfrak{T}}$ nachzuweisen. Man darf auch zusätzlich annehmen, daß t_1,t_2,t_4 sämtlich von Maximalrang sind, insbesondere darf man $\Delta(p,q)$ in der Gestalt (15) benutzen. Die Behauptung (17) folgt dann aus der Identität

$$\begin{aligned}&t_4^{[q]}z_2^{[q]}(t_1-t_2t_4^{-1}t_3)^{[q]}\sqcap t_3^{[p]}\\&\quad=t_1^{[q]}z_3^{[q]}(t_4-t_3t_1^{-1}t_2)^{[q]}\sqcap t_2^{[p]}.\end{aligned} \tag{18}$$

Zum Nachweis von (18) hat man zusätzlich zu den Regeln (4c), (4d) noch die Gleichung

$$A^{[p]}\sqcap B^{[q]}=A'^{[p]}\sqcap B'^{[q]},\qquad A,B\in\mathbb{C}^{(n,n)},\ p+q=n$$

zu berücksichtigen.

Wir benötigen Informationen über das Verhalten des Operators $\Delta(p,q)$ bezüglich der Abbildung

$$Z\mapsto I^{\downarrow}\langle Z\rangle=\begin{pmatrix}z_1-z_2z_4^{-1}z_3&z_2z_4^{-1}\\z_4^{-1}z_3&-z_4^{-1}\end{pmatrix},$$

wobei

$$I=\begin{pmatrix}0&-1\\1&0\end{pmatrix}\in\mathrm{Sp}(n,\mathbb{R}).$$

Im nachfolgenden Abschnitt 2.4 wird gezeigt:

Proposition 1. *Es sei $f=f_{\mathfrak{T}}$ eine Testfunktion* (6) *mit* $\det(t_1)\neq 0$. *Dann gilt*

$$\begin{aligned}\Delta(p,q)(f|_kI^{\downarrow})=(-1)^n\sum_{p+q=n}\sum_{\alpha+\beta=q}(-1)^{\alpha}\binom{q}{\alpha}C_{\alpha}\left(-k+\frac{n+\beta}{2}\right)\\\cdot(t_1z_2-t_2)^{[p]}\sqcap t_1^{[q]}z_2^{[q]}\\\cdot(1_n^{[\alpha]}\sqcap z_4^{-[\beta]}(t_1^{-1}[t_2]-t_4)^{[\beta]})f|_{k+1}I^{\downarrow}.\end{aligned} \tag{19}$$

Für k kann hier nach Wahl eines Zweiges von $\log\det(z_4)$ jede beliebige komplexe Zahl zugelassen werden.

3. Schritt: Mit zunächst noch unbestimmten Koeffizienten $\lambda_k(q)$ bilde man den Operator

$$\mathscr{D}_k = \sum_{p+q=n} \lambda_k(q) \Delta(p,q). \tag{20}$$

Aus (19) erhält man nach der Aufspaltung

$$(t_1 z_2 - t_2)^{[p]} = \sum_{\gamma+\delta=p} (-1)^\delta \binom{p}{\gamma} (t_1 z_2)^{[\gamma]} \sqcap t_2^{[\delta]}$$

und einer einfachen Umformung für unsere Testfunktion $f_{\mathfrak{T}}$, $\det(t_1) \neq 0$:

$$\begin{aligned}\mathscr{D}_k(f_{\mathfrak{T}}|_k I^\downarrow) = \sum_{p+q=n} \sum_{\gamma+\delta=p} \sum_{\alpha+\beta=q} (-1)^{\alpha+\beta+\delta+n} \lambda_k(q) \binom{p}{\gamma} \binom{q}{\alpha} C_\alpha\left(-k+\frac{n+\beta}{2}\right) \\ \cdot (t_1 z_2)^{[\gamma+q]} (1_n^{[\alpha+\gamma]} \sqcap z_4^{-[\beta]} (t_4 - t_1^{-1}[t_2])^{[\beta]}) \sqcap t_2^{[\delta]} f_{\mathfrak{T}}|_{k+1} I^\downarrow.\end{aligned} \tag{21}$$

Man setze $t = \alpha + \gamma$ und

$$a_k(t,\beta,\delta) = \sum_{\alpha+\gamma=t} (-1)^{\alpha+\beta+\delta+n} \lambda_k(\alpha+\beta) \binom{\alpha+\beta}{\alpha} \binom{\gamma+\delta}{\gamma} C_\alpha\left(-k+\frac{n+\beta}{2}\right).$$

Dann geht (21) über in

$$\begin{aligned}\mathscr{D}_k(f_{\mathfrak{T}}|_k I^\downarrow) = \sum_{t+\beta+\delta=n} a_k(t,\beta,\delta)(t_1 z_2)^{[t+\beta]} \\ \cdot (1_n^{[t]} \sqcap z_4^{-[\beta]} (t_4 - t_1^{-1}[t_2])^{[\beta]}) \sqcap t_2^{[\delta]} f_{\mathfrak{T}}|_{k+1} I^\downarrow.\end{aligned} \tag{22}$$

Andererseits gilt

$$(\mathscr{D}_k f_{\mathfrak{T}})|_{k+1} I^\downarrow = \sum_{p+q=n} \lambda_k(q) (t_1 z_2 z_4^{-1})^{[q]} (t_4 - t_1^{-1}[t_2])^{[q]} \sqcap t_2^{[p]} f_{\mathfrak{T}}|_{k+1} I^\downarrow. \tag{23}$$

Die angestrebte Gleichheit

$$\mathscr{D}_k(f_{\mathfrak{T}}|_k I^\downarrow) = (\mathscr{D}_k f_{\mathfrak{T}})|_{k+1} I^\downarrow \tag{24}$$

ist gewährleistet, falls die Koeffizienten $\lambda_k(q)$ für $t+\beta+\delta=n$ die Gleichungen

$$a_k(t,\beta,\delta) = 0 \qquad (1 \leqq t \leqq n) \tag{25}$$

$$a_k(0,\beta,\delta) = \lambda_k(\beta) \tag{26}$$

erfüllen; die letzteren sind nach Definition von $a_k(0,\beta,\delta)$ automatisch erfüllt.

Im Spezialfall $\beta = 0$ besagen die Gleichungen (25), daß

$$\sum_{\alpha+\gamma=t} (-1)^\alpha \lambda_k(\alpha) \binom{\gamma+n-t}{\gamma} C_\alpha\left(-k+\frac{n}{2}\right) = 0 \qquad (1 \leqq t \leqq n). \tag{27}$$

Falls man voraussetzt, daß k von $\frac{n}{2}, \frac{n+1}{2}, \ldots, \frac{2n-1}{2}$ verschieden ist, so gilt

$C_\alpha\left(-k+\frac{n}{2}\right)\neq 0$ für $0\leqq\alpha\leqq n$, und man erhält

$$\lambda_k(\alpha)=\binom{n}{\alpha}\frac{\lambda_k(0)}{C_\alpha\left(-k+\frac{n}{2}\right)}\qquad(0\leqq\alpha\leqq n).\tag{28}$$

Die Gleichungen (25) sind dann auch bei beliebigem β erfüllt: Wegen

$$C_{\alpha+\beta}\left(-k+\frac{n}{2}\right)=C_\beta\left(-k+\frac{n}{2}\right)C_\alpha\left(-k+\frac{n+\beta}{2}\right)$$

und

$$\binom{n}{\alpha+\beta}\binom{\alpha+\beta}{\alpha}\binom{\gamma+\delta}{\gamma}=\frac{n!}{t!\,\beta!\,\delta!}\binom{t}{\alpha}\qquad\text{mit } t=\alpha+\gamma$$

folgt nämlich für $t\geqq 1$:

$$a_k(t,\beta,\delta)=(-1)^{\beta+\delta+n}\frac{n!}{t!\,\beta!\,\delta!}\frac{1}{C_\beta\left(-k+\frac{n}{2}\right)}\sum_{\alpha+\gamma=t}(-1)^\alpha\binom{t}{\alpha}=0.$$

Setzt man noch

$$\tilde{C}_q(x)=\prod_{\substack{0\leqq p\leqq n\\ p\neq q}}C_p(x),$$

so erhält man mit der Normierung

$$\lambda_k(0)=\tilde{C}_0\left(-k+\frac{n}{2}\right):$$

Satz 2. *Der Operator*

$$\begin{aligned}\mathscr{D}_k&=\sum_{p+q=n}\binom{n}{q}\tilde{C}_q\left(-k+\frac{n}{2}\right)\Delta(p,q)\\&=\sum_{\alpha+\beta+p=n}(-1)^\beta\binom{n}{\alpha+\beta}\tilde{C}_{\alpha+\beta}\left(-k+\frac{n}{2}\right)\\&\quad\cdot z_2^{[\alpha]}\partial_4^{[\alpha]}\sqcap((1_n^{[p]}\sqcap z_2^{[\beta]}\partial_3^{[\beta]})(\mathrm{Ad}^{[p+\beta]}\partial_1)\partial_2^{[p+\beta]})\end{aligned}$$

genügt den Vertauschungsrelationen

$$\mathscr{D}_k(f\,|\,V)=(\mathscr{D}_k f)|V\tag{29a}$$

$$\mathscr{D}_k(f|_k M^\downarrow)=(\mathscr{D}_k f)|_{k+1}M^\downarrow\tag{29b}$$

$$\mathscr{D}_k(f|_k M^\uparrow)=(\mathscr{D}_k f)|_{k+1}M^\uparrow\tag{29c}$$

für alle $f\in\mathrm{Hol}(\mathfrak{H}_{2n})$, $M\in\mathrm{Sp}(n,\mathbb{R})$ *und* $V=\begin{pmatrix}0&1&&0\\1&0&&\\&&0&1\\0&&1&0\end{pmatrix}\in\mathrm{Sp}(2n,\mathbb{R})$. *Für* k

ist dabei eine beliebige komplexe Zahl zugelassen, wenn man für jedes $M = \begin{pmatrix} * & * \\ c & d \end{pmatrix}$ *aus* $\mathrm{Sp}(n, \mathbb{R})$ *einen festen Zweig von* $\log\det(cz_4+d)$ *zugrunde legt.*

Beweis. Die Formel (29a) gilt nach (17) bereits für $\Delta(p, q)$. Wegen (24) gilt (29b) jedenfalls für $M=I=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, alle Testfunktionen $f_{\mathfrak{T}}$ mit $\det(t_1)\neq 0$ und alle $k\in\mathbb{C}$, die von $\frac{n}{2}, \ldots, \frac{2n-1}{2}$ verschieden sind; daraus ergibt sich die Gültigkeit von (29b) im Falle $M=I$ für alle $k\in\mathbb{C}$ und alle $f\in\mathrm{Hol}(\mathfrak{H}_{2n})$. Für $j(M, z_4)^k = e^{k\log\det(cz_4+d)}$ gilt die übliche Kozykeleigenschaft i.allg. nicht mehr. Man hat nur

$$j(MN, z_4)^k = j(M, N\langle z_4\rangle)^k j(N, z_4)^k \mu_k(M, N) \tag{30}$$

zur Verfügung, wobei $\mu_k(M, N)\neq 0$ eine komplexe Zahl vom Betrag 1 ist, welche nur von k modulo $\mathbb{Z}$ abhängt: $\mu_k(M, N)=\mu_{k+1}(M, N)$. Daher genügt es, (29b) für Erzeugende der symplektischen Gruppe $\mathrm{Sp}(n, \mathbb{R})$ nachzuweisen. Für die Translationen

$$\begin{pmatrix} 1_n & S \\ 0 & 1_n \end{pmatrix}, \qquad S=S' \text{ reell}$$

gilt (29b) trivialerweise. Da die Translationen zusammen mit I ganz $\mathrm{Sp}(n, \mathbb{R})$ erzeugen, folgt (29b).

Vermöge V sind die Untergruppen $\mathrm{Sp}(n, \mathbb{R})^{\uparrow}$ und $\mathrm{Sp}(n, \mathbb{R})^{\downarrow}$ von $\mathrm{Sp}(2n, \mathbb{R})$ zueinander konjugiert, so daß (29c) aus (29a) und (29b) folgt.

Bemerkung. Für $n=1$ erhält man einen Operator, der in etwas anderer Form von Eichler und Zagier [8] eingeführt worden ist:

$$\mathscr{D}_k = (-k+\tfrac{1}{2})\partial_2 + z_2(\partial_1\partial_4 - \partial_2\partial_2).$$

Für $\nu\in\mathbb{N}$ führen wir den iterierten Operator

$$\mathscr{D}_k^{\nu} := \mathscr{D}_{k+\nu-1}\circ\ldots\circ\mathscr{D}_k$$

und dessen Komposition mit der Restriktion $z_2=0$ ein:

$$\mathscr{D}_{k,0}^{\nu}\colon \mathrm{Hol}(H_{2n})\to\mathrm{Hol}(H_n\times H_n)$$

wobei $(\mathscr{D}_{k,0}^{\nu} f)(z_1, z_4) = (\mathscr{D}_k^{\nu} f)\begin{pmatrix} z_1 & 0 \\ 0 & z_4 \end{pmatrix}$.

Diese Operatoren führen natürlich Modulformen in Modulformen über:

Satz 3. *Der Operator* $\mathscr{D}_{k,0}^{\nu}$ *vermittelt für* $\nu>0$ *eine lineare Abbildung*

$$\mathscr{D}_{k,0}^{\nu}\colon \mathfrak{M}_{2n}^k \to (\mathfrak{S}_n^{k+\nu}\otimes\mathfrak{S}_n^{k+\nu})^{\mathrm{sym}}.$$

Die Symmetrieeigenschaft des Bildes folgt aus (29a). Nur die Aussage, daß man Spitzenformen erhält, bedarf einer Begründung. Dazu braucht man nur zu überlegen, daß

$$\Delta(p, q)\, e^{\mathrm{Spur}(\mathfrak{T}Z)} = 0, \tag{31}$$

falls die letzte Zeile der symmetrischen Matrix $\mathfrak{T}$ gleich Null ist.

Zum Beweis von (31) darf man o.B.d.A. annehmen, daß t_1 Maximalrang hat, also $\Delta(p, q)$ in der Gestalt (15) vorliegt. Die Behauptung folgt dann aus dem elementaren Sachverhalt, daß

$$A^{[p]} \sqcap B^{[q]} = 0 \qquad (p+q=n),$$

falls A und B n-reihige Matrizen sind, deren letzte Spalte jeweils die Nullspalte ist.

Bemerkungen:

1) Es sollte darauf hingewiesen werden, daß Satz 3 vielerlei Varianten zuläßt (für Jacobische Modulformen, für andere Untergruppen von $\mathrm{Sp}(n, \mathbb{R}), \ldots$).

2) Mit den bereitgestellten Hilfsmitteln kann man auch Differentialoperatoren einführen, die einer allgemeineren Aufspaltung

$$Z = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix}, \qquad Z \in \mathfrak{H}_n,\ z_1 \in \mathfrak{H}_r,\ z_4 \in \mathfrak{H}_s,\ r+s=n$$

angepaßt sind. Solche Operatoren führen dann skalarwertige Funktionen (bzw. Modulformen) in vektorwertige Funktionen (bzw. Modulformen) über.

3) Den „Hauptterm" von $\mathscr{D}_k$ hat man in

$$\Delta(n, 0) = \delta(n, 0, 0) = \partial_2^{[n]}$$

zu sehen, die übrigen $\Delta(p, q)$ haben den Charakter von Korrekturtermen. Daher ist es in den Anwendungen meist günstiger, den Operator $\mathscr{D}_k$ anders zu normieren:

$$\tilde{\mathscr{D}}_k := \sum_{p+q=n} \binom{n}{q} C_q \left(-k+\frac{n}{2}\right)^{-1} \Delta(p, q). \tag{32}$$

Analog bilde man $\tilde{\mathscr{D}}_k^{\nu}$, $\tilde{\mathscr{D}}_{k,0}^{\nu}$; dabei hat man aber die Gewichte

$$k = \frac{n-2\nu+2}{2}, \frac{n-2\nu+3}{2}, \ldots, \frac{2n-1}{2}$$

auszuschließen.

2.4. *Beweis der Proposition 1*

Lemma 4. *Es sei $T = T'$ n-reihig mit $\det(T) \neq 0$. Dann gilt für $r \geqq 1$:*

$$(\partial_4 + \partial_3 T^{-1} \partial_2)^{[r]} \det(z_4)^{-\frac{n}{2}} e^{\mathrm{Spur}(T[z_2 z_4^{-1}])} = 0.$$

Beweis. Man darf o.B.d.A. annehmen, daß T reell und positiv ist (insbesondere existiert dann $T^{\frac{1}{2}}$), außerdem kann man die komplexe Variable Z durch eine reelle Variable

$$Y = Y' = \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} > 0, \qquad y_3' = y_2$$

ersetzen. Nach dem Vorbild [11, III, § 6] benutze man die Integralformel

$$\det(y_4)^{-\frac{n}{2}} = \pi^{-\frac{n^2}{2}} \int e^{-\mathrm{Spur}(y_4[P])}\, dP,$$

wobei über den Raum der n-reihigen reellen Matrizen zu integrieren ist. Die Variablensubstitution

$$P \mapsto P + y_4^{-1} y_2' T^{\frac{1}{2}}$$

führt zu

$$\det(y_4)^{-\frac{n}{2}} e^{\mathrm{Spur}(T[y_2] y_4^{-1})} = \pi^{-\frac{n^2}{2}} \int e^{-\mathrm{Spur}(y_4[P] + 2PT^{\frac{1}{2}} y_2)}\, dP. \tag{33}$$

Aus den Formeln

$$(\partial_4 + \partial_3 T^{-1} \partial_2)^{[r]} = \sum_{\alpha+\beta=r} \binom{r}{\alpha} \partial_4^{[\alpha]} \sqcap (\partial_3 T^{-1} \partial_2)^{[\beta]}$$

$$\partial_4^{[\alpha]} e^{-\mathrm{Spur}(y_4[P])} = (-PP')^{[\alpha]} e^{-\mathrm{Spur}(y_4[P])}$$

$$(\partial_3 T^{-1} \partial_2)^{[\beta]} e^{-\mathrm{Spur}(2PT^{\frac{1}{2}} y_2)} = (PP')^{[\beta]} e^{-\mathrm{Spur}(2PT^{\frac{1}{2}} y_2)}$$

$$0 = (PP' - PP')^{[r]} = \sum_{\alpha+\beta=r} \binom{r}{\alpha} (-PP')^{[\alpha]} \sqcap (PP')^{[\beta]}$$

und der Gleichung (33) folgt die Behauptung.

Lemma 5. *Es sei $T = T'$ eine n-reihige Matrix, $Z \in \mathfrak{H}_n$. Dann gilt*

$$\partial^{[q]} \det(Z)^{-k} e^{\mathrm{Spur}(TZ^{-1})} = (-1)^q \det(Z)^{-k} \sum_{\alpha+\beta=q} \binom{q}{\alpha} C_\alpha \left(k - \frac{q-1}{2}\right) \cdot Z^{-[q]} (Z^{[\alpha]} \sqcap T^{[\beta]}) Z^{-[q]} e^{\mathrm{Spur}(TZ^{-1})}.$$

Beweis. Das Verhalten der Operatoren $\partial^{[q]}$ unter der Substitution $Z \mapsto Z^{-1}$ ist bekannt [11, S. 214]. Speziell gilt

$$\partial^{[q]} \det(Z)^{-k} e^{\mathrm{Spur}(TZ^{-1})}$$
$$= (-1)^q \det(Z)^{-\frac{q-1}{2}} Z^{-[q]} (\partial^{[q]} h)(Z^{-1}) Z^{-[q]},$$

wobei $h(Z) = \det(Z)^{k-\frac{q-1}{2}} e^{\mathrm{Spur}(TZ)}$.

Die Behauptung folgt jetzt aus (10) und der Produktformel (9).

Zum Beweis von Proposition 1 ist

$$f_{\mathfrak{T}}|_k I^{\downarrow} = \det(z_4)^{-k} e^{\mathrm{Spur}(t_1(z_1 - z_2 z_4^{-1} z_3) + 2t_2' z_2 z_4^{-1} - t_4 z_4^{-1})}$$

zu untersuchen. Es ist zweckmäßig, diese Funktion aufzuspalten gemäß

$$f_{\mathfrak{T}}|_k I^{\downarrow} = g(z_2, z_4)\, h(z_4)\, e^{\mathrm{Spur}(t_1 z_1)} \tag{34}$$

mit

$$g(z_2, z_4) = \det(z_4)^{-\frac{n}{2}} e^{\mathrm{Spur}(-t_1[z_2 - t_1^{-1}[t_2]]z_4^{-1})}$$

$$h(z_4) = \det(z_4)^{-k+\frac{n}{2}} e^{\mathrm{Spur}((t_1^{-1}[t_2] - t_4)z_4^{-1})}.$$

Mit dieser Aufspaltung gilt

Lemma 6.

$$(\partial_4 - \partial_3 t_1^{-1}\partial_2)^{[q]}(gh) = \sum_{\alpha+\beta=q} \binom{q}{\alpha} ((\partial_4 - \partial_3 t_1^{-1}\partial_2)^{[\alpha]} g) \sqcap (\partial_4^{[\beta]} h). \tag{35}$$

Beweis. Nach (4b) und der Produktregel (9) gilt

$$(\partial_4 - \partial_3 t_1^{-1}\partial_2)^{[q]}(gh)$$

$$= \sum_{a+b=q} \sum_{b_1+b_2=b} \binom{q}{a} \binom{b}{b_1} (-\partial_3 t_1^{-1}\partial_2)^{[a]} \sqcap ((\partial_4^{[b_1]} g) \sqcap (\partial_4^{[b_2]} h)). \tag{36}$$

Da die Funktion h nicht von z_2 abhängt, darf man in (36) anders herum klammern:

$$\sum_{a+b=q} \sum_{b_1+b_2=b} \binom{q}{a} \binom{b}{b_1} ((-\partial_3 t_1^{-1}\partial_2)^{[a]} \sqcap \partial_4^{[b_1]} g) \sqcap (\partial_4^{[b_2]} h).$$

Wegen

$$\binom{q}{a}\binom{b}{b_1} = \binom{q}{a+b_1}\binom{a+b_1}{a}$$

folgt durch nochmalige Anwendung von (5b) die Behauptung.

Zur Bestimmung von $\Delta(p, q) f_{\mathfrak{T}}$ darf man $\Delta(p, q)$ in der Gestalt (16) benutzen. Der Beitrag, der von $(\partial_4 - \partial_3 t_1^{-1}\partial_2)^{[q]}$ geliefert wird, ist in Lemma 6 beschrieben. Nach Lemma 4 liefert dabei nur $\alpha = 0$ einen Beitrag zu (35), der Wert von $\partial_4^{[q]} h$ kann Lemma 5 entnommen werden. Man erhält also

$$(\partial_4 - \partial_3 t^{-1}\partial_2)^{[q]} f_{\mathfrak{T}}|_k I^{\downarrow} = (-1)^q \det(z_4)^{-k+\frac{n}{2}} \sum_{\alpha+\beta=q} C_\alpha \left(k - \frac{n}{2} - \frac{q-1}{2}\right)$$
$$\cdot z_4^{-[q]} (z_4^{[\alpha]} \sqcap (t_1^{-1}[t_2] - t_4)^{[\beta]}) z_4^{-[q]} e^{\mathrm{Spur}(\mathfrak{T} I^{\downarrow}\langle Z\rangle)}. \tag{37}$$

Andererseits zeigt man leicht, daß

$$\partial_2^{[p]} g(z_2, z_4) = (-1)^p (t_1 z_2 - t_2)^{[p]} z_4^{-[p]} g(z_2, z_4). \tag{38}$$

Die Behauptung der Proposition 1 folgt nun aus (37) und (38), wenn man die Regeln (4c) und (4d) sowie

$$C_\alpha \left(k - \frac{n}{2} - \frac{q-1}{2}\right) = (-1)^\alpha C_\alpha \left(-k + \frac{n}{2} + \frac{\beta}{2}\right), \qquad \alpha+\beta = q$$

beachtet.

3. Die Fourier-Jacobi-Entwicklung von $\mathscr{D}_k \Psi_{2n}^k$

Die Eisensteinreihen n-ten Grades werden für gerades Gewicht $k>n+1$ gegeben durch

$$\Psi_n^k(z)=\sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma^{(n)}} j(M,z)^{-k}=\sum_{(c,d)} \det(cz+d)^{-k}. \tag{39}$$

Dabei bezeichnet $\Gamma^{(n)}$ die Siegelsche Modulgruppe $\mathrm{Sp}(n,\mathbb{Z})$,

$$\Gamma_\infty^{(n)}=\left\{M=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in\Gamma^{(n)}\,\middle|\, c=0\right\},$$

und (c,d) durchläuft alle nicht assoziierten teilerfremden symmetrischen Matrizenpaare. Die Fourierentwicklung von Ψ_n^k schreiben wir in der Form

$$\Psi_n^k(z)=\sum_{t\geqq 0} a_n^k(t)\, e^{2\pi i \mathrm{Spur}(tz)}, \tag{40}$$

wobei $t=t'$ alle positiv semidefiniten halbganzen n-reihigen Matrizen durchläuft.

Neben diesen Eisensteinreihen werden noch Poincaréreihen vom Exponentialtyp benötigt. Diese sind für $z\in\mathfrak{H}_n$, $k>2n$, $k\cdot n$ gerade und positiv definites halbganzes $t^{(n)}$ gegeben durch

$$\mathfrak{g}_n^k(z,t)=\sum_{M\in A_n\backslash\Gamma^{(n)}} j(M,z)^{-k}\, e^{2\pi i \mathrm{Spur}(tM\langle z\rangle)}.$$

Mit A_n wird dabei die Untergruppe der Translationen bezeichnet:

$$A_n=\left\{M\in\Gamma^{(n)}\,\middle|\, M=\begin{pmatrix}\pm 1 & b\\ 0 & \pm 1\end{pmatrix}\right\}.$$

Für alle Spitzenformen $f\in\mathfrak{S}_n^k$ mit Fourierentwicklung

$$f(z)=\sum_{t>0} b(t,f) e^{2\pi i \mathrm{Spur}(tz)}$$

gilt

$$\langle f, \mathfrak{g}_n^k(*,t)\rangle = A_n^k\, b(t,f)\det(t)^{\frac{n+1}{2}-k}, \tag{41}$$

wobei A_n^k die numerische Konstante

$$A_n^k=\pi^{\frac{n(n-1)}{4}}(4\pi)^{\frac{n(n+1)}{2}-nk}\prod_{\nu=1}^{n}\Gamma\left(k-\frac{n+\nu}{2}\right)$$

ist, und $\langle\ ,\ \rangle$ das Peterssonsche Skalarprodukt bezeichnet, das für zwei Modulformen $f, g\in\mathfrak{M}_n^k$ mit $f\cdot g\in\mathfrak{S}_n^{2k}$ gegeben wird durch

$$\langle f,g\rangle=\int_{\Gamma^{(n)}\backslash\mathfrak{H}_n} f\cdot\bar{g}\det(y)^{k-n-1}\,dx\,dy.$$

Es soll jetzt gezeigt werden, daß die Sätze I.7 und I.12 hinsichtlich der Aufspaltung (5) auch für $\tilde{\mathscr{D}}_k^\nu \Psi_{2n}^k$ gelten:

Satz 7. *Es sei k gerade, $k>2n$ und $Z\in\mathfrak{H}_{2n}$. Dann gilt für alle $\nu\in\mathbb{N}$:*

$$\tilde{\mathscr{D}}_k^{\nu}\Psi_{2n}^k(Z)=\phi_1^{(\nu)}(Z)+\sum_{t>0}\sum_{w_1\in\mathbb{Z}_n^{(n,n)}/GL(n,\mathbb{Z})}\sum_{\substack{w_3\in\mathbb{Z}^{(n,n)}\\ \binom{w_1}{w_3}\text{primitiv}}}\sum_{M\in\Gamma_{(2)}^{(2)}\backslash\Gamma^{(n)}}$$
$$\cdot\det(2\pi i w_1 t w_3')^{\nu}a_n^k(t)j(M,z_4)^{-k-\nu}e^{2\pi i\mathrm{Spur}(tM\downarrow\langle Z\rangle[\binom{w_1}{w_3}])}. \tag{42}$$

Dabei bezeichnet $\mathbb{Z}_n^{(n,n)}$ die n-reihigen ganzen Matrizen vom Maximalrang n. In $\phi_1^{(\nu)}$ werden all die Glieder der Fourier-Jacobi-Entwicklung von $\tilde{\mathscr{D}}_k^{\nu}\Psi_{2n}^k$ bezüglich z_1 zusammengefaßt, die zu Exponentenmatrizen t vom Rang $<n$ gehören. Aus (31) folgt, daß $\phi_1^{(\nu)}$ für $\nu>0$ identisch verschwindet.

Für $\nu=0$ handelt es sich bei (42) um einen Spezialfall von Satz I.7. Durch Anwendung des Operators $\tilde{\mathscr{D}}_k^{\nu}$ erhält man daraus (42) auch für $\nu>0$: Man benutze die Vertauschungsrelation (29b) und die einfache Formel

$$\tilde{\mathscr{D}}_k^{\nu}(e^{2\pi i\mathrm{Spur}(tZ[\binom{w_1}{w_3}])})=\det(2\pi i w_1 t w_3')^{\nu}e^{2\pi i\mathrm{Spur}(tZ[\binom{w_1}{w_3}])}. \tag{43}$$

Wie in [4] erhält man daraus mit einem einfachen Umordnungsargument:

Satz 8. (*Voraussetzungen und Bezeichnungen wie in Satz 7*)

$$(\tilde{\mathscr{D}}_k^{\nu}\Psi_{2n}^k)\begin{pmatrix}z_1&0\\0&z_4\end{pmatrix}=\phi_1^{(\nu)}\begin{pmatrix}z_1&0\\0&z_4\end{pmatrix}+\phi_2^{(\nu)}\begin{pmatrix}z_1&0\\0&z_4\end{pmatrix}$$
$$+2\sum_{t>0}\sum_{w_1\in\mathbb{Z}_n^{(n,n)}/GL(n,\mathbb{Z})}(2\pi i)^{n\nu}a_n^k(t)e^{2\pi i\mathrm{Spur}(t[w_1']z_1)}$$
$$\cdot\sum_{\substack{w_3\in GL(n,\mathbb{Z})\backslash\mathbb{Z}_n^{(n,n)}\\ \binom{w_1}{w_3}\text{primitiv}}}\det(w_1 t w_3')^{\nu}\mathfrak{g}_n^{k+\nu}(z_4,t[w_3']). \tag{44}$$

Dabei wurden in $\phi_2^{(\nu)}$ diejenigen Summanden aus (42) zusammengefaßt, die zu Matrizen w_3 mit $\mathrm{Rang}(w_3)<n$ gehören. Es gilt

$$\phi_1^{(\nu)}=\phi_2^{(\nu)}=0 \quad \text{für } \nu>0.$$

Da wir $\tilde{\mathscr{D}}_k^{\nu}\Psi_{2n}^k$ durch einen Differentiationsprozeß aus Ψ_{2n}^k gewonnen haben, treten gegenüber dem Fall $\nu=0$ keine neuen Konvergenzprobleme auf (vgl. [4]).

4. Die Hauptformel von Garrett für $\tilde{\mathscr{D}}_k^{\nu}\Psi_{2n}^k$

Es sei D eine n-reihige Elementarteilermatrix

$$D=\begin{pmatrix}d_1&&0\\&\ddots&\\0&&d_n\end{pmatrix},\quad d_i|d_{i+1},\ d_i>0.$$

Einer solchen Matrix ordne man den Heckeoperator $T(D)$ zu, der für $f\in\mathfrak{M}_n^k$ gegeben wird durch

$$f\mapsto f|T(D):=\det(D)^{-k}\sum_{R\in\Gamma^{(n)}\backslash\Gamma^{(n)}\binom{D\ 0}{0\ D^{-1}}\Gamma^{(n)}}f|_kR. \tag{45}$$

Für die allgemeine Theorie der Heckeoperatoren sei auf die Literatur verwiesen [1, 11].

Für $f \in \mathfrak{S}_n^k$, $\nu \in \mathbb{N}$ bilden wir (Konvergenz vorausgesetzt) die unendliche Reihe

$$f|S_n^{(\nu)} = \sum_D \det(D)^\nu f|T(D), \tag{46}$$

wobei D alle Elementarteilermatrizen durchläuft.

Falls (46) für alle $f \in \mathfrak{S}_n^k$ konvergiert, ist $S_n^{(\nu)}$ ein bezüglich $\langle\ ,\ \rangle$ Hermitescher Operator auf $\mathfrak{S}_n^k$, insbesondere besitzt dann $\mathfrak{S}_n^k$ eine Orthogonalbasis, die aus lauter Eigenfunktionen von $S_n^{(\nu)}$ besteht.

Mit der numerischen Konstanten

$$\mu_n^k = 2^{n\left(\frac{n}{2}+\frac{3}{2}-k\right)+1} i^{kn} \pi^{\frac{n(n+1)}{2}} \prod_{\nu=1}^{n} \frac{\Gamma\left(k-\frac{n+\nu}{2}\right)}{\Gamma\left(k-\frac{\nu-1}{2}\right)}$$

gilt nun

Satz 9. (Garrettsche Hauptformel)

Es sei k gerade, $k>2n$ und $\nu \in \mathbb{N}$, $f_1, \ldots, f_d$ eine Orthogonalbasis von $\mathfrak{S}^{k+\nu}$, die aus Eigenfunktionen von $S_n^{(\nu)}$ besteht:

$$f_i|S_n^{(\nu)} = \lambda(f_i, \nu) f_i.$$

Dann gilt für die Aufspaltung (5):

$$(\tilde{\mathscr{D}}_k^\nu \Psi_{2n}^k)\begin{pmatrix} z_1 & 0 \\ 0 & z_4 \end{pmatrix} = \delta_{0,\nu} F(z_1, z_4)$$

$$+ \mu_n^{k+\nu} \prod_{i=0}^{\nu-1} C_n(-k-i) \sum_{j=1}^{d} \lambda(f_j, \nu) \frac{f_j(z_1)\overline{f_j(-\bar{z}_4)}}{\langle f_j, f_j\rangle}. \tag{47}$$

Dabei ist $F(z_1, z_4)$ eine Funktion aus $(\mathfrak{S}_n^k)^\perp \otimes (\mathfrak{S}_n^k)^\perp$, δ = Kroneckersymbol.

Beweis. Für $\nu = 0$ handelt es sich um die „Main Formula" von Garrett ([12]) für unsere Standardaufspaltung (5). In [12] wird auch die Funktion $F(z_1, z_4)$ explizit beschrieben, sie ist hier aber nicht weiter von Interesse.

Ausgangspunkt in Garretts Beweis von (47) im Falle $\nu \neq 0$ ist eine geschickte Aufspaltung der Summation

$$\Psi_{2n}^k(Z) = \sum_{R \in \Gamma^{(2n)}_\infty \backslash \Gamma^{(2n)}} j(R, Z)^{-k}.$$

Er zeigt nämlich, daß man diese in der Form

$$\Psi_{2n}^k(Z) = \sum_M \sum_H j(G_{\tilde{M}} H, Z)^{-k} \tag{48}$$

schreiben kann.

Dabei durchläuft

a) M alle r-reihigen Elementarteilermatrizen

$$M=\begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_r \end{pmatrix}, \quad m_i | m_{i+1}, \quad \text{für } 0 \leqq r \leqq n,$$

$$\tilde{M}=\begin{pmatrix} 0 & 0 \\ 0 & M \end{pmatrix} \in \mathbb{Z}^{(n,n)}, \qquad G_{\tilde{M}}=\begin{pmatrix} 1_{2n} & 0 & \\ 0 & \tilde{M} & 1_{2n} \\ \tilde{M}' & 0 & \end{pmatrix} \in \Gamma^{(2n)}.$$

b) H eine (von M abhängige) Teilmenge von $\Gamma^{(n)\uparrow} \cdot \Gamma^{(n)\downarrow}$.

Wegen der (noch zu beweisenden) Formel

$$\tilde{\mathscr{D}}_k^{\nu} j(G_{\tilde{M}} H, Z)^{-k} = \prod_{i=0}^{\nu-1} C_n(-k-i) \det(M)^{\nu} j(G_{\tilde{M}} H, Z)^{-k-\nu} \tag{49}$$

ist Garretts Beweis wortwörtlich auch für $\nu > 0$ richtig (mit $k+\nu$ statt k und einem zusätzlichen Faktor $\det(M)^{\nu}$), diejenigen M mit $\operatorname{Rang}(\tilde{M}) < n$ liefern keinen Beitrag mehr. Ein Beitrag von Nichtspitzenformen tritt für $\nu > 0$ in (47) nicht mehr auf; dieser rührte für $\nu = 0$ nämlich genau von den M mit $\operatorname{Rang}(\tilde{M}) < n$ her.

Zur Begründung von (49) ist anzumerken:

Wegen (29b), (29c) darf man $H = 1_{4n}$ annehmen. Es genügt daher, folgendes zu zeigen:

Lemma 10. *Für* $M \in \mathbb{R}^{(n,n)}$ *und* $G_M = \begin{pmatrix} 1_{2n} & 0 & \\ 0 & M & 1_{2n} \\ M' & 0 & \end{pmatrix}$ *gilt:*

$$\tilde{\mathscr{D}}_k j(G_M, Z)^{-k} = C_n(-k) \det(M) j(G_M, Z)^{-k-1}. \tag{50}$$

Da (50) analytisch in M ist, kann man sich auf reguläre M beschränken; für diese gilt insbesondere

$$\begin{pmatrix} 0 & M^{-1} \\ -M' & 0 \end{pmatrix} \in \operatorname{Sp}(n, \mathbb{R}).$$

Für die Hilfsfunktion $h_k(Z) := \det(z_1 + z_2 + z_2' + z_4)^{-k}$ zeigt man durch direktes Nachrechnen

$$h_k|_k \begin{pmatrix} 0 & M^{-1} \\ -M' & 0 \end{pmatrix}^{\uparrow} = \det(M)^k j(G_M, Z)^{-k}. \tag{51}$$

Wieder wegen (29b) braucht nur noch $\tilde{\mathscr{D}}_k h_k$ bestimmt zu werden.

Bemerkung. a) Es sei f eine auf $\mathfrak{H}_{2n}$ holomorphe Funktion, die nur von $z_1 + z_2 + z_2' + z_4$ abhängt. Dann gilt

$$\Delta(p, q) f = 0 \quad \text{für } q > 0.$$

b) $\partial_2^{[h]} \det(z_2 + z_2')^{\alpha} = C_h(\alpha) \det(z_2 + z_2')^{\alpha} (z_2 + z_2')^{-[h]} \quad (\alpha \in \mathbb{C},\ 0 \leqq h \leqq n).$

Der (elementare) Beweis sei dem Leser überlassen. Aus der Bemerkung folgt

$$\tilde{\mathscr{D}}_k h_k = C_n(-k)\, h_{k+1},$$

in Verbindung mit (51) ergibt sich die Behauptung des Lemmas.

Die Garrettsche Hauptformel besagt, daß die Funktion $\tilde{\mathscr{D}}^{\nu}_{k,0} \Psi^k_{2n}$ als Kernfunktion der Abbildung $S^{(\nu)}_n$ bezüglich $\langle\,,\rangle$ interpretiert werden kann, genauer gilt für alle $f \in \mathfrak{S}^{k+\nu}_n$

$$(\tilde{\mu}^{k+\nu}_{n,\nu} S^{(\nu)}_n f)(z) = \left\langle f, \tilde{\mathscr{D}}^{\nu}_k \Psi^k_{2n} \begin{pmatrix} -\bar{z} & 0 \\ 0 & * \end{pmatrix} \right\rangle, \tag{52}$$

wobei noch

$$\tilde{\mu}^{k+\nu}_{n,\nu} = \mu^{k+\nu}_n \prod_{i=0}^{\nu-1} C_n(-k-i)$$

gesetzt wurde.

Die Fourierentwicklungen

$$f(z) = \sum_{t>0} b(t,f)\, e^{2\pi i \,\mathrm{Spur}(tz)}$$

$$(\tilde{\mu}^{k+\nu}_{n,\nu} S^{(\nu)}_n f)(z) = \sum_{t>0} d_\nu(t,f)\, e^{2\pi i \,\mathrm{Spur}(tz)}$$

stehen wegen (44) und (41) in einem einfachen Zusammenhang:

Für alle halbganzen $T^{(n)} > 0$ gilt

$$d_\nu(T) = 2 A^{k+\nu}_n (2\pi i)^{n\nu} \sum_{t>0} \sum_{w_1 \in \mathbb{Z}^{(n,n)}_+ / Gl(n,\mathbb{Z})} a^k_n(t) \sum_{\substack{w_3 \in Gl(n,\mathbb{Z}) \backslash \mathbb{Z}^{(n,n)}_+ \\ \binom{w_1}{w_3} \text{primitiv}}}$$

$$\cdot \det(w_1\, t\, w_3')^{\nu}\, b(t[w_3'], f) \det(t[w_3'])^{\frac{n+1}{2} - k - \nu}, \tag{53}$$

wobei t und w_1 der zusätzlichen Bedingung $t[w_1'] = T$ unterliegen.

Auf dem schon in [4] eingeschlagenen Weg wollen wir die Eigenwerte $\lambda(f,\nu)$ genauer beschreiben und insbesondere zeigen, daß diese stets von Null verschieden sind. Dazu benötigen wir die folgenden einer Spitzenform $f(z) = \sum_{t>0} b(t,f)\, e^{2\pi i\, \mathrm{Spur}(tz)} \in \mathfrak{S}^l_n$ und einer Exponentenmatrix $T > 0$ zugeordneten Dirichletreihen:

$$D(s,f,T) := \sum_{w \in Sl(n,\mathbb{Z}) \backslash \mathbb{Z}^{(n,n)}_+} b(T[w']) \det(w)^{-s-l+1} \quad (\mathrm{Re}(s) > n+1)$$

$$L(s,f,T)_k := \underbrace{\sum_{t>0} \sum_{w_1 \in \mathbb{Z}^{(n,n)}_+ / Sl(n,\mathbb{Z})}}_{t[w_1'] = T} a^k_n(t) \sum_{\substack{w_3 \in Sl(n,\mathbb{Z}) \backslash \mathbb{Z}^{(n,n)}_+ \\ \binom{w_1}{w_3} \text{primitiv}}}$$

$$\cdot\, b(t[w_3'], f) \det(t[w_3'])^{-s} \quad \left(\mathrm{Re}(s) > \frac{n+l}{2}\right).$$

Diese Dirichletreihen konvergieren in den angegebenen Bereichen absolut, $\mathbb{Z}^{(n,n)}_+$ bezeichnet die Menge der ganzen n-reihigen Matrizen mit positiver De-

terminante, k ist eine gerade Zahl, $k>n+1$. Ist $f\in\mathfrak{S}_n^l$ insbesondere eine Eigenfunktion aller Heckeoperatoren, so stehen beide Dirichletreihen in einer einfachen Beziehung zu der Standardzetafunktion $D_f(s)=\zeta(s)D_f^*(s)$ von Andrianov [1]; diese ist durch ein Eulerprodukt definiert:

$$D_f^*(s)=\prod_p\left\{\prod_{i=1}^{n}(1-\alpha_{i,p}^{-1}p^{-s})(1-\alpha_{i,p}p^{-s})\right\}^{-1},$$

wobei die $\alpha_{i,p}$ die p-Parameter von f sind (vgl. [1]).

Satz 11. *Es sei* $f(z)=\sum_{t>0}b(t,f)e^{2\pi i\mathrm{Spur}(tz)}\in\mathfrak{S}_n^l$ *eine Eigenfunktion aller Heckeoperatoren,* k *gerade mit* $k>n+1$ *und* T *eine* f*-Kernform* (d.h. $b(T,f)\neq 0$ *und* $b(T[w^{-1}],f)=0$ *für alle* $w\in\mathbb{Z}_n^{(n,n)}$ *mit* $|\det(w)|>1$). *Dann gilt (jeweils im Bereich absoluter Konvergenz):*

a) $$D(s,f,T)=b(T,f)B^n(s,T)D_f^*(s), \tag{54a}$$

b) $$L(s,f,T)_k=\tilde{a}_n^k(T)\sum_{w\in Sl(n,\mathbb{Z})\backslash\mathbb{Z}^{(n,n)}}b(T[w'],f)\det(T[w'])^{-s} \tag{54b}$$

$$=\tilde{a}_n^k(T)\det(T)^{-s}b(T,f)B^n(2s-l+1)D_f^*(2s-l+1). \tag{54c}$$

Beweis. Die Aussage a) ist ein schwieriges Resultat von Andrianov [1], $B^n(s,T)$ ist dabei eine von f unabhängige Dirichletreihe mit Eulerprodukt, die in Satz I.15 genauer beschrieben ist.

Die Aussage b) ist eine Umformulierung der in (I.23), ..., (I.28) durchgeführten Rechnungen. Der modifizierte Fourierkoeffizient $\tilde{a}_n^k(T)$ ist in (I.33) explizit angegeben.

Ist nun $f\in\mathfrak{S}_n^l$ Eigenfunktion aller Heckeoperatoren, $k=l-v>2n$ und gerade, dann gilt nach Definition

$$d_v(t,f)=\tilde{\mu}_{n,v}^l\lambda(f,v)b(t,f) \tag{55}$$

für alle halbganzen $t>0$. Für f-Kernformen $T>0$ gilt andererseits nach (53) und (54c)

$$d_v(T,f)=2A_n^l(2\pi i)^{nv}\det(T)^{\frac{v}{2}}L\left(l-\frac{n+1}{2}-\frac{v}{2},f,T\right)_k$$

$$=2A_n^l(2\pi i)^{nv}\tilde{a}_n^k(T)\det(T)^{\frac{n+1}{2}-k}b(T,f)B^n(k-n,T)D_f^*(l-n-v), \tag{56}$$

insbesondere gilt

$$d_v(T,f)\neq 0.$$

In [4] hatte ich die Vermutung aufgestellt, daß die Größe

$$\beta^*(n,k)=a_n^k(T)\det(T)^{\frac{n+1}{2}-k}B^n(k-n,T)$$

von T unabhängig sei; das ist in der Tat richtig, wie ich in [5] zeigen konnte, es gilt nämlich

$$\beta^*(n,k)=(-1)^{\frac{nk}{2}}2^{n\left(k-\frac{n-1}{2}\right)}\prod_{i=0}^{n-1}\frac{\pi^{k-i/2}}{\Gamma(k-i/2)}\zeta(k-n)\zeta(k)^{-1}\prod_{j=1}^{n}\zeta(2k-2j)^{-1}.$$

Die genaue Gestalt des Eigenwerts $\lambda(f, v)$ ergibt sich nun durch Vergleich von (55) und (56). Zusammenfassend erhält man

Satz 12. *Es seien l und v natürliche Zahlen mit $l-v$ gerade und $l-v>2n$. Dann vermittelt der Operator $S_n^{(v)}$ einen Automorphismus von $\mathfrak{S}_n^l$. Ist $f\in\mathfrak{S}_n^l$ eine Eigenfunktion aller Heckeoperatoren, so wird der Eigenwert $\lambda(f, v)$ gegeben durch*

$$\lambda(f, v)=(\tilde{\mu}_{n,v}^l)^{-1}\, 2\, A_n^l (2\pi i)^{nv}\, \beta^*(n, l-v)\, D_f^*(l-n-v), \tag{57}$$

insbesondere ist dieser Eigenwert stets von Null verschieden.

5. Die Modulformen von Cohen-Zagier

Wir definieren ein Polynom Q_k^v (und entsprechend $\tilde{Q}_k^v$) in $n(2n+1)$ komplexen Variablen $t_{i,j}$ $(1\leqq i\leqq j\leqq 2n)$ vermöge

$$(\mathscr{D}_k^v f_{\mathfrak{T}})\begin{pmatrix} z_1 & 0 \\ 0 & z_4 \end{pmatrix}=Q_k^v(\mathfrak{T})\, e^{\mathrm{Spur}(t_1 z_1+t_4 z_4)}.$$

Dabei werden die Variablen t_{ij} in einer $2n$-reihigen symmetrischen Matrix $\mathfrak{T}$ angeordnet. Das Polynom Q_k^v ist homogen vom Grade nv:

$$Q_k^v(a\mathfrak{T})=a^{nv} Q_k^v(\mathfrak{T}), \qquad a\in\mathbb{C}.$$

Daraus folgt

Bemerkung. Die Abbildung

$$(2\pi i)^{-nv}\mathscr{D}_{k,0}^v\colon \mathfrak{M}_{2n}^k\to(\mathfrak{M}_n^{k+v}\otimes\mathfrak{M}_n^{k+v})^{\mathrm{sym}}$$

führt Modulformen mit rationalen Fourierkoeffizienten in Modulformen mit rationalen Fourierkoeffizienten (bezüglich des Variablenpaares (z_1, z_4)) über.

Für halbganze n-reihige Matrizen $R, T\geqq 0$ definieren wir rationale Zahlen $\alpha_k^v(R, S)$ durch die (endliche) Summe

$$\alpha_k^v(R, S)=\sum_{A\in\mathbb{Z}^{(n,n)}} \tilde{Q}_k^v\begin{pmatrix} R & A \\ A' & T \end{pmatrix} a_{2n}^k\begin{pmatrix} R & A \\ A' & T \end{pmatrix}.$$

Dies sind genau die Fourierkoeffizienten von $(2\pi i)^{-nv}\tilde{\mathscr{D}}_{k,0}^v\Psi_{2n}^k$. Wir erhalten

Satz 13. *Es sei $l-v$ gerade, $l-v>2n$. Für jedes halbganze $R\geqq 0$ ist*

$$\mathscr{C}_n^{l,v}(z, R)=\sum_{T\geqq 0}\alpha_{l-v}^v(R, T)\, e^{2\pi i\,\mathrm{Spur}(Tz)}$$

eine Modulform n-ten Grades vom Gewicht l mit rationalen Fourierkoeffizienten. Für $v>0$ erhält man Spitzenformen, für alle

$$f(z)=\sum_{t>0} b(t, f)\, e^{2\pi i\,\mathrm{Spur}(tz)}\in\mathfrak{S}_n^l$$

gilt

$$\langle f, \mathscr{C}_n^{l,v}(*, R)\rangle=(2\pi i)^{-nv} d_v(f, R).$$

Für $v>0$ erzeugt (unter obigen Bedingungen)

$$\{\mathscr{C}_n^{l,\nu}(*,R)|R \text{ halbganz}, R>0\}$$

den Raum $\mathfrak{S}_n^l$ linear.

Bemerkungen:

1) Für $v=0$ war in [4] gezeigt worden, daß

$$\{\mathscr{C}_n^{l,0}(*,R)|R \text{ halbganz}, R\geqq 0\}$$

sogar ganz $\mathfrak{M}_n^l$ erzeugt (l gerade, $l>2n$).

2) Für $n=1$ erhält man die wohlbekannten Funktionen

$$C_{l,l-1}, C_{l,l-3}, \ldots, C_{l,3} \quad (l \text{ gerade}, l>2)$$

von Cohen-Zagier ([6, 29]). Bis auf eine Konstante stimmt nämlich für gerade l und v, $l-v>2$ die Funktion $\mathscr{C}_1^{l,\nu}(*,r)$ überein mit $C_{l,l-\nu-1}|T(r)$, wobei $T(r)$ den r-ten Heckeoperator bezeichnet.

3) Auf das hier offengelassene Problem der expliziten Beschreibung des Polynoms Q_k^ν hoffe ich an anderer Stelle näher eingehen zu können.

4) An einigen Stellen taucht in der Literatur das Problem auf, Modulformen von ungeradem Gewicht mit ganzrationalen Fourierkoeffizienten zu konstruieren ([22], in [3] wird diese Schwierigkeit dadurch umgangen, daß von vornherein nur gerade Gewichte behandelt werden). Für gerades n, ungerades l, $l-1>2n$ werden durch die Funktionen

$$\mathscr{C}_n^{l,1}(z,R)=\sum_{T>0}\sum_{A\in\mathbb{Z}^{(n,n)}}\det(A)\,a_{2n}^{l-1}\begin{pmatrix}R & A\\ A' & T\end{pmatrix}e^{2\pi i\,\mathrm{Spur}(TZ)}$$

einfache Beispiele für Modulformen ungeraden Gewichts mit rationalen Fourierkoeffizienten von beschränktem Nenner gegeben. Verschwinden alle $\mathscr{C}_n^{l,1}(*,R)$, $R>0$ halbganz, identisch, so folgt $\mathfrak{S}_n^l=\{0\}$. Eine weitere Möglichkeit, solche Modulformen zu konstruieren, wird durch Thetareihen mit harmonischen Koeffizienten gegeben (s.u.).

6. Rationalitätsaussagen für die Standard-Zetafunktion von Andrianov

Der Raum $\mathfrak{S}_n^l$ besitzt eine Basis mit rationalen Fourierkoeffizienten ([3, 7, 22]; für den hier interessierenden Fall großer Gewichte könnte man auch Satz 13 heranziehen). Daher ist für alle Körperautomorphismen σ von $\mathbb{C}$ mit

$$f(z)=\sum_{t>0}b(t,f)\,e^{2\pi i\,\mathrm{Spur}(tz)}$$

auch f^σ eine Spitzenform vom Gewicht l,

$$f^\sigma(z):=\sum_{t>0}b(t,f)^\sigma\,e^{2\pi i\,\mathrm{Spur}(tz)}.$$

Mit ρ sei die komplexe Konjugation bezeichnet. Für $f \in \mathfrak{S}_n^l$ gilt

$$f^\rho(z) = \overline{f(-\bar{z})}.$$

Die Heckeoperatoren vertauschen mit der Operation von $\mathrm{Aut}(\mathbb{C})$:

$$(f^\sigma)|T = (f|T)^\sigma \tag{58}$$

für alle $\sigma \in \mathrm{Aut}(\mathbb{C})$ und alle Heckeoperatoren T ([25, §7]). Mit f ist dann auch f^σ eine Eigenfunktion der Heckeoperatoren.

Einer Idee Garretts [12] folgend, beweisen wir Rationalitätsaussagen für spezielle Werte der Standardzetafunktion $D_f(s)$. Dabei ist von entscheidender Bedeutung, daß $(2\pi i)^{-n\nu}\tilde{\mathscr{D}}_{k,0}^\nu \Psi_{2n}^k$ als Funktion von z_1 und z_4 rationale Fourierkoeffizienten hat. Zur Abkürzung setzen wir

$$\begin{aligned}\lambda^*(f, \nu) &:= (2\pi i)^{-n\nu}\tilde{\mu}_{n,\nu}^l \lambda(f, \nu) \\ &= 2A_n^l \beta^*(n, l-\nu) D_f^*(l-n-\nu).\end{aligned}$$

Für eine Eigenfunktion $f \in \mathfrak{S}_n^l$ aller Heckeoperatoren gilt dann einerseits

$$\left\langle f^\sigma, (2\pi i)^{-n\nu}\tilde{\mathscr{D}}_k^\nu \Psi_{2n}^k \begin{pmatrix} -\bar{z} & 0 \\ 0 & * \end{pmatrix} \right\rangle = \lambda^*(f^\sigma, \nu) f^\sigma. \tag{59}$$

Wir wählen andererseits eine Orthogonalbasis $f_1, \ldots, f_d$ von $\mathfrak{S}_n^l$, die aus Eigenfunktionen der Heckealgebra besteht mit $f_1 = f$.

Dann gilt nach der Garrettschen Hauptformel:

$$\begin{aligned}(2\pi i)^{-n\nu}\tilde{\mathscr{D}}_k^\nu \Psi_{2n}^k \begin{pmatrix} z_1 & 0 \\ 0 & z_4 \end{pmatrix} &= \sum_{i=1}^d \lambda^*(f_i, \nu) \frac{f_i^\rho(z_1) f_i(z_4)}{\langle f_i, f_i \rangle} + \delta_{0,\nu} F(z_1, z_4) \\ &= \sum_{i=1}^d \lambda^*(f_i, \nu) \frac{f_i(z_1) f_i^\rho(z_4)}{\langle f_i, f_i \rangle} + \delta_{0,\nu} F(z_1, z_4) \\ &= \sum_{i=1}^d \lambda^*(f_i, \nu)^{\sigma\rho} \frac{f_i^{\sigma\rho}(z_1) f_i^{\rho\sigma\rho}(z_4)}{\langle f_i, f_i \rangle^{\sigma\rho}} + \delta_{0,\nu} F(z_1, z_4)^{\sigma\rho}.\end{aligned}$$

Dabei wurde benutzt, daß $(2\pi i)^{-n\nu}\tilde{\mathscr{D}}_{k,0}^\nu \Psi_{2n}^k$ ein Element von $(\mathfrak{S}_n^l \otimes \mathfrak{S}_n^l)^{\mathrm{sym}}$ ist mit rationalen Fourierkoeffizienten. Hieraus folgt

$$\left\langle f^\sigma, (2\pi i)^{-n\nu}\tilde{\mathscr{D}}_k^\nu \Psi_{2n}^k \begin{pmatrix} -\bar{z} & 0 \\ 0 & * \end{pmatrix} \right\rangle = \sum_{i=1}^d \lambda^*(f_i, \nu)^\sigma \frac{\langle f^\sigma, f_i^{\rho\sigma\rho} \rangle}{\langle f_i, f_i \rangle^\sigma} f_i^\sigma(z). \tag{60}$$

Dabei ist zu beachten, daß im Falle $\nu = 0$ mit F auch F^σ in $(\mathfrak{S}_n^l)^\perp \otimes (\mathfrak{S}_n^l)^\perp$ gelegen ist (zumindest für große Gewichte, vgl. [13, Theorem 2.5]).

Vergleich von (59) und (60) ergibt

$$\lambda^*(f^\sigma, \nu) = \lambda^*(f, \nu)^\sigma \frac{\langle f^\sigma, f^{\rho\sigma\rho} \rangle}{\langle f, f \rangle^\sigma}.$$

Man rechnet direkt nach, daß $\lambda^*(f, \nu)$ von der Gestalt

$$(\text{rationale Zahl}) \cdot \pi^{-d(l,n,\nu)} D_f(l-n-\nu)$$

ist, wobei

$$d(l, n, v)=2nl+l-v-vn-\frac{3n(n+1)}{2}.$$

Wir erhalten

Satz 14. *Es sei $l-v$ gerade, $l-v>2n$, $f\in\mathfrak{S}_n^l$ eine Eigenfunktion aller Heckeoperatoren. Dann gilt für alle $\sigma\in\mathrm{Aut}(\mathbb{C})$ mit $d=d(l,n,v)$:*

$$\pi^{-d}\frac{D_{f^\sigma}(l-n-v)}{\langle f^\sigma, f^{\rho\sigma\rho}\rangle}=\left(\pi^{-d}\frac{D_f(l-n-v)}{\langle f,f\rangle}\right)^\sigma. \tag{61}$$

Bemerkungen:

1) Nimmt man zusätzlich an, daß f total reelle algebraische Fourierkoeffizienten hat, so vereinfacht sich wegen $f^{\rho\sigma\rho}=f^\sigma$ die obige Formel, es folgt insbesondere, daß dann der Quotient

$$\pi^{-d}\frac{D_f(l-n-v)}{\langle f,f\rangle}$$

in dem von den Fourierkoeffizienten von f erzeugten algebraischen Zahlkörper liegt.

Im übrigen überlegt man sich leicht, daß in $\mathfrak{S}_n^l$ jeder Eigenraum der Heckealgebra eine Basis besitzt, deren Elemente total reelle algebraische Fourierkoeffizienten haben.

2) Ähnliche Sätze sind für gerades n von Harris [14] und Sturm [25] bewiesen worden, für $n=1$ ebenfalls von Sturm [24] und Zagier [29]. Für ungerades $n>1$ scheint Satz 14 aber neu zu sein.

7. Harmonische Formen

Eine Polynomfunktion $P\colon \mathbb{C}^{(m,n)}\to\mathbb{C}$ heißt harmonische Form vom Gewicht $v\in\mathbb{N}$, falls gilt

a) $P(XA)=\det(A)^v P(X)$ für alle $A\in Gl(n,\mathbb{C})$, $X\in\mathbb{C}^{(m,n)}$, (62a)

b) $\Delta P=\sum_{i,j}\frac{\partial^2 P}{\partial^2 x_{ij}}=0.$ (62b)

Bei festen m, n und v bilden diese Funktionen einen endlich dimensionalen $\mathbb{C}$-Vektorraum $H_v(m,n)$, dessen Dimension man wie folgt berechnen kann: Nach [16] ist $H_v(m,n)$ Darstellungsraum einer explizit angebbaren irreduziblen Darstellung der orthogonalen Gruppe $O(m,\mathbb{C})$, daher kann man die gesuchte Dimension aus den Dimensionsformeln von Weyl [28] ablesen (man vergleiche dazu auch [27, §6]). Eine einfache Folgerung aus [16, 28] ist

Bemerkung. Es sei m gerade, $m=2(n+t)$, $t\geqq 0$; dann gilt mit einer geeigneten Konstanten c

$$\dim H_v(m,n)\sim c\,v^{\frac{1}{2}n(n+1)+n(2t-1)}\quad\text{für } v\gg 0. \tag{63}$$

(Die Einzelheiten seien dem Leser überlassen.)

Es sei nun S eine positiv-definite quadratische Form in m Variablen; eine Polynomfunktion Q in der Variablen $X=X^{(m,n)}$ heiße S-harmonisch, falls die Funktion $X\mapsto Q(S^{-\frac{1}{2}}X)$ im gewöhnlichen Sinne harmonisch ist, also (62b) erfüllt; es ist dann Q genau dann S-harmonisch, wenn

$$\mathrm{Spur}(\partial X' S^{-1}\partial X)Q=0.$$

Daraus folgt, daß für $\sigma\in\mathrm{Aut}(\mathbb{C})$ mit Q auch Q^σ eine S-harmonische Form vom Gewicht v ist, falls S rationale Koeffizienten hat; dabei gehe Q^σ aus Q durch Anwendung von σ auf die Koeffizienten von Q hervor.

Nach einer bekannten Schlußweise (vgl. [23, Lemma 3]) folgt daraus:

Bemerkung. Die quadratische Form $S>0$ habe rationale Koeffizienten. Dann besitzt der Vektorraum $H_v^S(m,n)$ aller S-harmonischen Formen vom Gewicht v eine Basis, die aus Polynomen mit rationalen Koeffizienten besteht.

Wir benötigen noch eine Charakterisierung harmonischer Formen durch die Gaußtransformation. Diese ist für ein Polynom P in der Variablen $X=X^{(m,n)}$ definiert durch

$$\hat{P}(X)=\int_{\mathbb{R}^{(m,n)}} P(U+X)\,e^{-\pi\mathrm{Spur}(U'U)}\,dU.$$

Nach [19] gilt:

Eine Polynomfunktion P auf $\mathbb{C}^{(m,n)}$, die (62a) erfüllt, ist genau dann harmonisch, wenn $P=\hat{P}$.

Ziel dieses Paragraphen ist der Beweis von

Satz 15. *Für jedes feste $X_0\in\mathbb{C}^{(m,n)}$ ist die Polynomfunktion*

$$X\mapsto P_{X_0}^v(X):=Q_{m/2}^v((X_0,X)'(X_0,X))$$

eine harmonische Form vom Gewicht v in der Variablen $X=X^{(m,n)}$; außerdem ist $P_*^v(*)$ symmetrisch:

$$P_{X_0}^v(X)=P_X^v(X_0). \tag{64}$$

Beweis. Die Eigenschaft (62a) braucht nur für $A\in Gl(n,\mathbb{R})$ nachgewiesen werden; für reelles A ist aber (62a) ebenso wie die Symmetrieeigenschaft (64) eine unmittelbare Folge der Vertauschungseigenschaften (29a) und (29b) des Operators $\mathscr{D}_{m/2}^v$. Es braucht also nur noch (62b) bzw. $\hat{P}_{X_0}^v=P_{X_0}^v$ gezeigt werden.

Man betrachte für $z_4\in\mathfrak{H}_n$, $W\in\mathbb{C}^{(m,n)}$ die Funktion

$$f(z_4,W)=\sum_{G\in\mathbb{Z}^{(m,n)}} P_{X_0}^v(G+W)\,e^{\pi i\mathrm{Spur}((G+W)'(G+W)z_4)}.$$

Man entwickle nun f in eine Fourierreihe bezüglich W

$$f(z_4,W)=\sum_{G\in\mathbb{Z}^{(m,n)}} A(G,z_4)\,e^{2\pi i\mathrm{Spur}(G'W)}$$

und berechne die Fourierkoeffizienten $A(G,z_4)$ auf zwei verschiedene Weisen:

Nach [11, S. 161] gilt für $z_4=iy_4$:

$$A(G,iy)=\det(y_4)^{-\frac{m}{2}}e^{-\pi\mathrm{Spur}(G'Gy_4^{-1})}\int_{\mathbb{R}^{(m,n)}} P_{X_0}^v(Uy_4^{-\frac{1}{2}})\,e^{-\pi\mathrm{Spur}(1_m[U+iGy_4^{-\frac{1}{2}}])}\,dU.$$

Wegen

$$P^v_{X_0}(U\, y_4^{-\frac{1}{2}}) = \det(y_4)^{-\frac{v}{2}} P^v_{X_0}(U)$$

erhält man nach einer Substitution

$$A(G, i y_4) = \det(y_4)^{-\frac{m+v}{2}} e^{-\pi \mathrm{Spur}(G' G y_4^{-1})} \hat{P}^v_{X_0}(-iG y_4^{-\frac{1}{2}}). \tag{65}$$

Andererseits betrachte man für $Z \in \mathfrak{H}_{2n}$, $W \in \mathbb{C}^{(m,n)}$ die Funktion

$$\begin{aligned} g(Z, W) &= \sum_{G \in \mathbb{Z}^{(m,n)}} e^{\pi i \mathrm{Spur}(Z[(X_0, G+W)'])} \\ &= e^{\pi i \mathrm{Spur}(z_1[X_0'] + 2 z_2' X_0' W)} \vartheta_{2W, 2X_0 z_2}(z_4), \end{aligned}$$

wobei wie üblich mit $\vartheta_{A,B}$ die Thetakonstante

$$\vartheta_{A,B}(z_4) = \sum_{G \in \mathbb{Z}^{(m,n)}} e^{\pi i \mathrm{Spur}(z_4[G' + \frac{1}{2}A'] + B'G)}$$

bezeichnet wird ($A, B \in \mathbb{C}^{(m,n)}$).

Die Reziprozitätsformel

$$\vartheta_{A,B}(-z_4^{-1}) = e^{-\frac{\pi}{2}\mathrm{Spur}(A'B)} \det\left(\frac{z_4}{i}\right)^{\frac{m}{2}} \vartheta_{B,-A}(z_4)$$

liefert nach kurzer Rechnung

$$\begin{aligned} g(Z, W) &= \det\left(\frac{z_4}{i}\right)^{-\frac{m}{2}} \sum_{G \in \mathbb{Z}^{(m,n)}} e^{\pi i \mathrm{Spur}(I^{\downarrow} \langle Z \rangle [(X_0, G)'] + 2G'W)} \\ &= i^{\frac{mn}{2}} \Big(\sum_{G \in \mathbb{Z}^{(m,n)}} e^{\pi i \mathrm{Spur}(Z[(X_0, G)'] + 2G'W)} \Big)\Big|_{\frac{m}{2}} I^{\downarrow}. \end{aligned}$$

Anwendung des Operators $\mathscr{D}^v_{\frac{m}{2},0}$ ergibt

$$\begin{aligned} &\mathscr{D}^v_{\frac{m}{2},0}\, g(*, W) \begin{pmatrix} z_1 & 0 \\ 0 & z_4 \end{pmatrix} \\ &\quad = e^{\pi i \mathrm{Spur}(z_1[X_0'])} \sum_{G \in \mathbb{Z}^{(m,n)}} P^v_{X_0}(G+W) e^{\pi i \mathrm{Spur}(z_4[G'+W'])} = e^{\pi i \mathrm{Spur}(z_1[X_0'])} f(z_4, W). \end{aligned}$$

Wegen (29b) ist das aber gleich

$$i^{\frac{mn}{2}} \det(z_4)^{-\frac{m}{2}-v} \sum_{G \in \mathbb{Z}^{(m,n)}} P^v_{X_0}(G)\, e^{\pi i \mathrm{Spur}(z_1[X_0'] - z_4^{-1}[G'] + 2G'W)},$$

es folgt

$$A(G, i y_4) = i^{\frac{mn}{2}} \det(z_4)^{-\frac{m}{2}-v} e^{-\pi i \mathrm{Spur}(z_4^{-1}[G'])} P^v_{X_0}(G).$$

Durch Vergleich mit (65) ergibt sich für $z_4 = i \cdot 1_n$:

$$\hat{P}^v_{X_0}(-iG) = i^{-vn} P^v_{X_0}(G) = P^v_{X_0}(-iG) \quad \text{für alle } G \in \mathbb{Z}^{(m,n)}.$$

Da nun $\hat{P}^v_{X_0}$ und $P^v_{X_0}$ zwei Polynomfunktionen auf $\mathbb{C}^{(m,n)}$ sind, die auf dem Gitter $i\mathbb{Z}^{(m,n)}$ übereinstimmen, folgt

$$\hat{P}^v_{X_0} = P^v_{X_0}.$$

Bemerkung. Für den Fall $n=1$ werden ähnliche Sätze von Waldspurger [26] sowie Eichler-Zagier [8] auf eher algebraischem Wege bewiesen; sie benutzen nämlich eine explizite Beschreibung der Polynome $P^v_{X_0}(X)$ durch Gegenbauer-Polynome. Die Tatsache, daß man auf diesem Wege harmonische Formen erhält, erscheint dort als ein glücklicher Zufall (bzw. als „morally certain" [8, Theorem 7.2]).

8. Thetareihen mit harmonischen Koeffizienten

Für eine harmonische Form $P \in H_v(m,n)$, eine positiv definite quadratische Form S in m Variablen und $z \in \mathfrak{H}_n$ bildet man die Thetareihe n-ten Grades

$$\vartheta^n_{S,P}(z) = \sum_{G \in \mathbb{Z}^{(m,n)}} P(S^{\frac{1}{2}} G)\, e^{\pi i \operatorname{Spur}(S[G]z)}.$$

Ist S gerade und unimodular, so erhält man eine Modulform (für $v>0$ eine Spitzenform) vom Gewicht $\frac{m}{2}+v$ [11]. Notwendig und hinreichend für die Existenz eines solchen S ist $m \equiv 0 \bmod 8$.

Der Raum $B_{n,l}(m)$, der von allen $\vartheta^n_{S,P}$ (S gerade, unimodular und m-reihig, $P \in H_{l-\frac{m}{2}}(m,n)$) aufgespannt wird, besitzt zwei wichtige Invarianzeigenschaften:

1) $B_{n,l}(m)$ ist invariant unter allen Heckeoperatoren.

2) Für alle $\sigma \in \operatorname{Aut}(\mathbb{C})$ gilt $B_{n,l}(m)^\sigma = B_{n,l}(m)$.

Die erste Eigenschaft ist von Freitag [10] unter Verwendung seiner Resultate [9] über vektorwertige singuläre Modulformen gezeigt worden.

Die zweite Invarianzeigenschaft folgt unmittelbar aus der entsprechenden Eigenschaft von $H^S_v(m,n)$, dem Raum der S-harmonischen Formen, $S = S^{(m)} > 0$ gerade und unimodular.

Unter dem *Basisproblem* versteht man die Frage, unter welchen Bedingungen an n, m und l der Teilraum $B_{n,l}(m)$ mit dem ganzen Raum $\mathfrak{M}^l_n$ $\left(\text{für } l=\frac{m}{2}\right)$ beziehungsweise $\mathfrak{S}^l_n$ $\left(l>\frac{m}{2}\right)$ übereinstimmt.

Mit dem Basisproblem verbunden ist eine allgemeinere Frage:

Wegen (58) operieren die Heckeoperatoren auf $\mathfrak{S}^l_n(\mathbb{Q})$, dem Raum der Spitzenformen mit rationalen Fourierkoeffizienten. Mit $\mathscr{T}^l_n$ sei die natürliche Darstellung der abstrakten $\mathbb{Q}$-Heckealgebra auf $\mathfrak{S}^l_n(\mathbb{Q})$ bezeichnet. Es ist ein wichtiges – leider kaum direkt angreifbares – Problem in der Arithmetik der Modulformen, den Raum $\mathfrak{S}^l_n(\mathbb{Q})$ als $\mathscr{T}^l_n$-Modul zu studieren. Die Frage nach der Reduzibilität der Darstellung $\mathscr{T}^l_n$ ist äquivalent zur Frage nach der Existenz nichttrivialer Teilräume $\mathfrak{A}$ von $\mathfrak{S}^l_n$, die invariant unter den Heckeoperatoren sind und $\mathfrak{A}^\sigma = \mathfrak{A}$ für alle $\sigma \in \operatorname{Aut}(\mathbb{C})$ erfüllen.

Der Zusammenhang mit dem Basisproblem ergibt sich aus der (trivialen)

Bemerkung. Falls $\mathscr{T}_n^l$ irreduzibel ist, folgt für alle m mit $\frac{m}{2}<l$:

$$B_{n,l}(m)=\{0\} \quad \text{oder} \quad B_{n,l}(m)=\mathfrak{S}_n^l.$$

Interessant ist in dieser Hinsicht der Fall $n\equiv 0 \bmod 4$, $m=2n$:

Aus dem asymptotischen Verhalten (63) der Dimension von $H_\nu(2n,n)$ folgt jedenfalls, daß

$$B_{n,l}(2n)\subsetneqq \mathfrak{S}_n^l \quad \text{für } l\gg 0$$

gelten muß, da bekanntlich die Dimension von $\mathfrak{S}_n^l$ asymptotisch mit l^N, $N=\frac{1}{2}n(n+1)$ anwächst.

Die Frage, wann eine Thetareihe $\vartheta_{S,P}^n$ nicht identisch verschwindet, stellt im allgemeinen ein schwieriges Problem dar. Für $m\geqq 2n$ hat Raghavan – einer Idee von Maaß folgend – in [21] gezeigt, daß es für eine geeignete natürliche Zahl $h=h(S,n)$ unter h aufeinander folgenden Gewichten $\nu, \nu+1, \ldots, \nu+h-1$ mindestens ein Gewicht $\nu+j$ gibt, für das mindestens eine Thetareihe $\vartheta_{S,P}^n$, $P\in H_{\nu+j}(m,n)$ nicht verschwindet (Raghavan hat dieses Resultat nur für $\nu=1$ formuliert, sein Beweis ist aber für alle ν gültig).

Daraus erhält man

Satz 15. *Es sei $n\equiv 0 \bmod 4$. Dann gibt es eine natürliche Zahl h derart, daß es für $l\gg 0$ unter den Gewichten $l, l+1, \ldots, l+h-1$ mindestens ein Gewicht l' gibt mit*

$$\{0\}\neq B_{n,l'}(2n)\subsetneqq \mathfrak{S}_n^{l'},$$

insbesondere ist für diese Gewichte die Darstellung $\mathscr{T}_n^{l'}$ reduzibel.

Für $n=1$ hat J. Buhler durch numerische Rechnungen festgestellt, daß $\mathscr{T}_1^l$ für l gerade, $l\leqq 200$ in der Tat irreduzibel ist (vgl. [18, S. 252]).

Für $n=2$ jedoch ist $\mathscr{T}_2^l$ für gerades l im allgemeinen reduzibel:

Die Spezialschar von Maaß stellt nämlich für großes gerades l einen nichttrivialen Teilraum von $\mathfrak{S}_2^l$ dar, der die geforderten Invarianzeigenschaften besitzt [2].

9. Das Basisproblem

Ziel dieses Abschnitts ist der Nachweis von „$B_{n,l}(m)=\mathfrak{S}_n^l$“ unter den Nebenbedingungen $m>4n$, $l>\frac{m}{2}$.

Wir beginnen nach dem Vorbild von [4] mit dem Siegelschen Hauptsatz:

Für $k>2n+1$, $2k=m\equiv 0 \bmod 8$ gilt

$$\Psi_{2n}^k(Z)=\gamma(m)^{-1}\sum_{i=1}^{h(m)} \#\,\mathrm{Aut}(S_i)^{-1}\,\vartheta_{S_i,1}^{2n}(Z). \tag{65}$$

Dabei durchläuft S_i $(1\leqq i\leqq h(m))$ ein vollständiges Repräsentantensystem aller Klassen gerader unimodularer quadratischer Formen $S=S^{(m)}>0$,

$$\mathrm{Aut}(S)=\{U\in Gl(m,\mathbb{Z})\mid S[U]=S\},$$
$$\gamma(m)=\sum_{i=1}^{h(m)}\#\,\mathrm{Aut}(S_i)^{-1}.$$

Wir wenden nun den Operator $\tilde{\mathscr{D}}_k^\nu$ auf beide Seiten von (65) an. Dabei ist zu beachten:

$$\begin{aligned}(\tilde{\mathscr{D}}_k^\nu\,\vartheta_{S,1}^{2n})\begin{pmatrix}z_1&0\\0&z_4\end{pmatrix}&=(\pi i)^n\sum_{G,H\in\mathbb{Z}^{(m,n)}}\tilde{Q}_k^\nu((G,H)'S(G,H))\,e^{\pi i\,\mathrm{Spur}(S[G]z_1+S[H]z_4)}\\&=(\pi i)^n\sum_{G,H\in\mathbb{Z}^{(m,n)}}\tilde{P}^\nu_{S^{1/2}G}(S^{\frac12}H)\,e^{\pi i\,\mathrm{Spur}(S[G]z_1+S[H]z_4)}\\&=(\pi i)^n\sum_{G\in\mathbb{Z}^{(m,n)}}\vartheta^n_{S,\tilde{P}^\nu_{S^{1/2}G}}(z_4)\,e^{\pi i\,\mathrm{Spur}(S[G]z_1)}.\end{aligned}$$

Für $f\in\mathfrak{S}_n^{k+\nu}$ ist die Funktion

$$X\mapsto\mathscr{P}^\nu_{S,f}(X):=\langle\vartheta^n_{S,\tilde{P}^\nu_X},f\rangle$$

eine harmonische Form von Gewicht ν in der Variablen $X=X^{(m,n)}$, und es gilt

$$\left\langle f,(\tilde{\mathscr{D}}_k^\nu\,\vartheta_{S,1}^{2n})\begin{pmatrix}-\bar z&0\\0&*\end{pmatrix}\right\rangle=(\pi i)^{n\nu}\,\vartheta^n_{S,\mathscr{P}^\nu_{S,f}}(z).$$

Es sei nun $f\in\mathfrak{S}_n^{k+\nu}$ eine Eigenfunktion aller Heckeoperatoren.

Dann gilt einerseits

$$\left\langle f,(\tilde{\mathscr{D}}_k^\nu\,\Psi_{2n}^k)\begin{pmatrix}-\bar z&0\\0&*\end{pmatrix}\right\rangle=\tilde{\mu}^{k+\nu}_{n,\nu}\,\lambda(f,\nu)f(z),\tag{66}$$

wegen (65) ist dies andererseits gleich

$$(\pi i)^{n\nu}\gamma(m)^{-1}\sum_{i=1}^{h(m)}\#\,\mathrm{Aut}(S_i)^{-1}\,\vartheta^n_{S_i,\mathscr{P}^\nu_{S_i,f}}(z).\tag{67}$$

Daraus erhält man als vielleicht wichtigstes Ergebnis dieser Arbeit

Satz 16. *Es sei $m\equiv 0 \bmod 8$, $m>4n$. Dann gilt für alle $l>\frac{m}{2}$*

$$B_{n,l}(m)=\mathfrak{S}_n^l.$$

Darüber hinaus hat man für jede Eigenform $f\in\mathfrak{S}_n^l$ die folgende explizite Darstellung $\left(\nu=l-\frac{m}{2}\right)$

$$\tilde{\mu}^l_{n,\nu}\,\lambda(f,\nu)f=(\pi i)^{n\nu}\gamma(m)^{-1}\sum_{i=1}^{h(m)}\#\,\mathrm{Aut}(S_i)^{-1}\,\vartheta^n_{S_i,\mathscr{P}^\nu_{S_i,f}}.\tag{68}$$

Bemerkungen. 1) Für $n=1$ ist ein solcher Satz erstmals von Waldspurger [26] gezeigt worden (man vergleiche auch [8]). Die Darstellung (68) einer Eigen-

form tritt in [8] und [26] nicht explizit auf, kann aber aus den dort angegebenen Resultaten leicht abgeleitet werden.

Den Fall $v=0$, $n \geqq 1$ hatte ich schon in [4] behandelt.

2) Verzichtet man auf die explizite Darstellung (68), so kann man obigen Satz – ohne die Hauptformel von Garrett zu benutzen – auch wie folgt bewiesen:

Nach [10] ist $B_{n,l}(m)$ und damit auch $B_{n,l}(m)^{\perp}$ invariant unter allen Heckeoperatoren. Eine Eigenform $0 \neq f \in B_{n,l}(m)^{\perp}$ müßte wegen (65) die Gleichung

$$\left\langle f, \tilde{\mathscr{D}}^{v}_{m/2} \Psi^{m/2}_{2n} \begin{pmatrix} -\bar{z} & 0 \\ 0 & * \end{pmatrix} \right\rangle = 0$$

erfüllen $\left(v = l - \frac{m}{2}\right)$ – im Widerspruch zu (56).

Erratum zu [4]

S. 32, Zeile 3: „$a^{k-r/2}_{n-r}(R_4 - T^{-1}[R_2])$“ statt „$a^{k-r/2}_{n-r}(R_4 - T[R_2])$“.

S. 34, Zeile 15: „2 für $s'>0$“ statt „1/2 für $s'>0$“, die nachfolgenden Formeln sind entsprechend zu korrigieren.

S. 37, Zeile 4: „$w \in Sl(n,\mathbb{Z}) \backslash \mathbb{Z}^{(n,n)}_{+}$“ statt „$w \in GL(n,\mathbb{Z}) \backslash \mathbb{Z}^{(n,n)}_{n}$“.

S. 41, Zeile 30: „$\det(t[w'_3])$“ statt „$\det(t[w'_1])$“.

S. 41, Zeile 31 ist zu ersetzen durch:
„wobei t und w_1 noch der zusätzlichen Bedingung $t[w'_1] = T$ unterliegen“.

S. 42, Zeile 9: „$r-s=s=1$“ statt „$r-s=s$“.

Literatur

1. Andrianov, A.N.: The multiplicative arithmetic of Siegel modular forms. Russian Math. Surveys **34**, 75–148 (1979)
2. Andrianov, A.N.: Modular descent and the Saito-Kurokawa conjecture. Invent. Math. **53**, 267–280 (1979)
3. Baily, W.L.: Automorphic forms with integral Fourier coefficients. In: Several Complex Variables I Maryland 1970, pp. 1–8. Lecture Notes in Math. **155**. Berlin-Heidelberg-New York: Springer 1970
4. Böcherer, S.: Über die Fourier-Jacobi-Entwicklung Siegelscher Eisensteinreihen. Math. Z. **183**, 21–46 (1983)
5. Böcherer, S.: Über die Fourierkoeffizienten Siegelscher Eisensteinreihen. Manuscripta math. **45**, 273–288 (1984)
6. Cohen, H.: Sums involving the values at negative integers of L-functions of quadratic characters. Math. Ann. **217**, 271–285 (1975)
7. Eichler, M.: Zur Begründung der Theorie der automorphen Funktionen in mehreren Variablen. Aequationes Math. **3**, 93–111 (1969)
8. Eichler, M., Zagier, D.: On the theory of Jacobi forms I. Preprint MPI Bonn 1983
9. Freitag, E.: Thetareihen mit harmonischen Koeffizienten zur Siegelschen Modulgruppe. Math. Ann. **254**, 27–51 (1980)

10. Freitag, E.: Die Wirkung von Heckeoperatoren auf Thetareihen mit harmonischen Koeffizienten. Math. Ann. **258**, 419–440 (1982)
11. Freitag, E.: Siegelsche Modulfunktionen. Grundlehren der mathematischen Wissenschaften 254. Berlin-Heidelberg-New York: Springer 1983
12. Garrett, P.B.: Pullbacks of Eisenstein series; applications. Preprint 1980. Eine revidierte Fassung erscheint in: Proceedings of the Katata Conference (1983) on Automorphic Forms in Several Variables. Progress in Math.
13. Harris, M.: The rationality of holomorphic Eisenstein series. Invent. Math. **63**, 303–310 (1981)
14. Harris, M.: Special values of zeta functions attached to Siegel modular forms. Ann. Sci. E.N.S. **14**, 77–120 (1981)
15. Hecke, E.: Analytische Arithmetik der positiven quadratischen Formen. In: Mathematische Werke, S. 789–898. Göttingen: Vandenhoeck und Ruprecht 1970
16. Kashiwara, M., Vergne, M.: On the Segal-Shale-Weil representations and harmonic polynomials. Invent. Math. **44**, 1–47 (1978)
17. Klingen, H.: Zum Darstellungssatz für Siegelsche Modulformen. Math. Z. **102**, 30–43 (1967)
18. Kohnen, W.: Modular forms of half-integral weight on $\Gamma_0(4)$. Math. Ann. **248**, 249–266 (1980)
19. Maaß, H.: Spherical functions and quadratic forms. J. Indian Math. Soc. **20**, 117–162 (1956)
20. Maaß, H.: Zetafunktionen mit Größencharakteren und Kugelfunktionen. Math. Ann. **134**, 1–32 (1957)
21. Raghavan, S.: Cusp forms of degree 2 and 3. Math. Ann. **224**, 149–156 (1976)
22. Shimura, G.: On the Fourier coefficients of modular forms of several variables. Nachr. Akadem. Wiss. Göttingen 1975, 261–268
23. Shimura, G.: The critical values of certain zeta functions associated with modular forms of half integral weight. J. Math. Soc. Japan **33**, 649–671 (1981)
24. Sturm, J.: Special values of zeta functions, and Eisenstein series of half integral weight. Amer. J. Math. **102**, 219–240 (1980)
25. Sturm, J.: The critical values of zeta functions associated to the symplectic group. Duke Math. J. **48**, 327–350 (1981)
26. Waldspurger, J.L.: Engendrement par des séries theta de certains espaces de formes modulaires. Invent. Math. **50**, 135–168 (1979)
27. Weissauer, R.: Vektorwertige Siegelsche Modulformen kleinen Gewichtes. J. Reine Angew. Math. **343**, 184–202 (1983)
28. Weyl, H.: The classical groups. Princeton: University Press 1946
29. Zagier, D.: Modular forms whose Fourier coefficients involve zeta-functions of quadratic fields. In: Modular Functions of One Variable VI, pp. 105–169. Lecture Notes in Math. **627**. Berlin-Heidelberg-New York: Springer 1977

Eingegangen am 2. März 1984

Math. Z., 189, 111-129 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Maximal Subgroups of the Finite Orthogonal and Unitary Groups Stabilizing Anisotropic Subspaces

Roger H. Dye

School of Mathematics, The University, Newcastle upon Type, NE1 7RU, England

1. Introduction

Let V be a vector space of dimension $n \geqq 3$ over a field K. Suppose that G is one of the following groups acting on V: the general linear group $GL_n(K)$, the symplectic group $Sp_n(K)$, an orthogonal group $O_n(K)$, a unitary group $U_n(K)$. Suppose, further, that in the last three cases the form f fixed by G, respectively alternating, quadratic and hermitian, is non-degenerate. Of course n must be even if G is $Sp_n(K)$, or if G is $O_n(K)$ and K has characteristic 2. Let W be a subspace of V and consider the cases $G \neq GL_n(K)$. Then W has a conjugate subspace $W^\perp$ with respect to f. We say that W is totally singular if the restriction of f to W is the zero form, and that W is non-isotropic if the restriction to W, and hence also to $W^\perp$, is non-degenerate. If f is quadratic or hermitian then a singular vector is a zero of f; we say that W is *anisotropic* if W contains no singular vectors other than the zero vector. When f exists two subspaces are called isomorphic if the restrictions of f to them are equivlent: by Witt's theorem the isomorphism classes are the orbits of subspaces under G. Our use of terms is one of those standardly encountered.

Suppose that the following situation occurs. V can be written as the direct sum

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_s, \tag{1}$$

where $s \geqq 2$ and the V_i all have the same dimension r. Suppose, in the cases $G \neq GL_n(K)$, that the V_i are all isomorphic, non-isotropic, and pairwise conjugate. Let F be the stabilizer in G of the set of subspaces $\{V_1, V_2, \ldots, V_s\}$. King has asked the question: when is F a maximal subgroup of G? He has been able [13, pp. 351, 352, 373], [15, p. 416], [16, 193], to answer his question except in the following two circumstances:

(I) $G = GL_n(K)$ and $r = 1$;

(II) G is $O_n(K)$ or $U_n(K)$ and the V_i are anisotropic.

He shows that, with these cases excepted, F is nearly always maximal in G. The exceptions to the maximality all occur when F is strictly contained in the

stabilizer of a simple geometrical configuration, and this can happen only for the smallest finite fields. To find techniques that will cope with the case of anisotropic V_i in general is probably very difficult. For the action of F on V is determined by the action of the orthogonal or unitary group of the restriction of f to V_i. The structure of groups of anisotropic forms depends essentially on K and is hard to elucidate; see the remarks of [3, pp. 34, 51, 71] and [4, p. 71 et seq]. In this paper we settle completely King's omitted cases when K is finite, perhaps the situation when the maximal subgroups of G are of greatest interest to group theorists. This also explains our restriction $n \geqq 3$. For all the subgroups of $PSL_2(q)$ have long been known [2, pp. 285, 286], the subgroups of $PGL_2(q)$ are subgroups of $PSL_2(q^2)$, a subgroup of $GL_2(q)$ fixing a direct sum contains all the scalar maps, and the orthogonal and unitary subgroups of $PGL_2(q)$ appear in other simple guises.

Suppose, henceforth, that K is the finite field $GF(q)$. Then there are the following possibilities for isomorphism type of anisotropic subspaces (see [2, pp. 158, 197], [4, p. 16], or [9, pp. 7, 12]): (i) if f is quadratic one type of dimension 2, and one or two types of dimension 1 according as q is even or odd; (ii) if f is hermitian one type of dimension 1. Each of these types of anisotropic subspace is non-isotropic except for the type of dimension 1 when f is quadratic and q is even.

From (i) it follows that if f is quadratic then one of the possibilities for (II) occurs when $r=2$ so that $n=2s$. We prove that (Theorem 2) in this case F is maximal in $G=O_n(q)$ except when $q=3$ or $q=s=2$, and identify the type of f. If $q=s=2$ then $F=G=O_4(2)$, while if $q=3$ then F lies strictly in the stabilizer of a direct sum of type (II) of 1-dimensional subspaces. The proof uses the action of symmetries of $O_n(q)$, a maximality result of King for the stabilizer of a subspace, and the fact (Theorem 1) that the stabilizer in $O_3(q)$ of an anisotropic 2-dimensional subspace is maximal in $O_3(q)$, except when $q=2$ when it is $O_3(2)$: the case of even q when $O_3(q)$ fixes a nonsingular but degenerate form necessarily intrudes here. From (ii), the only other possibility for (II) when f is quadratic is to have q odd and $r=1$. Then (see §2.5 below) f is a constant times a quadratic form $\tilde{f}$ defined over $GF(p)$ where p is the characteristic of $K=GF(q)$. We have $F \leqq O_n(p) \leqq O_n(q)$, where $O_n(p)$ is the group of $\tilde{f}$, and the V_i determine, on restriction to $GF(p)$, a direct sum of type (II) for $\tilde{f}$ with stabilizer F in $O_n(p)$. We show (Theorem 3) that, apart from the case $n=p=3$ when $F=G=O_3(3)$, if $n \neq 4$ then F is maximal in $O_n(p)$ if and only if $p \equiv \pm 3 \pmod 8$, and we determine the subgroups of $O_n(p)$ containing F when $p \equiv \pm 1 \pmod 8$. *This is a very surprising result:* $|F| = 2^n \cdot n!$, independent of p, yet p can be as large as we like, so $|O_n(p)|$ can be enormous but have a tiny maximal subgroup. If $n=4$ and $p \neq 3$ then F is never maximal in $O_n(p)$; this is due to the existence of triads of desmic tetrahedra in $PG(3, p)$. If $n=4$ and $p=3$ then $PO_4(3)$ *is* the group of a triad of desmic tetrahedra and F is maximal in $O_4(3)$; a fluke. Our proof relies heavily on Wagner's catalogue of irreducible groups generated by reflections [19, p. 239], [20, p. 191], [21, p. 475].

If f is hermitian then, by (ii), the only possibility for (II) is to have $r=1$. We show (Theorem 5) that F is maximal in $G=U_n(q)$: here, of course, q must be a square for the set-up to be possible. This time the proof uses Wagner's

description [18, p. 387] of groups generated by homolgies of order $m>2$. This same description allows us to deal with (I). There F is maximal in $G=GL_n(K)$ provided q is not 2, 3 or 4. If $q=2$ then F fixes a 1-dimensional subspace: a bigger configuration giving maximal subgroups of $O_n(2)$ and $Sp_n(2)$ is discussed in [7]. If $q=3$ then F is the stabilizer of the V_i in an $O_n(3)$, and if $q=4$ then F is the stabilizer of the V_i in a $U_n(4)$. We have already described when F is maximal in these groups. Thus, although there are two types of proofs the description of the whole situation fits neatly together.

Key [10, 11] has determined, in relation to (I), when $F\cap SL_n(q)$ is maximal in $SL_n(q)$-actually she considers the projective counterpart but we have already mentioned that it is easy to pass from one setting to the other. She is able to establish maximality for $q>4$ except for the cases when q is odd and $PSL_n(q)$ contains no homologies. She further shows [11, p. 299] that if $n>3$ is odd then $F\cap SL_n(3)$ is maximal in $SO_n(3)$, while $F\cap SL_n(4)$ is maximal in $SU_n(4)$ except when $n=3$ or $n=6$ [10, p. 377].

The conditions on the V_i that King imposed are the simplest that ensure that F is not reducible. He had earlier discussed when the stabilizer H in G of a subspace W is maximal in G, settling, for finite K, all cases except when G is $O_n(q)$ or $U_n(q)$ and W and $W^\perp$ are anisotropic, non-isotropic, and non-isomorphic [12–14]. If $n\geqq 3$ then, by (i) and (ii), the only case remaining is $G=O_3(q)$, one of W, $W^\perp$ has dimension 1, and the other has dimension 2 and is anisotropic. This is precisely the case we had to settle in Theorem 1.

The interlocking scenery has been described. It remains to present the proofs.

2. The Orthogonal Cases

As in § 1, V is a vector space of dimension n over $K=GF(q)$.

2.1. The polar bilinear form associated with a quadratic form $Q(x)$ on V is

$$B(x,y)=Q(x+y)-Q(x)-Q(y). \tag{2}$$

If a is a nonsingular vector then the map

$$S_a\colon x\to x-B(x,a)[Q(a)]^{-1}a \tag{3}$$

fixes $Q(x)$. If $B(x,a)\not\equiv 0$, which always happens when $Q(x)$ and thus, by definition, $B(x,y)$ is non-degenerate, then S_a is a symmetry and has order 2. Its image R_a acting on the projective space $PG(n-1,q)$ whose points are the 1-dimensional subspaces of V is a reflection with centre $\langle a\rangle$. If $B(x,a)\equiv 0$ then S_a and R_a are identity maps. S_a and R_a depend only on $\langle a\rangle$; they are unaltered if one replaces a by λa where $0\neq\lambda\in K$.

$Q(x)$ is called *nonsingular* if there is no singular vector $b\neq 0$ conjugate to the whole of V; i.e. $B(x,b)\equiv 0$. When q is odd nonsingularity and non-degeneracy are the same thing; but if q is even and n is odd then $Q(x)$ must be degenerate, yet may be nonsingular. In particular, if q is even and $n=3$, then

coordinates $x=(x_1, x_2, x_3)$ may be taken so that if $Q(x)$ is nonsingular then [2, p. 197]

$$Q(x)=x_1x_2+x_3^2. \tag{4}$$

It is very easy to check from (2) and (4) that $V^\perp$, the *nucleus* of $Q(x)$, is $\langle(0,0,1\rangle$; for a 1-dimensional subspace $W\neq V^\perp$ we have $W^\perp$ is the join of W to $V^\perp$. If q is odd and $n=3$ then there are two non-equivalent types of nonsingular $Q(x)$; the canonical form for one type is the scalar multiple of that of the other by a particular non-square in $GF(q)$ [2, p. 158]. A check using (2) shows that the $Q(x)$ given by (4) is non-degenerate when q is odd. It is thus, to within a scalar factor, a canonical form for nonsingular quadratics. The image, for all q, of the $Q(x)$ of (4) on $PG(2,q)$ is a conic.

We shall use the terms isomorphic, non-isotropic and anisotropic in the $PG(n-1,q)$ setting with their obvious interpretations.

2.2. Our first result is

Theorem 1. *Let $O_3(q)$ be the orthogonal group of a nonsingular quadratic form defined on a 3-dimensional vector space V over $GF(q)$. Then:*

(i) *the stabilizer in $O_3(q)$ of an anisotropic 2-dimensional subspace of V is a maximal subgroup of $O_3(q)$ except when $q=2$;*

(ii) *the stabilizer in $PO_3(q)$ of an anisotropic line of $PG(2,q)$ is a maximal subgroup of $PO_3(q)$ except when $q=2$;*

(iii) *if $q=2$ then the stabilizers of* (i) *and* (ii) *are respectively $O_3(2)$ and $PO_3(2)$.*

Proof. Let $Q(x)$ be the quadratic form, and W a 2-dimensional anisotropic subspace of V. Let $\mathscr{C}$ be the conic of $PG(2,q)$ corresponding to $Q(x)$ and l the image of W. Let H be the stabilizer in $O_3(q)$ of W and PH its image in $PO_3(q)$.

Now $PO_3(q)\cong PGL(V_2,q)$, where V_2 is a vector space of dimension 2: since each of its point is some $(\lambda^2, -\mu^2, \lambda\mu)\mathscr{C}$ can be considered as an abstract projective line $PG(V_2,q)$ with coordinates (λ,μ). So $\mathscr{C}$ has $q+1$ points, and thus $q+1$ tangent lines and $(q+1)q/2$ chords. There being q^2+q+1 lines in a plane, $q(q-1)/2$ are anisotropic. $PO_3(q)$ is transitive on the anisotropic lines [5, pp. 40, 41], so PH has index $q(q-1)/2$ in $PO_3(q)$. If $q=2$ then this index is 1; so (iii) follows. Further, since $|PGL_2(q)|=q(q^2-1)$ we have $|PH|=2(q+1)$.

Now suppose that $q>2$, so that $|PH|<|PO_3(q)|$. Since H contains the subgroup Λ of scalar maps of $O_3(q)$ and there is, by a fundamental homomorphism theorem, a bijective correspondence between subgroups of $PO_3(q)$ and subgroups of $O_3(q)$ containing Λ we need only establish (ii). So suppose that

$$PH<J\leqq PO_3(q). \tag{5}$$

Take first the case of q odd. Then $Q(x)$ is non-degenerate and $l^\perp$ is a nonsingular point A arising from $W^\perp=\langle a\rangle$. If $w\in W$ then, by (3), S_w fixes W. Moreover the restriction of S_w to W is a symmetry of the orthogonal group $O(W)$ of $Q|_W$, and any symmetry of $O(W)$ so arises. Since $O(W)$ is generated by

its symmetries [3, p. 20] we see, by Witt's theorem, that PH has 2 orbits of (anisotropic) points on l (see § 1). One orbit corresponds to the set of squares and the other to the set of nonsquares of K, so that [2, p. 46] each orbit has size $(q+1)/2$. Now [3, pp. 17, 20] $H \cong O(W) \times \langle S_a \rangle$, and by (3), R_a moves a point not in $\{A\} \cup l$ to a different point on its join to A. Hence, by considering the action on lines through A, we see that the orbits under PH of points of the plane not in $\{A\} \cup l$ each have size $q+1$.

Suppose that $q+1 \geqq 5$. Let

$$X = \{A\} \cup \{\text{the } (q+1)/2 \text{ points of } l \text{ isomorphic to } A\}.$$

A is conjugate to $(q+1)/2 \geqq 3$ members of X, but each other point of X is conjugate to at most 2 points of X, namely A and possibly a point on l. Hence if J fixed X it would fix A and thus $l = A^{\perp}$, in contradiction to (5). Hence the orbit of A under J contains a point not in $\{A\} \cup l$. Thus J contains a reflection centred off $\{A\} \cup l$: by (3) this reflection cannot fix A. Thus if

$$E = \langle R_b : R_b \in J \rangle$$

then, since $O(W)$ is generated by symmetries and thus PH is generated by reflections,

$$PH < E \leqq J \leqq PO_3(q). \tag{6}$$

From our discussion of orbits E can fix no point and thus, because of the polarity, no line. Hence E acts irreducibly. It it were not primitive then E would fix a triangle [18, p. 388]. Then PH would fix this triangle. But apart from A the only orbits of points under PH of size at most 3 consist of $(q+1)/2 = 3$ collinear points on l when $q = 5$. There is no possibility for the 3 non-collinear vertices of the triangle; E acts primitively. There are two conjugacy classes of reflections in $PO_3(q)$ corresponding to the two isomorphism types of anisotropic points. There are reflections of both these classes with centres on l and thus in PH and E. Hence E has at least two classes of reflections. Mitchell [17] determined the subgroups of $PGL_3(q)$, and Wagner [19, p. 241], giving credit to Mitchell, has recently listed the primitive subgroups generated by reflections. The only possibility for E, which fixes a conic $\mathscr{C}$ and has more than one reflection class is [19, p. 241] $E \cong PGL_2(q')$ where $GF(q') \leqq GF(q)$. Mitchell's discussion shows that there is a base of V such that with respect to this base E arises from 3×3 matrices with entries in $GF(q')$. The $GF(q')$-linear combinations of the base elements give rise to the points of a $PG(2, q') \subseteq PG(2, q)$; and $E = PO_3(q')$ fixing a conic $\mathscr{C}'$ in $PG(2, q')$. Thus $|PH| = 2(q+1)$ divides $|E| = q'(q'+1)(q'-1)$. Now $q = (q')^{\alpha}$ for some positive integer α. Hence $q+1$ and q' are relatively prime so that $(q+1) | ((q')^2 - 1$. Thus $(q')^{\alpha} + 1 \leqq (q')^2 - 1 < (q')^2$. Hence we must have $\alpha = 1$, so that $q = q'$ and $|E| = |PO_3(q)|$. Thus, by (6) $J = PO_3(q)$, and PH is maximal in $PO_3(q)$, by (5).

If $q = 3$ then we may take [2, 158]

$$Q(x) = \alpha(x_1^2 + x_2 + x_3^2)$$

where $\alpha = \pm 1$. There are $q(q-1)/2 = 3$ anisotropic lines. If e_1, e_2, e_3 is the base of the coordinate system we may take $W = \langle e_1, e_2 \rangle$. Then PH is the stabilizer of the point $(0,0,1)$. The solutions of $Q(x) = \alpha$ are $x = \pm e_i$ for $i = 1,2,3$. Hence $PO_3(3)$ permutes the points $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$, and by the symmetry of the form of $Q(x)$, permutes them doubly transitively: the symmetric group Σ_3 acts. Hence PH, the stabilizer of one of the points, is maximal in $PO_3(3)$.

Now consider the case $q > 2$ and even. Then $l^\perp = A$ is the nucleus of $\mathscr{C}$. From (4) it follows that the join of A to a point $\langle w \rangle$ on l meets $\mathscr{C}$ in a unique point; this is one way we can think of $\mathscr{C}$ as a $PG(V_2, q)$. Here, $PO_3(q) \cong PGL(V_2, q) = PSL(V_2, q) \cong SL(V_2, q)$ [4, p. 40]. Recall that $|PH| = 2(q+1)$. Wagner has given [18, p. 397], a convenient non-overlapping catalogue of the subgroups of $PSL_2(q) = PGL_2(q)$. They are: cyclic groups of order z where z divides $q \pm 1$; dihedral groups of order $2z$; semidirect products of order $2^c m$ where $c \geqq 1$, $2^c | q$ and $m | (q-1)$ and $m | (2^c - 1)$; $PSL(V_2, q')$ where $GF(q') \leqq GF(q)$. Since $|J| > |PH|$ and $|PH| \,|\, |J|$ order considerations show that we must have $J = PSL(V_2, q')$. Here we can have matrices for J with entries in $GF(q')$ by a choice of base of V. Hence $|PH| = 2(q+1)$ divides $|PSL_2(q')| = q'(q'+1)(q'-1)$. The numerical argument used above when q was odd still applies, and shows that $q = q'$. Thus $|J| = |PSL_2(q)| = |PO_3(q)|$. So, from (5), $J = PO_3(q)$ and PH is maximal in $PO_3(q)$.

2.3. For use in the proof of Theorem 2 we establish

Proposition 1. *Let f be a quadratic form on an n-dimensional vector space V over $GF(q)$ where $q > 2$ and q is even. Let $O_n(q)$ be the group of f.*

(i) *If $n = 3$ and f is nonsingular then $O_3(q)$ is generated by its symmetries, and the stabilizer H in $O_3(q)$ of an anisotropic 2-dimensional subspace W is generated by the symmetries of $O_3(q)$ with centres in W.*

(ii) *If $n \geqq 4$, f is non-degenerate, and U is a 3-dimensional subspace such that $f|_U$ is nonsingular then the symmetries of $O_n(q)$ whose centres are in U but not at the nucleus of $f|_U$ generate the subgroup of $O_n(q)$ that fixes $U^\perp$ vectorwise.*

Proof. (i) The conjugate of a symmetry is a symmetry. So the symmetries of $O_3(q)$ generate a normal subgroup N of $O_3(q)$. But $O_3(q) \cong PO_3(q)$, and $PO_3(q) \cong PSL_2(q)$ which is simple [4, p. 39]. Hence $N = O_3(q)$. The group $S = \langle S_w : w \in W \rangle$ fixes W, by (3). Its restriction to W contains all the symmetries of the group $O(W)$ of $f|_W$, and so is $O(W)$ [3, p. 42] of order $2(q+1)$ [4, p. 70]. This, by the proof of Theorem 1, is $|PH| = |H|$. Hence, by orders, $S = H$.

(ii) Let the nucleus of $f|_U$ be $\langle a \rangle = U \cap U^\perp$. Let

$$L = \{S_u : u \in U\}.$$

By (3), each element of L fixes each vector of $U^\perp$. The restriction L_U of L to U contains every symmetry of the group $O(U)$ of $f|_U$. Hence, by (i), $L|_U = O(U)$. Any member of $O_n(q)$ fixing U and $U^\perp$ vectorwise fixes each vector of $U + U^\perp$

$=\langle a\rangle^{\perp}$ and so is either the identity or S_a. Suppose that the element g of $O_n(q)$ fixes $U^{\perp}$ vectorwise. Then there is an element h in L such that $g|_U = h|_U$. Hence $gh^{-1} \in \langle S_a\rangle$. Thus $g \in L$, and L is the vectorwise stabilizer of $U^{\perp}$. Each g in L fixes a and so commutes with S_a; this is easily checked by (3). Further g and gS_a have the same effect on U and just one of them is in the special orthogonal group $SO_n(q)$. Thus $L = \bar{L} \times \langle S_a\rangle$ where $\bar{L} = SO_n(q) \cap L$, and $\bar{L} \cong L|_U = O(U)$.

Suppose that

$$M = \{S_u : u \in U \setminus \langle a\rangle\} \leqq L.$$

Since $S_a|_U$ is the identity, $M|_U = L|_U = O(U)$. Thus $|M|$ is a multiple of $|\bar{L}| = |L|/2$. Thus if $M \neq L$ then $|\bar{L}| = |M| = |L|/2$. Now $M \neq \bar{L}$, as M contains symmetries which are transvections and are not in $SO_n(q)$ [4, p. 65]. Hence if $M \neq \bar{L}$ then $|\bar{L} : M \cap \bar{L}| = 2$. But $\bar{L} \cong O(U)$ which is simple; a contradiction. Hence $M = L$ as required.

We remark that (i) is true but (ii) is false when $q = 2$.

2.4. Our first major result is

Theorem 2. *Suppose that f is a non-degenerate quadratic form on a vector space V of dimension $n \geqq 4$ over $GF(q)$ with group $O_n(q)$. Suppose that V is the direct sum of mutually conjugate, isomorphic, anisotropic subspaces $V_1, V_2, \ldots, V_s$ of dimension 2. Then:*

(i) *the stabilizer in $O_n(q)$ of $\{V_1, V_2, \ldots, V_s\}$ is a maximal subgroup of $O_n(q)$ except when $q = 3$, or $q = 2$ and $n = 4$;*

(ii) *the stabilizer in $PO_n(q)$ of the set of images of the V_i in $PG(n-1, q)$ is a maximal subgroup of $PO_n(q)$ except when $q = 3$, or $q = 2$ and $n = 4$;*

(iii) *if $q = 3$ then the stabilizers of* (i) *and* (ii) *lie strictly in the stabilizers respectively of a direct sum of n mutually conjugate isomorphic nonsingular 1-dimensional subspaces and a self-polar simplex of isomorphic nonsingular points;*

(iv) *if $q = 2$ and $n = 4$ then the stabilizers of* (i) *and* (ii) *are respectively $O_4(2)$ and $PO_4(2)$.*

Moreover:

(v) *a direct sum of such V_i can occur if and only if, according as s is even or odd f has Witt index s or $s-1$.*

Proof. Clearly $n = 2s$. There being only one type of 2-dimensional anisotropic subspace (§ 1) the $f|_{V_i}$ are all equivalent, say to the form $Q(\alpha, \beta)$ in the variables α, β. We have (1). So, by the mutual conjugacy of the V_i, we may choose a base $e_1, e_2, \ldots, e_n$ and thus coordinates $x = (x_1, x_2, \ldots, x_n)$ such that $V_i = \langle e_{2i-1}, e_{2i}\rangle$ for $i = 1, \ldots, n/2$, and

$$f(x) = Q(x_1, x_2) + Q(x_3, x_4) + \ldots + Q(x_{n-1}, x_n). \tag{7}$$

(v) The canonical form for a quadratic on V with Witt index $s = n/2$ is

$$x_1 x_2 + x_3 x_4 + \ldots + x_{n-1} x_n. \tag{8}$$

A non-degenerate quadratic not of index s has index $s-1$. Suppose that q is odd. Then (8) has determinant $(-1)^s 2^{-2s}$. Dickson's discussion of canonical

forms [2, pp. 156–158] shows that f is equivalent to (8) if and only if $D(-1)^s 2^{2s}$ is a square in K, where D is the determinant of f. By (7), $D=\lambda^s$ where λ is the determinant of $Q(\alpha,\beta)$. Taking $s=1$ we see $(-1)\lambda 2^2$ is a non-square. Since $D(-1)^s 2^{2s}=(-\lambda 2^2)^s$ we see that f is equivalent to (8) if and only if s is even. Suppose, next, that q is even. The Arf invariant of (8) is 0, so f is equivalent to (8) if and only if its Arf invariant Δ is in $X=\{\lambda^2+\lambda:\ \lambda\in K\}$ [6, p. 36]. We may [2, p. 199] have, for some k in K, $Q(\alpha,\beta)=\alpha^2+\alpha\beta+k\beta^2$. From (7), $\Delta=sk$. The case $s=1$ shows that $k\notin X$. If s is even then $\Delta=0\in X$; if s is odd then $\Delta=k\notin X$. So (v) follows.

For essentially the same reason as for Theorem 1 we need only establish the vector space parts of the rest of the statement of the theorem. As in § 1, call the stabilizer F.

By (3) the only subspaces fixed by a symmetry S_a of $O_n(q)$ are those in $\langle a\rangle^\perp$ and those through $\langle a\rangle$. Thus if $S_a\in F$ then $\langle a\rangle$ is in some V_i. For otherwise S_a fixes each $V_i\cap\langle a\rangle^\perp$, which is not $\langle 0\rangle$, and thus fixes V_i, the unique member of $\{V_1,\ldots,V_s\}$ containing $V_i\cap\langle a\rangle^\perp$: this forces every V_i, and thus V, to be contained in $\langle a\rangle^\perp$; a nonsense. Since a symmetry with centre in a V_i is in F, by (3), we see that the symmetries of F are precisely those with centres in the V_i. Any element of $O_n(q)$ fixing the set of non-zero vectors in $V_1\cup V_2\cup\ldots\cup V_s$ must be in F; for this set determines the V_i uniquely, as the only vectors of the set not conjugate to one in V_i are those in q of the 1-dimensional subspaces of V_i.

If $q\neq 2$ then there is $\lambda\neq 0,\ 1$ in K. There are a_1, a_2 in V_1, V_2 respectively such that $f(a_1)=\lambda$ and $f(a_2)=1-\lambda$. Then, by (7), $f(a_1+a_2)=1$. If $q=2$ and $s\geqq 3$ then there are a_1, a_2, a_3 in V_1, V_2, V_3 such that $f(a_1)=f(a_2)=f(a_3)=1$. Then $f(a_1+a_2+a_3)=1$. Thus except when $q=2$ and $n=4$ there are non-singular vectors not in any V_i and so, by the previous paragraph, symmetries not in F. If $q=2$ and $n=4$ then any x in V not in $V_1\cup V_2$ can be written $x=a_1+a_2$ with $0\neq a_i\in V_i$. Since V_i is anisotropic $f(a_i)=1$. Thus $f(a_1+a_2)=0$. Hence every nonsingular vector is in $V_1\cup V_2$. Thus, by the last sentence of the previous paragraph, $O_4(2)=F$; we have (iv).

Suppose that $q\neq 3$, and we exclude the case $q=2$ and $n=4$. Then we know that $F<O_n(q)$. Let

$$F<J\leqq O_n(q). \tag{9}$$

By the previous paragraph but one J does not fix $V_1\cup V_2\cup\ldots\cup V_s$ and so contains a symmetry S_a whose centre $\langle a\rangle$ is in no V_i. Thus

$$a=a_1+a_2+\ldots+a_s, \tag{10}$$

where $a_i\in V_i$, and at least two of $a_1,\ldots,a_s$ are not 0. Let $O(V_i)$ be the group generated by the symmetries with centres in V_i: it is isomorphic to the group of $f|_{V_i}$, acts as such on V_i, and fixes $V_i^\perp$ vectorwise. $O(V_i^\perp)$ denotes the subgroup of $O_n(q)$ fixing V_i vectorwise; it is isomorphic to the group of $f|_{V_i^\perp}$. We have $O(V_i)<J$. Denote the polar bilinear form associated with f by B.

Consider the possibility $n=4$. Then neither of a_1, a_2 is 0 and $W=\langle a_2, V_1\rangle$ has dimension 3. Any vector in W is λa_2+v_1 for some λ in K and v_1 in V_1. This is in $W^\perp$ if and only if $B(\lambda a_2+v_1,\ \mu a_2+\hat{v}_1)=0$ for all μ in K and $\hat{v}_1$ in V_1. By the conjugacy of V_1 and V_2 this requirement is $2\lambda\mu f(a_2)+B(v_1,\hat{v}_1)=0$.

Thus, since $f(a_2) \neq 0$, we require $2\lambda = 0$ and $v_1 = 0$ since $B|_{V_1}$ is non-degenerate. Hence $W \cap W^\perp$ is $\langle 0 \rangle$ or $\langle a_2 \rangle$ according as q is odd or even. So $f|_W$ is nonsingular. Let

$$M = \langle O(V_1), S_{a_2}, S_a \rangle \leqq J.$$

Let $O(W)$ be the group of $f|_W$. The stabilizer in $O(W)$ of V_1 is the restriction of $\langle O(V_1), S_{a_2} \rangle$ to W. For q odd we saw this in the course of proving Theorem 1; for q even it follows from Proposition 1(i) and the fact that $S_{a_2}|_W$ is the identity. Since M fixes W, it follows from Theorem 1 that $M|_W = O(W)$. If q is even then [5, p. 40] the orbit of $\langle a \rangle$ under M contains all anisotropic 1-dimensional subspaces apart from the nucleus $\langle a_2 \rangle$, and thus M contains the symmetries centred on these. Hence, by Proposition 1, J contains the stabilizer of W in $O_4(q)$. If q is odd then M has 2 orbits of anisotropic 1-dimensional subspaces each of which have members in V_1. Thus M contains all symmetries centred on W. These, together with the symmetry centred on $W^\perp$, which is in V_2, generate the stabilizer in $O_4(q)$ of W [3, p. 17], which is thus in J. For both q odd and even J contains elements not fixing W: one takes a symmetry whose centre in V_2 is not $\langle a_2 \rangle$ or $W^\perp$. The stabilizer of W in $O_4(q)$ is maximal [13, pp. 352, 373]. Thus $J = O_4(q)$, and, by (9), F is maximal in $O_4(q)$.

We may now prove (i) by induction on s. So suppose that $s \geqq 3$. One possibility for (10) is that at least one of $a_1, \ldots, a_s$ is 0. In view of the symmetry of (7) – the symmetric group Σ_s acts on the ordered pairs of coordinates $(x_1, x_2), (x_3, x_4), \ldots, (x_{n-1}, x_n)$ and is in F – we may assume that $a_1 = 0$. Now

$$V_1^\perp = V_2 \oplus V_3 \oplus \ldots \oplus V_s,$$

and $V_2, \ldots, V_s$ are mutually conjugate and anisotropic for $f|_{V_1^\perp}$. Clearly, from (7), if $F_1 = F \cap O(V_1^\perp)$ then F_1 is the stabilizer of $\{V_2, V_3, \ldots, V_s\}$ in $O(V_1^\perp)$ which is the group of $f|_{V_1^\perp}$. Now $J \cap O(V_1^\perp)$ contains F_1 and S_a, which is not in F_1 since its centre is not in one of $V_2, \ldots, V_s$. Hence, by the induction hypothesis, or (iv) if $q = 2$ and $s = 3$, we have $J \cap O(V_1^\perp) = O(V_1^\perp)$. Thus $J \geqq O(V_1^\perp) \times O(V_1)$, which is the stabilizer N of V_1 in $O_n(q)$. The map, in coordinate terms, $x_1 \leftrightarrow x_3$, $x_2 \leftrightarrow x_4$ is in F, by (7). This map does not fix V_1 so that, by (9), $J > N$. By (v) the index v_2 of $V_1^\perp$ is $s-1$ or $s-2$ according as s is odd or even, while the index v_1 of V_1 is 0. We deduce from [13, pp. 352, 373] that $J = O_n(q)$, except when $n = 6$, $q = 2$. But this exceptional case cannot occur with our possibility. For if it did then, by (7) and (10), $1 = f(a) = f(a_2) + f(a_3)$. Then one of $f(a_2)$, $f(a_3)$ is 0 and the other 1; and so one of a_2, a_3 is 0 giving, by (10), a to be in V_2 or V_3, which is not allowed.

The other possibility is that no, a_i is 0. Suppose that one of $a - a_1$, $a - a_2, \ldots, a - a_s$ is nonsingular. By the symmetry of (7) we may assume that $b = a - a_1$ is nonsingular. Now, by (10), $b = a_2 + a_3 + \ldots + a_s$ and so b is in $V_1^\perp$ but not in any of $V_2, \ldots, V_s$. Thus $S_b \notin F$. Let $U = \langle V_1, b \rangle$ of dimension 3. The argument for W when $n = 4$ with b in place of a_1 shows that $f|_U$ is nonsingular, with nucleus $\langle b \rangle$ if q is even. If $q = 2$ then U would contain at least 5 nonsingular vectors, namely 3 on V_1, a and b. The canonical form (4) shows that over $GF(2)$ there are only 4 non-singular vectors; $(0, 0, 1)$, $(1, 0, 1)$, $(0, 1, 1)$

and $(1,1,0)$ in the coordinate system for (4). Thus $q>2$. Let

$$Y=\langle O(V_1),\quad S_a\rangle \leqq J, \tag{11}$$

and let $O(U)$ be the group of $f|_U$. Then Y fixes U, but not V_1. By Proposition 1(i), if q is even then $Y|_U$ strictly contains the stabilizer of V_1 in $O(U)$ and so, by Theorem 1, $Y|_U=O(U)$. Hence Y permutes transitively the anisotropic 1-dimensional subspaces of U other than the nucleus $\langle b\rangle$ [5, p. 40]. Thus Y contains all symmetries S_u of $O_n(q)$ with $u\in U\setminus\langle b\rangle$. Hence, by Proposition 1(ii), Y contains the subgroup of $O_n(q)$ fixing $U^\perp$ vectorwise. Thus Y, and hence J contains S_b. We are back to our first possibility so that $J=O_n(q)$. If q is odd then the stabilizer in $O(U)$ of V_1 is isomorphic to, and acts as, $O(V_1)\times\langle S_b\rangle$ [3, p. 17]. Now $(-S_b)$ fixes b and V_1, so that $(-S_b)|_U\in O(V_1)|_U$. Concentrate on the projective plane $PG(U,q)$ and let l be the image of V_1. It follows that the stabilizer in $PO(U)$ of l is isomorphic to $PO(V_1)$. It then follows from (11) that the projective image of Y on $PG(U,q)$ strictly contains this stabilizer and so, by Theorem 1, is $PO(U)$. Thus Y permutes transitively the anisotropic 1-dimensional subspaces of U of each type. V_1 contains members of each type. Thus Y contains all symmetries with centres in U. In particular Y, and thus J, contains S_b. We are, again, back to the earlier possibility and $J=O_n(q)$.

We have now to deal with the case when $a-a_1, a-a_2, \ldots, a-a_s$ are all singular. By (7)

$$f(a)-f(a_1)=f(a)-f(a_2)=\ldots=f(a)-f(a_s)=0. \tag{12}$$

Write $v=[f(a)]^{-1}$. The image of $v_1\neq 0$ in V_1 under S_a is, by (3),

$$c=(v_1-B(v_1,a)\,va_1)-(B(v_1,a)\,va_2)-(B(v_1,a)\,va_3)-\ldots-(B(v_1,a)\,va_s),$$

and $S_c\in J$. Now $B(v_1,a)=B(v_1,a_1)$. Thus if $v_1\notin\langle a_1\rangle^\perp$ and we are not to be back to an earlier possibility we must have, by the analogue of (12) and (12),

$$f(v_1)=f(c)=f(-B(v_1,a)\,va_2)=v^2f(a_2)[B(v_1,a_1)]^2=v[B(v_1,a_1)]^2.$$

Thus the restriction to V_1 of the quadratic form $f(x)-v[B(x,a_1)]^2$ vanishes at all 1-dimensional subspaces of V_1 other than $\langle a_1\rangle^\perp$. There are q of these. Hence if $q>2$ the restriction is identically zero $f|_{V_1}$ is a scalar multiple of a linear form squared, contradicting V_1 is anisotropic. Hence $q=2$. Then $Q(\alpha,\beta)=\alpha^2+\alpha\beta+\beta^2$. Each V_i contains 3 nonsingular vectors which are permuted transitively by $O(V_i)$. By (7), F contains $O(V_1)\times O(V_2)\times\ldots\times O(V_s)$ which has for orbit of a the set X of all 3^s vectors that are the sum of s nonsingular vectors one from each V_i. If we are not to have an earlier possibility we see that these 3^s vectors and only these are centres of symmetries in J/F. Since J cannot fix $V_1\cup V_2\cup\ldots\cup V_s$ the orbit of e_1 under J must be the union Z of X with $\{V_1\cup V_2\ldots\cup V_s\setminus\{0\}\}$. Since $f(a_i)=f(a)=1$, we see, from (7) and (10), that $f(a)=s$, so s must be odd. The vectors of Z conjugate to e_1 are, by (2) and (7), e_1, those on $V_2,\ldots,V_s$, and those in X whose sum expression starts with e_1; altogether $1+3(s-1)+3^{s-1}=3^{s-1}+3s-2$. A vector x conjugate to

$e_2+e_4+\ldots+e_n$ satisfies, in coordinate terms, $x_1+x_3+x_5+\ldots+x_{n-1}=0$. The only such x in V_i is e_{2i}. If $x\in X$ then an even number r of $x_1, x_3, \ldots, x_{2s-1}$ must be 1, and the rest must be 0; further if $x_{2i-1}=0$ we must take $x_{2i}=1$. The number of such x in X is thus the sum for r even of $\binom{s}{r}2^r$, which is $[3^s+(-1)^s]/2$ or $(3^s-1)/2$ by the binomial theorem. Thus $e_2+e_4+\ldots+e_n$ is conjugate to $s+(3^s-1)/2$ members of Z. The transitivity of J on Z makes this equal to $3^{s-1}+3s-2$. Hence $3^{s-1}=4s-3$. Simple induction shows that $3^{s-1}>4s-3$ if $s\geqq 4$. Hence $s=3$. But when $s=3$ we have seen in the proof of (iv) that any nonsingular vector in $V_1^{\perp}$ is in V_2 or V_3, so the orbit of a under $O(V_1)\times O(V_2)\times O(V_3)$ contains all nonsingular vectors of V not in a V_i. Thus J contains all symmetries of $O_6(2)$ so that $J=O_6(2)$ [3, p. 42].

This in all cases $J=O_n(q)$, and, by (9), F is maximal in $O_n(q)$. The induction is complete.

Finally (iii). When $q=3$ we may take $Q(\alpha,\beta)=\alpha^2+\beta^2$ [2, p. 158] so that

$$f(x)=x_1^2+x_2^2+\ldots+x_n^2. \tag{13}$$

The only x in V_i for which $f(x)=1$ are $\pm e_{2i-1}$ and $\pm e_{2i}$. Hence F permutes $\langle e_1\rangle, \langle e_2\rangle, \ldots, \langle e_n\rangle$, but ties them in imprimitivity blocks of size 2. From (13), the stabilizer $\hat{F}$ of $\{\langle e_1\rangle, \ldots, \langle e_n\rangle\}$ acts on the $\langle e_i\rangle$ as Σ_n: just take any permutation of $x_1, \ldots, x_n$ to obtain an element of Σ_n. So $F<\hat{F}$. We show immediately below that $\hat{F}$ is a maximal subgroup of $O_n(3)$. As V is the direct sum of the $\langle e_i\rangle$ which, by (13), are isomorphic, anisotropic and mutually conjugate, the proof of the theorem is complete.

We should point out that the detail of the proof has essentially established the fact: *F is the semidirect product of $O(V_1)\times O(V_2)\times\ldots\times O(V_s)$ with Σ_s.*

2.5. By § 1, the other case we must consider when f is quadratic is to have (1) with $r=1$, and $K=GF(q)$ with q odd. Then we may take a base $e_1, \ldots, e_n$ so that $V_i=\langle e_i\rangle$ and for some α in K,

$$f(x)=\alpha(x_1^2+x_2^2+\ldots+x_n^2). \tag{14}$$

Any element of the group Σ_n of permutations of $e_1, \ldots, e_n$ is in the stabilizer F of $\{\langle e_1\rangle, \ldots, \langle e_n\rangle\}$ in $O_n(q)$, and Σ_n acts as the symmetric group on $\langle e_1\rangle, \ldots, \langle e_n\rangle$. Hence if $g\in F$ then there is an element σ in Σ_n so that $\rho^{-1}g$ fixes each $\langle e_i\rangle$. By (14), $\sigma^{-1}g$ is a diagonal matrix with each diagonal element being ± 1. If Λ_n denotes the group of such matrices then Λ_n is an elementary abelian group of order 2^n, and F is the semidirect product $\Lambda_n.\ \Sigma_n$ of order $2^n n!$. Further each element of F has for each of its matrix entries 0, 1, or -1. Thus $F\leqq O_n(p)$, where p is the characteristic of K, so $GF(p)\leqq K$, and $O_n(p)$ is the group of the quadratic form $x_1^2+\ldots+x_n^2$ over $GF(p)$. Clearly $O_n(p)\leqq O_n(q)$. If we consider the sub-vector space of V consisting of the $GF(p)$-linear combinations of the e_i then F is the stabilizer in $O_n(p)$ of $\{\langle e_1\rangle, \ldots, \langle e_n\rangle\}$; the $\langle e_i\rangle$ are still isomorphic over $GF(p)$.

2.6. We need some notation, due to Wagner in the projective setting. If q is odd and $O_n(q)$ is the orthogonal group of the non-degenerate form f then we have already remarked that $O_n(q)$ has two conjugacy classes of symmetries; the

value of f at the centres of the symmetries one class are squares in $GF(q)$ and the values at the centres for the other class are non-squares. The subgroup of $O_n(q)$ generated by the symmetries of the first class is denoted by $R^+O_n(q)$ and that generated by the symmetries of the second class by $R^-O_n(q)$. Their images in $PO_n(q)$, generated by the corresponding classes of reflections, are denoted by $PR^+O_n(q)$ and $PR^-O_n(q)$. The subgroups of $PGL_n(q)$, $PU_n(q)$ (for q a square), and $PO_n(q)$ generated by their respective reflections are called $PRL_n(q)$, $PRU_n(q)$, and $PRO_n(q)$ by Wagner.

Always $PRO_n(q)=PO_n(q)$. If n is even and the determinant of f is a square then $PR^+O_n(q)$, $PR^-O_n(q)$ and $PSO_n(q)$ are 3 distinct subgroups of $PO_n(q)$ each of index 2. If n is odd then we may change the determinant of f from a non-square to a square by multiplying f by a constant: $O_n(q)$ is the group of the new form. If n is odd and the determinant of f is a square then, excluding the case $q=n=3$, $PR^-O_n(q)=PO_n(q)$ and $PR^+O_n(q)=P\Omega_n(q)$, of index 2 in $PRO_n(q)$. All this is recorded in [20, p. 192], with more besides. Since its reflections generate $PO_n(q)$ and $PR^+O_n(q)$ is a proper normal subgroup it follows that, excluding the case $n=q=3$, for the kinds of f just mentioned $PR^+O_n(q)$ does not contain any generating symmetry of $PR^-O_n(q)$. Similarly, if n is even then $PR^-O_n(q)$ contains no generating symmetry of $PR^+O_n(q)$ when f has square determinant.

2.7. The discussion of §§ 2.5, 2.6 motivates

Theorem 3. *Suppose that f is a non-degenerate quadratic form on a vector space V of dimension $n\geqq 3$ over the field $GF(p)$ where p is an odd prime. Suppose that V can be written as the direct sum of mutually conjugate, isomorphic, anisotropic subspaces $V_1, V_2, \ldots, V_n$ of dimension 1, whose point images in $PG(n-1,q)$ are the vertices $X_1, \ldots, X_n$ of a self-polar simplex. Let F be the stabilizer of $\{V_1, V_2, \ldots, V_n\}$ in $O_n(p)$ and PF the stabilizer of $\{X_1, X_2, \ldots, X_n\}$ in $PO_n(p)$. Then:*

(i) *F has order $2^n \cdot n!$ and PF order $2^{n-1} \cdot n!$;*

(ii) *if $p \equiv \pm 3 (\mathrm{mod}\, 8)$ and $n \neq 4$ then F is a maximal subgroup of $O_n(p)$ and PF is a maximal subgroup of $PO_n(p)$, except when $n=p=3$;*

(iii) *if $p \equiv \pm 1 (\mathrm{mod}\, 8)$ and $n \neq 4$ and we consider $O_n(p)$ as the group of the form $\alpha^{-1}f$ where α is a value of $f|_{V_1}$, then the subgroups of $O_n(p)$ strictly containing F are $R^+O_n(p)$ and $O_n(p)$, and the subgroups of $PO_n(p)$ strictly containing PF are $PR^+O_n(p)$ and $PO_n(p)$;*

(iv) *if $n=4$ then F is maximal in $O_4(p)$, and PF is maximal in $PO_4(p)$, only when $p=3$;*

(v) *if $n=p=3$ then $F=O_3(3)$ and $PF=PO_3(3)$.*

Proof. Again we need only establish either the vector space or the projective space results; this time the projective setting is the convenient one: it follows from (15) below that $R^+O_n(p)$ contains the product of the symmetries centred on $e_1, \ldots, e_n$ and thus contains $-I_n$.

We may take a base $e_1, \ldots, e_n$ so that, with respect to this base, f is given by (14) with $\alpha \in GF(p)$. Now $\alpha = f(e_1)$. We henceforth consider $O_n(p)$ as the group of

$$\alpha^{-1}f(x) = x_1^2 + x_2^2 + \ldots + x_n^2, \tag{15}$$

whose discriminant is a square, namely 1. We denote the images in $PO_n(p)$ of Σ_n and Λ_n of §2.5 by $P\Sigma_n$ and $P\Lambda_n$; here $P\Sigma_n \cong \Sigma_n$.

The symmetry S_{e_i} has matrix $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ where $\lambda_i = -1$ and $\lambda_j = 1$ if $j \neq i$, by (3). Hence it is in F, and Λ_n is generated by symmetries. The only symmetries of Λ_n are those centred on the $\langle e_i \rangle$. The transposition $e_1 \leftrightarrow e_2$ of Σ_n is, it is readily checked by (3), S_a where $a = e_1 - e_2$. Thus Σ_n is generated by symmetries. Hence $F = \Lambda_n \cdot \Sigma_n$ is generated by symmetries and PF by reflections. If $S_b \in F \setminus \Lambda_n$ then there is an e_i not in $\langle b \rangle^\perp$. This e_i is, by (3), moved to a vector in $\langle e_i, b \rangle \setminus \langle e_i \rangle$. Hence $b \in \langle e_i, e_j \rangle$ for some $j \neq i$, and S_b interchanges $\langle e_i \rangle$ and $\langle e_j \rangle$. If $b = \lambda_i e_i + \lambda_j e_j$ then, by (3), $v(\lambda_j^2 - \lambda_i^2)\, e_i - 2v\lambda_i\lambda_j e_j \in \langle e_j \rangle$ where $v^{-1} = \lambda_i^2 + \lambda_j^2 \neq 0$. This forces $\lambda_j = \pm\lambda_i$ so that S_b has centre $\langle e_i + e_j \rangle$ or $\langle e_i - e_j \rangle$. Thus the centres of the symmetries of F are the orbits of $\langle e_1 \rangle$ and $\langle e_1 + e_2 \rangle$ under F. The values of $\alpha^{-1} f$ at e_1 and $e_1 + e_2$ are respectively 1 and 2. Hence if 2 is a square in $GF(p)$ then $F \leqq R^+ O_n(p)$, while if 2 is a non-square then F contains members of the generating set of both $R^+ O_n(p)$ and $R^- O_n(p)$. Now 2 is a square if and only if 2, considered as an integer, is a quadratic residue mod p. This is so if and only if $p \equiv \pm 1 \pmod 8$ [8, p. 75].

By considering the action of $P\Sigma_n$ we see that any orbit of points of $PG(n-1, p)$ under PF contains a point $(\alpha_1, \alpha_2, \ldots, \alpha_r, 0, \ldots, 0)$ where $\alpha_1 \alpha_2 \ldots \alpha_r \neq 0$. By considering the action of $P\Sigma_n$ on the possible zero coordinates and the action of $P\Lambda_n$ on the non-zero coordinates we see that if $r > 1$ then this orbit has size at least $2^{r-1}\binom{n}{r}$. If $1 < r < n$ then $2^{r-1}\binom{n}{r} \geqq 2^{r-1}\, n \geqq 2n$, while if $r = n$ then $2^{r-1}\binom{n}{r} = 2^{n-1} > n$ since $n \geqq 3$. Thus the only orbit under PF of size at most n is $\{X_1, X_2, \ldots, X_n\}$. So PF fixes only one simplex. Thus if g is in the normalizer of PF in $PO_n(p)$ then $PF = gPFg^{-1}$ fixes the simplex $\{gX_1, gX_2, \ldots, gX_n\}$ and so $g \in PF$. Thus PF is self-normalizing. Any subspace U fixed by PF contains all points of an orbit and thus some point $(\alpha_1, \alpha_2, \ldots, \alpha_r, 0, 0, \ldots, 0)$ with $\alpha_1 \alpha_2 \ldots \alpha_r \neq 0$. Take r to be the least possible value for this to occur. Operating by $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ in Λ_n where $\lambda_r = -1$ and $\lambda_j = 1$ for $j \neq r$ we see that $(\alpha_1, \alpha_2, \ldots, -\alpha_r, 0, \ldots, 0)$ is in U and hence so is $(\alpha_1, \alpha_2, \ldots, \alpha_{r-1}, 0, 0, \ldots, 0)$. By the minimality of r we must have $r = 1$. Then U contains $X_1, X_2, \ldots, X_n$ and so is $PG(n-1, p)$; PF acts irreducibly. That PF is irreducible and generated by reflections could be obtained from Wagner's case III(i) of his list of irreducible groups generated by reflections [19, p. 239], [20, p. 191], [21, p. 475]; we needed more information.

We proved (v) in the course of the proof of Theorem 1. So suppose that we do not have $n = p = 3$. If $n \geqq 4$ and $c = (1, 1, 1, 1, 0, \ldots, 0)$ then $S_c \in R^+ O_n(p)$ but is not, by our above discussion in PF. If $n = 3$ and $p > 3$ then V has [2, p. 48] at least $p^2 - p \geqq 20$ vectors c with $\alpha^{-1} f(c) = 1$. Only 18 of these can be the centres of the symmetries in F; we again have an S_c in $R^+ O_n(p)$ but not PF. Hence $PF < PO_n(p)$, and if $p \equiv \pm 1 \pmod 8$ $PF < PR^+ O_n(p)$. So suppose that

$$PF < J \leqq PO_n(p). \tag{16}$$

Let R be the group generated by the reflections of J. Then $PF \leqq R$ and $R \trianglelefteq J$. Since PF is self-normalizing we have

$$PF < R \leqq J \leqq PO_n(p). \tag{17}$$

Because PF is, R is irreducible.

Take the case $n=3$, so that $p>3$. Then R fixes a conic and acts primitively; the only triangle left fixed by PH is, we have seen above, $\{X_1, X_2, X_3\}$ and this, by (17), is not fixed by R. Hence, by [19, p. 241], R is one of $PSL_2(q')$, $PGL_2(q')$, A_5, where $GF(q') \leqq GF(p)$ and A_5 denotes the alternating group of degree 5. Since $|A_5|=60$ and $|PF|=2^2 \cdot 6=24$ we cannot, by (17), have $R=A_5$. Since p is a prime R must be either $PSL_2(p)$ or $PGL_2(p)$. Now $PO_3(p) \cong PGL_2(p)$ so that R has index 1 or 2 in $PO_3(p)$. As $p>3$ the only normal subgroup of $PGL_2(p)$ of index 2 is its simple subgroup $PSL_2(p)$ [4, p. 41], which is [4, p. 112] isomorphic to $P\Omega_3(p)$. Hence $P\Omega_3(p)$ is the only subgroup of $PO_3(p)$ of index 2. Thus R is either $PO_3(p)$ or $P\Omega_3(p)$. The latter is $PR^+O_3(p)$ by § 2.5, and, by the last sentences of § 2.5, can only occur when $p \equiv \pm 1 \pmod 8$. Now (ii) and (iii) follow for $n=3$.

Consider now the case $n>4$. Wagner's list of irreducible subgroups generated by reflections gives the following possibilities for R [19, p. 239], [20, p. 191], [21, p. 475]; (a) $PRL_n(q')$ or $PRU_n(q')$ where $GF(q') \leqq GF(p)$; (b) $PR^-\hat{O}_n(q')$, $PR^+\hat{O}_n(q')$, or $P\hat{O}_n(q')$ where $\hat{O}_n(q')$ is an n-dimensional orthogonal group over $GF(q') \leqq GF(p)$; (c) $P\Omega_5(3)$ if $n=5$, $PR^-O_6^-(3)$ or $PO_5(3)$ if $n=6$, $Sp_6(2)$ if $n=7$, or $O_8^+(2)$ if $n=8$; (d) Σ_{n+1} or Σ_{n+2}. Here $O_8^+(2)$ means the group of a form of index 4 and $O_6^-(3)$ the group of a form of index 2. In cases (a) or (b) we must have $q'=p$. Then $PRU_n(q')$ does not exist as q' is not a square, while $PRL_n(p)$ has $PSL_n(p)$ as a subgroup of index 1 or 2 [19, p. 240]. A Sylow p-subgroup of $PSL_n(p)$ and thus $PRL_n(p)$ has order $p^{n(n-1)/2}$ [4, p. 40], while a Sylow p-subgroup of $PO_n(p)$ has order $p^{n(n-1)/4}$ if n is even and $p^{(n-1)^2/4}$ if n is odd [4, p. 63]. Order considerations and (16) show that (a) is not possible. Now $|PF| = 2^{n-1}n!$, which is $2^7 \cdot 3 \cdot 5$, $2^9 \cdot 3^2 \cdot 5$, $2^9 \cdot 3^2 \cdot 5$, $2^{10} \cdot 3^2 \cdot 5 \cdot 7$, $2^{14} \cdot 3^2 \cdot 5 \cdot 7$ according as $n=5, 6, 7, 8$. The orders of $P\Omega_5(3)$, $PR^-O_6^-(3)$, $PO_5(3)$, $Sp_6(2)$, $O_8^+(2)$ are respectively $2^6 \cdot 3^4 \cdot 5$, $2^8 \cdot 3^6 \cdot 5 \cdot 7$, $2^7 \cdot 3^4 \cdot 5$, $2^9 \cdot 3^4 \cdot 5 \cdot 7$, $2^{13} \cdot 3^5 \cdot 5^2 \cdot 7$ [2, pp. 309, 310]. Thus possibility (c) cannot occur as in no case does $|PF|$ divide $|R|$. If (d) occurred then $2^{n-1}n!$ would divide $(n+2)!$ so 2^{n-1} would divide $(n+1)(n+2)$. Only one of $(n+1)$, $(n+2)$ is even and $2^{n-1} > n+2$ since $n \geqq 5$, so 2^{n-1} does not divide $(n+1)(n+2)$; thus possibility (d) cannot occur. We must thus have possibility (b) with $q'=p$. Wagner's description shows that $\hat{O}_n(p)$ fixes a quadratic form $\hat{f}$ on V, and so R fixes the corresponding quadric in $PG(n-1, q)$. By (17), PF fixes this quadric. But it is clear on considering the actions of $P\Lambda_n$ and $P\Sigma_n$ that the only quadric fixed by PF is that arising from f and its multiples. Hence $\hat{f}$ is a multiple of f and $\hat{O}_n(p)=O_n(p)$. Hence R is one of $PR^-O_n(p)$, $PR^+O_n(p)$ or $PO_n(p)$. We know that $F \leqq PR^+O_n(p)$ if $p \equiv \pm 1 \pmod 8$, while F contains symmetries of both $PR^-O_n(p)$ and $PR^+O_n(p)$ if $p \equiv \pm 3 \pmod 8$. Thus if n is odd so that $PR^-O_n(p)=PO_n(p)$ then, by the final remarks of § 2.6, R can be $PR^+O_n(p)$ if and only of $p \equiv \pm 1 \pmod 8$. The same remarks show that if n is even then $R \neq PR^-O_n(p)$, and so again, R can be $PR^+O_n(p)$ if

and only $p \equiv \pm 1 \pmod 8$. Now (ii) and (iii) follow from (17) and (16) since R has index 1 or 2 in $PO_n(p)$.

Now take $n=4$. Then each element of PA_n and $P\Sigma_n$, and thus each element of PF, either fixes or interchanges the tetrahedra T_1 and T_2 with respective vertices $(1,1,1,1)$, $(1,1,-1,-1)$, $(1,-1,1,-1)$, $(1,-1,-1,1)$ and $(1,1,1,-1)$, $(1,1,-1,1)$, $(1,-1,1,1)$, $(-1,1,1,1)$. T_1, T_2 and $\{X_1, X_2, X_3, X_4\}$ from a desmic set $\mathscr{D}$. The reflection with centre $(1,1,1,1)$ fixes T_1 and interchanges $\{X_1, X_2, X_3, X_4\}$ and T_2; the reflection with centre $(-1,1,1,1)$ fixes T_2 and interchanges $\{X_1, X_2, X_3, X_4\}$ and T_1. Thus the stabilizer L of $\mathscr{D}$ in $PO_4(p)$ acts transitively on the 3 tetrahedra and has order $3|PF| = 576$; we have $PF < L$ of index 3, so that PF is maximal in L. Thus PF can be maximal in $PO_4(p)$ if and only if $|PO_4(p)| = 576$. Since p^2 divides $|PO_4(p)|$ [4, p. 63], and the only odd prime dividing 576 is 3, this can happen only when $p=3$. It does happen when $p=3$ [2, p. 160]. Thus (iv) is established; (i) has been established in §2.5 so the proof is complete.

To find all subgroups of $PO_4(p)$ containing PF when $p>3$ would take us too far from our present purpose. The reason for including (iii) is that it is a dividend of the argument used for proving maximality. We should record the attendant

Theorem 4. (i) *If n is odd then the direct sum situation of Theorem 3 occurs for any non-degenerate f: if $p \equiv \pm 1 \pmod 8$ then PF is maximal in $P\Omega_n(p)$.*

(ii) *If n is even then the direct sum situation of Theorem 3 occurs for non-degenerate f of Witt index $n/2$ if this is even, and for those of Witt index $n/2$ or $n/2-1$ if $n/2$ is odd according as $p \equiv 1$ or $p \equiv -1 \pmod 4$.*

Proof. The situation occurs provided f can be taken as (14). If n is odd this is always so [2, p. 156]. If n is even then (14) has a square for its discriminant. It is thus equivalent to the form (8) with Witt index $n/2$ if and only if $(-1)^{n/2}$ is a square. This is so when $n/2$ is even, or if -1 is a square when $n/2$ is odd. Now -1 is a square in $GF(p)$ if and only if -1 is a quadratic residue mod p, i.e. $p \equiv 1 \pmod 4$ [8, p. 69]. The theorem now follows from these remarks and Theorem 3(iii).

3. The Unitary and General Linear Cases

3.1. It is possible to give a nearly simultaneous proof of our two results. They are now stated.

Theorem 5. *Suppose that f is a non-degenerate hermitian form on a vector space V of dimension $n \geqq 3$ over a $GF(q^2)$ with group $U_n(q^2)$. Then V can be written as the direct sum of mutually conjugate anisotropic 1-dimensional subspaces $V_1, V_2, \ldots, V_n$. Further:*

(i) *the stabilizer in $U_n(q^2)$ of $\{V_1, V_2, \ldots, V_n\}$ is a maximal subgroup of $U_n(q^2)$;*

(ii) *the stabilizer in $PU_n(q^2)$ of a self-polar simplex of $PG(n-1, q^2)$ is a maximal subgroup of $PU_n(q^2)$.*

Theorem 6. (i) *If a vector space V of dimension $n \geqq 3$ over $GF(q)$ is written as the direct sum of 1-dimensional subspaces $V_1, V_2, \ldots, V_n$, then the stabilizer in $GL_n(q)$ of $\{V_1, V_2, \ldots, V_n\}$ is a maximal sugroup of $GL_n(q)$ except when $q=2,3$ or 4;*

(ii) *The stabilizer in $PGL_n(q)$ with $n \geqq 3$ of a simplex of $PG(n-1,q)$ is a maximal subgroup of $PGL_n(q)$ except when $q=2,3$, or 4;*

(iii) *if $q=2,3,4$ then the stabilizer of* (i) *is contained respectively in the stabilizer of a vector, an $O_n(3)$, a $U_n(4)$;*

(iv) *if $q=2,3,4$ then the stabilizer of* (ii) *is contained respectively in the stabilizer of a point, a $PO_n(3)$, a $PU_n(4)$.*

3.2. In the case of Theorem 5 coordinates may be taken so that [9, p. 7]

$$f(x)=x_1^{q+1}+x_2^{q+1}+\ldots+x_n^{q+1}. \tag{18}$$

Thus we may take $V_i=\langle e_i\rangle$ where $e_1, e_2, \ldots, e_n$ is the base for the coordinate system. Conversely given the V_i, we may obtain this coordinate set-up. Since (§ 1) any two 1-dimensional anisotropic subspaces are isomorphic we may take the simplex of Theorem 5(ii) to be $X_1, X_2, \ldots, X_n$ arising from $V_1, \ldots, V_n$. More trivially we can have the same for Theorem 6(ii). We need only establish Theorem 5(ii) and Theorem 6(ii) and (iv).

For Theorem 5 write $U_n(q^2)=G$ and $PU_n(q^2)=PG$, and for Theorem 6 write $G=GL_n(q)$ and $PGL_n(q)=PG$. In both cases let PF be the stabilizer of $\{X_1, X_2, \ldots, X_n\}$ in PG, and F the stabilizer of $\{V_1, V_2, \ldots, V_n\}$ in G. If $PG=PGL_n(2)$ then e_i is the only non-zero vector of $\langle e_i\rangle$ so F permutes $e_1, e_2, \ldots, e_n$. Thus PF fixes the point $(1,1,\ldots,1)$. If $PG=PGL_n(3)$ then the non-zero vectors of $\langle e_i\rangle$ are $\pm e_i$; for each of these $x_j^2=0$ if $j\neq i$ and $x_i^2=1$. Thus F fixes the quadratic form $x_1^2+x_2^2+\ldots+x_n^2$. If $PG=PGL_n(4)$ then the non-zero vectors of $\langle e_i\rangle$ are e_i, we_i and w^2e_i where w is a primitive cube root of 1. Hence F fixes the hermitian form $x_1^3+x_2^3+\ldots+x_n^3$ over $GF(4)$. Theorem 6(iv) now follows.

Suppose, henceforth, that PG is not one of the cases of Theorem 6(iv). Now F contains all elements of the symmetric group Σ_n permuting $e_1, \ldots, e_n$. Thus, cf. §2.5, F is the semidirect product $\Lambda_n \cdot \Sigma_n$ where Λ_n consists of the diagonal matrices in F. If $PG=PGL_n(q)$ then $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is in F whenever $\lambda_1\lambda_2\ldots\lambda_n\neq 0$, and is the product of n matrices of which the ith is $\mathrm{diag}(1,1,1,\lambda_i,1,\ldots,1)$. This last matrix induces a homology in $PG(n-1,q)$ whose order is the order of λ_i considered as an element of the multiplicative group $GF(q)^x$ of $GF(q)$. Now $GF(q)^x$ is cyclic of order $q-1$. Hence $P\Lambda_n$, the image of Λ_n on $PG(n-1,q)$ is generated by homologies of order $m=q-1$. As $q\neq 2,3,4$, $m>3$. If $PG=PU_n(q^2)$ then, by (18), $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is in F provided $\lambda_i^{q+1}=1$ for each i. Now $GF(q^2)^x$ is cyclic of order q^2-1, so, for each i, there are $q+1$ such λ_i forming a cyclic group of order $q+1$. Thus writing $\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ as a product as before we see that $P\Lambda_n$ is generated by homologies of period $m=q+1\geqq 3$. We may repeat the essential argument of the fourth paragraph of the proof of Theorem 3, replacing -1 by a primitive mth root w of 1, to show that $\{X_1, X_2, \ldots, X_n\}$ is the only simplex fixed by PF, and that PF acts irreducibly. In our present case orbits have size at least

$$m^{r-1}\binom{n}{r}!$$

Suppose that U is a subspace of $PG(n-1,q)$ fixed by $P\Lambda_n$. Then if the point $(x_1, x_2, \ldots, x_n)$ is in U with $x_i \neq 0$ so is $(x_1, x_2, \ldots, x_{i-1}, wx_i, x_{i+1}, \ldots, x_n)$ and thus, on subtraction, the point X_i is in U Hence U is spanned by some of $X_1, X_2, \ldots, X_n$. Any subspace spanned by some of $X_1, X_2, \ldots, X_n$ is fixed by $P\Lambda_n$. In particular the only points fixed by $P\Lambda_n$ are $X_1, X_2, \ldots, X_n$. If g is in the normalizer of $P\Lambda_n$ in PG then $g^{-1}P\Lambda_n g = P\Lambda_n$ fixes each of $g^{-1}X_1$, $g^{-1}X_2, \ldots, g^{-1}X_n$. Hence g must permute $X_1, X_2, \ldots, X_n$ and thus be in PF. The normalizer of $P\Lambda_n$ in PG is PF.

Clearly if $PG = PGL_n(q)$ then $PF < PGL_n(q)$ as PF does not permute the points of $PG(n-1,q)$ transitively. $PU_n(q^2)$ permutes, by Witt's theorem, the anisotropic points transitively. There are such points other than the X_i; for example $(1,1,0,0,\ldots,0)$ if q is odd and $(1,1,1,0,\ldots,0)$ if q is even. Hence $PF < PG$.

So suppose that

$$PF < J \leqq PG. \tag{19}$$

Let R be the subgroup of J generated by its homologies of order m. Then

$$P\Lambda_n \leqq R \trianglelefteq J. \tag{20}$$

If $P\Lambda_n = R$ then J is in the normalizer in PG of $P\Lambda_n$ which is PF; this contradicts (19). Hence $P\Lambda_n < R$. Since PF acts irreducibly so, by (19), does J. Suppose that R does not act irreducibly. Then by Clifford's Theorem [1, pp. 534, 535] R is completely reducible. The projective space is the join of proper disjoint subspaces each of the same (projective) dimension r, each fixed by R, and on each of which R acts irreducibly. Any such subspace is, by (20), fixed by $P\Lambda_n$ and so is spanned by $r+1$ of $X_1, X_2, \ldots, X_n$. Suppose that $r > 0$ and one such subspace W is $\langle X_{i_0}, X_{i_1}, \ldots, X_{i_r}\rangle$. There is an element σ of PF, arising from one of Σ_n that fixes each of $X_{i_0}, X_{i_1}, \ldots, X_{i_{r-1}}$ and takes X_{i_r} to X_{i_s}, where i_s is none of $i_0, i_1, \ldots, i_r$. By (19), $\sigma \in J$ and hence, by (20), $\sigma R \sigma^{-1} = R$. Thus R fixes σW and thus $W \cap \sigma W = \langle X_{i_0}, \ldots, X_{i_{r-1}}\rangle$. This contradicts the result that $R|_W$ is irreducible. Hence $r = 0$. Then R must fix each of $X_1, X_2, \ldots, X_n$ in order that we have R completely reducible. But $P\Lambda_n$ is the subgroup of PG fixing each X_i, so that $R \leqq P\Lambda_n$. This contradicts $P\Lambda_n < R$. Hence R acts irreducibly.

We now have that R is an irreducible group generated by homologies of order m where $m \geqq 3$. We have seen that $m = 3$ only when $PG = PU_n(q^2)$, and then when $m = q+1 = 3$, i.e. $q = 2$. Wagner [18, p. 387] has given 4 possibilities for R. Write $q = p^a$ or $q^2 = p^a$ according as PG is $PGL_n(q)$ or $PU_n(q^2)$, where p is the (prime) characteristic. Two of Wagner's possibilities relate to the case $m = 3$ and p odd. This does not occur for us. Hence, by [18, p. 387], either

(A) $PGL_n(p^b) \trianglerighteq R \trianglerighteq PSL_n(p^b)$ for some $b \mid a$, where $m \mid p^b - 1$

or

(B) $P\hat{U}_n(p^{2b}) \trianglerighteq R \trianglerighteq PS\hat{U}_n(p^{2b})$ for some $2b \mid a$, where $m \mid p^b + 1$.

Here the groups over $GF(p^b)$ or $GF(p^{2b})$ act on projective spaces arising from linear groups acting on vector spaces whose members are respectively $GF(p^b)$

and $GF(p^{2b})$ – linear combinations of some base of V, and $\hat{U}_n(p^{2b})$ fixes an hermitian form $\hat{f}$ on its such sub vector space.

Take, first, the case $PG = PGL_n(q)$. Then $m = q-1 = p^a - 1$. If (A) occurs then since $(p^a-1)|(p^b-1)$ we have $b \geqq a$. Since $b|a$ this forces $a=b$. Hence $R \geqq PSL_n(q)$, which acts on our $PG(n-1,q)$. Now $P\Lambda_n$ has members in every coset of $PSL_n(q)$ in $PGL_n(q)$; for $\mathrm{diag}(\lambda_1, 1, 1, \ldots, 1)$ has determinant λ_1. Hence, by (20), $R \geqq PGL_n(q)$. Then, by (19), $J = PGL_n(q) = PG$ and so, by (19), PF is maximal in PG. If (B) occurs then $p^a - 1 \leqq p^b + 1$, so that $p^b(p^{a-b}-1) \leqq 2$. This can happen only if $p=2$ and $b = 1 = a-b$. Then $q = p^a = 4$, a case we have excluded. Thus PF is maximal in $PGL_n(q)$.

Now suppose that $PG = PU_n(q^2)$. Then $m = q+1 = p^{a/2}+1$. If (A) occurs then $(p^{a/2}+1)|(p^b-1)$, so that $b > a/2$. Since $b|a$ we must have $b=a$. Hence $R \geqq PSL_n(q^2)$, which acts on our $PG(n-1,q^2)$. But $PSL_n(q^2)$ permutes the points of $PG(n-1,q^2)$ transitively, which R, a subgroup of $PU_n(q^2)$ does not. Hence (A) cannot occur. Thus (B) occurs. Thus $p^{a/2}+1 \leqq p^b+1$ so that $b \geqq a/2$. Since $2b|a$ we must have $2b = a$. Hence

$$P\hat{U}_n(q^2) \geqq R \geqq PS\hat{U}_n(q^2), \tag{21}$$

where $\hat{U}_n(q^2)$ fixes $\hat{f}$ defined on V. Now $\hat{f}$ defines an hermitian variety $\mathscr{F}$ in $PG(n-1,q^2)$. The only hermitian variety fixed by $PS\hat{U}_n(q^2)$ is $\mathscr{F}$, so, by (21), $\mathscr{F}$ is the only hermitian variety fixed by R. By (19), R fixes the hermitian variety corresponding to f. Hence $\hat{f}$ is a multiple of f and so $\hat{U}_n(q^2) = U_n(q^2)$. It follows from (21) that provided R contains a member of each coset of $PSU_n(q^2)$ in $PU_n(q^2)$ then $R = PU_n(q^2)$. Then, by (20), $J = PU_n(q^2) = PG$ so that PF is maximal in PG, by (19). Each member of $PU_n(q^2)$ arises from a matrix of $U_n(q^2)$ whose determinant λ satisfies $\lambda^{q+1} = 1$. There is a matrix in Λ_n with determinant λ if $\lambda^{q+1} = 1$, namely $\mathrm{diag}(\lambda, 1, 1, \ldots, 1)$. Thus R does contain a member of each coset; our proof is complete.

References

1. Clifford, A.H.: Representations induced in an invariant subgroup. Ann. of Math. **38**, 533–550 (1937)
2. Dickson, L.E., Linear groups. Reprint. New York: Dover 1958
3. Dieudonné, J.: Sur les groupes classiques (third ed.). Paris: Hermann 1967
4. Dieudonné, J.: La géométrie des groupes classiques (third ed.). Berlin Heidelberg New York: Springer 1971
5. Dye, R.H.: On the transitivity of the orthogonal and symplectic groups in projective space. Proc. Cambridge, Philos Soc. **68**, 33–43 (1970)
6. Dye, R.H.: On the Arf invariant. J. Algebra **53**, 36–39 (1978)
7. Dye, R.H.: Symmetric groups as maximal subgroups of orthogonal and symplectic groups over the field of two elements. J. London Math. Soc. (2), **20**, 227–237 (1979)
8. Hardy, G.H., Wright, E.M.: An introduction to the theory of numbers (fourth ed.). Oxford: Clarendon Press 1959
9. Kantor, W.M.: Classical groups from a non-classical viewpoint. Lecture Notes, Mathematics Institute, Oxford, 1978
10. Key, J.D.: Some maximal subgroups of $PSL(n,q)$, $n \geqq 3$, $q = 2^r$. Geometriae Dedicata **4**, 377–386 (1975); erratum, ibid **6**, 389 (1977)

11. Key, J.D.: Some maximal subgroups of certain projective unimodular groups. J. London Math. Soc. (2), **19**, 291–300 (1979)
12. King, O.H.: Some maximal subgroups of the classical groups. J. Algebra **68**, 109–120 (1981)
13. King, O.H.: Maximal subgroups of the classical groups associated with non-isotropic subspaces of a vector space. J. Algebra **73**, 350–375 (1981)
14. King, O.H.: Maximal subgroups of the orthogonal group over a field of characteristic two. J. Algebra **76**, 540–548 (1982)
15. King, O.H.: Imprimitive maximal subgroups of the general linear, special linear, symplectic and general symplectic groups. J. London Math. Soc. (2), **25**, 416–424 (1982)
16. King, O.H.: Imprimitive maximal subgroups of the orthogonal, special orthogonal, unitary and special unitary groups. Math. Z. **182**, 193–203 (1983)
17. Mitchell, H.H.: Determination of the ordinary and modular ternary linear groups. Trans. Amer. Math. Soc. **12**, 207–242 (1911)
18. Wagner, A.: Collineation groups generated by homologies of order greater than 2, Geometriae Dedicata **7**, 387–398 (1978)
19–21. Wagner, A.: Determination of the finite primitive reflection groups over an arbitrary field of characteristic not 2, Part I, Geometriae Dedicata **9**, 239–253 (1980); Part II, ibid **10**, 191–203 (1981); Part III, ibid **10**, 475–523 (1981)

Received April 5, 1984

Math. Z. 189, 131–142 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

CR-Distributions and Analytic Continuation at Generating Edges

Emil J. Straube*

Department of Mathematics, Indiana University, Bloomington, IN 47405, USA

1. Introduction

The original Edge-of-the-Wedge theorem, due to Bogoljubov, was discovered in connection with the verification of the dispersion relations in quantum field theory. It dealt with analytic continuation across the edge $\mathbb{R}^n \subset \mathbb{C}^n$. This theorem has then seen many generalizations, but up to about 1974, the edge was always understood to be an open subset of euclidean space $\mathbb{R}^n \subset \mathbb{C}^n$. For statements, proofs and applications, see [2, 5, 7, 13]. For physical aspects, compare also [2] as well as chapter *V* of [12]. Around 1974 it was realized at about the same time by Bedford ([1]) and Pinčuk ([6]), that the theorem holds also for curved edges, as long as they are totally real, or, what is equivalent, as long as they are generating (note that the manifold in question has real dimension n). It is the purpose of this paper to show that the Edge-of-the-Wedge theorem remains true (Theorem 2.1) if the edge is allowed to be a *generating* manifold of arbitrary dimension (but $\geqq n$, this is required to be generating). Naturally, only the case of distribution boundary values needs attention, the case of continuous boundary values trivially reduces to the case of an n-dimensional edge. We will show that such a reduction is also possible in the distribution case. The reduction is made possible by the fact that convergence in the distribution sense to a boundary value on the edge implies convergence in a space of distribution-valued C^∞-*functions* (Theorem 4.1). The "extra" variables (reflecting that the dimension of the edge is more than n) enter only as variables of *functions*, and the convergence result implies that the *traces* on certain n-dimensional manifolds converge to a boundary value at an n-dimensional (generating) edge. This essentially completes the reduction. The proof of Theorem 4.1 also applies to families of CR-distributions (distribution solutions of the tangential CR-equations), and we single out the regularity properties of CR-distributions (Remark 4.2). We also briefly discuss a sufficient condition for hypoellipticity of the tangential CR-equations (Remark 4.3).

* Supported by a grant from Swiss National Science Foundation

2. Edge-of-the-Wedge Theorem for Generating Edges

We consider C^∞-submanifolds M of $\mathbb{C}^n$. By $T_P^{\mathbb{R}}(M)$ we denote the real tangent space to M at P, obtained by identifying $\mathbb{R}^{2n}$ with $\mathbb{C}^n$. We adopt the convention that the point $(x_1, y_1, x_2, y_2, \ldots, x_n, y_n)$ corresponds to $(z_1, \ldots, z_n)$ with $z_j = x_j + iy_j$, $j = 1, \ldots, n$. The complex tangent space is the largest complex subspace contained in the real tangent space, i.e.

$$T_P^{\mathbb{C}}(M) := T_P^{\mathbb{R}}(M) \cap i \cdot T_P^{\mathbb{R}}(M). \tag{1}$$

The multiplication by i is of course the one induced by $\mathbb{C}^n$.

The manifold M is called *generating*, if for all $P \in M$

$$T_P^{\mathbb{R}}(M) + i \cdot T_P^{\mathbb{R}}(M) = \mathbb{C}^n. \tag{2}$$

In passing, we note that (2) implies that the dimension of the complex tangent space is independent of $P \in M$: a generating manifold is a so called CR-manifold.

M is called *totally real*, if $T_P^{\mathbb{C}}(M) = \{0\}$, $\forall P \in M$, that is, if there are no complex tangents. Obviously, if $\dim_{\mathbb{R}} M = n$, M is totally real if and only if it is generating.

Let $\rho_1, \ldots, \rho_{2n}$ be C^∞, real valued functions, defined in an open neighborhood Ω of 0 in $\mathbb{C}^n$, such that the mapping

$$\rho := (\rho_1, \ldots, \rho_{2n}) \tag{3}$$

is a diffeomorphism of Ω onto the image, which we assume to be a ball around 0 for simplicity, with $\rho(0) = 0$. For $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$, ($k$ fixed from now on) we set

$$M_t : \{z \in \Omega / \rho_j(z) = t_j, j = 1, \ldots, k\} \tag{4}$$

M_0 will be denoted simply by M; this will be the edge. For an arbitrary function φ on M, we define a function $\tilde{\varphi}$ on Ω by

$$\tilde{\varphi} := \varphi \circ \rho^{-1} \circ \Pi_k \circ \rho, \tag{5}$$

where Π_k is the projection of $\mathbb{R}^{2n}$ onto $\mathbb{R}^{2n-k} \subset \mathbb{R}^{2n}$ given by

$$\Pi_k(t_1, \ldots, t_{2n}) = (0, \ldots, 0, t_{k+1}, \ldots, t_{2n}). \tag{6}$$

Since $\rho(\Omega)$ is a ball, $\Pi_k(\rho(z)) \in \rho(\Omega)$ and hence $\in \rho(M)$, so $\tilde{\varphi}$ is indeed well defined. The sets U_j, U_{-j}, $j = 1, \ldots, 2n$, are defined as follows

$$U_{\pm j} := \{z \in \Omega / \rho_j(z) \gtrless 0\}; \tag{7}$$

furthermore, we set

$$U_{j_1 \ldots j_l} := \bigcap_{s=1}^{l} U_{j_s}. \tag{8}$$

Finally, we denote by $\mathcal{D}(M)$ the space of C^∞-functions on M with compact support, endowed with the usual topology that turns it into a strict inductive

limit of Fréchet spaces. (The topology plays no role in the *formulation* of the theorem below.)

With this notation, we have the following:

Theorem 2.1. *Assume M is generating. There exists a neighborhood V of M such that to every pair of functions f^+, f^-, which are analytic in $U_{12\ldots k}$ and $U_{-1-2\ldots-k}$ respectively and which have the property that for all $\varphi\in\mathscr{D}(M)$*

$$\lim_{\substack{t\to 0\\ t>0}} \int_{M_t} f^+\tilde{\varphi}\,d\sigma_t = \lim_{\substack{t\to 0\\ t<0}} \int_{M_t} f^-\tilde{\varphi}\,d\sigma_t, \tag{9}$$

there exists a common analytic continuation f of f^+ and f^- to V.

Here, $d\sigma_t$ is the euclidean volume element on M_t; $t\gtrless 0$ means that $t_j\gtrless 0$, $j=1,\ldots,k$.

Remarks. 1) Pinčuk's theorem is the case $k=n$, and $dz_1\wedge\ldots\wedge dz_n$ instead of $d\sigma_t$. But on M_t, $dz_1\wedge\ldots\wedge dz_n=\Psi_t d\sigma_t$. Now a weakly (and hence strongly) convergent sequence of distributions evaluated at a convergent sequence of test functions converges to "the limit evaluated at the limit", hence convergence as given by (9) implies convergence in Pinčuk's formulation. Furthermore, Ψ_t is never zero on $\operatorname{supp}\tilde{\varphi}\cap M_t$ for t small, since M is totally real. Thus the argument works in the other direction also, showing that the two formulations are equivalent in the case $k=n$. Compare Sect. 4 for details.

2) The function f *is of course unique*, by uniqueness of analytic continuation.

3) It is essential that M be generating. In the following example, M is a 2 dimensional submanifold of $\mathbb{C}^2$ which is generating everywhere except at $0\in M$. Yet this suffices to make the theorem fail.

Example. We take $\rho\colon\mathbb{C}^2\to\mathbb{R}^4$ to be

$$\begin{aligned}
\rho_1(z_1,z_2)&:=x_2\\
\rho_2(z_1,z_2)&:=y_2-(y_1)^2-(x_1)^2\\
\rho_3(z_1,z_2)&:=x_1\\
\rho_4(z_1,z_2)&:=y_1.
\end{aligned} \tag{10}$$

M is the manifold defined by $\rho_1=\rho_2=0$. Then M is generating everywhere except at 0, where the tangent space coincides with the complex plane $\{(z_1,z_2)/z_2=0\}$. The set U_{12} is given by

$$U_{12}=\{z/x_2>0, y_2>(x_1)^2+(y_1)^2\}$$

and for U_{-1-2} *we have* (11)

$$U_{-1-2}=\{z/x_2<0, y_2<(x_1)^2+(y_1)^2\}.$$

By $(z)^{\frac{1}{2}}$ we denote that branch of the square root where $0<\arg z<2\pi$ (i.e. the complex plane is cut along the positive x-axis). By setting $(0)^{\frac{1}{2}}:=0$, we obtain a function defined and continuous on $\mathbb{C}\setminus\{z/y=0, x>0\}$ and analytic in the interior of this set, which we also denote by $(\ )^{\frac{1}{2}}$. The functions f^+, f^- are now

defined as

$$\begin{aligned} f^+(z) &:= (z_2)^{\frac{1}{2}}, \quad z \in U_{12} \\ f^-(z) &:= (z_2)^{\frac{1}{2}}, \quad z \in U_{-1-2}. \end{aligned} \tag{12}$$

Then, f^+ and f^- are indeed analytic in U_{12} and U_{-1-2} respectively. Let us verify (9). The continuity of $(\ )^{\frac{1}{2}}$ (also at 0!) implies that both limits are equal to

$$\int_M (z_2)^{\frac{1}{2}} \varphi(z) d\sigma_M,$$

so (9) also holds. However, a common analytic continuation of f^+ and f^- to some neighborhood of M (or to some neighborhood of 0, for that matter) would yield an analytic continuation of the square root to a full neighborhood of $0 \in \mathbb{C}$, which is impossible. □

4) The theorem entails the usual *boundary uniqueness result:* if two functions h^+ and g^+ are analytic in $U_{12\ldots k}$ and

$$\lim_{\substack{t \to 0 \\ t>0}} \int h^+ \tilde{\varphi} d\sigma_t = \lim_{\substack{t \to 0 \\ t>0}} \int g^+ \tilde{\varphi} d\sigma_t \quad \forall \varphi \in \mathscr{D}(M) \tag{13}$$

then $h^+ \equiv g^+$ in $U_{12\ldots k}$. (Set $f^+ := h^+ - g^+$, $f^- \equiv 0$ and apply the theorem. By analytic continuation $f^+ = 0^+$ in all of $U_{12\ldots k}$). □

5) The theorem is *local* in nature. If there exists an analytic continuation to some neighborhood of every point in M, then a continuation to a neighborhood of M is easily constructed, again by using uniqueness of analytic continuation (all local extensions continue f^+, f^-!).

6) The case $k=1$ is somewhat special. In this case, instead of a theorem on analytic continuation into some open set, we rather have a theorem on the removal of a singularity. Roughly, if two functions admit the same distribution boundary value from either side of a hypersurface, then there exists an analytic continuation across the surface, i.e. the surface is a removeable singularity. This result can also be obtained from the structure of analytic functions admitting a distribution boundary value on a hypersurface. This is carried out in [9], where these functions are characterized by either the existence of an analytic primitive which is continuous up to the surface or by at most polynomial growth in 1/(distance to the surface) near the surface. For results along these lines, (including harmonic functions) see also [10].

We postpone the proof of Theorem 2.1 until Sect. 5. In Sect. 4 we will show that convergence of f^+, f^- as in (9) implies *convergence in a space of distribution-valued C^∞-functions.* This will be a consequence of a "partially elliptic" a priori estimate. Since we did not find a reference for precisely the statement we need and since we must introduce some notation anyway, Sect. 3 is devoted to that estimate, which is completely analogous to a priori estimates for an elliptic operator.

3. A "Partially Elliptic" Estimate

On an open subset $\Omega \subseteq \mathbb{R}^N$ we consider the following system for $u \in \mathscr{D}'(\Omega)$:

$$L_j u := \sum_{l=1}^{m} a_{jl}(x) \frac{\partial u}{\partial x_l} = v_j \qquad j = 1, 2, \ldots, q \tag{1}$$

where $m \leqq N$ and the a_{jl} are C^∞-functions in some neighborhood of $\bar{\Omega}$ with values in $\mathbb{C}$. We will need the following non isotropic Sobolev spaces: for $s, r \in \mathbb{R}$ we set

$$H_{s,r}(\mathbb{R}^N) := \{u \in S'(\mathbb{R}^N) / \|u\|_{s,r}^2 := \int |\hat{u}(\xi)|^2 (1 + |\xi_1|^2 + \ldots + |\xi_m|^2)^s (1 + |\xi_{m+1}|^2 + \ldots + |\xi_N|^2)^r d\xi < \infty\}. \tag{2}$$

The hat denotes the Fourier transform, S' is the space of tempered distributions. The superscript "loc" denotes the usual localized versions of the spaces, i.e.

$$H_{s,r}^{\mathrm{loc}}(\Omega) := \{u \in \mathscr{D}'(\Omega) \mid \varphi u \in H_{s,r}(\mathbb{R}^N), \forall \varphi \in \mathscr{D}(\Omega)\}. \tag{3}$$

The topology induced by the seminorms $\|u\|_\varphi := \|\varphi u\|_{s,r}$, $\varphi \in \mathscr{D}(\Omega)$, turns these spaces into Fréchet spaces. Finally, we set

$$Lu := (L_1 u, L_2 u, \ldots, L_q u)$$

and

$$\|v\|_{s,r} := \left(\sum_{j=1}^{q} \|v_j\|_{s,r}^2 \right)^{\frac{1}{2}}, \quad v_j \in H_{s,r}(\mathbb{R}_N) \qquad j = 1, \ldots, q. \tag{4}$$

The estimate we have in mind is:

Lemma 3.1. *Assume that the columns of the matrix (a_{jl}) are linearly independent over $\mathbb{R}$ on $\bar{\Omega}$. If $u \in H_{s-1,r}^{\mathrm{loc}}(\Omega)$ and $Lu \in H_{s-1,r}^{\mathrm{loc}}(\Omega)$, then $u \in H_{s,r}^{\mathrm{loc}}(\Omega)$. Moreover, $\forall \varphi \in \mathscr{D}(\Omega)$ there exists a constant C such that*

$$\|\varphi u\|_{s,r} \leqq C \left[\|\varphi Lu\|_{s-1,r} + \sum_{l=1}^{m} \left\| \left(\frac{\partial \varphi}{\partial x_l} \right) u \right\|_{s-1,r} + \|\varphi u\|_{s-1,r} \right]. \tag{5}$$

Remark. "$Lu \in H_{s-1,r}^{\mathrm{loc}}(\Omega)$" and φLu have their obvious meaning.

Proof of Lemma. The proof can be extracted from the proof of local regularity of elliptic operators, as presented in [4], Chap. 6, Sect. B. We indicate the necessary adaptions. The assumption on the linear independence of the columns of (a_{jl}) implies $a \cdot \xi \neq 0$ for $\xi \neq 0 (\xi \in \mathbb{R}^m)$, hence

$$\left(\sum_{j=1}^{q} \left| \sum_{l=1}^{m} a_{jl}(x) \xi_l \right|^2 \right)^{\frac{1}{2}} \geqq \lambda \left(\sum_{l=1}^{m} |\xi_l|^2 \right)^{\frac{1}{2}} \tag{6}$$

for some constant $\lambda > 0$ (independent of x!). From here on, the proof proceeds exactly as the proof of estimate 6.25 and Lemma 6.28 in [4], with 6.24 replaced by (6). The non isotropic versions of the theorems used in these proofs are

easily seen to be true. We get $u \in H^{\mathrm{loc}}_{s,r}(\Omega)$ and

$$\left\| \frac{\partial}{\partial x_l}(\varphi u) \right\|_{s-1,r} \leqq C_1 [\|L(\varphi u)\|_{s-1,r} + \|\varphi u\|_{s-1,r}] \qquad l = 1, \ldots, m. \tag{7}$$

Note that $L(\varphi u) = \varphi L u + [L, \varphi] u \in H_{s-1,r}(\mathbb{R}^N)$, since the commutator is a multiplication by a function in $\mathscr{D}(\Omega)$. If we use this equality again, calculate the commutator and estimate $\|\varphi u\|_{s,r}$ by $C_2 \left[\left\| \frac{\partial}{\partial x_1}(\varphi u) \right\|_{s-1,r} + \ldots + \left\| \frac{\partial}{\partial x_m}(\varphi u) \right\|_{s-1,r} + \|\varphi u\|_{s-1,r} \right]$, we see that (7) implies (5). □

4. Convergence of CR-Distributions

Let $P \in M$ and $(\xi_{k+1}, \ldots, \xi_{2n}) \in U$ be arbitrary coordinates on M in a neighborhood of P. Then the ξ_j, $j = k+1, \ldots, 2n$ serve also as coordinates on M_t "by translation of M along the $\rho_1, \ldots, \rho_k$-lines", that is $z_t(\xi)$ is the unique point on M_t such that $\Pi_k \circ \rho(z_t(\xi)) = \rho(z(\xi))$.

The integrals appearing in (2.9) become

$$\int_{M_t} f^+ \tilde{\varphi} \, d\sigma_t = \int_U f^+(z_t(\xi)) \varphi(\xi) \left(\frac{d\sigma_t}{d\xi} \right)(\xi) \, d\xi = \left\langle f_t^+, \varphi \left(\frac{d\sigma_t}{d\xi} \right) \right\rangle \qquad \varphi \in \mathscr{D}(U) \tag{1}$$

where $\frac{d\sigma_t}{d\xi}$ denotes the density of the euclidean volume element on M_t with respect to the given coordinates, f_t^+ has the obvious meaning and $\langle \ , \ \rangle$ is the duality between $\mathscr{D}'(U)$ and $\mathscr{D}(U)$. Note that

and

$$\left.\begin{aligned} \frac{d\sigma_t}{d\xi} &\xrightarrow[t \to 0]{} \frac{d\sigma_0}{d\xi} \\ \left(\frac{d\sigma_t}{d\xi} \right)^{-1} &\xrightarrow[t \to 0]{} \left(\frac{d\sigma_0}{d\xi} \right)^{-1} \end{aligned}\right\} \text{ in } C^\infty(U). \tag{2}$$

Therefore ([8]: theorème XI, chapitre III) the limits in (1) as $t \to 0$ exist *if and only if*

$$f_0^+ := \lim_{\substack{t \to 0 \\ t > 0}} f_t^+ \in \mathscr{D}'(U) \qquad . \tag{3}$$

exists (we also use the fact that weak convergence of sequences implies their strong convergence in $\mathscr{D}'(U)$).

Let us split the coordinates into two groups:

$$\xi_1 := (\xi_{k+1}, \ldots, \xi_{2n-k}) \in U_1, \xi_2 := (\xi_{2n-k+1}, \ldots, \xi_{2n}) \in U_2,$$

$U = U_1 \times U_2$. With this setup, our result on the convergence of f_t^+ is:

Theorem 4.1. *Assume that for all* $\xi \in U$

$$T^{\mathbb{C}}_{z(\xi)}(M) \cap \mathrm{sp}_{\mathbb{R}} \left[\frac{\partial z}{\partial \xi_{2n-k+1}}(\xi), \ldots, \frac{\partial z}{\partial \xi_{2n}}(\xi) \right] = \{0\}. \tag{4}$$

If the limit (3) *exists, then the convergence is in fact in* $C^\infty(U_1, \mathscr{D}'(U_2))$.

Here, $\mathrm{sp}_{\mathbb{R}}[\ ,\ ]$ denotes the linear span over $\mathbb{R}$. Note that the theorem asserts in particular that the limit $f_0^+ \in C^\infty(U_1, \mathscr{D}'(U_2))$! This is the space of $\mathscr{D}'(U_2)$-valued C^∞-functions on U_1 with the usual topology of locally uniform convergence of the functions and their derivatives. Then $\tau \in C^\infty(U_1, \mathscr{D}'(U_2))$ means that there exists $g \in C^\infty(U_1, \mathscr{D}'(U_2))$ such that

$$\langle \tau, \chi \rangle = \int \langle g(\xi_1), \chi(\xi_1, .) \rangle_{\xi_2} d\xi_1 \qquad \forall \chi \in \mathscr{D}(U_1 \times U_2). \tag{5}$$

Proof of Theorem 4.1. Assume that the limit (3) exists. The topology of $C^\infty(U_1, \mathscr{D}'(U_2))$ is local with respect to the variables in U_1 by definition, and with respect to the variables in U_2 because the topology in $\mathscr{D}'(U_2)$ is local. Therefore, to prove convergence in that topology, it suffices to show that every point has a neighborhood $\tilde{U}_1 \times \tilde{U}_2$, such that the restrictions to this neighborhood converge in $C^\infty(\tilde{U}_1, \mathscr{D}'(\tilde{U}_2))$. This will follow from the fact that f^+ is analytic. More precisely we use that f_t^+ is the restriction to M_t of the analytic function f^+ and hence is a solution of *the tangential CR-equations*

$$(D_X + iD_{iX}) f_t^+ = 0 \tag{6}$$

for an arbitrary complex tangent vector X of M_t (D denotes the Lie derivative).

It suffices to argue near P. It is an easy consequence of (2.1) that complex tangent vectors to M_t at $z_t(\xi)$ are exactly the solutions $(\zeta_1, \ldots, \zeta_n)$ of

$$\sum_{s=1}^{n} \frac{\partial \rho_j}{\partial z_s}(z_t(\xi))\, \zeta_s = 0 \qquad j = 1, \ldots, k. \tag{7}$$

M generating means that this system for $t=0$ has rank k at all points. Therefore, for some $\tilde{U}_1 \Subset U_1$, $\tilde{U}_2 \Subset U_2$ and t_0 small enough, there exist smooth (also in t) complex vector fields $X_{j,t}(\xi)$ on M_t, $1 \leqq j \leqq n-k$, such that they span the complex tangent space at $z_t(\xi)$, $|t| < t_0$, $\xi \in \overline{\tilde{U}_1 \times \tilde{U}_2}$. In local coordinates, assume

$$X_{j,t}(\xi) = \sum_{s=k+1}^{2n} \alpha_{j,t}^s(\xi) \frac{\partial z_t}{\partial \xi_s}$$

and

$$iX_{j,t}(\xi) = \sum_{s=k+1}^{2n} \beta_{j,t}^s(\xi) \frac{\partial z_t}{\partial \xi_s}. \tag{8}$$

Then the tangential CR-equations (6) become: for $1 \leqq j \leqq n-k$

$$\sum_{s=k+1}^{2n-k} (\alpha_{j,t}^s(\xi) + i\beta_{j,t}^s(\xi)) \frac{\partial f_t^+(\xi)}{\partial \xi_s} = -\sum_{s=2n-k+1}^{2n} (\alpha_{j,t}^s(\xi) + i\beta_{j,t}^s(\xi)) \frac{\partial f_t^+(\xi)}{\partial \xi_s}, \tag{9}$$

where we have split the sum in anticipation of an application of Lemma 3.1: we claim that the columns of the left side are independent over $\mathbb{R}$! Assume they are not for some (ξ, t). Then, the square matrix (γ_j^s) with

$$\gamma_j^s := \begin{cases} \alpha_{j,t}^s(\xi) & 1 \leqq j \leqq n-k \\ \beta_{j-(n-k)}^s(\xi) & n-k < j \leqq 2n-2k \end{cases} \qquad s = k+1, \ldots, 2n-k \tag{10}$$

is singular (note that the α's and β's are real) and there exists a nontrivial real linear combination of its rows which is zero. Let $\lambda_1, \ldots, \lambda_{2n-2k}$ be the coefficients of such a combination. Then

$$\sum_{s=1}^{n-k} \lambda_s X_{s,t}(\xi) + \sum_{s=1}^{n-k} \lambda_{n-k+s}(iX_{s,t}(\xi)) \in \mathrm{sp}_{\mathbb{R}}\left[\frac{\partial z_t(\xi)}{\partial \xi_{2n-k+1}}, \ldots, \frac{\partial z_t(\xi)}{\partial \xi_{2n}}\right].$$

(4), which is a linear independence condition, and therefore holds on M_t, for t small enough, implies that the above *complex* tangent vector must be zero. But this means that

$$\sum_{s=1}^{n-k} (\lambda_s + i\lambda_{n-k+s}) X_{s,t}(\xi) = 0. \tag{11}$$

Since not all of $\lambda_1, \ldots, \lambda_{2n-2k}$ are 0, this contradicts the fact that $X_{1,t}(\xi), \ldots, X_{n-k,t}(\xi)$ span the complex tangent space (and therefore are independent over $\mathbb{C}$). This proves our claim. Therefore, Lemma 3.1 applies to system (9), with $m = 2n-2k$, $q = n-k$ and

$$L_{j,t} = \sum_{s=k+1}^{2n-k} (\alpha^s_{j,t}(\xi) + i\beta^s_{j,t}(\xi)) \frac{\partial}{\partial \xi_s}, \quad j = 1, \ldots, n-k. \tag{12}$$

t here plays the role of a parameter. The constant C in the crucial estimate (3.5) can be chosen to be independent of $t (|t| < t_0)$; this is seen by inspection of the proof! Note that by continuity it follows from (3) that f_0^+ is a solution of (9) for $t = 0$: *the boundary value of an analytic function is a* CR*-distribution.*

Since $\tilde{U}_1 \times \tilde{U}_2$ is relatively compact in $U_1 \times U_2$, it follows from (3) and the local structure of distributions that there exist $s_0, r_0 \in \mathbb{R}$ such that

$$f_0^+ = \lim_{\substack{t \to 0 \\ t > 0}} f_t^+ \quad \text{in } H^{\mathrm{loc}}_{s_0,r_0}(\tilde{U}_1 \times \tilde{U}_2). \tag{13}$$

Taking into account the smoothness of the α^s_j and β^s_j (note that here t is *not* restricted to $t > 0$!), we conclude from (13) that the right hand side of (9) converges in $H^{\mathrm{loc}}_{s_0,r_0-1}(\tilde{U}_1 \times \tilde{U}_2)$; therefore, the same is true for the left hand side:

$$L_{j,t} f_t^+ \xrightarrow[t \to 0]{} L_{j,0} f_0^+ \quad \text{in } H^{\mathrm{loc}}_{s_0,r_0-1}(\tilde{U}_1 \times \tilde{U}_2). \tag{14}$$

Thus we have $L_{j,0} f_0^+ \in H^{\mathrm{loc}}_{s_0,r_0-1}$, $1 \leqq j \leqq n-k$, and $f_0^+ \in H^{\mathrm{loc}}_{s_0,r_0} \subset H^{\mathrm{loc}}_{s_0,r_0-1}$ and hence by Lemma 3.1 we get $f_0^+ \in H^{\mathrm{loc}}_{s_0+1,r_0-1}$. Next we want to show that convergence also takes place in $H^{\mathrm{loc}}_{s_0+1,r_0-1}$. Fix $\varphi \in \mathscr{D}(\tilde{U}_1 \times \tilde{U}_2)$. First we observe that (13), (14) and (3.5) imply that there is a constant $C_1 = C_1(\Psi)$ for all $\Psi \in \mathscr{D}(\tilde{U}_1 \times \tilde{U}_2)$ such that

$$\|\Psi f_t^+\|_{s_0+1,r_0-1} \leqq C_1, \quad 0 \leqq |t| < t_0. \tag{15}$$

(Here it is important that the constant in (3.5) is independent of t). Next, we apply (3.5) for L_0 to $(f_t^+ - f_0^+)$ and get, upon adding and subtracting $L_t f_t^+$ to $L_0(f_t^+ - f_0^+)$:

$$\|\varphi(f_t^+ - f_0^t)\|_{s_0+1,r_0-1} \leqq C\Big[\|\varphi(L_t f_t^+ - L_0 f_0^+)\|_{s_0,r_0-1} + \|\varphi(L_t - L_0) f_t^+\|_{s_0,r_0-1}$$
$$+ \sum_{l=1}^{m} \left\| \frac{\partial \varphi}{\partial x_l}(f_t^+ - f_0^+) \right\|_{s_0,r_0-1} + \|\varphi(f_t^+ - f_0^+)\|_{s_0,r_0-1}\Big]. \quad (16)$$

By (13) and (14), the first and the last two terms on the right side of (16) tend to 0 as $t \to 0$. The smoothness of the coefficients of L_t (in t) and (15) imply

$$\|\varphi(L_t - L_0) f_t^+\|_{s_0,r_0-1} \leqq C_2(\varphi)|t|. \quad (17)$$

Thus indeed

$$\|\varphi(f_t^+ - f_0^+)\|_{s_0+1,r_0-1} \xrightarrow[t\to 0]{} 0. \quad (18)$$

Since $\varphi \in \mathscr{D}(\tilde{U}_1 \times \tilde{U}_2)$ was arbitrary, this proves convergence in $H^{\mathrm{loc}}_{s_0+1,r_0-1}$. Substituting this into (13) and repeating the argument we get convergence in $H^{\mathrm{loc}}_{s_0+2,r_0-2}$, and eventually, by iteration, convergence in $H^{\mathrm{loc}}_{s_0+j,r_0-j}(\tilde{U}_1 \times \tilde{U}_2)$, *for all integers* j! Fix $N \in \mathbb{N}$. By the usual Fourier transform argument (proving, so to speak, a non isotropic version of the Sobolev lemma), we get that $H^{\mathrm{loc}}_{s_0+j,r_0-j}(\tilde{U}_1 \times \tilde{U}_2)$ is continuously imbedded in $C^N(\tilde{U}_1, H^{\mathrm{loc}}_{r_0-j}(\tilde{U}_2))$ for $s_0 + j > N + \frac{1}{2}(2n-2k)$. This space itself is continuously imbedded in $C^N(\tilde{U}_1, \mathscr{D}'(\tilde{U}_2))$, so that instead of (3) we now have convergence in $C^N(\tilde{U}_1, \mathscr{D}'(\tilde{U}_2))$. Since N was arbitrary, we have indeed

$$f_0^+ = \lim_{\substack{t\to 0 \\ t>0}} f_t^+ \quad \text{in } C^\infty(\tilde{U}_1, \mathscr{D}'(\tilde{U}_2)) \quad (19)$$

and the proof of Theorem 4.1 is complete. □

Remarks. 1) In the above proof, all that was used of the analyticity of f^+ was the fact that f_t^+ was a distribution solution of the tangential CR-equations on M_t (we did *not* use that for $t > 0$, f_t^+ is C^∞!). *So Theorem* 4.1 *remains true with* f_t^+ *replaced by a family* τ_t *of* CR-*distributions on* M_t. "CR-distribution" refers to a distribution solution of the tangential CR-equations (6).

2) The above proof (slightly simplified) also applies to the situation where all τ_t are CR on the same manifold, say M, and yields these *regularity properties for* CR-*distributions:*

a) Any CR-distribution τ on $U \subset M$ is in fact an element of $C^\infty(U_1, \mathscr{D}'(U_2))$, as defined by (5). (Compare also, for example, [11], Prop. 5.2, for this result.)

b) If $(\tau_n)_1^\infty$ is a sequence of CR-distributions on $U \subset M$, converging in $\mathscr{D}'(U)$, then convergence takes place in $C^\infty(U_1, \mathscr{D}'(U_2))$.

c) **Corollary.** *Assume that the sequence* $(\tau_n)_1^\infty$ *of* CR-*distributions is bounded in* $\mathscr{D}'(U_1 \times U_2)$. *Then, there exists a subsequence which converges in* $C^\infty(U_1, \mathscr{D}'(U_2))$.

Proof. Bounded sets in $\mathscr{D}'(U_1 \times U_2)$ are sequentially compact. Together with b) above, this implies the corollary. □

3) Property 2a) above holds on arbitrary CR-manifolds. However, *there are manifolds where the system* (6) *is hypoelliptic*, that is, all distribution solutions are C^∞-functions with respect to *all* variables. Assume that M has the property

that near $P \in M$, every sufficiently differentiable CR-function can be extended as an analytic function to a *full* neighborhood of P in $\mathbb{C}^n$. A sufficient condition for this to happen is given in [3], Theorem 1.2, in terms of the image of the Levi form. Now it is an important fact that, locally, for every CR-distribution τ there exists a CR-function f of prescribed (but finite) differentiability, such that $\tau = (\Delta/M)^N f$. Here, Δ/M is a differential operator on M such that for functions h which are analytic in a neighborhood of P, the restriction to M of $\left(\frac{\partial^2}{\partial z_1^2} + \ldots + \frac{\partial^2}{\partial z_n^2}\right) h$ is equal to Δ/M applied to the restriction of h. This fact follows from [11], Theorem 3.1. A slightly refined version is given in [9] (for the case of tangential CR-equations). At present, all that matters is that τ has a CR-primitive f of a sufficient differentiability. By our assumption, f will be the restriction of an analytic function and therefore smooth (M is smooth!); so τ is also smooth. Thus (6) is hypoelliptic near P.

5. Proof of Theorem 2.1[1]

As mentioned earlier, it suffices to prove the theorem locally. Fix $P \in M$. For simplicity, assume that $\rho(P) = 0$. On M, we take coordinates $(\xi_{k+1}, \ldots, \xi_{2n})$ given by

$$\xi_j(z) := \rho_j(z), \quad j = k+1, \ldots, 2n \tag{1}$$

M is generating, so $\dim_{\mathbb{R}} T^{\mathbb{C}}(M) = 2n - 2k$. Therefore, there exist $j_1, \ldots, j_k$ such that

$$\mathrm{sp}_{\mathbb{R}}\left[\frac{\partial z}{\partial \xi_{j_1}}(P), \ldots, \frac{\partial z}{\partial \xi_{j_k}}(P)\right] \cap T_p^{\mathbb{C}}(M) = \{0\}. \tag{2}$$

This implies that the vectors $\frac{\partial z}{\partial \xi_{j_1}}(P), \ldots, \frac{\partial z}{\partial \xi_{j_k}}(P)$ are independent over $\mathbb{C}$. Using again that M is generating, we conclude from this that this set of vectors can be augmented by $n-k$ of the vectors $\left\{\frac{\partial z}{\partial \xi_j}(P)/k+1 \leqq j \leqq 2n\right\}$ to a set of n vectors which, over $\mathbb{C}$, spans all of $\mathbb{C}^n$. Possibly after permuting the $\rho_{k+1}, \ldots, \rho_{2n}$ (which changes nothing in the theorem!), we may therefore assume

$$\mathrm{sp}_{\mathbb{R}}\left[\frac{\partial z}{\partial \xi_{2n-k+1}}, \ldots, \frac{\partial z}{\partial \xi_{2n}}\right] \cap T^{\mathbb{C}}(M) = \{0\} \tag{3}$$

and

$$\mathrm{sp}_{\mathbb{C}}\left[\frac{\partial z}{\partial \xi_{n+1}}, \ldots, \frac{\partial z}{\partial \xi_{2n}}\right] = \mathbb{C}^n \tag{4}$$

in a certain neighborhood of P. Let us take this neighborhood to be of the form $U = U_{1,1} \times U_{1,2} \times U_2$, where

$$(\xi_{k+1}, \ldots, \xi_n) \in U_{1,1}, \quad (\xi_{n+1}, \ldots, \xi_{2n-k}) \in U_{1,2}$$

and

$$(\xi_{2n-k+1}, \ldots, \xi_{2n}) \in U_2.$$

[1] See note added in proof

(2.9) and (3) imply that the assumptions of Theorem 4.1 are fulfilled with $U_1 := U_{1,1} \times U_{1,2}$ (see the discussion at the beginning of Sect. 4). Therefore, with the notation as in Sect. 4, we have

$$f_0^+ = \lim_{\substack{t \to 0 \\ t > 0}} f_t^+ \quad \text{in } C^\infty(U_{1,1} \times U_{1,2}, \mathscr{D}'(U_2)). \tag{5}$$

We need one more notation: for $\theta \in U_{1,1}$ we denote by $f_{t,\theta}^+$ the following function on $U_{1,2} \times U_2$:

$$f_{t,\theta}^+(\xi_{n+1}, \ldots, \xi_{2n}) := f_t^+(\theta_{k+1}, \ldots, \theta_n, \xi_{n+1}, \ldots, \xi_{2n}) \tag{6}$$

From (5), we have a fortiori convergence in $C^\infty(U_{1,1}, \mathscr{D}'(U_{1,2} \times U_2))$. In particular, the convergence is locally uniform on $U_{1,1}$, whence

$$f_{t,\theta}^+ \xrightarrow[\substack{t \to 0, t > 0 \\ \theta \to 0, \theta > 0}]{} f_{0,0}^+ \quad \text{in } \mathscr{D}'(U_{1,2} \times U_2). \tag{7}$$

The above arguments also apply to the function f^-; thus (with analogous notation):

$$f_{t,\theta}^- \xrightarrow[\substack{t \to 0, t < 0 \\ \theta \to 0, \theta < 0}]{} f_{0,0}^-. \tag{8}$$

The equality of the limits (2.9) yields

$$f_{0,0}^+ = f_{0,0}^- \tag{9}$$

(7), (8) and (9) together with the discussion at the beginning of Sect. 4 will put us in a position to apply Pinčuk's theorem ([6], Theorem 1; see also Remark 2.1 for the statement of the theorem) provided the manifold

$$M := \{z / \rho_1(z) = \ldots = \rho_n(z) = 0\}$$

is generating near P. This is indeed the case by (4). Therefore, we conclude from [6], Theorem 1, that there exists an analytic function f, defined in some ball B around P, such that

$$f = \begin{cases} f^+ & \text{in } U_{123\ldots n} \cap B \\ f^- & \text{in } U_{-1-2\ldots n} \cap B \end{cases}. \tag{10}$$

By uniqueness of analytic continuation, these equalities extend to $U_{12\ldots k} \cap B$ and $U_{-1-2\ldots-k} \cap B$ respectively, so that f is indeed a common analytic continuation of f^+ and f^-. The ball B can be chosen independently of f^+, f^-, see [6], §1, Remark 4. As P was arbitrary, we can invoke Remark 2.5 and end up with a neighborhood V of M having the desired property. This concludes the proof of Theorem 2.1. □

Acknowledgement. The author thanks Professor K. Osterwalder for introducing him to some concepts and ideas in several complex variables which have their origin in mathematical physics.

References

1. Bedford, E.: Holomorphic continuation at a totally real edge. Math. Ann. **230**, 213–235 (1977)
2. Bros, J., Iagolnitzer, D.: Causality and Local Analyticity: Mathematical Study. Ann. Inst. H. Poincaré, Sect. A, **18**, 147–184 (1973)
3. Boggess, A., Polking, J.: Holomorphic Extension of CR-Functions. Duke Math. J. **49**, 757–784 (1982)
4. Folland, G.: Introduction to Partial Differential Equations. Math. Notes #17. Princeton: Princeton Univ. Press 1976
5. Kerzman, N.: Four Lectures on Several Complex Variables. Math. Inst., Univ. of Amsterdam, Roeterstraat 15, 1018 WB Amsterdam, 1983
6. Pinčuk, S.: Bogoljubov's Theorem on the Edge-of-the-Wedge for Generic Manifolds. Math. USSR Sbornik **23**, 441–455 (1974)
7. Rudin, W.: Lectures on the Edge-of-the-Wedge Theorem. Reg. Conf. Series in Math., no 6, AMS 1971
8. Schwartz, L.: Théorie des distributions. Paris: Hermann 1978
9. Straube, E.: Cauchy-Riemann Distributions and Boundary Values of Analytic Functions. ETH dissertation, no. 7351, Zurich 1983
10. Straube, E.: Harmonic and Analytic Functions Admitting a Distribution Boundary Value. To appear in Ann. Scuola Norm. Sup. Pisa
11. Trèves, F.: Approximation and Representation of Functions and Distributions Annihilated by a System of Complex Vector Fields. Editions de l'école polytechnique, centre de mathématiques, Palaiseau, France, 1981
12. Vladimirov, V.: Methods of the Theory of Functions of Many Complex Variables. M.I.T. Press, Boston 1966
13. Webster, S.: On the Reflection Principle in Several Complex Variables. Proc. AMS **71**, I 26–28 (1978)

Received December 6, 1983

Note added in proof: The proof we give here formally applies only for $n>1$, since Pinčuk's Edge-of-the-Wedge theorem in [6] is stated only for $n>1$. However, the case $n=1$ is classical, compare for ex. A. Martineau: Distributions et valeurs au bord des fonctions holomorphes, in Theory of Distributions, Proc. Int. Summer Inst. Lisbon, Sept. 1964, chapitre I, theorème VIII', and the remark following it. Also, the argument involving the Cauchy integral representation, familiar from the case of continuous boundary values, extends to the distribution case. Compare for ex. Remark 2.6 in J. Cima – E. Straube: Boundary Uniqueness Theorems in $\mathbb{C}^n$, preprint.

Math. Z. 189, 143–150 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Degree of Best Approximation by Trigonometric Blending Functions

Werner Haußmann, Kurt Jetter and Bernd Steinhaus

Department of Mathematics, University of Duisburg, D-4100 Duisburg, Federal Republic of Germany

1. Introduction

In the last years, there was a considerable interest in approximation of multivariate functions with respect to *infinite dimensional vector spaces* of approximating functions, and, in particular, with respect to *blending functions*. For a recent survey we refer to Cheney [3].

Let us first state the approximation problem in the following framework. Suppose, we are given two real normed vector spaces $(X, \|.\|_X)$ and $(Y, \|.\|_Y)$ and subspaces U and V, respectively. If the tensor product $X \otimes Y$ is provided with a cross-norm α satisfying

$$\alpha(x \otimes y) = \|x \otimes y\|_\alpha = \|x\|_X \|y\|_Y \quad \text{for all } x \in X,\ y \in Y,$$

(see Gilbert-Leih [6]), then the following approximation problem is considered: Given an $h \in X \otimes_\alpha Y$ (completion of $X \otimes Y$ under α), does there exist an element $w^* \in U \otimes Y + X \otimes V \subset X \otimes_\alpha Y$ such that

$$\|h - w^*\|_\alpha \leqq \|h - w\|_\alpha \quad \text{for all } w \in U \otimes Y + X \otimes V.$$

w^* is called a best approximation (or proximum) to h with respect to $U \otimes Y + X \otimes V$ in the norm α. The approximation constant of h with respect to $U \otimes Y + X \otimes V$ and the norm α is defined by

$$E^\alpha_{U \otimes Y + X \otimes V}(h) := \inf_{w \in U \otimes Y + X \otimes V} \|h - w\|_\alpha.$$

Among the main questions of interest are those concerning existence and uniqueness of a proximum w^* as well as its characterization, and the calculation of (bounds for) the approximation constants under various norms for individual functions or classes of functions.

There is a large number of papers which deal with problems of the type described above, especially concerning existence of best approximations, algorithmic aspects and projection constants. Besides the survey papers by Cheney

[2,3] and the references given there, we mention von Golitschek [7], von Golitschek-Cheney [8], Halton-Light [11,12] and Haußmann-Zeller [14,15]. These papers mostly deal with approximation of functions in the bivariate Chebyshev norm as well as in the bivariate L_p norms (especially $p=1$), but there are also investigations concerning mixed norm approximation problems (Cheney-McCabe-Phillips [4], Haußmann-Zeller [16]).

For the following, let us adopt these notations. In the case of function spaces X, Y, we shall call

$$\mathscr{B}(U,X;V,Y):=U\otimes Y+X\otimes V$$

the vector space of *blending functions* (corresponding to $U\subset X$ and $V\subset Y$). This concept was introduced by Gordon [9,10], who was the first to consider blending interpolation and approximation methods. Further, let $\mathbb{N}:=\{1,2,3,\ldots\}$, and $\mathbb{N}_0:=\mathbb{N}\cup\{0\}$. We shall use the (ordinary resp. partial) differential operators

$$D^r:=\frac{d^r}{dx^r} \quad \text{for } r\in\mathbb{N}_0, \qquad D^{r,s}:=\frac{\partial^{r+s}}{\partial x^r\,\partial y^s} \quad \text{for } r,s\in\mathbb{N}_0.$$

Here D^0 resp. $D^{0,0}$ mean the corresponding identity operators. For $r,s\in\mathbb{N}_0$, we define the following function spaces:

$$\mathscr{C}^r_{2\pi}:=\mathscr{C}^r_{2\pi}(\mathbb{R}):=\{\varphi:\mathbb{R}\to\mathbb{R}\,|\,D^\rho\varphi \text{ continuous and } 2\pi\text{-periodic for } 0\leqq\rho\leqq r\},$$

$$\mathscr{C}^{r,s}_{2\pi}:=\mathscr{C}^{r,s}_{2\pi}(\mathbb{R}^2):=\{\varphi:\mathbb{R}^2\to\mathbb{R}\,|\,D^{\rho,\sigma}\varphi \text{ continuous, and } 2\pi\text{-periodic with respect to both variables, for } 0\leqq\rho\leqq r, 0\leqq\sigma\leqq s\}.$$

For $r=s=0$, we write $\mathscr{C}_{2\pi}(\mathbb{R})$ and $\mathscr{C}_{2\pi}(\mathbb{R}^2)$ instead of $\mathscr{C}^0_{2\pi}(\mathbb{R})$ and $\mathscr{C}^{0,0}_{2\pi}(\mathbb{R}^2)$, respectively. In addition, we shall use $f^{(r)}:=D^rf$ and $h^{(r,s)}:=D^{r,s}h$ for functions $f\in\mathscr{C}^r_{2\pi}$ and $h\in\mathscr{C}^{r,s}_{2\pi}$ $(r,s\in\mathbb{N}_0)$. By $\mathscr{T}_k$ we denote the vector space of all trigonometric polynomials of degree $\leqq k\in\mathbb{N}_0$. Finally, for $1\leqq p\leqq\infty$ and $S:=[-\pi,\pi]$, we consider the classes $L_p(S)$ and $L_p(S^2)$ endowed with

$$\|\varphi\|_p:=\Big(\int_E|\varphi|^p\Big)^{1/p} \qquad (1\leqq p<\infty)$$

and

$$\|\varphi\|_\infty:=\operatorname*{ess\,sup}_E|\varphi|,$$

where $E=S$ in the univariate and $E=S^2$ in the bivariate case. For $\mathscr{C}_{2\pi}$-functions, $\|.\|_\infty$ is the usual *Chebyshev norm.*

In the present paper we shall prove the following estimate for the approximation constant of an $h\in\mathscr{C}^{r,s}_{2\pi}(\mathbb{R}^2)$, $r,s\geqq1$: Let $1\leqq p\leqq\infty$, then

$$E^p_{\mathscr{B}(\mathscr{T}_{m-1},\mathscr{C}_{2\pi};\mathscr{T}_{n-1},\mathscr{C}_{2\pi})}(h)\leqq\frac{K_r}{m^r}\cdot\frac{K_s}{n^s}\,\|D^{r,s}h\|_p,$$

where the (so-called Favard) constants K_q satisfy $\frac{\pi^2}{8} \leqq K_q \leqq \frac{\pi}{2}$ for any $q \in \mathbb{N}$. In addition, we can show that the approximation constant of a function $h \in \mathscr{C}_{2\pi}(\mathbb{R}^2)$ is independent of the spaces of coefficient functions within a large scale. To be more precise, for $h \in \mathscr{C}_{2\pi}(\mathbb{R}^2)$ and $m, n \in \mathbb{N}$ we have

$$E^p_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})}(h) = E^p_{\mathscr{B}(\mathscr{T}_{m-1}, X; \mathscr{T}_{n-1}, Y)}(h)$$

whenever X and Y are dense subspaces of $\mathscr{C}_{2\pi}$ (with respect to the L_p norm, $1 \leqq p \leqq \infty$) satisfying $\mathscr{T}_{m-1} \subset X$, $\mathscr{T}_{n-1} \subset Y$. Moreover, using a result due to von Golitschek-Cheney [8], we get

$$E^\infty_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})}(h) = E^\infty_{\mathscr{B}(\mathscr{T}_{m-1}, M(S); \mathscr{T}_{n-1}, M(S))}(h)$$

for $h \in \mathscr{C}_{2\pi}(\mathbb{R}^2)$ and $m, n \in \mathbb{N}$, where we denote the vector space of all bounded real functions on S by $M(S)$. I.e. the degree of uniform approximation of a continuous function cannot be improved using discontinuous functions.

2. Approximation of Differentiable Functions

Our main estimate for the approximation constant of an $h \in \mathscr{C}^{r,s}_{2\pi}(\mathbb{R}^2)$, $r, s \geqq 1$, reads as follows:

Theorem 1. *Let $h \in \mathscr{C}^{r,s}_{2\pi}(\mathbb{R}^2)$, $r, s \in \mathbb{N}$. Then for $m, n \in \mathbb{N}$ and $1 \leqq p \leqq \infty$ we have*

$$E^p_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})}(h) \leqq \frac{K_r}{m^r} \cdot \frac{K_s}{n^s} \| D^{r,s} h \|_p,$$

where the Favard constants are given by

$$K_r := \frac{4}{\pi} \sum_{\rho=0}^{\infty} \left(\frac{(-1)^\rho}{2\rho+1} \right)^{r+1} \quad \text{for } r \in \mathbb{N}.$$

In order to prove this, we shall employ the functions

$$D_r(z) := \frac{1}{2\pi} \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \frac{e^{ikz}}{(ik)^r} \quad \text{for } z \in \mathbb{R}, r \in \mathbb{N}.$$

Let us first list some of their properties (see Hardy-Rogosinski [13, pp. 28–30] and Lorentz [18, Chap. 8]). Obviously, this series is uniformly convergent on $\mathbb{R}$ for $r \geqq 2$. The series defining D_1 converges on $\mathbb{R}$ with uniformly bounded partial sums and the convergence is uniform on compact subsets of $\mathbb{R} \setminus 2\pi\mathbb{Z}$. In addition, for $N \in \mathbb{Z}$ we have

$$D_1(z) = \frac{1}{2\pi} \sum_{\substack{k \in \mathbb{Z} \\ k \neq 0}} \frac{e^{ikz}}{ik} = \begin{cases} \dfrac{\pi - z + 2\pi N}{2\pi} & \text{for } 2\pi N < z < 2\pi(N+1), \\ 0 & \text{for } z = 2\pi N, \end{cases}$$

and

$$D_r'(z)=D_{r-1}(z) \qquad (z\in\mathbb{R})$$

for $r\geqq 2$ (but $z\neq 2\pi N$, $N\in\mathbb{Z}$, if $r=2$). Moreover, the following fundamental convolution formula holds true: If $f\in\mathscr{C}_{2\pi}^r$, then

$$f=a_0+f'*D_1=a_0+f^{(r)}*D_r, \tag{*}$$

where $a_0=\frac{1}{2\pi}\int_S f(t)\,dt$ is the mean value of f, and the convolution operator $*$ is defined by

$$(f*g)(x):=\int_S f(x-t)\,g(t)\,dt$$

for 2π-periodic functions $f,g\in L_1(S)$. Correspondingly, in the bivariate case we shall use

$$(h*\varphi)(x,y):=\int_{S^2} h(x-t,y-u)\,\varphi(t,u)\,dt\,du$$

for $h,\varphi\in L_1(S^2)$, 2π-periodic with respect to both variables.

A blending type analogon of formula (*) is given by the following

Lemma 2. (a) *For any* $h\in\mathscr{C}_{2\pi}^{1,1}(\mathbb{R}^2)$ *we have*

$$h(x,y)=-\frac{1}{4\pi^2}\int_{S^2} h(t,u)\,dt\,du+\frac{1}{2\pi}\int_S h(t,y)\,dt$$
$$+\frac{1}{2\pi}\int_S h(x,u)\,du+\int_{S^2} h^{(1,1)}(t,u)\,D_1(x-t)\,D_1(y-u)\,dt\,du.$$

(b) *If* $h\in\mathscr{C}_{2\pi}^{r,s}(\mathbb{R}^2)$, $r,s\in\mathbb{N}$, *then*

$$\int_{S^2} h^{(1,1)}(t,u)\,D_1(x-t)\,D_1(y-u)\,dt\,du=\int_{S^2} h^{(r,s)}(t,u)\,D_r(x-t)\,D_s(y-u)\,dt\,du.$$

Indeed, in order to prove (a), we use (*) repeatedly and get

$$\int_{S^2} h^{(1,1)}(t,u)\,D_1(x-t)\,D_1(y-u)\,dt\,du$$
$$=\int_S\left[h^{(0,1)}(x,u)-\frac{1}{2\pi}\int_S h^{(0,1)}(t,u)\,dt\right]D_1(y-u)\,du$$
$$=h(x,y)-\frac{1}{2\pi}\int_S h(x,u)\,du-\frac{1}{2\pi}\int_S h(t,y)\,dt+\frac{1}{4\pi^2}\int_{S^2} h(t,u)\,dt\,du.$$

From this, (b) follows by repeated integration by parts.

Proof of Theorem 1. Let $r,s\in\mathbb{N}$, $h\in\mathscr{C}_{2\pi}^{r,s}(\mathbb{R}^2)$, and

$$h^*(x,y):=\int_{S^2} h^{(r,s)}(t,u)\,D_r(x-t)\,D_s(y-u)\,dt\,du.$$

With the aid of Lemma 2 and the differentiation formula for parameter integrals we see that $h-h^*\in\mathscr{B}(\mathscr{T}_0,\mathscr{C}_{2\pi}^r;\mathscr{T}_0,\mathscr{C}_{2\pi}^s)$, in particular $h-h^*\in$

$\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})$ for all $m, n \in \mathbb{N}$. Since we approximate with these spaces of blending functions, we have for $m, n \in \mathbb{N}$

$$E^p_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})}(h) = E^p_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})}(h^*).$$

Now, let u^* and v^* be L_1 proxima for D_r resp. D_s with respect to T_{m-1} resp. T_{n-1}. Then for

$$w(x, y) := u^*(x) D_s(y) + D_r(x) v^*(y) - u^*(x) v^*(y)$$

we have $w \in \mathscr{B}(\mathscr{T}_{m-1}, L_1(S); \mathscr{T}_{n-1}, L_1(S))$. Consider

$$w^*(x, y) := \int_{S^2} h^{(r,s)}(t, u) w(x-t, y-u) \, dt \, du.$$

By construction, $w^* \in \mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})$. Hence for $1 \leqq p \leqq \infty$

$$\begin{aligned} E^p_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{C}_{2\pi}; \mathscr{T}_{n-1}, \mathscr{C}_{2\pi})}(h^*) &\leqq \|h^* - w^*\|_p \\ &= \|h^{(r,s)} * (D_r D_s - w)\|_p \\ &= \|h^{(r,s)} * (D_r - u^*)(D_s - v^*)\|_p \\ &\leqq \|D_r - u^*\|_1 \, \|D_s - v^*\|_1 \, \|D^{r,s} h\|_p \\ &= \frac{K_r}{m^r} \cdot \frac{K_s}{n^s} \|D^{r,s} h\|_p, \end{aligned}$$

since we have (see Lorentz [18, Chap. 8])

$$E^1_{\mathscr{T}_{m-1}}(D_r) = \frac{4}{\pi} \cdot \frac{1}{m^r} \sum_{\rho=0}^{\infty} \left(\frac{(-1)^\rho}{2\rho+1} \right)^{r+1} = \frac{K_r}{m^r}. \quad \square$$

3. Remarks

(i) Achyeser-Krein [1] and Favard [5] proved the following univariate estimate for the class of functions W_r^* defined below (see also Timan [19, p. 292] for an estimate in the L_p norm for $p \geqq 1$). Let $r \in \mathbb{N}$ and

$$W_r^* := \{f: \mathbb{R} \to \mathbb{R} \mid f \in \mathscr{C}^{r-1}_{2\pi}(\mathbb{R}), D^{r-1} f \text{ absolutely continuous}, |D^r f| \leqq 1 \text{ a.e.}\}.$$

Then for any $f \in W_r^*$ we have

$$E^\infty_{\mathscr{T}_{m-1}}(f) \leqq \frac{K_r}{m^r}. \qquad (**)$$

This estimate is *optimal* for W_r^*, i.e. there is no constant $C_r < K_r$ such that $(**)$ holds true for all $f \in W_r^*$ with K_r replaced by C_r.

(ii) In addition, the constants K_r are *optimal for the class of functions*

$$\{g: \mathbb{R} \to \mathbb{R} \mid g \in \mathscr{C}^r_{2\pi}(\mathbb{R}), \|D^r g\|_\infty \leqq 1\},$$

too. Now consider the class of bivariate functions

$$Z^*_{r,s} := \{h: \mathbb{R}^2 \to \mathbb{R} \mid h \in \mathscr{C}^{r,s}_{2\pi}(\mathbb{R}^2), \|D^{r,s} h\|_\infty \leqq 1\}.$$

Then, by Theorem 1, we have

$$E^{\infty}_{\mathscr{B}(\mathscr{T}_{m-1},\mathscr{C}_{2\pi};\mathscr{T}_{n-1},\mathscr{C}_{2\pi})}(h)\leqq\frac{K_r}{m^r}\cdot\frac{K_s}{n^s}$$

for all $h\in Z^*_{r,s}$. Furthermore, this estimate is optimal for the class $Z^*_{r,s}$ (consider a bivariate function of product type, see Haußmann-Zeller [16]).

(iii) Denote by $\mathscr{T}:=\bigcup_{k=0}^{\infty}\mathscr{T}_k$ the vector space of all trigonometric polynomials. The next theorem shows that the estimate of Theorem 1 even holds true if we approximate with respect to the smaller space $\mathscr{B}(\mathscr{T}_{m-1},\mathscr{T};\mathscr{T}_{n-1},\mathscr{T})$ rather than $\mathscr{B}(\mathscr{T}_{m-1},\mathscr{C}_{2\pi};\mathscr{T}_{n-1},\mathscr{C}_{2\pi})$.

Theorem 3. *Let $h\in\mathscr{C}^{r,s}_{2\pi}(\mathbb{R}^2)$ for $r,s\in\mathbb{N}$. Then for any $m,n\in\mathbb{N}$ and any $1\leqq p\leqq\infty$*

$$E^{p}_{\mathscr{B}(\mathscr{T}_{m-1},\mathscr{T};\mathscr{T}_{n-1},\mathscr{T})}(h)\leqq\frac{K_r}{m^r}\cdot\frac{K_s}{n^s}\|D^{r,s}h\|_p.$$

This will be proved by showing that for any $h\in\mathscr{C}_{2\pi}(\mathbb{R}^2)$ we have

$$E^{p}_{\mathscr{B}(\mathscr{T}_{m-1},\mathscr{T};\mathscr{T}_{n-1},\mathscr{T})}(h)=E^{p}_{\mathscr{B}(\mathscr{T}_{m-1},\mathscr{C}_{2\pi};\mathscr{T}_{n-1},\mathscr{C}_{2\pi})}(h)$$

since $\mathscr{T}$ is dense in $\mathscr{C}_{2\pi}$, and the bivariate L_p norm is a cross-norm associated with the univariate L_p norm (see Gilbert-Leih [6]). Indeed, we have the following

Lemma 4. *Let X and Y be normed vector spaces, U and V finite dimensional subspaces, respectively. Let α be any cross-norm on $X\otimes Y$. Given $h\in X\otimes_\alpha Y$ (completion under α), then for the approximation constant we have*

$$E^{\alpha}_{\mathscr{B}(U,X;V,Y)}(h)=E^{\alpha}_{\mathscr{B}(U,X_1;V,Y_1)}(h),$$

whenever X_1 and Y_1 are dense subspaces of X and Y, respectively.

In fact, by the definition of the approximation constant, for every $k\in\mathbb{N}$ we can find an element

$$h_k=\sum_{i=1}^{M}u_i\otimes g_{i,k}+\sum_{j=1}^{N}f_{j,k}\otimes v_j$$

in $\mathscr{B}(U,X;V,Y)$ satisfying

$$E^{\alpha}_{\mathscr{B}(U,X;V,Y)}(h)\leqq\|h-h_k\|_\alpha<E^{\alpha}_{\mathscr{B}(U,X;V,Y)}(h)+\frac{1}{k}.$$

Here $\{u_1,\ldots,u_M\}$ and $\{v_1,\ldots,v_N\}$ are fixed bases of U and V, respectively, and $f_{j,k}\in X$, $g_{i,k}\in Y$. Since $\overline{X_1}=X$ and $\overline{Y_1}=Y$, there exists an element $H_k\in\mathscr{B}(U,X_1;V,Y_1)$ such that

$$\|h_k-H_k\|_\alpha<\frac{1}{k}.$$

Now the conclusion is immediate. □

(iv) For the case $p=\infty$, von Golitschek-Cheney [8] proved the following

Proposition 5. *Let E and F be compact topological spaces, and G and H finite dimensional subspaces of $C(E)$ and $C(F)$, respectively. If $h \in \mathscr{C}(E \times F)$, then*

$$E^{\infty}_{\mathscr{B}(G, C(E); H, C(F))}(h) = E^{\infty}_{\mathscr{B}(G, M(E); H, M(F))}(h),$$

where by $M(A)$ we denote the vector space of all bounded functions from a set A into $\mathbb{R}$.

This, together with Lemma 4 yields for every $h \in \mathscr{C}_{2\pi}(\mathbb{R}^2)$

$$E^{\infty}_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{T}; \mathscr{T}_{n-1}, \mathscr{T})}(h) = E^{\infty}_{\mathscr{B}(\mathscr{T}_{m-1}, M(S); \mathscr{T}_{n-1}, M(S))}(h).$$

The corresponding formula with $M(S)$ replaced by $L_{\infty}(S)$ follows with the aid of a result due to Light-Cheney [17, p. 390].

In the case $1 \leqq p < \infty$ we get

$$E^{p}_{\mathscr{B}(\mathscr{T}_{m-1}, \mathscr{T}; \mathscr{T}_{n-1}, \mathscr{T})}(h) = E^{p}_{\mathscr{B}(\mathscr{T}_{m-1}, L_p(S); \mathscr{T}_{n-1}, L_p(S))}(h)$$

for every $h \in L_p(S^2)$, an immediate consequence of Lemma 4.

References

1. Achyeser, N., Krein, M.: Sur la meilleure approximation des fonctions périodiques dérivables au moyen de sommes trigonométriques. Compt. Rend. (Doklady) Acad. Sci. URSS **15**, 107–111 (1937)
2. Cheney, E.W.: Best approximation in tensor product spaces. In: Numerical analysis (G.A. Watson, Ed.), 25–32. Lect. Notes Math. **773**. Berlin-Heidelberg-New York: Springer 1980
3. Cheney, E.W.: The best approximation of multivariate functions by combinations of univariate ones. In: Approximation Theory IV (C.K. Chui, L.L. Schumaker, J.D. Ward, Eds.), 1–26. New York-London: Academic Press 1983
4. Cheney, E.W., McCabe, J.H., Phillips, G.M.: A mixed-norm bivariate approximation problem with applications to Lewanowicz operators. In: Multivariate approximation (D.C. Handscomb, Ed.), 315–323. London-New York: Academic Press 1978
5. Favard, J.: Sur les meilleurs procédés d'approximation de certaines classes de fonctions par des polynomes trigonométriques. Bull. Sci. Math. **61**, 209–224 and 243–256 (1937)
6. Gilbert, J.E., Leih, T.J.: Factorization, tensor products, and bilinear forms in Banach space theory. In: Notes in Banach spaces (H.E. Lacey, Ed.), 185–305. Austin-London: University of Texas Press 1980
7. von Golitschek, M.: Approximation of functions of two variables by the sum of two functions of one variable. In: Numerical methods of approximation theory, Vol. 5 (L. Collatz, G. Meinardus, H. Werner, Eds.). Internat. Ser. Numer. Math. **52**, 117–124 (1980)
8. von Golitschek, M., Cheney, E.W.: The best approximation of bivariate functions by separable functions. Contemporary Mathematics, Vol. 21, 125–136. Providence: Amer. Math. Soc. 1983
9. Gordon, W.J.: Distributive lattices and the approximation of multivariate functions. In: Approximation with special emphasis on spline functions (I.J. Schoenberg, Ed.), 223–277. New York: Academic Press 1969
10. Gordon, W.J.: Blending-function methods of bivariate and multivariate interpolation and approximation. SIAM J. Numer. Anal. 8, 158–177 (1971)
11. Halton, E.J., Light, W.A.: Existence questions in tensor product spaces. In: Approximation Theory IV (C.K. Chui, L.L. Schumaker, J.D. Ward, Eds.), 499–504. New York-London: Academic Press 1983
12. Halton, E.J., Light, W.A.: Projections on tensor product spaces. University of Lancaster Math. Dept. Report, Lancaster, England, 1983

13. Hardy, G.H., Rogosinski, W.W.: Fourier series. Cambridge: University Press 1968
14. Haußmann, W., Zeller, K.: Blending interpolation and best L^1-approximation. Arch. Math. (Basel) **40**, 545–552 (1983)
15. Haußmann, W., Zeller, K.: Uniqueness and non-uniqueness in bivariate L^1-approximation. In: Approximation Theory IV (C.K. Chui, L.L. Schumaker, J.D. Ward, Eds.), 509–514. New York-London: Academic Press 1983
16. Haußmann, W., Zeller, K.: Mixed norm multivariate approximation with blending functions. In: Proceedings of the International Conference on Constructive Function Theory, Varna 1984 (Bl. Sendov et al., Eds.). To appear
17. Light, W.A., Cheney, E.W.: Some best-approximation theorems in tensor product spaces. Math. Proc. Camb. Philos. Soc. **89**, 385–390 (1981)
18. Lorentz, G.G.: Approximation of functions. New York: Holt, Rinehart and Winston 1966
19. Timan, A.F.: Theory of approximation of functions of a real variable. New York: MacMillan 1963

Received May 2, 1984

Math. Z. 189, 151–165 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

The Hasse-Witt Matrix of a Formal Group

Harry A.W.M. Kneppers

Wiskundig Seminarium der V.U., Postbus 7161, 1007 MC Amsterdam, The Netherlands

In [5], Ditters has introduced a natural lifting of the Hasse-Witt matrix, containing the structure constants (in the sense of Lazard-Dieudonné) of a formal group.

This new matrix can be regarded as a matrix of the semilinear *Frobenius* map. From this point of view, the relation with the classical Hasse-Witt matrix of an algebraic curve, as defined by Manin in [14], is obvious.

The classification up to isomorphism of smooth commutative formal groups is a matter of finding normal forms for the Hasse-Witt matrix of a formal group. Assuming the ground field to be algebraically closed, we shall derive some preliminary reductions for arbitrary dimension. Furthermore, we shall give intrinsic definitions for the discrete isomorphism invariants. The general results, together with the definitions of the discrete parameters, enable us to realize the covariant classification of two-dimensional smooth commutative formal groups in a way which is more transparent and needs fewer subcases than the cumbersome techniques of [11] Chap. 2.

1. Generalities

Let k be a perfect field of characteristic $p>0$. With each smooth commutative formal group G over k, there is the well-known topological group $C=C_p(G)$ of *p-typical curves* or *hyperexponential vectors* in G. For definitions, see [3] § III.3, [6] § 2, [9] § (27.7.1), [11] § 1.3 or [12] § VI.1.

(1.1) Let $D=D(k)$ be the *Dieudonné ring* over k, i.e. the elements of D are of the form

$$\sum_{i=1}^{<\infty} a_{-i}F^i + \sum_{i=0}^{\infty} V^i a_i,$$

with unique coefficients a_{-i}, a_i in the *Witt ring* $W=W(k)$ over k and D has multiplication rules $Fa=a^\sigma F$, $aV=Va^\sigma$ and $FV=VF=p\in W$, where $a\in W$ and a^σ is the image of a under the *Frobenius endomorphism* of W. (cf. [9] § (28.3.5).)

(1.2) The group C of p-typical curves in G has a natural structure of left D-module. Moreover, C is a *reduced* D-module, in other words, C is a left D-module with the properties:

(i) C has no V-torsion,

(ii) C is complete for its V-adic topology, or equivalently: $C = \varprojlim C/V^i C$, the $C/V^i C$ having the discrete topology.

Cf. [9] §(27.7) or [12] §VI.1–VI.3.

(1.3) The left D-module $C_p(G)$ is the *covariant Dieudonné module* of G. The importance of this module is expressed by the well-known

Cartier's theorem. *The (covariant) functor $G \to C_p(G)$ is an equivalence between the category of smooth commutative formal groups and the category of reduced D-modules.*

Cf. [4] §III.3.8, [9] §(27.7.14) or [12] §IV.7.12.

In the sequel, we shall always assume our formal groups to be *smooth* and *commutative.*

(1.4) In view of the fact that W/pW is isomorphic to k, the modules C/VC and C/pC have a natural structure of k-vector space. The k-vector space C/VC is called the *tangent space* of the corresponding formal group. Its rank is the *V-dimension* of C. The V-dimension of C is equal to the dimension of the corresponding formal group. Cf. [9] §(27.7.20) or [12] §VI.3.10. The rank of the k-vector space C/pC is the *height* of C or equivalently, the height of the corresponding formal group. Cf. [9] §(28.2.1).

(1.5) Let D^+ be the subring of D which consists of all elements of the (unique) form

$$\sum_{i=0}^{\infty} V^i a_i.$$

A left D^+-module C is a reduced D-module if and only if it satisfies the properties (i) and (ii) in (1.2) and moreover $pC \subset VC$. Cf. [9] §(28.3.6) or [12] §VI.3.2. Note that Dieudonné and Honda regard C as a right $(D^+)^\circ$-module, where $(D^+)^\circ = W_\sigma[[T]]$ is the opposite ring of D^+. Cf. [3] §III.2 and [10]. In view of this observation, a *distinguished* $(D^+)^\circ$-module in terms of [3] §III.4 (nrs. (13) and (14)) is the same as a reduced D-module of finite V-dimension.

(1.6) As usual, we shall write $[a] \in W$ for the image of an element $a \in k$ under the *Teichmüller* map: $k \to W$. In other words: The action of $[a]$ on a D-module is the *homothety* defined by a.

Let C be a reduced D-module of V-dimension $n < \infty$. Then C has a *V-basis*, i.e. an ordered set of elements $(\gamma_1, \ldots, \gamma_n)$ such that each element γ of C can *uniquely* be expressed as

$$\gamma = \sum_{s=0}^{\infty} V^s \sum_{i=1}^{n} [c_i^{(s)}] \gamma_i.$$

So a V-basis generates C over D^+. In fact, any ordered set of cardinality n which generates C over D^+ is a V-basis.

(1.7) Let M and N be finitely generated left modules over a ring R, with (ordered) sets of generators $(m_1, \ldots, m_s)$ resp. $(n_1, \ldots, n_t)$. Furthermore, let f: $M \to N$ be a semilinear map in the sense of Bourbaki [1] § 1.13.

Definition. An element $A = (a_{ij})$ of $M_{t \times s}(R)$ is called a *matrix* of f relative to $(m_1, \ldots, m_s)$ and $(n_1, \ldots, n_t)$ if

$$f(m_j) = \sum_{i=1}^{t} a_{ij} n_i \qquad (j = 1, \ldots, s).$$

In other words:

$$\begin{pmatrix} f(m_1) \\ \vdots \\ f(m_s) \end{pmatrix} = {}^t A \begin{pmatrix} n_1 \\ \vdots \\ n_t \end{pmatrix}.$$

This extends the definition given for free modules by Bourbaki in [1] § 10.6.

(1.8) Let B and C be reduced D-modules of finite V-dimension with V-bases $(\beta_1, \ldots, \beta_m)$ resp. $(\gamma_1, \ldots, \gamma_n)$.

Definition. The *Teichmüller matrix* of a semilinear map f: $B \to C$ relative to $(\beta_1, \ldots, \beta_m)$ and $(\gamma_1, \ldots, \gamma_n)$ is the unique matrix $A \in M_{n \times m}(D^+)$ of f relative to $(\beta_1, \ldots, \beta_m)$ and $(\gamma_1, \ldots, \gamma_n)$ such that the entries of A are of the form

$$a_{ij} = \sum_{s=0}^{\infty} V^s [a_{ij}^{(s)}].$$

The existence and the unicity of such a matrix are immediate consequences of the properties of a V-basis (cf. (1.6)).

(1.9) Extend the Frobenius endomorphism σ of W to the whole of D in the usual way ($F^\sigma = F$ and $V^\sigma = V$). Then $F a \gamma = a^\sigma F \gamma$ ($a \in D$, $\gamma \in C$), so that the action of F on C is a σ-semilinear map F: $C \to C$.

Definition. The *Hasse-Witt matrix* (shortly: HW matrix) of a reduced D-module C relative to a V-basis $(\gamma_1, \ldots, \gamma_n)$ is the Teichmüller matrix of the σ-semilinear map F: $C \to C$ relative to $(\gamma_1, \ldots, \gamma_n)$.

So the coefficients of the V^i in the Hasse-Witt matrix are the *higher Hasse-Witt matrices* of C relative to $(\gamma_1, \ldots, \gamma_n)$ (cf. [5]). The entries of the higher Hasse-Witt matrices are the *structure constants* of C relative to $(\gamma_1, \ldots, \gamma_n)$ (cf. [9] § (27.4.1) and [12] § V.5.17).

If G is the completion at the origin of the Jacobian of an algebraic curve Γ then the reduction modulo (p, V) of the Hasse-Witt matrix of $C_p(G)$ relative to a given V-basis, or equivalently, the matrix of the (p-linear) Frobenius on the tangent space of G (relative to the reduction mod VC of our given V-basis), is the *Hasse-Witt matrix* of the curve Γ (cf. [5] and [14]).

It is well-known that C is determined by any of its HW matrices. The explicit description of C in terms of a (not necessarily Teichmüller) matrix of F looks as follows:

(1.10) **Proposition.** *Let C be a reduced D-module of V-dimension n. Furthermore, let L be a matrix of the σ-semilinear map F: $C \to C$ relative to a V-basis*

$(\gamma_1, \ldots, \gamma_n)$. *A combination* $\sum_{i=1}^{n} x_i\gamma_i$ *(all* x_i *in* D^+*) is zero if and only if*

$$(x_1, \ldots, x_n) \in (D^+)^n(pI_n - V.^tL),$$

where I_n *is the* $n \times n$ *identity matrix.*

Proof. If $(x_1, \ldots, x_n) = (y_1, \ldots, y_n)(pI_n - V.^tL)$ then

$$\sum_{i=1}^{n} x_i\gamma_i = (x_1, \ldots, x_n)\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix} = (y_1, \ldots, y_n)(pI_n - V.^tL)\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}$$

and since L is a matrix of F relative to $(\gamma_1, \ldots, \gamma_n)$ we have

$$\begin{pmatrix}p\gamma_1\\ \vdots\\ p\gamma_n\end{pmatrix} = VI_n\begin{pmatrix}F\gamma_1\\ \vdots\\ F\gamma_n\end{pmatrix} = V.^tL\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}.$$

Conversely, assume $\sum_{i=1}^{n} x_i\gamma_i = 0$. It is easily verified that $(x_1, \ldots, x_n)$ can be written as

$$(x_1, \ldots, x_n) = (y_1, \ldots, y_n) + (z_1, \ldots, z_n)(pI_n - V.^tL)$$

with y_i having the form $y_i = \sum_{s=0}^{\infty} V^s[y_i^{(s)}]$ and $z_i \in D^+$. Now $\sum_{i=1}^{n} y_i\gamma_i = \sum_{i=1}^{n} x_i\gamma_i = 0$, so that all $y_i^{(s)}$ are zero (cf. (1.6)), hence $(y_1, \ldots, y_n)$ is zero. □

(1.11) **Corollary** (cf. [3] § III.4 Prop. 7 and [10] § 4). *Let* C *be a reduced* D*-module of* V*-dimension* n*. Furthermore, let* L *be a matrix of* F *relative to a* V*-basis* $(\gamma_1, \ldots, \gamma_n)$*. Then*

$$C = (D^+)^n/(D^+)^n(pI_n - V.^tL),$$

or, which amounts to the same,

$$C = D^n/D^n(F - {}^tL).$$

(1.12) **Corollary.** *Let* B *and* C *be reduced* D*-modules with* V*-bases* $(\beta_1, \ldots, \beta_m)$ *resp.* $(\gamma_1, \ldots, \gamma_n)$*. Furthermore, let* P *and* Q *be matrices of* τ*-semilinear maps:* $B \to C$ *relative to* $(\beta_1, \ldots, \beta_m)$ *and* $(\gamma_1, \ldots, \gamma_n)$*, where* τ *is a ring endomorphism of* D*. Let* L *be a matrix of* F *relative to* $(\gamma_1, \ldots, \gamma_n)$*. Then* P *and* Q *represent the same* τ*-semilinear map if and only if*

$${}^tP \equiv {}^tQ \bmod M_{m\times n}(D^+)(pI_n - V.^tL).$$

Proof. P and Q represent the same τ-semilinear map if and only if (cf. (1.7)) $({}^tP - {}^tQ)\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix} = \begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$ if and only if (cf. (1.10)) each row of ${}^tP - {}^tQ$ is in $(D^+)^n(pI_n - V.^tL)$ if and only if ${}^tP - {}^tQ \in M_{m\times n}(D^+)(pI_n - V.^tL)$. □

(1.13) **Corollary.** *Let B and C be reduced D-modules and assume K is a matrix of $F\colon B\to B$ relative to a V-basis $(\beta_1,\dots,\beta_m)$ of B and L is a matrix of $F\colon C\to C$ relative to a V-basis $(\gamma_1,\dots,\gamma_n)$ of C. Then a matrix P in $M_{n\times m}(D^+)$ represents a morphism $f\colon B\to C$ of D-modules relative to $(\beta_1,\dots,\beta_m)$ and $(\gamma_1,\dots,\gamma_n)$ if and only if*

$$ {}^tP^{\sigma}.{}^tL\equiv{}^tK.{}^tP \bmod M_{m\times n}(D^+)(pI_n-V.{}^tL).$$

Proof. First, assume P represents a morphism $f\colon B\to C$ relative to $(\beta_1,\dots,\beta_m)$ and $(\gamma_1,\dots,\gamma_n)$. Then $f(F\beta)=Ff(\beta)$ for all $\beta\in B$, so that the σ-semilinear maps fF and Ff are the same. Now ${}^t({}^tP^{\sigma}.{}^tL)$ represents Ff and ${}^t({}^tK.{}^tP)$ represents fF relative to $(\beta_1,\dots,\beta_m)$ and $(\gamma_1,\dots,\gamma_n)$ (cf. [1] §10.6). Corollary (1.12) gives the desired relation. Conversely, assume ${}^tK.{}^tP={}^tP^{\sigma}.{}^tL+R(pI_n-V.{}^tL)$ for some $R\in M_{m\times n}(D^+)$. We must show that $f\left(\sum_{i=1}^{m}x_i\beta_i\right)=(x_1,\dots,x_m){}^tP\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}$ is a well-defined morphism $f\colon B\to C$. Assume $\sum_{i=1}^{m}x_i\beta_i=0$. In view of (1.10) we may write

$$(x_1,\dots,x_m)=(y_1,\dots,y_m)(pI_m-V.{}^tK).$$

Now

$$(x_1,\dots,x_m){}^tP=(y_1,\dots,y_m)({}^tP+VR)(pI_n-V.{}^tL)\in(D^+)^n(pI_n-V.{}^tL),$$

so that $(x_1,\dots,x_m){}^tP\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}=0$. □

2. General Results

(2.1) **Lemma.** *Let C be a reduced D-module of V-dimension $n<\infty$ and let L be a matrix of F relative to a V-basis $(\gamma_1,\dots,\gamma_n)$. Furthermore, let Λ be an invertible element of $M_{n\times n}(D^+)$. Then the set $(\delta_1,\dots,\delta_n)$, defined by $\begin{pmatrix}\delta_1\\ \vdots\\ \delta_n\end{pmatrix}={}^t\Lambda\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}$, is another V-basis of C and ${}^t({}^t\Lambda^{\sigma}.{}^tL.({}^t\Lambda)^{-1})$ is a matrix of F relative to $(\delta_1,\dots,\delta_n)$.*

Proof. Obviously, the set $(\delta_1,\dots,\delta_n)$ generates C over D^+, so that it is a V-basis (cf. (1.6)).

Furthermore,

$$\begin{pmatrix}F\delta_1\\ \vdots\\ F\delta_n\end{pmatrix}={}^t\Lambda^{\sigma}\begin{pmatrix}F\gamma_1\\ \vdots\\ F\gamma_n\end{pmatrix}={}^t\Lambda^{\sigma}.{}^tL\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}={}^t\Lambda^{\sigma}.{}^tL.({}^t\Lambda)^{-1}\begin{pmatrix}\delta_1\\ \vdots\\ \delta_n\end{pmatrix}.$$ □

Warnings. (i) In general, even if L is the HW matrix relative to $(\gamma_1,\dots,\gamma_n)$, the matrix ${}^t({}^t\Lambda^{\sigma}.{}^tL.({}^t\Lambda)^{-1})$ is *not* the HW matrix relative to $(\delta_1,\dots,\delta_n)$, since the coefficients of its entries do not belong to the image of the Teichmüller map.

(ii) In spite of [1] §10.7, $({}^t\Lambda)^{-1}$ is not always the same as ${}^t(\Lambda^{-1})$ (a counterexample is easily found). Nevertheless, Λ is invertible if and only if ${}^t\Lambda$ is invertible (if and only if the reduction of Λ modulo (p, V) is invertible in $M_{n\times n}(k)$).

(2.2) **Definition.** Let C be a reduced D-module. The *order* $\mathrm{ord}(\gamma)$ of an element γ of C is the order of γ in the V-filtration of C, i.e.

$$\mathrm{ord}(\gamma)=\sup\{i\,|\,\gamma\in V^iC\}.$$

For any natural number s, we define

$$\mathrm{ord}(F^sC)=\sup\{i\,|\,F^sC\subset V^iC\}.$$

Obviously, the numbers $\mathrm{ord}(F^sC)$ are isomorphism invariants of C.

The following properties are easily verified:

(i) $\mathrm{ord}(V^s\gamma)=s+\mathrm{ord}(\gamma)$.

(ii) $\mathrm{ord}(F^s\gamma)\geqq\mathrm{ord}(F^sC)+\mathrm{ord}(\gamma)$.

(iii) If $a\in W$ then $\mathrm{ord}(a\gamma)\geqq\mathrm{ord}(\gamma)$.

(iv) If $(\gamma_1,\ldots,\gamma_n)$ is a V-basis of C then $\mathrm{ord}(\gamma_i)=0$ for $i=1,\ldots,n$, since the classes modulo V of the γ_i constitute a basis of the n-dimensional k-vector space C/VC.

(v) If $(\gamma_1,\ldots,\gamma_n)$ is a V-basis of C then

$$\mathrm{ord}(F^sC)=\min\{\mathrm{ord}(F^s\gamma_i)\,|\,i=1,\ldots,n\}.$$

(vi) $\mathrm{ord}(F^{s_1+s_2}C)\geqq\mathrm{ord}(F^{s_1}C)+\mathrm{ord}(F^{s_2}C)$.

(2.3) *Notation.* Let C be a reduced D-module. Then $h_1=h_1(C)=\mathrm{ord}(FC)$ (cf. (2.2)). For a formal group G we shall write $h_1(G)=h_1(C_p(G))$. Obviously, h_1 is an isomorphism invariant of C resp. of G.

In view of (2.2) property (vi), we have $\mathrm{ord}(F^sC)\geqq sh_1$.

Note that $h_1=\infty$ if and only if $F\gamma=0$ for all γ in C, in other words C is the covariant Dieudonné module of the n-dimensional additive formal group (where n is the V-dimension of C).

(2.4) Assume C has finite V-dimension and $h_1<\infty$. Let M be the HW matrix of C relative to a V-basis $(\gamma_1,\ldots,\gamma_n)$. Since FC is contained in $V^{h_1}C$ but not in $V^{h_1+1}C$, the matrix M is zero modulo V^{h_1} but nonzero mod V^{h_1+1}.

Theorem (Ditters, cf. [6] §3). *Assume k is algebraically closed. Let C be a reduced D-module of V-dimension $n<\infty$ and assume $h_1=h_1(C)<\infty$. Then there exists a unique matrix $N_{h_1}=N_{h_1}(C)$ in $M_{n\times n}(W)$ which has a (nonzero) Hasse-Witt normal form (cf.* [8] *p.* 489) *such that there exists a V-basis $(\delta_1,\ldots,\delta_n)$ with HW matrix $M\equiv V^{h_1}N_{h_1}$ mod V^{h_1+1}.*

Furthermore, the matrix N_{h_1} is an isomorphism invariant of C.

Proof. Let L be the HW matrix of C relative to a V-basis $(\gamma_1,\ldots,\gamma_n)$. Furthermore, let $A\in M_{n\times n}(k)$ be the reduction of $V^{-h_1}L$ modulo (p,V). Then there exists an invertible matrix S in $M_{n\times n}(k)$ such that $N_{h_1}:=S^{-1}AS^{(p^{h_1+1})}$ has a

Hasse-Witt normal form. (Cf. [8] Sect. 3 Satz 11 for the case $h_1=0$, the general case follows from a straightforward generalization of this theorem). Let $\Lambda\in M_{n\times n}(D^+)$ be a lifting of S (hence Λ is invertible). Furthermore, let M be the HW matrix relative to the V-basis $\begin{pmatrix}\delta_1\\ \vdots\\ \delta_n\end{pmatrix}={}^t\Lambda\begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix}$. We claim that $M\equiv V^{h_1}N_{h_1} \bmod V^{h_1+1}$. Indeed, in view of (2.1) and (1.12) we have: ${}^tM\equiv{}^t(\Lambda^\sigma.{}^tL.({}^t\Lambda)^{-1}) \bmod M_{n\times n}(D^+)(pI_n-V.{}^tM)$, so that

$$V^{-h_1}M\equiv S^{-1}AS^{(p^{h_1+1})}\equiv N_{h_1} \bmod (p,V).$$

Next, we shall show that N_{h_1} is an isomorphism invariant (hence unique). Let B be another reduced D-module and assume we have an isomorphism Φ: $B\to C$. Let M' and M be HW matrices of B resp. C relative to V-bases $(\beta_1,\ldots,\beta_n)$ resp. $(\gamma_1,\ldots,\gamma_n)$ and assume

$$M\equiv V^{h_1}N_{h_1} \quad\text{and}\quad M'\equiv V^{h_1}N'_{h_1} \bmod V^{h_1+1},$$

where both N_{h_1} and N'_{h_1} have a Hasse-Witt normal form. Furthermore, let P be a matrix of Φ relative to $(\beta_1,\ldots,\beta_n)$ and $(\gamma_1,\ldots,\gamma_n)$. Then P is invertible and

$${}^tP^\sigma.{}^tM\equiv{}^tM'.{}^tP \bmod M_{n\times n}(D^+)(pI_n-V.{}^tM)$$

(see (1.13)). Writing $S\in M_{n\times n}(k)$ for the reduction of P modulo (p,V) (hence S is invertible), we have $N_{h_1}S^{(p^{h_1+1})}=SN'_{h_1}$. In view of the uniqueness of the Hasse-Witt normal form we have $N'_{h_1}=N_{h_1}$. □

For a formal group G we define $N_{h_1}(G)=N_{h_1}(C_p(G))$.

(2.5) Note that the matrix N_{h_1} is not nilpotent if and only if its $(1,1)$-entry is 1 (cf. [8], loc. cit.). On the other hand, N_{h_1} is not nilpotent if and only if $\operatorname{ord}(F^sC)=sh_1$ for all s.

Theorem. *Assume k is algebraically closed. Let C be a reduced D-module of finite V-dimension such that h_1 is finite and N_{h_1} is not nilpotent. Then C is isomorphic to the direct sum of $D/D(F-V^{h_1})$ and a reduced D-module C'.*

Proof. We must show that F has a matrix of the form $V^{h_1}\oplus M$ (direct sum of matrices) relative to some V-basis $(\gamma_1,\ldots,\gamma_n)$. Let $s>0$ and assume $L^{(s)}$ is a matrix of F relative to a V-basis $(\gamma_1^{(s)},\ldots,\gamma_n^{(s)})$ with the properties:

$$L^{(s)}\equiv V^{h_1}N_{h_1} \bmod V^{h_1+1} \quad\text{and}\quad L^{(s)}\equiv V^{h_1}\oplus M^{(s)} \bmod V^{h_1+s}.$$

For $s=1$, the existence of such a matrix and V-basis follows from (2.4) plus the assumption that N_{h_1} is not nilpotent. Put $\Lambda=1+V^sT$, where $T=(\tau_{ij})\in M_{n\times n}(W)$. Whatever value T has, Λ is invertible and $L^{(s+1)}:={}^t(\Lambda^\sigma.{}^tL^{(s)}.({}^t\Lambda)^{-1})$ is a matrix of F relative to $\begin{pmatrix}\gamma_1^{(s+1)}\\ \vdots\\ \gamma_n^{(s+1)}\end{pmatrix}:={}^t\Lambda\begin{pmatrix}\gamma_1^{(s)}\\ \vdots\\ \gamma_n^{(s)}\end{pmatrix}$. Note that $\gamma_i^{(s+1)}\equiv\gamma_i^{(s)} \bmod V^sC$, so that the sequences $(\gamma_i^{(s)})_{s\in\mathbb{N}}$ converge V-adically in C. Once we have shown that T can be

chosen such that $L^{(s+1)} \equiv V^{h_1} \oplus M^{(s+1)} \bmod V^{h_1+s+1}$, the proof of the theorem is done by taking limits.

Since $s>0$, we have $({}^t\Lambda)^{-1} = \sum_{i=0}^{\infty} (-V^s \cdot {}^tT)^i$. So $L^{(s+1)} \equiv L^{(s)} \bmod V^{h_1+s}$. Moreover, writing U and $Q \in M_{n\times n}(W)$ for the coefficients of V^{h_1+s} in $L^{(s)}$ resp. $L^{(s+1)}$, we have

$$Q = N_{h_1} T^{\sigma^{h_1+1}} - TN_{h_1} + U.$$

It is an elementary exercise to show that T can be chosen such that the first row and the first column of Q are zero. Essentially, this is based on [3] § III.5 Lemma 1. For details, cf. [11] Lemma 1.10.4. □

Note that, with the notations of the theorem, we have $h_1(C') \geqq h_1(C)$. Moreover, equality holds if and only if $N_{h_1}(C)$ is *not* the direct sum of I_1 and the $(n-1)\times(n-1)$-zero matrix. If this is the case, we have $N_{h_1}(C) = I_1 \oplus N_{h_1}(C')$. So if $N_{h_1} = I_n$, we find that C is isomorphic to $(D/D(F-V^{h_1}))^n$. Consequently, if $h_1 = 0$ and $N_{h_1} = I_n$ then C is isomorphic to the covariant Dieudonné module of the n-dimensional multiplicative formal group.

More generally, each reduced D-module of finite V-dimension splits into a number of D-modules having V-dimension 1 and a D-module C with nilpotent N_{h_1}.

(2.6) Theorem (2.5) gives at once the following well-known result:

Corollary. *Two one-dimensional formal groups G and H are isomorphic if and only if $h_1(G) = h_1(H)$.*

In the one-dimensional case we have the isomorphism $C = D/D(F - V^{h_1})$. It follows that the *height* of G is $h_1 + 1$.

So this corollary expresses the fact that the one-dimensional formal groups over an algebraically closed field of positive characteristic are classified by their height.

(2.7) Next we shall consider the case of a nilpotent N_{h_1}.

Theorem. *Assume k is algebraically closed. Let C be a reduced D-module of V-dimension $n<\infty$ such that $h_1<\infty$ and N_{h_1} nilpotent. Let m be the number of nonzero columns in N_{h_1}. Then there exists a V-basis $(\gamma_1, \ldots, \gamma_n)$ with HW matrix M such that:*

(i) $M \equiv V^{h_1} N_{h_1} \bmod V^{h_1+1}$ *and*

(ii) *The first m columns of M are equal to the first m columns of $V^{h_1} N_{h_1}$.*

Proof. Let L be a matrix of F relative to a V-basis $(\gamma_1, \ldots, \gamma_n)$ such that $L \equiv V^{h_1} N_{h_1} \bmod V^{h_1+1}$ (cf. (2.4)). Then for $i = 1, \ldots, m$ we have: $F\gamma_i \equiv V^{h_1}\gamma_{k_i} \bmod V^{h_1+1} C$ for some $k_i > i$ and for $i > m$ we have $f\gamma_i \equiv 0 \bmod V^{h_1+1} C$. Put $\delta_j = \gamma_j$ if $j \notin \{k_i | i = 1, \ldots, m\}$ and recursively: $\delta_{k_i} = V^{-h_1} F\delta_i$ for $i = 1, \ldots, m$. Then $\delta_j \equiv \gamma_j \bmod VC$ $(j = 1, \ldots, n)$, so that $(\delta_1, \ldots, \delta_n)$ is a V-basis of C. It is obvious that $F\delta_i = V^{h_1}\delta_{k_i}$ $(i = 1, \ldots, m)$ and that $F\delta_i \equiv 0 \bmod V^{h_1+1}$ if $i > m$, so that $(\delta_1, \ldots, \delta_n)$ is the required V-basis. □

3. Normal Forms in Dimension 2

Throughout this section, the field k is assumed to be algebraically closed. We shall reconsider the classification of reduced D-modules C having V-dimension 2. In contrast with [11] Chap. 2, we can start with the discrete parameters h, the height of C, and $\mathrm{ord}(F^s C)$, $s \geqq 0$. Moreover, Theorem (2.7) saves a lot of work.

(3.1) *Notations.* $h_1 = h_1(C) = \mathrm{ord}(FC)$. If $h_1 < \infty$: let h_2 be such that $2h_1 + h_2 = \mathrm{ord}(F^2 C)$, so that $0 \leqq h_2 \leqq \infty$ (see (2.3)). Let h be the height of C (cf. (1.4)). We shall see that $h \geqq 2h_1 + h_2 + 2$. Note that if $h_1 = \infty$ or $h_2 = \infty$ we have $p^2 C = V^2 F^2 C = 0$, so that C has W-torsion, hence C has infinite height (cf. [9] §(28.3.10)). For finite h_1 and h_2, put $h_3 = h - (2h_1 + h_2 + 2)$.

For a two-dimensional formal group G define $h_i(G) = h_i(C_p(G))$ $(i = 1, 2, 3)$.

Obviously, h_1, h_2 and h_3 are isomorphism invariants of C resp. of G. If $h_1 = \infty$, there is nothing more to be said about C (see (2.3)). If $h_1 < \infty$ then N_{h_1} is either $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

(3.2) Assume $h_1 < \infty$. First we shall consider the case that N_{h_1} is not nilpotent. In (2.5) we have seen that $h_2 = 0$ if and only if N_{h_1} is not nilpotent and then C is isomorphic to $D/D(F - V^{h_1}) \oplus C'$ for some C' with $h_1' := h_1(C') \geqq h_1$. Since C' has V-dimension 1, we have: $C' = D/D(F - V^{h_1'})$. It follows that in this case the height h of C is $h_1 + h_1' + 2$ (cf. (2.6)), so that $h_3 = h_1' - h_1$ $(\geqq 0!)$. Summarizing, we have

Proposition. *If $h_1 < \infty$ and $h_2 = 0$ then C has a V-basis with HW matrix*

$$N = V^{h_1} \begin{pmatrix} 1 & 0 \\ 0 & V^{h_3} \end{pmatrix}.$$

(3.3) Assume $h_1 < \infty$ and $h_2 > 0$, so that $N_{h_1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. We are now in the position to apply (2.7).

Lemma. *Let (γ_1, γ_2) be a V-basis with HW matrix $V^{h_1} \begin{pmatrix} 0 & x' \\ 1 & y' \end{pmatrix}$ (cf. (2.7)). Then $F\gamma_1$ has order h_1 and $F\gamma_2$ has order $h_1 + h_2$.*

Proof. As we have seen in (2.2), we have $\mathrm{ord}(\gamma_1) = \mathrm{ord}(\gamma_2) = 0$. Note that $F\gamma_1 = V^{h_1}\gamma_2$. Using (2.2), we find:

$$\mathrm{ord}(F\gamma_1) = \mathrm{ord}(V^{h_1}\gamma_2) = h_1$$

and

$$\mathrm{ord}(F^2\gamma_1) = \mathrm{ord}(V^{h_1} F\gamma_2) = h_1 + \mathrm{ord}(F\gamma_2).$$

Furthermore, $\mathrm{ord}(F^2\gamma_2) \geqq h_1 + \mathrm{ord}(F\gamma_2)$. Consequently, we have

$$2h_1 + h_2 = \mathrm{ord}(F^2 C) = \min\{\mathrm{ord}(F^2\gamma_1), \mathrm{ord}(F^2\gamma_2)\} = h_1 + \mathrm{ord}(F\gamma_2),$$

hence $\mathrm{ord}(F\gamma_2) = h_1 + h_2$. □

So the HW matrix of C relative to (γ_1, γ_2) has the form:

$$M = V^{h_1}\begin{pmatrix} 0 & V^{h_2}x \\ 1 & V^{h_2}y \end{pmatrix},$$

where x and y are elements of D^+ whose coefficients are in the image of the Teichmüller map: $k \to W$. Moreover, if $h_2 < \infty$ then x and y are *not both zero* mod V.

Note that for $h_2 = \infty$ we have

$$N := M = V^{h_1}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

So if G is a two-dimensional formal group with $h_2(G) = \infty$ then G is a formal group of "streak" type (cf. [2] § 10).

(3.4) Assume $h_1 < \infty$ and $0 < h_2 < \infty$.

Lemma. *Let* (γ_1, γ_2) *be a* V*-basis with HW matrix* $V^{h_1}\begin{pmatrix} 0 & V^{h_2}x \\ 1 & V^{h_2}y \end{pmatrix}$ (*cf.* (3.3)). *Put* $v = \mathrm{ord}(x\gamma_1)$. *Then* pC *is generated over* D^+ *by* $\{V^{h_1+h_2+v+1}\gamma_1, V^{h_1+1}\gamma_2\}$.

Proof. In view of the unicity of x and the fact that $\mathrm{ord}(\gamma_1) = 0$, we have $x = V^v x'$ for some $x' \in D^+$. Now $p\gamma_1 = VF\gamma_1 = V^{h_1+1}\gamma_2$ and $p\gamma_2 = VF\gamma_2 = V^{h_1+h_2+v+1}x'\gamma_1 + V^{h_1+h_2+1}y\gamma_2$, so that pC is contained in the D^+-submodule of C generated by $\{V^{h_1+h_2+v+1}\gamma_1, V^{h_1+1}\gamma_2\}$.

Conversely, $V^{h_1+1}\gamma_2 = p\gamma_1 \in pC$; if $v = \infty$ (in other words $x = 0$) we are done and if $v < \infty$ (hence x' invertible in D^+) then $V^{h_1+h_2+v+1}x'\gamma_1 = p\gamma_2 - V^{h_1+h_2+1}y\gamma_2 \in pC$, hence we also have $V^{h_1+h_2+v+1}\gamma_1 \in pC$. □

(3.5) **Corollary.** *Under the assumptions of* (3.4), *the height of* C *is* $2h_1 + h_2 + v + 2$, *so that* $h_3 = v$ ($\geqq 0$!).

Proof. If $v = \infty$ then $\{V^i\gamma_1 \mid i \geqq 0\}$ is k-linearly independent in C/pC, hence the height of C is infinite.

If $v < \infty$ it is a straightforward verification that

$$\{V^i\gamma_1 \mid 0 \leqq i \leqq h_1 + h_2 + v\} \cup \{V^i\gamma_2 \mid 0 \leqq i \leqq h_1\}$$

constitutes a k-basis of C/pC. □

(3.6) **Proposition.** *Assume* $h_1 < \infty$, $0 < h_2 < \infty$ *and* $h_3 = \infty$. *Then there exists a* V*-basis with HW matrix*

$$N = V^{h_1}\begin{pmatrix} 0 & 0 \\ 1 & V^{h_2} \end{pmatrix}.$$

Proof. Let (γ_1, γ_2) be a V-basis with HW matrix $M = V^{h_1}\begin{pmatrix} 0 & 0 \\ 1 & V^{h_2}y \end{pmatrix}$ (cf. (3.4) and (3.5)). Note that y is nonzero mod(p, V) (cf. (3.3)). Put $\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{\sigma^{h_1+1}} \end{pmatrix}$, with $\lambda \in D^+$. We are done if we can show that λ can be chosen such that Λ is invertible and ${}^t\Lambda^{\sigma} \cdot {}^tM = {}^tN \cdot {}^t\Lambda$ (cf. (2.1)). In other words: λ must be nonzero

$\mathrm{mod}(p, V)$ and it must satisfy

$$\lambda^{\sigma^{2h_1+h_2+2}} y = \lambda^{\sigma^{h_1+1}}.$$

Now [3] § III.5 Lemma 1 shows the existence of a solution $\mathrm{mod}\, V$ and once we have a solution $\mathrm{mod}\, V^i$ for some $i>0$, the existence of a solution $\mathrm{mod}\, V^{i+1}$ follows again from [3], loc. cit. Taking limits (in the V-adic topology) we find the desired λ. □

So if G is a two-dimensional formal group with $h_3(G)=\infty$ then G is a formal group of "tree" type (cf. [2] § 10).

(3.7) So far, we have treated the case $h=\infty$ and the case $h_2=0$. Next, assume that $h=2h_1+h_2+h_3+2$ is finite and that $h_2>0$.

Lemma. *There exists a V-basis (δ_1, δ_2) such that F has a matrix of the form*

$$V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2} z \end{pmatrix}$$

relative to (δ_1, δ_2), for some $z \in D^+$.

Proof. Let (γ_1, γ_2) be a V-basis with HW matrix $V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3}x \\ 1 & V^{h_2}y \end{pmatrix}$ (cf. (3.4) and (3.5)). Recall that x is nonzero $\mathrm{mod}(p, V)$. We proceed in the same way as in (3.6), the only difference is that we want λ to be such that $\lambda^{\sigma^h} x = \lambda$. □

(3.8) **Proposition.** *Assume $h<\infty$ and $h_2>0$. Let $L=V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2}z \end{pmatrix}$ be a matrix of F relative to a V-basis (γ_1, γ_2) of C (cf. (3.7)). Put $d=\min\{h_2, h_3\}$ and let $c \in D^+$ be the unique element of the form $c=\sum_{i=0}^{d-1} V^i[c_i]$ such that $c \equiv z \bmod (p, V^d)$ (if $d=0$ then $c=0$). Then C has a V-basis with HW matrix*

$$N = V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2}c \end{pmatrix}.$$

Proof. First of all, we shall show that z may be assumed to be such that $z \equiv c \bmod V^d$. Let $u \in D^+$ be such that $z \equiv c + up \bmod V^d$. In view of (1.12), the matrix L' defined by ${}^tL' = {}^tL - V^{h_1+h_2}\begin{pmatrix} 0 & 0 \\ 0 & u \end{pmatrix}(pI_2 - V.{}^tL)$ also is a matrix of F relative to (γ_1, γ_2). An easy calculation shows that L' is of the form $L' = V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3}x \\ 1 & V^{h_2}y \end{pmatrix}$, with x and y in D^+ such that $x \equiv 1 \bmod V^d$ and $y \equiv c \bmod V^d$. Next, consider the same transformation as in (3.7), but now λ must be such that $\lambda \equiv 1 \bmod V^d$. Since $x \equiv 1 \bmod V^d$, such a choice of λ is possible. This transformation leads to a matrix of the form $V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2}z' \end{pmatrix}$, with $z' \equiv c \bmod V^d$.

So from now on, let us assume that $z \equiv c \bmod V^d$. The proof will be completed by transformation with $\Lambda = I_2 + \begin{pmatrix} V^{h_2}\lambda & V^{h_2+h_3}\mu \\ \xi & V^{h_2}\eta \end{pmatrix}$, where λ, μ, ξ and η

must be chosen such that

$$ {}^t\Lambda^{\sigma}.{}^tL = {}^tN.{}^t\Lambda. $$

Note that $\Lambda - I_2$ is topologically nilpotent, so that Λ is invertible. Let $y \in D^+$ be such that $z = c + V^d y$. A straightforward computation shows that the $(1,1)$-entry of the equation ${}^t\Lambda^{\sigma}.{}^tL = {}^tN.{}^t\Lambda$ gives:

$$ \mu = \xi^{\sigma^{h_1+h_2+h_3+1}}, $$

the $(2,1)$-entry gives

$$ \lambda = \eta^{\sigma^{h_1+h_2+h_3+1}} - c^{\sigma^{h_2+h_3}}\mu, $$

the $(1,2)$-entry gives

$$ \lambda^{\sigma^{h_1+1}} + \xi^{\sigma^{h_1+h_2+1}} z - \eta = 0 $$

and the $(2,2)$-entry gives

$$ V^d y + V^{h_3}(\mu^{\sigma^{h_1+1}} - \xi) + V^{h_2}(\eta^{\sigma^{h_1+h_2+1}} z - c^{\sigma^{h_2}}\eta) = 0. $$

The first two of these four equations enable us to eliminate λ and μ. For the remaining equations in ξ and η: assuming the existence of a solution modulo V^s, it is an elementary (though tricky in the case $h_2 = h_3$) exercise to show the existence of a solution modulo V^{s+1}. For details, see [11] §2.2. Taking limits, we find the desired λ, μ, ξ and η. □

Note that if $h_3 = 0$ we have the nice form $N = V^{h_1}\begin{pmatrix} 0 & V^{h_2} \\ 1 & 0 \end{pmatrix}$. So a two-dimensional formal group G with $h_3(G) = 0$ is a formal group of "braid" type (cf. [2] § 10).

Also note, that if $h_3 > 0$ then $c_0 \neq 0$ (cf. (3.3)).

(3.9) **Definition.** The matrices N in (3.2), (3.3), (3.6) and (3.8) are called *normalized Hasse-Witt matrices* of C. In the case $h_1 = \infty$ put $N = 0$.

Note that if $h = \infty$ or if $h < \infty$ and $d = \min\{h_2, h_3\} = 0$ then C has a unique normalized HW matrix. In the following proposition we shall show that for *positive* d the normalized HW matrix is unique *modulo* the action of a finite group. The group and its action will be described explicitly.

(3.10) Assume $h < \infty$ and $d = \min\{h_2, h_3\} > 0$.

Proposition. *Let B and C be reduced D-modules of V-dimension 2. Assume*

$$ M = V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2}b \end{pmatrix} \quad \text{and} \quad N = V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2}c \end{pmatrix} $$

are normalized HW matrices of B resp. C, so that we have $b = \sum_{i=0}^{d-1} V^i[b_i]$ and $c = \sum_{i=0}^{d-1} V^i[c_i]$. Then B and C are isomorphic if and only if there exist elements $\lambda_0, \ldots, \lambda_{d-1}$ in F_{p^h}, with $\lambda_0 \neq 0$ and such that

$$ \sum_{i+j=s} (\lambda_i^{p^{h_1+h_2+j+1}} c_j - b_j^{p^i}\lambda_i) = 0 $$

for $s = 0, \ldots, d-1$.

Proof. Given such $\lambda_0, \ldots, \lambda_{d-1}$, put $\lambda = \sum_{i=0}^{d-1} V^i[\lambda_i]$. A straightforward verification shows that $\lambda^{\sigma^{h_1+h_2+1}} c \equiv b\lambda \bmod (p, V^d)$. Furthermore, put $\Lambda = \begin{pmatrix} \lambda^{\sigma^{-h_1-1}} & 0 \\ 0 & \lambda \end{pmatrix}$. Since $\lambda_0 \neq 0$, the matrix Λ is invertible. Now $L = {}^t({}^t\Lambda^{\sigma}.{}^tN.({}^t\Lambda)^{-1})$ is a matrix of F relative to some V-basis of C (cf. (2.1)). Using the fact that $\lambda^{\sigma^h} = \lambda$, a straightforward calculation shows that $L = V^{h_1}\begin{pmatrix} 0 & V^{h_2+h_3} \\ 1 & V^{h_2}z \end{pmatrix}$ with $z = \lambda^{\sigma^{h_1+h_2+1}} c \lambda^{-1}$, hence $b \equiv z \bmod (p, V^d)$. In view of (3.8), the matrix M is a HW matrix of C, so that B and C are isomorphic (cf. (1.11)).

Conversely, assume B and C are isomorphic. Say $\Phi\colon B \to C$ is an isomorphism. Let (β_1, β_2) and (γ_1, γ_2) be V-bases of B resp. C with HW matrices M resp. N. Furthermore, let $P = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$ be the Teichmüller matrix of Φ relative to (β_1, β_2) and (γ_1, γ_2) (cf. (1.8)). Then P is invertible and

$$ {}^tP^{\sigma}.{}^tN - {}^tM.{}^tP = Q(pI_2 - V.{}^tN) $$

for some $Q = \begin{pmatrix} q & r \\ s & t \end{pmatrix} \in M_{2\times 2}(D^+)$ (cf. (1.13)). Now the $(1,1)$-entry of this matrix equation gives $-V^{h_1}x \equiv pq \bmod V^{h_1+h_2+h_3}$. Since P is the Teichmüller matrix of Φ, the coefficients of x are of the form $[x_i]$, hence

$$ x \equiv 0 \bmod V^{h_2+h_3} \quad \text{and} $$

$$ q \equiv 0 \bmod V^{h_1+h_2+h_3}. $$

Using this, we find that the $(1,2)$-entry of the matrix equation gives $V^{h_1}(w^{\sigma^{h_1+1}} - z) \equiv 0 \bmod (p, V^{h_1+d})$, hence

$$ w^{\sigma^{h_1+1}} \equiv z \bmod V^d. $$

Furthermore, the $(2,1)$-entry gives $ps \equiv 0 \bmod V^{h_2+d}$, so that $s \equiv 0 \bmod V^{h_2+d}$. In an analogous way, the $(2,2)$-entry gives $t \equiv 0 \bmod V^d$. Now a closer look at the $(2,1)$-entry gives

$$ V^{h_1+h_2+h_3}(z^{\sigma^{h_1+h_2+h_3+1}} - w) \equiv 0 \bmod (p, V^{h_1+h_2+h_3+d}) $$

and since $z \equiv w^{\sigma^{h_1+1}} \bmod V^d$ we find

$$ w^{\sigma^h} \equiv w \quad \text{and} \quad z^{\sigma^h} \equiv z \bmod V^d. $$

The $(2,2)$-entry gives $V^{h_1+h_2}(z^{\sigma^{h_1+h_2+1}} c - bz) \equiv 0 \bmod (p, V^{h_1+h_2+d})$, hence

$$ z^{\sigma^{h_1+h_2+1}} c - bz \equiv 0 \bmod (p, V^d). $$

Since P is the Teichmüller matrix of Φ, we may put $z \equiv \sum_{i=0}^{d-1} V^i[\lambda_i] \bmod V^d$. Then $P \equiv \begin{pmatrix} \lambda_0^{p^{-h_1-1}} & 0 \\ y_0 & \lambda_0 \end{pmatrix} \bmod (p, V)$, hence $\lambda_0 \neq 0$. Since $z^{\sigma^h} \equiv z \bmod V^d$, the λ_i are in F_{p^h}

and the relation which we just have derived from the (2,2)-entry of the matrix equation gives the desired relations between the b_i, c_i and λ_i. □

(3.11) We conclude with an explicit description of the module space S_{h_1,h_2,h_3} of isomorphism classes of 2-dimensional formal groups G with $h_i(G)=h_i$ $(i=1,2,3)$.

If one of the h_i is infinite then the dimension of S_{h_1,h_2,h_3} is zero (cf. (2.3) for $h_1=\infty$, (3.3) for $h_2=\infty$ and (3.6) for $h_3=\infty$).

If h_1, h_2 and h_3 are finite, or equivalently, the *height* $h=2h_1+h_2+h_3+2$ is finite, put $d=\min\{h_2,h_3\}$. If $d=0$ then again the dimension of S_{h_1,h_2,h_3} is zero (cf. (3.2) for $h_2=0$ and (3.8) for $h_3=0$).

For positive d, consider the variety

$$X_d=\{(c_0,\ldots,c_{d-1})\in k^d \mid c_0\neq 0\}.$$

In view of (3.8), there exists a surjection from X_d to S_{h_1,h_2,h_3}.

Let R be the noncommutative ring $k[V]/(V^d)$ with multiplication rule $aV=Va^p$ $(a\in k)$. The elements of R may uniquely be written in the form $c=\sum_{i=0}^{d-1} V^i c_i$ $(c_i\in k)$. Moreover, c is a unit in R if and only if $c_0\neq 0$. There is an obvious one-to-one correspondence between the elements of R^* and the elements of X_d.

Let Γ be the subgroup $(F_{p^h}[V]/(V^d))^*$ of R^*, so that Γ is a finite group of order $p^{hd}-p^{h(d-1)}$. Define the endomorphism σ of Γ by $\left(\sum_{i=0}^{d-1} V^i\lambda_i\right)^\sigma=\sum_{i=0}^{d-1} V^i\lambda_i^p$. Furthermore, define a left action l of Γ on R^* as follows: if $\lambda\in\Gamma$ and $c\in R^*$ then $l(\lambda,c)=\lambda^{\sigma^{h_1+h_2+1}}c\lambda^{-1}$.

Thus we also have a left action of Γ on X_d. In view of (3.10), two elements b and c of X_d correspond to the same element of S_{h_1,h_2,h_3} if and only if b and c are in the same orbit under the action of Γ. In other words: S_{h_1,h_2,h_3} is the quotient of the variety X_d under the action of Γ.

Comments. In [11], one can find the explicit relations between this *covariant* classification of two-dimensional smooth commutative formal groups over an algebraically closed field of positive characteristic and Manin's *contravariant* classification (cf. [13] §III.8). The covariant classification turns out to be more detailed and more explicit then the contravariant classification. Taking $W(k)$ as base ring instead of k, the question of finding normal forms for the Hasse-Witt matrix is more tedious, since F and V do not commute. See [6] and [7] for results in this area.

References

1. Bourbaki, N.: Eléments de Mathématique: Algèbre, chap. 2, Algèbre Linéaire. Paris: Hermann, 1962
2. Dieudonné, J.: Lie Groups and Lie Hyperalgebras over a Field of Characteristic $p>0$ (IV). Amer. J. of Math. 77, 429-452 (1955)

3. Dieudonné, J.: Introduction to the Theory of Formal Groups. New York: Dekker 1973
4. Ditters, E.J.: Groupes Formels. Cours 3e cycle 1973–1974. Orsay: Univ. Paris XI, 1974
5. Ditters, E.J.: On the Classification of Smooth Commutative Formal Groups. Journées de Géométrie Algébrique de Rennes. Astérisque **63**, pp. 67–71. Paris: Soc. Math. de France (1979)
6. Ditters, E.J.: On the Covariant Dieudonné Module of an Abelian Variety of Dimension Two over $W(k)$. Math. Z. **176**, 135–150 (1981)
7. Ditters, E.J.: The Formal Group of an Abelian Variety, defined over $W(k)$. Report no. **144**, Vrije Universiteit (1980)
8. Hasse, H., Witt, E.: Zyklische Unverzweigte Erweiterungskörper von Primzahlgrade p über einem Algebraischen Funktionenkörper der Charakteristik p. Monatsh. Math. Phys. **43**, 477–492 (1936)
9. Hazewinkel, M.: Formal Groups and Applications. New York: Academic Press 1978
10. Honda, T.: On the Theory of Commutative Formal Groups. J. Math. Soc. of Japan **22**, 213–246 (1970)
11. Kneppers, H.A.W.M.: The Covariant Classification of Two-Dimensional Smooth Commutative Formal Groups over an Algebraically Closed Field of Positive Characteristic. Mathematical Centre Tracts **157**. Amsterdam: Mathematical Centre 1983
12. Lazard, M.: Commutative Formal Groups. Lecture Notes in Math. **443**. Berlin-Heidelberg-New York: Springer 1975
13. Manin, Yu.I.: The Theory of Commutative Formal Groups over Fields of Finite Characteristic. Russ. Math. Surveys **18**, 1–83 (1963)
14. Manin, Ju.I.: The Hasse-Witt Matrix of an Algebraic Curve. Amer. Math. Soc. Transl. (2) **45**, 245–264 (1965)

Received June 27, 1983

Math. Z. 189, 167–184 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Die zweite Variation von Minimalflächen im $\mathbb{R}^p$ mit polygonalem Rand

Friedrich Sauvigny

Institut für Mathematik der Technischen Universität Clausthal,
Erzstr. 1, D-3392 Clausthal-Zellerfeld, Bundesrepublik Deutschland

§ 0. Einleitung

Sei Γ ein Polygon im $\mathbb{R}^p$ $(p \geqq 2)$ mit $N+3$ Eckpunkten $(N \geqq 1)$. Sei D das Dirichletintegral über die Einheitskreisscheibe $B \subset \mathbb{C}$ und $T \subset \mathbb{R}^N$ ein gewisser Parameterbereich. Eine Minimalfläche $x(u, v)$ zu Γ kann man dann gemäß einer Arbeit von I. Marx und M. Shiffman [11] (s. [7] im Fall $p \geqq 2$) einbetten in eine Schar von Quasiminimalflächen

$$x(u, v, \tau)\colon \bar{B} \times T \to \mathbb{R}^p.$$

Der Punkt $\mathring{\tau} \in T$ mit $x(u, v, \mathring{\tau}) = x(u, v)$ ist dann kritischer Punkt der Shiffmanschen Funktion

$$\Theta(\tau) := D(x(\ , \tau)), \qquad \tau \in T.$$

In den Arbeiten [6, 7] von E. Heinz wurde Θ als analytisch in T erkannt. Weiterhin wurde in [6], Satz 2 bewiesen, daß $x(u, v, \tau)$ analytisch von den Parametern abhängt, und es wurden parameterabhängige singuläre Entwicklungen in den Ecken des Polygons angegeben.

Hierauf aufbauend werden wir in Satz 1 die zweite Variation von Quasiminimalflächen berechnen (für den Fall $p = 3$ vergleiche [11] 4., 5.). Wir modifizieren hierzu in Hilfssatz 3 die Flächenschar durch innere Variationen und glätten sie an den Ecken und in den Verzweigungspunkten ab.

Gemäß Hilfssatz 5 bleibt dann nur die zweite Variation des Flächenintegrals unter zulässigen Variationen im Sinne von Definition 1 zu berechnen, was wir in § 1 ausführen. Das Funktional der zweiten Variation wird in Definition 3 und 4 erklärt.

In § 4, Satz 2 erhalten wir dann ein Kriterium für den positiv semidefiniten Charakter der Hesseschen Matrix von Θ. Kombinieren wir dies mit den Überlegungen von E. Heinz [8] zur Degeneration der Hesseschen Matrix, so können wir in Satz 3 ein Kriterium für ihren positiv definiten Charakter beweisen. Im Fall $p = 3$ ist diese Aussage – ohne vollständigen Beweis – in [11] Theorem II enthalten.

In einer anschließenden Arbeit soll Satz 3 genutzt werden, um einen Eindeutigkeitssatz für Minimalflächen zu Polygonen im $\mathbb{R}^p$ zu beweisen.

§1. Die zweite Variation des Flächeninhalts bei zulässigen Variationen – grundlegende Definitionen

Es sei

$$B := \{w = u + iv \in \mathbb{C} \,|\, |w| < 1\}$$

die Einheitskreisscheibe und Γ ein geschlossenes, doppelpunktfreies Polygon im $\mathbb{R}^p$ $(p \geqq 2)$ mit den Eckpunkten $e_1, \ldots, e_{N+3} (N \geqq 1)$. Nun sei $x(u, v)$: $\bar{B} \to \mathbb{R}^p$ eine Minimalfläche zu Γ mit nicht notwendig monotonen Randbedingungen (s. [11, 8]).

Den Punkten

$$w_k := e^{i\mathring{\tau}_k} \in \partial B \qquad (k = 1, \ldots, N+3)$$

$(0 < \mathring{\tau}_1 < \mathring{\tau}_2 < \ldots < \mathring{\tau}_N < \pi;\ \mathring{\tau}_{N+k} = \frac{\pi}{2}(1+k),\ k = 1, 2, 3)$ entsprechen die Eckpunkte $\{e_k\}_{k=1,\ldots,N+3}$ des Polygons Γ.

Setzen wir noch

$$\begin{aligned} \gamma_k &:= \{e^{i\varphi} | \mathring{\tau}_k < \varphi < \mathring{\tau}_{k+1}\} \\ \Gamma_k &:= \{x = \lambda(e_{k+1} - e_k) \in \mathbb{R}^p | \lambda \in \mathbb{R}\} \end{aligned} \tag{1}$$

$(\mathring{\tau}_{N+4} = \mathring{\tau}_1 + 2\pi;\ e_{N+4} = e_1)$, so gilt also für $k = 1, \ldots, N+3$

$$x(\gamma_k) \subset e_k + \Gamma_k. \tag{2}$$

Die Menge der Verzweigungspunkte von x in

$$\bar{B} \setminus \{w_1, \ldots, w_{N+3}\}$$

ist gemäß [8] §2 endlich, und wir bezeichnen sie mit

$$\{w_{N+4}, \ldots, w_L\} \qquad L \geqq N+3.$$

(Im Fall $L = N+3$ ist die leere Menge gemeint).

Zu $\delta > 0$ erklären wir

$$V_\delta := \bigcup_{k=1}^{L} \{w \in \bar{B} \,|\, |w - w_k| < \delta\}.$$

In $\bar{B} \setminus V_\delta$ ist die Metrik der Minimalfläche

$$dx^2 = E(u, v)(du^2 + dv^2)$$

regulär. Wichtig ist nun

Definition 1. *Sei* $\varepsilon_0, \delta > 0$. *Eine Schar von Abbildungen*

$$t(u, v, \varepsilon) = t(w, \varepsilon)\colon \bar{B} \times (-\varepsilon_0, +\varepsilon_0) \to \mathbb{R}^p \in C^\infty(\bar{B} \times (-\varepsilon_0, +\varepsilon_0))$$

mit $t(w,0)\equiv 0$ *in* $\bar{B}$ *heißt zulässige Variation der Minimalfläche* x, *falls für alle* $|\varepsilon|<\varepsilon_0$ *gilt:*

$$t(\gamma_k,\varepsilon)\subset\Gamma_k,$$
$$t(w,\varepsilon)=0 \text{ in } V_\delta.$$

Bemerkung. Eine zulässige Variation verschwindet in der Umgebung der singulären Punkte von x und variiert in einem linearen Raum.

$$\frac{\partial}{\partial w}:=\frac{1}{2}\left(\frac{\partial}{\partial u}-i\frac{\partial}{\partial v}\right).$$

Definition 2. *Ist* $z(w)\colon \bar{B}\to\mathbb{C}^p$ *ein beliebiges Vektorfeld, so nennen wir*

$$z^*(w):=z(w)-\frac{1}{E}(z\cdot x_u)\,x_u-\frac{1}{E}(z\cdot x_v)\,x_v$$
$$=z(w)-\frac{1}{|x_w|^2}(z\cdot x_{\bar{w}})\,x_w-\frac{1}{|x_w|^2}(z\cdot x_w)\,x_{\bar{w}}$$

die Normalkomponente von z *(bezüglich der Minimalfläche* x*).*

Unter $H^{1,2}(B)$ verstehen wir den Sobolevraum der im schwachen Sinne einmal differenzierbaren Funktionen, welche zusammen mit ihren ersten Ableitungen quadratisch integrierbar sind; mit $L^2(B)$ bezeichnen wir die quadratintegrablen Funktionen.

Für $\varphi(u,v)\in H^{1,2}(B)$ ist mit φ_u auch $\varphi_u^*\in L^2(B)=$*im folgenden ist mit* φ_u^* ... *immer* $(\varphi_u)^*$ ... *gemeint* =und wir können folglich erklären

Definition 3. *Für alle* $\varphi\in C^1(B)\cap H^{1,2}(B)$ *heißt*

$$\mathbf{I}_x(\varphi):=\int\int_B\{|\varphi_u^*|^2+|\varphi_v^*|^2+\frac{2}{E}\,((x_u\cdot\varphi_u)(x_v\varphi_v)-(x_u\varphi_v)(x_v\varphi_u))\}\,du\,dv$$

die zweite Variation von x *(in Richtung* φ*).*

Das Flächenfunktional bezeichnen wir mit

$$A(x):=\int\int_B(x_u^2x_v^2-(x_u\cdot x_v)^2)^{1/2}\,du\,dv.$$

Für die Ableitung nach dem Parameter ε verwenden wir $'$.

Hilfssatz 1. *Sei* $t(u,v,\varepsilon)$ *eine zulässige Variation der Minimalfläche* x, $\varphi(u,v):=t'(u,v,o)$ *und*

$$\bar{x}(u,v):=x(u,v)+t(u,v,\varepsilon),\quad (u,v)\in\bar{B},\qquad |\varepsilon|<\varepsilon_0.$$

Dann ist die Funktion $\varepsilon\to A(\bar{x})$ *in* $\varepsilon=0$ *beliebig oft differenzierbar, und es gilt für* $\varepsilon\to 0$*:*

$$A(\bar{x})=A(x)+\tfrac{1}{2}\varepsilon^2\,\mathbf{I}_x(\varphi)+o(\varepsilon^2).$$

Beweis. Da $t(\ ,\varepsilon)$ in V_δ verschwindet, ist $A(\bar{x})$ differenzierbar.

1. Wir berechnen: $\bar{x}_u = x_u + t_u$, $\bar{x}_v = x_v + t_v$

$$\bar{E} = \bar{x}_u \cdot \bar{x}_u = E + 2x_u t_u + t_u^2,$$
$$\bar{F} = \bar{x}_u \cdot \bar{x}_v = x_u t_v + x_v t_u + t_u t_v,$$
$$\bar{G} = \bar{x}_v \cdot \bar{x}_v = E + 2x_v t_v + t_v^2$$

$\bar{W}^2 := \bar{E}\bar{G} - \bar{F}^2 = E^2 \cdot g(u, v, \varepsilon)$ mit

$$g(u,v,\varepsilon) = 1 + \frac{2}{E}(x_u t_u + x_v t_v) + \frac{1}{E}(t_u^2 + t_v^2)$$
$$+ \frac{4}{E^2}(x_u t_u)(x_v t_v) - \frac{1}{E^2}(x_u t_v + x_v t_u)^2 + \ldots.$$

Da nun $t_u(u,v,0) \equiv 0 \equiv t_v(u,v,0)$ gilt, und beachten wir $\varphi(u,v) := t'(u,v,0)$, so sehen wir:

$$g(u,v,0) = 1,$$
$$g'(u,v,0) = \frac{2}{E}(x_u \cdot \varphi_u + x_v \cdot \varphi_v),$$
$$g''(u,v,0) = \frac{2}{E}(x_u \cdot t_u'' + x_v \cdot t_v'')$$
$$+ \frac{2}{E}(\varphi_u^2 + \varphi_v^2) + \frac{8}{E^2}(x_u \cdot \varphi_u)(x_v \cdot \varphi_v)$$
$$- \frac{2}{E^2}(x_u \cdot \varphi_v + x_v \cdot \varphi_u)^2.$$

Nun ist $\bar{W}(u,v,\varepsilon) = E(g(u,v,\varepsilon))^{1/2}$ und somit

$$\bar{W}' = \tfrac{1}{2} E g^{-1/2} g',$$
$$\bar{W}'' = \tfrac{1}{2} E g^{-1/2} g'' - \tfrac{1}{4} E g^{-3/2}(g')^2.$$

Wir erhalten dann:

$$\bar{W}'(u,v,0) = x_u \varphi_u + x_v \varphi_v,$$
$$\bar{W}''(u,v,0) = x_u t_u'' + x_v t_v'' + \varphi_u^2 + \varphi_v^2$$
$$+ \frac{4}{E}(x_u \cdot \varphi_u)(x_v \cdot \varphi_v) - \frac{1}{E}(x_u \varphi_v + x_v \varphi_u)^2 - \frac{1}{E}(x_u \varphi_u + x_v \varphi_v)^2$$
$$= x_u t_u'' + x_v t_v'' + |\varphi_u^*|^2 + |\varphi_v^*|^2 + \frac{2}{E}(x_u \varphi_u)(x_v \varphi_v)$$
$$- \frac{2}{E}(x_u \varphi_v)(x_v \varphi_u).$$

2. Nun ist

$$\frac{d}{d\varepsilon} A(\bar{x})\Big|_{\varepsilon=0} = \int\int_B \bar{W}'(u,v,0)\,du\,dv$$
$$= \int\int_B ((x_u\varphi)_u + (x_v\varphi)_v - \Delta x\cdot\varphi)\,du\,dv$$
$$= \int_{\partial B} x_r\cdot\varphi\,ds = 0.$$

Genauso sehen wir $(t''(w,0)\in\Gamma_k, w\in\gamma_k)$

$$\int\int_B (x_u t_v'' + x_v t_v'') = 0,$$

und wir erhalten

$$\frac{d^2}{d\varepsilon^2} A(\bar{x})\Big|_{\varepsilon=0} = \mathbf{I}_x(\varphi). \qquad \text{w.z.b.w.}$$

Bemerkung 1. In allgemeinerer Situation wird in [10] S. 45–54 die zweite Variation berechnet; allerdings mit fixierten Randwerten (s. auch [1, 16]).

Bemerkung 2. Sind $t_j(u,v,\varepsilon)(j=1,2)$ zwei zulässige Variationen der Minimalfläche x mit

$$t_1'(u,v,0) = t_2'(u,v,0) \text{ in } \bar{B},$$

so gilt für die Flächenscharen

$$x_j(u,v,\varepsilon) = x(u,v) + t_j(u,v,\varepsilon),$$
$$A(x_1)''|_{\varepsilon=0} = A(x_2)''|_{\varepsilon=0}.$$

Hiermit beweisen wir:

Hilfssatz 2. *$t(u,v,\varepsilon)$ sei zulässige Variation der Minimalfläche x und $\bar{x}(u,v) := x(u,v) + t(u,v,\varepsilon)$.*

Für die Normalkomponente des Vektorfelds $t'(\ ,0)$

$$\psi(u,v) := (t'(u,v,0))^*, \qquad (u,v)\in\bar{B}$$

ist $\psi|_{\partial B} = 0$ erfüllt, und es gilt:

$$\frac{d^2}{d\varepsilon^2} A(\bar{x})\Big|_{\varepsilon=0} = \mathbf{I}_x(\psi).$$

Beweis. 1. Für $w\in\gamma_k$ ist $t(w,\varepsilon)\in\Gamma_k$, also $t'(w,0)\in\Gamma_k$ und somit $\psi(w) = (t'(w,0))^* = 0$ erfüllt.

2. Die Funktion

$$c(w) := |x_w|^{-2}(t'(w,0)\cdot x_{\bar{w}})$$

erfüllt die Bedingungen (vgl. [8] (5.6)):

$$c\in C^\infty(\bar{B}), c\Big|_{V_\delta} = 0 \text{ und } \operatorname{Re}\left(\frac{c(w)}{w}\right) = 0 \text{ auf } \partial B. \tag{3}$$

Mit der Glättungsfunktion $\eta(w)\in C^\infty(\mathbb{C})$ mit

$$\eta(w)=\begin{cases}1 & |w|\leqq\frac{1}{3}\\ 0 & |w|\geqq\frac{2}{3}\end{cases}$$

erklären wir für $\varepsilon\in\mathbb{R}$; $w\in\bar{B}$:

$$h(w,\varepsilon):=w+\eta(w)\,\varepsilon c(w)+(1-\eta(w))(e^{\varepsilon\frac{c(w)}{w}}-1)\,w.$$

Wegen (3) und $f(w,0)=w$ ist für alle $|\varepsilon|<\varepsilon_1$

$$h(w,\varepsilon)\colon \bar{B}\to\bar{B}\qquad C^\infty\text{-diffeomorph}$$

erfüllt, und es gilt:

$$\begin{aligned}&h|_{V_\delta}=\mathrm{id}|_{V_\delta}\\ &h'(w,0)=c(w),\ w\in\bar{B}.\end{aligned}\tag{4}$$

3. In $\tilde{t}(w,\varepsilon):=(x(\ )+\varepsilon\psi(\ ))\circ h(w,\varepsilon)-x(w)$ erhalten wir eine zulässige Variation von x mit

$$\begin{aligned}\tilde{t}'(w,0)&=\psi(w)+x_w\,h'(w,0)+x_{\bar{w}}\overline{h'(w,0)}\\ &=(t'(w,0))^*+cx_w+\bar{c}x_{\bar{w}}=t'(w,0).\end{aligned}$$

Wir setzen

$$\tilde{x}(u,v)=x+\tilde{t}=(x(\ )+\varepsilon\psi(\ ))\circ h(w,\varepsilon).$$

Beachten wir Bemerkung 2 zu Hilfssatz 1 und die Parameterinvarianz von A, so erhalten wir insgesamt mit Hilfssatz 1:

$$\frac{d^2}{d\varepsilon^2}A(\tilde{x})\bigg|_{\varepsilon=0}=\frac{d^2}{d\varepsilon^2}A(\tilde{x})\bigg|_{\varepsilon=0}=\frac{d^2}{d\varepsilon^2}A(x+\varepsilon\psi)\bigg|_{\varepsilon=0}=\mathbf{I}_x(\psi).\qquad\text{w.z.b.w.}$$

§2. Das Funktional der zweiten Variation

Wegen Hilfssatz 2 hat das Funktional $\mathbf{I}_x$ sinnvollerweise den folgenden Definitionsbereich

Definition 4.

$$\begin{aligned}\mathfrak{N}_x:=\{&\psi(u,v)\colon \bar{B}\to\mathbb{R}^p\in C^1(B)\cap C^0(\bar{B})\cap H^{1,2}(B)|\\ &\psi\cdot x_u=0=\psi\cdot x_v \text{ in } B,\ \psi=0 \text{ auf } \partial B\}\end{aligned}$$

nennen wir den Normalraum an die Minimalfläche x.

Schreiben wir die im Anhang dargestellte Gauß'sche Krümmungsgleichung (25) in komplexer Form (im Fall $p=3$ s. [8] (3.2)), so erhalten wir für die Gaußsche Krümmung K der Minimalfläche x; $x_{ww}{}^*=(x_{ww})^*$:

$$EK=-2\frac{|x_{ww}{}^*|^2}{|x_w|^2}.\tag{5}$$

Parallel zu [8] (2.13), (2.15) überlegen wir: Ist $w_0 \in \{w_1, \ldots, w_L\}$ ein singulärer Punkt von x, so gilt

$$x_w(w) = a(w-w_0)^\alpha + O(|w-w_0|^{\alpha+\sigma}), \qquad w \to w_0$$

mit $a \in \mathbb{C}^p \setminus \{0\}$, $\alpha \in (-1, +\infty)$, $\sigma > 0$. Wegen

$$x_{ww}{}^* = O(|w-w_0|^{\alpha-1+\sigma})$$

folgt

$$\frac{|x_{ww}{}^*|}{|x_w|} = O(|w-w_0|^{-1+\sigma}), \qquad w \to w_0.$$

Somit hat wegen (5) die Minimalfläche x endliche Gesamtkrümmung

$$-\int\int_B EK \, du \, dv < +\infty. \tag{6}$$

Bemerkung. Für all $\psi \in \mathfrak{N}_x$ schreibt sich die *zweite Variation* in der Form:

$$\mathbf{I}_x(\psi) = \int\int_B \left\{ |\psi_u{}^*|^2 + |\psi_v{}^*|^2 + \frac{2}{E}((x_{uu}\cdot\psi)(x_{vv}\cdot\psi) - (x_{uv}\cdot\psi)^2) \right\} du \, dv.$$

Der Gaußschen Krümmungsgleichung (25) entnehmen wir:

$$\begin{aligned} 0 &\leqq \frac{1}{E}\{(x_{uv}\cdot\psi)^2 - (x_{uu}\cdot\psi)(x_{vv}\cdot\psi)\} \\ &\leqq |\psi|^2 \frac{1}{E}\{|x_{uv}{}^*|^2 + |x_{uu}{}^*||x_{vv}{}^*|\} \\ &= |\psi|^2(-EK). \end{aligned} \tag{7}$$

Wegen (6), (7) existiert der zweite Anteil von $\mathbf{I}_x(\psi)$.

Spezialfall $p=3$. Sei $\xi(u,v)$ die Normale an x, $\psi = \zeta \cdot \xi$ mit $\zeta: \bar{B} \to \mathbb{R}$, $\zeta|_{\partial B} = 0$.

$$\psi_u = \zeta_u \xi + \zeta \xi_u, \psi_u{}^* = \zeta_u \xi; h_{11} = x_{uu} \cdot \xi, \ldots.$$

Wir kommen so zum wohlbekannten Integral (s. [2]):

$$\mathbf{I}_x(\zeta\xi) = \int\int_B \{|\nabla\zeta|^2 + 2EK\zeta^2\} \, du \, dv.$$

§ 3. Die zweite Variation von Quasiminimalflächen

Wie in [8] § 1 bezeichnen wir:

$$T := \{\tau = (\tau_1, \ldots, \tau_N) \in \mathbb{R}^N | 0 < \tau_1 < \ldots < \tau_N < +\pi\},$$

$$\tau_{N+k} = \frac{\pi}{2}(1+k), \quad k = 1, 2, 3, \quad \tau_{N+4} = \tau_1 + 2\pi.$$

$\mathscr{F}(\tau)$ sei die Menge aller Funktionen

$$x(w)=(x^1(u,v),\ldots,x^p(u,v))\colon \bar{B}\to\mathbb{R}^p\in C^2(B)\cap C^0(\bar{B}),$$

welche die Kreisbögen

$$[e^{i\tau_k},e^{i\tau_{k+1}}]:=\{e^{i\varphi}|\tau_k<\varphi<\tau_{k+1}\}$$

in die Geraden $e_k+\Gamma_k$ $(k=1,\ldots,N+3)$ abbilden.

Die eindeutig bestimmte Lösung des Variationsproblems

$$D(x):=\int\int_B (x_u^2+x_v^2)\,du\,dv\to\text{Minimum},\qquad x\in\mathscr{F}(\tau)$$

nennen wir *Quasiminimalfläche* $x(w,\tau)$, $w\in\bar{B}$.

Diese löst ein gemischtes elliptisches Randwertproblem (s. [8] § 1). Ist nun $\mathring{\tau}=(\mathring{\tau}_1,\ldots,\mathring{\tau}_N)\in T$ kritischer Punkt der analytischen *Shiffman'schen Funktion*

$$\Theta(\tau)=D(x(\ ,\tau))\colon T\to\mathbb{R},\tag{8}$$

so entspricht $\mathring{\tau}$ eine Minimalfläche $x(w)=x(w,\mathring{\tau})$ mit „Shiffmanschen Randbedingungen" (s. [6–8]).

Wie in § 1 beschrieben, genügt x den Randbedingungen (2), und bildet die Punkte $w_k=e^{i\mathring{\tau}_k}(k=1,\ldots,N+3)$ auf die Eckpunkte $\{e_k\}_k$ des berandenden Polygons Γ ab.

Berücksichtigen wir noch die Verzweigungspunkte außerhalb der Ecken, so erhalten wir in

$$\{w_1,\ldots,w_L\}\qquad (L\geqq N+3)$$

alle singulären Punkte von x.

Zu festem $\alpha\in\mathbb{R}^N$ betrachten wir die Schar der Quasiminimalflächen ($\varepsilon_0>0$ sei jeweils hinreichend klein gewählt):

$$y(w,\varepsilon):=x(w,\mathring{\tau}+\varepsilon\alpha),\qquad w\in\bar{B},\quad \varepsilon\in(-\varepsilon_0,+\varepsilon_0).$$

Die Ableitung nach dem Parameter $\varepsilon\,\dfrac{d}{d\varepsilon}$ bezeichnen wir mit $'$.

In diesem Abschnitt werden wir im wesentlichen die zweite Variation des parameterinvarianten Flächenfunktionals der Schar $y(\ ,\varepsilon)$ berechnen. Daher können wir $y(\ ,\varepsilon)$ durch innere Abbildungen zu einer Schar $z(\ ,\varepsilon)$ abändern, so daß letztere in einem linearen Raum variiert.

Hierzu konstruieren wir zu hinreichend kleinem $\delta>0$ folgende Diffeomorphismen von $\bar{B}$:

Sei $\chi(u,v)\in C^\infty(\mathbb{R}^2)$ mit $\chi(w)=\begin{cases}1 & |w|\leqq 1\\ 0 & |w|\geqq 2\end{cases}$.

Zum Parameter ε betrachten wir die Schar von Abbildungen

$$f(w,\varepsilon):=w\exp\left\{i\varepsilon\sum_{j=1}^{N}\alpha_j\chi(\delta^{-1}(w-e^{i\mathring{\tau}_j}))\right\}.$$

Für $|\varepsilon| \leqq \varepsilon_0(\alpha, \delta)$ ist

$$f(w, \varepsilon)\colon \bar{B} \to \bar{B} \qquad C^\infty\text{-diffeomorph,}$$

und genügt für $k = 1, \ldots, N$ und $w \in \bar{B}$ der Bedingung:

$$\begin{aligned} f(w, \varepsilon) &= w, && \text{falls } |w - e^{i\vartheta_k}| \geqq 2\delta \\ f(w, \varepsilon) &= w e^{i\varepsilon\alpha_k}, && \text{falls } |w - e^{i\vartheta_k}| \leqq \delta. \end{aligned} \tag{9}$$

Wir gehen nun über zur Flächenschar

$$z(w, \varepsilon) := y(f(w, \varepsilon), \varepsilon), \qquad w \in \bar{B}, \qquad |\varepsilon| < \varepsilon_0. \tag{10}$$

Nach [8], § 2 (2.8) gilt für $|w - e^{i\vartheta_k}| < 2\delta$, $|\varepsilon| < \varepsilon_1(\alpha)$ und $k = 1, \ldots, N$

$$y(w, \varepsilon) = \operatorname{Re}\left\{ \sum_{j=1}^{p_k} g_j^{(k)}(w, \varepsilon)(w - e^{i\vartheta_k} e^{i\varepsilon\alpha_k})^{\rho_j + 1} \right\} + e_k. \tag{11}$$

Dabei ist $-1 < \rho_1 < \ldots < \rho_{p_k} \leqq 0$ und $g_j^{(k)}$ in

$$|w - e^{i\vartheta_k}| < 2\delta, \qquad |\varepsilon| < \varepsilon_1(\alpha) \text{ holomorph.}$$

Wegen (9) transformieren sich die singulären Faktoren in (11) wie folgt: Für $|w - e^{i\vartheta_k}| < \delta$, $w \in \bar{B}$ und $|\varepsilon| < \varepsilon_0$ gilt

$$(f(w, \varepsilon) - e^{i\vartheta_k} e^{i\varepsilon\alpha_k})^{\rho_j + 1} = e^{i\varepsilon(\rho_j + 1)\alpha_k}(w - e^{i\vartheta_k})^{\rho_j + 1}.$$

Hilfssatz 3. *Für $w \in \bar{B}$ mit $|w - e^{i\vartheta_k}| < \delta$ und $|\varepsilon| < \varepsilon_0(\alpha, \delta)$ gilt:*

$$z(w, \varepsilon) = \operatorname{Re}\left\{ \sum_{j=1}^{p_k} h_j^{(k)}(w, \varepsilon)(w - e^{i\vartheta_k})^{\rho_j + 1} \right\} + e_k$$

mit holomorphen Funktionen $h_j^{(k)}$ $(k = 1, \ldots, N + 3)$.

Bemerkung. Im Unterschied zu $y'(w, 0)$ hat nun $z'(w, 0)$ beschränktes Dirichletintegral!

Setzen wir nun

$$t(w, \varepsilon) := z(w, \varepsilon) - x(w), w \in \bar{B}, \qquad |\varepsilon| < \varepsilon_0,$$

so gilt

$$t(\gamma_k, \varepsilon) \subset \Gamma_k, \qquad k = 1, \ldots, N + 3. \tag{12}$$

Die in B reellanalytischen Funktionen bilden den Raumm $C^\omega(B)$. In Hinsicht auf Satz 1 beweisen wir schon jetzt

Hilfssatz 4. *Die Vektorfelder $y'(w, 0)$ und $z'(w, 0)$ haben die gleiche Normalkomponente $\psi(w)$, $w \in \bar{B}$.*

Diese gehört der Regularitätsklasse

$$\psi(w) := (y'(w, 0))^* = (z'(w, 0))^* \in C^\omega(B) \cap C^0(\bar{B}) \cap H^{1,2}(B)$$

an und erfüllt folgende Randbedingung

$$\psi = 0 \quad \textit{auf} \quad \partial B.$$

Beweis. 1.Wegen

$$\begin{aligned} z'(w,0) &= y_w(w,0) f'(w,0) + y_{\bar{w}}(w,0) \overline{f'(w,0)} + y'(w,0) \\ &= 2\,\mathrm{Re}(x_w f'(w,0)) + y'(w,0) \end{aligned}$$

stimmen $z'(w,0)$ und $y'(w,0)$ bis auf ein tangentiales Vektorfeld überein. Also folgt $(z'(w,0))^* = (y'(w,0))^*$, $w \in \bar{B}$.

2. Ist $w_0 \in \bar{B}$ ein Verzweigungspunkt von x außerhalb der Ecken, so gilt

$$x_w(w) = (w - w_0)^n c(w).$$

Dabei ist $c(w)$ analytische Funktion um den Punkt w_0 mit $c(w_0) \neq 0$ und $n \in \mathbb{N}$. Beachten wir Definition 2, so erhalten wir

$$\psi(w) = (y'(w,0))^* = y' - \frac{1}{|c|^2}(y' \cdot \bar{c})\, c - \frac{1}{|c|^2}(y' \cdot c)\, \bar{c}.$$

Somit ist ψ reellanalytisch um den Punkt w_0.

3. In den Ecken ist gemäß Hilfssatz 3 für $k = 1, \ldots, N+3$ die asymptotische Entwicklung

$$\begin{aligned} z'(w,0) &= 0(|w - e^{i\vartheta_k}|^{\sigma}) \\ z'_w(w,0) &= 0(|w - e^{i\vartheta_k}|^{\sigma - 1}), \qquad w \to e^{i\vartheta_k} \end{aligned} \tag{13}$$

mit einem $\sigma > 0$ richtig.

Zusammen mit 2. lesen wir hieraus ab

$$\psi(w) = (z'(w,0))^* = (y'(w,0))^* \in C^0(\bar{B}) \cap C^{\omega}(B).$$

Wegen (12) gilt

$$z'(w,0) \in \Gamma_k, \qquad w \in \gamma_k,$$

und somit folgt

$$\psi = 0 \text{ auf } \partial B.$$

4. Es bleibt noch $D(\psi) < +\infty$ nachzuweisen.

Beachten wir Definition 2, so erhalten wir mit Konstanten c_1, c_2:

$$|\psi_w| \leqq c_1 |z'_w| + c_2 |z'| \frac{|x_{ww}|}{|x_w|}$$

Da $\dfrac{|x_{ww}|}{|x_w|} = O(|w - e^{i\vartheta_k}|^{-1})$ richtig ist, liefert (13)

$$\psi_w = O(|w - e^{i\vartheta_k}|^{\sigma - 1}),\ w \to e^{i\vartheta_k}. \quad \text{w.z.b.w.}$$

Glätten wir nun noch $t(w,\varepsilon)$ in den singulären Punkten

$$\{w_1,\ldots,w_L\}$$

ab, so erhalten wir eine im Sinne von Definition 1 zulässige Variation der Minimalfläche x.

Für alle $n\in\mathbb{N}$ sei

$$U_n:=\bigcup_{\nu=1}^{L}\left\{w\in\bar{B}\,\middle|\,|w-w_\nu|<\frac{1}{n}\right\}$$

und

$$B_n:=\bar{B}\setminus U_n$$

erklärt.

Dann konstruieren wir eine beschränkte Funktionenfolge

$$\overset{n}{\eta}(w)\colon \bar{B}\to\mathbb{R}\in C^\infty(\bar{B})$$

mit $D(\overset{n}{\eta})\leqq\frac{1}{n}$ $(n\in\mathbb{N})$, so daß gilt:

(i) $\overset{n}{\eta}=0$ in einer Umgebung der singulären Punkte $\{w_k\}_{k=1,\ldots,L}$,

(ii) $\overset{n}{\eta}=1$ auf B_n.

Für alle $n\in\mathbb{N}$ ist dann

$$\overset{n}{t}(w,\varepsilon):=\overset{n}{\eta}(w)(z(w,\varepsilon)-x(w)),\qquad w\in\bar{B};\quad|\varepsilon|<\varepsilon_0$$

im Sinne von Definition 1 eine zulässige Variation der Minimalfläche x. Nennen wir $g(w,\varepsilon)\colon\bar{B}\to\bar{B}$ (C^∞-diffeomorph) die Umkehrfunktion von $f(w,\varepsilon)$, $|\varepsilon|<\varepsilon_0$, so setzen wir:

$$\overset{n}{y}(w,\varepsilon):=(x(\ )+\overset{n}{t}(\ ,\varepsilon))\circ g(w,\varepsilon),\ w\in\bar{B},\qquad|\varepsilon|<\varepsilon_0. \tag{14}$$

Wichtig ist nun der folgende

Hilfssatz 5. *Es gilt:*

$$\frac{d^2}{d\varepsilon^2}D(y(\ ,\varepsilon))\bigg|_{\varepsilon=0}=\lim_{n\to\infty}\left\{\frac{d^2}{d\varepsilon^2}D(\overset{n}{y}(\ ,\varepsilon))\bigg|_{\varepsilon=0}\right\}.$$

Beweis. 1. Für $n\geqq n_0$ ist wegen (9) die Abbildung

$$f(\ ,\varepsilon)\colon U_n\to f(U_n,\varepsilon)\text{ konform.} \tag{15}$$

Da $\overset{n}{\eta}=1$ auf $\bar{B}\setminus U_n$ gilt, folgt

$$\overset{n}{\eta}\circ g(w,\varepsilon)=1\ \text{ auf }\bar{B}\setminus f(U_n,\varepsilon).$$

Für alle $w\in\bar{B}\setminus f(U_n,\varepsilon)$ gilt:

$$\begin{aligned}\overset{n}{y}(w,\varepsilon)&=x\circ g(w,\varepsilon)+\overset{n}{t}\circ g(w,\varepsilon)\\&=x\circ g+\overset{n}{\eta}\circ g(z\circ g-x\circ g)=z\circ g=y(w,\varepsilon).\end{aligned}$$

Da das Dirichletintegral unter der Abbildung (15) invariant ist, folgt:

$$D(y(\ ,\varepsilon))-D(\overset{n}{y}(\ ,\varepsilon))=\iint\limits_{f(U_n,\varepsilon)}(|\nabla y|^2-|\nabla \overset{n}{y}|^2)=\iint\limits_{U_n}|\nabla z(\ ,\varepsilon)|^2-\iint\limits_{U_n}|\nabla(x+\overset{n}{t}(\ ,\varepsilon))|^2. \quad (16)$$

2. Nach Hilfssatz 3 gibt es eine Majorante $\zeta(w)\in L^2(B)$, so daß für alle $w\in B$, $|\varepsilon|<\varepsilon_0$, $\nu=0,1,2$ gilt:

$$|z_w^{(\nu)}(w,\varepsilon)|\leqq\zeta(w). \quad (17)$$

Somit können wir für alle $n\geqq n_0$ das erste Integral in (16) nach dem Parameter durchdifferenzieren und erhalten

$$\lim_{n\to\infty}\left\{\frac{d^2}{d\varepsilon^2}\iint\limits_{U_n}|\nabla z(\ ,\varepsilon)|^2\right\}=0.$$

3. Beachten wir die Definition von $\overset{n}{t}(\ ,\varepsilon)$ und (17), so zeigt man für alle $w\in B$, $|\varepsilon|<\varepsilon_0$, $n\geqq n_0$, $\nu=0,1,2$:

$$|(x(w)+\overset{n}{t}(w,\varepsilon))_w^{(\nu)}|\leqq c_1\zeta(w)+c_2|\overset{n}{\eta}_w|$$

mit reellen Konstanten c_1, c_2.

Wir differenzieren nun das zweite Integral in (16) durch und erhalten

$$\left|\frac{d^2}{d\varepsilon^2}\iint\limits_{U_n}|\nabla(x+\overset{n}{t}(\ ,\varepsilon))|^2\right|\leqq c_3\iint\limits_{U_n}\zeta^2+c_4\iint\limits_{U_n}\zeta|\overset{n}{\eta}_w|+c_5\,D(\overset{n}{\eta})$$

$$\leqq c_6\iint\limits_{U_n}\zeta^2+c_7\,D(\overset{n}{\eta});$$

mit reellen Konstanten c_i.

Da $\zeta\in L^2(B)$ und $\lim\limits_{n\to\infty}D(\overset{n}{\eta})=0$ gilt, folgt

$$\lim_{n\to\infty}\left\{\frac{d^2}{d\varepsilon^2}\iint\limits_{U_n}|\nabla(x+\overset{n}{t}(\ ,\varepsilon))|^2\right\}=0.$$

4. Die Identität (16) zusammen mit 2. und 3. ergibt nun die Behauptung. w.z.b.w.

Unsere Überlegungen fassen wir nun zusammen zu

Satz 1. *Zu $\alpha\in\mathbb{R}^N$ betrachten wir die Schar der Quasiminimalflächen*

$$y(w,\varepsilon):=x(w,\mathring{\tau}+\varepsilon\alpha),\ w\in\bar{B},\qquad |\varepsilon|<\varepsilon_0.$$

I. Die Normalkomponente ψ des Vektorfeldes $y'(w,0)$ liegt im Normalraum $\mathfrak{N}_x$ an die Minimalfläche und ist in B analytisch, d.h.

$$\psi(w):=(y'(w,0))^*\in\mathfrak{N}_x\cap C^\omega(B).$$

Weiterhin existiert $\mathbf{I}_x(\psi)$.

II. Für die differenzierbare Funktion $\varepsilon \to D(y(\ ,\varepsilon))$ *gilt:*

$$\frac{d^2}{d\varepsilon^2} D(y(\ ,\varepsilon))\Big|_{\varepsilon=0} \geqq 2\mathbf{I}_x(\psi).$$

Beweis. Die Aussage I. entnehmen wir Hilfssatz 4 und der Bemerkung zu Definition 4. Es bleibt die in II. angegebene Ungleichung nachzuweisen.

1. Hierzu betrachten wir wegen Hilfssatz 5 zunächst die Funktionen ($n \geqq n_0$)

$$\overset{n}{y}(w,\varepsilon) = (x(\ ) + \overset{n}{t}(\ ,\varepsilon)) \circ g(w,\varepsilon),\ w \in \bar{B}, \qquad |\varepsilon| < \varepsilon_0.$$

Da das Flächenfunktional A parameterinvariant ist, folgt:

$$D(\overset{n}{y}(\ ,\varepsilon)) \geqq 2A(\overset{n}{y}(\ ,\varepsilon)) = 2A(x + \overset{n}{t}(\ ,\varepsilon)), \qquad |\varepsilon| < \varepsilon_0. \tag{18}$$

Für festes $n \geqq n_0$ ist nun $\overset{n}{t}(\ ,\varepsilon)$, $|\varepsilon| < \varepsilon_0$ eine zulässige Variation gemäß Definition 1, und es gilt:

$$(\overset{n}{t}'(w,0))^* = \overset{n}{\eta}(z'(w,0))^* = \overset{n}{\eta}(y'(w,0))^* = \overset{n}{\eta}\psi.$$

Mit Hilfssatz 2 berechnen wir

$$\frac{d^2}{d\varepsilon^2} A(x + \overset{n}{t}(\ ,\varepsilon))\Big|_{\varepsilon=0} = \mathbf{I}_x(\overset{n}{\eta}\psi).$$

Für $\varepsilon = 0$ stimmen die Funktionen in (18) überein, und wir erhalten so aus dieser Ungleichung

$$\frac{d^2}{d\varepsilon^2} D(\overset{n}{y}(\ ,\varepsilon))\Big|_{\varepsilon=0} \geqq 2I_x(\overset{n}{\eta}\psi).$$

Anwendung von Hilfssatz 5 ergibt nun

$$\frac{d^2}{d\varepsilon^2} D(y(\ ,\varepsilon))\Big|_{\varepsilon=0} \geqq 2 \lim_{n\to\infty} \mathbf{I}_x(\overset{n}{\eta}\psi).$$

Wir sehen dann die behauptete Ungleichung ein, wenn wir noch zeigen:

2. $$\lim_{n\to\infty} \mathbf{I}_x(\overset{n}{\eta}\psi) = \mathbf{I}_x(\psi).$$

Es ist

$$\mathbf{I}_x(\overset{n}{\eta}\psi) = \int\int_B (|(\overset{n}{\eta}\psi)_u^*|^2 + |(\overset{n}{\eta}\psi)_v^*|^2) + \int\int_B \overset{n}{\eta}{}^2 \frac{2}{E}((x_{uu}\psi)(x_{vv}\psi) - (x_{uv}\psi)^2)$$

richtig, und wir können somit im zweiten Integral zur Grenze übergehen. Da $\overset{n}{\eta} = 1$ auf $\bar{B} \setminus U_n$ gilt, bleibt zu zeigen

$$\lim_{n\to\infty} \int\int_{U_n} |(\overset{n}{\eta}\psi)_u^*|^2 + |(\overset{n}{\eta}\psi)_v^*|^2 = 0. \tag{19}$$

Hierzu beachten wir

$$|(\overset{n}{\eta}\psi)_u{}^*|^2+|(\overset{n}{\eta}\psi)_v{}^*|^2\leqq|(\overset{n}{\eta}\psi)_u|^2+|(\overset{n}{\eta}\psi)_v|^2=4|(\overset{n}{\eta}\psi)_w|^2$$

$$\leqq 4(\overset{n}{\eta}|\psi_w|+|\overset{n}{\eta}_w||\psi|)^2$$

$$\leqq c(|\psi_w|^2+|\overset{n}{\eta}_w|^2).$$

Da $\psi\in H^{1,2}(B)$ und $\lim\limits_{n\to\infty}\mathrm{D}(\overset{n}{\eta})=0$ richtig ist, folgt (19). w.z.b.w.

§ 4. Über den positiv (semi-)definiten Charakter der Hesseschen Matrix der Shiffman'schen Funktion

Zunächst formulieren wir

Satz 2. *Dem kritischen Punkt $\mathring{\tau}\in T$ der Shiffmanschen Funktion $\Theta(\tau)$ entspreche eine Minimalfläche $x(w)=x(w,\mathring{\tau})$, $w\in\bar{B}$, mit nichtnegativer zweiter Variation, d.h.*

$$\mathbf{I}_x(\psi)\geqq 0 \quad \textit{für alle } \psi\in\mathfrak{N}_x.$$

Dann ist die Matrix

$$M=(\Theta_{\tau_j\tau_k}(\mathring{\tau}))_{j,k=1,\ldots,N}$$

positiv semidefinit.

Beweis. Für alle $\alpha\in\mathbb{R}^N\setminus\{0\}$ betrachten wir die Funktion

$$\Theta_\alpha(\varepsilon):=\Theta(\mathring{\tau}+\varepsilon\alpha), \qquad |\varepsilon|<\varepsilon_0.$$

Aus Satz 1 ersehen wir

$$\Theta_\alpha''(0)\geqq 2\,\mathbf{I}_x(\psi)\geqq 0. \quad \text{w.z.b.w.}$$

Bemerkungen. 1. Wir setzen in Satz 2 nicht voraus, daß die Minimalfläche x verzeigungspunktfrei ist.

2. Absolute Minima des Plateauproblems haben nicht negative zweite Variation (s. z.B. [14], Hilfssatz 8).

Da eine Minimalfläche x zum Polygon Γ endliche Gesamtkrümmung hat, können wir erklären (s. § 2 (6)).

Definition 5. *Eine Minimalfläche $x(u,v)\colon \bar{B}\to\mathbb{R}^p$ heißt strikt stabil, falls für alle nicht verschwindenden Funktionen $\zeta(u,v)\colon \bar{B}\to\mathbb{R}\in H^{1,2}(B)\cap C^0(\bar{B})$ mit $\zeta|_{\partial B}=0$ gilt:*

$$\iint\limits_B|\nabla\zeta|^2\,du\,dv>\iint\limits_B(-2EK\,\zeta^2)\,du\,dv.$$

Hilfssatz 6. *Sei $\psi(u,v)\in\mathfrak{N}_x\cap C^\omega(B)$ und $\zeta(u,v):=|\psi(u,v)|$. Dann ist $\zeta\in H^{1,2}(B)\cap C^0(\bar{B})$ und $\zeta|_{\partial B}=0$ erfüllt, und es gilt:*

$$\mathbf{I}_x(\psi)\geqq\iint\limits_B(|\nabla\zeta|^2+2EK\,\zeta^2)\,du\,dv.$$

Beweis. 1. Die reell-analytische Funktion ζ^2 verschwindet in einer abgeschlossenen Jordanschen Nullmenge. Auf dem Komplement

$$P := \{(u, v) \in B | \zeta(u, v) > 0\}$$

können wir ζ differenzieren, und es gilt

$$|\zeta_u| \leqq |\psi_u|, \qquad |\zeta_v| \leqq |\psi_v|.$$

Wir sehen etwa durch Abglätten in der Nullstellenmenge $B \setminus P$ ein, daß ζ schwach differenzierbar ist, also $\zeta \in H^{1,2}(B)$ gilt.

2. Sei nun

$$\xi(u, v) := \frac{\psi(u, v)}{\zeta(u, v)} : P \to S^{p-1} := \{x \in \mathbb{R}^p | |x| = 1\}$$

erklärt, so gilt in P:

$$\psi(u, v) = \zeta(u, v)\, \xi(u, v).$$

Wegen $\psi_u{}^* = (\zeta_u \xi + \zeta \xi_u)^* = \zeta_u \xi + \zeta \xi_u{}^*$ und $\xi \cdot (\xi_u{}^*) = \xi \cdot \xi_u = 0$, folgt in P:

$$|\psi_u{}^*|^2 + |\psi_v{}^*|^2 \geqq |\nabla \zeta|^2.$$

Zusammen mit der Abschätzung (7) im Anschluß an Definition 4 sehen wir die behauptete Ungleichung ein. w.z.b.w.

Anschließend an [8] §2 (2.5)–(2.7) nennen wir eine Minimalfläche zum Polygon Γ *strikt verzweigungspunktfrei*, falls für alle $w_0 \in \bar{B}$ gilt

$$\lim_{\substack{w \to w_0 \\ w \neq w_0}} |x_w(w)| \in (0, \infty].$$

Wir bemerken, daß eine in diesem Sinne verzweigungspunktfreie Minimalfläche an einer Ecke den spitzen Winkel überstreicht!

Im Fall $p = 3$ können wir aus Satz 2 und der Arbeit von E. Heinz [8], Satz 3 sofort folgern:

Satz 3. *Dem kritischen Punkt $\mathring{\tau} \in T$ der Shiffmanschen Funktion $\Theta(\tau)$ entspreche eine strikt stabile, strikt verzweigungspunktfreie Minimalfläche $x(w) = x(w, \mathring{\tau})$, $w \in \bar{B}$.*

Dann ist die Matrix

$$M = (\Theta_{\tau_j \tau_k}(\mathring{\tau}))_{j,k=1,\ldots,N}$$

positiv definit.

Beweis (für $p \geqq 2$). 1. Wie in Satz 1 betrachten wir zu festem $\alpha \in \mathbb{R}^N$ die Schar der Quasiminimalflächen

$$y(w, \varepsilon) := x(w, \mathring{\tau} + \varepsilon \alpha), \; w \in \bar{B}, \qquad |\varepsilon| < \varepsilon_0.$$

Wir setzen $\psi(w) := (y'(w, 0))^* \in \mathfrak{N}_x \cap C^\omega(B)$ und $\zeta(u, v) := |\psi(u, v)|$. Für die Funktion

$$\Theta_\alpha(\varepsilon) := \Theta(\mathring{\tau} + \varepsilon \alpha) = D(y(\,\cdot\,, \varepsilon)), \qquad |\varepsilon| < \varepsilon_0$$

gilt nach Satz 1 und Hilfssatz 6:

$$\Theta_\alpha''(0) \geqq 2\,\mathbf{I}_x(\psi) \geqq 2 \iint\limits_B (|\nabla \zeta|^2 - 2EK\zeta^2)\, du\, dv \geqq 0 \tag{20}$$

2. Ist für $\alpha \in \mathbb{R}^N$ $\Theta_\alpha''(O) = 0$, so folgt aus (20):

$$\mathbf{I}_x(\psi) = 0.$$

Da x strikt stabil ist, folgt $0 = \psi = (y'(w,0))^*$ und somit nach Definition 2:

$$y'(w,0) = 2\,\mathrm{Re}(a(w)\, x_w) \tag{21}$$

mit

$$a(w) := \frac{1}{|x_w|^2}(y' \cdot x_{\bar{w}}).$$

Da M positiv semidefinit ist, folgt aus

$$0 = \Theta_\alpha''(0) = \sum_{j,k=1}^{N} \Theta_{\tau_j \tau_k}(\mathring{\tau})\, \alpha_j \alpha_k$$

$$M(\alpha) = 0.$$

Dieses lineare Gleichungssystem für α ist nach [8], Hilfssatz 7, Beweis (I) äquivalent mit

$$0 = x_w \cdot (y'(w,0))_w.$$

Setzen wir nun (21) ein, so erhalten wir

$$0 = x_w \cdot (\bar{a} x_{\bar{w}})_w = \bar{a}_w |x_w|^2.$$

Die analytische Funktion a erfüllt

$$\mathrm{Re}\left(\frac{a(w)}{w}\right) = 0 \qquad \text{auf } \partial B,$$

und verschwindet in drei Randpunkten (s. [8] (5.3)–(5.6), (5.10)). Mit den Überlegungen in [8] § 5 folgt $a = 0$ in $\bar{B}$ und somit

$$0 = y'(w,0) = \sum_{k=1}^{N} x_{\tau_k}(w, \mathring{\tau})\, \alpha_k .$$

Da x verzweigungspunktfrei ist, sind gemäß [8] (5.7), (5.8) die Funktionen $\{x_{\tau_k}(\ , \mathring{\tau})\}_{k=1,\ldots,N}$ linear unabhängig, und wir erhalten

$$\alpha = (\alpha_1, \ldots, \alpha_N) = (0, \ldots, 0).$$

3. Aus (20) und 2. folgt für alle $\alpha \in \mathbb{R}^N \setminus \{0\}$

$$\Theta_\alpha''(0) > 0.$$

Somit ist M positiv definit. w.z.b.w.

Anhang: Die Gaußsche Krümmungsgleichung für Minimalflächen im $\mathbb{R}^p$

Diese grundlegende Identität wollen wir in unserem Spezialfall notieren. Dabei halten wir uns an die Darstellung in [3] (siehe auch [5, 9, 13]).

Auf einer offenen Menge $U \subset \mathbb{R}^2$ sei eine verzweigungspunktfreie Minimalfläche

$$x(u^1, u^2)\colon U \to \mathbb{R}^p$$

mit erster Fundamentalform

$$ds^2 = g_{jk}\, du^j\, du^k = E(u^1, u^2)\, \delta_{jk}\, du^j\, du^k \tag{22}$$

gegeben. Wie üblich wird über doppelt auftretende Indizes $j, k, l, m \in \{1, 2\}$ summiert; $\lambda, \mu \in \{1, \ldots, p-2\}$; δ_{jk} ist das Kroneckersymbol. Lokal ist es möglich, die Vektoren $x_j(u^1, u^2) := x_{u^j}(u^1, u^2)$ zu einem im $\mathbb{R}^p$ orthogonalen, beliebig oft differenzierbaren System

$$\{x_1, x_2, n_1, \ldots, n_{p-2}\}$$

mit

$$n_\lambda \cdot n_\mu = \delta_{\lambda\mu}$$

zu ergänzen. Nach [3] (6.1, 18) schreiben sich die Ableitungsgleichungen wie folgt ($x_{jk} := x_{u^j u^k}$):

$$x_{jk} = \Gamma_{jk}^l\, x_l + h_{jk}^\lambda\, n_\lambda.$$

Multiplikation mit x_m bzw. n_μ liefert

$$\begin{aligned} \Gamma_{jk}^m &= \frac{1}{E}\, x_{jk} \cdot x_m \\ h_{jk}^\mu &= x_{jk} \cdot n_\mu \end{aligned} \tag{23}$$

Mit dem Riemannschen Krümmungstensor R_{jklm} (s. [3] (6.3, 9)) können wir nun die Gaußsche Krümmung K erkären, und wir berechnen mit (23):

$$E^2 K(u^1, u^2) := R_{1212} = -\frac{E}{2}\, \Delta(\log E),$$

also

$$K(u^1, u^2) = -\frac{1}{2E}\, \Delta(\log E). \tag{24}$$

Für $\mu \in \{1, \ldots, p-2\}$ nennen wir

$$ds^2 = h_{jk}^\mu\, du^j\, du^k$$

die zweite Fundamentalform von x bezüglich der Normale n_μ.

Wegen $\Delta x = 0$ sehen wir aus (23), daß $(h_{jk}^\mu)_{j,k}$ eine symmetrische 2×2 Matrix mit verschwindender Spur ist.

Die Verbindung zwischen innerer Geometrie und Einbettung in den $\mathbb{R}^p$ wird durch die Gaußsche Krümmungsgleichung [3] (6.3, 17) hergestellt:

$$E^2 K = R_{1212} = \sum_{\lambda=1}^{p-2} (h_{22}^\lambda h_{11}^\lambda - h_{21}^\lambda h_{12}^\lambda) = -\sum_{\lambda=1}^{p-2} \{(h_{11}^\lambda)^2 + (h_{12}^\lambda)^2\}.$$

Mit (23) und Definition 2 folgt

$$E^2K(u^1,u^2)=-|x_{u^1u^1}{}^*||x_{u^2u^2}{}^*|-|x_{u^1u^2}{}^*|^2. \tag{25}$$

Man beachte $x_{u^1u^1}{}^*=(x_{u^1u^1})^*$ etc.

Literatur

1. Barbosa, J.L., do Carmo, M.: Stability of Minimal Surfaces and Eigenvalues of the Laplacian. Math. Z. **173**, 13–28 (1980)
2. Blaschke, W.: Vorlesungen über Differentialgeometrie, I. 4. Aufl. Berlin Heidelberg New York: Springer 1945
3. Brauner, H.: Differentialgeometrie. Braunschweig: Vieweg 1981
4. Courant, R.: Dirichlets principle, conformal mappings and minimal surfaces. (Reprint). Berlin-Heidelberg-New York: Springer 1977
5. Eisenhart, L.P.: Riemannian Geometry. 5. Aufl. Princeton: University Press 1964
6. Heinz, E.: Über die analytische Abhängigkeit der Lösungen eines linearen elliptischen Randwertproblems von Parametern. Nach. Akad. Wiss. Göttingen Math.-Phys. Kl. II (1979), Heft 1
7. Heinz, E.: Über eine Verallgemeinerung des Plateau'schen Problems. Manuscripta Math. **28**, 81–88 (1979)
8. Heinz, E.: Minimalflächen mit polygonalem Rand. Math. Z. **183**, 547–564 (1983)
9. Hicks, N.J.: Notes on Differential Geometry. Math. Studies 3. London; van Nostrand Reinhold Co. 1971
10. Lawson, H.B., jr.: Lectures on Minimal Submanifolds, I. Math. Lecture series 9, Berkeley, Ca., (1980)
11. Marx, I.: On the classification of unstable minimal surfaces with polygonal boundaries. Comm. Pure Appl. Math. **8**, 235–244 (1955)
12. Nitsche, J.C.C.: Vorlesungen über Minimalflächen. Berlin, Heidelberg, New York: Springer 1975
13. Sasaki, S.: On the total curvature of a closed curve. Japan. J. Math. **29**, 118–125 (1959)
14. Sauvigny, F.: Flächen vorgeschriebener mittlerer Krümmung mit eineindeutiger Projektion auf eine Ebene. Math. Z. **180**, 41–67 (1982)
15. Schwarz, H.A.: Gesammelte Math. Abhandlungen, I. Berlin: Springer 1890
16. Spruck, J.: Remarks on the stability of minimal submanifolds of $\mathbb{R}^n$. Math. Z. **114**, 169–174 (1975)

Eingegangen am 20. Juni 1984

Math. Z. 189, 185–192 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

The Isometry Groups of Manifolds of Nonpositive Curvature with Finite Volume

Takao Yamaguchi

Department of Mathematics (Faculty of Science and Engineering), Saga University, Saga 840, Japan

§ 0. Introduction

Let M be a complete Riemannian manifold with finite volume of nonpositive sectional curvature and Ricci tensor negative definite at some point of M. It is well known that the group $I(M)$ of isometries of M is finite ([2, 15]). It is natural to ask on which geometrical quantities the order $\mathrm{Ord}\, I(M)$ of $I(M)$ depend. In the case where M is compact, various estimates on $\mathrm{Ord}\, I(M)$ have been obtained by using the diameter, the volume or the injectivity radius of exponential map on M etc. ([1, 13, 14, 17]). The main purpose of this paper is to give some estimates on $\mathrm{Ord}\, I(M)$ for noncompact manifolds under an additional condition.

Suppose further that M is noncompact and has finite topological type, i.e., M is homeomorphic to the interior of a compact manifold with boundary. Then M has a bounded open submanifold M' which is a deformation retract of M. This assumption is automatically satisfied in the case where M has negative and bounded sectional curvature ([10]), or M is a visiblity manifold with bounded sectional curvature ([7]). Our starting point is to extend a result of Frankel [8] which states that if an isometry of a compact manifold with negative sectional curvature is homotopic to the identity, then it is the identity. This is done in Sect. 1. After choosing M' suitably, in Sect. 2 the order of $I(M)$ is estimated in terms of the geometrical quantities of M', that is the diameter of M', the injectivity radius of M' and the lower bound of the sectional curvature on M'. In Sect. 3, we shall attempt another approach. In Sect. 4, we shall also consider the special case that M has negative and bounded sectional curvature.

The author is very grateful to N. Innami for his valuable advice.

§ 1. Identity Criteria

From now on, let M be a complete Riemannian manifold with nonpositive sectional curvature and finite volume such that the universal covering space H

of M has no Euclidean factor in the de Rham decomposition, unless otherwise stated. We denote by Γ the deck transformation group. We now give some sufficient conditions under which an isometry will be the identity map.

Lemma 1.1. *If an isometry φ of H commutes with all elements of Γ, then φ is the identity map.*

Proof. By the hypothesis on φ, the convex function $\phi: H \to R$ given by $\phi(p) = d^2(p, \varphi p)$ is Γ-invariant. Therefore ϕ induces a convex function on H/Γ which is also denoted by ϕ. By [4, Proposition 2.2] ϕ must be constant since $M = H/\Gamma$ has finite volume. Hence it turns out from [19, Theorem 1] that φ must be the identity map.

The following theorem is an extended version of Frankel [8] and Lawson-Yau [15].

Theorem 1. *Suppose that M is noncompact and has finite topological type. Let M' be a bounded open submanifold which is a deformation retract of M. If an isometry f of M is continuously homotopic to the identity on M', then it is the identity.*

Proof. We proceed along the same line as in [8]. Let $h_t: M' \to M$ be a homotopy such that $h_0 = \mathrm{id}_{M'}$ and $h_1 = f|_{M'}$. Let H be the universal covering space of M. For every point q of M', the deck transformation group Γ is identified with the fundamental group $\pi_1(M, q)$ of M, and every element of $\pi_1(M, q)$ is represented by a loop in M' since M' is a deformation retract of M. From this, we see that $\pi^{-1}(M')$ is arcwise connected, where $\pi: H \to M$ is the covering map. Set $H' = \pi^{-1}(M')$. Let $\tilde{h}_t: H' \to H$ be the covering of h_t with $\pi \circ \tilde{h}_t = h_t \circ \pi$ and $\tilde{h}_0 = \mathrm{id}_{H'}$. Fix a point p of M' and a point x of H over p. Let $\tilde{f} \in I(H)$ be the unique covering isometry of f with $\tilde{f}(x) = \tilde{h}_1(x)$. We show that the restriction of $\tilde{f}$ to H' is equal to $\tilde{h}_1$. For every point y of H', let c be a curve in H' from x to y. Clearly two curves $\tilde{f} \circ c$ and $\tilde{h}_1 \circ c$ have the same initial point $\tilde{f}(x)$ and project to $f \circ \pi \circ c$. It follows from the uniqueness of lift that $\tilde{f} \circ c = \tilde{h}_1 \circ c$, and hence $\tilde{f}(y) = \tilde{h}_1(y)$. From this form of $\tilde{f}$, we see that $\tilde{f}$ commutes with all elements of Γ on H', and hence on H. Thus Lemma 1.1 implies $\tilde{f} = \mathrm{id}_H$ and $f = \mathrm{id}_M$.

Corollary 1.2. *Let M and M' be as in Theorem 1. If an isometry f of M satisfies that $d(f(p), p) < i_M(p)$ for all points p of M', then f is the identity.*

We denote by i_M the injectivity radius function on M.

Proof. The unique minimal geodesic from p to $f(p)$ provides the homotopy $h_t: M' \to M$ between $\mathrm{id}_{M'}$ and $f|_{M'}$.

Using Corollary 1.2, we give a characterization of the identity transformation. Recall that if an isometry f of a Riemannian manifold fixes a point p and if the differential df_p at p fixes all the tangent vectors at p, then f is the identity. For given two points p and q, we denote by $P_{p,q}$ the parallel translation along a minimal geodesic from p to q. We consider the distance d^* on the group of orthogonal transformations of a Euclidean vector space V. By definition,

$$d^*(A, B) = \sup \{ \measuredangle (Ax, Bx); x \in V, |x| = 1 \},$$

where $\measuredangle(x, y)$ denotes the angle between x and y. It should be noted that the metric d^* is Finsler rather than Riemannian. For a discussion of these Finsler metric, see Buser-Karcher [5, §7.3].

Lemma 1.3. *Given Λ, d, $r>0$, set*

$$\delta_0 = r/(2\cosh \Lambda d), \quad \delta_1 = \Lambda r/(2\sinh \Lambda d).$$

Let M be a complete manifold with $K \leqq 0$. For a point $p \in M$, let $-\Lambda^2$ be a lower bound of K on the ball $B(p, d+r/2)$ of radius $d+r/2$ around p. Then if an isometry f of M satisfies $d(p, f(p)) < \delta_0$ and $d^*(P_{f(p),p} \circ df_p, \mathrm{id}) < \delta_1$ for a minimal geodesic γ from p to $f(p)$, then $d(f(p), q) < r$ for every point q of $B(p, d)$.

Proof. For every point q of $B(p, d)$, let $\sigma\colon [0, l] \to M$ be a minimal geodesic from p to q with $l = d(p, q)$. We denote by P the parallel translation along γ. We set $v = \dot{\sigma}(0)$ and $q' = \exp_{f(p)} lP(v)$. Notice that for every non zero tangent vector w, the geodesic hypersurface consisting of geodesics perpendicular to w has no focal points along the geodesic of direction w since the sectional curvature is nonpositive. It follows from the Berger comparison theorem (Cf. [6, p. 31]) that

$$d(q, q') \leqq d(p, f(p)) \cosh \Lambda l \leqq \delta_0 \cosh \Lambda d = r/2.$$

On the other hand, the Rauch comparison theorem implies

$$\begin{aligned} \Lambda d(q', f(q)) &\leqq \cosh^{-1}(\cosh^2 \Lambda l - \sinh^2 \Lambda l \cos \measuredangle(P(v), df_p(v)) \\ &\leqq \cosh^{-1}(\cosh^2 \Lambda d - \sinh^2 \Lambda d \cos \delta_1) \leqq \Lambda r/2. \end{aligned}$$

Hence we have $d(q, f(q)) \leqq d(q, q') + d(q', f(q)) \leqq r$.

From Corollary 1.2 and Lemma 1.3, we get immediately the following

Proposition 1.4. *Let M and M' be as in Theorem 1. We set d = the diameter of M', r = the infimum of i_M on M' and $-\Lambda^2$ = the infimum of K on the $r/2$-neighborhood of M'. Let δ_0 and δ_1 be the constants given in Lemma 1.3. If an isometry f of M satisfies $d(p, f(p)) < \delta_0$ and $d^*(P_{f(p),p} \circ df_p, \mathrm{id}) < \delta_1$ for some point p of M', then f is the identity.*

§2. Main Estimate

Throughout this section, suppose further that M is noncompact and has finite topological type. Let M' be a bounded open submanifold which is a deformation retract of M. Since M has finite volume and $K \leqq 0$, the standard volume comparison argument shows that $i_M(p)$ goes to zero as p tends to infinity. Therefore we may take M' so that

$$\sup_{M'} i_M > \sup_{M - M'} i_M.$$

From now on, let M' be chosen so as to satisfy this inequality. We set

$$n = \dim M, \quad d = \operatorname{diam}(M'), \quad r = \inf_{M'} i_M,$$

$-\Lambda^2$ = the infimum of K on the d-neighborhood of M'.

Using these quantities, the order $\mathrm{Ord}\, I(M)$ is estimated by the following

Theorem 2. $\mathrm{Ord}\, I(M) < (8e^{2\Lambda d}/\Lambda r)^{n^2}$.

Let p be a fixed point of M' with $i_M(p) > \sup_{M-M'} i_M$, and let x be a point of the universal covering space H over p. Let O_x denote the group of orthogonal transformations of the tangent space H_x to H at x. We define a distance $\bar{d}$ on the product $H \times O_x$ by

$$\bar{d}((y_1, A_1), (y_2, A_2)) = \max\{d(y_1, y_2), \frac{1}{\Lambda} d^*(A_1, A_2)\}, (y_i, A_i) \in H \times O_x.$$

Let $\mathscr{V}$ denote the volume measure on $H \times O_x$ induced from the Riemannian product metric of H and the standard bi-invariant metric on O_x. Then for every positive ε and for every point (y, A) of $H \times O_x$ we have

$$\begin{aligned} B((y, A), \varepsilon) &= B(y, \varepsilon) \times B(A, \Lambda\varepsilon) \\ \mathscr{V}(B((y, A), \varepsilon)) &= \mathrm{Vol}\,(B(y, \varepsilon))\, \mathrm{Vol}\,(B(A, \Lambda\varepsilon)) \\ &= \mathrm{Vol}\,(B(y, \varepsilon))\, v(\Lambda\varepsilon), \end{aligned}$$

where $v(\Lambda\varepsilon)$ denotes the volume of a $\Lambda\varepsilon$-ball in O_x with respect to the distance d^* defined after Corollary 1.2. Since each isometry f of M maps the point p to a point of M', we may choose a covering isometry $\tilde{f} \in I(H)$ of f such that $d(\tilde{f}(x), x) \leqq d$. We consider a map Φ of the isometry group $I(M)$ to $H \times O_x$ defined by

$$\Phi(f) = (\tilde{f}(x), P_{\tilde{f}(x), x} \circ d\tilde{f}_x).$$

We set $\delta = \min\{\delta_0, \delta_1/(\Lambda(1 + \Lambda d))\}$, where δ_i $(i = 1, 2)$ are the constants given in Lemma 1.3.

Lemma 2.1. *$\bar{d}(\Phi(f), \Phi(g)) < \delta$ implies $f = g$.*

Proof. Set $h = f^{-1}g$ and $\tilde{h} = \tilde{f}^{-1}\tilde{g}$. We apply Lemma 1.3 to $\tilde{h}$. We have easily $d(\tilde{h}(x), x) < \delta \leqq \delta_0$. The fact that the image of a parallel field under an isometry is again a parallel field implies

$$\begin{aligned} &d^*(\mathrm{id}, P_{\tilde{f}^{-1}(x), x} \circ P_{\tilde{h}(x), \tilde{f}^{-1}(x)} \circ d\tilde{h}_x) \\ &\quad = d^*(P_{\tilde{f}(x), x} \circ d\tilde{f}_x, P_{\tilde{g}(x), x} \circ d\tilde{g}_x) < \Lambda\delta. \end{aligned}$$

Let $\sigma: [0, 1] \to H$ be the geodesic from $\tilde{h}(x)$ to x. We consider the geodesic variation $\alpha: [0, 1] \times [0, 1] \to H$ defined by $\alpha(t, s) = \exp t(\exp_{\tilde{f}^{-1}(x)}{}^{-1}(\sigma(s)))$. The norm $\|J_s(t)\|$ of the Jacobi field $J_s(t) = \dfrac{\partial\alpha}{\partial s}(t, s)$ is convex in t since the sectional curvature K is nonpositive. It follows that $\|J_s(t)\| \leqq t\|J_s(1)\| \leqq \delta t$ for every $0 \leqq t \leqq 1$. We also observe

$$\begin{aligned} \left\| \frac{\partial\alpha}{\partial t}(t, s) \right\| &= d(\tilde{f}^{-1}(x), \sigma(s)) \\ &\leqq \max\{d(\tilde{f}^{-1}(x), \tilde{h}(x)), d(\tilde{f}^{-1}(x), x)\} < d. \end{aligned}$$

Hence we have the following estimate for the area F of α:

$$F=\int_0^1\int_0^1\left\|\frac{\partial\alpha}{\partial t}\wedge\frac{\partial\alpha}{\partial s}\right\|dt\,ds\leqq\int_0^1\int_0^1 d\delta t\,dt\,ds=d\delta/2.$$

Since the path dependence of parallel translations is controlled by the sectional curvature and the area of the homotopy. (See [5, §6.2] for detail), we get

$$d^*(P_{\tilde f^{-1}(x),x}\circ P_{h(x),\tilde f^{-1}(x)},P_{h(x),x})\leqq\tfrac{4}{3}\Lambda^2F\leqq\Lambda^2d\delta.$$

The above two inequalities yield $d^*(\mathrm{id},P_{h(x),x}\circ d\tilde h_x)<\Lambda\delta+\Lambda^2d\delta\leqq\delta_1$. Then Lemma 1.3 implies $d(\tilde h,\mathrm{id})<r$ on $B(x,d)$, and hence $d(h,\mathrm{id})<r$ on M' since the covering map is distance nonincreasing. Thus from Corollary 1.2 we conclude $h=\mathrm{id}$ and $f=g$.

Proof of Theorem 2. Let $m=\mathrm{Ord}\,I(M)$ and let $f_1,\dots,f_m$ be the pairwise distinct isometries of M. By Lemma 2.1, the balls $B(\Phi(f_i),\delta/2)$ are disjoint and contained in $B(x,d+\delta/2)\times O_x$. Hence

$$\begin{aligned}m&\leqq\mathscr{V}(B(x,d+\delta/2)\times O_x)/\mathscr{V}(B(\Phi(f_i),\delta/2)\\&=(\mathrm{Vol}\,(B(x,d+\delta/2))\,v(\pi))/(\mathrm{Vol}\,(B(x,\delta/2))\,v(\Lambda\delta/2)).\end{aligned}$$

According to the volume comparison theorem due to Bishop (Cf. [3]) and Gromov [12],

$$\mathrm{Vol}\,(B(x,d+\delta/2))/\mathrm{Vol}\,(B(x,\delta/2))\leqq s_n(\Lambda,d+\delta/2)/s_n(\Lambda,\delta/2),$$

where $s_n(\Lambda,t)=\int_0^t\sinh^{n-1}\Lambda s\,ds$, and by [5, §7.6],

$$v(\pi)/v(\Lambda\delta/2)\leqq2(2\pi/\Lambda\delta)^{\dim O_x}.$$

Therefore we get the estimate for $\mathrm{Ord}\,I(M)$ in terms of n, Λ, r and d: More precisely, we have

$$\begin{aligned}s_n(\Lambda,d+\delta/2)/s_n(\Lambda,\delta/2)&\leqq\int_0^{d+\delta/2}e^{(n-1)\Lambda s}ds\Big/\int_0^{\delta/2}(\Lambda s)^{n-1}ds\\&=\frac{n}{n-1}[(e^{(n-1)\Lambda(d+\delta/2)}-1)/(\Lambda\delta/2)^n]\\&\leqq e^{n\Lambda d}/(\Lambda\delta/2)^n=(2e^{\Lambda d}/\Lambda\delta)^n\end{aligned}$$

and

$$\delta=\frac{r}{2}\min\{(\cosh\Lambda d)^{-1},((1+\Lambda d)\sinh\Lambda d)^{-1}\}\geqq r(2(1+\Lambda d)\,e^{\Lambda d})^{-1},$$

from which we estimate

$$\begin{aligned}m&\leqq2(2\pi/\Lambda\delta)^{n(n-1)/2}(2e^{\Lambda d}/\Lambda\delta)^n\leqq(2\pi)^{(n^2+1)/2}\,e^{n\Lambda d}/(\Lambda\delta)^{n(n+1)/2}\\&\leqq(2\pi)^{(n^2+1)/2}\,e^{n\Lambda d}(2(1+\Lambda d)\,e^{\Lambda d}/(\Lambda r))^{n(n+1)/2}\\&\leqq8^{n^2}e^{2n^2\Lambda d}/(\Lambda r)^{n^2}.\end{aligned}$$

Remark. From the above proof, we see that in the case where M is compact, the same estimate for $\operatorname{Ord} I(M)$ is still valid if we set $M' := M$.

§ 3. Another Approach

In this section, we apply a method of Lemaire [16] and Adachi-Sunada [1] to our situation. Let M and M' be as in Sect. 2. We set

$$d = \operatorname{diam}(M'), \quad V = \operatorname{Vol}(M'),$$

$-\Lambda^2 =$ the infimum of K on the $5d$-neighborhood of M'.

Theorem 3. *The order* $\operatorname{Ord} I(M)$ *is estimated explicitly in terms of* n, d, V *and* Λ, *e.g.,*

$$\operatorname{Ord} I(M) < (c_1 e^{5n\Lambda d}/(V\Lambda^n))^{(c_1 e^{3n\Lambda d}/(V\Lambda^n))},$$

where c_1 *depends only on the dimension* n.

Let p be a fixed point of M' such that $i_M(p) > \sup_{M-M'} i_M$, and let x be a point of the universal covering space H over p.

Lemma 3.1. *The deck transformation group* Γ *is generated by those elements* γ_i *that* $d(x, \gamma_i x) \leqq 2d$.

Proof. Every loop at p is homotopic to a loop $c: [0,1] \to M'$ in M'. For a positive integer N and for each integer i with $0 \leqq i \leqq N$, set $t_i = i/N$ and choose a minimal geodesic σ_i from p to $c(t_i)$. Let c_i denote the product loop $\sigma_{i+1}^{-1} \circ c \,|\, [t_i, t_{i+1}] \circ \sigma_i$. Then c is homotopic to the product of c_i. Since the set of length of all geodesic loops at p is discrete, the geodesic loops homotopic to c_i are of length not greater than $2d$ for a sufficiently large N.

Proof of Theorem 3. For each isometry f of M, let $\tilde{f}$ be a covering isometry such that $d(x, \tilde{f}(x)) \leqq d$. Define an automorphism ρ_f of Γ by $\tilde{f} \circ \gamma = \rho_f(\gamma) \circ \tilde{f}$, $\gamma \in \Gamma$. We consider the map ρ of $I(M)$ into the group of automorphisms of Γ given by $\rho(f) = \rho_f$. By Lemma 1.1, ρ is one to one. Let T be the set of generators of Γ obtained in Lemma 3.1. Set

$$W = \bigcup_{f \in I(M)} \tilde{f}(B(x, 2d)), \qquad S = \{\gamma \in \Gamma;\ \gamma B(x, d) \cap W \neq \emptyset\}.$$

Since W is bounded, the discreteness of Γ implies that S is finite. By the construction, ρ_f carries T into S for every $f \in I(M)$. Since ρ_f is uniquely determined by the values on T, we get the following estimate:

$$\operatorname{Ord} I(M) \leqq (\# S)^{(\# T)}.$$

Let U denote the intersection of $\pi^{-1}(M')$ and the fundamental domain of Γ about x, where π is the covering map. If $\gamma_i \in T$, $\gamma_i U \subset B(x, 3d)$. This implies

$$\# T \leqq \operatorname{Vol}(B(x, 3d))/\operatorname{Vol}(U) = \operatorname{Vol}(B(x, 3d))/V.$$

On the other hand, every element γ of S satisfies $d(x, \gamma x) \leqq 4d$ since W is contained in $B(x, 3d)$. Hence the same argument applies to get the following bound:

$$\#S \leqq \mathrm{Vol}(B(x, 5d))/V.$$

By the result of Bishop, $\mathrm{Vol}(B(x,s)) \leqq \omega_{n-1} s_n(\Lambda, s)/\Lambda^{n-1}$ for all $s \in (0, 5d)$, where ω_{n-1} denotes the volume of unit $(n-1)$-sphere. Therefore we have the required estimate for $\mathrm{Ord}\, I(M)$ in terms of n, d, V and Λ.

§ 4. Remark for the Case of Negative Curvature

Let M be a complete noncompact n-manifold with $-\Lambda^2 \leqq K \leqq 0$. Suppose further that $K<0$ or M is a visibility manifold (see [7] for the definition). It is known in [7] and [10] that M has finite topological type. Let $\varepsilon_1 = \varepsilon_1(n, \Lambda) > 0$ be the Margulis constant. This means that for every $0 < \varepsilon \leqq \varepsilon_1$ and every point p of M, the subgroup of $\pi_1(M, p)$ generated by those geodesic loops at p whose length is not greater than ε has a nilpotent subgroup of finite index. The constant ε_1 can be written as $\varepsilon_1 = \mathrm{const}_n \Lambda^{-1}$. For detail see [5, 10, 18]. We set $M_\varepsilon = i_M^{-1}((0, \varepsilon])$ and denote by M' the complement of all unbounded connected components of $M_{\varepsilon_1/2}$. M' satisfies all the conditions in Sect. 2 (Cf. [9]). In particular, it is known that $M - M_{\varepsilon_1}$ is non-empty. As is easily checked, the assumption $K<0$ implies that i_M is Lipschitz continuous with Lipschitz constant 1. It follows from the volume comparison theorem that $V = \mathrm{Vol}(M')$ is not smaller than the volume of an $\varepsilon_1/2$-ball in Euclidean n-space. Therefore we have from Theorem 3

$$\begin{aligned}
\mathrm{Ord}\, I(M) &\leqq (c_2 e^{5n\Lambda d}/(\Lambda \varepsilon_1)^n)(c_2 e^{3n\Lambda d}/(\Lambda \varepsilon_1)^n) \\
&\leqq (e^{5n\Lambda d + c_3})(e^{3n\Lambda d + c_3}) \\
&\leqq e^{(5n\Lambda d + c_3)(e^{3n\Lambda d + c_3})} \\
&\leqq e^{e^{3n\Lambda d + c_4}}
\end{aligned}$$

where c_i depend only on n.

The following should be pointed out. According to Gromov [10], for a simply connected complete n-manifold H with $-\Lambda^2 \leqq K < 0$, the volume $v(\Gamma)$ of the fundamental domain of any (not necessarily freely acting) discrete subgroup Γ of $I(H)$ has a positive lower bound $C = C(n, \Lambda)$ depending only on n and Λ. Let Γ be a freely acting properly discontinuous subgroup of $I(H)$ such that the volume of $M = H/\Gamma$ is finite. Let Γ^* denote the normalizer of Γ in $I(H)$. Then $I(M)$ is isomorphic as a group to Γ^*/Γ by means of the map $f \mapsto \tilde{f} \cdot \Gamma$, where $f \in I(M)$ and $\tilde{f} \in I(H)$ is a covering isometry of f. It follows that Γ^* is discrete and the order of $I(M)$ is equal to $v(\Gamma)/v(\Gamma^*)$. Together with the above result of Gromov, this implies that $\mathrm{Ord}\, I(M) \leqq \mathrm{Vol}(M)/C$. In our situation, the order $\mathrm{Ord}\, I(M)$ is estimated by the data on the bounded domain M'.

References

1. Adachi, T., Sunada, T.: Energy spectrum of certain harmonic mappings, Preprint, Nagoya University
2. Averous, G., Kobayashi, S.: On automorphisms of spaces of nonpositive curvature with finite volume, Differential geometry and Relativity. Dortrecht: D. Reidel Publishing Co., 19-26 (1976)
3. Bishop, R.L., Crittenden, R.J.: Geometry of manifolds. New York: Academic Press 1964
4. Bishop, R.L., O'Neill, B.: Manifolds of negative curvature. Trans. Am. Math. Soc., **145**, 1-49 (1969)
5. Buser, P., Karcher, H.: Gromov's almost flat manifolds, Astérisque **81** (1981)
6. Cheeger, J., Ebin, D.: Comparison theorems in Riemannian geometry. Amsterdam: North-Holland 1975
7. Eberlein, P.: Lattices in spaces of nonpositive curvature, Ann. Math. **111**, 435-476 (1980)
8. Frankel, T.: On theorems of Hurwitz and Bochner, J. Math. Mech. **15**, 373-377 (1966)
9. Fukaya, K.: A finiteness theorem for negatively curved manifolds. Preprint, Tokyo University
10. Gromov, M.: Manifolds of negative curvature. J. Differ. Geom. **13**, 223-230 (1978)
11. Gromov, M.: Structures metriques pour les variétés riemanniennes, rédigé par J. Lafontaine et P. Pansu. Paris: Cedic-Fernand Nathan 1981
12. Heintze, E.: Mannigfaltigkeiten negativer Krümmung. Preprint, Bonn, 1976
13. Huber, H.: Über die Isometriegruppe einer kompakten Mannigfaltigkeit mit negativer Krümmung. Helv. Phys. Acta **45**, 277-288 (1972)
14. Im Hof, H.C.: Über die Isometriegruppe bei kompakten Mannigfaltigkeiten negative Krümmung. Comment. Math. Helv., **48**, 14-30 (1973)
15. Lawson, H.B., Yau, S.T.: Compact manifolds of nonpositive curvature. J. Differ. Geom. **7**, 211-288 (1972)
16. Lemaire, L.: Harmonic mappings of uniform bounded dilatation. Topology **16**, 199-201 (1977)
17. Maeda, M.: The isometry groups of compact manifolds with nonpositive curvature. Proc. Japan Acad., Ser. A **51**, 790-794 (1975)
18. Thurston, W.: The geometry and topology of 3-manifolds. Lecture notes, Princeton University, 1978
19. Wolf, J.A.: Homogeneity and bounded isometries in manifolds of negative curvature. Ill. J. Math. **8**, 14-18 (1964)

Received December 5, 1983, received in final form September 19, 1984

Math. Z. 189, 193–210 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Compact Hermitian Manifolds of Constant Holomorphic Sectional Curvature

Andrew Balas

Mathematics Department, Hill Center, Rutgers University, New Brunswick, NJ 08903, USA

§ 0. Introduction

Although compact Kähler manifolds of constant holomorphic sectional curvature have been classified, little is known of the more general Hermitian case. S. Kobayshi has shown that a compact Hermitian manifold with negative holomorphic sectional curvature bounded away from zero, is hyperbolic in the sense of Kobayashi (Kobayashi [5]), which implies that every holomorphic function from the complex plane to the manifold is constant. But few other results are known. This paper is an attempt to extend our knowledge of compact Hermitian manifolds of constant holomorphic sectional curvature.

The first question that arises is whether there are any such manifolds that are not Kähler. We answer this in the affirmative by giving examples of compact non-Kähler manifolds of vanishing curvature, in every dimension above 2 (Proposition (3.1)). These manifolds are of the form G/Γ, where G is a non-abelian Lie group, and Γ is a discrete subgroup. It is not known whether there are such examples where the holomorphic sectional curvature is a non-zero constant.

It is easy to show that a Hermitian manifold of constant holomorphic sectional curvature has $s+\hat{s}$ constant, where s and $\hat{s}$ are the two sectional curvatures of Hermitian geometry (Corollary (2.4)). Using the methods of S. Kobayashi, H. Wu, and P. Gauduchon, we show that a compact Hermitian manifold cannot have $s+\hat{s}$ equal to a nonnegative constant unless its plurigenera meet certain conditions. This leads to the main result of this paper, Corollary (3.5): Let M be a compact Hermitian manifold of constant holomorphic sectional curvature $=k$. Let P_m be the mth plurigenus and Q_m be the mth dual plurigenus of M. Then

a) $k>0 \Rightarrow P_m=0$, $\forall_m>0$;

b) $k=0 \Rightarrow$ either $P_m=0$, $\forall m>0$, or $P_m=Q_m$, $\forall m>0$, and $P_m\in\{0,1\}$.

We obtain sharper results if we assume the metric is conformal to a balanced metric, as well as $s+\hat{s}$ being constant (Theorem (3.3)). (We say a metric is balanced if its contracted torsion tensor vanishes.)

Using the methods of S.I. Goldberg, we extend one of his results to show that a Hermitian manifold (compact or not) with holomorphic torsion tensor and constant holomorphic sectional curvature has vanishing curvature if the holomorphic sectional curvature is identically zero, and is Kähler otherwise (Theorem (3.2)). It should be noted, however, that holomorphic torsion is quite a strong assumption.

§1 develops the results needed from Hermitian geometry. §2 works with Hermitian manifolds of constant holomorphic sectional curvature. §3 presents the main results of the paper. §4 is a short section on Einstein conditions.

I wish to thank Professor Gauduchon for his many helpful suggestions on this paper.

§1. Hermitian Manifolds

Let M be a connected complex manifold of complex dimension n, with complex structure J, and Hermitian metric h. We define the Hermitian connection ∇ on (M, J, h) by requiring that J and h be parallel with respect to ∇ and that the $(1,1)$-part of the torsion vanish. (This connection is also called the "canonical connection" and the "Chern connection.") The quadruple (M, J, h, ∇), where ∇ is the Hermitian connection, is called a Hermitian manifold. For the rest of this paper M will always denote a connected Hermitian manifold. We define the curvature tensors $R(\cdot,\cdot)$ and $R(\cdot,\cdot,\cdot,\cdot)$ by

$$R(X, Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z$$

and

$$R(X, Y, Z, W) = h(R(Z, W)Y, X),$$

respectively, for vector fields X, Y, Z and W. We define the torsion tensor by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

We define $R^{\beta}_{\alpha\gamma\bar{\delta}}$, $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ and $T^{\gamma}_{\alpha\beta}$ in local coordinates $z^1, \ldots, z^n$ by

$$R(Z_\gamma, \bar{Z}_\delta)Z_\alpha = \sum R^{\beta}_{\alpha\gamma\bar{\delta}} Z_\beta,$$

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} = R(Z_\alpha, \bar{Z}_\beta, Z_\gamma, \bar{Z}_\delta),$$

and

$$T(Z_\alpha, Z_\beta) = \sum T^{\gamma}_{\alpha\beta} Z_\gamma,$$

respectively. Then we have (Kobayashi and Nomizu [6], Chapt. IX)

$$\Gamma^{\beta}_{\gamma\alpha} = \sum h^{\beta\bar{\delta}} \frac{\partial h_{\alpha\bar{\delta}}}{\partial z^\gamma}, \tag{1}$$

$$R^{\beta}_{\alpha\gamma\bar{\delta}} = -\frac{\partial \Gamma^{\beta}_{\gamma\alpha}}{\partial \bar{z}^\delta}, \tag{2}$$

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} = -\sum h_{\varepsilon\bar{\beta}} R^{\varepsilon}_{\alpha\gamma\bar{\delta}}, \tag{3}$$

and

$$T^{\gamma}_{\alpha\beta} = \Gamma^{\gamma}_{\alpha\beta} - \Gamma^{\gamma}_{\beta\alpha}. \tag{4}$$

The metric h is called Kähler if the torsion vanishes. This is equivalent to:

$$d\Phi = 0, \tag{5}$$

where $\Phi = i\sum h_{\alpha\bar{\beta}} dz^\alpha \wedge d\bar{z}^\beta$ is the Kähler form. We define the contracted torsion tensor $\tau_{\mathbf{R}}$ by

$$\tau_{\mathbf{R}} = \sum (\tau_\alpha dz^\alpha + \bar{\tau}_\alpha d\bar{z}^\alpha), \tag{6}$$

where $\tau_\alpha = \sum T^\gamma_{\alpha\gamma}$. We will denote the $(1,0)$-part of $\tau_{\mathbf{R}}$ by τ. The metric h is called balanced if τ vanishes. (We also say T and M are balanced in this case.) Elsewhere in the literature this condition is called "semi-Kähler."

From the identity

$$\tau_{\mathbf{R}} = J\delta\Phi$$

we obtain the following:

(1.1) *Remark* (Gauduchon). A compact Hermitian manifold is balanced if and only if either of the two following equivalent conditions holds:

a) $d\Phi^{n-1} = 0$,

b) $\delta\Phi = 0$. □

It follows from condition a) that a balanced compact Hermitian surface is Kähler.

When h is Kähler, the components $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ of R have the Kähler symmetries:

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} = R_{\gamma\bar{\beta}\alpha\bar{\delta}} = R_{\alpha\bar{\delta}\gamma\bar{\beta}} = R_{\gamma\bar{\delta}\alpha\bar{\beta}}.$$

These symmetries do not in general hold in the Hermitian case (however, see (7) below for an identify satisfied even this case) and we have four different Ricci-type contractions of the curvature tensor. Define $R_{\alpha\bar{\beta}}$, $S_{\alpha\bar{\beta}}$, $\hat{R}_{\alpha\bar{\beta}}$, $\hat{S}_{\alpha\bar{\beta}}$ by

$$R_{\alpha\bar{\beta}} = \sum R^\gamma_{\gamma\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} R_{\gamma\bar{\delta}\alpha\bar{\beta}}$$

$$S_{\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} R_{\alpha\bar{\beta}\gamma\bar{\delta}}$$

$$\hat{R}_{\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} R_{\alpha\bar{\delta}\gamma\bar{\beta}}$$

$$\hat{S}_{\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} R_{\gamma\bar{\beta}\alpha\bar{\delta}}.$$

We can represent R locally as a matrix Ω of $(1,1)$-forms.

$$\Omega = (\Omega^\alpha_\beta), \qquad \Omega^\alpha_\beta = \sum R^\alpha_{\beta\gamma\bar{\delta}} dz^\gamma \wedge d\bar{z}^\delta.$$

Then the Ricci form $-2i\sum R_{\alpha\bar{\beta}} dz^\alpha \wedge d\bar{z}^\beta$ is a constant multiple of the trace of this matrix. Thus the Ricci form is closed.

If we let $H = (h_{\alpha\bar{\beta}})$, then

$$\Omega H = (\Omega_{\alpha\bar{\beta}}), \qquad \Omega_{\alpha\bar{\beta}} = \sum R_{\alpha\bar{\beta}\gamma\bar{\delta}} dz^\gamma \wedge dz^\delta$$

is skew-Hermitian. As a consequence,

$$\bar{R}_{\alpha\bar{\beta}\gamma\bar{\delta}} = R_{\beta\bar{\alpha}\delta\bar{\gamma}}. \tag{7}$$

From (7) we see that

$$\bar{\hat{R}}_{\alpha\bar{\beta}} = \hat{S}_{\beta\bar{\alpha}}. \tag{8}$$

There are two scalar curvatures in Hermitian geometry. We define them by

$$s = -\sum h^{\alpha\bar{\beta}} h^{\gamma\bar{\delta}} R_{\alpha\bar{\beta}\gamma\bar{\delta}}$$

and

$$\hat{s} = -\sum h^{\alpha\bar{\delta}} h^{\gamma\bar{\beta}} R_{\alpha\bar{\beta}\gamma\bar{\delta}}.$$

For the Hermitian connection, only $\Gamma^{\gamma}_{\alpha\beta}$ and $\Gamma^{\bar{\gamma}}_{\bar{\alpha}\bar{\beta}}$ are nonzero: Christoffel symbols of mixed parity with respect to conjugation vanish. Thus, if $K^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s}$ are the local components of a tensor K, then

$$K^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s|\bar{\delta}} = \frac{\partial}{\partial \bar{z}^{\delta}} K^{\alpha_1\ldots\alpha_r}_{\beta_1\ldots\beta_s},$$

where the solidus denotes covariant differentiation with respect to the Hermitian connection.

Thus, the first Bianchi identity

$$R^{\varepsilon}_{\gamma\alpha\bar{\delta}} - R^{\varepsilon}_{\alpha\gamma\bar{\delta}} = T^{\varepsilon}_{\gamma\alpha|\bar{\delta}} \tag{9}$$

shows that the derivatives of the torsion components measure the lack of symmetry of the curvature components. Lowering indices, we have

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} - R_{\gamma\bar{\beta}\alpha\bar{\delta}} = \sum h_{\varepsilon\bar{\beta}} T^{\varepsilon}_{\gamma\alpha|\bar{\delta}}. \tag{10}$$

From (7) and (10) we obtain

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} - R_{\alpha\bar{\delta}\gamma\bar{\beta}} = \sum h_{\alpha\bar{\varepsilon}} T^{\varepsilon}_{\bar{\delta}\bar{\beta}|\gamma}, \tag{11}$$

where $T^{\bar{\gamma}}_{\bar{\alpha}\bar{\beta}} = \bar{T}^{\gamma}_{\alpha\beta}$. From (10) and (11) we obtain

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} - R_{\gamma\bar{\delta}\alpha\bar{\beta}} = \sum h_{\varepsilon\bar{\beta}} T^{\varepsilon}_{\gamma\alpha|\bar{\delta}} - \sum h_{\gamma\bar{\varepsilon}} T^{\bar{\varepsilon}}_{\bar{\beta}\bar{\delta}|\alpha}. \tag{12}$$

(1.2) *Remark.* Clearly the $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ have the Kähler symmetries if and only if the $T^{\gamma}_{\alpha\beta}$ are holomorphic.

Contracting (10), (11), and (12), we obtain

$$R_{\alpha\bar{\beta}} - S_{\alpha\bar{\beta}} = -\tau_{\alpha|\bar{\beta}} + \sum h^{\gamma\bar{\delta}} h_{\alpha\bar{\varepsilon}} T^{\bar{\varepsilon}}_{\bar{\delta}\bar{\beta}|\gamma} \tag{13}$$

$$R_{\alpha\bar{\beta}} - \hat{R}_{\alpha\bar{\beta}} = -\tau_{\alpha|\bar{\beta}} \tag{14}$$

$$R_{\alpha\bar{\beta}} - \hat{S}_{\alpha\bar{\beta}} = -\tau_{\bar{\beta}|\alpha} \tag{15}$$

$$S_{\alpha\bar{\beta}} - \hat{R}_{\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} h_{\alpha\bar{\varepsilon}} T^{\varepsilon}_{\bar{\delta}\bar{\beta}|\gamma} \tag{16}$$

$$S_{\alpha\bar{\beta}} - \hat{S}_{\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} h_{\varepsilon\bar{\beta}} T^{\varepsilon}_{\gamma\alpha|\bar{\delta}} \tag{17}$$

$$\hat{R}_{\alpha\bar{\beta}} - \hat{S}_{\alpha\bar{\beta}} = -\tau_{\bar{\beta}|\alpha} + \tau_{\alpha|\bar{\beta}}. \tag{18}$$

(1.3) *Remark.* Note that $R_{\alpha\bar{\beta}}$, $\hat{R}_{\alpha\bar{\beta}}$ and $\hat{S}_{\alpha\bar{\beta}}$ agree if and only if the τ_{α} are holomorphic. The τ_{α} being holomorphic is evidently not sufficient to ensure that $R_{\alpha\bar{\beta}} = S_{\alpha\bar{\beta}}$, however.

Contracting (14), we obtain

$$s - \hat{s} = -\sum h^{\alpha\bar{\beta}} \tau_{\alpha|\bar{\beta}} = -\sum h^{\alpha\bar{\beta}} \frac{\partial \tau_{\alpha}}{\partial \bar{z}^{\beta}}. \tag{19}$$

Next we give the formulas for the change of the curvatures when the metric undergoes a conformal change. Let $h'=ah$ be a new metric, where a is a positive function. We distinguish by primes the curvature and other tensors with respect to this new metric h'. Thus $h'_{\alpha\bar\beta}=ah_{\alpha\bar\beta}$, and $h'^{\alpha\bar\beta}=a^{-1}h^{\alpha\bar\beta}$. From (3) we obtain

$$R'_{\alpha\bar\beta\gamma\bar\delta}=aR_{\alpha\bar\beta\gamma\bar\delta}+\left(\frac{\partial^2 a}{\partial z^\gamma \partial \bar z^\delta}-a^{-1}\frac{\partial a}{\partial z^\gamma}\frac{\partial a}{\partial \bar z^\delta}\right)h_{\alpha\bar\beta}. \tag{20}$$

Contracting (20), we obtain

$$s'=a^{-1}s+na^{-1}L(\log a) \tag{21}$$

$$\hat s'=a^{-1}\hat s+a^{-1}L(\log a) \tag{22}$$

where $L=-\sum h^{\alpha\bar\beta}\dfrac{\partial^2}{\partial z^\alpha \partial \bar z^\beta}$ is the complex Laplacian.

Now let M be a *compact* Hermitian manifold with Hermitian metric h. We extend h to act on pairs of (p,q)-forms in the usual way. We define the usual volume form on M as

$$dV=\frac{1}{n!}\Phi^n=i^n\hat H dz^1\wedge\ldots\wedge dz^n\wedge d\bar z^1\wedge\ldots\wedge d\bar z^n, \tag{23}$$

where $\hat H$ is the determinant of $H=(h_{\alpha\bar\beta})$. We define a global inner product on the space $\mathscr{A}^{p,q}$ of global (p,q)-forms by setting

$$\langle\phi,\bar\psi\rangle_h=\int_M h(\phi,\bar\psi)\,dV.$$

We write $d=\partial+\bar\partial$, where $\partial:\mathscr{A}^{p,q}\to\mathscr{A}^{p+1,q}$ and $\bar\partial:\mathscr{A}^{p,q}\to\mathscr{A}^{p,q+1}$. We denote by δ the adjoint of d with respect to the inner product $\langle\ ,\ \rangle_h$. Thus $\langle d\phi,\bar\psi\rangle_h=\langle\phi,\overline{\delta\psi}\rangle_h$ for $\phi\in\mathscr{E}^{p-1}$ and $\psi\in\mathscr{E}^p$. We denote other adjoints by a star superscript: ∂^*, $\bar\partial^*$, Δ^*, etc. The Hodge star operator is a linear map

$$*:\mathscr{A}^{p,q}\to\mathscr{A}^{n-p,n-q}$$

with the following properties:

i) $h(\phi,\bar\psi)\dfrac{\Phi^n}{n!}=\phi\wedge *\bar\psi$

ii) $\overline{(*\psi)}=*(\bar\psi)$

iii) $**\psi=(-1)^{p+q}\psi$

$\forall\phi,\psi\in\mathscr{A}^{p,q}$. Then $\delta=-*d*$.

On a compact Hermitian manifold,

$$s-\hat s=+\tfrac12\delta\tau_{\mathbf R}+\tfrac12\|\tau_{\mathbf R}\|^2.$$

This gives us the following:

(1.4) **Proposition** (Gauduchon). *On a compact Hermitian manifold,*

a) $\int_M (s-\hat s)\,dV\geqq 0$, *with equality holding if and only if* $\tau=0$;

b) $\delta\tau_{\mathbf{R}}=0\Rightarrow s\geqq\hat{s}$

c) $\tau=0\Leftrightarrow s=\hat{s}$. □

From part c) we see that a compact Hermitian manifold is balanced if its curvature vanishes. Also from part c) we obtain

(1.5) *Remark.* A compact Hermitian manifold is balanced if and only if the τ_α are holomorphic.

Proof. τ_α holomorphic $\Rightarrow R_{\alpha\bar{\beta}}=\hat{R}_{\alpha\bar{\beta}}$ by (14)

$\Rightarrow s=\hat{s}$

$\Rightarrow \tau=0$ by (1.4)c. □

For details of the following exposition, see Gauduchon [1]. We assume M is a compact connected Hermitian manifold of complex dimensiohn $n>1$.

Definition. The metric h on the compact Hermitian manifold M is said to be standard, or to have null eccentricity, it the contracted torsion tensor is coclosed ($\delta\tau=0$).

The complex Laplacian

$$L=-\sum h^{\alpha\bar{\beta}}\frac{\partial^2}{\partial z^\alpha\partial\bar{z}^\beta}$$

is a second-order elliptic differential operator, with C^∞ coefficients, and satisfying $L(1)=0$. Hence, according to the theorem of *E.* Hopf,

$$Lf\geqq 0 \text{ (or } Lf\leqq 0) \text{ on all of } M$$
$$\Rightarrow Lf=0 \text{ and } f \text{ is constant,}$$

for any $f\in C^\infty(M)$.

L is not in general self-adjoint. In fact, Gauduchon shows (Gauduchon [1], pg. 130) that L is self-adjoint if and only if M is balanced.

L^* has a kernel of dimension 1. Thus the conditions

1) $L^*(f_0)=0$

2) $\int_M f_0 dV=\operatorname{vol} M$,

where $\operatorname{vol} M$ denotes the volume of M, uniquely define a function f_0, called the eccentricity function. Gauduchon has shown that f_0 is positive everywhere on M (Gauduchon [2]). M has null eccentricity if and only if the eccentricity function f_0 is identically 1 (Gauduchon [1] pp. 130, 131). Since f_0 is positive everywhere, we can multiply any given metric by $f_0^{1/n-1}$ to obtain a conformally related metric whose eccentricity function is identically one. Thus

(1.6) **Theorem of Null Eccentricity** (Gauduchon). *Every conformal class of metrics on a compact Hermitian manifold contains a metric of null eccentricity. It is unique up to multiplication by a constant.* □

Definition. The fundamental constant $C(M,h)$ of a compact Hermitian manifold M is defined to be

$$C(M,h)=\int_M f_0 s dV,$$

where s is the scalar curvature.

Let $h'=ah$, $a>0$, be a conformal change of metric. Then

a) $dV'=a^n dV$

b) $L'^*(f)=a^{-n}L^*(a^{n-1}f)$, $\forall f\in C^\infty$

c) $f_0'=ca^{1-n}f_0$, $c=(\int_M a^n dV)(\int_M af_0 dV)^{-1}$,

where the primes indicate the entities with respect to the metric h'.

(1.7) *Remark.* Using these formulas, together with (21) and (22), one can easily show that

$$\int_M f_0(cs+\hat{c}\hat{s})\,dV$$

is a conformal invariant for any real constants c and $\hat{c}$. In particular, the fundamental constant $C(M,h)$ is a conformal invariant of the metric h.

(1.8) **Proposition.** *Let* $Q=\int_M f_0(cs+\hat{c}\hat{s})\,dV$, *with* $nc+\hat{c}\neq 0$. *Then for some conformally related metric* h',

$$cs'+\hat{c}\hat{s}'=f_1 Q,$$

where f_1 *is a positive function. (If* $nc+\hat{c}=0$, *then* $cs+\hat{c}\hat{s}$ *has the same sign for every metric in the conformal class of* h.)

Proof. Let g be the conformally related metric of null eccentricity, and let s_g and $\hat{s}_g$ denote the scalar curvatures for this metric. Then

$$\int_M (cs_g+\hat{c}\hat{s}_g)\,dV_g=Q.$$

Since L and L^* are elliptic operators on a compact complex manifold, we have the following orthogonal decomposition of the space $\mathscr{E}^0=C^\infty(M)$, called the Hodge decomposition:

$$\mathscr{E}^0=\operatorname{Ker}L^*+\operatorname{Im}L,$$

where $\operatorname{Ker}L^*$ denotes the kernel of L^*, and $\operatorname{Im}L$ denotes the image of L. Let us call $f=cs_g+\hat{c}\hat{s}_g$. Then

$$f=\tilde{c}+L(\log\tilde{f}), \tag{24}$$

where L is the complex Laplacian with respect to g, $\tilde{c}$ is a constant (since g has null eccentricity), and $\tilde{f}>0$. Integrating both sides of (24) over M with respect to dV_g, we obtain

$$Q=\tilde{c}(\operatorname{vol}M)_g.$$

Now set $h'=\tilde{f}^b g$, where $b\in\mathbf{R}$ will be determined. h' is conformal to h. We distinguish the scalar curvature, etc. with respect to h' by primes. Using (21) and (22), we obtain

$$\begin{aligned} cs' &= c\tilde{f}^{-b}s_g+nc\tilde{f}^{-b}L(\log\tilde{f}^b)\\ &= c\tilde{f}^{-b}s_g+bnc\tilde{f}^{-b}L(\log\tilde{f}), \end{aligned}$$

and

$$\hat{c}\hat{s}'=\hat{c}\tilde{f}^{-b}\hat{s}_g+b\hat{c}\tilde{f}^{-b}L(\log\tilde{f}).$$

So

$$cs' + \hat{c}\hat{s}' = \tilde{f}^{-b}(cs_g + \hat{c}\hat{s}_g) + (nc + \hat{c})b\tilde{f}^{-b}L(\log \tilde{f})$$
$$= \tilde{f}^{-b}\tilde{c} + \tilde{f}^{-b}L(\log \tilde{f}) + (nc + \hat{c})b\tilde{f}^{-b}L(\log \tilde{f}).$$

Set $b = -1/(nc + \hat{c})$. Then $cs' + \hat{c}\hat{s}' = f_1 Q$, where

$$f_1 = \frac{\tilde{f}^{1/(nc+\hat{c})}}{(\mathrm{vol}\, M)_g}. \quad \square$$

(1.9) **Corollary.** *For some h' in the conformal class of h,*

$$s' = f\, C(M, h),$$

where $f > 0$. $\square$

Let $K = \Lambda^n T^*M$ be the canonical line bundle. A holomorphic section of $K^{\otimes m}(K^{*\otimes m})$ is called a holomorphic (m,n)-form ((m,n)-tensor), and the vector space of all such sections is denoted $\Gamma(M, K^{\otimes m})(\Gamma(M, K^{*\otimes m})$. We define

$$P_m = \dim_{\mathbb{C}} \Gamma(M, K^{\otimes m})$$
$$Q_m = \dim_{\mathbb{C}} \Gamma(M, K^{*\otimes m}).$$

If $\xi \in \Gamma(K^{\otimes m})$ and $\sigma \in \Gamma(K^{*\otimes m})$, then $\xi(\sigma)$ is a global holomorphic function and hence constant. Note that if either ξ or σ has a zero on M, then $\xi(\sigma) = 0$. Now, if $K^{\otimes m}$ has a nowhere zero holomorphic section, then $K^{\otimes m}$ is trivial and $P_m = 1$. In this case $K^{*\otimes m}$ is also trivial and so $Q_m = 1$. If $P_m > 1$, then every $\xi \in \Gamma(K^{\otimes m})$ has a zero. Therefore, if $\sigma \in \Gamma(K^{*\otimes m})$, we have $\xi(\sigma) = 0$ for all $\xi \in \Gamma(K^{\otimes m})$, which implies that σ vanishes identically. Hence $Q_m = 0$. Similarly, $Q_m > 1 \Rightarrow P_m = 0$.

(1.10) **Theorem** (c.f. Kobayashi and Wu [7], Corollary 2, and Gauduchon [1], Classification Theorem). *Let M be a compact Hermitian manifold. Then*

a) $s > 0 \Rightarrow P_m = 0$, $\forall m > 0$, *and every holomorphic (m,n)-tensor has at least one zero on M;*

b) $s < 0 \Rightarrow Q_m = 0$, $\forall m > 0$, *and every holomorphic (m,n)-form has at least one zero on M;*

c) $s = 0 \Rightarrow P_m = Q_m$, *and* $P_m \in \{0, 1\}, \forall m > 0$. *A holomorphic (m,n)-form or (m,n)-tensor is either never zero or else vansishes identically on M. If there exists a least positive integer m_0 such that $P_{m_0} = Q_{m_0} = 1$, and if $P_m = Q_m = 1$, then $m = m_1 m_0$, where m_1 is a positive integer.* $\square$

(1.11) **Corollary** (Plurigenera Theorem, Gauduchon [1]). *Let $C(M,h)$ be the fundamental constant of the compact Hermitian manifold M. Then*

a) $C(M,h) > 0 \Rightarrow P_m = 0$, $\forall m > 0$;

b) $C(M,h) < 0 \Rightarrow Q_m = 0$, $\forall m > 0$;

c) $C(M,h) = 0 \Rightarrow P_m = Q_m$ and $P_m \in \{0, 1\}$, $\forall m > 0$. $\square$

It is clear that except for the case where $P_m = Q_m = 0, \forall m > 0$, the sign of $C(M,h)$ is independent of h.

§ 2. Hermitian Manifolds of Constant Holomorphic Sectional Curvature

Definition. Let $Z_1, Z_2, \ldots, Z_n$ be a basis of T_pM and let $X = \sum a^\alpha Z_\alpha$. Then the *holomorphic sectional curvature* $H(X)$ at X is defined to be

$$H(X) = \frac{-R(X, \bar{X}, X, \bar{X})}{h^2(X, \bar{X})} = \frac{-1}{h^2(X, \bar{X})} \sum R_{\alpha\bar{\beta}\gamma\bar{\delta}} a^\alpha \bar{a}^\beta a^\gamma \bar{a}^\delta. \tag{25}$$

Remark. The property $R_{\alpha\bar{\beta}\gamma\bar{\delta}} = \bar{R}_{\beta\bar{\alpha}\delta\bar{\gamma}}$ implies that $H(X) \in \mathbf{R}$. Moreover, it is easily shown that this definition is equivalent to the usual one of

$$H(X) = h(R(X, JX)JX, X) \quad \text{for } X \text{ a unit vector.}$$

We have the following theorem for a Kähler manifold of constant holomorphic sectional curvature.

(2.1) **Theorem.** *A Kähler manifold M has constant holomorphic sectional curvature k if and only if*

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} = -\tfrac{1}{2}k(h_{\alpha\bar{\beta}} h_{\gamma\bar{\delta}} + h_{\alpha\bar{\delta}} h_{\gamma\bar{\beta}}). \tag{26}$$

Proof. This is Proposition 7.6, p. 169, of Kobayashi and Nomizu [6]. □

The simple expression (26) is possible because the curvature coefficients $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ of a Kähler metric have the Kähler symmetries. Let us define, then, on a Hermitian manifold,

$$K(X, Y, Z, W) = \tfrac{1}{4}(R(X, Y, Z, W) + R(Z, W, X, Y) + R(X, W, Z, Y) + R(Z, Y, X, W))$$

for vectors X, Y, Z and W. K has the Kähler symmetries; it is the Kähler-symmetric part of R. As before.

$$K_{\alpha\bar{\beta}\gamma\bar{\delta}} = K(Z_\alpha, \bar{Z}_\beta, Z_\gamma, \bar{Z}_\delta)$$

and

$$K_{\alpha\bar{\beta}} = -\sum h^{\gamma\bar{\delta}} K_{\alpha\bar{\beta}\gamma\bar{\delta}}$$

by definition. Note that

$$\sum h^{\alpha\bar{\beta}} K_{\alpha\bar{\beta}} = 1/2\ (s + \hat{s}).$$

Here is the Hermitian analogue of Theorem (2.1).

(2.2) **Theorem.** *Let M be a compact Hermitian manifold with metric h. Let $K_{\alpha\bar{\beta}\gamma\bar{\delta}}$ be the Kähler part of the local curvature components. Then M has constant holomorphic sectional curvature $= k$ if and only if, at every point of M,*

$$K_{\alpha\bar{\beta}\gamma\bar{\delta}} = -\tfrac{1}{2}k(h_{\alpha\bar{\beta}} h_{\gamma\bar{\delta}} + h_{\alpha\bar{\delta}} h_{\gamma\bar{\beta}}).$$

Proof. First we prove a lemma.

(2.3) **Lemma.** *Let M be a Hermitian manifold of constant holomorphic sectional curvature $= k$. Let X and Y be orthonormal vectors at some point $p \in M$. Then*

i) $R(X,\bar{X},X,\bar{X})=R(Y,\bar{Y},Y,\bar{Y})=-k$,

ii) $R(X,\bar{X},Y,\bar{Y})+R(Y,\bar{Y},X,\bar{X})+R(Y,\bar{X},X,\bar{Y})+R(X,\bar{Y},Y,\bar{X})=-2k$,

iii) $R(X,\bar{Y},X,\bar{Y})=R(Y,\bar{X},Y,\bar{X})=0$,

iv) $R(X,\bar{X},X,\bar{Y})+R(X,\bar{Y},X,\bar{X})=0$,

and

v) $R(Y,\bar{Y},Y,\bar{X})+R(Y,\bar{X},Y,\bar{Y})=0$.

Proof of Lemma. i) is by the definition of holomorphic sectional curvature $=k$. For the others, let $a,b\in\mathbf{C}$ such that $|a|^2+|b|^2=1$. Then $aX+bY$ is a unit vector, so

$$\begin{aligned}
-k&=R(aX+bY,\bar{a}\bar{X}+\bar{b}\bar{Y},aX+bY,\bar{a}\bar{X}+\bar{b}\bar{Y})\\
&=|a|^4R(X,\bar{X},X,\bar{X})+|b|^4R(Y,\bar{Y},Y,\bar{Y})\\
&\quad+|a|^2|b|^2(R(X,\bar{X},Y,\bar{Y})+R(Y,\bar{Y},X,\bar{X})+R(X,\bar{Y},Y,\bar{X})+R(Y,\bar{X},X,\bar{Y}))\\
&\quad+a^2\bar{b}^2R(X,\bar{Y},X,\bar{Y})+\bar{a}^2b^2R(Y,\bar{X},Y,\bar{X})\\
&\quad+a^2\bar{a}\bar{b}(R(X,\bar{X},X,\bar{Y})+R(X,\bar{Y},X,\bar{X}))\\
&\quad+\bar{a}^2ab(R(R(Y,\bar{X},X,\bar{X})+R(X,\bar{X},Y,\bar{X}))\\
&\quad+\bar{a}b^2\bar{b}(R(Y,\bar{Y},Y,\bar{X})+R(Y,\bar{X},Y,\bar{Y}))\\
&\quad+a\bar{b}^2b(R(Y,\bar{Y},X,\bar{Y})+R(X,\bar{Y},Y,\bar{Y})).
\end{aligned}\tag{27}$$

Setting, successively,

$$a=\frac{1}{\sqrt{2}}\qquad b=\frac{1}{\sqrt{2}},$$

$$a=\frac{-1}{\sqrt{2}}\qquad b=\frac{1}{\sqrt{2}},$$

$$a=\frac{1}{\sqrt{2}}\qquad b=\frac{i}{\sqrt{2}}$$

and

$$a=\frac{-1}{\sqrt{2}}\qquad b=\frac{i}{\sqrt{2}}$$

and eliminating, we obtain ii), as well as

$$R(X,\bar{Y},X,\bar{Y})+R(Y,\bar{X},Y,\bar{X})=0.$$

Then the first three terms of (27) are

$$\begin{aligned}
&|a|^4(-k)+|b|^4(-k)+|a|^2|b|^2(-2k)\\
&\quad=(-k)(|a|^4+2|a|^2|b|^2+|b|^4)\\
&\quad=(-k)(|a|^2+|b|^2)^2\\
&\quad=-k,
\end{aligned}$$

so that

$$\begin{aligned} &a^2\bar{b}^2R(X,\bar{Y},X,\bar{Y})+\bar{a}^2b^2R(Y,\bar{X},Y,\bar{X})\\ &\quad+a^2\bar{a}\bar{b}(R(X,\bar{X},X,\bar{Y})+R(X,\bar{Y},X,\bar{X}))\\ &\quad+\bar{a}^2ab(R(Y,\bar{X},X,\bar{X})+R(X,\bar{X},Y,\bar{X}))\\ &\quad+\bar{a}b^2\bar{b}(R(Y,\bar{Y},Y,\bar{X})+R(Y,\bar{X},Y,\bar{Y}))\\ &\quad+a\bar{b}^2b(R(Y,\bar{Y},X,\bar{Y})+R(X,\bar{Y},Y,\bar{Y}))=0. \end{aligned} \tag{28}$$

By judicious choice of a and b we can show that each summand must be zero, which will prove the lemma. Specifically, we take $a=+\frac{1}{2}\sqrt{2}$ and $b=-\frac{1}{2}\sqrt{2}i$ to obtain, from (28),

$$R(X,\bar{X},X,\bar{Y})+R(X,\bar{Y},X,\bar{X})+R(Y,\bar{Y},X,\bar{Y})+R(X,\bar{Y},Y,\bar{Y})=0.$$

Setting $a=\frac{1}{2}\sqrt{2}$ and $b=\frac{1}{2}(1-i)$ yields

$$R(X,\bar{Y},X,\bar{Y})-R(Y,\bar{X},Y,\bar{X})=0.$$

Hence iii) holds. Finally, setting $a=\frac{1}{2}\sqrt{3}$, $b=\frac{1}{2}$ and then $a=\frac{1}{2}\sqrt{3}$, $b=\frac{1}{2}i$ yield iv) and v). □

Back to Proof of (2.2). To prove sufficiency, note that the holomorphic sectional curvature sum (25) is Kähler-symmetric. Thus only the $K_{\alpha\bar{\beta}\gamma\bar{\delta}}$ contribute to the holomorphic sectional curvature. The rest of the proof is the same as in the Kähler case.

To prove necessity, assume we have an orthonormal basis of the tangent space. Then we are trying to prove

$$K_{\alpha\bar{\beta}\gamma\bar{\delta}}=-\tfrac{1}{2}k(\delta_{\alpha\beta}\delta_{\gamma\delta}+\delta_{\alpha\delta}\delta_{\gamma\beta}). \tag{29}$$

The lemma gives this conclusion when $\alpha,\beta,\gamma,\delta$ represent only one or two distinct indices: each left hand side of i)–v) is a multiple of a Kähler-symmetric curvature coefficient, and each right hand side of i)–v) is the same multiple of the right hand side of (29).

We must now show that if $\alpha,\beta,\gamma,\delta$ are 3 or 4 distinct numbers, then $K_{\alpha\bar{\beta}\gamma\bar{\delta}}=0$. Accordingly, let Z_α, Z_β, and Z_γ be distinct members of an orthonormal basis. Then Z_α and $\frac{1}{2}\sqrt{2}(Z_\beta+Z_\gamma)$ are orthonormal, and so, by the lemma,

$$\begin{aligned} K\left(Z_\alpha,\bar{Z}_\alpha,\frac{1}{\sqrt{2}}(Z_\beta+Z_\gamma),\frac{1}{\sqrt{2}}(\bar{Z}_\beta+\bar{Z}_\gamma)\right)&=-\tfrac{1}{2}k\\ &=\tfrac{1}{2}(K_{\alpha\bar{\alpha}\beta\bar{\beta}}+K_{\alpha\bar{\alpha}\beta\bar{\gamma}}+K_{\alpha\bar{\alpha}\gamma\bar{\beta}}+K_{\alpha\bar{\alpha}\gamma\bar{\gamma}}). \end{aligned}$$

Since $K_{\alpha\bar{\alpha}\beta\bar{\beta}}=K_{\alpha\bar{\alpha}\gamma\bar{\gamma}}=-\frac{1}{2}k$, we have

$$K_{\alpha\bar{\alpha}\beta\bar{\gamma}}+K_{\alpha\bar{\alpha}\gamma\bar{\beta}}=0.$$

Consideration of $K\left(Z_\alpha,\bar{Z}_\alpha,\frac{1}{\sqrt{2}}(Z_\beta+iZ_\gamma),\frac{1}{\sqrt{2}}(\bar{Z}_\beta-i\bar{Z}_\gamma)\right.$ shows that

$$K_{\alpha\bar{\alpha}\beta\bar{\gamma}} - K_{\alpha\bar{\alpha}\gamma\bar{\beta}} = 0,$$

whence $K_{\alpha\bar{\alpha}\beta\bar{\gamma}} = 0$. Similarly, consideration of

$$K\left(\frac{1}{\sqrt{2}}(Z_\alpha + Z_\beta), \frac{1}{\sqrt{2}}(\bar{Z}_\alpha + \bar{Z}_\beta), \frac{1}{\sqrt{2}}(Z_\alpha + Z_\gamma), \frac{1}{\sqrt{2}}(\bar{Z}_\alpha + \bar{Z}_\gamma)\right)$$

shows that

$$K_{\alpha\beta\alpha\bar{\gamma}} + K_{\beta\bar{\alpha}\gamma\bar{\alpha}} = 0,$$

and consideration of

$$K\left(\frac{1}{\sqrt{2}}(Z_\alpha + Z_\beta), \frac{1}{\sqrt{2}}(\bar{Z}_\alpha + \bar{Z}_\beta), \frac{1}{\sqrt{2}}(Z_\alpha + iZ_\gamma), \frac{1}{\sqrt{2}}(\bar{Z}_\alpha - i\bar{Z}_\gamma)\right)$$

shows that

$$K_{\alpha\beta\alpha\bar{\gamma}} - K_{\beta\bar{\alpha}\gamma\bar{\alpha}} = 0;$$

whence $K_{\alpha\beta\alpha\bar{\gamma}} = 0$.

Finally, we show that $K_{\alpha\beta\gamma\delta} = 0$ when all of $\alpha, \beta, \gamma, \delta$ are distinct by considering $K\left(Z_\alpha, \bar{Z}_\beta, \frac{1}{\sqrt{2}}(Z_\gamma + Z_\delta), \frac{1}{\sqrt{2}}(\bar{Z}_\gamma + \bar{Z}_\delta)\right)$ and $K\left(Z_\alpha, \bar{Z}_\beta, \frac{1}{\sqrt{2}}(Z_\gamma + iZ_\delta), \frac{1}{\sqrt{2}}(\bar{Z}_\gamma - i\bar{Z}_\delta)\right)$. □

(2.4) **Corollary.** *On a Hermitian manifold of constant holomorphic sectional curvature* $= k$,

a) $K_{\alpha\bar{\beta}} = (1/2)\, k(n+1)\, h_{\alpha\bar{\beta}}$

b) $(1/2)(s + \hat{s}) = (1/2)kn(n+1)$.

Proof. Contract $K_{\alpha\bar{\beta}\gamma\bar{\delta}}$. □

(2.5) *Remark.* Thus we see that on a Hermitian manifold of constant holomorphic sectional curvature, the sum of the scalar curvatures is constant, and the $K_{\alpha\bar{\beta}}$ satisfy an Einstein condition.

(2.6) **Corollary.** *On a compact Hermitian manifold of constant holomorphic sectional curvature* $= k$,

$$\int_M s\, dV \geqq (1/2)\, kn(n+1)(\mathrm{vol}\, M),$$

with equality holding if and only if M *is balanced.*

Proof. Since

$$\int_M s\, dV \geqq \int_M \hat{s}\, dV$$

on a compact Hermitian manifold, we have

$$\int_M s\, dV \geqq 1/2 \int_M (s + \hat{s})\, dV = (1/2)\, kn(n+1)(\mathrm{vol}\, M).$$

Equality in the line above means

$$\int_M s\,dV = \int_M \hat{s}\,dV,$$

which occurs of and only if M is balanced. □

It is sometimes convenient to use the following restatement of the last corollary.

(2.7) **Corollary.** *Let M be a compact Hermitian manifold of constant holomorphic sectional curvature $=k$. Then*

$$\int_M \gamma \wedge \phi^{n-1} \geqq (1/4\pi)\,k(n+1)!\ (\mathrm{vol}\ M),$$

where

$$\gamma = (i/2\pi) \sum R_{\alpha\beta}\, dz^\alpha \wedge d\bar{z}^\beta.$$

Equality holds if and only if M is balanced. □

The next theorem places restrictions on the Hermitian manifolds that can carry metrics of constant negative holomorphic sectional curvature.

Definition. A complex manifold M is called *hyperbolic* if the Kobayashi pseudodistance d_M is a true distance on M, i.e., if $p \neq q \Rightarrow d_M(p,q) > 0$ for $p, q \in M$. For the definition of the Kobayashi pesudodistance, see Kobayashi [5], p. 45.

Remark. If M is a hyperbolic manifold, then every holomorphic map $f\colon \mathbf{C} \to M$ is constant (Kobayashi [5], p. 60).

(2.8) **Theorem.** *Let M be a Hermitian manifold with holomorphic sectional curvature $\leqq k < 0$. Then M is hyperbolic.*

Proof. This is Theorem 4.11, p. 61, of Kobayashi [5].

§ 3. Compact Hermitian Manifolds of Constant Holomorphic Sectional Curvature

The following proposition shows that there are examples of compact non-Kähler Hermitian manifolds of constant zero holomorphic sectional curvature in every dimension above 2.

(3.1) **Proposition.** *Let G be a complex Lie group and $\Gamma \subset G$ a discrete subgroup. Then there is a G-invariant Hermitian metric on $M = G/\Gamma$ with vanishing curvature. Moreover, it is Kähler if and only if G is abelian.*

Proof (See also Goldberg [4], §§ 6.6 and 6.7). Let $\mathfrak{G}$ be the Lie algebra of the complex Lie group G. Let $E_1, \ldots, E_n$ be a complex basis of $\mathfrak{G}$, and define θ^i, $i = 1, \ldots, n$, by $\theta = \sum \theta^i E_i$, where θ is the canonical one-form on G. Let $\pi\colon G \to G/\Gamma$ be the projection. Define $\phi^1, \ldots, \phi^n$ on M by requiring that $\pi^* \phi^i = \theta^i$. Since Γ is discrete, the ϕ^i are holomorphic one-forms everywhere independent on M. Define a connection ∇ by requiring $\nabla \phi^i = 0$. We will show that ∇ is the Hermitian connection of the metric $h = \sum \phi^i \bar{\phi}^i$. We denote by s_i the duals of the ϕ^i.

a) $\nabla h = 0$. This is obvious since the ϕ^i are parallel.

b) $\nabla J = 0$. Since the s_i are parallel, we have

$$\begin{aligned}\nabla_{s_j}(J(s_i)) &= (\nabla_{s_j} J)(s_i) - J(\nabla_{s_j} s_i) \\ &= (\nabla_{s_j} J)(s_i).\end{aligned}$$

On the other hand, since the s_i are holomorphic, we have

$$\begin{aligned}\nabla_{s_j}(J(s_i)) &= \nabla_{s_j}(i s_i) \\ &= i \nabla_{s_j} s_i = 0.\end{aligned}$$

Thus $\nabla J = 0$.

c) The $(1,1)$-part of the torsion vanishes. By definition,

$$T(s_i, \bar{s}_j) = \nabla_{s_i} \bar{s}_j - \nabla_{\bar{s}_j} s_i - [s_i, \bar{s}_j].$$

This vanishes because the s_i are parallel, and bracket preserves type on a complex manifold. This completes the proof that ∇ is the Hermitian connection of h.

Note that the torsion vanishes if and only if G is abelian.

The curvature vanishes for this connection because

$$R(s_i, \bar{s}_j) s_k = \nabla_{\bar{s}_i}(\nabla_{\bar{s}_j} s_k) - \nabla_{\bar{s}_j}(\nabla_{\bar{s}_i} s_k) - \nabla_{[s_i, \bar{s}_j]} s_k$$

and

$$\nabla_{\bar{s}_j} s_k = \nabla_{\bar{s}_i} s_k = [s_i, \bar{s}_j] = 0. \quad \square$$

These are the only known examples of compact non-Kähler Hermitian manifolds of constant holomorphic sectional curvature. It is significant that the curvature vanishes for all these manifolds. A compact Hermitian surface with vanishing curvature is Kähler, so these examples do not occur in dimension 2. They do occur in dimension 3, however; the Iwasawa manifold is an example. The Iwasawa manifold M is defined by

$$M = G/\Gamma, \text{ where } G = \left\{ \begin{bmatrix} 1 & z^1 & z^2 \\ 0 & 1 & z^3 \\ 0 & 0 & 1 \end{bmatrix} : z^i \in \mathbf{C} \right\}$$

$$\text{and } \Gamma = \left\{ \begin{bmatrix} 1 & a^1 & a^2 \\ 0 & 1 & a^3 \\ 0 & 0 & 1 \end{bmatrix} : a^i \text{ Gaussian integers} \right\}.$$

Thus M is the quotient of a nonabelian Lie group by a discrete subgroup, and hence is an example of a 3-dimensional compact non-Kähler Hermitian manifold of vanishing curvature.

We are now ready to prove the main results of this paper, which we present in order of increasing generality.

If the Hermitian manifold M has holomorphic torsion and constant holomorphic sectional curvature, then the curvature components $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ have the Kähler symmetries (Remark (1.2)), and we can draw strong conclusions about M. Notice we do not need M to be compact.

(3.2) **Theorem.** *Suppose M is a Hermitian manifold with holomorphic torsion components $T^{\gamma}_{\alpha\beta}$, and constant holomorphic sectional curvature $=k$.*

a) *If $k\neq 0$, then M is Kähler.*

b) *If $k=0$, M has vanishing curvature.*

Proof. It follows from Remark (1.2) and Theorem (2.2) that

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}} = -(1/2)\,k(h_{\alpha\bar{\beta}}h_{\gamma\bar{\delta}} + h_{\alpha\bar{\delta}}h_{\gamma\bar{\beta}}).$$

It is clear that the $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ vanish if $k=0$.

To prove a), we assume $k\neq 0$. The $R_{\alpha\bar{\beta}\gamma\bar{\delta}}$ are parallel, so the second Bianchi identity

$$R_{\alpha\bar{\beta}\gamma\bar{\delta}|\varepsilon} - R_{\alpha\bar{\beta}\varepsilon\bar{\delta}|\gamma} = \sum R_{\alpha\bar{\beta}\rho\bar{\delta}}\,T^{\rho}_{\varepsilon\gamma},$$

reduces to

$$\sum (h_{\alpha\bar{\beta}}h_{\rho\bar{\delta}} + h_{\alpha\bar{\delta}}h_{\rho\bar{\beta}})\,T^{\rho}_{\varepsilon\gamma} = 0.$$

Multiplying by $h^{\alpha\bar{\beta}}$ and contracting gives

$$\sum h_{\rho\bar{\delta}}\,T^{\rho}_{\varepsilon\gamma} = 0 \Rightarrow T^{\rho}_{\varepsilon\gamma} = 0, \qquad \forall \rho, \varepsilon, \gamma,$$

since $(h_{\rho\bar{\delta}})$ is positive definite. □

Remark. The manifolds $M = G/\Gamma$ of Proposition (3.1) have holomorphic torsion.

(3.3) **Theorem.** *Let (M,h) be a compact Hermitian manifold of dimension n and constant holomorphic sectional curavature $=k$. Suppose h is conformal to a balanced metric h'. Let P_m be the mth plurigenus, and Q_m the mth dual plurigenus (see §2 for definitions). Then*

a) *$k>0 \Rightarrow P_m = 0$, $\forall m>0$, and every holomorphic (m,n)-tensor has at least one zero on M;*

b) *$k<0 \Rightarrow Q_m = 0$, $\forall m>0$, and every holomorphic (m,n)-form has at least one zero on M;*

c) *$k=0 \Rightarrow P_m = Q_m$, and $P_m \in \{0,1\}$, $\forall m>0$. A holomorphic (m,n)-form or (m,n)-tensor is either never zero or else vanishes identically on M. If there exists a least positive integer m_0 such that $P_{m_0} = Q_{m_0} = 1$, and if $P_m = Q_m = 1$, then $m = m_1 m_0$, where m_1 is a positive integer.*

Proof. Recall that

$$s + \hat{s} = kn(n+1)$$

for a metric of constant holomorphic sectional curvature $=k$. Thus

$$\begin{aligned}
\int_M f_0(s+\hat{s})\,dV &= kn(n+1)(\operatorname{vol} M) \\
&\int_M (s'+\hat{s}')\,dV' \text{ by (1.7)} \\
&= 2\int_M s'\,dV' \text{ since } s' = \hat{s}' \text{ for } h' \text{ balanced} \\
&= 2\,C(M,h).
\end{aligned}$$

Thus $C(M,h) = \frac{1}{2}kn(n+1)(\operatorname{vol} M)$. The theorem now follows from (1.11). □

Remark. Note that the only property of constant holomorphic sectional curvature we have used in the proof is that $s+\hat{s}$ is constant. Therefore the hypothesis can be weakened to: M is a compact Hermitian manifold whose metric is conformal to a balanced metric, and whose scaler curvatures satisfy $s+\hat{s}=k$, constant.

Finally we come to the case where we make no assumptions on the torsion of M.

(3.4) **Theorem.** *Let M be a compact Hermitian manifold of constant holomorphic sectional curvature $=k$. Let $C(M,h)$ be the fundamental constant of (M,h). Then*

a) $k>0 \Rightarrow C(M,h)>0$;

b) $k=0 \Rightarrow C(M,h)\geqq 0$, *with equality holding if and only if h is conformal to a balanced metric.*

Proof. Let f_0 be the eccentricity function for h. The conformally related metric

$$h'=f_0^{1/n-1}h$$

has null eccentricity, i.e., $f_0'\equiv 1$. By (1.7),

$$\begin{aligned}\int_M (s'+\hat{s}')\,dV' &= \int_M f_0(s+\hat{s})\,dV \\ &= kn(n+1)(\mathrm{vol}\,M).\end{aligned}$$

(To obtain the last equality we used (2.4b) and the fact that

$$\int_M f_0\,dV=\mathrm{vol}\,M.)$$

Since the fundamental constant is a conformal invariant, we have

$$\begin{aligned}C(M,h)=C(M,h')&=\int_M s'\,dV'\geqq(1/2)\int_M (s'+\hat{s}')\,dV' \\ &=(1/2)\,kn(n+1)(\mathrm{vol}\,M).\end{aligned}$$

(The inequality used was (1.4a).) a) and b) follow from the inequality above.

Now suppose $k=0$ and $C(M,h)=0$. Then we must have

$$\int_M s'\,dV'=\int_M \hat{s}'\,dV',$$

and so by (1.4a), (M,h') is balanced. □

(3.5) **Corollary.** *Let M be a compact Hermitian manifold of constant holomorphic sectional curvature $=k$. Let P_m be the mth plurigenus and Q_m be the mth dual plurigenus of M. Then*

a) $k>0 \Rightarrow P_m=0, \forall m>0$;

b) $k=0 \Rightarrow$ *either* $P_m=0, \forall m>0$, *or* $P_m=Q_m, \forall m>0$, *and* $P_m\in\{0,1\}$. □

Remark. As before, note the last theorem and its corollary use only the fact that $s+\hat{s}$ is constant.

§ 4. Einstein Conditions

Definition. We say that the metric h satisfies an Einstein condition if one of the Ricci-type tensors has components $fh_{\alpha\bar{\beta}}$, where f is a real function.

The core of the following proof is due to Goldberg [3].

(4.1) **Theorem.** *Suppose that on the Hermitian manifold M of dimension $n>1$, an Einstein condition is satisfied by any one of $R_{\alpha\bar{\beta}}$, $\hat{R}_{\alpha\bar{\beta}}$, or $\hat{S}_{\alpha\bar{\beta}}$, for some f not identically zero. Then $\tau=0$ if and only if f is constant, and either of these conditions implies M is Kähler.*

Proof. We start with the second Beanchi identity:

$$R_{\alpha\bar{\beta}\varepsilon\bar{\delta}|\gamma}-R_{\alpha\bar{\beta}\gamma\bar{\delta}|\varepsilon}=\sum R_{\alpha\bar{\beta}\rho\bar{\delta}}T^{\rho}_{\gamma\varepsilon}. \tag{30}$$

If

$$R_{\alpha\bar{\beta}}=fh_{\alpha\bar{\beta}}, \tag{31}$$

we multiply (30) by $h^{\alpha\bar{\beta}}$ and contract, obtaining

$$R_{\varepsilon\bar{\delta}|\gamma}-R_{\gamma\bar{\delta}|\varepsilon}=\sum R_{\rho\bar{\delta}}T^{\rho}_{\gamma\varepsilon}.$$

We make the substitution (31), then multiply by $h^{\varepsilon\bar{\delta}}$ and contract again, obtaining

$$(n-1)f_{|\gamma}=f\tau_{\gamma}. \tag{32}$$

Note

$$\hat{R}_{\alpha\bar{\beta}}=fh_{\alpha\bar{\beta}} \tag{33}$$

if and only if

$$\hat{S}_{\alpha\bar{\beta}}=fh_{\alpha\bar{\beta}}. \tag{34}$$

If either of these conditions holds, we multiply (30) by $h^{\alpha\bar{\delta}}$ and contract. We next make the substitution (33) or (34), then multiply by $h^{\varepsilon\bar{\beta}}$ and contract. Again we obtain (32). Thus $\tau=0\Rightarrow f$ is constant.

Conversely, if f is constant, (32) becomes

$$f\tau_{\gamma}=0,$$

with $f\neq 0$. Hence $\tau=0$.

Finally, $\tau=0\Rightarrow R_{\alpha\bar{\beta}}=\hat{R}_{\alpha\bar{\beta}}=\hat{S}_{\alpha\bar{\beta}}$ (Remark (1.3)). Thus the Ricci form $\sum R_{\alpha\bar{\beta}}dz^{\alpha}\wedge d\bar{z}^{\beta}$, which is closed, is proportional to the Kähler form. Hence M is Kähler. □

(4.2) **Theorem.** *Let M be a compact Hermitian manifold with constant zero holomorphic sectional curvature. Suppose further that either $s=0$ or $\hat{s}=0$. Then M is balanced.*

Proof. M has constant zero holomorphic sectional curvature implies

$$s+\hat{s}=0.$$

Hence $s=\hat{s}=0$, and M is balanced by (1.4c). □

Bibliography

1. Gauduchon, P.: Fibrés hermitiens à endomorphisme de Ricci non négatif. Bull. Soc. Math. Fr. **105**, 113–140 (1977)
2. Gauduchon, P.: Le théorème de l'excentricité nulle. Comptes Rendus (Acad. Sc. Paris), t. **285**, Serie A (1977), 387–390
3. Goldberg, S.I.: Tensorfields and curvature in Hermitian manifolds with torsion. Ann. Math. **63**, 64–76 (1956)
4. Goldberg, S.I.: Curvature and Homology. New York: Academic Press 1962
5. Kobayashi, S.: Hyperbolic Manifolds and Holomorphic Mappings. New York: Dekker 1970
6. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry, Vol. II. New York: Wiley 1969
7. Kobayashi, S., Wu, H.: On holomorphic sections of certain Hermitian vector bundles. Math. Ann. **189**, 1–4 (1970)

Received November 4, 1983; received in final form July 2, 1984

Math. Z. 189, 211–226 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Charakterisierung symmetrischer R-Räume durch ihre Einheitsgitter

Ottmar Loos

Institut für Mathematik der Universität, Innrain 52, A-6020 Innsbruck, Österreich

Herrn Professor Dr. Max Koecher zum 60. Geburtstag gewidmet

Einleitung

Symmetrische Räume, die sich als Quotienten halbeinfacher Lie-Gruppen nach parabolischen Untergruppen darstellen lassen, also R-Räume im Sinne von [16] sind, bilden einen interessanten Typ von symmetrischen Räumen und sind von einer Reihe von Autoren untersucht worden. D. Ferus hat sie als die kompakten symmetrischen Untermannigfaltigkeiten des Euklidischen Raumes charakterisiert ([5]). In der vorliegenden Arbeit werden die symmetrischen R-Räume innerhalb der Klasse aller kompakten symmetrischen Räume durch die einfache innergeometrische Eigenschaft gekennzeichnet, daß die Einheitsgitter ihrer maximalen Tori bezüglich einer unter allen Symmetrien invarianten Riemannschen Metrik kubische Gitter sind. Der Beweis beruht auf der Korrespondenz zwischen symmetrischen R-Räumen und kompakten Jordan-Tripelsystemen. Diese Ergebnisse wurden in [10] angekündigt.

Hier ist eine ausführlichere Inhaltsangabe. Es sei M ein kompakter zusammenhängender symmetrischer Raum im Sinne von [9]; die Symmetrie (Spiegelung) am Punkte x von M werde mit S_x bezeichnet. Überdies sei $M = X(\mathbb{R})$ die Menge der reellen Punkte eines R-Raumes, d.h., einer über $\mathbb{R}$ definierten homogenen projektiven rationalen Varietät X. Ein Diffeomorphismus von M läßt sich auf höchstens eine Weise zu einem Automorphismus von X fortsetzen. Daher ist folgende Definition sinnvoll: X heiße mit M verträglich, und das Paar (M, X) heiße ein symmetrischer R-Raum, wenn alle Symmetrien von M Automorphismen von X sind. Sei G die algebraische Einskomponente der Automorphismengruppe von X. In Satz 1 wird gezeigt: Nach Wahl eines Basispunktes $o \in M$ gibt es genau einen Homomorphismus γ der multiplikativen Gruppe von $\mathbb{R}$ in die Isotropiegruppe von o in G derart, daß $\gamma(t)$ auf dem Tangentialraum $T_o(M)$ wie $t \cdot \mathrm{Id}$ operiert und $\gamma(-1) = S_o$ ist. Die Gewichte von $\mathrm{Ad}\,\gamma$ auf der Lie-Algebra von G sind $0, \pm 1$. Es gibt eine Cartan-Involution θ von G mit $\theta(\gamma(t)) = \gamma(t^{-1})$, und der Gewichtsraum $\mathfrak{V}^+$ zum Gewicht $+1$ wird mit dem Tripelprodukt $\{xyz\} = -[[x, \theta y], z]$ zu einem kompakten Jordan-Tripelsystem. Man erhält auf diese Weise eine Bijektion zwischen Isomorphieklassen symmetrischer R-Räume und kompakter Jordan-Tripelsysteme (Satz 2).

Vermöge der Auswertungsabbildung der Vektorfelder wird $T_o(M)$ zu einem Jordan-Tripelsystem $J(M,X)\cong\mathfrak{B}^+$. Dann ist das Lie-Tripelsystem $L(M)$ von M mit $J(M,X)$ durch

$$[uvw]=-\{uvw\}+\{vuw\} \tag{i}$$

verknüpft, und ein Rahmen $e_1,\ldots,e_r$ von $J(M,X)$ (d.h., ein maximales Orthogonalsystem primitiver Tripotente) spannt einen maximalen abelschen Unterraum $\mathfrak{A}$ von $L(M)$ auf mit der Eigenschaft, daß das Einheitsgitter $\Gamma(A)=\{v\in\mathfrak{A}: \operatorname{Exp}(v)=o\}$ des maximalen Torus $A=\operatorname{Exp}\mathfrak{A}$ durch

$$\Gamma(A)=\sum\mathbb{Z}\pi e_i \tag{ii}$$

gegeben ist (Satz 3). Man erhält so eine Bijektion zwischen symmetrischen R-Raumstrukturen auf M und Jordan-Tripelstrukturen auf $T_o(M)$, die (i) und (ii) erfüllen (Satz 4). Eine Folgerung hieraus ist: Ein symmetrischer R-Raum (M,X) besitzt eine kanonische unter allen Symmetrien invariante Riemannsche Metrik g mit $g(e_i,e_j)=\delta_{ij}$ (Satz 5). Insbesondere sind die Einheitsgitter der maximalen Tori von M kubisch, d.h., sie besitzen eine $\mathbb{Z}$-Basis aus orthogonalen Vektoren gleicher Länge. Diese rein geometrische Eigenschaft ist bereits charakteristisch für symmetrische R-Räume: Eine unter allen Symmetrien invariante Metrik eines kompakten symmetrischen Raumes M, bezüglich der die Einheitsgitter der maximalen Tori kubisch sind, ist bis auf einen skalaren Faktor die kanonische Metrik einer eindeutigen symmetrischen R-Raumstruktur X von M (Satz 6). Der Beweis beruht auf Satz 4 und besteht in der Konstruktion des Jordan-Tripelsystems $J(M,X)$ aus $L(M)$ und einer Orthonormalbasis des Einheitsgitters.

§ 1. Korrespondenz zwischen symmetrischen R-Räumen und kompakten Jordan-Tripelsystemen

Sei $\operatorname{Aut}(M)$ die Automorphismengruppe eines kompakten symmetrischen Raumes M, also die Gruppe aller Diffeomorphismen f von M mit $f(S_x y)=S_{f(x)}f(y)$ für alle $x,y\in M$, und sei K die von den „Transvektionen" $S_x S_y$ erzeugte Untergruppe. Dann ist $\operatorname{Aut}(M)$ eine Lie-Transformationsgruppe von M und K ein kompakter zusammenhängender Normalteiler, der transitiv auf M operiert ([9]). Im allgemeinen ist $\operatorname{Aut}(M)$ nicht kompakt.

Lemma 1. *Eine kompakte Lie-Transformationsgruppe von M, die alle Spiegelungen enthält, besteht aus Automorphismen von M.*

Beweis. Sei K' eine solche Gruppe. Wähle eine unter K' invariante Riemannsche Metrik auf M. Für $x\in M$ und $f\in K'$ sind dann $f\circ S_x$ und $S_{f(x)}\circ f$ Isometrien, die wegen $T_x(S_x)=-\mathrm{Id}$ auf $T_x(M)$ übereinstimmen und daher gleich sind. Also ist f ein Automorphismus von M.

Sei nun $o\in M$ ein fest gewählter Punkt, K_o die Isotropiegruppe von o in K, und σ der durch Konjugation mit S_o gegebene Automorphismus von $\operatorname{Aut}(M)$ bzw. K. Dann liegt K_o zwischen der Fixpunktmenge von σ und ihrer Einskom-

ponente. Die (± 1)-Eigenraumzerlegung von $\mathfrak{k}=L(K)$ unter σ ist

$$\mathfrak{k}=\mathfrak{k}_o\oplus\mathfrak{m}$$

mit $\mathfrak{k}_o=L(K_o)$. Es gilt $\mathfrak{k}_o=[\mathfrak{m},\mathfrak{m}]$ und $\mathfrak{k}$ ist die Standardeinbettung des Lie-Tripelsystems $\mathfrak{m}$.

Lemma 2. *K ist die Einskomponente von* $\mathrm{Aut}(M)$.

Beweis. $\mathfrak{k}$ ist die direkte Summe seines Zentrums $\mathfrak{z}$ und des halbeinfachen Ideals $\mathfrak{k}'=[\mathfrak{k},\mathfrak{k}]$. Beide sind unter σ invariant. Da $\mathfrak{k}_o$ vermöge der adjungierten Darstellung treu auf $\mathfrak{m}$ operiert, gilt $\mathfrak{z}\cap\mathfrak{k}_o=0$, also $\mathfrak{m}=\mathfrak{z}\oplus\mathfrak{m}'$ mit $\mathfrak{m}'=\mathfrak{k}'\cap\mathfrak{m}$, und $\mathfrak{k}'=\mathfrak{k}_o\oplus\mathfrak{m}'$. Die Lie-Algebra $\mathfrak{d}=\mathrm{Der}(M)$ von $\mathrm{Aut}(M)$ zerfällt unter σ in $\mathfrak{d}_o\oplus\mathfrak{m}$ ([9]) mit $\mathfrak{k}_o\subset\mathfrak{d}_o$, und wir können $\mathfrak{d}_o$ mit einer Unteralgebra der Derivationsalgebra des Lie-Tripelsystems $\mathfrak{m}$ identifizieren. Nun ist jede Derivation des halbeinfachen Lie-Tripelsystems $\mathfrak{m}'$ inner, also gilt $\mathrm{Der}(\mathfrak{m})=\mathrm{Der}(\mathfrak{z})\oplus\mathrm{Der}(\mathfrak{m}')=\mathrm{Der}(\mathfrak{z})\oplus\mathfrak{k}_o$ mit $[\mathrm{Der}(\mathfrak{z}),\mathfrak{m}']=0$. Für ein $Z\in\mathfrak{d}_o\cap\mathrm{Der}(\mathfrak{z})$ ist $\mathrm{Int}(\exp(tZ))$ eine Einparametergruppe von Automorphismen des Torus $\exp(\mathfrak{z})$ von K. Da die Automorphismengruppe eines Torus diskret ist, folgt $[Z,\mathfrak{z}]=0$ und daraus wegen $[Z,\mathfrak{m}']=0$ auch $[Z,\mathfrak{m}]=0$, also $Z=0$. Somit gilt $\mathfrak{d}_o=\mathfrak{k}_o$ und das Lemma ist bewiesen.

Der Automorphismengruppenfunktor eines reellen R-Raumes X ist durch eine reelle affine algebraische Gruppe $\mathbf{Aut}(X)$ darstellbar. Bezeichnet G ihre algebraische Einskomponente und $P\subset G$ die Isotropiegruppe eines Punktes $o\in X(\mathbb{R})$, so ist G halbeinfach und adjungiert (zentrumsfrei), P parabolisch, und die Operation von $\mathbf{Aut}(X)$ auf G durch innere Automorphismen identifiziert $\mathbf{Aut}(X)$ mit einer offenen Untergruppe von $\mathbf{Aut}(G)$ (siehe [4]). Da der Grundkörper die Charakteristik 0 hat, können wir $\mathbf{Aut}(X)$ vermöge der adjungierten Darstellung mit einer offenen Untergruppe von $\mathbf{Aut}(\mathfrak{g})$ identifizieren, wobei $\mathfrak{g}=L(G)$ die Lie-Algebra von G sei. Schließlich gilt $X=G/P$ und weil P parabolisch ist, auch $X(\mathbb{R})=G(\mathbb{R})/P(\mathbb{R})$ ([1], [4]). Wir können $\mathrm{Aut}(X)=(\mathbf{Aut}(X))\,(\mathbb{R})$ und $G(\mathbb{R})$ als Lie-Transformationsgruppen von $X(\mathbb{R})$ auffassen. Diese Gruppen haben endlich viele topologische Zusammenhangskomponenten, im allgemeinen ist aber $G(\mathbb{R})$ nicht topologisch zusammenhängend.

Nun sei (M,X) ein symmetrischer R-Raum. Ein *Automorphismus* von (M,K) ist ein Automorphismus f der Varietät X, der auf $M=X(\mathbb{R})$ ein Automorphismus des symmetrischen Raumes M ist. Die Automorphismengruppe von (M,X) ist also $\mathrm{Aut}(M,X)=\mathrm{Aut}(M)\cap\mathrm{Aut}(X)$; sie enthält alle Symmetrien S_x. Die Gruppe K ist in der topologischen Einskomponente $G(\mathbb{R})^0$ von $G(\mathbb{R})$ enthalten. Mit σ werde wieder der durch Konjugation mit S_o definierte Automorphismus von $\mathbf{Aut}(X)$ bzw. G bezeichnet.

Satz 1. *Sei (M,X) ein symmetrischer R-Raum.*

(a) *$K=\mathrm{Aut}(M)\cap G(\mathbb{R})^0$ ist maximal kompakt in $G(\mathbb{R})^0$ und $\mathrm{Aut}(M,X)$ ist maximal kompakt in $\mathrm{Aut}(X)$. Es gibt genau einen involutorischen Automorphismus θ von G, dessen Fixpunktmenge in $G(\mathbb{R})^0$ gleich K ist und der mit σ vertauscht.*

(b) *$P^+=\theta(P)$ ist eine zu $P^-=P$ entgegengesetzte parabolische Untergruppe, und $H=P^+\cap P^-$ ist eine Levi-Untergruppe von $P^\pm$. Das unipotente Radikal $U^\pm$ von*

$P^{\pm}$ ist abelsch. Die Lie-Algebra $\mathfrak{g}$ zerfällt in

$$\mathfrak{g}=\mathfrak{V}^{+}\oplus\mathfrak{h}\oplus\mathfrak{V}^{-} \tag{1}$$

mit $\mathfrak{h}=L(H)$ und $\mathfrak{V}^{\pm}=L(U^{\pm})$, und es gilt

$$[\mathfrak{h},\mathfrak{V}^{\varepsilon}]\subset\mathfrak{V}^{\varepsilon},\quad [\mathfrak{V}^{\varepsilon},\mathfrak{V}^{\varepsilon}]=0,\quad [\mathfrak{V}^{+},\mathfrak{V}^{-}]\subset\mathfrak{h}. \tag{2}$$

(c) *Es gibt genau einen Homomorphismus γ der multiplikativen Gruppe $\mathbb{R}_m$ von $\mathbb{R}$ nach P mit den Eigenschaften*

$$\gamma(t)\ \textit{operiert auf}\ T_o(M)\ \textit{wie}\ t\cdot\mathrm{Id}, \tag{3}$$

$$\gamma(-1)=S_o. \tag{4}$$

Für γ gilt weiter

$$\theta(\gamma(t))=\gamma(t^{-1}), \tag{5}$$

der Zentralisator von γ ist H, und die Gewichte von $\mathrm{Ad}\,\gamma$ *sind* $0,\pm 1$.

Beweis. (a) Sei K' eine maximal kompakte Untergruppe von $\mathrm{Aut}(X)$, die $K\cup K\cdot S_o$ umfaßt. Nach Lemma 1 ist $K'\subset\mathrm{Aut}(M)$ und nach Lemma 2 ist $(K')^0=K$. Nach bekannten Sätzen über maximal kompakte Untergruppen von reellen Lie-Gruppen mit endlich vielen Komponenten ist daher K maximal kompakt in $G(\mathbb{R})^0=\mathrm{Aut}(X)^0$. Für $g\in\mathrm{Aut}(M)\cap G(\mathbb{R})^0$ gilt $gKg^{-1}=K$ nach Lemma 2. Andrerseits ist K als maximal kompakte Untergruppe der reellen halbeinfachen zentrumsfreien Lie-Gruppe $G(\mathbb{R})^0$ sein eigener Normalisator. Also ist $g\in K$ und daher $K=\mathrm{Aut}(M)\cap G(\mathbb{R})^0$. Weiter gilt $\mathrm{Aut}(X)=K'\cdot G(\mathbb{R})^0$, denn K' trifft jede topologische Komponente von $\mathrm{Aut}(X)$. Folglich $\mathrm{Aut}(X)\cap\mathrm{Aut}(M)=K'\cdot(G(\mathbb{R})^0\cap\mathrm{Aut}(M))=K'\cdot K=K'$. Es gibt genau eine Cartan-Involution θ von $\mathfrak{g}$ mit Fixpunktmenge $\mathfrak{k}$. Da G durch Ad isomorph zur algebraischen Einskomponente von $\mathbf{Aut}(\mathfrak{g})$ ist und diese von θ normalisiert wird, läßt sich θ zu einem wieder mit θ bezeichneten Automorphismus von G fortsetzen. Wegen $\sigma(\mathfrak{k})=\mathfrak{k}$ gilt dann $\sigma\theta=\theta\sigma$, und nach bekannten Tatsachen über Cartan-Involutionen ist die Fixpunktmenge von θ in $G(\mathbb{R})^0$ gleich K.

(b) Für die ersten beiden Aussagen siehe [2, 1.9]. Wegen $S_o(o)=o$ ist P^{-} und somit auch U^{-} stabil unter σ, und wegen $\sigma\theta=\theta\sigma$ gilt dasselbe für P^{+}, U^{+}, und H. Da $U^{+}\cdot H\cdot U^{-}$ offen in G ist, gilt (1). Weiter ist $G(\mathbb{R})^0=K\cdot P(\mathbb{R})^0$, denn K operiert transitiv auf M, und $K\cap P(\mathbb{R})=K_o$. Aus $\mathfrak{k}=\mathfrak{k}_o\oplus\mathfrak{m}$ folgt daher $\mathfrak{g}=\mathfrak{m}\oplus\mathfrak{h}\oplus\mathfrak{V}^{-}$. Aus (1) und $\sigma|\mathfrak{m}=-\mathrm{Id}$ sowie $\sigma(\mathfrak{V}^{+})=\mathfrak{V}^{+}$ folgt nun $\sigma|\mathfrak{V}^{+}=-\mathrm{Id}$ und daher ist $\mathfrak{V}^{+}$ abelsch. Dies gilt dann auch für $\mathfrak{V}^{-}=\theta(\mathfrak{V}^{+})$ und für $U^{\pm}$, denn diese Gruppen sind zusammenhängend. Da $U^{\pm}$ von H normalisiert wird, bleibt von (2) nur die letzte Inklusion zu beweisen. Nun liegt $[\mathfrak{V}^{+},\mathfrak{V}^{-}]$ im $(+1)$-Eigenraum von σ, und dieser ist nach (1) in $\mathfrak{h}$ enthalten.

(c) Nach (1) und (2) wird durch

$$\alpha(t)\cdot v=t^{\pm 1}v,\qquad \alpha(t)\cdot h=h \tag{6}$$

für $v\in\mathfrak{V}^{\pm}$ und $h\in\mathfrak{h}$ ein Homomorphismus $\alpha\colon\mathbb{R}_m\to\mathbf{Aut}(\mathfrak{g})^0$ definiert. Wegen des Isomorphismus $\mathrm{Ad}\colon G\to\mathbf{Aut}(\mathfrak{g})^0$ gibt es genau ein $\gamma\colon\mathbb{R}_m\to G$ mit $\mathrm{Ad}\,\gamma=\alpha$. Da

$\mathrm{Ad}\,H$ die Zerlegung (1) invariant läßt, vertauscht α mit $\mathrm{Ad}\,H$ und folglich γ mit H. Der Zentralisator von γ ist zusammenhängend und hat nach (6) dieselbe Lie-Algebra wie H. Also ist H der Zentralisator von γ, insbesondere ist γ ein Homomorphismus nach $H \subset P$. Die adjungierte Darstellung von H auf $\mathfrak{V}^+$ ist isomorph zur linearen Isotropiedarstellung auf $T_o(M)$ und daher gilt (3). Aus (6) folgt sofort $\theta\alpha(t)\theta = \alpha(t^{-1})$ und somit (5). Insbesondere ist $\gamma(-1)$ fix unter θ. Daher normalisiert $\gamma(-1)$ die Fixpunktmenge von θ in $G(\mathbb{R})^0$, d.i. K. Nun ist K sein eigener Normalisator in $G(\mathbb{R})^0$ und wegen $\mathrm{Aut}(X) = \mathrm{Aut}(M, X) \cdot G(\mathbb{R})^0$ ist $\mathrm{Aut}(M, X)$ der Normalisator von K in $\mathrm{Aut}(X)$. Es folgt $\gamma(-1) \in \mathrm{Aut}(M)$. Nach (3) haben $\gamma(-1)$ und S_o dieselbe Tangentialabbildung im Punkte o und sind daher gleich.

Es bleibt zu zeigen, daß γ durch (3) und (4) eindeutig bestimmt ist. Dazu habe β dieselben Eigenschaften, und sei $\beta(t) = \beta_H(t) \cdot \phi(t)$ die Zerlegung im semidirekten Produkt $P = H \cdot U^-$. Aus (3) folgt $\mathrm{Ad}\,\beta_H(t) \cdot v = tv$ für $v \in \mathfrak{V}^+$. Nun ist die adjungierte Darstellung von H auf $\mathfrak{V}^+ \cong U^+$ treu, denn die Bahn von o unter U^+ ist offen und dicht in X. Daher folgt $\beta_H = \gamma$. Weiter ist $\phi\colon \mathbb{R}_m \to U^-$ ein gekreuzter Homomorphismus von $\mathbb{R}_m$ in die Vektorgruppe $U^- \cong \mathfrak{V}^-$, also wegen $H^1(\mathbb{R}_m, \mathfrak{V}^-) = 0$ von der Form $\phi(t) = \exp((t-1)v)$. Nun zeigt $\beta(-1) = \gamma(-1) \cdot \exp(-2v) = S_o = \gamma(-1)$, daß $v = 0$, also $\beta = \gamma$.

Korollar. *Das Element $Z = \dot{\gamma}(1) \in \mathfrak{h}$ hat die Eigenwerte* $0, \pm 1$ *mit den Eigenräumen* $\mathfrak{h}$, $\mathfrak{V}^\pm$ *in der adjungierten Darstellung, und H ist die Isotropiegruppe von Z bei der Operation von G auf* $\mathfrak{g}$ *durch* Ad.

Beweis. Die erste Aussage ist klar. Offenbar gilt $\mathrm{Ad}\,h \cdot Z = Z$ für $h \in H$. Umgekehrt folgt aus $\mathrm{Ad}\,g \cdot Z = Z$, daß $\mathrm{Ad}\,g$ die Eigenräume von $\mathrm{ad}\,Z$ invariant läßt und daher mit $\alpha = \mathrm{Ad}\,\gamma$ vertauscht. Also vertauscht g mit γ und die Behauptung folgt aus (c).

Aus Satz 1 und [14, 3.10] geht hervor, daß die durch $\psi_t = \mathrm{Int}(\gamma(t))$ definierte Operation von $\mathbb{R}_m$ auf G eine elementare Operation im Sinne von [14] ist. Insbesondere wird $\mathfrak{V} = (\mathfrak{V}^+, \mathfrak{V}^-)$ mit den quadratischen Abbildungen $Q_x y = \frac{1}{2}[x, [x, y]]$ ein reelles Jordan-Paar. Die Cartan-Involution θ definiert eine Involution τ von $\mathfrak{V}$, d.h., einen Vektorraumisomorphismus $\tau = \theta | \mathfrak{V}^+ \colon \mathfrak{V}^+ \to \mathfrak{V}^-$ mit $Q_{\tau x} = \tau Q_x \tau$, und damit auf $\mathfrak{V}^+$ die Struktur eines reellen Jordan-Tripelsystems mit der trilinearen Komposition

$$\{x\,y\,z\} = L(x, y) \cdot z = -[[x, \theta y], z]. \tag{7}$$

Die Jordan-Tripelidentitäten

$$\{x\,y\,z\} = \{z\,y\,x\}, \tag{JT1}$$

$$\{u\,v\{x\,y\,z\}\} - \{x\,y\{u\,v\,z\}\} = \{\{u\,v\,x\}\,y\,z\} - \{x\{v\,u\,y\}\,z\} \tag{JT2}$$

folgen leicht aus der Jacobi-Identität in $\mathfrak{g}$ und der Kommutativität von $\mathfrak{V}^+$.

Lemma 3. *Sei* $\mathfrak{g}$ *eine Lie-Algebra, die wie in* (1) *zerfällt und* (2) *erfüllt. Dann gilt für die Killingform* κ *von* $\mathfrak{g}$:

$$\kappa(\mathfrak{h}, \mathfrak{V}^\varepsilon) = \kappa(\mathfrak{V}^\varepsilon, \mathfrak{V}^\varepsilon) = 0, \tag{8}$$

$$\kappa(x, y) = \mathrm{Spur}(\mathrm{ad}\, x \cdot \mathrm{ad}\, y)|\, \mathfrak{V}^+ + \mathrm{Spur}(\mathrm{ad}\, y \cdot \mathrm{ad}\, x)|\, \mathfrak{V}^-, \tag{9}$$

$$\kappa(h, h') = \kappa_{\mathfrak{h}}(h, h') + \mathrm{Spur}(\mathrm{ad}\, h \cdot \mathrm{ad}\, h')|\, \mathfrak{V}^+ \oplus \mathfrak{V}^-. \tag{10}$$

Hier sei $x \in \mathfrak{V}^+$, $y \in \mathfrak{V}^-$, $h, h' \in \mathfrak{h}$, und $\kappa_{\mathfrak{h}}$ die Killingform von $\mathfrak{h}$.

Beweis. Siehe [8, III, § 3].

Nun ist θ eine Cartan-Involution, also $\kappa(x, \theta x) < 0$, und aus (9) folgt $\mathrm{Spur}\, L(x, x) = -\frac{1}{2}\kappa(x, \theta x) > 0$. Daher ist τ eine positive Involution von $\mathfrak{V}$ im Sinne von [12] und $\mathfrak{V}^+$ mit $\{\ \}$ ein kompaktes Jordan-Tripelsystem. Wir wollen zeigen, daß hierdurch eine Bijektion zwischen Isomorphieklassen symmetrischer R-Räume und kompakter Jordan-Tripelsysteme induziert wird. Für den Beweis ist es etwas bequemer, Jordan-Paare mit Involution statt Jordan-Tripelsysteme zu verwenden, was nach [11, 1.13] nichts ausmacht. Sei also $\mathfrak{V}$ ein reelles Jordan-Paar mit positiver Involution τ, und sei $\overline{X}$ die nach [13] zu $\mathfrak{V}$ assoziierte reelle algebraische Varietät. Die reellen Punkte von $\overline{X}$ sind dann die Äquivalenzklassen $(x:y)$ von Paaren $(x, y) \in \mathfrak{V}^+ \times \mathfrak{V}^-$ nach der in [13, 2.1] definierten Äquivalenzrelation. Wegen der Positivität von τ ist die Spurform von $\mathfrak{V}$ nicht ausgeartet und somit $\mathfrak{V}$ separabel. Nach [13, 2.2, 2.8, 4.6] ist $\overline{X}$ rational, projektiv, und homogen, und die Automorphismengruppe von $\overline{X}$ ist kanonisch isomorph zur projektiven Gruppe $\overline{G} = \mathbf{PG}(\mathfrak{V})$ (siehe [14, 5.14]). Seien die Untergruppen $\overline{U}^{\pm}$ und $\overline{H} = \mathbf{Aut}(\mathfrak{V})$ wie in [14, 5.15] definiert, und $\overline{\gamma}$: $\mathbb{R}_m \to \overline{H}$ durch $\overline{\gamma}(t) = (t \cdot \mathrm{Id},\ t^{-1} \cdot \mathrm{Id})$ gegeben. Durch τ wird ein involutorischer Automorphismus $\overline{\theta}$ von $\overline{G}$ vermöge

$$\overline{\theta}(\exp(x)) = \exp(\tau x), \qquad \overline{\theta}(\exp(y)) = \exp(\tau^{-1} y),$$
$$\overline{\theta}(h_+, h_-) = (\tau^{-1} h_- \tau, \tau h_+ \tau^{-1})$$

für $x \in \mathfrak{V}^+$, $y \in \mathfrak{V}^-$, $(h_+, h_-) \in \overline{H}$ induziert. Dies folgt aus [14, 4.14].

Satz 2. (a) *Die Fixpunktmenge $\overline{K}$ von $\overline{\theta}$ in $\overline{G}(\mathbb{R}) = \mathrm{Aut}(\overline{X})$ ist maximal kompakt. Es gibt genau eine symmetrische Raumstruktur auf $\overline{M} = \overline{X}(\mathbb{R})$ derart, daß $\overline{K} = \mathrm{Aut}(\overline{M}) \cap \mathrm{Aut}(\overline{X})$, und die Spiegelung am Punkt $\overline{o} = (0:0)$ durch $(x:y) \mapsto (-x: -y)$ gegeben ist. Hiermit ist $(\overline{M}, \overline{X})$ ein symmetrischer R-Raum, und das zugehörige Jordan-Paar mit Involution ist kanonisch isomorph zu $(\mathfrak{V}, \tau)$.*

(b) *Sei (M, X) ein symmetrischer R-Raum mit Jordan-Paar $(\mathfrak{V}, \tau)$. Dann ist die Abbildung $(x:y) \mapsto \exp y \cdot \exp x \cdot P$ ein Isomorphismus $(\overline{M}, \overline{X}) \to (M, X)$ der symmetrischen R-Räume.*

Beweis. (a) Die Lie-Algebra von $\overline{G}$ ist die Koecher-Tits-Algebra $\bar{\mathfrak{g}} = \mathfrak{V}^+ \oplus \mathrm{Der}(\mathfrak{V}) \oplus \mathfrak{V}^-$; insbesondere gelten (1) und (2). Nun folgt aus Lemma 3 und der Positivität von τ, daß $\overline{\theta}$ eine Cartan-Involution von $\bar{\mathfrak{g}}$ ist. Identifizieren wir $\overline{G}(\mathbb{R})$ mit einer offenen Untergruppe von $\mathrm{Aut}(\bar{\mathfrak{g}})$, so entspricht die Operation von $\overline{\theta}$ auf $\overline{G}(\mathbb{R})$ dem Transponiert-Inversen bezüglich der positiv definiten Form $-\kappa(x, \overline{\theta} x)$ auf $\bar{\mathfrak{g}}$. Folglich ist die Fixpunktmenge $\overline{K}$ von $\overline{\theta}$ in $\overline{G}(\mathbb{R})$ maximal kompakt. Die Isotropiegruppe von $\overline{o}$ bei der natürlichen Operation von $\overline{G}$ auf $\overline{X}$ ist die parabolische Untergruppe $\overline{P} = \overline{H} \cdot \overline{U}^-$. Daher gilt $\overline{G}(\mathbb{R}) = \overline{K} \cdot \overline{P}(\mathbb{R})$, insbesondere ist $\overline{K}$ transitiv auf $\overline{X}(\mathbb{R})$. Es gilt $\overline{\theta}(\overline{\gamma}(t)) = \overline{\gamma}(t^{-1})$; daher ist $\overline{\gamma}(-1) \in \overline{K}$. Die Isotropiegruppe von $\overline{o}$ in $\overline{K}$ ist $\overline{K} \cap \overline{P}(\mathbb{R}) \subset \overline{H}(R)$ und wird

daher von $\bar{\gamma}(-1)$ zentralisiert. Schließlich induziert $\bar{\gamma}(-1)$ auf dem Tangentialraum im Punkte $\bar{o}$ die Abbildung $-\mathrm{Id}$, denn $\bar{\gamma}(t)$ operiert durch $(x:y)\mapsto(t\,x:t^{-1}\,y)$. Folglich wird $\bar{X}(\mathbb{R})$ mit den Symmetrien $S_{k(\bar{o})}=k\cdot\bar{\gamma}(-1)\cdot k^{-1}$ $(k\in\bar{K})$ zu einem symmetrischen Raum $\bar{M}$, und aus der Konstruktion ist klar, daß das zu $(\bar{M},\bar{X})$ assoziierte Jordan-Paar mit Involution gerade $(\mathfrak{V},\tau)$ ist.

(b) Dies folgt aus [13, 4.3] und der Definition der symmetrischen Raumstruktur in (a).

Bemerkung. Ein kompakter *hermitescher* symmetrischer Raum N ist bekanntlich projektiv algebraisch, d.h., $N=Y(\mathbb{C})$ ist die Menge der komplexen Punkte einer komplexen projektiven Varietät Y. Durch Restriktion der Skalare erhält man hieraus einen symmetrischen R-Raum $(N,\prod_{\mathbb{C}/\mathbb{R}} Y)$. Die zugehörigen Jordan-Paare mit Involution sind genau die komplexen Jordan-Paare mit positiver hermitescher Involution (vgl. [12]). Im folgenden Sinne sind die symmetrischen R-Räume genau die reellen Formen der kompakten hermiteschen symmetrischen Räume: Ist N, Y wie oben und X eine reelle Form von Y (also $X\bigotimes_{\mathbb{R}}\mathbb{C}=Y$) mit der Eigenschaft, daß alle Symmetrien von N an Punkten von $X(\mathbb{R})\subset N$ Automorphismen von X sind, so ist $M=X(\mathbb{R})$ ein symmetrischer Unterraum von $N=X(\mathbb{C})$, und (M,X) ist ein symmetrischer R-Raum. Umgekehrt sei (M,X) ein symmetrischer R-Raum mit Jordan-Paar $(\mathfrak{V},\tau)$, und sei θ^* die antiholomorphe Fortsetzung von θ zu einem Automorphismus von $G(\mathbb{C})$. Dann ist die Fixpunktmenge K^* von θ^* in $G(\mathbb{C})$ maximal kompakt, operiert transitiv auf $X(\mathbb{C})$, und $X(\mathbb{C})$ wird mit den Symmetrien $S_{k(o)}=k\cdot S_o\cdot k^{-1}$ $(k\in K^*)$ ein hermitescher symmetrischer Raum. Das zugehörige Jordan-Paar ist $\mathfrak{V}_{\mathbb{C}}$ mit der komplex-antilinearen Fortsetzung von τ als Involution. Im allgemeinen besitzt ein kompakter hermitescher symmetrischer Raum mehrere nicht isomorphe reelle Formen (M,X) (vgl. [12, § 11]).

§ 2. Das Jordan-Tripelsystem im Tangentialraum

Wir wollen nun das Jordan-Tripelsystem eines symmetrischen R-Raumes im Tangentialraum $T_o(M)$ realisieren und geometrische Eigenschaften von M dadurch ausdrücken. Sei zunächst $M=K/K_o$ ein symmetrischer Raum und $\mathfrak{k}=\mathfrak{k}_o\oplus\mathfrak{m}$ wie in § 1. Wir betrachten $\mathfrak{k}$ als Lie-Algebra von Vektorfeldern auf M. Die Auswertungsabbildung $\xi\mapsto\xi(o)$ ist ein Vektorraumisomorphismus $\mathfrak{m}\to T_o(M)$, durch den $T_o(M)$ zu einem Lie-Tripelsystem $L(M)\cong\mathfrak{m}$ gemacht wird. Es gilt also $[u\,v\,w]=[[\tilde{u},\tilde{v}],\tilde{w}](o)$ für $u,v,w\in T_o(M)$, wobei $\tilde{u}$ das eindeutig bestimmte Vektorfeld in $\mathfrak{m}$ mit $\tilde{u}(o)=u$ ist.

Nun sei (M,X) ein symmetrischer R-Raum. Dann ist die Auswertungsabbildung $\mathfrak{V}^+\to T_o(M)$ ebenfalls ein Vektorraumisomorphismus, mit dem wir die Jordan-Tripelstruktur von $\mathfrak{V}^+$ auf $T_o(M)$ übertragen. Das so erhaltene Jordan-Tripelsystem werde mit $J(M,X)$ bezeichnet. Hier gilt also $\{u\,v\,w\}=-[[\hat{u},\hat{v}],\hat{w}](o)$, wobei $\hat{u}$ das eindeutig bestimmte Vektorfeld in $\mathfrak{V}^+$ mit $\hat{u}(o)=u$ ist. Man beachte, daß K_o auf $T_o(M)$ durch Automorphismen von $L(M)$ und $J(M,X)$ operiert.

Sei allgemein V ein kompaktes Jordan-Tripelsystem. Ein Rahmen in V ist ein maximales Orthogonalsystem von primitiven Tripotenten (vgl. [12, §3, §11]).

Lemma 4. *Der Vektorraum V wird mit $[u\,v\,w]=-\{u\,v\,w\}+\{v\,u\,w\}$ zu einem Lie-Tripelsystem V^L, und der von $e_1,\dots,e_r$ aufgespannte Untervektorraum ist maximal abelsch in V^L.*

Beweis. Die erste Aussage besteht in einer einfachen Verifikation der Identitäten eines Lie-Tripelsystems (vgl. [15]). Die zweite folgt aus den Eigenschaften der Peirce-Zerlegung von V (siehe [12, 3.15, 5.2, 11.6]).

Satz 3. *Das Lie-Tripelsystem eines symmetrischen R-Raumes (M,X) ist $L(M)=J(M,X)^L$. Ist $e_1,\dots,e_r$ ein Rahmen von $J(M,X)$, so ist $A=\mathrm{Exp}(\sum \mathbb{R}e_i)$ ein maximaler Torus von M mit Einheitsgitter*

$$\Gamma(A)=\sum \mathbb{Z}\,\pi\,e_i. \tag{11}$$

Beweis. Aus Satz 1 folgt, daß $\tilde{u}=\hat{u}+\theta(\hat{u})$ für $u\in T_o(M)$. Wegen der Kommutativität von $\mathfrak{B}^{\pm}$ hat man

$$[[\tilde{u},\tilde{v}],\tilde{w}]=[[\hat{u}+\theta\hat{u},\hat{v}+\theta\hat{v}],\hat{w}+\theta\hat{w}]=[[\hat{u},\theta\hat{v}],\hat{w}]-[[\hat{v},\theta\hat{u}],\hat{w}]$$
$$+\theta([[\hat{u},\theta\hat{v}],\hat{w}]-[[\hat{v},\theta\hat{u}],\hat{w}]),$$

und Auswertung im Punkte o liefert die erste Behauptung. Nach Lemma 4 ist $\mathfrak{A}=\sum \mathbb{R}\,e_i$ maximal abelsch in $L(M)$ und daher $A=\mathrm{Exp}(\mathfrak{A})$ ein maximaler Torus von M (vgl. [9, vol. II, p. 55, Lemma 1.1]). Für $x=\sum t_i e_i\in\mathfrak{A}$ gilt $\mathrm{Exp}(x)=(\exp\tilde{x})(o)=o$ dann und nur dann, wenn $\exp\tilde{x}\in K_o=K\cap P(\mathbb{R})=K\cap H(\mathbb{R})=\{k\in K\colon \mathrm{Ad}\,k\cdot Z=Z\}$ nach Kor. zu Satz 1. Durch Induktion zeigt man, daß

$$(\mathrm{ad}\,\tilde{x})^{2n}\cdot Z=(-1)^{n+1}\,2^{2n-1}\widehat{[x^{(2n-1)},\theta\hat{x}]},$$
$$(\mathrm{ad}\,\tilde{x})^{2n+1}\cdot Z=(-1)^{n+1}\,2^{2n}\widehat{(x^{(2n+1)}-\theta(x^{(2n+1)}))},$$

wobei die ungeraden Potenzen in einem Jordan-Tripelsystem durch $x^{(1)}=x$, $x^{(2n+1)}=\frac{1}{2}\{x,x,x^{(2n-1)}\}$ definiert sind. Wegen der Orthogonalität der e_i ist $x^{(2n+1)}=\sum t_i^{2n+1}e_i$ und damit folgt

$$\mathrm{Ad}\exp\tilde{x}\cdot Z=Z+\tfrac{1}{2}\sum(1-\cos 2t_i)[\hat{e}_i,\theta\hat{e}_i]$$
$$+\tfrac{1}{2}\sum(\sin 2t_i)(\theta\hat{e}_i-\hat{e}_i). \tag{12}$$

Also gilt $\mathrm{Ad}\exp\tilde{x}\cdot Z=Z$ dann und nur dann, wenn $t_i\in\pi\mathbb{Z}$.

In Umkehrung dieses Sachverhaltes beweisen wir nun

Satz 4. *Sei M ein kompakter symmetrischer Raum. Sei $V=(T_o(M),\{\ \})$ ein kompaktes Jordan-Tripelsystem auf $T_o(M)$ mit $V^L=L(M)$ und der Eigenschaft, daß für jeden Rahmen $e_1,\dots,e_r$ von V das Einheitsgitter des maximalen Torus $\mathrm{Exp}(\sum \mathbb{R}e_i)$ durch (11) gegeben ist. Dann gibt es genau eine symmetrische R-Raumstruktur X auf M mit $J(M,X)=V$.*

(Hierbei ist die Eindeutigkeit von X so zu verstehen: Gilt auch $V=J(M,X')$,

dann gibt es einen Isomorphismus $X \cong X'$, der auf $M = X(\mathbb{R}) = X'(\mathbb{R})$ die Identität ist).

Beweis. Zum Jordan-Tripelsystem V bilden wir das Jordan-Paar (V, V) mit der positiven Involution $\tau = \mathrm{Id}$, und den nach Satz 2 zugehörigen symmetrischen R-Raum $(\bar{M}, \bar{X})$. Dann ist $J(\bar{M}, \bar{X}) \cong V$, und ein Isomorphismus $\phi\colon V \to J(\bar{M}, \bar{X})$ der Jordan-Tripelsysteme ist nach Satz 3 wegen der Voraussetzung $V^L = L(M)$ ein Isomorphismus $\phi\colon L(M) \to L(\bar{M})$ der Lie-Tripelsysteme. Das Bild eines Rahmens in V ist ein Rahmen in $J(\bar{M}, \bar{X})$, und nach Satz 3 und der Voraussetzung über V und M ist ϕ ein Isomorphismus der Einheitsgitter der zugehörigen maximalen Tori A und $\bar{A}$. Also gibt es einen Isomorphismus $\Phi_A\colon A \to \bar{A}$ mit $T_o(\Phi_A) = \phi | L(A)$. Das untenstehende Lemma 5 zeigt, daß sich Φ_A eindeutig zu einem Isomorphismus $\Phi\colon M \to \bar{M}$ der symmetrischen Räume mit $T_o \Phi = \phi$ fortsetzen läßt. Durch Übertragen vermöge Φ^{-1} erhält man eine symmetrische R-Raumstruktur X mit $J(M, X) = V$.

Zum Beweis der Eindeutigkeit sei $J(M, X) = J(M, X') = V$, und seien $(\mathfrak{B}, \tau)$ und $(\mathfrak{B}', \tau')$ die zu X und X' gehörigen Jordan-Paare mit Involution. Es genügt zu zeigen, daß (bei Identifikation von $\mathfrak{B}^{\pm}$ und $\mathfrak{B}'^{\pm}$ mit Lie-Algebren von Vektorfeldern auf M) $\mathfrak{B}^{\pm} = \mathfrak{B}'^{\pm}$ und $\tau = \tau'$ gilt. Ist nämlich $(\bar{M}, \bar{X})$ der symmetrische R-Raum zu $(\mathfrak{B}, \tau) = (\mathfrak{B}', \tau')$ wie in Satz 2, und sind $\alpha\colon (\bar{M}, \bar{X}) \to (M, X)$ und $\alpha'\colon (\bar{M}, \bar{X}) \to (M, X')$ die Isomorphismen nach Satz 2(b), so liefert $\alpha' \circ \alpha^{-1}\colon X \to X'$ den gewünschten Isomorphismus. Es genügt sogar, $Z = Z'$ zu zeigen (mit Z wie im Kor. zu Satz 1). In der Tat ist dann

$$\mathfrak{B}^{\pm} = \mathfrak{B}'^{\pm} = \{\xi \pm [Z, \xi] : \xi \in \mathfrak{m}\},$$

und τ und τ' sind durch $\xi + [Z, \xi] \mapsto \xi - [Z, \xi]$ gegeben. Nun ist Z der Gradient der Funktion

$$f(k \cdot o) = \kappa(Z, \mathrm{Ad}\, k \cdot Z) \qquad (k \in K)$$

auf M bezüglich der Riemannschen Metrik $\langle\!\langle\ ,\ \rangle\!\rangle$, die auf $T_o(M)$ durch $\langle\!\langle u, v \rangle\!\rangle = -\kappa(\tilde{u}, \tilde{v})$ definiert ist (siehe [7]). Daher genügt es zu zeigen, daß sich die Funktion f und die Metrik $\langle\!\langle\ ,\ \rangle\!\rangle$ durch V ausdrücken lassen. Da f K_o-invariant ist, brauchen wir nur $f | A$ zu betrachten. Nach (12) ist

$$f(\mathrm{Exp}(\textstyle\sum t_i e_i)) = \kappa(Z, Z + \frac{1}{2}\sum (1 - \cos 2t_i)[\hat{e}_i, \theta \hat{e}_i] + \frac{1}{2}\sum \sin 2t_i (\theta \hat{e}_i - \hat{e}_i)).$$

Nun gilt $\kappa(Z, Z) = \mathrm{Spur}(\mathrm{ad}\, Z)^2 = 2 \dim M$, und $\kappa(Z, \theta \hat{e}_i - \hat{e}_i) = 0$ nach Lemma 3. Weiter ist

$$\kappa(Z, [\hat{e}_i, \theta \hat{e}_i]) = \kappa([Z, \hat{e}_i], \theta \hat{e}_i) = \kappa(\hat{e}_i, \theta \hat{e}_i) = -2\,\mathrm{Spur}\, L(e_i, e_i)$$

nur von V abhängig. Daher ist f durch V eindeutig bestimmt. Weiter ist nach Lemma 3

$$\langle\!\langle u, v \rangle\!\rangle = -\kappa(\tilde{u}, \tilde{v}) = -\kappa(\hat{u} + \theta \hat{u}, \hat{v} + \theta \hat{v}) = -2\kappa(\hat{u}, \theta \hat{v}) = 4\,\mathrm{Spur}\, L(u, v)$$

ebenfalls durch V eindeutig bestimmt. Damit folgt die Behauptung.

Lemma 5. *Seien M und N kompakte symmetrische Räume, ϕ: $L(M) \to L(N)$ ein Isomorphismus der Lie-Tripelsysteme, A und B maximale Tori von M und N und Φ_A: $A \to B$ ein Isomorphismus mit $T_o \Phi_A = \phi | L(A)$. Dann gibt es genau einen Isomorphismus Φ: $M \to N$ mit $T_o \Phi = \phi$.*

Für den Fall kompakter Lie-Gruppen an Stelle von kompakten symmetrischen Räumen siehe [3, § 2, Prop. 8]. Die dort verwendeten Methoden führen auch im Fall symmetrischer Räume zum Ziel, wenn man die Eigenschaften des Zentrums eines symmetrischen Raumes benützt (vgl. [9, vol. I, chap. IV]). Die Einzelheiten werden übergangen.

§ 3. Symmetrische Räume mit kubischem Einheitsgitter

Der Beweis von Satz 4 zeigt, daß die $\mathbb{Z}$-Basis $\pi e_1, \ldots, \pi e_r$ des Einheitsgitters bezüglich der Metrik $\langle\!\langle\ \rangle\!\rangle$ auf M orthogonal ist, denn für orthogonale Tripotente gilt $L(e_i, e_j) = 0$. Durch Abändern der Metrik läßt sich noch mehr erreichen:

Satz 5. *Sei (M, X) ein symmetrischer R-Raum. Dann gibt es genau eine unter allen Symmetrien invariante Riemannsche Metrik g auf M, sodaß $g(e_i, e_j) = \delta_{ij}$ für jeden Rahmen $e_1, \ldots, e_r$ von $J(M, X)$.*

Beweis. Sei $e_1, \ldots, e_r$ ein Rahmen von $J(M, X)$. Definiere die quadratische Form q auf $\mathfrak{U} = \sum \mathbb{R} e_i$ durch $q(\sum t_i e_i) = \sum t_i^2$. Sei W die Weylgruppe von M bezüglich A. Da K_o durch Automorphismen von $J(M, X)$ operiert, ist das auch für die Operation von W auf dem Unter-Tripelsystem $\mathfrak{U}$ von $J(M, X)$ der Fall. Nun sind die e_i die bis auf Reihenfolge und Vorzeichen eindeutig bestimmten primitiven Tripotente von $\mathfrak{U}$. Also operiert W auf $\mathfrak{U}$ durch Permutation und Vorzeichenwechsel der e_i und läßt daher q invariant. Nach dem Satz von Chevalley (vgl. [6, Ch. X, Th. 6.10]) läßt sich q eindeutig zu einer K_o-invarianten quadratischen Form auf $T_o(M)$ fortsetzen, und diese liefert durch Verschieben mit K eine unter allen Symmetrien invariante Metrik g, denn eine K-invariante Metrik ist bereits unter allen Symmetrien invariant. Ist $e'_1, \ldots, e'_r$ ein weiterer Rahmen, so gibt es ein $k \in K_o$ mit $k(e_i) = \pm e'_{\pi(i)}$ mir einer Permutation π. Dies folgt aus der Konjugiertheit der maximalen Tori mit einem ähnlichen Schluß wie oben. Also sind die e'_i ebenfalls bezüglich g orthonormal.

Wir nennen g die *kanonische Metrik* von (M, X).

Bemerkung. Sei T die generische Spur des durch $J(M, X)$ definierten Jordan-Paares $\mathfrak{V}$ (vgl. [11, § 16]). Ist $\mathfrak{V}$ einfach und reduziert (im Sinne von [12, 11.9]), so gilt $T(e_i, e_i) = 1$ und die kanonische Metrik ist durch T gegeben. Ist dagegen $\mathfrak{V}$ einfach und nicht reduziert, so ist $T(e_i, e_i) = 2$, und die kanonische Metrik ist $\frac{1}{2} T$. Indem man $J(M, X)$ in einfache Ideale zerlegt, erhält man so einen weiteren Beweis von Satz 5.

Nach Satz 3 und Satz 5 ist das Einheitsgitter eines symmetrischen R-Raumes bezüglich der kanonischen Metrik ein *kubisches Gitter*, d.h., es besitzt eine $\mathbb{Z}$-Basis aus orthogonalen Vektoren gleicher Länge. Bemerkenswerterweise ist diese Eigenschaft schon charakteristisch für symmetrische R-Räume:

Satz 6. *Sei M ein kompakter symmetrischer Raum und sei g eine unter allen Spiegelungen invariante Riemannsche Metrik auf M, bezüglich der die Einheitsgitter der maximalen Tori von M kubisch sind. Dann gibt es genau eine symmetrische R-Raumstruktur X auf M derart, daß g bis auf einen Faktor die kanonische Metrik von (M, X) ist.*

Beweis. Nach Änderung von g um einen Faktor kann man annehmen, daß ein maximal abelscher Unterraum $\mathfrak{A}$ von $\mathfrak{M} := L(M)$ eine Orthonormalbasis $e_1, \ldots, e_r$ besitzt und das Einheitsgitter des maximalen Torus $A = \operatorname{Exp} \mathfrak{A}$ durch $\Gamma(A) = \sum \mathbb{Z} \pi e_i$ gegeben ist. Wir bezeichnen das durch g auf $T_o(M)$ definierte Skalarprodukt mit $(\ ,\ )$. Nach Satz 4 genügt es zu zeigen: Es gibt genau eine kompakte Jordan-Tripelstruktur $\{\ \}$ auf dem Vektorraum $\mathfrak{M}$ mit

$$[u\,v\,w] = -\{u\,v\,w\} + \{v\,u\,w\}, \tag{13}$$

$$e_1, \ldots, e_r \quad \text{ist ein Rahmen.} \tag{14}$$

Wir beweisen zuerst die Eindeutigkeit eines solchen Jordan-Tripelsystems. Es genüge also $\{\ \}$ den Bedingungen (13 und (14). Nach Satz 4 ist dann $(T_o(M), \{\ \}) = J(M, X)$, insbesondere operiert K_o durch Automorphismen von $\{\ \}$. Definiere

$$\langle x\,y\,z\rangle = \{x\,y\,z\} + \{y\,z\,x\} + \{z\,x\,y\}. \tag{15}$$

Dann ist $\langle x\,y\,z\rangle$ vollkommen symmetrisch in x, y, z und $6x^{(3)} = \langle x\,x\,x\rangle$, also ist $\langle x\,y\,z\rangle$ durch die kubische Abbildung $c(x) = x^{(3)}$ eindeutig bestimmt. Aus (13) und (15) folgt leicht

$$3\{x\,y\,z\} = \langle x\,y\,z\rangle + [y\,x\,z] + [y\,z\,x]. \tag{16}$$

Daher ist $\{\ \}$ durch das Lie-Tripelprodukt und die kubische Abbildung c eindeutig bestimmt. Nun ist $\mathfrak{M}$ die Vereinigung der $k \cdot \mathfrak{A}$ $(k \in K_o)$, und daher ist c schon durch seine Einschränkung auf $\mathfrak{A}$ festgelegt. Für $x = \sum t_i e_i$ ist aber wegen (14) notwendig $x^{(3)} = \sum t_i^3 e_i$.

Zum Beweis der Existenz wird die Wurzelraumzerlegung von $\mathfrak{M}$ benützt, an die deshalb kurz erinnert sei (siehe [9, vol. II, Ch. VI]). Für jede Linearform λ auf $\mathfrak{A}$ sei

$$\mathfrak{M}_\lambda = \{y \in \mathfrak{M} \colon [y\,x\,x] = -\lambda(x)^2\, y, \text{ für alle } x \in \mathfrak{A}\}. \tag{17}$$

Das Wurzelsystem von (M, A) ist dann $R = \{0 \neq \lambda \in \mathfrak{A}^* \colon \mathfrak{M}_\lambda \neq 0\}$. Es gilt die orthogonale Zerlegung

$$\mathfrak{M} = \mathfrak{A} \oplus \sum_{\lambda \in R} \mathfrak{M}_\lambda. \tag{18}$$

Wir identifizieren $\mathfrak{k}_o$ mit einer Unteralgebra der Derivationsalgebra des Lie-Tripelsystems $\mathfrak{M}$ vermöge der linearen Isotropiedarstellung. Dann gibt es zu jedem $y \in \mathfrak{M}_\lambda$ genau ein $D_y \in \mathfrak{k}_o$ mit

$$D_y(x) = \lambda(x)\, y, \tag{19}$$

$$[x\,y\,z] = \lambda(x)\, D_y(z) \tag{20}$$

für alle $x \in \mathfrak{A}$ und $z \in \mathfrak{M}$ (in der Terminologie von [9] sind D_y und y „related"). Man bemerke, daß die hier definierten Wurzeln zur Vereinfachung der Formeln die π-fachen der in [9] verwendeten sind.

Lemma 6. *Es gibt genau eine K_o-invariante quartische Form p auf $\mathfrak{M}$, die für $x = \sum t_i e_i \in \mathfrak{A}$ den Wert $p(x) = \frac{1}{4}\sum t_i^4$ annimmt. Der Gradient von p bezüglich $(\ ,\ )$ ist eine kubische Abbildung $c\colon x \mapsto x^{(3)}$ von $\mathfrak{M}$ in sich, die auf $\mathfrak{A}$ durch*

$$x^{(3)} = \sum t_i^3 e_i \tag{21}$$

gegeben ist.

Beweis. Die Weylgruppe W läßt das Skalarprodukt und das Gitter $\Gamma(A)$ invariant. Daher ist die Matrix eines Elements von W bezüglich der Basis $e_1, \ldots, e_r$ von $\mathfrak{A}$ orthogonal und ganzzahlig, und somit operiert W auf $\mathfrak{A}$ durch Permutation und Vorzeichenwechsel der e_i. Insbesondere ist p invariant unter W und läßt sich daher nach dem Satz von Chevalley zu einem K_o-invarianten Polynom p auf $\mathfrak{M}$ fortsetzen. Der Gradient $c = \operatorname{grad} p$ ist eine kubische Abbildung, und $c(x)$ steht senkrecht auf den Flächen $p = \text{const}$, also insbesondere wegen der K_o-Invarianz von p auf den Bahnen von K_o in $\mathfrak{M}$. Sei $x \in \mathfrak{A}$ mit $\lambda(x) \neq 0$ für alle $\lambda \in R$. Dann zeigt (19), daß der Tangentialraum an die Bahn von x unter K_o alle $\mathfrak{M}_\lambda$ enthält. Also liegt $c(x)$ im orthogonalen Komplement der $\mathfrak{M}_\lambda$ und das ist $\mathfrak{A}$. Nun ist (21) klar, denn die e_i sind orthonormal. Da die Menge der x mit der $\lambda(x) \neq 0$ für alle $\lambda \in R$ dicht in $\mathfrak{A}$ liegt, gilt (21) für alle $x \in \mathfrak{A}$.

Es sei $\langle x y z \rangle$ die durch Linearisieren von $x^{(3)}$ erhaltene vollkommen symmetrische Trilinearform mit

$$6x^{(3)} = \langle x x x \rangle. \tag{22}$$

Wir definieren nun das ternäre Produkt $\{x y z\}$ durch (16) und haben zu zeigen: die Jordan-Tripelidentitäten (JT1) und (JT2) sind erfüllt, die Spurform $\operatorname{Spur}(z \mapsto \{x x z\})$ ist positiv definit, und es gilt (13) und (14). Mit Hilfe der Jacobi-Identität für das Lie-Tripelprodukt folgt (13) leicht aus der Definition. Ebenfalls klar nach Definition ist (JT1). Die Verifikation von (JT2) ist dagegen langwierig.

Lemma 7. *Seien $\omega_i \in \mathfrak{A}^*$ durch $\omega_i(e_j) = \delta_{ij}$ definiert. Dann ist*

$$R \subset \{\pm\omega_i, \pm 2\omega_i, \pm\omega_i \pm \omega_j (i \neq j)\}.$$

Beweis. Für $\lambda \in R$ gilt $\lambda(\Gamma(A)) \subset \pi\mathbb{Z}$, denn $\operatorname{Exp} x \mapsto e^{2i\lambda(x)}$ ist ein wohldefinierter Homomorphismus $A \to S^1$ ([9, vol. II, Prop. 1.3, p. 59]). Ferner ist die inverse Wurzel $\lambda^* = \pi \cdot (2\vec{\lambda}/(\vec{\lambda}, \vec{\lambda})) \in \Gamma(A)$ ([9, p. 66f]), wobei $\vec{\lambda} \in \mathfrak{A}$ durch $(\vec{\lambda}, x) = \lambda(x)$ definiert sei. Schreiben wir $\lambda = \sum n_i \omega_i$, so bedeutet dies wegen $\vec{\omega}_i = e_i$ und $\Gamma(A) = \sum \mathbb{Z}\pi e_i$, daß $n_i \in \mathbb{Z}$ und $2n_i/\sum n_j^2 \in \mathbb{Z}$. Hieraus folgt durch elementare Überlegungen leicht, daß für λ nur die aufgezählten Möglichkeiten bestehen.

Wir setzen $\omega_0 := 0$ und definieren Untervektorräume V_{ij}^ε ($\varepsilon = \pm$, $0 \leqq i, j \leqq r$) von $\mathfrak{M}$ durch $V_{ii}^+ = \mathbb{R} e_i$, $V_{00}^\varepsilon = 0$, und $V_{ij}^\varepsilon = \mathfrak{M}_{\omega_i - \varepsilon\omega_j}$ in allen anderen Fällen. Man beachte, daß $V_{ij}^\varepsilon = V_{ji}^\varepsilon$ wegen $\mathfrak{M}_\lambda = \mathfrak{M}_{-\lambda}$. Insbesondere gilt also $V_{ii}^- = \mathfrak{M}_{2\omega_i}$ und $V_{i0} := V_{i0}^\varepsilon = \mathfrak{M}_{\omega_i}$. Nach (18) ist $\mathfrak{M}$ die Summe der V_{ij}^ε.

Lemma 8. *Für* $x=\sum t_i e_i \in \mathfrak{A}$ *und* $y \in V_{ij}^{\varepsilon}$ *gelten die Formeln*

$$\langle x\,x\,y\rangle = 2(t_i^2+\varepsilon t_i t_j+t_j^2)\,y, \tag{23}$$

$$\{x\,x\,y\} = (t_i^2+t_j^2)\,y, \tag{24}$$

$$\{x\,y\,x\} = 2\varepsilon t_i t_j\,y. \tag{25}$$

(Hier sei $t_0 := 0$).

Beweis. K_o operiert durch Automorphismen und daher $\mathfrak{k}_o$ durch Derivationen von $\langle\ \rangle$. Sei $y \in \mathfrak{M}_\lambda$ und $D_y \in \mathfrak{k}_o$ wie in (19). Durch Anwenden von D_y auf (22) folgt dann wegen (21)

$$6\lambda(x^{(3)})\,y = 6D_y(x^{(3)}) = 3\langle x, x, D_y x\rangle = 3\lambda(x)\langle x\,x\,y\rangle.$$

Für $\lambda = \omega_i - \varepsilon\omega_j$ liefert dies wegen (21)

$$\langle x\,x\,y\rangle = \frac{2(t_i^3-\varepsilon t_j^3)}{t_i-\varepsilon t_j}\,y = 2(t_i^2+\varepsilon t_i t_j+t_j^2)\,y.$$

Das gilt auch noch für $y \in V_{ii}^{+}$, was man durch Linearisieren von (21) sieht. Nach Definition von $\{\ \}$ ist $3\{x\,x\,y\} = \langle x\,x\,y\rangle - [y\,x\,x]$ und $3\{x\,y\,x\} = \langle x\,y\,x\rangle + 2[y\,x\,x]$, und nach (17) $[y\,x\,x] = -\lambda(x)^2\,y = -(t_i-\varepsilon t_j)^2\,y$. Auch dies bleibt für $y \in V_{ii}^{+}$ bestehen, weil $\mathfrak{A}$ abelsch ist. Hieraus folgen nun (24) und (25) durch leichte Rechnung.

Korollar. $\mathrm{Spur}(z \mapsto \{x\,x\,z\})$ *ist positiv definit.*

Beweis. Wegen der K_o-Invarianz genügt es, dies für $x \in \mathfrak{A}$ zu zeigen, und dann ist nach (24) $\mathrm{Spur}(z \mapsto \{x\,x\,z\}) \geqq \mathrm{Spur}_{\mathfrak{A}}(z \mapsto \{x\,x\,z\}) = 2\sum t_i^2$.

Unser Ziel ist nun der Nachweis der schwachen Assoziativität

$$\{x\,y\,\{x\,z\,x\}\} = \{\{x\,y\,x\}\,z\,x\}, \tag{JT3}$$

aus der nach den unten bewiesenen Lemmata 10 und 11 die Jordan-Tripel-Identität (JT2) folgen wird. Da K_o durch Automorphismen von $\{\ \}$ operiert und $\mathfrak{M}$ die Vereinigung der $k \cdot \mathfrak{A}$ $(k \in K_o)$ ist, genügt es, dies für $x = \sum t_i e_i \in \mathfrak{A}$ zu zeigen. Falls y und z in $\mathfrak{A}$ liegen, ist (JT3) klar. Falls $z = \sum s_k e_k \in \mathfrak{A}$ und $y \in \mathfrak{M}_\lambda$, etwa $\lambda = \omega_i - \varepsilon\omega_j$, bekommt man durch Linearisieren von (24) und (25) wegen $\{x\,z\,y\} - \{z\,x\,y\} = -[x\,z\,y] = 0$ die Formeln

$$\{x\,z\,y\} = (s_i t_i + s_j t_j)\,y,\ \{x\,y\,z\} = \varepsilon(s_i t_i + s_j t_j)\,y.$$

Da ferner $\{x\,z\,x\} = 2\sum s_k t_k^2 e_k$, folgt hieraus leicht (JT3). Schließlich seien $y \in \mathfrak{M}_\lambda$, $z \in \mathfrak{M}_\mu$, und D_y und D_z wie in (19). Dann erhält man unter Benutzung der Jacobi-Identität und (20):

$$\begin{aligned} 3\{x\,y\,z\} &= \langle x\,y\,z\rangle + [y\,x\,z] + [y\,z\,x] \\ &= \langle x\,y\,z\rangle - 2[x\,y\,z] + [x\,z\,y] \\ &= \langle x\,y\,z\rangle - 2\lambda(x)\,D_y z + \mu(x)\,D_z y. \end{aligned} \tag{26}$$

Sei nun $\lambda=\omega_i-\varepsilon\omega_j$. Zur Berechnung von $\langle x\,y\,z\rangle$ wende man auf die Gleichung (23) die Derivation D_z an. Das ergibt wegen der Symmetrie von $\langle x\,y\,z\rangle$:

$$\langle x,x,D_z\,y\rangle+2\langle x,y,D_z\,x\rangle=\langle x,x,D_z\,y\rangle+2\mu(x)\langle x\,y\,z\rangle=2(t_i^2+\varepsilon\,t_i\,t_j+t_j^2)\,D_z\,y,$$

also

$$\langle x\,y\,z\rangle=\frac{1}{\mu(x)}((t_i^2+\varepsilon\,t_i\,t_j+t_j^2)\,D_z\,y-\tfrac{1}{2}\langle x,x,D_z\,y\rangle). \tag{27}$$

Lemma 9. *Sei* $y\in\mathfrak{M}_\lambda$ *und* $z\in\mathfrak{M}_\mu$. *Dann ist* $D_y\,z+D_z\,y\in\mathfrak{M}_{\lambda-\mu}$ *und* $D_y\,z-D_z\,y\in\mathfrak{M}_{\lambda+\mu}$.

Beweis. Wende auf die Gleichung $[z\,x\,x]=-\mu(x)^2\,z$ die Derivation D_y an und benütze (19) und (20) sowie die Jacobi-Identität:

$$\begin{aligned}-\mu(x)^2\,D_y\,z&=[D_y\,z,x,x]+[z,D_y\,x,x]+[z,x,D_y\,x]\\&=[D_y\,z,x,x]+\lambda(x)([z\,y\,x]+[z\,x\,y])\\&=[D_y\,z,x,x]+\lambda(x)([x\,y\,z]+2[z\,x\,y])\\&=[D_y\,z,x,x]+\lambda(x)(\lambda(x)\,D_y\,z-2\mu(x)\,D_z\,y).\end{aligned}$$

Addition und Subtraktion der durch Vertauschen von y und z erhaltenen Gleichung ergibt

$$[D_y\,z\pm D_z\,y,x,x]=-(\lambda(x)\mp\mu(x))^2(D_y\,z\pm D_z\,y).$$

Das ist wegen (17) die Behauptung.

Zum Nachweis von (JT3) unterscheiden wir nun mehrere Fälle. Wir nennen eine Wurzel λ kurz, mittel, oder lang, je nachdem, ob $(\lambda,\lambda)=1,2,4$ ist. Es sei $y\in\mathfrak{M}_\lambda$ und $z\in\mathfrak{M}_\mu$. Nach (25) ist $\{x\,y\,x\}$ (bzw. $\{x\,z\,x\}$) ein Vielfaches von y (bzw. z).

(a) λ und μ sind kurz: Dann ist etwa $y\in V_{i0}$, $z\in V_{k0}$, und nach (25) gilt $\{x\,y\,x\}=0=\{x\,z\,x\}$.

(b) $\lambda=\mu$ ist mittel oder lang: Dann genügt es wegen (25), $\{x\,y\,z\}=\{x\,z\,y\}$ zu zeigen. Es ist $2\lambda\notin R$, denn $(2\lambda,2\lambda)=4(\lambda,\lambda)>4$. Also ist $\mathfrak{M}_{2\lambda}=0$ und nach Lemma 9 ist $D_y\,z=D_z\,y$. Nun zeigt (26) die Behauptung.

(c) $0\neq\lambda\pm\mu\notin R$: Dann ist $\mathfrak{M}_{\lambda\pm\mu}=0$ und wegen (26) und (27) ist $\{x\,y\,z\}=\{x\,z\,y\}=0$, also verschwinden beide Seiten von (JT3).

Eine Aufzählung aller Möglichkeiten zeigt nun, daß nach Eliminierung der durch (a), (b), (c) erledigten Fälle bis auf Permutation und Vorzeichenwechsel nur noch die folgenden verbleiben:

(d) $\lambda=\omega_1$, $\mu=2\omega_1$;

(e) $\lambda=\omega_1$, $\mu=\omega_1-\omega_2$;

(f) $\lambda=2\omega_1$, $\mu=\omega_1-\omega_2$;

(g) $\lambda=\omega_1+\omega_2$, $\mu=\omega_1-\omega_2$;

(h) $\lambda=\omega_1+\omega_2$, $\mu=\omega_2+\omega_3$.

In jedem dieser Fälle zeigt eine direkte Rechnung unter Verwendung von (23), (26), (27) sowie Lemma 9, daß (JT3) gilt.

Lemma 10. *Das Skalarprodukt* $(\ ,\ )$ *ist bezüglich* $\{\ \}$ *assoziativ:*

$$(\{x\,y\,z\},v)=(x,\{y\,z\,v\}). \tag{28}$$

Beweis. Nach Lemma 6 ist $(x^{(3)},v)=\frac{1}{6}(\langle x\,x\,x\rangle,v)$ die Richtungsableitung von p in Richtung v. Folglich ist $(\langle x\,y\,z\rangle,v)$ die durch vierte Ableitung von p erhaltene vollkommen symmetrische Quadrilinearform. Weiter gehört die Abbildung $z\mapsto[x\,y\,z]$ zu $\mathfrak{k}_o$ und ist daher schiefsymmetrisch. Hieraus und aus der Jacobi-Identität folgt nach einiger Rechnung (vgl. [9, vol. I, p. 146]) $([y\,x\,z],v)=(x,[z\,v\,y])$. Nun bekommt man (28) leicht aus (16).

Lemma 11. *Sei* $\{x\,y\,z\}$ *eine trilineare Komposition auf einem Vektorraum* V *über einem Körper der Charakteristik* $\neq 2,3$, *die* (JT1) *und* (JT3) *erfüllt. Falls* V *eine nichtausgeartete, im Sinne von* (28) *assoziative symmetrische Bilinearform besitzt, dann ist* V *ein Jordan-Tripelsystem.*

Beweis. Wir setzen $L_{x,y}z=P_{x,z}y=\{x\,y\,z\}$ und $P_x=\frac{1}{2}P_{x,x}$, und bezeichnen mit A^* das Transponierte von $A\in\operatorname{End}V$ bezüglich $(\ ,\ )$. Dann ist $L_{x,y}^*=L_{y,x}$ nach (28), und $2(P_x y,z)=(\{x\,y\,x\},z)=(x,\{y\,x\,z\})=(\{x\,z\,x\},y)=2(y,P_x z)$ wegen der Symmetrie von $(\ ,\ )$. Also gilt $P_x^*=P_x$ und durch Linearisieren auch $P_{x,y}^*=P_{x,y}$. Die Identität (JT3) ist äquivalent mit $L_{x,y}P_x=P_{P_x y,y}$. Durch Übergang zum Transponierten folgt also $L_{x,y}P_x=P_x L_{y,x}$ oder

$$\{x\,y\,\{x\,z\,x\}\}=\{x\,\{y\,x\,z\}\,x\}. \tag{29}$$

Durch vollständiges Linearisieren von (29) erhält man

$$\{u\,y\,\{x\,z\,v\}\}+\{v\,y\,\{u\,z\,x\}\}+\{x\,y\,\{u\,z\,v\}\}=\{u\,\{y\,v\,z\}\,x\}+\{u\,\{y\,x\,z\}\,v\}+\{x\,\{y\,u\,z\}\,v\}.$$

Als lineare Abbildung in v betrachtet besagt dies

$$L_{u,y}L_{x,z}+L_{\{xzu\},y}+L_{x,y}L_{u,z}=P_{u,x}P_{y,z}+L_{u,\{yxz\}}+L_{x,\{yuz\}}. \tag{30}$$

Durch Übergang zum Transponierten und Vertauschung von y, z mit u, x folgt

$$L_{x,z}L_{u,y}+L_{u,\{zxy\}}+L_{x,y}L_{u,z}=P_{u,x}P_{y,z}+L_{\{uzx\},y}+L_{\{uyx\},z}. \tag{31}$$

Subtrahiere (31) von (30):

$$[L_{u,y},L_{x,z}]=2L_{u,\{yxz\}}+L_{x,\{yuz\}}-2L_{\{uzx\},y}-L_{\{uyx\},z}. \tag{32}$$

Übergang zum Transponierten und Vertauschung von z, x mit u, y ergibt:

$$[L_{u,y},L_{x,z}]=2L_{\{uyx\},z}+L_{\{xzu\},y}-2L_{x,\{yuz\}}-L_{u,\{zxy\}}. \tag{33}$$

Durch zweimalige Addition von (33) zu (32) bekommt man

$$3[L_{u,y},L_{x,z}]=3L_{\{uyx\},z}-3L_{x,\{yuz\}}$$

und das ist die Behauptung.

Zum Abschluß des Beweises von Satz 6 bleibt nur noch zu zeigen, daß $e_1, \ldots, e_r$ ein Rahmen des Jordan-Tripelsystems $(T_o(M), \{\ \})$ ist. Lemma 8 zeigt, daß die e_i orthogonale Tripotente mit Peirce-Räumen V_{ij}^{ε} sind. Wegen $V_{ii}^{+} = \mathbb{R} e_i$ sind die e_i primitiv, und wegen $V_{00} = 0$ maximal. Damit ist Satz 6 bewiesen.

Literatur

1. Borel, A., Tits, J.: Groupes réductifs. Publ. Math. IHES No. **27**, 55–150 (1965)
2. Borel, A., Serre, J.-P.: Corners and arithmetic groups. Comm. Math. Helv. **48**, 436–491 (1973)
3. Bourbaki, N.: Groupes et Algèbres de Lie. Chapitre IX. Paris: Masson 1982
4. Demazure, M.: Automorphismes et déformations des variétés de Borel. Invent. Math. **39**, 179–186 (1977)
5. Ferus, D.: Symmetric submanifolds of Euclidean space. Math. Ann. **247**, 81–93 (1980)
6. Helgason, S.: Differential Geometry and Symmetric Spaces. New York: Academic Press 1962
7. Kobayashi, S., Takeuchi, M.: Minimal imbeddings of R-spaces. J. Differ. Geom. **2**, 203–215 (1968)
8. Koecher, M.: Gruppen und Lie-Algebren von rationalen Funktionen. Math. Z. **109**, 349–392 (1969)
9. Loos, O.: Symmetric Spaces. Amsterdam: W.A. Benjamin 1969
10. Loos, O.: Jordan triple systems, R-spaces, and bounded symmetric domains. Bull. Am. Math. Soc. **77**, 558–561 (1971)
11. Loos, O.: Jordan Pairs. Springer Lecture Notes No. 460. Berlin-Heidelberg-New York: Springer 1975
12. Loos, O.: Bounded symmetric domains and Jordan pairs. Lecture Notes, The University of California at Irvine, 1977
13. Loos, O.: Homogeneous algebraic varieties defined by Jordan pairs. Monatsh. Math. **86**, 107–129 (1978)
14. Loos, O.: On algebraic groups defined by Jordan pairs. Nagoya Math. J. **74**, 23–66 (1979)
15. Meyberg, K.: Zur Konstruktion von Lie-Algebren aus Jordan-Tripelsystemen. Manuscr. Math. **3**, 115–132 (1970)
16. Takeuchi, M.: Cell decompositions and Morse equalities on certain symmetric spaces. J. Fac. Sci. Univ. Tokyo, Sect. IA **12**, 81–192 (1965)

Eingegangen am 20. März 1984

Math. Z. 189, 227–237 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Regularity and Finiteness of Solutions to the Free Boundary Problem for Minimal Surfaces

Hans Wilhelm Alt[1] and Friedrich Tomi[2]

[1] Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, D-5300 Bonn 1, Federal Republic of Germany
[2] Mathematisches Institut, Universität Heidelberg, Im Neuenheimer Feld 288, D-6900 Heidelberg, Federal Republic of Germany

Dedicated to E. Heinz on occasion of his sixtieth birthday

0. Introduction

The free boundary value problem for minimal surfaces as it was posed and also solved by Courant [2] consists in the following task. Let S be a compact surface embedded in $\mathbb{R}^3$ such that not all components of $\mathbb{R}^3 \setminus S$ are simply connected. Let then a Jordan curve Γ in $\mathbb{R}^3 \setminus S$ be chosen which is not contractible in $\mathbb{R}^3 \setminus S$. One searches then for a disc type surface whose boundary curve lies on S and is linked with Γ and which has minimal area in the class of all such surfaces. One can also consider the corresponding minimization problem with constraint in which the admissible surfaces are restricted to the component of $\mathbb{R}^3 \setminus S$ determined by Γ. More generally, one can replace such a component by a threedimensional bordered Riemannian manifold M, the boundary ∂M taking the role of S. The constrained problem has been investigated by Meeks and Yau [9] and by the second author [11].

Whenever one knows that the boundary of a minimizing surface does not intersect its interior then the topological methods of Meeks and Yau yield the embedded character of this surface, undoubtful an optimal regularity result. In the constrained problem intersection of interior and boundary portions of a minimizing surface can be ruled out by suitable curvature conditions on the supporting boundary, like convexity or H-convexity [9, 11]. Such conditions are however not applicable in the unconstrained problem and they are never fulfilled in the exterior problem where a minimizing surface is to be determined in the unbounded component of $\mathbb{R}^3 \setminus S$, the physically most realistic situation. Even if one would discard all those solutions whose boundary intersects the interior as physically unrealistic there were still situations in which the class of all mathematically possible solutions has to be considered. This is e.g. the case when one investigates the question of the finiteness of the number of solutions.

Under the sole assumption of analycity of S we shall therefore prove the regularity, i.e. the immersed character of minimizing solutions without any

Supported by Deutsche Forschungsgemeinschaft, SFB 72

additional conditions (Theorem 1). Our proof is based on Osserman's result on the non-existence of true interior branch points in minimizing surfaces [10] and also uses several ideas of Gulliver [5] and of the first author [1] developed in connection with the regularity proof for solutions to Plateau's problem. We remark that under the same analycity hypothesis Hildebrandt and Nitsche [6] have already shown the regularity at the boundary.

On the basis of our regularity result we shall demonstrate that any analytic surface S can support only a finite number of minimizing solutions to the free boundary value problem which intersect the unbounded component of $\mathbb{R}^3 \setminus S$ (Theorem 2). The problem for solutions completely contained inside a bounded component of $\mathbb{R}^3 \setminus S$ has been treated by the second author [11], provided S satisfies a curvature condition.

1. The Regularity of Minimizing Surfaces

For the rest of the paper we assume that S is a compact analytic surface in $\mathbb{R}^3$ and Γ a closed Jordan curve in $\mathbb{R}^3 \setminus S$ which is not contractible in $\mathbb{R}^3 \setminus S$. The class of surfaces admissible in our variational problem will be denoted by $\mathscr{C}(S, \Gamma)$ and consists of all mappings f from the closed unit disc B in $\mathbb{R}^2$ into $\mathbb{R}^3$ of class $C^0(B) \cap H_2^1(B)$ such that $f|\partial B$ is a loop on S which is not contractible in $\mathbb{R}^3 \setminus \Gamma$. Each surface $f \in \mathscr{C}(S, \Gamma)$, $f = f(u_1, u_2)$, which is stationary for Dirichlet's integral on $\mathscr{C}(S, \Gamma)$ is a generalized conformal minimal surface orthogonal to S along $f(\partial B)$. In particular, f satisfies the system

$$f_{u_1} \cdot f_{u_2} = 0, \; |f_{u_1}| = |f_{u_2}| \geqq 0, \tag{1}$$

$$\Delta f = 0. \tag{2}$$

Points $z = u_1 + iu_2 \in B$ where f fails to be immersed are called branch points of f; they are exactly the zeros of the complex analytic vector function

$$f_z = \tfrac{1}{2}(f_{u_1} - if_{u_2}).$$

For non-constant maps f branch points are clearly isolated and have a well defined order which by definition is the order of the corresponding zero of f_z. It can easily be seen that the unit normal of f is real analytic even at a branch point and therefore there is a well defined tangent plane also at branch points. Let P denote the orthogonal projection of $\mathbb{R}^3$ onto the tangent plane of f at a branch point z_0, identified with the complex plane. Then the map $\Phi := P \cdot f$ has the expansion

$$\Phi(z) = a(z - z_0)^m + O(|z - z_0|^{m+1})$$

where $a \neq 0$ and $m - 1$ is the order of the branch point z_0. It has been shown by Gulliver [5, Lemma 2.2] that on some neighborhood U of z_0 one can introduce a (non-conformal) local parameter ζ of class $C^2(U) \cap C^\omega(U \setminus \{z_0\})$, $\zeta(z_0) = 0$, such that the surface f can be represented on U in the form

$$f_1 + if_2 = \zeta^m, \; f_3 = \psi(\zeta) \tag{3}$$

where f_1, f_2, f_3 are Euclidean coordinates with the origin at $f(z_0)$ and the third axis in direction of the surface normal of f at z_0. The function ψ is of class C^2 for $|\zeta|<r$, analytic for $0<|\zeta|<r$, and satisfies $\psi(\zeta)=O(|\zeta|^{m+1})$.

An important distinction between two types of branch points has to be made depending on the geometrical nature of the image of a small vicinity of the branch point: a branch point is called false if for some neighborhood this image is an embedded piece of surface in $\mathbb{R}^3$. Otherwise the branch point is a true branch point. One sees immediately from (3) that the branch point is false if and only if ψ is a function of ζ^m.

The boundary differentiability results of Jäger [7] imply that in case of an analytic supporting surface S any stationary solution to the free boundary value problem which is continuous up to the boundary is in fact analytic up to the boundary. Continuity up to the boundary of Sobolev class solutions has been proved by Jäger for globally minimizing surfaces and by Grüter-Hildebrandt-Nitsche [4] and Dziuk [3] for merely stationary solutions. A somewhat weaker analycity result had already been proved by Lewy [8]. By the boundary analycity all the above notions and developments concerning branch points remain valid on the boundary of the surface.

We can show regularity for those solutions to the free boundary value problem which furnish a strong relative minimum to Dirichlet's integral (or area) in the sense of the following

Definition 1. The surface $f\in\mathscr{C}(S,\Gamma)$ is a strong relative minimum for Dirichlet's integral D if there is an open subset U of $\mathbb{R}^3$ containing $f(B)$ such that $D(f)\leqq D(g)$ for alle surfaces $g\in\mathscr{C}(S,\Gamma)$ with $g(B)\subset U$.

From this minimizing property it follows already by the above mentioned result of Osserman [10] that f has no true interior branch points. Therefore we need only to consider interior false branch points and boundary branch points of either type. By the boundary analyticity we may assume that our surface f is defined on some open neighborhood B_0 of the closed unit disc B where, however, all branch points are contained in B.

Definition 2. Two points $z, w\in B_0$ are called equivalent, $z\sim w$, if there are fundamental systems of open connected neighborhoods $U_n(z)$ and $V_n(w)$, $n\in\mathbb{N}$, such that $f(U_n)=f(V_n)$ for all n.

We set

$$\tilde{\partial}B=\{z\in B\mid \exists z'\in\partial B \quad \text{with} \quad z'\sim z\}.$$

Lemma 1. *Let $z_0, w_0\in B_0$ such that $f(z_0)=f(w_0)$ and also the tangent planes of f at z_0 and w_0 coincide. Let U and V be coordinate neighborhoods of z and w_0 for which a representation* (3) *holds, respectively, and assume that there is an open subset U' of U such that $f(U')\subset f(V)$. Then it follows that $z\sim w$.*

Proof. Since we allow also orientation reversing local coordinate transformations we can assume that

$$f(U)=\{x=(x_1,x_2,x_3)\mid x_1+ix_2=\zeta^m,\ x_3=\varphi(\zeta),\ |\zeta|<r\},$$
$$f(V)=\{x=(x_1,x_2,x_3)\mid x_1+ix_2=\omega^n,\ x_3=\chi(\omega),\ |\omega|<s\}.$$

It follows from the hypothesis that there is an open set of complex numbers z contained in $\{w \mid 0<|w|<\min(\sqrt[n]{r}, \sqrt[m]{s})\}$ such that

$$\varphi(z^n)=\chi(\eta_n z^m) \tag{4}$$

for some n-th root η_n of unity. By the analycity of φ and χ in $0<|\zeta|<r$ and $0<|\omega|<s$, respectively, (4) must hold for all $z, |z|<\min(\sqrt[n]{r}, \sqrt[m]{s})$. If we let $U_\varepsilon = \{z \in U \mid |\zeta(z)|<\sqrt[n]{\varepsilon}\}$ and $V_\varepsilon=\{w \in V \mid |\omega(w)|<\sqrt[n]{\varepsilon}\}$ then it follows immediately that $f(U_\varepsilon)=f(V_\varepsilon)$ for all sufficiently small ε, which proves the lemma.

Corollary. *The equivalence relation $\sim$ is closed, i.e. from $z_k \to z$, $w_k \to w (k \to \infty)$ and $z_k \sim w_k$ for all $k \in \mathbb{N}$ it follows that $z \sim w$. In particular, it follows that $\tilde{\partial} B$ is a closed set.*

The following technical notion will be useful in the sequel:

Definition 3. An analytic arc in $\mathbb{R}^n$ $(n \geqq 2)$ emanating from a point p is the image of a closed interval $[0, \delta]$ under a non-constant map α which is real analytic on an open interval containing $[0, \delta]$ and with $\alpha(0)=p$.

Lemma 2. *Let $\Phi: U_0 \to \mathbb{C}$ be a real analytic map defined in some neighborhood U_0 of $0 \in \mathbb{C}$ which for $z \to 0$ has the asymptotic behavior*

$$\Phi(z)=a z^m+O(|z|^{m+1}), \qquad a \neq 0. \tag{5}$$

Let furthermore $\alpha: [0, \delta_0] \to C$ be an analytic arc emanating from $0 \in U_0$. Then there exist a positive number $\delta \leqq \delta_0$ and a neighborhood U of 0 such that $U \cap \Phi^{-1}(\alpha([0, \delta]))$ consists of m analytic arcs emanating from 0.

Proof. We can clearly assume that $a=1$ in (5). Replacing the map α by $\tilde{\alpha}(t) = \alpha(t^{2m})$ which parametrizes the same arc we may also suppose that the map α can be represented in the form

$$\alpha(t)=p(t)^m \gamma(t), \qquad p \geqq 0, \ |\gamma|=1,$$

with analytic functions p and γ. We now choose $U=\{z \in \mathbb{C} \mid |z| \leqq \delta\}$ such that $\Phi(z) \neq 0$ for $z \in U \setminus \{0\}$ and introduce polar coordinates for Φ, too; i.e. we define

$$\hat{\Phi}: [0, \delta] \times S^1 \to \mathbb{R}^+ \times S^1,$$
$$\hat{\Phi}(r, \zeta)=(|\Phi(r \zeta)|, |\Phi(r \zeta)|^{-1} \Phi(r \zeta)).$$

It follows from (5) that

$$\hat{\Phi}(0, \zeta)=(0, \zeta^m) \tag{6}$$

and that $\hat{\Phi}$ can analytically be continued across the boundary $\{0\} \times S^1$, allowing negative values for r. In fact, the representation

$$\Phi(r \zeta)=r^m \zeta^m (1+\Phi_1(r, \zeta)) \tag{7}$$

holds with an analytic map $\Phi_1, \Phi_1(0, \zeta)=0$. We can therefore define an analytic map

$$\tilde{\Phi}(r, \zeta)=(\sqrt[m]{|\Phi(r \zeta)|}, |\Phi(r \zeta)|^{-1} \Phi(r \zeta)).$$

Now clearly the pre-image of the arc α viewed in polar coordinates equals the pre-image of the arc (p^m, γ) under $\tilde{\Phi}$ and this is exactly the pre-image of (p, γ) under $\tilde{\Phi}$. By (6), $\tilde{\Phi}^{-1}(0, \gamma(0))$ consists of the m-th roots $\gamma_1, \ldots, \gamma_m$ of $\gamma(0)$ and from the properness of $\tilde{\Phi}$ on $[0, \delta] \times S^1$ it follows that for any given $\varepsilon_0 > 0$ there exists $\varepsilon > 0$ such that the preimage under $\tilde{\Phi}$ of an ε-neighborhood of $(0, \gamma(0))$ in $\mathbb{R}^+ \times S^1$ is contained in the union of ε_0-neighborhoods of $(0, \gamma_1), \ldots, (0, \gamma_m) \in [0, \delta] \times S^1$. We choose $\varepsilon_0 > 0$ so that $\tilde{\Phi}$ is an analytic diffeomorphism on each rectangle $|r| < \varepsilon_0$, $|\zeta - \gamma_j| < \varepsilon_0$ which is possible by (7). If we choose δ so small that $p(t) < \varepsilon$ and $|\gamma(t) - \gamma(0)| < \varepsilon$ for $0 \leqq t \leqq \delta$, it follows that $\tilde{\Phi}^{-1}((p, \gamma)([0, \delta]))$ consists of m analytic arcs.

Lemma 3. *$\tilde{\partial} B$ is the union of finitely many analytic arcs.*

Proof. By the corollary to Lemma 1, $\tilde{\partial} B$ is compact and therefore it suffices to prove the lemma locally. It is obvious from (3) that each point $z_0 \in B$ has a neighborhood $U(z_0)$ such that $f(z) \neq f(z_0)$ for $z \in U(z_0) \setminus \{z_0\}$. Therefore $f^{-1}(p)$ is discrete for all points $p \in (B)$. Since $f(z) = f(w)$ if $z \sim w$ it follows that there are only finitely many points in B equivalent to a given point z_0. Let now $z_0 \in \partial B$ and let $z_1, \ldots, z_n$ be all points on ∂B, possibly including z_0 itself, which are equivalent to z_0. We claim that for given neighborhoods $U_j(z_j)$ $(j = 1, \ldots, n)$ we can find a neighborhood U of z_0 such that

$$\{w \in \partial B \mid w \sim \zeta \text{ for some } \zeta \in U\} \subset \bigcup_{j=1}^{n} (U_j \cap \partial B). \tag{8}$$

If not, then there are sequences $\zeta_k \in B_0$, $w_k \in \partial B$ such that $\zeta_k \to z_0 (k \to \infty)$, $\zeta_k \sim w_k$, but $w_k \notin \bigcup_{j=1}^{n} U_j$.

We can assume that $w_k \to w \in \partial B$ where of course $w \notin \bigcup_{j=1}^{n} U_j$. Since our equivalence relation is closed it follows that $w \sim z_0$ and therefore $w = z_j$ for some j, a contradiction. Thus (8) must be true. Let P denote the orthogonal projection onto the common tangent plane of f at the points $z_0, z_1, \ldots, z_n$ and let us set $\Phi = P \circ f$. We can assume that $f(z_0) = 0$ and that the tangent plane is the (x_1, x_2)-plane. It follows from (8) that

$$\tilde{\partial} B \cap U \subset B \cap U \cap \Phi^{-1}\left(\bigcup_{j=1}^{n} \Phi(U_j \cap \partial B)\right). \tag{9}$$

We can clearly assume that $U_j \cap \partial B$ is the union of two circular arcs γ_j^+ and γ_j^- emanating from z_j in opposite directions. Therefore, $\bigcup_{j=1}^{n} \Phi(U_j \cap \partial B)$ is a finite union of analytic arcs emanating from the origin and by Lemma 2 we can choose $U_1, \ldots, U_n$ and U in such a way that the right hand side of (9) is also a union of analytic arcs $\alpha_1, \ldots, \alpha_N$ emanating from z_0. We shall prove the lemma by showing that, for sufficiently small U, $\tilde{\partial} B \cap U$ consists of exactly those arcs α_k in $B \cap U$ which contain some point of $\tilde{\partial} B$ different from z_0. Let α_k be such an arc. We can assume that $\Phi(\alpha_k) \subset \Phi(\gamma_j^+)$ for some j and that there exist points $z \in \alpha_k$ and $w \in \gamma_j^+$, $z \neq z_0$, such that $z \sim w$. We want to show that $\alpha_k \subset \tilde{\partial} B$. We can of course assume that each U_j is a coordinate neighborhood for which a

representation (3) holds. If U_j is small enough it follows from (3) that we can find an open sector

$$\Sigma = \{z_j + re^{i\theta} \mid r>0, |\theta-\theta_j|<\varepsilon\},$$

$$\theta_j = \arg z_j \pm \frac{\pi}{2},$$

such that $\gamma_j^+ \setminus \{z_j\} \subset \Sigma$ and $f(\Sigma \cap U_j)$ is a graph over the plane domain $\Phi(\Sigma \cap U_j)$. We may furthermore suppose that $\Phi(\Sigma \cap U_j)$ covers the intersection of $\Phi(U_j)$ with an open halfspace $H = \left\{re^{i\theta} \mid r>0, |\theta-\tau|<\frac{\pi}{2}\right\}$ where $e^{i\tau}$ is the direction of the analytic arc $\Phi(\gamma_j^+)$ at $\Phi(z_j)=0$. Since $\Phi(\alpha_k) \subset \Phi(\gamma_j^+)$ we can by continuity find an open sector

$$\Sigma_0 = \{z_0 + re^{i\theta} \mid r>0, |\theta-\theta_0|<\delta\}$$

where $e^{i\theta_0}$ is the direction of the arc α_k at z_0 such that, for small enough U, $\Phi(\Sigma_0 \cap U) \subset H$, $\alpha_k \setminus \{z_0\} \subset \Sigma_0$, and $f(\Sigma_0 \cap U)$ is a graph over $\Phi(\Sigma_0 \cap U)$. By assumption there exist points $z \in \alpha_k \setminus \{z_0\} \subset \Sigma_0 \cap U$ and $w \in \gamma_j^+ \setminus \{z_j\} \subset \Sigma \cap U_j$ which are equivalent. It follows from the analycity of minimal graphs that $f(\Sigma_0 \cap U) \subset f(\Sigma \cap U_j)$. Consequently, if $\eta \in \Sigma_0 \cap U$ and $\zeta \in \Sigma \cap U_j$ are points with $\Phi(\eta) = \Phi(\zeta)$ then necessarily $\eta \sim \zeta$. Hence we can conclude that $\alpha_k \subset \tilde{\partial} B$.

Lemma 4. *Let $\tilde{\partial}_1 B$ be the connected component of $\tilde{\partial} B$ which contains ∂B. Then $\mathring{B} \setminus \tilde{\partial}_1 B$ is connected.*

Proof. It follows from the analycity result of the previous lemma that $\mathring{B} \setminus \tilde{\partial}_1 B$ consists of finitely many connected components $B_1, \ldots, B_n$, each B_k having piecewise analytic boundary. By construction each B_k is simply connected and we have

$$f(\partial B) = \sum_{k=1}^{n} f(\partial B_k)$$

in the sense of 1-chains, where $f(\partial B_k) \subset f(\tilde{\partial} B) \subset f(\partial B) \subset S$. Since $f(\partial B)$ is not contractible in $\mathbb{R}^3 \setminus \Gamma$ it follows that $f(\partial B_j)$ is not contractible in $\mathbb{R}^3 \setminus \Gamma$ for some index j. Let $\tau: \mathring{B} \to B_j$ be a conformal map. The surface $f \circ \tau$ then belongs to $\mathscr{C}(S, \Gamma)$ and if n were greater 1 we would have

$$D(f \circ \tau) = D(f \mid B_j) < D(f),$$

contradicting the minimality of f.

Lemma 5. *$\tilde{\partial}_1 B = \partial B$ and consequently $\tilde{\partial} B \cap \mathring{B}$ is compact.*

Proof. If $\partial B \subsetneqq \tilde{\partial}_1 B$ then we cut out $\tilde{\partial}_1 B \cap \mathring{B}$ from B, i.e. we replace f by $f \circ \tau$ where $\tau: \mathring{B} \to \mathring{B} \setminus \tilde{\partial}_1 B$ is a conformal map. By Lemma 4, $\tilde{\partial}_1 B$ consists of ∂B together with a finite number of "trees" growing out of certain points on ∂B. Since we can shrink down each tree to its root it follows that the loop $f \circ \tau \mid \partial B$ is homotopic to $f(\partial B)$ on S. Therefore $f \circ \tau$ belongs to $\mathscr{C}(S, \Gamma)$, $f \circ \tau(B) \subset f(B)$, and $D(f \circ \tau) = D(f)$. It follows that $f \circ \tau$ is again a strong relative minimum in the sense of Definition 1 and for $f \circ \tau$ one has $\tilde{\partial}_1 B = \partial B$. It is conceivable that

the above cutting process makes branch points of f disappear. On the other hand new boundary branch points are created at those points which correspond to end points of a tree in $\tilde{\partial}_1 B \cap \mathring{B}$ since there the boundary map of $f \circ \tau$ runs back in its own trace. Therefore we have to show that strong relative minima which additionally satisfy $\tilde{\partial}_1 B = \partial B$ are free of branch points. But this is exactly what we are going to prove in Lemma 6 and Theorem 1.

Lemma 6. $\tilde{\partial} B = \partial B$.

Proof. If $\partial B \subsetneqq \tilde{\partial} B$ then the set

$$\partial_0 B = \{z \in \partial B \mid z \sim z_0 \text{ for some } z_0 \in \mathring{B}\}$$

is not empty. It is immediate from the definition of the relation $\sim$ that $\partial_0 B$ is open in ∂B and it follows from the corollary to Lemma 1 and from Lemma 5 that we can assume that $\partial_0 B$ is also closed. Hence $\partial B = \partial_0 B$, which means that f maps some neighborhood of ∂B into $f(\mathring{B})$. This implies that $f(\mathring{B})$ is a compact minimal surface in $\mathbb{R}^3$, a contradiction.

Now we are ready to prove

Theorem 1. *Each strong relative minimum $f \in \mathscr{C}(S, \Gamma)$ is immersed up to the boundary.*

Proof. As was already remarked at the beginning of this section, by the result of Osserman [10] a strong relative minimum has no true interior branch points. Let us show that there are no false branchpoints in $\mathring{B}$ either. Assuming that $z_0 \in \mathring{B}$ is a false branch point of order m we can conclude from Lemma 5 and from Satz 2.2 in [1] that there exist m Jordan arcs $\gamma_k : [0, 1] \to B$ such that

$$\begin{aligned}
&\gamma_k([0, 1[) \subset B^0, \\
&\gamma_k(0) = z_0, \quad \gamma_k(1) \in \partial B \ (k = 1, \ldots, m), \\
&\gamma_k(t) \sim \gamma_l(t) \quad (1 \leqq k \leqq l \leqq m), \\
&\gamma_k(]0, 1[) \cap \gamma_l(]0, 1[) = \varnothing \quad \text{for } k \neq l.
\end{aligned}$$

For the convenience of the reader let us sketch the proof for the existence of these curves. We start with an analytic Jordan arc $\gamma_1 = \gamma_1(t)$, $0 \leqq t \leqq 1$, $\gamma_1([0, 1[) \subset \mathring{B}$, joining z_0 with an arbitrary boundary point $w = \gamma(1)$. This arc can be chosen in such a way that for $t > 0$ it avoids branch points and points equivalent to branch points. Now we consider the set of all points in B equivalent to z_0 and the set of all curves $\gamma_2, \ldots, \gamma_N$ equivalent to γ_1 in the sense that

$$\gamma_j(t) \sim \gamma_1(t) \quad (j = 2, \ldots, N).$$

The numbering is such that $\gamma_1, \ldots, \gamma_m$ start from z_0. The existence of these curves for small t follows from (3) and globally by analytic continuation. The existence up to $t = 1$ is secured by Lemma 5. It may occur that this system $\{\gamma_j\}$ possesses points of intersection, i.e. there are values $t_0 > 0$, $t_1 > t_0$ and indices k, l such that $\gamma_k(t_0) = \gamma_l(t_1)$. This situation can however happen only a finite number of times. Let us choose t_0 minimal and then t_1 maximal with the

above property. Obviously the sets $\{\gamma_j(t_0)|j=1,\dots,N\}$ and $\{\gamma_j(t_1)|j=1,\dots,N\}$ are identical whence there is a permutation σ with

$$\gamma_j(t_0)=\gamma_{\sigma(j)}(t_1).$$

We define a new system $\tilde{\gamma}_j$ by

$$\tilde{\gamma}_j(t)=\begin{cases}\gamma_j(t), & 0<t\leqq t_0,\\ \gamma_j(t_0), & t_0\leqq t\leqq t_1,\\ \gamma_{\sigma(j)}(t), & t_1\leqq t\leqq 1.\end{cases}$$

According to our construction, in the new system points of intersection can only occur for values of t strictly greater than t_0. If necessary this process has to be repeated. After a finite number of times we obtain the desired system (for $j=1,\dots,m$).

After renumbering we can assume that $z_1=\gamma_1(1),\dots,z_m=\gamma_m(1)$ are consecutive points on ∂B in positive orientation. Denoting the arc on ∂B joining z_k and z_{k+1} by σ_k where $z_{m+1}:=z_1$, we see that $f(\sigma_k)$ is a closed loop on S and

$$f(\partial B)=\sum_{k=1}^{m} f(\sigma_k).$$

Let us choose k such that $f(\sigma_k)$ is not contractible in $\mathbb{R}^3\setminus\Gamma$ and let us denote by B_k the subdomain of B bounded by γ_k, σ_k, and γ_{k+1}. One can then construct a diffeomorphism

$$\tau: B\setminus[0,1]\to B_k\cup\mathring{\sigma}_k$$

such that τ has the one sided limits

$$\lim_{z\to t}\tau(z)=\gamma_k(t),\ (\operatorname{Im} z>0, 0\leqq t\leqq 1)$$
$$\lim_{z\to t}\tau(t)=\gamma_{k+1}(t),\ (\operatorname{Im} z<0, 0\leqq t\leqq 1).$$

Since $\gamma_k(t)\sim\gamma_{k+1}(t)$, the map $f\circ\tau$ is continuous throughout B and therefore $f\circ\tau\in\mathscr{C}(S,\Gamma)$. Obviously $f\circ\tau(B)\subset f(B)$ and $f\circ\tau$ has strictly smaller area than f, contradicting the minimum property of f.

Let us now show that there are no boundary branch points, either. We argue again by contradiction and assume that $z_0\in\partial B$ is a branch point of order $m\geqq 1$. Here it is convenient to assume that B is the closed upper half plane, $z_0=0$, $f(z_0)=0$, and that the oriented tangent plane at $z_0=0$ is the (x_1,x_2)-plane, identified with $\mathbb{C}$. We can also assume that the asymptotic tangent direction of the analytic curve $f(\mathbb{R}^+)$ at $z_0=0$ is the vector $(1,0,0)$. We know that the tangent planes of f and of the supporting surface S at the origin intersect orthogonally along the x_1-axis. We want to show that there exists a differentiable arc $\alpha:[0,1]\to B$, $\alpha(0)=z_0$, $\alpha(]0,1])\subset\mathring{B}$ such that $f(\alpha([0,1]))\subset S$. Again we denote by P the orthogonal projection onto the tangent plane and let $\Phi=P\circ f$. It follows from the representation (3) that there are numbers $r>0$, $R>0$, and

$\theta \in \left]\frac{\pi}{m+1}, \frac{2\pi}{m+1}\right[$ such that the image under Φ of the sector $S_{r,\theta} = \{\rho e^{i\varphi} \mid 0<\rho<r,\ 0<\varphi<\theta\}$ covers the half disc $H_R = \left\{\rho e^{i\varphi} \middle| \frac{\pi}{2}<\varphi<\frac{3\pi}{2}\right\}$ and that $f(S_{r,\theta})$ is a graph over $\Phi(S_{r,\theta})$. Since at the origin the surface S is orthogonal to the (x_1, x_2)-plane it is then obvious that $f(S_{r,\theta})$ intersects S along an analytic arc $\hat{\alpha}:]0,1] \to \mathbb{R}^3$ such that $\hat{\alpha}(t) \to 0\,(t \to 0)$ and $|\hat{\alpha}'(t)|^{-1}\hat{\alpha}'(t) \to (-1,0,0)\,(t \to 0)$. The arc $\alpha = (f \mid S_{r,\theta})^{-1}\hat{\alpha}$ then has the desired properties.

Next we want to show that f intersects S along α orthogonally. Indeed, let $B' = B \setminus \alpha([0,1])$ and $\tau: \mathring{B} \to \mathring{B}'$ be a conformal map. Then $f \circ \tau(\partial B)$ is homotopic to $f(\partial B)$ on S and hence $f \circ \tau \in \mathscr{C}(S, \Gamma)$ with $D(f \circ \tau) = D(f)$. It follows that $f \circ \tau$ is a strong relative minimum to our free boundary value problem and consequently $f \circ \tau$ is orthogonal to S along ∂B which shows that f is perpendicular to S along α.

Since we know already that $f \mid \mathring{B}$ is immersed, the curve α can analytically be continued in $\mathring{B}$ as an immersed curve. By analycity f must stay orthogonal to S along α and therefore α can be continued until it reaches ∂B. The curve α cannot have any double points since f is an embedding locally in $\mathring{B}$ and therefore intersects S in an embedded arc. It follows that $\mathring{B} \setminus \alpha$ consists of two simply connected domains B_1, B_2 where

$$f(\partial B) = f(\partial B_1) + f(\partial B_2).$$

Choose $k \in \{1,2\}$ such that $f(\partial B_k)$ is not contractible in $\mathbb{R}^3 \setminus \Gamma$ and let $\tau: \mathring{B} \to B_k$ be a conformal map. Then $f \circ \tau \in \mathscr{C}(S, \Gamma)$, and $f \circ \tau(B) \subset f(B)$ and $D(f \circ \tau) < D(f)$, a contradiction. The theorem is proved.

2. On the Finiteness of the Number of Solutions

Building upon Theorem 1 we shall in this section prove the following finiteness theorem,

Theorem 2. *Let S be a compact, embedded, analytic surface in $\mathbb{R}^3$ and Γ be a homotopically non-trivial Jordan curve in the unbounded component of $\mathbb{R}^3 \setminus S$. Then there are only finitely many geometrically different surfaces which minimize Dirichlet's integral in $\mathscr{C}(S, \Gamma)$.*

In order to prove this theorem we must recall some results from a paper [11] of the second author. Let $f \in \mathscr{C}(S, \Gamma)$ be a minimizing surface which by Theorem 1 is immersed. Though f need not be an embedding we can in complete analogy to Lemma 2 in [11] construct an analytic immersion

$$\Phi: B \times]-\delta, \delta[\to \mathbb{R}^3, \qquad \Phi = \Phi(z,t),$$

such that

i) $\Phi(z,0) = f(z)$

ii) $\frac{\partial}{\partial t}\Phi(z,0)$ is a unit normal along f

iii) $\Phi(\partial B \times]-\delta, \delta[) \subset S$.

It follows then that each surface $g: B \to \mathbb{R}^3$ with $g(\partial B) \subset S$ which is sufficiently close to f in C^1-norm can be uniquely represented in the form

$$g \circ \tau(z) = \Phi(z, u(z))$$

with some diffeomorphism $\tau: B \to B$ and some function $u: B \to \mathbb{R}$, both depending on g. This representation makes it possible to apply the analysis of § 2 in [11]. In particular, we obtain the conclusion of Proposition 6 in [11]: let $f \in \mathscr{C}(S, \Gamma)$ be a minimizing surface which is the C^1-limit of stationary surfaces geometrically different from f, then there exists a one-parameter family $F = F(t), |t| < \varepsilon$, of area minimizing surfaces in $\mathscr{C}(S, \Gamma)$ (not necessarily in conformal representation) such that $F(0) = f$, $F'(0)$ is a non-vanishing normal vector field along $F(0) = f$ and any solution to the free boundary value problem in $\mathscr{C}(S, \Gamma)$ sufficiently close to f belongs to the family F (after suitable reparametrization). It follows that the map

$$B \times]-\varepsilon, \varepsilon[\ni (z, t) \mapsto F(t)(z) \in \mathbb{R}^3$$

is an immersion for sufficiently small ε. It has also been proved in Proposition 5 of [11] that the set of minimizing surfaces in $\mathscr{C}(S, \Gamma)$ becomes compact in any C^k-topology if one introduces the normalization condition $f(0) \in \Gamma$. If we therefore assume that there are infinitely many geometrically different minimizing surfaces in $\mathscr{C}(S, \Gamma)$ then there is a sequence f_n of such surfaces which converges to some $f^* \in \mathscr{C}(S, \Gamma)$ which again must be minimizing. Let $\mathscr{C}^*$ denote the connected component of f^* in the set of minimizing surfaces. Clearly, $\mathscr{C}^*$ contains only such solutions to the free boundary value problem which are not isolated in the set of all solutions. From the above result on the structure of this set in the neighborhood of a non isolated solution it follows that

$$U^* := U \cap \bigcup_{f \in \mathscr{C}^*} f(B)$$

is a non-empty open subset of the unbounded component U of $\mathbb{R}^3 \setminus S$. But from the compactness of minimizing solutions we conclude on the other hand that U^* is bounded and relatively closed in U. It follows that $U^* = U$ and hence U would be bounded, an obvious contradiction.

Bibliography

1. Alt, H.W.: Verzweigungspunkte von H-Flächen. II. Math. Ann. **201**, 33–55 (1973)
2. Courant, R.: Dirichlet's principle, conformal mapping, and minimal surfaces. New York: Interscience 1950
3. Dziuk, G.: Über die Stetigkeit teilweise freier Minimalflächen. Manuscr. Math. **36**, 241–251 (1981)
4. Grüter, M., Hildebrandt, S., Nitsche, J.C.C.: On the boundary behavior of minimal surfaces with a free boundary which are not minima of the area. Manuscr. Math. **35**, 387–410 (1981)
5. Gulliver, R.D.: Regularity of minimizing surfaces of prescribed mean curvature. Ann. of Math. **97**, 275–305 (1973)

6. Hildebrandt, S., Nitsche, J.C.C.: Geometric properties of minimal surfaces with free boundaries. Math. Z. **184**, 497–509 (1983)
7. Jäger, W.: Behavior of minimal surfaces with free boundaries. Commun. Pure Appl. Math. **23**, 803–818 (1970)
8. Lewy, H.: On minimal surfaces with free boundaries. Commun. Pure Appl. Math. **4**, 1–13 (1951)
9. Meeks, W.H., Yau, S.T.: Topology of three dimensional manifolds and the embedding problems in minimal surface theory. Ann. of Math. **112**, 441–484 (1980)
10. Ossermann, R.: A proof of the regularity everywhere of the classical solution to Plateau's problem. Ann. of Math. **91**, 550–569 (1970)
11. Tomi, F.: A finiteness result in the free boundary value problem for minimal surfaces. To appear

Received July 2, 1984

Math. Z. 189, 239–261 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Modular Representations of $GL(n, p)$, Splitting $\Sigma(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty)$, and the β-family as Framed Hypersurfaces

D. Carlisle, P. Eccles, S. Hilditch, N. Ray, L. Schwartz, G. Walker, R. Wood
Department of Mathematics, The University, Manchester M139PL, Great Britain

In memory of M.H.A. Newman
Fielden Professor of Pure Mathematics 1945–1964

0. Introduction

In this paper we show how to split the p-localisation of the space

$$Y = \Sigma(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty)$$

into a bouquet of spaces whose cell structures are governed by the modular representations of the general linear group $GL(n,p)$, where p is a prime and n the number of copies of $\mathbb{C}P^\infty$ in the above product. Our main interest is in the case $n=2$, where detailed information on the modular representation theory is available.

To formulate the general result, let ρ run through the finite set of irreducible representations of $GL(n,p)$ over the finite field $\mathbb{F}_p$. Suppose $d(\rho)$ denotes the dimension of the underlying vector space of a module $V(\rho)$ affording the representation ρ, and $n(\rho,r)$ the number of times that $V(\rho)$ appears as a composition factor in the $GL(n,p)$ module of homogeneous polynomials of degree r in $\mathbb{F}_p[x_1,\ldots,x_n]$. In §2 we prove

2.13. Theorem. *There are CW-complexes $Y(\rho)$ of finite type such that Y and the bouquet $\bigvee_\rho d(\rho)\,Y(\rho)$ are homotopy equivalent after p-localisation. Moreover, $Y(\rho)$ has no cells in even positive dimensions, and $n(\rho,r)$ cells in dimension $2r+1$.*

The notation $d(\rho)\,Y(\rho)$ indicates that the space $Y(\rho)$ appears $d(\rho)$ times.

Motivation for this work arises in the problem of representing elements in the stable homotopy groups of spheres π_*^S by framed manifolds of some specified nature. We recall for example that the image of the classical J homomorphism, determined by Adams [1] and Mahowald [13], consists of those elements representable via the Pontrjagin-Thom construction by framings on standard spheres. A natural extension is to investigate elements which

can be represented by framings on hypersurfaces; i.e. smooth manifolds which admit codimension 1 embeddings in Euclidean space. For work in this area we refer to Eccles [8], Eccles-Walker [9], Ray-Schwartz [19] and Rees [20].

Our aim is to use the splitting of Theorem 2.13 in the case $n=2$ to produce certain elements in the unstable homotopy of $\Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$. These, in conjunction with the double transfer and some elementary hypersurface geometry, allow us to prove the following result in § 3.

3.7. Theorem. *For any odd prime p, the elements $\beta_1, \ldots, \beta_{p-2}$ and β_p of the β-family in $\hat{p}\pi_*^S$ can be represented by framed hypersurfaces. So too can the element γ_1, $p \geqq 5$, and the non-zero products $\alpha_1\beta_i$, $1 \leqq i \leqq p-2$.*

We refer to Smith [21] and Toda [24] for background information on the β-family, and to Miller-Ravenel-Wilson [16] for details about γ_1 and the products $\alpha_1\beta_i$. The omission of β_{p-1} in the theorem will be explained in § 3.

Our methods use the ideas of Cohen [4], Cooke-Smith [6], Mimura-Nishida-Toda [16], Mitchell-Priddy [17] and other authors for splitting the localised suspension of an H-space according to the action of certain self-maps in $\mathbb{F}_p$ homology. In addition, Knapp's work [12] is invaluable in locating the β-family in the image of the double transfer.

In § 1 we review the modular representation theory required for Theorem 2.13, and in § 2 we apply it to the proof. In § 3 we explain the homotopy viewpoint behind the hypersurface problem, and prove Theorem 3.7.

We would like to acknowledge the invaluable organisational suggestions provided by the referees, and to thank the Manchester algebraists for many helpful discussions on modular representation theory. Bob Sandling deserves special mention for directing our attention to Glover's work [10].

1. Modular Representations

As a general reference for this section we cite Curtis and Reiner [7], and for specific information on $GL(2,p)$ our source is Glover [10].

We shall let $A = \mathbb{F}_p[GL(n,p)]$ denote the group ring of $GL(n,p)$ over $\mathbb{F}_p$, and write L for the polynomial ring $\mathbb{F}_p[x_1, \ldots, x_n]$ viewed as a right $GL(n,p)$ module under linear substitution. We let L_k stand for the set of homogeneous polynomials in L of degree k, so that L_k is a right A-module. For future applications we make L into a graded ring by assigning to each x_k the grading 2.

Steenrod powers act as A-module maps on L by the formulae

$$\mathscr{P}^r(x_k^i) = \binom{i}{r} x_k^{i+r(p-1)},$$

extended to the whole of L by linearity and the Cartan formulae

$$\mathscr{P}^r(ab) = \sum_{u+v=r} \mathscr{P}^u(a)\mathscr{P}^v(b),$$

where a and b are homogeneous polynomials. For $p=2$, we interpret $\mathscr{P}^r$ as Sq^{2r}.

We quote the following basic result in representation theory.

1.1. Proposition. *There is a splitting of right $GL(n,p)$ modules*

$$A \xrightarrow{\cong} \bigoplus_{\rho} d(\rho) P(\rho),$$

where $P(\rho)$ is a projective indecomposable which occurs $d(\rho)$ times. Associating to $P(\rho)$ its top composition factor $V(\rho)$ sets up the bijection, up to isomorphism, between the projective indecomposables and the irreducible representations.

In the above splitting, $P(\rho)$ occurs $d(\rho)$ times since the irreducible representations ρ are absolutely irreducible. And $V(\rho)$ also occurs as a bottom composition factor of $P(\rho)$.

The splitting is actually effected by a set of primitive idempotents $e_{i\rho}$, $1 \leqq i \leqq d(\rho)$, acting on the left of A, so that $P(\rho) \cong e_{i\rho}A$. By definition, the identity element of A may be expressed as

$$1 = \sum_{i,\rho} e_{i\rho},$$

where $e_{i\rho}^2 = e_{i\rho}$, $e_{i\rho}e_{j\sigma} = 0$ if $(i,\rho) \neq (j,\sigma)$ and no $e_{i\rho}$ can be further decomposed. The expression for 1 and the corresponding splitting of A are not unique, but given

$$1 = \sum_{i,\rho} e_{i\rho} = \sum_{i,\rho} e'_{i\rho}$$

we have

1.2. Proposition. *There is an invertible $\delta \in A$ such that $e'_{i\rho} = \delta e_{i\rho} \delta^{-1}$ for all i, ρ.*

Proof. The projective indecomposables are determined up to isomorphism, so we can choose a right module isomorphism $\phi\colon A \to A$ mapping $P_{i\rho} = e_{i\rho}A$ onto $P'_{i\rho} = e'_{i\rho}A$. Then ϕ is given by left multiplication with $\delta = \phi(1)$. But $\delta e_{i\rho}\delta^{-1} = \phi(e_{i\rho})\delta^{-1}$ lies in $P'_{i\rho}$ and is idempotent, so $\delta e_{i\rho}\delta^{-1} = e'_{i\rho}$ by uniqueness.

Next let W be an arbitrary finitely generated right A module. Then we can decompose W as

$$W = \bigoplus_{i,\rho} W e_{i\rho}.$$

We emphasise that because A is non-commutative, the images $We_{i\rho}$ are not in general submodules.

Observe that, although a composition series for W will not usually be unique, the Jordan-Hölder theorem ensures that the composition factors are determined up to isomorphism. Thus we can define $n(\rho, W)$ as the number of times that $V(\rho)$ appears as a composition factor. The following elementary fact is then the key to part of Theorem 2.13.

1.3. Lemma. *The action of $e_{i\rho}$ on W is trivial unless $V(\rho)$ is a composition factor of W. In this case, $We_{i\rho}$ has dimension $n(\rho, W)$ over $\mathbb{F}_p$.*

Proof. Note first that $e_{i\rho}$, $1 \leqq i \leqq d(\rho)$, are associated with the same ρ, and hence act non-trivially on $V(\rho)$. Thus

$$V(\rho) = \bigoplus_{i=1}^{d(\rho)} V(\rho) e_{i\rho}$$

as a vector space. Since $\dim V(\rho) = d(\rho)$, each $V(\rho)e_{i\rho}$ is 1-dimensional. The proof is completed by induction, working up the composition series of W: each occurrence of $V(\rho)$ contributes a 1-dimensional summand of $We_{i\rho}$.

We now turn to the polynomial modules L_k, which for $0 \leqq k \leqq p-1$ are irreducible. Also, L_0 is the trivial 1-dimensional module and L_1 is the dual of the defining module for $GL(n,p)$. The determinant module D and its powers D^r afford all the 1-dimensional modules, of which there are $p-1$, since $D^r = D^s$ if $r \equiv s \bmod (p-1)$. In the case $n=2$ we have the following result.

1.4. Proposition (Glover [10]). *A complete set of inequivalent irreducible modules for $GL(2,p)$ over $\mathbb{F}_p$ is given by*

$$L_{i-1} \otimes D^j, \quad 1 \leqq i \leqq p, \ 0 \leqq j \leqq p-2.$$

Given an idempotent e_ρ, our ultimate aim is to find an explicit basis for the image $L_k e_\rho$ of e_ρ acting on the homogeneous polynomials of degree k. We cannot normally expect to do this canonically, since there will not be a canonical choice for e_ρ in A. If, however, a 1-dimensional representation $V(\rho)$ occurs as a bottom composition factor of L_k, then $V(\rho)$ is a submodule of L_k and is generated by an eigenpolynomial for the action of $GL(n,p)$. For example, the ring of absolute invariants $L^{GL(n,p)}$ is the well known Dickson algebra, which affords generators for all occurrences of the trivial module as a bottom composition factor in the L_k. These are relatively sparse, however, as we can illustrate in the case $n=2$ from Glover's calculation [10] of all composition series for the L_k's.

Although his computations exhibit certain periodic phenomena, the general results are too complicated to be worth reproducing here. We content ourselves with selecting certain specific items of information, and the appended table for $p=5$ may be helpful. Every column is labelled by an irreducible module, and each entry gives the number of times it occurs in the polynomial module labelling the corresponding row (see diagram).

For two variables x, y, it transpires that after the constant polynomials, the next absolute invariants are scalar multiples of

$$(x^{p^2}y - xy^{p^2})/(x^p y - xy^p),$$

which generates a trivial submodule of L_{p^2-p}. If $p=5$, this is L_{20}. However, the table indicates that the trivial representation also occurs as a composition factor in L_8, L_{12}, L_{16}. To illustrate the situation further, let us consider the determinant representation, which admits $x^p y - xy^p$ as an eigenpolynomial.

1.5. Proposition (Glover [10]). *For $GL(2,p)$, the determinant module D appears precisely once in L_{p+1}, L_{3p-1}, L_{4p-2}, ..., L_{p^2-2p+3}, L_{p^2-p+2}, L_{p^2+1} as a composition factor, and nowhere else up to degree p^2+1. It appears as a*

L_m \ $L_i \otimes D^j$	0,0	0,1	0,2	0,3	1,0	1,1	1,2	1,3	2,0	2,1	2,2	2,3	3,0	3,1	3,2	3,3	4,0	4,1	4,2	4,3	Cell dimension
0	1																				1
1					1																3
2									1												5
3													1								7
4																	1				9
5					1									1							11
6		1							1		1										13
7						1		1					1								15
8	1									1							1				17
9					1									2							19
10									1		1							1			21
11						1		1					1		1						23
12	1		1							1		1					1				25
13					2		1							2							27
14		1							1		2							1			29
15						1		1					1		2						31
16	1									1		1					1		1		33
17					2		1							2		1					35
18		1		1					2		2							1			37
19						2		2					1		2						39
20	1		1							1		2					1		1		41
21					2		1							2		2					43
22		1							2		2							1		1	45
23						2		2					2		2						47
24	2		1							2		2					1		1		49
25					3		2							2		2					51
26		1		1					3		2							1		1	53
27						2		2					3		2						55
28	1		1							2		2					2		1		57
29					3		2							3		2					59
30		2		1					3		3							1		1	61

Occurrences of $L_i \otimes D^j$ in L_m, $p=5$.

submodule only in L_{p+1} and L_{p^2+1}. The next occurrence is in L_{p^2+p} where it appears twice, both as a submodule and top composition factor.

The eigenpolynomials $x^{p^2}y - xy^{p^2}$ and $(x^p y - xy^p)^p$ generate D in L_{p^2+1} and L_{p^2+p} respectively.

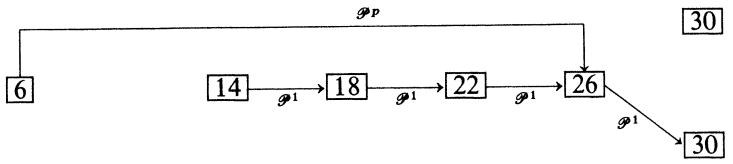
At this point we observe that by elementary calculations with Steenrod powers we obtain

$$\mathscr{P}^p(x^p y - xy^p) = x^{p^2}y - xy^{p^2},$$

$$\mathscr{P}^{p+1}(x^p y - xy^p) = (x^p y - xy^p)^p,$$

consistent with the fact that these operations are module maps and so must preserve eigenpolynomials.

For future reference, our information for the determinant representation can be usefully summarised in a diagram, which we exhibit for $p=5$. Up to p^2+p the pattern is similar for any prime p.



Each box represents a generator of $L_r e$ in the given degree r, where e is an idempotent associated to the determinant representation. In §2 we shall reinterpret this diagram as a skeleton of the cell complex $Y(D)$. For the moment, considering only the algebra, we have already commented on the action of the operations $\mathscr{P}^p$ and $\mathscr{P}^{p+1}=\mathscr{P}^1\mathscr{P}^p$, and also on the lack of canonical generators. However, we can show that with a careful choice of the idempotent e, we may take the polynomial

$$-(x^{3p-2}y-xy^{3p-2})+2(x^{2p-1}y^p-x^py^{2p-1})$$

as a generator of $L_{3p-1}e$ and then the presence of the Steenrod powers $\mathscr{P}^1$ in the above diagram can easily be checked in a purely algebraic way. We shall not go into any more details here.

In the two variable case, the first occurrence of a power of D is as the submodule generated by the corresponding power of the relative Dickson invariant x^py-xy^p. This, however, is false in the three variable case. To see this, note that

$$b=\det\begin{pmatrix} x_1 & x_2 & x_3 \\ x_1^p & x_2^p & x_3^p \\ x_1^{p^2} & x_2^{p^2} & x_3^{p^2} \end{pmatrix}$$

is the relative Dickson invariant analogous to x^py-xy^p, and it generates D as a submodule of L_{p^2+p+1}. However this is not the first occurrence of D as a composition factor. Writing

$$a=\det\begin{pmatrix} x_1 & x_2 & x_3 \\ x_1^p & x_2^p & x_3^p \\ x_1^{2p-1} & x_2^{2p-1} & x_3^{2p-1} \end{pmatrix},$$

it is easy to check that a is not an eigenpolynomial for the determinant representation but

$$\mathscr{P}^{p-1}a=\lambda b \quad .$$

for a non-zero scalar λ. This shows that D must occur as a composition factor in L_{3p}, and we can verify that this is the first occurrence of D: but for an arbitrary representation ρ it is difficult to obtain such information. In view of Theorem 2.13, the series

$$P(\rho,t)=\sum_{r\geqq 0} n(\rho,r)t^r$$

gives the Poincaré series for the complex $Y(\rho)$.

1.6. Problem. *Obtain a formula for* $P(\rho,t)$.

We can show that $P(\rho,t)$ is a rational function of t, a result which we understand has also been demonstrated by S. Mitchell, who has obtained formulae in certain special cases. One of us, D. Carlisle, has recently given a formula for $P(\rho,t)$ for all ρ in the three variable case.

In order to prove Theorems 2.13 and 3.7, we shall need not only L, but also its dual. Given a finite dimensional right A-module W, $W^* = \mathrm{Hom}_A(W, \mathbb{F}_p)$ is a left A-module whose composition series is the dual of that of W: in particular, the dual of a top composition factor in W is a bottom composition factor in W^* and vice-versa.

The study of L as an A-module is of course greatly facilitated by the compatibility of the action of A via linear substitution with the polynomial multiplication. The dual module L^* also has a product which is compatible with the left action of A. Thereby, L^* is isomorphic to the ring of divided polynomials in n variables over $\mathbb{F}_p$, given by R/pR where $R \subset \mathbb{Q}[u_1, \ldots, u_n]$ is the subring generated by the divided monomials $u_j^r/r!$. For notational convenience, we shall suppress this isomorphism.

Thus in the case $n=2$, L^* has generators u, v, dual to x, y, in grading 2, and u^p/p is a non-zero element of L^*_p. Note that for any $h \in L^*_k$, h^p/p is a well-defined element of L^*_{pk}. The action of A is again given by linear substitution, this time using the transposed matrix.

It is now easy to verify that $(u^p v - u v^p)/p$, the dual of the relative Dickson invariant, is not an eigenpolynomial for D if $p>2$. This class is nevertheless important, and we denote it by $\bar{h}_0$. We find that the classes $u^r v^s$ with $r+s = p-1$ and $r, s < p$ generate a submodule S of L^*_{p+1} of dimension $p-2$, and

$$\alpha(\bar{h}_0) = (\det \alpha)\bar{h}_0 \mod S, \quad \alpha \in GL(2,p).$$

Thus $\bar{h}_0$ generates a top composition factor isomorphic to D, in accordance with the facts about duality stated above. Applying [10], we now find that the first occurrence of D as a submodule of L^* is in degree p^2+p.

1.7. Proposition. *The divided polynomial*

$$\bar{h}_1 = \bar{h}_0^p/p$$

is an eigenpolynomial for D in $L^*_{p^2+p}$.

This result is proved by direct calculation, and is crucial for our work in § 3.

Referring to the diagram above, we may check that one of the cells in degree p^2+p, the image of $\mathscr{P}^1$, is given by a canonical polynomial in x, y, whilst the other is given by a canonical divided polynomial in u, v.

2. The Splitting of $\Sigma(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty)$

In this section we prove Theorem 2.13 and illustrate the result in the case $n=2$. The algebra in § 1 is linked to the topology by the isomorphism

$$H^*(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty; \mathbb{F}_p) \cong L$$

as graded $GL(n, p)$-modules. Here $\mathbb{C}P^\infty$ is viewed as an abelian topological group and an integer $n \times n$ matrix acts on the left of a column vector of points in $\mathbb{C}P^\infty$ as usual. The induced left action in homology is compatible with Pontrjagin multiplication arising from the topological group structure on the product of projective spaces. The induced right action in cohomology is compatible with the cup-product ring structure and Steenrod operations.

A formal sum of integer matrices acts on

$$Y = \Sigma(\mathbb{C}P^\infty \times \ldots \times \mathbb{C}P^\infty),$$

using the suspension coordinate to add the suspensions of the individual matrix maps. After reduction mod p the homology $H_*(Y; \mathbb{F}_p)$ and cohomology become dually related left and right modules over the group ring A, which we identify with L^* and L respectively after a shift of grading.

We write $\mathbb{Z}_{(p)}$ for the integers localised at p and $Y_{(p)}$ for the p-localisation of the space Y. The splitting of $Y_{(p)}$ exhibited in Theorem 2.13 is a geometric analogue of the group ring splitting exhibited in Proposition 1.1. To construct the spaces $Y(\rho)$ we follow a path well established by the authors cited in §0 but there are certain technical difficulties to circumvent.

As far as the case $n=1$ is concerned, several authors have shown how to decompose $\Sigma\mathbb{C}P^\infty$.

2.1. Theorem. *There is a homotopy equivalence of p-local spaces*

$$\Sigma\mathbb{C}P^\infty_{(p)} \xrightarrow{\simeq} \bigvee_{i=1}^{p-1} X_i,$$

where the homology of X_i is zero except for a copy of $\mathbb{Z}_{(p)}$ in each dimension $2j+1$ for $j \equiv i \bmod (p-1)$ and $j>0$.

A detailed exposition of the proof can be found in McGibbon [14]. It will help to understand our later construction if at this point we give a brief sketch of the ideas involved.

Each X_i arises as the direct limit of a carefully chosen sequence of self maps of $\Sigma\mathbb{C}P^\infty$, localised at p. In any particular dimension, the induced sequence in homology has a direct limit which is a torsion-free $\mathbb{Z}_{(p)}$ module of rank at most one, and hence is either 0, $\mathbb{Z}_{(p)}$ or $\mathbb{Q}$. Thus it is fairly simple to monitor how these self maps act in the limit. They are chosen to produce zero homology except in the specified dimensions, where they give a copy of $\mathbb{Z}_{(p)}$, so ensuring that X_i has the correct homotopy type.

The natural maps $\Sigma\mathbb{C}P^\infty_{(p)} \to X_i$ are then assembled using the suspension coordinate to form a map $\Sigma\mathbb{C}P^\infty_{(p)} \to \bigvee_{i=1}^{p-1} X_i$. This induces a homology isomorphism and therefore a homotopy equivalence of p-local spaces.

As far as two or more copies of $\mathbb{C}P^\infty$ are concerned, we start with the modular idempotents $e_{i\rho}$. These do not act on Y as they stand, and must first be lifted to formal sums of integer matrices. A lift f of a typical idempotent e is not unique and is no longer idempotent. However, we may still form the direct limit of the sequence

$$Y_{(p)} \xrightarrow{f} Y_{(p)} \xrightarrow{f} \ldots,$$

and its mod p homology is eL^* as required. Unfortunately the integral homology of the direct limit could in principle be quite bizarre, as the following example indicates.

2.2. Example. *Let* $V=\mathbb{Z}_{(p)}\oplus\mathbb{Z}_{(p)}$, *and let* $T\colon V\to V$ *be the module map represented by the matrix* $\begin{bmatrix}0 & p\\ 1 & 1\end{bmatrix}$. *Then if* $p>2$, *the direct limit* W *of the sequence*

$$V\xrightarrow{T} V\xrightarrow{T}\ldots$$

is a rank 2 *module with no non-zero infinitely divisible elements, yet is not a free* $\mathbb{Z}_{(p)}$*-module.*

Proof. Since T is a rational isomorphism, we see immediately that $W\otimes\mathbb{Q}\cong V\otimes\mathbb{Q}\cong\mathbb{Q}\oplus\mathbb{Q}$, and W therefore has rank 2.

We next consider the set $U\subset V$ defined by the condition that v is in U if given k there exists l such that $T^l(v)=p^k u$ for some u in V. It is easy to check that U is a T-stable submodule of V. In fact U is a direct summand. For suppose that λv is in U for some v in V and λ in $\mathbb{Z}_{(p)}$. Express λ as $p^r\mu$, where $r\geqq 0$ and μ is invertible in $\mathbb{Z}_{(p)}$. Given k, there is an l such that $T^l(\lambda v)=p^{k+r}u$ for some u in V. Thus $T^l(v)=p^k(\mu^{-1}u)$, so v is in U and U must be a direct summand since $\mathbb{Z}_{(p)}$ is principal ideal domain.

Now we observe that the set of infinitely divisible elements of W is the direct limit of

$$U\xrightarrow{T} U\xrightarrow{T}\ldots.$$

Thus to demonstrate that there are no such non-zero elements, we can proceed by showing $U=0$. Note that V cannot admit a T-stable summand of rank one because T has non-rational eigenvalues. Hence $U=0$ or V. But $U\neq V$ since the congruence

$$\begin{bmatrix}0 & p\\ 1 & 1\end{bmatrix}^l\begin{bmatrix}1\\ 0\end{bmatrix}\equiv\begin{bmatrix}0\\ 1\end{bmatrix}\bmod p$$

implies $T^l(1,0)$ is not divisible by p for any l. So indeed $U=0$.

On the other hand, W cannot be free of rank 2. For tensoring the original sequence with $\mathbb{F}_p$ turns T into a rank one idempotent, yielding $W\otimes\mathbb{F}_p\cong\mathbb{F}_p$.

In the lemmas which follow we explain how to avoid such problems by lifting an idempotent in such a way as to induce an almost idempotent in integral homology. We can then produce p-local spaces $Z(\rho)$, of finite $\mathbb{Z}_{(p)}$ type, such that $Y_{(p)}\simeq\bigvee_\rho d(\rho)Z(\rho)$. The section concludes by delocalising $Z(\rho)$, in order to produce complexes $Y(\rho)$ as specified by Theorem 2.13.

We begin with

2.3. Lemma. *Given any modular idempotent* e *in* A *and integer* $l>0$, e *admits a lift* f *to a formal sum of* $n\times n$ *integer matrices such that* $f_*\colon H_*(Y)\to H_*(Y)$ *satisfies* $f_*^2\equiv f_*\bmod p^l$.

Proof. Note first that $f_*^2\equiv f_*\bmod p$ for any lift f of e. Then f^{p^l} is a lift of e with the stated property.

2.4. Lemma. *Any linear transformation of the vector space of homogeneous polynomials of degree k in $\mathbb{Q}[x_1, \ldots, x_n]$ can be realised as a finite rational combination of linear substitutions $\sum v_j S_j$.*

Proof. Let $T(\alpha, \gamma)$ denote the linear transformation which sends a prescribed monomial $\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ to $\gamma = x_1^{\gamma_1} \ldots x_n^{\gamma_n}$, and all other monomials to zero. It suffices to demonstrate that $T(\alpha, \gamma)$ can be realised as $\sum v_i S_i$.

We first consider the vector space $\mathbb{Q}[x]$. Let $\lambda_i \in \mathbb{Q}$, $1 \leqq i \leqq k+1$, be distinct and define $T_i(x) = \lambda_i x$. Then for all $\mu_i \in \mathbb{Q}$,

$$(\sum \mu_i T_i)(x^s) = (\sum \mu_i \lambda_i^s) x^s.$$

Since the Vandermonde determinant

$$\det \begin{pmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 \ldots & \lambda_{k+1} \\ \vdots & \vdots & \vdots \\ \lambda_1^k & \lambda_2^k \ldots & \lambda_{k+1}^k \end{pmatrix}$$

is non-zero, given r we can find μ_i such that $(\sum \mu_i T_i) x^s = \delta_{r,s} x^s$ for $0 \leqq s \leqq k$. By applying such a transformation to each variable of α in turn, we realise $T(\alpha, \alpha)$ as $\sum v_i S_i$.

For the next step, observe that all monomials can be driven into $\beta = x_1^k$ by substituting x_1 for each of $x_2, x_3, \ldots, x_k$. So by composing with $T(\alpha, \alpha)$, we see that $T(\alpha, \beta)$ can be realised as required. It thus remains to realise $T(\beta, \gamma)$.

To do this, substitute for x_1 in $\beta = x_1^k$ a general linear form $\sum \eta_i x_i$. In the resulting multinomial expansion, we can apply $T(\gamma, \gamma)$ to kill all monomials except γ; hence a suitable choice of $\eta_i \in \mathbb{Q}$ will give $T(\beta, \gamma)$ as $\sum v_i S_i$, and the proof is complete.

2.5. Corollary. *For any endomorphism T: $H_k(Y) \to H_k(Y)$, there is an integer m, depending on k but not on T, such that $mT = g_*$ for some formal sum g of integer matrices acting on Y.*

Proof. Consider $T \otimes 1_{\mathbb{Q}}$. Identifying $H_*(Y) \otimes \mathbb{Q}$ *with* $\mathbb{Q}[x_1, \ldots, x_n]$, this becomes a linear transformation on the space of homogeneous polynomials of degree k.

Applying 2.4, we can multiply by a sufficiently large integer m, depending on k but not T, such that mT is induced by an integral combination of integral substitutions, i.e. $mT = g_*$ as sought.

2.6. Lemma. *For a given integer $k \geqq 0$ and idempotent e in A, there exists an integral lift f, depending on k, such that f_*: $H_k(Y) \to H_k(Y)$ is almost idempotent, in the sense that $f_*^2 = q f_*$ for some q prime to p.*

Proof. Referring to 2.5, let $m = r p^{l-1}$ where $l \geqq 1$ and $(r, p) = 1$, where r, l depend only on k and not on T. Applying 2.3, we may then select a lift f of e satisfying $f_*^2 \equiv f_* \bmod p^l$.

Choose a basis for $H_k(Y)$, and let f_* have integer matrix M. Since M is idempotent mod p^l, there is an integer matrix C, invertible over $\mathbb{Z}/p^l$, such that

$$CMC^{-1} = J \bmod p^l, \quad \text{where } J = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Hence $CM = JC + p^l B$, and $(r \det C)M = (r \operatorname{adj} C)JC + (pm \operatorname{adj} C)B$.

Now applying (2.5) to $(\operatorname{adj} C)B$, $(m \operatorname{adj} C)B = g_*$ by our initial choice of m. Therefore $(r \det C)f - pg$ has matrix $(r \operatorname{adj} C)JC$ acting on $H_k(Y)$, which is almost idempotent in the stated sense. To conclude, if s satisfies $sr \det C \equiv 1 \bmod p$, then

$$s[(r \det C)f - pg]$$

is a formal sum of integer matrices which reduces mod p to e and has the required property.

We need the following elementary results to clarify our later geometric constructions.

2.7. Lemma. *Let W be a $\mathbb{Z}_{(p)}$ module which embeds as a submodule of $\bigoplus_k \mathbb{Q}$, and which admits a module surjection onto $\bigoplus_k \mathbb{F}_p$. Then $W \cong \bigoplus_k \mathbb{Z}_{(p)}$.*

Proof. Choose $w_1, \ldots, w_k$ in W whose images under the surjection form a basis for $\bigoplus_k \mathbb{F}_p$: their images under the embedding in $\bigoplus_k \mathbb{Q}$ are rationally independent. It is now simple to check that $w_1, \ldots, w_k$ are a basis for W over $\mathbb{Z}_{(p)}$.

We can paraphrase the hypotheses of 2.7 by saying that the rational rank of W agrees with its mod p rank. This is what fails in Example 2.2.

2.8. Lemma. *Let V be a free $\mathbb{Z}_{(p)}$ module of finite rank. Let*

$$V \xrightarrow{T_1} V \xrightarrow{T_2} \ldots$$

be a sequence of $\mathbb{Z}_{(p)}$-module maps such that each T_i reduces mod p *to the same idempotent transformation E of rank k. Suppose further that the rational rank of each T_i is also k. Then the direct limit W of the sequence is a free $\mathbb{Z}_{(p)}$-module of rank k.*

Proof. We view V as embedded in the rational vector space $V \otimes \mathbb{Q}$. The direct limit of the sequence

$$V \otimes \mathbb{Q} \xrightarrow{T_1 \otimes 1} V \otimes \mathbb{Q} \xrightarrow{T_2 \otimes 1} \ldots$$

is a rational vector space of dimension at most k in which W is embedded as a $\mathbb{Z}_{(p)}$ submodule. On the other hand W projects onto the direct limit of the sequence

$$V \otimes \mathbb{F}_p \xrightarrow{T_1 \otimes 1} V \otimes \mathbb{F}_p \xrightarrow{T_2 \otimes 1} \ldots$$

which by hypothesis is the image of the idempotent E and is therefore isomorphic to $\bigoplus_k \mathbb{F}_p$. The conclusion follows from 2.7.

We are now ready to define our p-local spaces $Z(\rho)$.

2.9. Proposition. *Choose an idempotent e in A. Then there is a p-local space Z of finite $\mathbb{Z}_{(p)}$-type and a map $\psi\colon Y_{(p)}\to Z$ such that*

$$\psi_*\colon eL^*\to H_*(Z;\mathbb{F}_p) \quad \text{and} \quad \psi^*\colon H^*(Z;\mathbb{F}_p)\to Le$$

are isomorphisms.

Proof. We commence by constructing an integral lift $f=F_k$ of e, as in Lemma 2.6. Thus $f_*\colon H_k(Y)\to H_k(Y)$ is almost idempotent, and has the same rational and mod p rank.

Mimicking a trick in [6], take $G_k=F_kF_{k-1}\ldots F_1$. Then the telescope construction applied to

$$Y_{(p)}\xrightarrow{G_1} Y_{(p)}\xrightarrow{G_2}\cdots$$

produces a p-local space Z. Moreover, each

$$(G_k)_*\colon H_i(Y_{(p)};\mathbb{F}_p)\to H_i(Y_{(p)};\mathbb{F}_p), \qquad i\geqq 1$$

is the original e, whence the natural map $\psi\colon Y_{(p)}\to Z$ has the required properties.

Remark that G_i has the factor F_k for each $i\geqq k$, so that $(G_i)_*$ has the same rational and mod p rank. Thus 2.8 tells us that $H_i(Z)$ is a free $\mathbb{Z}_{(p)}$-module of finite rank.

We can now state the preliminary local version of Theorem 2.13, to be compared with 2.1.

2.10. Proposition. *For each ρ there is a CW complex $Z(\rho)$, of finite type over $\mathbb{Z}_{(p)}$, admitting a homotopy equivalence*

$$\phi\colon Y_{(p)}\xrightarrow{\simeq}\bigvee_{\rho} d(\rho)\,Z(\rho).$$

Also, $\phi^*\colon H^*(Z(\rho);\mathbb{F}_p)\to Le$ is an isomorphism, where e is a primitive idempotent associated with ρ.

Proof. Proposition 2.9 furnishes us with a space $Z(i,\rho)$ for each primitive idempotent $e_{i\rho}$ in an expression

$$1=\sum_{i,\rho} e_{i\rho} \quad \text{in } A.$$

The corresponding maps $\psi\colon Y_{(p)}\to Z(i,\rho)$ can be added using the suspension coordinate of Y to create a map

$$\phi\colon Y_{(p)}\to\bigvee_{i,\rho} Z(i,\rho)$$

which by construction induces an $\mathbb{F}_p$ homology isomorphism realising the splitting

$$H_*(Y_{(p)};\mathbb{F}_p)\cong L^*\cong\bigoplus_{i,\rho} e_{i\rho}L^*.$$

By the nature of the spaces involved, ϕ is then a homotopy equivalence.

Given a second expression

$$1=\sum_{i,\rho} e'_{i\rho} \quad \text{in } A,$$

then by Proposition 1.2 there is δ in A such that $e'_{i\rho} = \delta e_{i\rho} \delta^{-1}$ for all i, ρ. A lift Δ of δ to a formal sum of integer matrices can then be interpreted as a self map of Y carrying $e_{i\rho}L^*$ isomorphically onto $e'_{i\rho}L^*$ in $\mathbb{F}_p$ homology.

Choose $j\colon Z(i,\rho) \to Y_{(p)}$ to be a restricted inverse of ϕ. Then $\psi' \Delta j$ induces an isomorphism

$$H_*(Z(i,\rho); \mathbb{F}_p) \to H_*(Z'(i,\rho); \mathbb{F}_p),$$

whence $Z(i,\rho)$ and $Z'(i,\rho)$ have the same homotopy type. It is thus legitimate in our description of ϕ to replace the bouquet $\bigvee_i Z(i,\rho)$ by $d(\rho)Z(\rho)$ for each ρ in turn.

For our applications in §3 it is more convenient to work with CW-complexes of genuine finite type, and to prove Theorem 2.13 it remains to delocalise the spaces $Z(\rho)$.

2.11. Lemma. *The space Y is rationally homotopy equivalent to a bouquet of odd dimensional spheres.*

Proof. The inclusion map $S^2 \to \mathbb{C}P^\infty$ has degree 1, and its m-fold product $f\colon \times_m S^2 \to \mathbb{C}P^\infty$ has degree $m!$ Thus the composite ϕ_m

$$S^{2m+1} \to \Sigma(\times_m S^2) \xrightarrow{\Sigma f} \Sigma\mathbb{C}P^\infty,$$

where the first map splits off the top cell, induces an isomorphism of $(2m+1)$-th rational homology groups. So

$$\phi = \bigvee_m \phi_m \colon \bigvee_{m=1}^{\infty} S^{2m+1} \to \Sigma\mathbb{C}P^\infty$$

is a rational equivalence.

Hence so is

$$\phi \wedge 1 \colon \bigvee_{m=1}^{\infty} S^{2m+1} \wedge \mathbb{C}P^\infty \to \Sigma\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty,$$

and repeating our initial argument for each $S^{2n} \wedge \Sigma\mathbb{C}P^\infty$ gives a rational equivalence between a bouquet of odd dimensional spheres and $\Sigma\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty$. Since

$$\Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \simeq \Sigma\mathbb{C}P^\infty \vee \Sigma\mathbb{C}P^\infty \vee (\Sigma\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty),$$

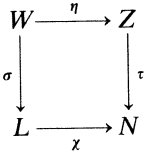
we have our result for $n=2$.

The general case follows by induction.

2.12. Proposition. *Let $Z(\rho) = Z$ be a p-local space of finite type over $\mathbb{Z}_{(p)}$ as constructed in* 2.9 *and* 2.10. *Then there is a CW-complex $Y(\rho) = W$ of finite type such that $W_{(p)} \simeq Z$. The cell structure of W can be chosen such that there are no cells in positive even dimensions, and $n(\rho, r)$ cells in dimension $2r+1$.*

Proof. From Lemma 2.11, Y is rationally equivalent to a bouquet of odd dimensional spheres, and by the construction of 2.9 $\psi\colon Y_{(p)} \to Z$ is a surjection in rational homology. Thus Z too admits a rationalisation of the form $\tau\colon Z \to \bigvee_i S^{2m_i+1}_{(0)} = N$. Choose $\bigvee_i S^{2m_i+1}\left[\frac{1}{p}\right] = L$, so that we have a second rationalisation $\chi\colon L \to N$.

Next consider the fibre square



defined by τ and χ, so that σ has the same fibre, say F, as τ. The homotopy exact sequence of τ shows that F is a simply connected torsion space, whence

$$H_i(N; H_j(F))=0 \quad \text{for all } i,j>0.$$

The homology spectral sequence for τ indicates that $\tilde{H}_*(F)$ is p primary, which in turn implies

$$H_i(L; H_j(F))=0 \quad \text{for all } i,j>0$$

since L is localised away from p. A comparison of the spectral sequences for σ and τ now shows that all differentials from x-axis to y-axis are surjective, and $\eta_*: H_*(W)\to H_*(Z)$ is therefore an injection whose image is a finitely generated free abelian subgroup of full rank. Thus W is of finite type, and $W_{(p)}\simeq Z$.

Finally, invoking Theorem A of [5] enables us to realise W as a minimal cell complex whose cells correspond to the generators of $H_k(W)$. So

$$\operatorname{rank} H_k(W)=\operatorname{rank} H_k(Z)=\operatorname{rank} e_\rho.$$

The proof is completed by appealing to Lemma 1.3.

The space X corresponding to $Z=Z(\rho)$ will be denoted by $Y(\rho)$. Combining 2.10 with 2.12, we arrive at the following result.

2.13. Theorem. *There are CW-complexes $Y(\rho)$ of finite type such that Y and the bouquet $\bigvee_\rho d(\rho)\,Y(\rho)$ are homotopy equivalent after p-localisation. Moreover, $Y(\rho)$ has no cells in even positive dimensions, and $n(\rho, r)$ cells in dimension $2r+1$.*

A convenient formulation provided by the theory of p-universal spaces, as expounded for example in [16], is given by

2.14. Corollary. *For any given $k>0$, there are maps of k-skeleta*

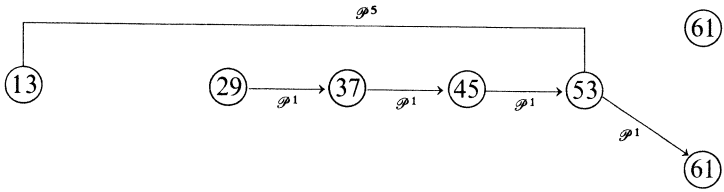
$$Y^{(k)} \leftrightarrows \bigvee_\rho d(\rho)\,Y(\rho)^{(k)}$$

which induce the splittings of mod p *cohomology and homology by a given set of primitive idempotents $e_{i\rho}$, and which are mutually inverse homotopy equivalences after p-localisation.*

Recent work of D. Notbohm indicates that with a careful choice of the complexes $Y(\rho)$, the corollary applies without the restriction to skeleta.

We conclude with a couple of remarks which are of particular relevance to the next section.

Firstly, as promised in §1, we exhibit the 61-skeleton of the complex $Y(D)$ associated to the determinant representation of $GL(2, 5)$.



The pattern is the same for any prime p up to dimension $2(p^2+p)+1$.

Analogous cell diagrams can be constructed for all the pieces $Y(D^k)$ associated to the 1-dimensional representations D^k. For example, the bottom cell of $Y(D^{-1})$ is represented cohomologically by the eigenpolynomial $(x^p y - x y^p)^{p-2}$.

Secondly, we pinpoint the complex $Y(D^{-1})$ as having a special role. This will be illustrated in action by the proof of (3.7), which builds on the above observation concerning the bottom cell when $n=2$.

2.15. Theorem. *The n-fold transfer is null homotopic on all pieces of the splitting*

$$Y_{(p)} \simeq \bigvee_{\rho} d(\rho) Z(\rho)$$

except for $Z(D^{-1})$.

A proof of this will appear in forthcoming works, both by the authors and by S. Mitchell.

3. Framed Hypersurfaces and the β-Family

Consider the usual inclusion $j_n: SO(n) \to \Omega^n S^n$ of the special orthogonal group into the identity component of the space of n-fold loops on S^n. In the limit we obtain the standard map $J: SO \to QS^0 = \lim \Omega^n S^n$ by $J(x) = j_n(x) * (-1)$, where $-1 \in QS^0$ is a fixed map of degree -1, $*$ denotes loop sum, and n is suitably large. So J maps into the component of zero degree loops.

Next, we can utilise the loop sum operation in QS^0 to extend J to a map $J_1: \Omega\Sigma SO \to QS^0$, whose image in homotopy consists of those elements of π_*^S representable by framed hypersurfaces. To interpret this geometrically, consider $\alpha \in \pi_r \Omega\Sigma SO \cong \pi_{r+1} \Sigma SO$. Applying the Pontrjagin-Thom construction, we can represent α by a map $f: M^r \to SO$, where $M^r \subset S^{r+1}$ is an oriented hypersurface. We then twist the canonical trivial framing of M^r in a high dimensional sphere by f, and so obtain $J_1(\alpha) \in \pi_r^S \cong \pi_r QS^0$. More details of these ideas can be found in [8], and of the underlying homotopy theory in [25].

In the now classical case of the J map, the groups $\pi_* SO$ are known by Bott periodicity; and the image of J in homotopy has been computed by Adams [1] and Mahowald [13]. However, the groups $\pi_* \Sigma SO$ are unknown, and the determination of the image of J_1 in homotopy seems to be a difficult problem.

In fact the problem has several variants, of equal interest and probably as intractable. For example, we can restrict attention to the subspace $\mathbb{R}P^\infty \subset SO$, arising as the limit of the usual embedding $\mathbb{R}P^n \to SO(n+1)$. The restriction $J_1': \Omega\Sigma\mathbb{R}P^\infty \to QS^0$ of the map J_1 has as image in homotopy those elements representable by hypersurfaces whose twisting function factors through $\mathbb{R}P^\infty$.

The maps J_1 and J_1' extend respectively to the *stable J-map* $J_\infty: QSO \to QS^0$ and the *real transfer* $t: Q\mathbb{R}P^\infty \to QS^0$. Again this utilises the loop sum operation in QS^0, and t is the obvious restriction of J_∞.

A further refinement involves iterating both the J-maps and the transfer. This time, consider matrix multiplication $SO \times SO \to SO$, and apply the adjoint of the Hopf construction to form $SO \wedge SO \to \Omega\Sigma SO$. We refer to the composite

$$SO \wedge SO \longrightarrow \Omega\Sigma(SO \wedge SO) \longrightarrow \Omega\Sigma SO \xrightarrow[J_1]{} QS^0$$

as the *double J-map:* it does not factor through the J-map. Once more we can apply the functor Q, and obtain the extension

$$Q(SO \wedge SO) \longrightarrow QSO \xrightarrow[J_\infty]{} QS^0.$$

Alternatively, we can construct the *double transfer*

$$t(2): Q(\mathbb{R}P^\infty \wedge \mathbb{R}P^\infty) \to QS^0$$

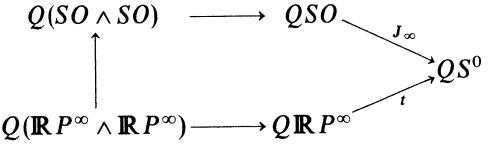
as the composite

$$Q(\mathbb{R}P^\infty \wedge \mathbb{R}P^\infty) \longrightarrow Q\mathbb{R}P^\infty \xrightarrow[t]{} QS^0,$$

where the first map is the natural extension of

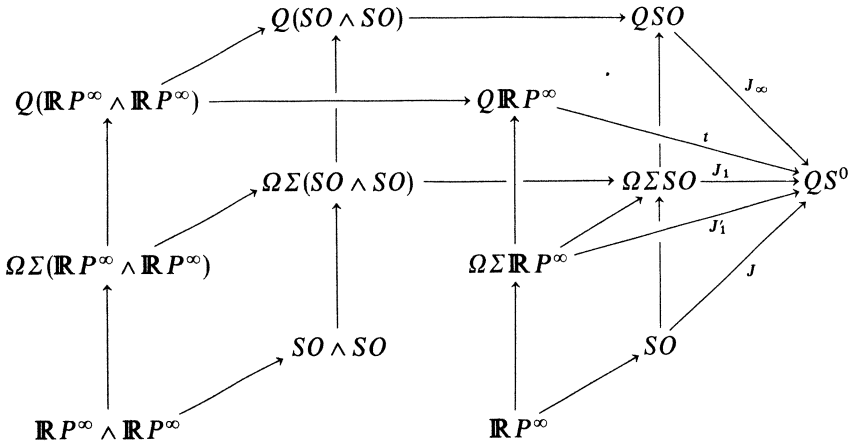
$$Q\mathbb{R}P^\infty \wedge \mathbb{R}P^\infty \xrightarrow[t \wedge 1]{} QS^0 \wedge \mathbb{R}P^\infty \longrightarrow Q\mathbb{R}P^\infty.$$

This compares with the double J-map via the expected pentagon



whose homotopy commutativity follows, for example, from [9, 2.2]. On homotopy groups, this pentagon tells us we can perform the double transfer on a singular framed manifold $f: M \to \mathbb{R}P^\infty \times \mathbb{R}P^\infty$ by suitably twisting the framing.

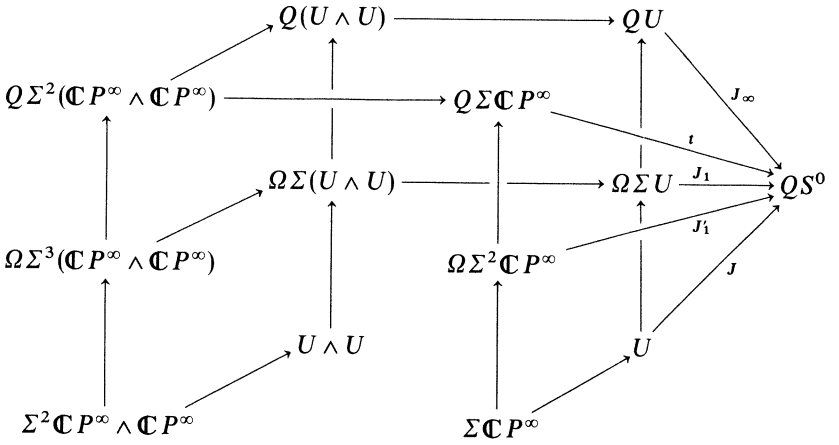
The following homotopy commutative diagram summarises the relationships between all the maps so far defined.



Note that we cannot insert the maps $\Omega\Sigma(\mathbb{R}P^\infty \wedge \mathbb{R}P^\infty) \to \Omega\Sigma\mathbb{R}P^\infty$ and $Q\mathbb{R}P^\infty \to QSO$, determined respectively by tensor product of line bundles and by inclusion, without destroying the commutativity of the diagram. Any map factoring through J_1 will have image in homotopy consisting of elements representable by framed hypersurfaces with special additional features depending on the map.

The Kahn-Priddy theorem [11] asserts that t and J_∞ are surjective in 2-primary homotopy, whilst it was proved in [20] that members of the μ-family in ${}_2\pi^S_*$ are not in the homotopy image of J_1. On the other hand, any element of π^S_* representable by the canonical framing on a compact Lie group [18] can also be represented by a framed hypersurface: for example $\nu \in \pi^S_8$, $\kappa \in \pi^S_{14}$ and $\beta_1 \in {}_3\pi^S_{10}$ are all in the homotopy image of J_1. It is also known that elements of Kervaire invariant 1 are not in the homotopy image of J'_1, but it remains an unsolved problem whether such elements lie in the homotopy image of J_1 in dimensions $\geqq 30$.

The whole edifice described above has a complex analogue founded on the unitary group $U \subset SO$, in which $\mathbb{R}P^\infty$ is replaced by the subspace $\Sigma\mathbb{C}P^\infty \subset U$ specified by complex reflection. Recalling from Knapp [12] and Stolz [22] that, up to sign, $J_\infty : QU \to QS^0$ restricts to the complex transfer $t : Q\Sigma\mathbb{C}P^\infty \to QS^0$, we have the diagram



Since hereafter we are concerned exclusively with the complex case, we can use the above notation without ambiguity: in particular

$$t(2) : Q\Sigma^2(\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty) \to QS^0$$

stands for the double complex transfer.

Note once more that we cannot insert maps $\Omega\Sigma^3(\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty) \to \Omega\Sigma^2\mathbb{C}P^\infty$ and $Q\Sigma\mathbb{C}P^\infty \to QU$, without destroying commutativity. Also, applying homotopy to the top pentagon expresses the fact that the double transfer on a singular framed manifold $f : M \to \mathbb{C}P^\infty \times \mathbb{C}P^\infty$ is represented by suitably reframing $M \times S^1 \times S^1$.

We are especially interested in the composite $\Sigma^2\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty \to QS^0$, which factors through $t(2)$. Many elements of $\pi_i\Sigma^2\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty$ will arise by suspen-

sion from $\pi_{i-1}\Sigma\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty$, so we wish to analyse the space $\Sigma\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty$. Since, as we have already implicitly assumed, there is the equivalence

$$\Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \simeq \Sigma\mathbb{C}P^\infty \vee \Sigma\mathbb{C}P^\infty \vee (\Sigma\mathbb{C}P^\infty \wedge \mathbb{C}P^\infty)$$

it suffices to study $\Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$, to which our splitting naturally applies. However, so far as producing framed hypersurfaces is concerned it is enough to start with elements in $\pi_i \Sigma^3 \mathbb{C}P^\infty \wedge \mathbb{C}P^\infty$.

To explain our approach to Theorem 3.7 we outline the argument given in [19]. Let f: $M^{2p+2} \to \Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$ be a singular oriented hypersurface representing a certain $x_0 \in \pi_{2p+3}\Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$, where p is henceforth restricted to be an odd prime. Using the H-space structure on $\mathbb{C}P^\infty \times \mathbb{C}P^\infty$ we can form the $(p-2)$-fold product g: $N^r \to \mathbb{C}P^\infty \times \mathbb{C}P^\infty$, where $N = \underset{p-2}{\times} M$ and $r = (p-2)(2p+2)$, which represents the Pontrjagin power $x_0^{p-2} \in \pi_{r+1}\Sigma(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$. Referring to 3.1 we deduce that the double transfer of x_0^{p-2}, which is easily computed to be β_1, is representable on $N \times S^1 \times S^1$ with a suitable framing.

Applying 3.2 below, which we shall use again later, finally establishes that $N \times S^1 \times S^1$ is a hypersurface.

3.2. Lemma. *Let $M^m \times \mathbb{R} \subset \mathbb{R}^{m+1}$ be an oriented hypersurface, and $K \times \mathbb{R}^n \subset \mathbb{R}^{k+n}$ be a framed submanifold. Then the cartesian product $\underset{a}{\times} K \underset{b}{\times} M$ is an oriented hypersurface if $b > 0$ and $m > n$.*

Proof. This involves the following well-known process. Iterating the given embeddings yields the sequences

$$(\underset{b}{\times} M) \times \mathbb{R} \subset (\underset{b-1}{\times} M) \times \mathbb{R}^{m+1} \subset \ldots \subset \mathbb{R}^{bm+1}$$
$$(\underset{a}{\times} K) \times \mathbb{R}^p \subset (\underset{a-1}{\times} K) \times \mathbb{R}^{k+n} \subset \ldots \subset \mathbb{R}^{ak+n}.$$

Since $b > 0$ and $m > n$, these may be combined to give the embedding

$$(\underset{a}{\times} K \underset{b}{\times} M) \times \mathbb{R} \subset (\underset{a}{\times} K) \times \mathbb{R}^{bm+1} \subset \mathbb{R}^{ak+bm+1}.$$

Embarking on our proof of 3.7, we must begin by explaining and refining some of Knapp's results.

3.3. Theorem (Knapp [12, p. 46]). *Let u and v be the standard generators of the divided power algebra $H_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty; \mathbb{Z})$, which we identify with the subring $R \subset \mathbb{Q}[u, v]$ generated by divided monomials in u and v (see § 1). Then the classes*

$$h_0 = (u^p v - u v^p)/p,$$
$$h_1 = (h_0^p - h_0(v^{p^2-1} + u^{p^2-1} - (uv)^{(p^2-1)/2}))/p$$

are stably spherical, and if

$$x_0 \in \pi^S_{2p+2}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty), \quad x_1 \in \pi^S_{2p^2+2p}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$$

have respective Hurewicz images h_0 and h_1, then

$$t(2)(x_0^{p-i-1} x_1^{i-1}) = \beta_i \quad (1 \leqq i \leqq p-1) \quad \text{and} \quad t(2)(x_0^{p-1} x_1^{p-2}) = \beta_p.$$

Knapp's point of view is to study the $Ad_*(\ )$ and $A_*(\ )$ groups of $\mathbb{C}P^\infty \times \mathbb{C}P^\infty$. Implicit in his work, and simpler to utilize for our purposes, is the following reformulation.

Let $PK_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \subset K_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$ denote the subgroup of primitives under the coaction

$$K_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty) \to K_*(K) \otimes K_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$$

of the complex K-theoretic co-operations. Then a simple computation shows that $PK_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$ may be identified via the Chern character as the subalgebra of $\mathbb{Q}[u,v]$ consisting of degree i homogeneous *numerical polynomials*, i.e. $\{k(u,v) | k$ is homogeneous of degree i, $k(\mathbb{Z} \times \mathbb{Z}) \subset \mathbb{Z}\}$. This is described implicitly by Knapp in [12, p. 43], can be deduced from Schwartz [23], and is proved in detail in Baker-Clarke-Ray-Schwartz [2].

Now by very definition there is a natural isomorphism for each p

$$A_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty_{(p)}) \xrightarrow{\cong} PK_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty_{(p)}).$$

Moreover, Knapp's lemma [12, 4.5] informs us that the Hurewicz map

$$\pi^S_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty_{(p)}) \to A_{2i}(\mathbb{C}P^\infty \times \mathbb{C}P^\infty_{(p)})$$

is an isomorphism mod torsion in dimensions $0 \leqq i \leqq 2p(p-1)^2+1$. This enables us to identify the classes x_0 and x_1 with the numerical polynomials h_0 and h_1 in 3.3.

Note that the mod p reduction of h_0 is the class $\bar{h}_0$ introduced in §1. Since each of u^{p^2-1}, v^{p^2-1} and $(uv)^{(p^2-1)/2}$ is divisible by p^2 in $H_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty; \mathbb{Z})$, the mod p reduction $\bar{h}_1$ of h_1 is the eigenpolynomial $\bar{h}_0^p/p$ of 1.7.

We can now describe the refinement of 3.3 which we shall need.

3.4. Proposition. *Let $x_0, x_1 \in \pi^S_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty)$ have* mod p *Hurewicz images $\bar{h}_0$ and $\bar{h}_1$.*

Then products of x_0 and x_1 transfer up to non-zero multiples as in 3.3.

Proof. A careful inspection of Knapp's proof of 3.3, viz [12, p. 49 and p. 65], reveals that the crucial property of h_1 which makes the double transfer calculations work is the presence of the term $u^{p^2}v^p/p^{p+1}$. Note that this term is itself $0 (\mathrm{mod}\, p)$ in $H_*(\mathbb{C}P^\infty \times \mathbb{C}P^\infty; \mathbb{Z})$.

Now let $k(u,v)$ be homogeneous and numerical, such that $k(u,v) = \bar{h}_0^p/p \,(\mathrm{mod}\, p)$. We can assume $k(u,v)$ is skew-symmetric, since it may be replaced by $\frac{1}{2}(p+1)(k(u,v)-k(v,u))$. From our introductory remarks, it suffices to show that $k(u,v)$ contains a term $a'u^{p^2}v^p/p^{p+1}$, where $a' \not\equiv 0 (\mathrm{mod}\, p)$. So to obtain a contradiction, let

$$k(u,v) = \sum_{j=1}^{p-1} \frac{(-1)^j (p-1)!}{(p-j)!\,j!} \frac{u^{j(p-1)+p} v^{(p-j)(p-1)+p}}{p^p} + a(u^{p^2}v^p - u^p v^{p^2})/p^p$$

$$+ \sum_m c_m \frac{u^m v^{p^2+p-m}}{p^p} + \sum_n c_n \frac{u^n v^{p^2+p-n}}{p^{p-1}} + d(u^{p^2+p} - v^{p^2+p})/p^{p+1}$$

where $a, b, c_m, c_n, d \in \mathbb{Z}$ and

$$0<m<p \quad \text{or} \quad p^2<m<p^2+p \quad \text{or} \quad p<m<p^2 \quad \text{with} \quad p \mid m$$
$$p<n<p^2 \quad \text{with} \quad p \nmid n.$$

Then $k(\lambda p, 1) \in \mathbb{Z}$ for each $1 \leqq \lambda \leqq p$, which implies

(i) $d = pc_0$ for some $c_0 \in \mathbb{Z}$,

(ii) $\sum_{m=0}^{p-1} c_m \lambda^m p^m \equiv 0 \pmod{p^p}$, $1 \leqq \lambda \leqq p$.

In matrix terms, Eqs. (ii) become

$$\begin{pmatrix} 1 & 1 & 1 \ldots 1 \\ 2^0 & 2^1 & 2^2 \ldots 2^{p-1} \\ \vdots & \vdots & \vdots \quad \vdots \\ p^0 & p^1 & p^2 \ldots p^{p-1} \end{pmatrix} \begin{pmatrix} c_0 p^0 \\ c_1 p^1 \\ \vdots \\ c_{p-1} p^{p-1} \end{pmatrix} \equiv 0 \pmod{p^p}.$$

Since the Vandermonde determinant $\not\equiv 0 \pmod{p}$, we deduce that $c_m p^m \equiv 0 \pmod{p^p}$ for each $m = 0, 1, \ldots, p-1$. Thus $p^{p-m} | c_m$, and by skew symmetry $p^{p-m} | c_{p^2+m}$, for each $1 \leqq m \leqq p$ (where $c_{p^2+p} = c_0$).

Substituting this information into $k(p-1, 1) - k(1, 1)$ yields a rational number with at least p^{p-1} in the denominator, contradicting the fact that $k(u, v)$ is numerical. Thus our original assumption that $u^{p^2} v^p$ occurs with coefficient a/p^p is false; rather it must have coefficient a'/p^{p+1}, $a' \not\equiv 0 \pmod{p}$, as sought.

Thus to construct our proof of 3.7, it suffices to identify x_0 and x_1 in mod p homology, which is exactly what our calculations in §2 are geared to do. We have to produce these classes in unstable homotopy, and so require two lemmas. The first produces x_0 explicitly without recourse to Knapp's work, whilst the second takes Knapp's x_1 and desuspends it into the required unstable group.

Recall that D stands for the determinant representation of $GL(2, p)$. From 1.6 and 2.14, $\bar{h}_1$ is clearly carried by a cell of $Y(D)$ for any choice of the primitive idempotent e_D.

3.5. Lemma. *The idempotent e_D can be chosen in such a way that the bottom cell of $Y(D)$ carries $\bar{h}_0$.*

Proof. Consider the subgroup $H \subset GL(2, p)$ consisting of matrices of the form $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ or $\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}$. Noe $e = \frac{1}{|H|} \sum_h (\det h)^{-1} h$ is an idempotent for the determinant representation of H, and we can assume that $e = e_D + f$ for some idempotent f orthogonal to e_D. Then the image of e, which contains the image of e_D, is easily seen to be generated by $\bar{h}_0$ in dimension $2p+3$. The result follows from 2.14.

3.6. Lemma. *Let p be an odd prime. Let W be an $(n-1)$-connected CW-complex and let λ be the number of values of k in the range $n \leqq k < \frac{1}{2}p(n+1) - 1$ for which a k-cell exists. Then if $t \leqq p(n+1) - 3$, any element in the p-primary part of the*

stable homotopy group $\pi_t^S(W)$ can be represented by an unstable map $S^{t+2\lambda} \to \Sigma^{2\lambda} W$.

Proof. We proceed by induction on λ. Consider the case $\lambda = 1$. Here it suffices to take W as a wedge of n-spheres V with the next cell occurring in dimension greater than $\frac{1}{2}p(n+1)-2$. Let Z be the quotient complex W/V, which is $(\frac{1}{2}p(n+1)-2)$-connected.

The argument is based on the following diagram.

$$\begin{array}{ccccc}
\pi_t V & \longrightarrow & \pi_{t+2}\Sigma^2 V & \longrightarrow & \pi_t^S(V) \\
\downarrow & & \downarrow & & \downarrow \\
\pi_t W & \longrightarrow & \pi_{t+2}\Sigma^2 W & \longrightarrow & \pi_t^S(W) \\
\downarrow & & \downarrow & & \downarrow \\
\pi_t Z & \longrightarrow & \pi_{t+2}\Sigma^2 Z & \longrightarrow & \pi_t^S(Z) \\
\downarrow & & \downarrow & & \downarrow \\
\pi_t \Sigma V & \longrightarrow & \pi_{t+2}\Sigma^3 V & \longrightarrow & \pi_t^S(\Sigma V).
\end{array}$$

The horizontal sequences are induced by suspension and the vertical sequences arise from the cofibration $V \to W \to Z$. Thus the right hand vertical sequence is exact.

Now consider a p-primary element α in $\pi_t^S(W)$. Its image β in $\pi_t^S(Z)$ pulls back to β_1 in $\pi_t Z$ since we are in the geometric stable range for Z. The image γ_1 of β_1 in $\pi_t \Sigma V$ maps trivially to $\pi_t^S(\Sigma V)$ by the exactness of the right hand sequence. Now γ_1 is determined by its projection on to the individual spheres of V together with certain Whitehead products (which vanish on suspension). Using the p-primary stable range of spheres we see that γ_1 vanishes after a single suspension. From the properties of a cofibration it now follows that the double suspension of β_1 lifts to an element α_2 in $\pi_{t+2}\Sigma^2 W$. It remains to observe that α_2 can be chosen to stabilize to the original α. Suppose that α_2 stabilizes to α'. Then $\alpha - \alpha'$ maps to 0 in $\pi_t^S(Z)$ and so pulls back to some element ω in $\pi_t^S(V)$. Again using the p-primary stable range for spheres, we deduce that ω pulls back to ω_2 in $\pi_{t+2}\Sigma^2 V$. Thus $\alpha_2 + i_*\omega_2$ stabilizes to α as required to complete the proof for $\lambda = 1$.

The inductive step follows the same line of argument. Suppose W has λ dimensions (in the given range) where cells exist. Then Z has $\lambda - 1$ such dimensions, so we may take as our inductive hypothesis that the result is proved for Z with $2\lambda - 2$ suspensions. We now replace V, W, Z in the above diagram by their $(2\lambda-2)$-suspensions and proceed as in the case $\lambda = 1$.

This lemma is designed to serve our present purpose and is not intended to be the best desuspension result of its kind. A few remarks may make it more digestible. For n odd, $\hat{p}\pi_i S^n$ maps surjectively on $p\pi_t^S(S^n)$ for $t \leqq p(n+1)-3$. The lemma asserts that for a general W we still have this result, but at the

cost of a double suspension for each dimension in which a cell of W occurs up to the geometric dimension for desuspension. Notice that if $\lambda=0$ then W is in fact $(\frac{1}{2}p(n+1)-2)$-connected and the lemma reduces to part of the Freudenthal suspension theorem.

We are now fully equipped to prove our main result.

3.7. Theorem. *The elements $\beta_1,\dots,\beta_{p-2}$ and β_p of the β-family in ${}_p\pi_*^S$ can be represented by framed hypersurfaces. So too can the element γ_1, $p\geqq 5$, and the non-zero products $\alpha_1\beta_i$ for $1\leqq i\leqq p-2$.*

Proof. Recall from 2.14 that we have the maps of skeleta

$$Y(D)^{(k)} \xrightarrow{i} \Sigma(\mathbb{C}P^\infty\times\mathbb{C}P^\infty)^{(k)} \xrightarrow{\pi} Y(D)^{(k)}$$

for any $k\geqq 2p^2+2p+2$. Choosing $Y(D)$ as in 3.5, we can include the bottom cell via a map $S^{2p+3}\to Y(D)^{(k)}$. Composing with i yields $x_0\in\pi_{2p+3}\Sigma(\mathbb{C}P^\infty\times\mathbb{C}P^\infty)$ whose mod p Hurewicz image is $\bar{h}_0$. Thus $\bar{h}_0$ may be represented by a singular oriented hypersurface M^{2p+2} in $\mathbb{C}P^\infty\times\mathbb{C}P^\infty$.

Next, choose Knapp's x_1 as in 3.3, and consider $\pi_*(x_1)\in\pi^S_{2p^2+2p+1}(Y(D)^{(k)})$. Since $Y(D)$ satisfies the hypotheses for W in 3.6 with $\lambda=\frac{1}{2}(p-1)$, we deduce that $\pi_*(x_1)$ is realisable unstably in $\pi_{2p^2+2p+p}\Sigma^{p-1}Y(D)^{(k)}$. We obtain $i_*\pi_*(x_1)\in\pi_{2p^2+2p+p}\Sigma^p(\mathbb{C}P^\infty\times\mathbb{C}P^\infty)$, which has mod p Hurewicz image $\bar{h}_1$. Thus $\bar{h}_1$ may be represented by a singular manifold K^{2p^2+2p} in $\mathbb{C}P^\infty\times\mathbb{C}P^\infty$, where K is framed in codimension p.

Therefore we may examine the Pontrjagin products $x_0^{p-i-1}x_1^{i-1}$, $1\leqq i\leqq p-1$, and $x_0^{p-1}x_1^{p-2}\in\pi_*^S(\mathbb{C}P^\infty\times\mathbb{C}P^\infty)$ in the light of 3.2 and 3.4. The former proves that they are representable respectively on the singular hypersurfaces $N(i)=\underset{i-1}{\times}K\underset{p-i-1}{\times}M$ and $N(p)=\underset{p-2}{\times}K\underset{p-1}{\times}M$, whilst the latter confirms that they map respectively under the double transfer to β_i, $1\leqq i\leqq p-2$, and to β_p.

We conclude via 3.1 that $S^1\times S^1\times N(i)$ and $S^1\times S^1\times N(p)$ are hypersurface representatives for β_i and β_p. The representability of $\alpha_1\beta_i$, $i<p-1$, is also immediate, since α_1 arises by framing S^{2p-3}.

So far as γ_1 is concerned, recall from [15] that $\gamma_1=-\alpha_1\beta_{p-1}$. Our discussion above shows that β_{p-1} is representable on a manifold N frameable in codimension $p-2$. By 3.2, $S^{2p-3}\times N$ is a hypersurface, and it may be framed to represent γ_1.

Using the methods of §1 it is a straightforward matter to check that the elements $x_0^{p-i-1}x_1^{i-1}$, $1\leqq i\leqq p-2$, and $x_0^{p-1}x_1^{p-2}$ are carried on the piece $Y(D^{-1})$ of $\Sigma(\mathbb{C}P^\infty\times\mathbb{C}P^\infty)$. This is a good illustration of (2.15) at work.

To prove that β_{p-1} is also representable on a hypersurface, 3.4 shows that it would suffice to verify the following.

3.8. Conjecture. *The homology class*

$$\bar{h}_1=(u^pv-uv^p)^p/p^{p+1}\in H_{2p^2+2p+1}(Y(D);\mathbb{F}_p)$$

(see 1.6*) is spherical.*

Both F.P. Peterson and the authors have established this result for $p=5$, using a Postnikov analysis. For $p=3$ it is trivial.

References

1. Adams, J.F.: On the groups $J(X)$: IV. Topology **5**, 21–71 (1966)
2. Baker, A., Clarke, F., Ray, N., Schwartz, L.: Kummer congruences and the stable homotopy of BU. In preparation
3. Baker, A., Ray, N.: Some infinite families of U hypersurfaces. Math. Scand. **50**, 149–166 (1982)
4. Cohen, F.: Splitting certain suspension via self-maps. Ill. J. Math. **20**, 336–347 (1976)
5. Cooke, G.: Embedding certain complexes up to homotopy type in euclidean space. Ann. of Math. **90**, 144–156 (1969)
6. Cooke, G., Smith, L.: Mod$p$ decompositions of co H-spaces and applications. Math. Z. **157**, 155–177 (1977)
7. Curtis, M., Reiner, I.: Representations of finite groups and associative algebras. New York: Interscience 1962
8. Eccles, P.J.: Filtering framed bordism by embedding codimension. J. London Math. Soc. (2), **19**, 163–169 (1979)
9. Eccles, P.J., Walker, G.: The elements β_1 are representable as framed hypersurfaces. J. Lon. Math. Soc. **22**, 153–160 (1980)
10. Glover, D.J.: A study of certain modular representations. J. Algebra **51**, 425–475 (1978)
11. Kahn, D.S., Priddy, S.B.: Applications of the transfer to stable homotopy theory. Bull. Am. Math. Soc. **78**, 981–987 (1972)
12. Knapp, K.: Some applications of K-theory to framed bordism: e-invariant and transfer. Habilitationsschrift, Univ. Bonn (1979)
13. Mahowald, M.E.: The order of the image of the J-homomorphism. Bull. Am. Math. Soc. **76**, 1310–1313 (1970)
14. McGibbon, C.A.: Stable properties of rank 1 loop structures. Topology **20**, 109–118 (1981)
15. Miller, H.R., Ravenel, D.C., Wilson, W.S.: Periodic phenomena in the Adams-Novikov spectral sequence. Ann. of Math. **106**, 469–516 (1977)
16. Mimura, M., Nishida, G., Toda, H.: Mod$p$ decomposition of compact Lie groups. Pub. RIMS. Kyoto Univ. **13**, 627–680 (1977)
17. Mitchell, S.A., Priddy, S.B.: Stable splittings derived from the Steinberg module. Topology **22**, 285–298 (1983)
18. Ossa, E.: Lie groups as framed manifolds. Topology **21**, 315–324 (1982)
19. Ray, N., Schwartz, L.: Embedding certain complexes via unstable homotopy theory. AMS Contemp. Math. **19**, 331–338 (1983)
20. Rees, E.: Framings on hypersurfaces. J. Lond. Math. Soc. **22**, 161–167 (1980)
21. Smith, L.: On realizing complex bordism modules. Amer. J. Math. **92**, 793–856 (1970)
22. Stolz, S.: Beziehungen zwischen Transfer und J-Homomorphismus. Diplomarbeit, Univ. Bonn (1979)
23. Schwartz, L.: Opérations d'Adams en K-homologie. Bull. Soc. Math. France **109**, 237–257 (1981)
24. Toda, H.: p primary components of homotopy groups IV. Mem. Coll. Sci. Univ. Kyoto, **32**, 297–332 (1959)
25. Whitehead, G.W.: Elements of homotopy theory. Berlin-Heidelberg-New York: Springer 1978

Received April 12, 1984, received in final form August 16, 1984

Math. Z. 189, 263-271 (1985)

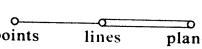
Mathematische Zeitschrift
© Springer-Verlag 1985

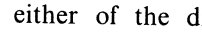
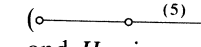
Finite C_3 Geometries in Which All Lines Are Thin

Sarah Rees

Department of Mathematics, University of Illinois, P.O. Box 4348, Chicago, IL 60680, USA

1. Introduction

This paper classifies the finite C_3 geometries (geometries described by the diagram 

) in which all lines are thin (incident with precisely two points). Our interest in such geometries stems from a result of Tits ([11], Theorem 1) which shows that every geometry described by a Coxeter diagram is a building or a quotient of a building precisely if the same is true of all residues described by either of the diagrams C_3 () or H_3 (

). Thus we search for a classification of geometries of types C_3 and H_3; in particular we are interested in finding all geometries of either type which are not quotients of buildings.

By Tits [10], free constructions yield examples of infinite geometries of both types which are unrelated to buildings. Only two finite geometries of type H_3 exist; the geometry of vertices, edges and faces of the icosahedron (a building) and a quotient of this by a group of order two. (The non-existence of any finite generalised pentagon other than the ordinary pentagon ([4]) severely restricts the structure of finite H_3 geometries.) The classification of finite geometries of type C_3, however, is more difficult. For present results see [1, 5-7].

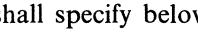
In this paper we shall see that infinitely many finite C_3 geometries with thin lines exist. Given any latin square we can form an associated C_3 geometry with just three points. More complicated geometries of the same type can be formed by joining together these simple geometries. In Theorem A, the main result of this paper, we see that all finite C_3 geometries with thin lines which are not buildings must arise in this way.

We end the paper with a description of another construction which yields C_3 geometries with some thin and some thick lines. The construction seems to be of interest here since these geometries too arise from Latin squares.

2. Concepts and Notation

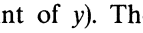
I assume that the reader is familiar with buildings and with the concept of geometries described by Coxeter diagrams. If not, he is referred to [2, 8 or 11]

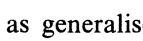
for definitions, but may be consoled by the fact that such knowledge is not really central to the understanding of this paper, but rather provides the motivation for it.

I recall here that by a C_3 *geometry* S I mean a collection of related points, lines and planes (the *varieties* of the geometry) whose structure is described, in a way which we shall specify below, by the diagram 

(C_3). The varieties of S are related by a symmetric and reflexive binary relation I which relates no two distinct points, no two distinct lines and no two distinct planes. The incidence graph of I is required to be connected; any pair of incident varieties is also required to be incident with at least two others.

The C_3 diagram describes the structure of the geometry inductively; appropriate subdiagrams describe the structure of the residue geometries of S. We define the *residue* of a variety V of S, written $Res(V)$, to be the set of varieties of S distinct from V but incident with V, and may consider $\mathrm{Res}(V)$ as a geometry equipped with the restriction of I. The portion 

of the diagram describes the residues of points of S, telling us that the set of lines and planes in the residue of a point of S has the structure of the set of points and lines of a generalised quadrangle (a geometry of points and lines, no two points incident with more than one common line and any point q non-incident with a line y being collinear with a unique point of y). The portion 

of the diagram tells us that every point incident with a line x must be incident with every plane also incident with x. Finally, the portion 

of the diagram describes the residues of planes as generalised projective planes (any pair of points in such a residue is incident with a unique line of the residue, any pair of lines with a unique point).

The buildings of type C_3 are precisely the rank 3 polar spaces (Tits, [8]). A set $\mathscr{P}$ of points and a set $\mathscr{L}$ of lines, defined as subsets of $\mathscr{P}$, are the points and lines of a polar space of rank r precisely if they satisfy the following axioms, due to Buekenhout and Shult ([3]):

(i) No point is collinear with all others.

(ii) If p is a point and x a line then p is collinear with either one or all points of x.

(iii) If $A_1 \subsetneqq A_2 \subsetneqq \ldots \subsetneqq A_s$ is a chain of singular subspaces of $\mathscr{P}$ then $s \leqq r$; moreover there exists such a chain with $s=r$.

(by a *singular subspace* of $\mathscr{P}$ we mean simply a non-empty subset of $\mathscr{P}$ of which any two points p and q are incident with some line all of whose points are in the subset.)

Note that the vertices and edges of a complete r-partite graph with at least 2 vertices in each of the r classes form the points and lines of a rank r polar space with two points on each line. The singular subspaces of such a geometry are the complete subgraphs of the graph. All polar spaces with two points on each line arise in precisely this way.

We define the *collinearity graph* of a C_3 geometry S to be the graph whose vertex set is the set of points of S and edge set is the set of pairs of collinear

points of S. We define the *radical* of S, written $Rad(S)$, to be the set of points of S collinear with all others. Finally, we say that a line of S is *thin* if it is incident with exactly two points of S and *thick* if it is incident with three or more points of S.

We also require in this paper the following terminology.

Where n is any positive integer, by an $n \times n$ *grid* I mean the direct product of two copies of the set $\{1,2,\ldots,n\}$, that is the set of all ordered pairs (i,j) where $i,j \in \{1,2,\ldots,n\}$. The subset of all pairs (i,k), for fixed i and $k=1,2,\ldots,n$, is called the *i-th row* of the grid, the subset of all pairs (k,j), for fixed j and $k=1,2,\ldots,n$, the *j-th column.*

A *latin square* on an $n \times n$ grid is a partition of the set into n sets each of size n, called *colours* and labelled by the set $\{1,2,\ldots,n\}$, such that no two elements of the grid in the same row or in the same column have the same colour.

3. Statement of Results

Before stating the main result of this paper, I give a couple of examples of C_3 geometries with thin lines.

Example 1. Define a geometry S as follows:

S contains 3 points, a, b and c, $3n$ lines, each incident with 2 of a, b and c, *and* n^2 planes, each incident with all three of a, b and c and with 3 lines, where n is some integer greater than 1. S is defined with respect to a latin square L on an $n \times n$ grid.

The lines of S incident with a and b can be labelled $R1, R2, \ldots, Rn$. the lines incident with b and c $C1, C2, \ldots, Cn$, and the lines incident with c and a $L1, L2. \ldots, Ln$. The planes can be labelled by the elements (i,j) of the $n \times n$ grid. The plane (i,j) is defined to be incident with the 3 points a, b and c and with the lines Ri, Cj and Lk, where the element (i,j) of the grid has colour k in the latin square L.

S is a geometry of type C_3.

A more general example can be constructed by joining together geometries of the type described in Example 1 in a particular way.

Example 2. Define a geometry S as follows:

S contains $r+s+t$ points, $(rs+st+rt)n$ lines and $rstn^2$ planes, for some positive integers r,s,t,n, where if $n=1$ we require each of r, s and t to be greater than 1.

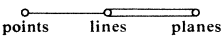
The points of S fall into three classes, X, Y and Z of sizes r, s and t respectively. Each triple (x,y,z) where x is a point in X, y a point in Y and z a point in Z, labels a latin square $L_{(x,y,z)}$ on an $n \times n$ grid.

For any $x \in X$, $y \in Y$ and $z \in Z$, we define n lines $R_{x,y}1$, $R_{x,y}2, \ldots, R_{x,y}n$ incident with x and y, n lines $C_{y,z}1$, $C_{y,z}2, \ldots, C_{y,z}n$ incident with y and z and n lines $L_{z,x}1$, $L_{z,x}2, \ldots, L_{z,x}n$ incident with z and with x. The lines $R_{x,y}i$, $C_{y,z}j$ and $L_{z,x}k$ are the lines of a plane incident also with the three points x, y and z precisely if the (i,j)-th entry of the latin square $L_{(x,y,z)}$ has colour k.

For any choice of $n \times n$ latin squares $L_{(x,y,z)}$, S is a geometry of type C_3.

Note that if $n=1$, the above description reduces to that of a rank 3 polar space (for remember that in this case we have ensured that each of X, Y and Z contains at least two points). If $r=s=t=1$, we have the geometry of Example 1.

In fact all C_3 geometries with thin lines arise as described above. For we have

Theorem A. *Suppose that S is a finite geometry described by* 

in which all lines are thin. Then

(i) *The collinearity graph of S is a complete tripartite graph, $K_{r,s,t}$.*

(ii) *Any two collinear points of S are incident with the same number n of lines.*

(iii) *If $n=1$, then S is a rank 3 polar space, and each of r, s and t is greater than 1*

(iv) *If $n>1$, then any set of three pairwise collinear points of S together with all lines incident with two of the points and all planes incident with all three forms a C_3 geometry as in Example 1. S itself arises as in Example 2.*

4. Proof of the Theorem

From now on, unless otherwise stated, we assume that S is a finite C_3 geometry in which all lines are thin.

We give first a series of elementary results.

Lemma 1. *The residue of a plane in S is a projective plane on three points. The residue of a point is a generalised quadrangle whose points and lines are the vertices and edges of a complete bipartite graph.*

Proof. Immediate.

Corollary 2. *Suppose that p is a point of S. Then the lines of* Res(p) *fall into two classes, two lines being coplanar if they are in opposite classes. Hence if A, B are two disjoint subsets of lines in* Res(p) *such that no two lines in A are coplanar and no two lines in B are coplanar, and if*

either (i) *$A \cup B$ is the set of all lines in* Res(p)

or (ii) *some line in A is coplanar with some line in B,*

then every line in A is coplanar with every line in B.

Proof. Immediate from Lemma 1.

Lemma 3. *Any pair of collinear points of S is incident with the same number n of lines. If some point of S is incident with all planes of S then $n>1$.*

Proof. Let p, q and r be three points incident with a plane of S. Let n_{pq} be the number of lines incident with the pair of points p, q; similarly define n_{pr}, n_{qr}. p, q and r, being coplanar, are pairwise collinear; thus n_{pq}, n_{pr} and n_{qr} are all non-zero.

Two lines incident with both p and q cannot be coplanar; neither can two lines incident with both p and r. However, since p, q and r are coplanar, some line incident with both p and q is coplanar with some line incident with both p and r. Thus, by Corollary 2, we see that every line incident with both p and q is coplanar with every line incident with both p and r; any two such lines are incident with a plane that is incident with each of p, q and r. The axioms of the generalised quadrangle Res(p) clearly forbid two lines of Res(q) to be incident with more than one plane of S. Thus we see that precisely $n_{pq}n_{pr}$ planes of S are incident with the three points p, q and r.

Counting differently, we see that $n_{pq}n_{pr}=n_{pq}n_{qr}=n_{pr}n_{qr}$. Thus $n_{pq}=n_{pr}=n_{qr}$. Now, using the connectivity of S, we see that any pair of collinear points of S is incident with the same number, n, of lines.

Finally, suppose that p is a point of S incident with all planes, and let q, r be two other points of S, incident with a line x. Since, as above, every line through p and q is coplanar with x, we see that precisely $n=n_{pq}$ planes are incident with both x and p. Thus, by our choice of p, x itself is incident with exactly n planes. Hence n must be greater than 1.

Lemma 4. *Suppose that P is a plane of S and p a point of S non-incident with P. Then p cannot be collinear with every point incident with P.*

Proof. Suppose the contrary. Let q, r, s be the points of Res(P) and x, y, z the lines, with xIq, r, yIr, s, zIs, q.

x and z, being coplanar, are in distinct classes of lines in Res(q). No two lines incident with both p and q can be coplanar; thus all such lines lie in the same class of lines of Res(q). Hence either every line incident with p and q is coplanar with x and not with z or vice versa. We shall assume the former.

Now we have a plane incident with p and x, and thus a line incident with p and r and coplanar with x. Hence all lines incident with p and r are coplanar with x and not with y.

But finally, there exists a line incident with p and s, coplanar with either y or z. In the first case, since we have a plane incident with p and y, we can find a line incident with p and r and coplanar with y; in the second case, we find a line incident with p and q and coplanar with z. Thus we have a contradiction.

Corollary 5. *If p is a point in the radical of S then p is incident with every plane of S. Hence* $|\mathrm{Rad}(S)| \leqq 3$.

Proof. Immediate from Lemmas 4 and 1.

Corollary 6. *Any three pairwise collinear points of S are coplanar.*

Proof. Suppose that p, q, r are three pairwise collinear points of S, and suppose that no plane is incident with all three points. Then if x is a line incident with q and r some plane P must be incident with x, thus also with q, r and some third point s.

Now if y is a line incident with p and q in Res(q) we see, from the fact that Res(q) is a generalised quadrangle, that y must be coplanar with a line incident with both q and P. Since p, q and r are not coplanar, y must therefore be coplanar with the line of Res(P) incident with q and s.

Thus p is collinear with each of q, r and s, a situation which is forbidden by Lemma 4.

Lemma 7. *Suppose that any two collinear points of S are incident with more than one common line. Then suppose that $\mathscr{P}$ is a set of three pairwise collinear points of S, $\mathscr{L}$ is the set of all lines of S incident with two points of $\mathscr{P}$ and π is the set of all planes of S incident with all three points of $\mathscr{P}$. Let S' be the union of $\mathscr{P}$, $\mathscr{L}$ and π, and define incidence in S' as in S.*

Then S' is a C_3 geometry with thin lines and radical of size 3.

Proof. First note that, by Corollary 6, π is non-empty. Note also that the residue in S' of a plane in π is certainly a projective plane, and that every point in S' incident with a given line is incident with every plane of S' incident with that line.

Now if p, q and r are the points of $\mathscr{P}$, no two lines incident with p and q are coplanar, no two lines incident with p and r are coplanar, but some line incident with p and q is coplanar with some line incident with p and r. Thus, by Corollary 2, we see that every line incident with p and q is coplanar with every line incident with p and r, in a plane incident with p, q and r. Thus the residue in S' of p is a generalised quadrangle.

The connectedness of S' and the fact that any pair of incident varieties of S' is incident with at least two other varieties is now clear. Thus we see that S' is a C_3 geometry with thin lines and radical $\mathscr{P}$ of size 3.

Lemma 8. *Suppose that S is any C_3 geometry. Then, if p is a point and x a line of S non-incident with p, p is collinear with at least one point of* Res(x).

Proof. Certainly it is sufficient to show that any chain $pIzIqIyIrIx$ where p, q and r are points and x, y and z are lines, can be replaced by a chain $pIwIsIx$, where s is a point and w a line.

In verifying this I use a shorthand notation shown to me by AE Brouwer. I call points varieties of type 1, lines varieties of type 2 and planes varieties of type 3, for the sake of brevity. For fixed varieties A of type i and B of type j, the sequence $i, i_1, i_2, \ldots, i_r, j$ means "there exists a chain $AIv_1Iv_2I\ldots Iv_rIB$, with each v_k of type i_k". We say that the sequence $i, i_1, i_2, \ldots, i_r, j$ *implies* the sequence $i, j_1, j_2, \ldots, j_s, j$ if whenever, for A of type i and B of type j, there exists a chain $AIv_1Iv_2I\ldots v_rIB$ with each v_k of type i_k there exists also a chain $AIw_1Iw_2I\ldots Iw_sIB$ with each w_k of type j_k. Clearly we hope to show that 121212 implies 1212.

We make use of the following:

(i) that, since the residue of a point of S is a generalised quadrangle, the sequences 213 and 312 imply 2323 and 3232 respectively

(ii) that, since in the residue of a line every point and every plane are incident, 121 implies 131, 123 implies 13 and 321 implies 31

(iii) that, since the residue of a plane is a projective plane, 232 implies 212 and 131 implies 121.

Thus, in the following list of sequences, each sequence implies the next: 121212, 121312, 1232312, 132312, 1323232, 1321232, 13132, 12132, 123232, 13232, 13212, 1312, 1212.

Corollary 9. *Let S be a finite C_3 geometry with thin lines and empty radical. Then the set of points and pairs of collinear points of S, with incidence defined by inclusion, is the geometry of points and lines of a rank 3 polar space.*

Proof. By Lemma 8, we can verify the axioms of Buekenhout and Shult ([3]) to show that we do indeed have a polar space. Lemma 4 tells us that the points of any plane of S form a maximal singular subspace; hence our polar space has rank 3.

We are now ready to prove our theorem.

Proof of Theorem A. Suppose that S is a finite C_3 geometry with thin lines. We know by Corollary 5 that the radical of S has size at most three and that every point of the radical is incident with all planes. By Lemma 3, any pair of collinear points of S is incident with n lines, where n is some positive integer.

If $n=1$, then the radical of S is empty (by Lemma 3) and the points and lines of S can be identified with the vertices and edges of its collinearity graph, the planes (by Corollary 6) with the triangles in the graph. By Corollary 9, S is a rank 3 polar space, and so a building, by [8].

So now suppose that $n>1$, and consider the four cases $|\mathrm{Rad}(S)|=3$, $|\mathrm{Rad}(S)|=2$, $|\mathrm{Rad}(S)|=1$ and $\mathrm{Rad}(S)=\varnothing$.

Clearly if $\mathrm{Rad}(S)$ has size three the collinearity graph of S is simply a triangle.

If $\mathrm{Rad}(S)$ has size two, let a, b be the two points of the radical and X the set of remaining points. Since all planes of S are incident with a, b and exactly one point of X, all lines of S must be incident with at least one of a and b and at most one point of X. Thus the collinearity graph of S is tripartite, with parts $\{a\}, \{b\}$ and X. Since X is not in the radical of S, X must contain at least two points.

If $\mathrm{Rad}(S)$ has size 1, let a be the point in its radical. By Corollary 2 the lines of $\mathrm{Res}(a)$ fall into two classes, two such lines being coplanar precisely if they are in opposite classes. If b is a point distinct from a then the set of lines incident with a and b is contained within one class of lines of $\mathrm{Res}(a)$. Thus we can also divide the points of S distinct from a into two classes X and Y, any plane of S being incident with a, a point of X and a point of Y, and every point of X being collinear with every point of Y. Since no plane of S is incident with more than one point of either X or Y, no two points of X are collinear and no two points of Y are collinear. Thus the collinearity graph of S is tripartite, with parts $\{a\}, X$ and Y. Since X and Y are not contained in the radical of S they must each have size greater than 1.

If $\mathrm{Rad}(S)$ is empty we use Corollary 9 to show that the collinearity graph of S is a rank 3 polar space, and hence a complete tripartite graph with at least two vertices in each class.

It remains to prove that statement (iv) of the theorem holds when $n>1$.

Choose a, b, c, three pairwise collinear points of S. By Lemma 7, these three points, together with all lines incident with two of them and all planes incident with all three form a C_3 geometry S' with thin lines; we need to verify that S' must arise as in Example 1. We know already that any two points of S' are incident with n lines. Label the lines incident with a and b $R1, R2, \ldots, Rn$, the

lines incident with b and c $C1, C2, \ldots, Cn$ and the lines incident with c and a $L1, L2, \ldots, Ln$. By Corollary 2, each pair of lines Ri, Cj is incident with a plane of S', this incident with a third line Lk. The axioms of the generalised quadrangles which are the residues of points of S' forbid two lines of S' to be incident with more than one common plane; thus we see that each pair Ri, Cj specifies a unique line Lk, and that no other pair Ri, Cj' or Ri', Cj specifies the same line. Thus we see that for any fixed Lk, the pairs (i,j) such that Ri and Cj specify Lk correspond to the entries of one particular colour of a latin square on an $n \times n$ grid.

Now that we know the shape of the collinearity graph and the local structure of S it becomes clear that S must arise as in Example 2. It may not be immediately obvious that, once the integer n is specified, the set of $n \times n$ latin squares which define our geometry may be chosen completely arbitrarily and independently. However, a moment's thought shows that the actual structure of each latin square does not in any way affect the structure of the residues of the geometry; all that is important is that each pair of coplanar lines in the residue of a point is incident with a unique and distinct plane. Any choice of latin squares ensures this.

5. Comments on the Results

We see from Theorem A that there are an awful lot of C_3 geometries with thin lines. The following points are perhaps worthy of comment.

(i) There exist a lot of C_3 geometries with thin lines which are not quotients of buildings. For example, if S is a C_3 geometry with thin lines and radical of size three, associated with a latin square L on an $n \times n$ grid, then we can show that S is a quotient of a building precisely if the set of permutations $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that the colours of the i-th row of the grid associated with L are a permutation under α_i of the colours of the first row forms a subgroup of the symmetric group S_n.

(ii) The existence of totally thin C_3 subgeometries in these geometries (geometries in which every line is incident with two points and with two planes, every incident point, plane pair with two lines – such subgeometries might play the equivalent role in S to *apartments* in a building, see [11]) is completely unrelated to whether or not the geometries are quotients of buildings. For example, if S is a C_3 geometry with thin lines and radical of size three, defined with respect to a latin square L on an $n \times n$ grid, then S contains completely thin C_3 subgeometries precisely if we can find a rectangle of entries in the $n \times n$ grid associated with L such that each pair of entries at the opposite ends of either diagonal of the rectangle has the same colour.

(iii) Recall that the structures induced on triples of pairwise collinear points in such a geometry S are completely independent of each other.

6. C_3 Geometries with Some Thin and Some Thick Lines

Given an $n \times n$ latin square L with $n \geqq 3$ we can also define with respect to L a C_3 geometry with some thin and some thick lines (incident each with n points) in which every point and every plane are incident.

We define such a geometry S as follows:

Let the points of S be $0, 1, 2, 3, \ldots, n$. In S we have n^2+2 lines, n^2 thin lines indexed by the elements of an $n \times n$ grid and 2 thick lines called R and C. We have $2n$ planes, labelled $R1, R2, \ldots, Rn$ and $C1, C2, \ldots, Cn$.

The thin line (i,j) is incident with the points 0 and k, where the element (i,j) has colour k in the latin square L. The thick lines R and C are each incident with the points $1, 2, \ldots, n$.

The plane Ri is incident with all the points of S and with the lines R and all (i,k) where $k=1,2,\ldots,n$; the plane Cj is incident with all the points of S and with the lines C and all (k,j) where $k=1,2,\ldots,n$.

As in the case of geometries with all lines thin we see that S is a quotient of a building precisely if the set of permutations $\alpha_1, \alpha_2, \ldots, \alpha_n$ which we can associate with the rows of the latin square form a group. Similarly, S contains totally thin C_3 subgeometries precisely if we can find a rectangle of entries in the $n \times n$ grid associated with L such that the entries at opposite ends of each diagonal of the rectangle have the same colour in the latin square.

Observing that the partition of the $n \times n$ grid into the colours of a latin square is essentially a partition of the points of the grid generalised quadrangle into ovoids, whereas the partition of the lines of the grid into rows and columns is essentially a partition of the lines of the grid generalised quadrangle into spreads, we see easily that a similar construction would give us geometries from any generalised quadrangle which could be partitioned by both ovoids and spreads. In fact, although we shall not prove this here, it is possible to show that all C_3 geometries with some thin and some thick lines in which every point and every plane are incident arise in this way.

References

1. Aschbacher, M.: Finite geometries of type C_3 with flag transitive group. To appear in Geom. Dedicata
2. Buekenhout, F.: Diagrams for geometries and groups. J. Comb. Theory, Ser. A**27**, 121–151 (1979)
3. Buekenhout, F., Shult, E.E.: On the foundations of polar geometry. Geom. Dedicata **3**, 155–170 (1974)
4. Feit, W., Higman, G.: The non-existence of certain generalised polygons. Journal of Algebra **1**, 114–131 (1964)
5. Ott, U.: On finite geometries of type B_3. To appear in J. Comb. Theory
6. Rees, S.: C_3 geometries arising from the Klein quadric. To appear in Geom. Dedicata
7. Rees, S.: A classification of a class of C_3 geometries. Preprint
8. Tits, J.: Buildings of spherical type and finite *BN*-pairs, Lecture Notes in Mathematics no. **386**. Berlin, Heidelberg, New York: Springer 1974
9. Tits, J.: Buildings and Buekenhout geometries, in Finite Simple Groups, ed. M. Collins, 309–320. New York: Academic Press 1981
10. Tits, J.: A local characterisation of buildings. Preprint and earlier version of [11]
11. Tits, J.: A local approach to buildings, in The Geometric Vein (The Coxeter Festschrift), ed. Davis, Grünbaum, Sherk, 519–547. Berlin, Heidelberg, New York: Springer 1981

Received May 15, 1984

Math. Z. 189, 273-288 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Über gewisse Dirichlet-Reihen mit Eulerscher Produktzerlegung

Hans Petersson*

R.A. Rankin zur Vollendung seines siebzigsten Lebensjahres am 27. Oktober 1985 gewidmet

Beziehungen zwischen Dirichlet-Reihen und Primzahlen drücken sich dadurch aus, daß die betreffende Dirichlet-Reihe eine Eulersche Produktzerlegung gestattet. Im folgenden sollen drei, wie es scheint, noch nicht bekannte Beispiele konkreter Dirichlet-Reihen aufgewiesen werden, für die solche Zerlegungen bestehen. Eine von ihnen hängt mit einer berühmten Untersuchung von Gauss [1] zusammen; die beiden anderen sind, wenn auch merklich komplizierter, zu der erstgenannten analog. Alle entsprechen durch Übertragung der Koeffizienten gewissen binären Thetareihen der Stufe 2, die aus der Systematik der einfachen Thetareihen der Thetagruppe Γ_ϑ hervorgehen.

Die Gestalt der Faktoren in der Eulerschen Produktzerlegung stimmt mit derjenigen überein, welche als ein Endresultat der Heckeschen Operatorentheorie vermöge der Metrik (durch Betrachtung der Eigenformen aller Hecke-Operatoren) zustande kommt (vgl. [3, 4]). Im Vergleich zu dieser relativ komplizierten Beweistechnik sind die Deduktionen in den vorliegenden Beispielsfällen überwiegend elementar.

Als Ergänzung ergibt sich aufgrund bekannter Sätze für jede der betrachteten Dirichlet-Reihen die holomorphe Fortsetzbarkeit in die volle Ebene sowie das Bestehen einer Riemannschen Funktionalgleichung. Im Falle des dritten Beispiels macht dies besondere Überlegungen erforderlich, die anhangsweise in §4 dargestellt werden.

Zur Terminologie und zu den theoretischen Grundlagen der vorliegenden Untersuchung sei, soweit Modulfunktionen in Betracht kommen, bemerkt, daß sich diese in dem Ergebnisbericht [5] des Literaturverzeichnisses, der im folgenden mit EB zitiert wird, vollständig zusammengestellt vorfinden. §1 enthält davon das Wichtigste über die Modulgruppe und über Modulformen, §2 die Einführung der einfachen Thetareihen der Thetagruppe und der zweite Teil von Anhang E die Systematik ihrer Relationen. Ein Symbolverzeichnis und ein Sachverzeichnis erleichtern die Orientierung über die zahlreichen Zitate.

* Vor Drucklegung dieser Arbeit, am 9.11.84 verstorben

§ 1. Zur Entwicklung des ersten Beispiels gehen wir aus von der Modulform (vgl. EB (1.35), Satz 2.3, Satz 2.4, (2.20))

$$\varphi(\tau)=\varphi_1(\tau)=\eta^3(\tau)\vartheta_3(\tau)=\vartheta_1^{(1)}(\tau)\vartheta_3(\tau)=\sum_{\substack{n=1\\ n\equiv 1(2)}}^{\infty}\beta(n)e^{\pi i n\frac{\tau}{4}}, \tag{1.1}$$

wo in der Schreibweise von EB (2.35)

$$\beta(n)=\sum_{m_1,m_2} m_1 \quad udB \quad m_1\equiv 1 \bmod 4, \qquad m_2\equiv 0 \bmod 2, \qquad m_1^2+m_2^2=n. \tag{1.2}$$

Ersichtlich ist

$$\beta(1)=1, \qquad \beta(n)=0 \quad \text{für } n\equiv 3 \bmod 4, \qquad \beta(p)=2m_1=2m_1(p)$$

für eine Primzahl $p\equiv 1 \bmod 4$, wenn $p=m_1^2+m_2^2$, m_1, $m_2\in\mathbb{Z}$ und $m_1\equiv 1 \bmod 4$. Eine andere Darstellung für $\beta(n)$ ergibt sich wie folgt: Es sei $\mathfrak{K}:=\mathbb{Q}(i)$ der Gaußsche Zahlkörper, $\mathfrak{o}$ die Hauptordnung in $\mathfrak{K}$ und $\mathfrak{o}_4:=\{\mu\in\mathfrak{o}\,|\,\mu=m_1+m_2 i;\ m_1,m_2\in\mathbb{Z},\ m_1\equiv 1 \bmod 4,\ m_2\equiv 0 \bmod 2\}$; $\mathfrak{o}_4$ enthält mit je zwei Zahlen aus $\mathfrak{o}$ deren Produkt. Es gilt dann mit ν' als algebraischer Konjugierter von ν in $\mathfrak{K}$

$$\beta(n)=\sum_{\mu\in\mathfrak{o}_4,\ \mu\mu'=n}\mu \qquad (n\in\mathbb{N},\ n\equiv 1 \bmod 2). \tag{1.3}$$

Zusätzlich sei bemerkt, daß die Gliederzahl dieser Summe durch

$$\alpha(n)=\sum_{\mu\in\mathfrak{o}_4,\ \mu\mu'=n}1=\sum_{d>0,\,d|n}\left(\frac{-1}{d}\right) \qquad (n\in\mathbb{N},\ n\equiv 1 \bmod 2) \tag{1.4}$$

gegeben und daher distributiv ist.

Im folgenden soll zunächst bewiesen werden, daß die der Modulform (1.1) $\varphi(\tau)$ zugeordnete Dirichlet-Reihe

$$D(s)=D(s,\varphi):=\sum_{\substack{n=1\\ n\equiv 1(4)}}^{\infty}\beta(n)n^{-s}, \tag{1.5}$$

die nach (1.2) für $\mathrm{Re}\,s>3/2$ absolut konvergiert, eine Eulersche Produktzerlegung gestattet. Dazu ist zu zeigen, daß gilt

$$\beta(l)\beta(m)=\beta(lm) \qquad \text{für } l,m\in\mathbb{N},\ lm\equiv 1 \bmod 2,\ (l,m)=1. \tag{1.6}$$

Nach (1.4) verschwinden beide Seiten von (1.6), wenn nicht $l\equiv m\equiv 1 \bmod 4$. Dies treffe zu; man erhält dann

$$\beta(l)\beta(m)=\sum_{\substack{\lambda,\,\mu\in\mathfrak{o}_4\\ \lambda\lambda'=l,\ \mu\mu'=m}}\lambda\mu. \tag{1.7}$$

Nach (1.4) hat diese Summe ebenso viele Glieder wie die Darstellung (1.3) für $n=lm$, und alle diese erfüllen die Summationsbedingung von (1.3). Die Relation (1.6) ist also bewiesen, wenn gezeigt wird, daß in (1.7) kein Glied doppelt auf-

tritt. Ist nun $\lambda_1 \mu_1$ irgendein Glied der Summe (1.7) mit dem gleichen Wert $\lambda\mu$ wie in (1.7), so hat man

$$\lambda\mu = \lambda_1 \mu_1 \quad (\lambda, \lambda_1, \mu, \mu_1 \in \mathfrak{o}_4, \lambda\lambda' = \lambda_1 \lambda_1' = l, \mu\mu' = \mu_1 \mu_1' = m).$$

Wegen $(l, m) = 1$ ist λ zu μ_1 teilerfremd und geht daher in λ_1 auf; aus Symmetriegründen sind λ und λ_1 assoziiert, stimmen also, da in $\mathfrak{o}_4$ enthalten, überein, was die Behauptung liefert. Diese hat für $\operatorname{Re} s > 3/2$ die Gestalt

$$D(s) = \prod_{p>2} E_p(p^{-s}), \quad \text{wo} \quad E_p(w) = \sum_{k=0}^{\infty} \beta(p^k) w^k$$

und p die ungeraden Primzahlen durchläuft.

Die Bestimmung von $E_p(w)$ im Falle $p \equiv 3 \bmod 4$ ergibt sich unmittelbar aus (1.2,4). Danach verschwindet $\beta(p^k)$ für ungerade k, während für gerade $k = 2h$ $(h \in \mathbb{N})$ gilt

$$p^k = p^{2h} = ((-p)^h)^2 + O^2, \quad \beta(p^k) = (-p)^h,$$

woraus $E_p(w) = (1 + pw^2)^{-1}$ folgt $(p \equiv 3 \bmod 4)$.

Im Falle $p \equiv 1 \bmod 4$ betrachten wir bei gegebenen $a, b \in \mathbb{C}$ die für hinreichend kleine w holomorphe Funktion von w:

$$(1 - aw - bw^2)^{-1} = \sum_{k=0}^{\infty} g_k w^k. \tag{1.8}$$

Hier gilt einerseits

$$g_0 = 1, \quad g_1 = a, \quad g_2 = a^2 + b, \ldots, g_k = a g_{k-1} + b g_{k-2}$$

$(k \geqq 2)$, andrerseits, wenn

$$\left.\begin{matrix} \rho \\ \rho^* \end{matrix}\right\} = \frac{a}{2} \pm \sqrt{\frac{a^2}{4} + b} \tag{1.9}$$

gesetzt wird:

$$(1 - aw - bw^2)^{-1} = (1 - \rho w)^{-1} (1 - \rho^* w)^{-1}$$

und daher

$$g_k = \sum_{j=0}^{k} \rho^j \rho^{*k-j} \quad (k \in \mathbb{N}). \tag{1.10}$$

Zunächst sollen a und b so bestimmt werden, daß $g_1 = \beta(p)$, $g_2 = \beta(p^2)$ zutrifft; dabei wird entsprechend (1.2) gesetzt

$$p = m_1^2 + m_2^2 \quad (m_1, m_2 \in \mathbb{Z}, m_1 \equiv 1 \bmod 4, m_2 > 0). \tag{1.11}$$

Daraus folgt erstens $\beta(p) = a = 2m_1$.

Nach (1.4) gibt es genau drei Darstellungen von p^2, die (1.2) entsprechen. Eine von ihnen erhält man mit $m_1 = p$, $m_2 = 0$, die anderen in der Gestalt

$$p^2 = (m_1^2 - m_2^2)^2 + (2\varepsilon m_1 m_2)^2 \quad (\varepsilon = \pm 1),$$

so daß nun folgt

$$\beta(p^2) = 2(m_1^2 - m_2^2) + p = 4m_1^2 - p = a^2 - p;$$

$\beta(p^2)=g_2$ besagt zweitens $b=-p$. Mit diesen Werten a, b ergibt sich für die Zahlen (1.9) ρ, ρ^* bei geeigneter Vorzeichenwahl

$$\rho=m_1+im_2\in\mathfrak{o}_4, \qquad \rho^*=m_1-im_2=\rho'\in\mathfrak{o}_4, \qquad p=\rho\rho'.$$

Die letzte Gleichung bedeutet die Zerlegung des Hauptideals (p) in ein Produkt verschiedener Primideale, die durch Zahlen aus $\mathfrak{o}_4$ dargestellt werden.

Nach (1.3) hat man $\beta(p^k)=\Sigma\nu$ unter den Bedingungen

$$\nu\in\mathfrak{o}_4, \qquad \nu\nu'=p^k=\rho^k\rho'^k \qquad (k\in\mathbb{N});$$

ν läßt sich daher in der Gestalt $\nu=\rho^j\rho'^{j'}$ mit $j,j'\in\mathbb{N}_0$ schreiben, und hier gilt $j+j'=k$, also $\nu=\rho^j\rho'^{k-j}$; umgekehrt genügen diese ν den obigen Bedingungen. (1.10) ergibt demnach

$$g_k=\beta(p^k)\ (k\in\mathbb{N}), \qquad \text{d.h. } E_p(w)=(1-2m_1w+pw^2)^{-1},$$

wenn $p\equiv 1 \bmod 4$, und man erhält als erstes Resultat

Satz 1. *Die Dirichlet-Reihe* (1.5) $D(s)$, *die der Modulform* (1.1) $\varphi(\tau)=\vartheta_1^{(1)}(\tau)\,\vartheta_3(\tau)$ *durch die Koeffizientenformel* (1.2) *entspricht, konvergiert für* $\operatorname{Re} s>3/2$ *absolut und gestattet die Eulersche Produktzerlegung*

$$D(s)=\prod_{p>2} E_p(p^{-s}) \qquad \textit{mit } E_p(w)=(1-2m_1w+pw^2)^{-1}.$$

Hier ist $p=m_1^2+m_2^2$ mit $m_1, m_2\in\mathbb{Z}$, $m_1\equiv 1 \bmod 4$, wenn $p\equiv 1 \bmod 4$, dagegen $m_1=0$ für $p\equiv 3 \bmod 4$.

Eine andere Formulierung des gleichen Sachverhalts besagt das

Korollar 1. *Es sei* $\mathfrak{P}_4$ *die Menge der ungeraden Primzahlen* ω *des Gaußschen Zahlkörpers* $\mathbb{Q}(i)$, *für die* $\operatorname{Re}\omega\equiv 1 \bmod 4$ *zutrifft. Dann gilt mit* $\mathbf{N}$ *als Norm in* $\mathbb{Q}(i)$

$$D(s)=\sum_{\substack{n=1\\ n\equiv 1\,(4)}} \beta(n)\,n^{-s}=\prod_{\omega\in\mathfrak{P}_4}\left(1-\frac{\omega}{(\mathbf{N}\omega)^s}\right)^{-1}.$$

Auf eine Beziehung zu Heckeschen Größencharakteren, die sich hier andeutet, sei hingewiesen.

Das Problem des analytischen Verhaltens der Funktion $D(s)$ läßt sich durch Anwendung von [6] §2, Satz 4 unmittelbar klären. Zur Erläuterung soll die in [6] §2 mit $D(s; f, \Gamma)$ bezeichnete Funktion hier $F(s; f, \Gamma)$ genannt werden. Die Entwicklung

$$\varphi(\tau)=\sum_{\substack{n=1\\ n\equiv 1\,(4)}}^{\infty} \beta(n)\,e^{\pi i n\frac{\tau}{4}}=\sum_{k=0}^{\infty}\beta(4k+1)\,e^{\pi i(k+\frac{1}{4})\tau}$$

zeigt, daß der Fall $f=\varphi$, $\Gamma=\Gamma_\vartheta$, $N=2$, $\kappa=1/4$ vorliegt, so daß gilt

$$F(s;\varphi,\Gamma_\vartheta)=4^s D(s);$$

und $\varphi|T=-\varphi$ ergibt nun nach [6] §2, Satz 4 folgendes Resultat:

Korollar 1'. *Sowohl $D(s)$ als auch die Funktion*

$$\Xi(s) := \left(\frac{\pi}{4}\right)^{-s} \Gamma(s)\, D(s)$$

ist in die volle komplexe s-Ebene holomorph fortsetzbar, und $D(s)$ genügt dort der Funktionalgleichung $\Xi(2-s)=\Xi(s)$.

(Berichtigung.) Der Text in [6] § 2 unmittelbar nach Satz 4 bedarf einer Korrektur: S. 19, Z. 8 v.u.: $r \leftrightarrow r-s$ ist abzuändern in: $s \leftrightarrow r-s$.

§ 2. Das zweite Beispiel einer Dirichlet-Reihe mit Eulerscher Produktzerlegung entsteht aus der Modulform

$$\varphi(\tau)=\varphi_3(\tau) := \eta^4(\tau)=\vartheta_1^{(1)}(\tau)\,\eta(\tau). \tag{2.1}$$

Diese kann nach einer analytischen Identität (s. EB S. 272) in der Gestalt

$$\varphi(\tau)=\vartheta_3^{(1)}(\tau)\,\vartheta_3(\tau)=\sum_{\substack{n=1\\ n\equiv 1\,(6)}}^{\infty} \beta(n)\, e^{\pi i n \frac{\tau}{3}} \tag{2.2}$$

mit

$$\vartheta_3^{(1)}(\tau) := \sum_{\substack{m=-\infty\\ m\equiv 1\,(3)}}^{+\infty} m e^{\pi i m^2 \frac{\tau}{3}}$$

dargestellt werden, so daß gilt

$$\beta(n)=\sum_{m_1,\, m_2} m_1 \quad udB \quad m_1 \equiv 1 \bmod 3, \quad m_1^2+3m_2^2=n, \tag{2.3}$$

wenn $n\in\mathbb{N}$, $n\equiv 1 \bmod 6$. Das Thetaprodukt (2.2) liefert zunächst die Summationsbedingung $n\equiv 1 \bmod 3$; die Exponenten von $e^{2\pi i\tau}$ in der Fourier-Entwicklung von $\eta^4(\tau)$ haben aber die Form $k+\frac{1}{6}$ mit $k\in\mathbb{N}_0$, was auf $n=6k+1\equiv 1 \bmod 6$ führt.

Wir benutzen die Formel (2.3) auch für $n\equiv -1 \bmod 6$. Da

$$\alpha(n) := \sum_{m_1,\, m_2} 1 \quad udB \quad m_1 \equiv 1 \bmod 3, \quad m_1^2+3m_2^2=n \tag{2.4}$$

für $n\in\mathbb{N}$, $n\equiv \pm 1 \bmod 6$ nach EB Satz 6.2 durch

$$\alpha(n)=\sum_{d>0,\, d\mid n} \left(\frac{d}{3}\right) \tag{2.5}$$

wiedergegeben wird, gilt $\alpha(n)=\beta(n)=0$ für $n\equiv -1 \bmod 6$.

Es sei $\mathfrak{K} := \mathbb{Q}(\sqrt{-3})$, $\mathfrak{o}$ die Hauptordnung in $\mathfrak{K}$,

$$\mathfrak{o}_1 := \{m_1+m_2\sqrt{-3} \mid m_1, m_2\in\mathbb{Z}\},$$

$$\mathfrak{o}_3 := \{m_1+m_2\sqrt{-3} \mid m_1, m_2\in\mathbb{Z}, m_1\equiv 1 \bmod 3\}$$

und $\xi_3 := \exp\frac{\pi i}{3}$ (vgl. EB (1.10)). Die Zahlen $1, \xi_3$ bilden eine Ganzheitsbasis von $\mathfrak{K}$. Eine Zahl $\mu = n_1 + n_2\xi_3$ $(n_1, n_2 \in \mathbb{Z})$ liegt genau dann in $\mathfrak{o}_1$, wenn $n_2 \equiv 0 \bmod 2$. Die Norm n von μ ist genau dann gerade, wenn n_1 und n_2 beide gerade sind. Im anderen Falle liegt von den Zahlen $\mu, \mu\xi_3, \mu\xi_3^2$ genau eine in $\mathfrak{o}_1$; wenn überdies $n \equiv 1 \bmod 3$, so liegt von den Zahlen $\mu\xi_3^j$ $(j = 0, \ldots, 5)$ genau eine in $\mathfrak{o}_3$. Das Produkt zweier Zahlen aus $\mathfrak{o}_3$ ist in $\mathfrak{o}_3$ enthalten. Für $\mu \in \mathfrak{K}$ bezeichne μ' die Konjugierte von μ in $\mathfrak{K}$.

Nun gilt, wenn $n \in \mathbb{N}$, $n \equiv 1 \bmod 3$

$$\beta(n) = \sum_{\mu \in \mathfrak{o}_3,\ \mu\mu' = n} \mu, \qquad \alpha(n) = \sum_{\mu \in \mathfrak{o}_3,\ \mu\mu' = n} 1,$$

insbesondere $\beta(1) = \alpha(1) = 1$, und man beweist zunächst in genauer Analogie zu der entsprechenden Überlegung in § 1:

$$\beta(l)\,\beta(m) = \beta(lm) \quad \text{für} \quad l, m \in \mathbb{N}, \qquad (l, 6) = (m, 6) = (l, m) = 1.$$

Daraus folgt, daß die abermals für $\operatorname{Re} s > 3/2$ absolut konvergente Dirichlet-Reihe $D(s) = \sum_{\substack{n=1 \\ n \equiv 1\,(6)}}^{\infty} \beta(n)\, n^{-s}$ mit den Koeffizienten $\beta(n)$ nach (2.3) eine Eulersche Zerlegung von der Gestalt

$$D(s) = \prod_{p > 3} E_p(p^{-s}) \qquad \text{mit} \quad E_p(w) = \sum_{k=0}^{\infty} \beta(p^k)\, w^k$$

gestattet; man hat jetzt die Euler-Faktoren $E_p(w)$ zu bestimmen.

Für $p \equiv -1 \bmod 6$ ergibt sich nach (2.3,4) unmittelbar

$$E_p(w) = (1 + pw^2)^{-1}.$$

Es sei $p \equiv 1 \bmod 6$, $p = m_1^2 + 3m_2^2$ $(m_1, m_2 \in \mathbb{Z})$, $m_1 \equiv 1 \bmod 3$, $m_2 > 0$. Man erhält zunächst $\beta(p) = 2m_1$. Die nicht-triviale Darstellung von p^2 in der Gestalt $p^2 = n_1^2 + 3n_2^2$ mit $n_1, n_2 \in \mathbb{Z}$ entsteht nach (2.4,5) aus

$$p^2 = (m_1^2 + 3m_2^2)^2 = (m_1^2 - 3m_2^2)^2 + 12m_1^2(\varepsilon m_2)^2 \qquad (\varepsilon = \pm 1)$$

und liefert, da $m_2 \neq 0$,

$$\beta(p^2) = p + 2(m_1^2 - 3m_2^2) = 4m_1^2 - p;$$

also genau dann $\beta(p^2) = a^2 - p = g_2$, wenn $b = -p$ gesetzt wird. Bei geeigneter Vorzeichenwahl ergeben sich dann für die Zahlen (1.9) ρ, ρ^* die Werte

$$\rho = m_1 + m_2\sqrt{-3} \in \mathfrak{o}_3, \qquad \rho^* = m_1 - m_2\sqrt{-3} = \rho' \in \mathfrak{o}_3,$$

so daß $p = \rho\rho'$. Nach der Argumentation von § 1, die genau analogisiert werden kann, gewinnt man hieraus zusammenfassend

Satz 2. *Die Dirichlet-Reihe* $D(s)=\sum_{\substack{n=1\\ n\equiv 1(6)}}^{\infty}\beta(n)n^{-s}$, *die der Modulform* (2.1) $\eta^4(\tau)$ *durch die Koeffizientenformel* (2.3) *entspricht, konvergiert für* $\mathrm{Re}\,s>3/2$ *absolut und gestattet die Eulersche Produktzerlegung*

$$D(s)=\prod_{p>3}E_p(p^{-s}) \quad \text{mit} \quad E_p(w)=(1-2m_1w+pw^2)^{-1}. \tag{2.6}$$

Hier ist $p=m_1^2+3m_2^2$ *mit* $m_1, m_2\in\mathbb{Z}$, $m_1\equiv 1 \bmod 3$, *wenn* $p\equiv 1 \bmod 6$, *dagegen* $m_1=0$, *wenn* $p\equiv -1 \bmod 6$.

Eine andere Formulierung des gleichen Sachverhalts entsteht unter Benutzung des oben definierten Ringes $\mathfrak{o}_1$ in $\mathbb{Q}(\sqrt{-3})$ und besagt

Korollar 2. *Es sei* $\mathfrak{P}_3$ *die Menge der zu* 6 *teilerfremden Primzahlen* ω *in* $\mathfrak{o}_1$, *für die* $\mathrm{Re}\,\omega\equiv 1 \bmod 3$ *zutrifft. Dann gilt mit* N *als Norm in* $\mathbb{Q}(\sqrt{-3})$

$$D(s)=\prod_{\omega\in\mathfrak{P}_3}\left(1-\frac{\omega}{(\mathsf{N}\omega)^s}\right)^{-1}.$$

Hinsichtlich des analytischen Verhaltens der in Satz 2 nochmals definierten Dirichlet-Reihe $D(s)$ ergibt sich nach [6] Satz 4 das genaue Analogon von Korollar 1′, d.h.

Korollar 2′. *Die Funktion*

$$\Xi(s):=\left(\frac{\pi}{3}\right)^{-s}\Gamma(s)D(s) \quad (\beta(n) \textit{ nach } (2.3))$$

ist in die ganze s-Ebene holomorph fortsetzbar und genügt der Funktionalgleichung $\Xi(2-s)=\Xi(s)$.

Zu der Produktzerlegung (2.6) von Satz 2 ist zu bemerken, daß ein Resultat dieser Art von vornherein zu erwarten war, da $\eta^4(\tau)$ bis auf konstante Faktoren das einzige Differential erster Gattung der Hauptkongruenzgruppe $\Gamma[6]$ darstellt. Eine Produktzerlegung der zugehörigen Dirichlet-Reihe folgt dann daraus, daß sich $\eta^4(\tau)$ bei Ausübung der Heckeschen Operatoren bis auf konstante Faktoren reproduziert; die Euler-Faktoren stimmen ersichtlich überein (vgl. [3]). - Über die in EB und hier oben eingeführte einfache Thetareihe $\vartheta_3^{(1)}(\tau)$ sei, wie bereits in EB (§2, S. 33) bemerkt, daß ihre dritte Potenz die bis auf konstante Faktoren einzige ganze Spitzenform der Formenklasse von $\vartheta_3^9(\tau)$ darstellt. Ihre Fourier-Entwicklung

$$\vartheta_3^{(1)3}(\tau)=\sum_{n=1}^{\infty}c_{3,n}e^{\pi i n\tau}$$

hat die Koeffizienten

$$c_{3,n}=\sum_{m_1,m_2,m_3}m_1m_2m_3 \quad udB \quad \left\{\begin{matrix} m_1\equiv m_2\equiv m_3\equiv 1 \bmod 3\\ m_1^2+m_2^2+m_3^2=3n\end{matrix}\right\}.$$

§3. Das dritte Beispiel für Dirichlet-Reihen mit Eulerscher Produktzerlegung steht in enger Beziehung zu der Modulform

$$\varphi_2(\tau) := \eta^3(2\tau)\,\vartheta_3(\tau) = \sum_{\substack{n=1 \\ n\equiv 1(2)}}^{\infty} \beta_2(n) e^{\pi i n \frac{\tau}{2}} \tag{3.1}$$

(vgl. EB S. 271ff.), wo

$$\beta_2(n) := \sum_{m_1, m_2} m_1 \quad udB \quad m_1 \equiv 1 \bmod 4, \qquad m_1^2 + 2m_2^2 = n. \tag{3.2}$$

Die Lösbarkeit der letzten Gleichung bei ungeradem n erfordert $n \equiv 1, 3 \bmod 8$ gemäß $m_2 \equiv 0$ bzw. $1 \bmod 2$. Setzt man

$$\alpha_2(n) = \sum_{m_1, m_2} 1 \quad udB \quad m_1 \equiv 1 \bmod 4, \qquad m_1^2 + 2m_2^2 = n \tag{3.3}$$

so erhält man, wie bekannt,

$$\alpha_2(n) = \sum_{d>0, d|n} \left(\frac{-2}{d}\right) \qquad (n \in \mathbb{N},\ n \equiv 1 \bmod 2), \tag{3.4}$$

also insbesondere für Primzahlen $n = p > 2$

$$\alpha_2(p) = \begin{cases} 2, & \text{wenn } p \equiv 1, 3 \bmod 8 \\ 0, & \text{wenn } p \equiv 5, 7 \bmod 8 \end{cases}.$$

Es sei $\mathfrak{K} := \mathbb{Q}(\sqrt{-2})$, $\mathfrak{o}$ die Hauptordnung in $\mathfrak{K}$,

$$\mathfrak{o}_4 := \{\mu = m_1 + m_2\sqrt{-2} \mid m_1, m_2 \in \mathbb{Z},\ m_1 \equiv 1 \bmod 4\}$$

und $-\mathfrak{o}_4 := \{-\mu \mid \mu \in \mathfrak{o}_4\}$. Dann gilt nach (3.2,3) für $n \in \mathbb{N}$, $n \equiv 1 \bmod 2$

$$\alpha_2(n) = \sum_{\mu \in \mathfrak{o}_4,\, \mu\mu' = n} 1, \qquad \beta_2(n) = \sum_{\mu \in \mathfrak{o}_4,\, \mu\mu' = n} \mu; \tag{3.5}$$

der Akzent kennzeichnet die Konjugierte der betr. Zahl in $\mathfrak{K}$. Wenn $\lambda = l_1 + l_2\sqrt{-2} \in \mathfrak{o}_4$, $\mu = m_1 + m_2\sqrt{-2} \in \mathfrak{o}_4$, so ist

$$\begin{aligned} &\lambda\mu \in \mathfrak{o}_4, && \text{falls } l_2 m_2 \equiv 0 \bmod 2, \\ &\lambda\mu \in -\mathfrak{o}_4, && \text{falls } l_2 m_2 \equiv 1 \bmod 2. \end{aligned} \tag{3.6}$$

Mit $l := \lambda\lambda'$, $m := \mu\mu'$ besagt $l_2 m_2 \equiv 0 \bmod 2$, daß l oder m (oder beide) $\equiv 1 \bmod 8$, dagegen $l_2 m_2 \equiv 1 \bmod 2$, daß $l \equiv m \equiv 3 \bmod 8$. Wird $\varepsilon_{l,m}$ dementsprechend erklärt, also für $l, m \in \mathbb{N}$, $l \equiv 1, 3$ und $m \equiv 1, 3 \bmod 8$ durch

$$\varepsilon_{l,m} := \begin{cases} +1, & \text{falls } l \text{ oder } m \text{ (oder beide)} \equiv 1 \bmod 8 \\ -1, & \text{falls } l \equiv m \equiv 3 \bmod 8 \end{cases},$$

so folgt nun $\varepsilon_{l,m}\lambda\mu \in \mathfrak{o}_4$ aus $\lambda \in \mathfrak{o}_4$, $\mu \in \mathfrak{o}_4$.

Es sei $l, m \in \mathbb{N}$, $lm \equiv 1 \bmod 2$, $(l, m) = 1$. Wenn nicht sowohl l als auch

$m \equiv 1, 3 \bmod 8$ ist, so gilt $\beta_2(l)\beta_2(m) = \beta_2(lm)$, weil beide Seiten nach (3.4) verschwinden. Wenn andrerseits sowohl l als auch $m \equiv 1, 3 \bmod 8$ ist, so gilt nach (3.5)

$$\varepsilon_{l,m}\beta_2(l)\beta_2(m) = \sum_{\substack{\lambda,\mu \in \mathfrak{o}_4 \\ \lambda\lambda' = l,\, \mu\mu' = m}} \varepsilon_{l,m}\lambda\mu.$$

Für die Glieder $\nu := \varepsilon_{l,m}\lambda\mu$ dieser Summe gilt $\nu \in \mathfrak{o}_4$, $\nu\nu' = n := lm$ und nach (3.4) stimmt die Anzahl dieser ν mit $\alpha_2(n)$ überein. Aufgrund der in § 1 angewendeten Schlußweise ergibt sich daher

$$\varepsilon_{l,m}\beta_2(l)\beta_2(m) = \beta_2(lm) \qquad (l, m \in \mathbb{N},\ lm \equiv 1 \bmod 2,\ (l,m) = 1).$$

Hinsichtlich der hier betrachteten l, m ($\equiv 1$ oder $3 \bmod 8$) kann die Relation

$$\xi_4^{l-1+m-1} = \varepsilon_{l,m}\xi_4^{lm-1} \qquad \left(\xi_4 := \exp\frac{\pi i}{4}\right)$$

benutzt werden; sie führt auf das folgende erste Resultat:

In der Bezeichnung (3.2) $\beta_2(n)$ sei

$$\beta^*(n) := \xi_4^{n-1}\beta_2(n) \qquad (n \in \mathbb{N},\ n \equiv 1 \bmod 2). \tag{3.7}$$

Dann ist $\beta^*(1) = 1$ und

$$\beta^*(k)\beta^*(l) = \beta^*(kl) \qquad (k, l \in \mathbb{N},\ kl \equiv 1 \bmod 2,\ (k,l) = 1).$$

Dies besagt, daß die Dirichlet-Reihe

$$D^*(s) = \sum_{\substack{n=1 \\ n \equiv 1\,(2)}}^{\infty} \beta^*(n)n^{-s} = \sum_{\substack{n=1 \\ n \equiv 1\,(2)}}^{\infty} \xi_4^{n-1}\beta_2(n)n^{-s}, \tag{3.8}$$

die offenbar für $\operatorname{Re} s > 3/2$ absolut konvergiert, eine Eulersche Produktzerlegung gestattet. Von dieser sind nun die Faktoren zu bestimmen.

Wir schreiben die Faktoren in der Gestalt

$$E_p^*(w) := \sum_{k=0}^{\infty} \beta^*(p^k)w^k \qquad (p \text{ Primzahl} > 2)$$

und erhalten zunächst nach (3.4) mit $\varepsilon_p := \left(\frac{-1}{p}\right)$:

$$E_p^*(w) = (1 - \varepsilon_p p w^2)^{-1}, \qquad \text{wenn } p \equiv 5, 7 \bmod 8.$$

In den Fällen $p \equiv 1, 3 \bmod 8$, die jetzt untersucht werden, hat man zunächst $\beta_2(p)$ und $\beta_2(p^2)$ zu berechnen. Mit

$$p = m_1^2 + 2m_2^2 \qquad (m_1, m_2 \in \mathbb{Z}),\ m_1 \equiv 1 \bmod 4,\ m_2 > 0 \tag{3.9}$$

ergibt sich $\beta_2(p) = 2m_1$. Nach (3.4) ist $\alpha_2(p^2) = 3$; eine der Darstellungen von p^2 in der Gestalt $k_1^2 + 2k_2^2$ mit $k_1, k_2 \in \mathbb{Z}$, $k_1 \equiv 1 \bmod 4$ liefert $k_1 = \varepsilon_p p$. Die beiden

anderen entstehen aus

$$p^2=(m_1^2+2m_2^2)^2=(m_1^2-2m_2^2)^2+2m_1^2(2m_2)^2$$

und liefern, da $m_1^2-2m_2^2\equiv p \bmod 4$ und $m_2\neq 0$, zusammen mit der obigen

$$\beta_2(p^2)=\pm(p+2(m_1^2-2m_2^2))=\pm(4m_1^2-p) \quad \text{für } p\equiv 2\mp 1 \bmod 8.$$

Im Falle $p\equiv 1 \bmod 8$ ist $\beta^*(p^k)=\beta_2(p^k)$ $(k\in\mathbb{N})$ und, wenn $a=2m_1$, $b=-p$ gesetzt wird (s. (1.8)):

$$\beta_2(p)=a=g_1, \quad \beta_2(p^2)=a^2+b=g_2,$$

wodurch ρ, ρ^*, abgesehen von der Reihenfolge, die Werte $m_1\pm m_2\sqrt{-2}$ erhalten. Wir schreiben $\rho=m_1+m_2\sqrt{-2}$, haben also $p=\rho\rho'$ (vgl. (1.9)). Wegen $m_2\equiv 0 \bmod 2$ reproduzieren sich die Relationen ρ, $\rho'\in\mathfrak{o}_4$ bei Multiplikation, und man findet daher nach dem Vorgehen in §1

$$E_p^*(w)=(1-2m_1w+pw^2)^{-1} \quad (p\equiv 1 \bmod 8).$$

Im Falle $p\equiv 3 \bmod 8$, den wir abschließend betrachten, gilt (3.9) mit $m_2\equiv 1 \bmod 2$. Ferner gilt, wie bewiesen,

$$\beta_2(p)=2m_1, \quad \beta^*(p)=2m_1 i, \quad \beta_2(p^2)=\beta^*(p^2)=-4m_1^2+p.$$

In der Terminologie von §1 wird also $\beta^*(p)=a$, $\beta^*(p^2)=a^2+b$ genau dann erreicht, wenn $a=2m_1 i$, $b=p$. Mit diesen Werten, die jetzt fest gewählt werden, gilt ρ^* oder $\rho=m_1 i+m_2\sqrt{2}$; es kann

$$\rho=i\sigma', \quad \rho^*=i\sigma, \quad \sigma:=m_1+m_2\sqrt{-2}\in\mathfrak{o}_4 \quad (\sigma\sigma'=p)$$

definiert werden, und das Verfahren von §1 liefert nun folgendes Resultat:

$$\mu\in\mathfrak{o}_4, \quad \mu\mu'=p^k \quad (k\in\mathbb{N}) \tag{3.10}$$

besteht genau dann, wenn $\mu=\varepsilon'\sigma^j\sigma'^{k-j}$ für $0\leqq j\leqq k$ $(j\in\mathbb{Z})$ und ein zugehöriges $\varepsilon'=\pm 1$ zutrifft.

Es zeigt sich, daß ε' daraus berechnet werden kann. Zu diesem Zweck führen wir die folgenden Zahlmengen ein:

$$\mathfrak{o}_{4,f}:=\{\lambda=l_1+l_2\sqrt{-2}\in\mathfrak{o}_4 \mid l_1, l_2\in\mathbb{Z},\ l_2\equiv f \bmod 2\}$$
$$-\mathfrak{o}_{4,f}:=\{-\lambda \mid \lambda\in\mathfrak{o}_{4,f}\} \quad (f\in\mathbb{Z}).$$

Zunächst ergibt sich: Wenn $\lambda\in\mathfrak{o}_{4,f}$, $\mu\in\mathfrak{o}_{4,g}$ $(f,g\in\mathbb{Z})$, so gilt $\lambda\mu\in(-1)^{fg}\mathfrak{o}_{4,f+g}$. Sind ferner $\gamma_1,\gamma_2,\gamma_3,\ldots\in\mathfrak{o}_{4,1}$, so gilt

$$\gamma_1\gamma_2\cdots\gamma_g\in\left(\frac{2}{2g-1}\right)\mathfrak{o}_{4,g} \quad (g\in\mathbb{N}).$$

Die rechte Seite ist periodisch in g mit der Periode 4; σ und σ' liegen in $\mathfrak{o}_{4,1}$.

In (3.10) ist einerseits $\mu \in \mathfrak{o}_{4,k}$, andrerseits aufgrund der Darstellung von μ durch σ, σ': $\mu \in \varepsilon' \left(\frac{2}{2k-1}\right) \mathfrak{o}_{4,k}$. Daraus folgt $\varepsilon' = \left(\frac{2}{2k-1}\right)$, so daß nun

$$\beta_2(p^k) = \sum_{\mu \in \mathfrak{o}_4, \mu\mu' = p^k} \mu = \left(\frac{2}{2k-1}\right) \sum_{j=0}^{k} \sigma^j \sigma'^{k-j}.$$

Diese Formel liefert die abschließende Darstellung von $\beta^*(p^k)$ vermöge der elementaren Relation

$$\xi_4^{p^k-1} \left(\frac{2}{2k-1}\right) = i^k \qquad (k \in \mathbb{N}).$$

Danach ist

$$\beta^*(p^k) = i^k \sum_{j=0}^{k} \sigma^j \sigma'^{k-j} = \sum_{j=0}^{k} \rho^{*j} \rho^{k-j},$$

und das besagt nach § 1

$$E_p^*(w) = (1 - 2m_1 i w - p w^2)^{-1} \qquad (p \equiv 3 \bmod 8).$$

Damit sind die Faktoren der Eulerschen Produktzerlegung der Dirichlet-Reihe (3.8) $D^*(s)$ bestimmt. Die zusammenfassende Formulierung der Resultate soll am Ende von § 4 erfolgen.

§ 4. Die folgenden Ausführungen über die Dirichlet-Reihe $D^*(s)$ und die zugehörige Fourier-Reihe $\varphi^*(\tau)$ (Def. s.w.u. (4.1)) dienen dazu, diese Funktionen mitsamt einer gewissen Parallelkonstruktion ($D_*(s)$ und $\varphi_*(\tau)$) in den Apparat von [6], Satz 4 einzuordnen. Das Endresultat betrifft dann alle vier Funktionen $\varphi_*(\tau)$, $\varphi^*(\tau)$, $D_*(s)$, $D^*(s)$ gemeinsam.

Mit $\xi_m := \exp \frac{\pi i}{m}$ $(m \in \mathbb{N})$ und (vgl. EB (1.35))

$$\eta(\tau) := \left(\exp \pi i \frac{\tau}{12}\right) \prod_{m=1}^{\infty} (1 - e^{2\pi i m \tau})$$

gilt $\eta(\tau+1) = \xi_{12} \eta(\tau)$ und (vgl. (3.1))

$$\varphi_2\left(\frac{\tau+1}{2}\right) = \sum_{\substack{n=1 \\ n \equiv 1(2)}}^{\infty} \beta_2(n) \xi_4^n e^{\pi i n \frac{\tau}{4}} = \eta^3(\tau+1) \vartheta_3\left(\frac{\tau+1}{2}\right),$$

so daß

$$\varphi^*(\tau) := \vartheta_3\left(\frac{\tau+1}{2}\right) \eta^3(\tau) = \sum_{\substack{n=1 \\ n \equiv 1(2)}}^{\infty} \beta^*(n) e^{\pi i n \frac{\tau}{4}}. \tag{4.1}$$

Daraus erhält man, indem man τ durch $\tau - 2$ ersetzt

$$\varphi_*(\tau) := \vartheta_3\left(\frac{\tau-1}{2}\right) \eta^3(\tau) = \sum_{\substack{n=1 \\ n \equiv 1(2)}}^{\infty} \beta_*(n) e^{\pi i n \frac{\tau}{4}}, \tag{4.2}$$

wo (vgl. (3.7))

$$\beta_*(n) := \left(\frac{-1}{n}\right)\beta^*(n) = \xi_4^{-n+1}\beta_2(n) \qquad (n\in\mathbb{N},\ n\equiv 1 \bmod 2).$$

Die oben formulierte Aufgabe bedeutet im wesentlichen, daß Formenklassen $\{\Gamma, -r, v\}$ der Modulgruppe anzugeben sind, die $\varphi^*(\tau)$ und $\varphi_*(\tau)$ als ganze Spitzenformen enthalten. Hier bezeichnet Γ eine Kongruenzgruppe in der Modulgruppe, es ist $r=2$, und v besteht aus gewissen Einheitswurzeln. Die Bestimmung von Γ beruht auf der näheren Kenntnis der Konjugationsklasse von $\Gamma^0[4]$ in ${}_1\Gamma$, weshalb wir jetzt diese untersuchen (zur Terminologie vgl. EB, § 1 und Anh. C).

Danach besteht mit $\Gamma := \Gamma^0[4]$ die disjunkte Zerlegung

$${}_1\Gamma = \bigcup_{j=0}^{3} \Gamma U^j \cup \Gamma T \cup \Gamma U^2 T.$$

Die sämtlichen Konjugierten von $\Gamma^0[4]$ in ${}_1\Gamma$ finden sich also unter den Gruppen

$$\Gamma^{0,j}[4] := U^{-j}\Gamma^0[4]U^j \qquad (j=0,1,2,3),$$
$$\Gamma_0[4] = T^{-1}\Gamma^0[4]T, \qquad (U^2T)^{-1}\Gamma^0[4]U^2T.$$

$L=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\in\Gamma^{0,j}[4]$ bedeutet $L\in{}_1\Gamma$ und zusätzlich

$$\beta + j\delta - j(\alpha + j\gamma) \equiv 0 \bmod 4. \tag{4.3}$$

Wir erweitern die obige Bedeutung von $\Gamma^{0,j}[4]$ auf alle $j\in\mathbb{Z}$ und beweisen zunächst, daß gilt

$$\Gamma^{0,1}[4] = \Gamma^{0,-1}[4], \qquad \Gamma^{0,2}[4] = \Gamma^0[4].$$

Für $j=\pm 1$ folgt aus (4.3): $\beta+\delta+\alpha+\gamma\equiv 0 \bmod 2$, so daß sich $\Gamma^{0,\pm 1}[4]\subset\Gamma_\vartheta$ herausstellt (zu Γ_ϑ s. EB (1.6)). Daher ist $\alpha\equiv\delta \bmod 2$. Gilt (4.3) mit $j=1$, so nach Addition von $-2\delta+2\alpha$ auch mit $j=-1$, was $\Gamma^{0,1}[4]\subset\Gamma^{0,-1}[4]$, also die erste Behauptung liefert. Für $j=2$ besagt (4.3) soviel wie

$$\beta + 2\delta - 2\alpha \equiv 0 \bmod 4.$$

Da dies für alle $L\in\Gamma^0[4]$ zutrifft, erhält man $\Gamma^0[4]\subset\Gamma^{0,2}[4]$, woraus die zweite Behauptung folgt. Aus dieser folgt weiter

$$(U^2T)^{-1}\Gamma^0[4]U^2T = \Gamma_0[4].$$

Danach besteht die Konjugationsklasse von $\Gamma^0[4]$ in ${}_1\Gamma$ aus den drei Gruppen

$$\Gamma^0[4], \qquad \Gamma^{0,1}[4], \qquad \Gamma_0[4]. \tag{4.4}$$

Daß diese paarweise verschieden sind, ist u.a. wie folgt zu sehen: Die Spitze ∞ hat in $\Gamma^0[4]$ und in $\Gamma^{0,1}[4]$ die Breite 4, in $\Gamma_0[4]$ die Breite 1. $\Gamma^0[4]\neq\Gamma^{0,1}[4]$ folgt daraus, daß $\Gamma^{0,1}[4]$, nicht aber $\Gamma^0[4]$ in Γ_ϑ enthalten ist. Das war die Behauptung. Aus ihr ergibt sich u.a., daß $\Gamma^0[2]$ der Normalisator von $\Gamma^0[4]$ in

${}_1\Gamma$ ist. Eine spezielle Konsequenz dieser Aussagen, die später benutzt wird, besagt

$$\Gamma^{0,1}[4]=T^{-1}\Gamma^{0,1}[4]\,T. \tag{4.5}$$

In der Tat ist $T^{-1}LT=\begin{pmatrix}\delta & -\gamma\\ -\beta & \alpha\end{pmatrix}$; gilt (4.3) mit $j=1$ für L, so mit $j=-1$ für $T^{-1}LT$. Es gilt aber *nicht* $T\in\Gamma^{0,1}[4]$.

Zur Untersuchung des Verhaltens von $\vartheta_3\left(\frac{\tau-1}{2}\right)$ bei Modulsubstitutionen schreiben wir $S:=\begin{pmatrix}1 & -1\\ 0 & 2\end{pmatrix}$, so daß

$$\vartheta_3\left(\frac{\tau-1}{2}\right)=(\vartheta_3(\tau)|S)\sqrt{2}\qquad(|:=|_r,\ r=1/2)$$

und fragen nach den $L=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\in{}_1\Gamma$, für die $SL=L^*S$ mit $L^*\in\Gamma_\vartheta$ zutrifft. Eine kurze Rechnung ergibt als notwendige und hinreichende Bedingung hierfür

$$\alpha+\beta-\gamma-\delta\equiv 0 \bmod 4,\qquad \text{d.h. } L\in\Gamma^{0,1}[4]$$

mit der Konsequenz $L^*\in\Gamma[2]$.

Da die Symbole $w(S,L)$ und $w(L^*,S)$ offenbar verschwinden, erhält man (zum Folgenden vgl. EB §2, §4)

$$\vartheta_3(\tau)|S|L=\vartheta_3(\tau)|L^*S=v_3(L^*)\,\vartheta_3(\tau)|S$$

d.h.

$$\vartheta_3\left(\frac{\tau-1}{2}\right)\Bigg|L=v_3(L^*)\,\vartheta_3\left(\frac{\tau-1}{2}\right)\qquad(L\in\Gamma^{0,1}[4])$$

mit

$$v_3(L^*)=\left(\frac{2\gamma}{\gamma+\delta}\right)_*\xi_4^{\gamma+\delta-1}\qquad(\gamma+\delta\equiv 1 \bmod 2). \tag{4.6}$$

Die Anwendung von EB Satz 1.11 erweist $\vartheta_3\left(\frac{\tau-1}{2}\right)$ als ganze Modulform $\{\Gamma^{0,1}[4],\ -1/2,\ v_*\}$, wo $v_*(L)=v_3(L^*)$ für $L\in\Gamma^{0,1}[4]$.

Das entsprechende Resultat über $\vartheta_3\left(\frac{\tau+1}{2}\right)$ ist analog zu gewinnen. Für $S':=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}$ erhält man $S'L=L_1^*S'$ mit $L\in{}_1\Gamma$, $L_1^*\in\Gamma_\vartheta$ genau dann, wenn $L\in\Gamma^{0,1}[4]$, und dann ist $L_1^*\in\Gamma[2]$. Dies führt auf

$$\vartheta_3(\tau)|S'L=\vartheta_3(\tau)|L_1^*S'=v_3(L_1^*)\,\vartheta_3(\tau)|S';$$

also stellt $\vartheta_3\left(\frac{\tau+1}{2}\right)$ eine ganze Modulform der Klasse $\{\Gamma^{0,1}[4],\ -\frac{1}{2},\ v^*\}$ dar, wo

$$v^*(L)=v_3(L_1^*)=\left(\frac{2\gamma}{-\gamma+\delta}\right)_*\xi_4^{-\gamma+\delta-1}\qquad(L\in\Gamma^{0,1}[4]). \tag{4.7}$$

Ein Zusammenhang zwischen $\vartheta_3\left(\frac{\tau-1}{2}\right)$ und $\vartheta_3\left(\frac{\tau+1}{2}\right)$ wird durch Transformation mit $T=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ vermittelt. Zunächst gilt

$$\vartheta_3\left(\frac{\tau-1}{2}\right)\Big| T=\sqrt{2}\,\vartheta_3(\tau)|S_1 \qquad \left(S_1:=\begin{pmatrix}-1 & -1\\ 2 & 0\end{pmatrix}\right);$$

S_1 kann in der Gestalt $S_1=-\dot{U}^{-2}S'$ geschrieben werden. Nach EB (1.23) und (2.10) folgt nun

$$\vartheta_3(\tau)|S_1=(-i)\,\vartheta_3(\tau)|S'=-i\vartheta_3\left(\frac{\tau+1}{2}\right)\sqrt{2}-1$$

und demnach (vgl. (4.5))

$$\vartheta_3\left(\frac{\tau-1}{2}\right)\Big| T=-i\vartheta_3\left(\frac{\tau+1}{2}\right). \tag{4.8}$$

Damit kann abschließend das vollständige Resultat der Untersuchungen von § 3, § 4 formuliert werden (zur Terminologie vgl. EB §§ 1, 2, 4).

Satz 3. *Die beiden einfachen Thetareihen*

$$\vartheta_3\left(\frac{\tau-1}{2}\right)=\vartheta_3(2\tau)-i\vartheta_2(2\tau), \qquad \vartheta_3\left(\frac{\tau+1}{2}\right)=\vartheta_3(2\tau)+i\vartheta_2(2\tau)$$

stellen ganze Modulformen der Gruppe $\Gamma^{0,1}[4]$ *dar.* $L=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}\in\Gamma^{0,1}[4]$ *bedeutet* $L\in{}_1\Gamma$ *und* $\beta+\delta-\alpha-\gamma\equiv 0 \bmod 4$; $\Gamma^{0,1}[4]$ *hat in* ${}_1\Gamma$ *den Index* 6. *Es gilt*

$$\vartheta_3\left(\frac{\tau-1}{2}\right)\in\{\Gamma^{0,1}[4],-1/2,v_*\}^0, \qquad \vartheta_3\left(\frac{\tau-1}{2}\right)\in\{\Gamma^{0,1}[4],-1/2,v^*\}^0$$

mit

$$v_*(L)=\left(\frac{2\gamma}{\gamma+\delta}\right)_*\xi_4^{\gamma+\delta-1}, \qquad v^*(L)=\left(\frac{2\gamma}{-\gamma+\delta}\right)_*\xi_4^{-\gamma+\delta-1}$$

und ferner (4.8).

Die beiden Funktionen

$$\varphi_*(\tau):=\vartheta_3\left(\frac{\tau-1}{2}\right)\eta^3(\tau), \qquad \varphi^*(\tau):=\vartheta_3\left(\frac{\tau+1}{2}\right)\eta^3(\tau)$$

stellen ganze Spitzenformen der Gruppe $\Gamma^{0,1}[4]$ dar. Genauer gilt

$$\varphi_*(\tau)\in\{\Gamma^{0,1}[4],-2,v_*v_5^3\}^+, \qquad \varphi^*(\tau)\in\{\Gamma^{0,1}[4],-2,v^*v_5^3\}^+$$

(zu $^+$ s. EB S. 13, zu v_5 s. EB (4.14)). Die beiden Spitzen ∞ und 0 haben wegen

$$0=L_2\infty; \qquad L_2:=\begin{pmatrix}0 & -1\\ 1 & 2\end{pmatrix}\in\Gamma^{0,1}[4]$$

in $\Gamma^{0,1}[4]$ die Breite 4 (vgl. (4.5)), und sowohl von $v_*v_5^3$ als auch von $v^*v_5^3$ ist der Drehrest κ im Unendlichen gleich 1/2. Die Fourier-Entwicklungen von φ_* und von φ^* weisen die Gestalt

$$\varphi_*(\tau)=\sum_{\substack{n=1\\ n\equiv 1(2)}}^{\infty}\beta_*(n)\exp\pi i n\frac{\tau}{4},\qquad \varphi^*(\tau)=\sum_{\substack{n=1\\ n\equiv 1(2)}}^{\infty}\beta^*(n)\exp\pi i n\frac{\tau}{4}$$

auf, wo (vgl. (3.2))

$$\beta_*(n)=\xi_4^{1-n}\beta_2(n),\qquad \beta^*(n)=\xi_4^{n-1}\beta_2(n).$$

Es gilt mit $|:=|_2$

$$\varphi_*(\tau)|T=\xi_4^3\varphi^*(\tau).$$

Satz 4. *Die beiden Dirichlet-Reihen*

$$D_*(s):=\sum_{\substack{n=1\\ n\equiv 1(2)}}^{\infty}\beta_*(n)n^{-s},\qquad D^*(s)=\sum_{\substack{n=1\\ n\equiv 1(2)}}^{\infty}\beta^*(n)n^{-s}$$

sind für $\mathrm{Re}\,s>3/2$ *absolut konvergent; sowohl sie als auch die Funktionen*

$$\Xi_*(s):=\left(\frac{\pi}{4}\right)^{-s}\Gamma(s)D_*(s),\qquad \Xi^*(s):=\left(\frac{\pi}{4}\right)^{-s}\Gamma(s)D^*(s)$$

sind in die volle s-Ebene holomorph fortsetzbar. Es besteht die Funktionalgleichung

$$\Xi^*(2-s)=\xi_4\,\Xi_*(s).$$

Satz 5. *Sowohl* $D_*(s)$ *als auch* $D^*(s)$ *gestattet eine Eulersche Produktzerlegung*

$$D_*(s)=\prod_{p>2}E_{*p}(p^{-s}),\qquad D^*(s)=\prod_{p>2}E_p^*(p^{-s}),$$

wo

$$E_p^*(w)=(1-\xi_4^{p-1}2m_1w+\varepsilon_p p w^2)^{-1},$$
$$E_{*p}(w)=(1-\xi_4^{1-p}2m_1w+\varepsilon_p p w^2)^{-1}.$$

Hier ist $m_1=m_1(p)$ *durch* $p=m_1^2+2m_2^2$ $(m_1,m_2\in\mathbb{Z})$, $m_1\equiv 1 \bmod 4$ *definiert, falls* $p\equiv 1,3 \bmod 8$; *dagegen durch* $m_1=0$, *falls* $p\equiv 5,7 \bmod 8$. *Ferner ist* $\varepsilon_p:=\left(\frac{2}{p}\right)$ *und* $\xi_4:=\exp\frac{\pi i}{4}$.

Zu den quadratischen Polynomen in der Variablen w, durch die sich die Euler-Faktoren $E_p(w)$, $E_p^*(w)$ im obigen Text ausdrücken, sei folgendes bemerkt: Bei Benutzung der Bezeichnungen des Textes gilt in §1 für $p\equiv 1 \bmod 4$, in §2 für $p\equiv 1 \bmod 6$, in §3 für $p\equiv 1 \bmod 8$

$$E_p^{-1}(w)=(1-\rho w)(1-\rho' w)\quad\text{mit}\quad \rho\rho'=p;$$

in §3 für $p\equiv 3 \bmod 8$

$$E_p^{*-1}(w)=(1-i\sigma w)(1-i\sigma' w)\quad\text{mit}\quad \sigma\sigma'=p.$$

ρ,ρ', σ,σ' haben den gleichen Betrag $\sqrt{p}$, daher die Nullstellen der obigen Polynome den Betrag $p^{-1/2}$.

Literatur

1. Gauss, C.F.: Theoria residuorum biquadratorum Commentatio prima, Gött. Nachr. 1828; Werke Bd. II, S. 90, 91
2. Hecke, E.: Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen II. Math. Z. **6**, 11 (1920)
3. Hecke, E.: Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung I, II. Math. Ann. **114**, 1, 316 (1937)
4. Petersson, H. (Sen.): Konstruktion der sämtlichen Lösungen einer Riemannschen Funktionalgleichung ..., I. Math. Ann. **116**, 401 (1939); II, Math. Ann. **117**, 39 (1940)
5. Petersson, H.: Modulfunktionen und quadratische Formen. Berlin-Heidelberg-New York: Springer 1982; im Text zitiert mit EB
6. Petersson, H.: Über die Primformen der Hauptkongruenzgruppen, Abhandl. Math. Sem. Univ. Hamburg 38, S. 8 (1972)

Eingegangen am 5. Dezember 1983

Math. Z. 189, 289-296 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Complete Intersections with Middle Picard Number 1 Defined over Q

Tomohide Terasoma
Department of Mathematics, Faculty of Science, University of Tokyo, Tokyo 113, Japan

§ 0. Introduction

Let X be a smooth algebraic variety over the complex number field $\mathbf{C}$. The space of algebraic cycles of codimension i, denoted by $C^i(X)$, is the $\mathbf{Q}$-subvectorspace of $H^{2i}(X, \mathbf{Q})$ spanned by the cohomology classes of subvarieties of codimension i on X. The dimension of $C^i(X)$ is called the i-th Picard number and will be noted by $\rho^i(X)$.

In this paper, a complete intersection always means a nonsingular complete intersection of hypersurfaces in a projective space. For $d \in \mathbf{Z}$, $\geqq 1$, $\mathbf{a} = (a_1, \ldots, a_d)$ ($a_i \in \mathbf{Z}$, $\geqq 2$, $i = 1, \ldots, d$), a complete intersection is said to be of type $(N; \mathbf{a})$ if it is a complete intersection of hypersurfaces of degree $a_1, \ldots, a_d$ in $\mathbf{P}^{N+d}$.

It is known that if X is an odd dimensional complete intersection, the cohomology of even degree is generated by the powers of the class of the hyperplane section, and if X is an even dimensional one, the powers of the class of the hyperplane section generate the cohomology $H^{2i}(X, \mathbf{Q})$ except for $2i = \dim X$. Thus, in studying the algebraic cycles on a complete intersection, the only nontrivial case is when the dimension of X is even ($\dim X = 2n$) and the codimension i is n. In this case $\rho^n(X)$ will be called the middle Picard number of X.

First we recall the following theorem which is a generalization of the classical Noether's theorem.

Theorem ([7]. XIX, Deligne). *A generic complete intersection X of type $(2n; a_1, \ldots, a_d)$ has middle Picard number* 1, *provided that*

$$(2n; a_1, \ldots, a_d) \neq (2; 3), (2n; 2), (2n; 2, 2). \quad \square \qquad (*)$$

(A complete intersection is said to be generic, when its defining equations have algebraically independent coefficients over the prime field $\mathbf{Q}$.)

Shioda has asked if there existed a complete intersection with middle Picard number 1 defined over $\mathbf{Q}$ for every type satisfying (*). He has given an explicit example of such a hypersurface under certain assumptions on the

degree and dimension (see [5]). The purpose of this paper is to give the affirmative answer to this problem.

Theorem 1. *For any type* $(2n; \mathbf{a})$ *satisfying* (*), *there exists a complete intersection* X *of type* $(2n; \mathbf{a})$ *defined over* $\mathbf{Q}$ *such that* $\rho^n(X)=1$. □

We will prove the above theorem by studying the étale cohomology of X with the action of Galois group. Most of the paper is devoted to finding complete intersections for which the image of the Galois group is open in the orthogonal group of the space of the vanishing cycles. To find such complete intersections, we first compare the image of the algebraic fundamental group of a rational curve over $\mathbf{Q}$ in a representation with that of the fundamental group of one of its $\mathbf{Q}$-valued points. Next we apply the above principle to the theory of vanishing cycles.

I would like to thank Prof. T. Shioda for suggesting this problem and giving valuable advice. I would also like to thank K. Kato and T. Oda for useful advice. In particular, K. Kato pointed out me the relevance of Hilbert's irreducibility theorem to this problem (see § 1).

§ 1. Hilbert's Irreducibility Theorem and an Approximation Theorem

In this section we will compare the fundamental group of a rational curve over $\mathbf{Q}$ with the Galois group corresponding to one of its $\mathbf{Q}$-valued points, using Hilbert's irreducibility theorem.

Let A be a subring of the rational function field $\mathbf{Q}(T)$ (T is a variable over $\mathbf{Q}$) which is generated by T and $1/f(T)$ for some polynomial $f(T)$, and let $\bar{\eta}$ be a fixed geometric generic point of $\operatorname{Spec} A$. Let $\pi_1=\pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{\eta})$ be the algebraic fundamental group of $\operatorname{Spec} A$. with the base point $\bar{\eta}$ (cf. [3], p. 39 for definition). Namely, for any finite étale covering $X \to \operatorname{Spec} A$, $\operatorname{Hom}_A(\bar{\eta}, X)$ is a finite discrete π_1-set, and the correspondence $X \rightsquigarrow \operatorname{Hom}_A(\bar{\eta}, X)$ defines an equivalence of the category of finite étale coverings of $\operatorname{Spec} A$ and the category of finite discrete π_1-sets.

For any geometric point $\bar{t}$ over a closed point t of $\operatorname{Spec} A$, and any path γ connecting $\bar{t}$ and $\bar{\eta}$, there is an induced isomorphism:

$$\gamma*: \pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{t}) \to \pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{\eta}).$$

(For the definition of a "path", see [1], p. 220.) We have then the composed homomorphism;

$$\delta(t, \gamma): \operatorname{Gal}(\kappa(\bar{t})/\kappa(t)) = \pi_{1\,\mathrm{alg}}(\bar{t}, t) \xrightarrow{\iota} \pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{t}) \xrightarrow{\gamma*} \pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{\eta}),$$

where $\kappa(t)$ is the residue field of the point t and ι is the natural homomorphism. Given an ℓ-adic continuous representation of $\pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{\eta})$;

$$\varphi: \pi_{1\,\mathrm{alg}}(\operatorname{Spec} A, \bar{\eta}) \twoheadrightarrow G \subset \mathrm{GL}(m, \mathbf{Q}_\ell),$$

where G is a closed subgroup of $\mathrm{GL}(m, \mathbf{Q}_\ell)$, we consider the composed homomorphism:

$$\alpha(t, \gamma, \varphi): \mathrm{Gal}(\kappa(\bar{t})/\kappa(t)) \xrightarrow{\delta(t,\gamma)} \pi_{1\,\mathrm{alg}}(\mathrm{Spec}\, A, \bar{\eta}) \to G.$$

Theorem 2. *Let A, $\bar{\eta}$ and φ be given as above. Then there exists a sequence of distinct $\mathbf{Q}$-valued points $\{t_i\}_{i\in\mathbf{N}}$ of $\mathrm{Spec}\, A$ and a path γ_i connecting $\bar{\eta}$ and a geometric point $\bar{t}_i$ over t_i for each i such that the map*

$$\alpha(t_i, \gamma_i, \varphi): \mathrm{Gal}(\kappa(\bar{t}_i)/\kappa(t_i)) \to G$$

is surjective.

Remark. Whether $\alpha(t_i, \gamma_i, \varphi)$ is surjective or not does not depend on the choice of the path γ_i.

Proof. Since G is an ℓ-adic Lie group, there exists a filtration by compact open subgroups,

$$G \supset G_0 \supset G_1 \supset G_2 \ldots \supset G_n \supset \supset \ldots$$

satisfying the following conditions (cf. [4]).

(1) $\bigcap_n G_n = \{e\}$.

(2) There exists a natural number i_0 such that G_i/G_{i+1} is an abelian ℓ-torsion group for all i greater than i_0.

By the condition (2), ℓ-th power map in G_i induces the following homomorphism.

$$G_i/G_{i+1} \to G_{i+1}/G_{i+2}$$
$$[x] \mapsto [x^\ell]$$

(3) The above homomorphism is an isomorphism.

If $i > i_0$, any representatives of G/G_i generate a dense subgroup of G. Therefore it is sufficient to show the surjectivity of the following composed map;

$$\mathrm{Gal}(\kappa(\bar{t}_i)/\kappa(t_i)) \xrightarrow{\alpha(t_i, \gamma_i, \varphi)} G \to G/G_i.$$

$\varphi^{-1}(G_i)$ is an open subgroup of $\pi_{1\,\mathrm{alg}}(\mathrm{Spec}\, A, \bar{\eta})$, which corresponds to a finite étale covering, say $\mathrm{Spec}\, B$, of $\mathrm{Spec}\, A$. Let y be a generator of the quotient field of B over $\mathbf{Q}(T)$, integral over $\mathbf{Q}[T]$. The defining equation of y is a monic polynomial with coefficient in $\mathbf{Q}[T]$:

$$Y^m + a_1(T)\, Y^{m-1} + \ldots + a_m(T) = 0. \qquad (**)$$

Now, let's consider a $\mathbf{Q}$-valued point $(T - t_0)A$ of $\mathrm{Spec}\, A$ ($t_0 \in \mathbf{Q}$). For each prime ideal $\mathfrak{p}$ of B that divides $(T - t_0)$, we obtain the following natural homomorphism $\gamma|_B$:

$$\begin{array}{ccc} \mathrm{Aut}(B/A) & \longrightarrow & \pi_{1\,\mathrm{alg}}(\mathrm{Spec}\, A, \bar{\eta})/\pi_{1\,\mathrm{alg}}(\mathrm{Spec}\, B, \bar{\eta}) \\ \uparrow{\scriptstyle \gamma|_B} & & \\ \mathrm{Aut}((B/\mathfrak{p})/(A/T - t_0))) & \longrightarrow & \pi_{1\,\mathrm{alg}}(\mathrm{Spec}(A/(T-t_0)), \bar{t}_0)/\pi_{1\,\mathrm{alg}}(\mathrm{Spec}(B/\mathfrak{p}), \bar{t}_0). \end{array}$$

Since some path connecting $\bar{\eta}$ to $\bar{t}_0$ induces the above $\gamma|_B$, it is enough to show the surjectivity of $\gamma|_B$. Therefore it is enough to show that there is a sequence of distinct $\mathbf{Q}$-valued points $\{(T-t_i)A\}_{i\in\mathbf{N}}$ of $\operatorname{Spec} A$ such that (**) remains irreducible when T is specialized to t_i. But $\operatorname{Spec} A$ is the complement of a finite set of closed points of the affine line $A^1_{\mathbf{Q}}$, thus we can find such a sequence by the following Hilbert's irreducibility theorem: (cf. [2], p. 148 Cor 3):

Theorem. *Let $f(T,X)\in\mathbf{Q}[T,X]$ be an irreducible polynomial with respect to X. Then there exists a sequence of distinct rational numbers $\{t_i\}_{i\in\mathbf{N}}$ such that for each i, $f(t_i,X)$ is an irreducible polynomial over $\mathbf{Q}$, and has the same degree as that of $f(T,X)$ with respect to X.*

This completes the proof of Theorem 2.

§ 2. Review of Deligne's Results on Vanishing Cycles

In this section we will review the theory of vanishing cycles and Deligne's results (cf. [1], [7]). Let $Y\subset\mathbf{P}^N$ be a smooth projective odd dimensional algebraic variety over an algebraically closed field k and ℓ a prime number distinct from the characteristic of k. By the above embedding, we identify a closed point p of $\mathbf{Proj}(H^0(Y,\mathcal{O}(1))$ with a hyperplane H_p in $\mathbf{P}^N$. A projective line D in $\mathbf{Proj}(H^0(Y,\mathcal{O}(1)))$ defines a Lefschetz pencil if

(1) *For some $0,\infty\in D(k)$, $H_0\cap H_\infty$ intersects properly with Y.*

(2) *The second projection $Y=\{(y,t)\in Y\times D\mid y\in H_t\}\xrightarrow{\mathrm{pr}_2} D$ is smooth outside of a finite set of closed points $S=\{s_1,\ldots,s_r\}$ in D. And for each $s\in S$, $Y_s=Y\cap H_s$ has only one isolated singularity y_s, and $\tilde{Y}-\{y_s\}_{s\in S}$ is smooth over D.*

(3) *At $y_s(s\in S)$, $\tilde{Y}\xrightarrow{\mathrm{pr}_2} D$ has only quadratic ordinary singularity.*

It is known that the set of points which correspond to Lefschetz pencils is Zariski open dense subset in the Grassmanian variety, when the characteristic of k is 0 (cf. [7]).

For a geometric generic point $\bar{\eta}$ in $D-S$, we write $Y_{\bar{\eta}}$ for $Y\cap H_{\bar{\eta}}$. This is a smooth projective variety over $\kappa(\bar{\eta})$. Let's define the space of vanishing cycles of $Y_{\bar{\eta}}$ by

$$\mathrm{Ev}(Y_{\bar{\eta}})=\mathrm{Coker}(H^n_{\text{ét}}(Y,\mathbf{Q}_\ell)\to H^n_{\text{ét}}(Y_{\bar{\eta}},\mathbf{Q}_\ell)),$$

where $n=\dim Y-1$.

When n is even we can also define the space of vanishing cycles in $H^n_{\text{ét}}(Y_{\bar{\eta}},\mathbf{Q}_\ell)$ as follows.

First we will define vanishing cycles in the local case. Let D be a strictly henselian discrete valuation ring, and s (resp. y) its closed (resp. generic) point. Let Y be a generically smooth, proper flat scheme of relative dimension n over D, and suppose that the structure morphism satisfies the above condition (2), (3) at the closed point s. Denote $\bar{D}$ by the normal closure of D in the separable

closure $\bar{y}$ of y. For any D-scheme T, we will denote the fiber product of Y and T over D by Y_T.

Now, consider the following diagrams.

$$\begin{array}{ccccc} s & \longrightarrow & \bar{D} & \longleftarrow & \bar{y} \\ \downarrow & & \downarrow & & \downarrow \\ s & \longrightarrow & D & \longleftarrow & y \end{array} \qquad \begin{array}{ccccc} Y_s & \xrightarrow{\bar{i}} & Y_{\bar{D}} & \xleftarrow{\bar{j}} & Y_{\bar{y}} \\ \downarrow & & \downarrow{\scriptstyle \varphi} & & \downarrow \\ Y_s & \xrightarrow{i} & Y & \xleftarrow{j} & Y_y \end{array}$$

Let Λ be a torsion ring prime to the residual characteristic and F an étale sheaf of Λ-module over Y. We define $\mathbf{R}\phi(F)$ the mapping cone of $\bar{i}^*\varphi^*F \to \bar{i}^*\mathbf{R}\bar{j}_*\bar{j}^*\varphi^*F$ in the category $K(Y_s \times_s y, \Lambda)$ of the complexes of étale sheaves of Λ-modules on Y_s with a compatible action of $\mathrm{Gal}(\bar{y}/y)$ (cf. [7], p. 91). When $F=\Lambda$, we can obtain the long exact sequence:

$$\cdots \to H^i_{\text{ét}}(Y_s, \Lambda) \to H^i_{\text{ét}}(Y_{\bar{y}}, \Lambda) \to H^i_{\text{ét}}(Y_s, \mathbf{R}\phi_{\bar{y}}(\Lambda)) \to \cdots$$

It is known that if i is not equal to n, $\mathbf{R}^i\phi_{\bar{y}}(\Lambda)$ is zero, and $\mathbf{R}^n\phi_{\bar{y}}(\Lambda)$ has a support on the unique isolated singularity x in Y_s with a fiber $\mathbf{R}^n\phi_{\bar{y}}(\Lambda)_x$ free Λ-module of rank 1 (cf. [7], p. 112, p. 184). By the Poincare duality theorem for $H^n_{\text{ét}}(Y_{\bar{y}}, \mathbf{Q}_\ell)$, there exists one dimensional $\mathbf{Q}_\ell$-vectorspace which is the orthogonal complement to the kernel of the following map:

$$H^n_{\text{ét}}(Y_{\bar{y}}, \mathbf{Q}_\ell) \to H^n_{\text{ét}}(Y_s, \mathbf{R}\phi_{\bar{y}}(\mathbf{Q}_\ell)) = \mathbf{R}^n\phi_{\bar{y}}(\mathbf{Q}_\ell)_x = \mathbf{Q}_\ell.$$

We call its generator δ a vanishing cycle if it satisfies $(\delta, \delta) = 2 \cdot (-1)^{n/2}$. This is determined up to sign.

In the global case, we will define the set of vanishing cycles R as a subset of $H^n_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell)$ as follows.

$R = \{\gamma_*\, \delta \mid \gamma$ varies all the paths connecting $\bar{\eta}$ and $\bar{y}$, where $\bar{y}$ varies all the generic geometric points of the henselizations D^h_s at $s \in S$, δ is a vanishing cycle in $H^n_{\text{ét}}(Y_{\bar{y}}, \mathbf{Q}_\ell)\}$, where n is the relative dimension of Y over D, and γ_*: $H^n_{\text{ét}}(Y_{\bar{y}}, \mathbf{Q}_\ell) \to H^n_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell)$ is induced from γ. The $\mathbf{Q}_\ell$-vectorspace spanned by R gives a splitting of $\mathrm{Ev}(Y_{\bar{\eta}})$ to $H^n_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell)$ (cf. [1], p. 222).

For any geometric point $\bar{\eta}$ of $D-S$, the pair $((D-S)^\sim, \varphi)$, where $(D-S)^\sim$ is a profinite étale covering and φ is an element of $\mathrm{Hom}_{(D-S)}(\bar{\eta}, (D-S)^\sim)$. is called the étale universal covering with a base point $\bar{\eta}$, if the following universal property holds: For any connected finite étale covering F with an element ψ of $\mathrm{Hom}_{(D-S)}(\bar{\eta}, F)$, there is a unique morphism $(D-S)^\sim \xrightarrow{\alpha} F$ with $\alpha \circ \varphi = \psi$. (We can define the classical universal covering with a base point in the same way. In this case we omit the finiteness condition for coverings.)

Once we fix a universal covering with a base point $((D-S)^\sim, \varphi)$ we can consider the monodromy action:

$$M: \pi_{1\,\text{alg}}(D-S, \bar{\eta}) \to \mathrm{Aut}(H^n_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell)).$$

The monodromy action of $\pi_{1\,\mathrm{alg}}(D-S,\bar{\eta})$ fixes the intersection form ψ and under this action, $\mathrm{Ev}(Y_{\bar{\eta}})$ is a stable subspace. Thus we can define the following homomorphism:

$$M\colon \pi_{1\,\mathrm{alg}}(D-S,\bar{\eta})\to \mathrm{Aut}(\mathrm{Ev}(Y_{\bar{\eta}}),\psi).$$

In the case of characteristic 0, $\pi_{1\,\mathrm{alg}}(D-S,\bar{\eta})$ is a profinite completion of a free group and the images of the generators under M are reflections with respect to some elements in R by means of Picard-Lefschetz formula.

Moreover Deligne has shown the following theorem ([1], p. 227).

(***) *In the above situation, if* ch $k\neq 2$ (*we continue to assume n is even*), *one of the following cases occurs.*

(1) Im M *is an* ℓ*-adic open subgroup in* $\mathrm{Aut}(\mathrm{Ev}(Y_{\bar{\eta}}),\psi)$.

(2) Im M *is a finite group.*

§ 3. The Proof of Theorem 1

Let $\bar{k}$ be an algebraically closed field, and $\bar{X}$ a proper smooth algebraic variety over $\bar{k}$, and ℓ a prime number distinct from the characteristic of $\bar{k}$. Then we can define the cycle map from the Chow group of codimension i to the $2i$-th étale cohomology (cf. [6], p. 143, [3], p. 275):

$$cl^i_\ell\colon CH^i(\bar{X})\otimes_{\mathbf{Q}}\mathbf{Q}_\ell\to H^{2i}_{\text{ét}}(\bar{X},\mathbf{Q}_\ell).$$

The image of the above map is called the space of (ℓ-adic) algebraic cycles of codimension i, and denoted by $C^i_\ell(\bar{X})$. When the characteristic is zero, $\dim_{\mathbf{Q}_\ell} C^i_\ell(\bar{X})=\dim_{\mathbf{Q}} C^i(\bar{X}\otimes_{\bar{k}}\mathbf{C})$, for any embedding of $\bar{k}\subset\mathbf{C}$, by the comparison theorem of étale cohomology and classical cohomology. So we can deduce $\rho^i(\bar{X}\otimes_{\bar{k}}\mathbf{C})=1$ from showing $\dim_{\mathbf{Q}_\ell} C^i_\ell(\bar{X})=1$.

For a given type $(2n;\ a_1,\ldots,a_d)$, we can find a smooth complete intersection Y over $\mathbf{Q}$ with the type $(2n+1;\ a_1,\ldots,a_{d-1})$. (In the case of $(2n;\ a_1)$, we take Y as $\mathbf{P}^{2n+1}$.) In fact, the parameter space of the complete intersections of the type $(2n+1;\ a_1,\ldots,a_{d-1})$ is $\mathbf{Q}$-unirational variety, so $\mathbf{Q}$-valued points corresponding to nonsingular varieties are Zariski dense.

We will find hypersurface sections of Y with middle Picard number 1 defined over $\mathbf{Q}$.

Step 1. There is a Lefschetz pencil $\{Y_t\}_{t\in D}$, on $\bar{Y}=Y\otimes_{\mathbf{Q}}\bar{\mathbf{Q}}$, defined over $\mathbf{Q}$ whose smooth elements are of type $(2n;\ a_1,\ldots,a_d)$.

In fact, using Veronese embedding, we may assume that the hyperplane sections are complete intersections of type $(2n;\ a_1,\ldots,a_d)$. The set of the points which correspond to Lefschetz pencils is an open dense subset in the Grassmanian. The Grassmanian is a $\mathbf{Q}$-rational variety, so there exists a Lefschetz pencil defined over $\mathbf{Q}$.

Step 2. Let D be the parameter space of the above Lefschetz pencil and $\bar{Y}$ a blown up variety of $\bar{Y}$ along $\bar{Y}_0\cap\bar{Y}_\infty$. We consider the fibration of the above

Lefschetz pencil: $\bar{Y} \xrightarrow{\varphi} \bar{D} = D \otimes_{\mathbf{Q}} \bar{\mathbf{Q}}$. Let $\bar{\eta}$ be a generic geometric point of $\bar{D}$ and S the set of the points in $\bar{D}$ where the above φ is not smooth. We will show that the image of the monodromy homomorphism;

$$M\colon \pi_{1\,\mathrm{alg}}(\bar{D}-S, \bar{\eta}) \to \mathrm{Aut}(H^{2n}_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell))$$

cannot be a finite group, when $(2n; a_1, \ldots, a_d)$ satisfies the condition (*).

Fix an embedding $\kappa(\bar{\eta}) \to \mathbf{C}$, and take a classical and étale universal covering with a base point $(((\bar{D}-S)\otimes_{\kappa(\bar{\eta})}\mathbf{C})^{\sim}, \varphi_{\mathbf{C}})$ and $((\bar{D}-S)^{\sim}, \bar{\varphi})$ which are compatible with the above embedding. Then there is a natural isomorphism compatible with the monodromy actions:

$$H^i_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell) \to H^i(Y_{\bar{\eta}} \otimes_{\bar{\eta}} \mathbf{C}, \mathbf{Q}_\ell).$$

In the above situation, the homomorphisms of the following diagram are defined and the diagram commutes, where $\bar{\eta}_{\mathbf{C}}$ is the image of $\bar{\eta}$ defined by the above inclusion $\kappa(\bar{\eta}) \to \mathbf{C}$.

$$\begin{array}{ccc}
\pi_{1\,\mathrm{alg}}(\bar{D}-S, \bar{\eta}) & \xrightarrow{M} & \mathrm{Aut}(H^{2n}_{\text{ét}}(Y_{\bar{\eta}}, \mathbf{Q}_\ell)) \\
\uparrow z & & \uparrow \\
\pi_1((\bar{D}-S)\otimes_{\bar{\eta}}\mathbf{C}, \overline{\eta_{\mathbf{C}}}) & \xrightarrow{M'} & \mathrm{Aut}(H^{2n}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C}, \mathbf{Q}_\ell)) \\
 & \searrow & \uparrow \\
 & & \mathrm{Aut}(H^{2n}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C}, \mathbf{Q}))
\end{array}$$

The image of ι is a dense subgroup with respect to the profinite topology. Put

$$\mathrm{Ev}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C})_{\mathbf{Q}} = \mathrm{Coker}(H^{2n}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C}, \mathbf{Q}) \to H^{2n}(Y\otimes_{\mathbf{Q}}\mathbf{C}, \mathbf{Q})).$$

Now suppose that the image of M is a finite group, then so is the image of the classical monodromy homomorphism M'

$$M'\colon \pi_1((\bar{D}-S)\otimes_{\bar{\eta}}\mathbf{C}, \bar{\eta}_{\mathbf{C}}) \to \mathrm{Aut}(\mathrm{Ev}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C})_{\mathbf{Q}}).$$

$\mathrm{Im}\,M'$ is also an indecomposable reflection group by the classical Picard-Lefschetz formula. If $\mathrm{Im}\,M'$ is finite, the intersection form on $\mathrm{Ev}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C})_{\mathbf{Q}}$ is positive definite if n is even, and negative definite if n is odd, by the standard argument on the reflection group over a real vectorspace (see [8] ch 5. §4. no. 8). But this implies that the Hodge type of $H^{2n}(Y_{\bar{\eta}}\otimes_{\bar{\eta}}\mathbf{C}, \mathbf{Q})$ is of pure Hodge type (n, n). Then the type of the complete intersection $Y_{\bar{\eta}}$ must be either $(2; 3)$, $(2n; 2)$ or $(2n; 2, 2)$ (cf. [7], p. 54). This contradicts to the first assumption.

Step 3. We can choose $\mathbf{Q}$-valued points of D such that the corresponding fibers Y_{t_i} are of type $(2n; a_1, \ldots, a_d)$ and $C^n_\ell(Y_{t_i}) = \mathbf{Q}_\ell$.

By the statement (***) and step 2, the image of M is ℓ-adic open in $\mathrm{Aut}(\mathrm{Ev}(Y_{\bar{\eta}}), \psi)$. Therefore by Theorem 2, there exists a sequence of distinct $\mathbf{Q}$-

valued points $\{t_i\}_{i\in\mathbf{N}}$ of $D-S$, and paths γ_i connecting $\bar{\eta}$ and $\bar{t}_i$ such that the image of the natural homomorphism as below contains an open subgroup of $\mathrm{Aut}(\mathrm{Ev}(Y_{\bar{t}_i}), \psi)$.

$$\begin{array}{ccccc}
\pi_{1\,\mathrm{alg}}(\bar{D}-S, \bar{\eta}) & \longrightarrow & \pi_{1\,\mathrm{alg}}(D-S, \bar{\eta}) & \longrightarrow & \mathrm{Aut}(\mathrm{Ev}(Y_{\bar{\eta}})) \\
 & & \uparrow & & \uparrow \\
 & & \pi_{1\,\mathrm{alg}}(D-S, \bar{t}_i) & & \mathrm{Aut}(\mathrm{Ev}(Y_{\bar{t}_i})) \\
 & & \uparrow & \nearrow & \\
 & & \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) & &
\end{array}$$

Therefore the image of the following natural homomorphism M'' also contains an open subgroup of $\mathrm{Aut}(\mathrm{Ev}(Y_{\bar{t}_i})\otimes_{\mathbf{Q}_\ell}\mathbf{Q}_\ell(n), \psi)$:

$$M'': \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{Aut}(\mathrm{Ev}(Y_{\bar{t}_i})\otimes_{\mathbf{Q}_\ell}\mathbf{Q}_\ell(n), \psi).$$

This implies that for any open subgroup H of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, there is no nonzero element in $\mathrm{Ev}(Y_{\bar{t}_i})\otimes_{\mathbf{Q}_\ell}\mathbf{Q}_\ell(n)$ fixed under the action of H. So for any open subgroup H of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$,

$$H^{2n}_{\text{ét}}(Y_{\bar{t}_i}, \mathbf{Q}_\ell(n))^H \xleftarrow{\cong} H^{2n}_{\text{ét}}(\bar{Y}, \mathbf{Q}_\ell(n))^H = \mathbf{Q}_\ell.$$

But Y_{t_i} is defined over $\mathbf{Q}$, thus all the algebraic cycles in $H^{2n}_{\text{ét}}(Y_{\bar{t}_i}, \mathbf{Q}_\ell(n))$ are defined over some algebraic number field K. Therefore all the algebraic cycles are invariant under the action of some open subgroup H of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ corresponding to K. Therefore, $C^i_\ell(Y_{\bar{t}_i}) = \mathbf{Q}_\ell$. This completes the proof of Theorem 1.

References

1. Deligne, P: La conjecture de Weil II. Publ. IHES No. 52, p. 137–428 (1980)
2. Lang, S: Diophantine Geometry. New York; Interscience 1962
3. Milne, J.S.: Etale cohomology. Princeton: Princeton University Press 1980
4. Serre, J.P.: Groupes Analytiques p-adiques. Seminaire de Bourbaki No. 270 (1963/64)
5. Shioda, T.: Algebraic cycles on a certain hypersurface, in "Algebraic Geometry, Proceedings of the Japan-France Coference, Tokyo/Kyoto, 1982", to appear in Lect. Notes in Math. No. 1016. Berlin Heidelberg New York: Springer 1983
6. SGA $4\frac{1}{2}$: Lect. Notes in Math. No. 569. Berlin Heidelberg New York: Springer 1977
7. SGA 7II: Lect. Notes in Math. No. 340. Berlin Heidelberg New York: Springer 1973
8. Bourbaki, N: Groupes et Algèbres de Lie, Chap. 4–6. Paris: Hermann 1968

Received November 8, 1983; received in final form May 21, 1984

Math. Z. 189, 297–318 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Boundaries of and Selectors for Upper Semi-Continuous Multi-Valued Maps

R.W. Hansell[1], J.E. Jayne[2], I. Labuda[3] and C.A. Rogers[2]

[1] The Department of Mathematics, University of Connecticut, Storrs, Connecticut 06268, USA
[2] The Department of Mathematics, University College London, Gower Street, London WC1E 6BT, England
[3] Institute of Mathematics, Polish Academy of Sciences, ul. Mielżyńskiego 27/29m. 5, 61–725 Poznań, Poland

1. Introduction

A set-valued function F from a topological space X into a topological space Y is said to be upper semi-continuous at a point x_0 of X if, for each open set U in Y containing $F(x_0)$, there is a neighbourhood N of x_0 with $F(N) \subset U$. The function F is said to be upper semi-continuous on X, if F is upper semi-continuous at each point of X, or, equivalently, if $\{x: F(x) \subset U\}$ is an open set in X, whenever U is an open set in Y.

Choquet [5, p. 70] stated, without proof, that, if X and Y are metric spaces and F is upper semi-continuous at a point x_0 of X, then it is possible to choose a compact subset $K(x_0)$ of $F(x_0)$, with the property that, for each open set G containing $K(x_0)$, there is a neighbourhood N of x_0 with

$$F(N) \subset G \cup F(x_0).$$

Dolecki ([9], [10]) introduced the set

$$H(x_0) = \bigcap_{V \in \mathscr{V}} \mathrm{cl}\,(F(V) \setminus F(x_0)), \tag{1}$$

where $\mathscr{V}$ is the set of all neighbourhoods of x_0, and where cl is used to denote the closure operation. Suppose F is upper semi-continuous at a point x_0 having a countable neighbourhood base: Dolecki [9] showed that $H(x_0)$ is compact in the case when Y is a metric space; Dolecki and Rolewicz [11, Lemma 2] showed that $H(x_0)$ is a subset of $F(x_0)$ when Y is a first countable Hausdorff space; and Dolecki and Lechicki [13, Theorem 4.1] showed that $H(x_0)$ has Choquet's property in the case when Y is metric.

Jayne and Rogers ([18] and [19]) have introduced and studied the set

$$K(x_0) = \mathrm{proj}_Y\{[\{x_0\} \times F(x_0)] \cap \mathrm{cl}\, E(x_0)\}, \tag{2}$$

where

$$E(x_0) = [(X \setminus \{x_0\}) \times (Y \setminus F(x_0))] \cap T, \tag{3}$$

and T is the graph

$$T = \bigcup_{x \in X} [\{x\} \times F(x)] \tag{4}$$

of F. As the set $H(x_0)$ can be expressed in the alternative form

$$H(x_0) = \text{proj}_Y \{[\{x_0\} \times Y] \cap \text{cl}\, E(x_0)\}, \tag{5}$$

with $E(x_0)$ and T defined by (3) and (4), it is clear that

$$K(x_0) = H(x_0) \cap F(x_0).$$

Thus $K(x_0)$ and $H(x_0)$ coincide whenever $H(x_0) \subset F(x_0)$. As we have remarked, this holds, in particular, when F is upper semi-continuous at x_0, Y is metric and the neighbourhoods of x_0 have a countable base. We regard $K(x_0)$ as the boundary set of F at x_0.

Our main objective, in this paper, is to obtain conditions that ensure that this boundary set $K(x_0)$ is compact. Previous work by Jayne and Rogers [18, 19] obtains powerful selection results for upper semi-continuous set-valued maps in cases when this boundary is compact. Our results do lead to two improved selection results; these are discussed towards the end of the introduction. We believe that the results of the present paper will lead, sooner or later, to further significant advances in this subject.

A space Y is said to be Dieudonné complete if it is homeomorphic to a closed subset of some cartesian product $\prod_{\alpha \in A} M_\alpha$ of metric spaces M_α, $\alpha \in A$. The space Y is Dieudonné complete, if, and only if, it admits a complete uniformity that is compatible with its topology, see [16, p. 568]. Dolecki and Lechicki [13, Th. 7.6] show that, if x_0 is a point with a countable base for its neigbourhoods in a Hausdorff space X, if F is a set-valued map from X to a Dieudonné complete space Y, and if F is upper semi-continuous at x_0, then $H(x_0)$ is compact. They give an example, see § 5 below, satisfying these conditions, where $H(x_0)$ is not contained in $F(x_0)$ and $K(x_0)$ is not compact. However they do show that, if in addition, $F(x_0)$ is closed, then $K(x_0)$ coincides with $H(x_0)$ and is a compact subset of $F(x_0)$.

Fremlin (see Pryce [22] or Floret [17, p. 30]) says that a Hausdorff space Y is angelic, if each relatively countably compact set A in Y is relatively compact, and each point of the closure of such a set A is the limit of a convergent sequence of points of A. It is easy to verify (see [17] or [22]) that in such an angelic space the families of relatively compact, of relatively countably compact and of relatively sequentially compact sets coincide, as do the families of compact, of countably compact and of sequentially compact sets. Many classes of locally convex topological vector spaces are angelic, see, for example [17, p. 39]. In particular, all metrizable locally convex spaces are angelic in their weak topology.

Our first main theorem shows that, *if x_0 is a point with a countable base for its neighbourhoods in a Hausdorff space X, if F is a set-valued map from X to an angelic space Y, and, if F is upper semi-continuous at x_0, then $K(x_0)$ coincides with $H(x_0)$ and is compact.*

In a second theorem we relax the condition that X be first-countable at x_0, but impose a weak condition on the value of F at x_0. We need to introduce some extra concepts.

The point x is said to be a cluster point of a sequence $(x_n)_{n=1}^{\infty}$ if each neighbourhood of x contains x_n for some arbitrarily large integers n. The point x_0 is said to be a q-point of the space X, if there is a sequence $(N_i)_{i=1}^{\infty}$ of open neighbourhoods of x_0, with the property that, each sequence $(x_i)_{i=1}^{\infty}$, with $x_i \in N_i$, for $i \geqq 1$, has a cluster point in X, see Michael [21]. Note in particular, that each point of a first-countable space is a q-point, and that each point of a locally compact space is also a q-point.

If Y is a topological space, the $\mathscr{G}_\delta$-topology on Y is the new topology generated by taking all the original $\mathscr{G}_\delta$-sets in Y as a base for the new topology. A Souslin-$\mathscr{F}$ set is a set S with a representation

$$S = \bigcup_{i_1, i_2, \ldots} \bigcap_{n=1}^{\infty} F(i_1, i_2, \ldots, i_n),$$

the sets $F(i_1, i_2, \ldots, i_n)$ being all closed, and the union being taken over all sequences $(i_n)_{n=1}^{\infty}$ of positive integers. We show that such a Souslin-$\mathscr{F}$ set S is necessarily closed in the $\mathscr{G}_\delta$-topology. Consider any point y not in S. Then, corresponding to each infinite sequence $\mathbf{i} = i_1, i_2, \ldots$, we can choose an integer $n(\mathbf{i}, y)$ with

$$y \notin \bigcap_{n=1}^{n(\mathbf{i}, y)} F(i_1, i_2, \ldots, i_n).$$

As $\mathbf{i}$ varies, there is only a countable number of distinct intersections of this form. The union of these intersections is an $\mathscr{F}_\sigma$-set containing S but not containing y. Thus S is $\mathscr{G}_\delta$-closed. A set is said to be a Borelian-$\mathscr{F}$ set if it belongs to the minimal family, containing the closed sets, and closed under countable unions and countable intersections; such sets are always Souslin-$\mathscr{F}$ sets, and so are $\mathscr{G}_\delta$-closed.

A space Y is said to be realcompact if it is homeomorphic to a closed subset of some cartesian product $\prod_{\alpha \in A} \mathbb{R}_\alpha$ of real lines. Clearly, a realcompact space is Dieudonné complete. In fact, many of the best known Dieudonné complete spaces are realcompact. Corson [6, Lemma 8] shows that a Banach space, with its weak topology, is Dieudonné complete, if, and only if, it is realcompact. He shows that the Hewitt realcompactification of a Banach space Y with its weak topology is the linear subspace of Y^{**} with its weak-star topology, obtained by taking all y^{**} in Y^{**} with the property that, for each countable set C in Y^*, there is a y in Y that coincides with y^{**} on C. As Wheeler has remarked, see [27, proof of Th. 5.2], it is easy to verify that the linear subspace of Y^{**}, defined in this way, is the $\mathscr{G}_\delta$-closure of Y in Y^{**} taken with its weak-star topology. Thus Y can only be realcompact if it is $\mathscr{G}_\delta$-closed in Y^{**} with its weak-star topology.

We remark that Edgar [15] has shown that, when various types of Banach spaces are taken with their weak topologies, they become realcompact. Further, Arhangel'skii [2] and Chigogidze [4] have characterized those completely

regular Hausdorff spaces Z, for which the space $C_p(Z)$, of all continuous real-valued functions on Z, taken with the topology of pointwise convergence, is realcompact.

We prove that, *if x_0 is a q-point of a regular Hausdorff space X, and if F is an upper semi-continuous set-valued map of X into a Dieudonné complete space Y, then $H(x_0)$ is compact, and if $F(x_0)$ is closed in the $\mathcal{G}_\delta$-topology on Y, then $K(x_0)$ coincides with $H(x_0)$.*

In §4 we explain, in some detail, the extent to which our results apply to various classes of locally convex topological vector spaces. In particular, we study the subspaces of the space of real-valued functions, with the topology of point-wise convergence, on various topological spaces, obtained by taking the Baire measurable functions, the Baire measurable functions of class 1 and the continuous functions. We also consider Banach spaces with their weak topologies and dual Banach spaces with their weak-star topologies.

In §5 we give some very simple examples of upper semi-continuous set-valued functions, designed to show that some of the known results can not be improved in certain directions.

In §6 we return to the selection problem discussed by Jayne and Rogers in [18, 19] and obtain some extensions of their results. *If F is an upper semi-continuous set-valued map, with arbitrary non-empty values, from a metric space X into the Banach space $c_0(\Gamma)$ with its weak topology, then F has a selector f of the second Borel class using the norm topology of $c_0(\Gamma)$. It is possible to choose f, an $\mathcal{F}_\sigma$-set X_1 in X and its complementary $\mathcal{G}_\delta$-set $X_2 = X \setminus X_1$, so that the restrictions of f to X_1 and X_2 are of the first Borel class, using the weak topology. Further, if F takes only weakly compact values, then F has a selector of the first Borel class, using the weak topology.*

A Hausdorff space Y is said to be K-analytic if it can be expressed as the continuous image of some $\mathcal{F}_{\sigma\delta}$-set in some compact Hausdorff space. The class of all Banach spaces that are K-analytic in their weak topologies includes the class of all weakly compactly generated Banach spaces, which in its turn includes all separable Banach spaces and all reflexive Banach spaces, see, for example, [23, pp. 50–53]. Recall that the Baire sets of a topological space are the sets of the minimal family of sets that is closed under the operations of countable union and complementation and contains all the sets on which some continuous real-valued function takes the value zero. Note that Edgar [14, Th. 2.3] gives the remarkable result that the Baire sets for a Banach space (indeed for a locally convex space) Y with its weak topology are the minimal family of sets that is closed under the operations of countable union and complementation and contains all the sets of the form $\{y: g(y) > r\}$ where r is real and g is a continuous *linear* real-valued functional on Y. A function f from a topological space X to a topological space Y is said to be Baire measurable if the inverse image in X of any Baire set in Y is a Baire set in X. Note that if X is metric, the Baire sets in X are just the Borel sets in X. We prove that, *if F is an upper semi-continuous set-valued map, with arbitrary non-empty values, from a metric space X to a Banach space Y, with its weak topology, and if Y is K-analytic in its weak topology, then F has a Baire measurable selector, using the weak topology on Y.*

2. The First Countable Case

In this section our aim is to prove the first result italicized in the Introduction. Throughout this section, we suppose that x_0 is a point with a countable base for its neighbourhoods in a Hausdorff space X, that F is a set-valued map from X to a Hausdorff space Y and that F is upper semi-continuous at x_0. We work with the sets $H(x_0)$ and $K(x_0)$ defined in the introduction and also with a set $L(x_0)$ defined in a more complicated way. Let $L(x_0)$ be the set of all points of $F(x_0)$ that can be obtained as cluster points of sequences (y_n) in Y, with

$$y_n \in F(x_n) \setminus F(x_0), \qquad n \geqq 1,$$

and with (x_n) a sequence converging to x_0 in X. Comparing the definitions we see that

$$L(x_0) \subset K(x_0) \subset H(x_0).$$

We show, in Lemma 5 below, that, if $L(x_0)$ is relatively compact, or if Y is regular, then $H(x_0)$ is the closure of $L(x_0)$. We show, in Theorem 1 below, that, if Y is angelic, then the sets $L(x_0)$, $K(x_0)$ and $H(x_0)$ coincide and are compact.

We first prove two lemmas that take us towards a type of "local boundedness" for $F(x) \setminus F(x_0)$ and a type of "boundedness" for $H(x_0)$.

Lemma 1. *If $\{S_\alpha : \alpha \in A\}$ is any discrete family of sets in Y, then there is a neighbourhood N of x_0 such that*

$$[F(N) \setminus F(x_0)] \cap S_\alpha \neq \varnothing,$$

for at most a finite number of α in A. If the sets S_α, $\alpha \in A$, of the discrete family are open, then

$$H(x_0) \cap S_\alpha \neq \varnothing,$$

for at most a finite number of α in A.

Lemma 2. *If h is any continuous real-valued function defined on Y, then h is bounded on $H(x_0)$ and there is a neighbourhood N of x_0 with h bounded on $F(N) \setminus F(x_0)$.*

Proof of Lemma 1. Let $N_1, N_2, \ldots$ be a decreasing sequence of open neighbourhoods of x_0 forming a base for the neighbourhoods of x_0. Suppose that, for each $i \geqq 1$, we have

$$[F(N_i) \setminus F(x_0)] \cap S_\alpha \neq \varnothing,$$

for infinitely many α in A. Then we can choose a sequence $\alpha(1), \alpha(2), \ldots$ of distinct elements of A with

$$[F(N_i) \setminus F(x_0)] \cap S_{\alpha(i)} \neq \varnothing,$$

$i = 1, 2, \ldots$. We can then choose pairs of points x_i, y_i, $i = 1, 2, \ldots$ so that

$$x_i \in N_i \quad \text{and} \quad y_i \in [F(x_i) \setminus F(x_0)] \cap S_{\alpha(i)},$$

for $i \geqq 1$. As the sets $S_{\alpha(i)}$, $i \geqq 1$, form a discrete family, the set

$$J = \{y_1, y_2, \ldots\}$$

is closed in Y. As F is upper semi-continuous at x_0 the set

$$F^{-1}(J) = \{x \in X : F(x) \cap J \neq \varnothing\}$$

is closed at x_0. As

$$y_i \in F(x_i), \quad y_i \notin F(x_0),$$

we have

$$x_i \in F^{-1}(J), \quad x_0 \notin F^{-1}(J).$$

This is impossible as the sequence $x_1, x_2, \ldots$ converges to x_0, while $F^{-1}(J)$ is closed at x_0. It follows that, for some $i \geqq 1$, we have

$$[F(N_i) \setminus F(x_0)] \cap S_\alpha \neq \varnothing$$

for at most a finite number of α in A.

Now suppose that S_α is open for each α in A. Suppose that

$$H(x_0) \cap S_\alpha \neq \varnothing,$$

for all α in an infinite subset A' of A. By the definition of $H(x_0)$ it follows that

$$[F(N_i) \setminus F(x_0)] \cap S_\alpha \neq \varnothing,$$

for all $i \geqq 1$ and for all α in the infinite set A'. This is impossible by the part of this lemma that we have already established. It follows that

$$H(x_0) \cap S_\alpha \neq \varnothing,$$

for at most a finite number of α in A.

Proof of Lemma 2. Define open intervals I_n and J_n in $\mathbb{R}$, for all n in the set $\mathbb{Z}$ of all integers, by taking

$$I_n = (n - \tfrac{3}{5}, n + \tfrac{1}{5}),$$

and

$$J_n = (n - \tfrac{1}{5}, n + \tfrac{3}{5}).$$

Then each family $\{I_n : n \in \mathbb{Z}\}$ and $\{J_n : n \in \mathbb{Z}\}$ is a discrete family of open sets in $\mathbb{R}$, and taken together, these two families cover $\mathbb{R}$. As h is continuous, the families $\{h^{-1}(I_n) : n \in \mathbb{Z}\}$ and $\{h^{-1}(J_n) : n \in \mathbb{Z}\}$ are open discrete families in Y that together cover Y. By Lemma 1, we can choose a neighborhood N of x_0 such that $F(N) \setminus F(x_0)$ meets a finite number only of the sets of these two families. This ensures that h is bounded on $F(N) \setminus F(x_0)$. Similarly h is bounded on $H(x_0)$.

We now investigate the properties associated with the set $L(x_0)$. Recall that a set A is said to be *relatively countably compact* in a space Y, if every sequence of points of A has a cluster point in Y.

Lemma 3. *If $(x_n), (y_n)$ are sequences with*

$$y_n \in F(x_n) \setminus F(x_0), \qquad n \geqq 1,$$

and with (x_n) converging to x_0, then (y_n) has a cluster point in $L(x_0)$. Further, $\{y_n\}$ is a relatively countably compact subset of Y.

Lemma 4. *If G is any open set in Y that contains $L(x_0)$, then there is a neighbourhood N of x_0 with*

$$F(N) \setminus F(x_0) \subset G.$$

Lemma 5. *If $L(x_0)$ is relatively compact, or if Y is regular, then $H(x_0)$ is the closure of $L(x_0)$.*

Proof of Lemma 3. Take J to be the closure in Y of the set of points $\{y_n\}$ of the sequence (y_n). Suppose $F(x_0)$ does not meet J. As F is upper semi-continuous at x_0, the closure of $F^{-1}(J)$ does not contain x_0. As this set contains the sequence (x_n) converging to x_0, it must also contain x_0. Hence, $F(x_0)$ meets J. As $F(x_0)$ contains none of the points of the sequence (y_n), it follows that $F(x_0)$ contains at least one cluster point of this sequence. By the definition of $L(x_0)$, any such cluster point of (y_n) in $F(x_0)$ belongs to $L(x_0)$. Hence (y_n) has at least one cluster point in $L(x_0)$. Applying this same argument to any subsequence of (y_n) we find that any such subsequence has a cluster point in $L(x_0)$. Hence $\{y_n\}$ is a relatively countably compact subset of Y.

Proof of Lemma 4. Let $N_1, N_2, \ldots$ be a decreasing sequence of open neighbourhoods of x_0 forming a base for the neighbourhoods of x_0. Suppose that for all $i \geqq 1$, the set $F(N_i) \setminus F(x_0)$ is not contained in G. Then, for each $i \geqq 1$, we can choose x_i, y_i so that

$$x_i \in N_i \quad \text{and} \quad y_i \in [F(x_i) \setminus F(x_0)] \cap [Y \setminus G].$$

Then $x_i \to x_0$ as $i \to \infty$. By Lemma 3, the sequence (y_n) has a cluster point in $L(x_0)$. As $Y \setminus G$ is a closed set containing $\{y_n\}$, this cluster point is in $Y \setminus G$ which is contained in $Y \setminus L(x_0)$. This contradiction shows that

$$F(N_i) \setminus F(x_0) \subset G$$

for all sufficiently large i.

Proof of Lemma 5. Consider any point e not in the closure $\operatorname{cl} L(x_0)$ of $L(x_0)$. Both in the case when $\operatorname{cl} L(x_0)$ is compact in the Hausdorff space Y, and also in the case when Y is a regular Hausdorff space, we can choose open sets G and G' with

$$\operatorname{cl} L(x_0) \subset G, \qquad G \cap G' = \varnothing, \qquad e \in G'.$$

By Lemma 4, we can choose a neighbourhood N of x_0 with

$$F(N) \setminus F(x_0) \subset G.$$

Now $N \times G'$ is a neighbourhood of (x_0, e) in $X \times Y$ that does not meet the set

$$E(x_0) = \bigcup_{\xi \in X} \{\xi\} \times [F(\xi) \setminus F(x_0)].$$

Hence, by the alternative definition of $H(x_0)$ (see Eq. (5)), the point e does not belong to $H(x_0)$. We conclude that $H(x_0)$ is contained in the closure of $L(x_0)$. The definition ensures that $H(x_0)$ is closed. Thus $H(x_0)$ is the closure of $L(x_0)$.

To take our investigation further we need to introduce extra conditions on our space Y. We shall need the following conditions in turn:

(a) If E is a countable subset of Y and E is relatively countably compact, then each point of the closure of E is the limit of some convergent sequence in E.

(b) If E is a countable subset of Y and E is relatively countably compact, then the closure of E is countably compact, *i.e.* each countable subset of $\mathrm{cl}\, E$ has a cluster point in $\mathrm{cl}\, E$.

(c) If E is a sequentially compact set in Y (*i.e.*, each sequence of points of E has a subsequence converging to some point of E), then E is compact in Y.

Note that an angelic space satisfies Conditions (a), (b) and (c).

Lemma 6. *If Y satisfies Condition* (a) *and* (x_n), (y_n) *are sequences with*

$$y_n \in F(x_n) \setminus F(x_0), \qquad n \geqq 1,$$

and with (x_n) *converging to* x_0, *then each cluster point of* (y_n) *belongs to* $L(x_0)$.

Lemma 7. *If Y satisfies Condition* (a), *then each cluster point of each countable subset of* $L(x_0)$ *belongs to* $L(x_0)$.

Lemma 8. *If Y satisfies Conditions* (a) *and* (b), *then* $L(x_0)$ *is sequentially compact and so is countably compact.*

Theorem 1. *If Y satisfies Conditions* (a), (b) *and* (c), *then the sets* $L(x_0)$, $K(x_0)$ *and* $H(x_0)$ *coincide and are compact.*

Proof of Lemma 6. By Lemma 3, the sequence (y_n) is relatively countably compact. Consider any cluster point y of (y_n). After removing any occurrences of the point y from the sequence (y_n) and removing the corresponding x_n from the sequence (x_n) we may suppose that

$$y \notin \{y_1, y_2, \ldots\}.$$

By Condition (a) there will be a subsequence, $(y_{n(r)})$ say, of (y_n) converging to y. Applying Lemma 3 to the sequences $(x_{n(r)})$, $(y_{n(r)})$ we find that $(y_{n(r)})$ has a cluster point in $L(x_0)$. As $(y_{n(r)})$ converges to y, the cluster point must be y. Hence y belongs to $L(x_0)$, as required.

Proof of Lemma 7. Let $(l^{(i)})$ be a sequence of points of $L(x_0)$. Let $N_1, N_2, \ldots$ be a decreasing sequence of open sets forming a base for the neighbourhoods of x_0. By the definition of $L(x_0)$, for each $i \geqq 1$, $j \geqq 1$, we can choose a pair of points $x_j^{(i)}$, $y_j^{(i)}$, with

$$y_j^{(i)} \in F(x_j^{(i)}) \setminus F(x_0),$$
$$x_j^{(i)} \in N_{i+j},$$

and with $l^{(i)}$ a cluster point of $(y_j^{(i)})$, for each $i \geqq 1$.

Let (ξ_k), (η_k) be the sequences obtained by reordering the double sequences $(x_j^{(i)})$, $(y_j^{(i)})$ as a single sequence, taking together the elements for which $i+j$ has a common value, and ordering these groups in increasing order of $i+j$.

Then $\xi_k \to x_0$ as $k \to \infty$ and each point $l^{(i)}$ is a cluster point of (η_k). Thus any cluster point of $(l^{(i)})$, say l, is also a cluster point of (η_k). By Lemma 6, it follows that l belongs to $L(x_0)$. Thus each cluster point of each countable subset of $L(x_0)$ belongs to $L(x_0)$.

Proof of Lemma 8. Let $(l^{(i)})$ be any sequence of points of $L(x_0)$. Construct the sequences (ξ_k) and (η_k) as in the proof of Lemma 7. By Lemma 3, $\{\eta_k\}$ is a relatively countably compact subset of Y. As in the proof of Lemma 7, each point $l^{(i)}$, $i \geqq 1$, belongs to the closure of the set $\{\eta_k\}$. By Condition (b), the closure of the set $\{\eta_k\}$ is countably compact. As the sequence $(l^{(i)})$ is contained in this closure of $\{\eta_k\}$, the set $\{l^{(i)}\}$ is relatively countably compact. By Lemma 7, each cluster point of $(l^{(i)})$ is contained in $L(x_0)$. Thus the closure of $\{l^{(i)}\}$ is a countably compact set contained in $L(x_0)$. Hence, $L(x_0)$ is itself countably compact.

We have seen that each sequence $(l^{(i)})$ of distinct points of $L(x_0)$ has a cluster point, say l, in $L(x_0)$, and that the set $\{l^{(i)}\}$ is relatively countably compact. As l can coincide with at most one of the points $l^{(i)}$, $i \geqq 1$, we may suppose this point removed from the sequence without changing its other properties. By Condition (a), the point l is necessarily the limit point of some convergent subsequent of $(l^{(i)})$. This shows that $L(x_0)$ is sequentially compact.

Proof of Theorem 1. By Lemma 8, the set $L(x_0)$ is sequentially compact. By Condition (c), this set $L(x_0)$ is compact. Now Lemma 5 applies and

$$H(x_0) = \operatorname{cl} L(x_0) = L(x_0).$$

Hence $L(x_0)$, $K(x_0)$ and $H(x_0)$ coincide and are compact.

3. The q-Point Case

In this section we prove the second italicized statement in the Introduction. We first prove some lemmas that will make the final proof easy. Throughout this section, x_0 will be a q-point in a regular Hausdorff space X, and F will be an upper semi-continuous set-valued map from X to a Hausdorff space Y.

Lemma 9. *Lemma 1 of § 2 holds under the conditions of this section.*

Lemma 10. *Lemma 2 of § 2 holds under the conditions of this section.*

Lemma 11. *If $F(x_0)$ is closed in the $\mathscr{G}_\delta$-topology on a regular Hausdorff space Y, then $K(x_0)$ coincides with $H(x_0)$.*

The next lemma is well-known, *cf.* [17, p. 51], where the realcompact case is proved. The idea of the proof also gives the Dieudonné complete case.

Lemma 12. *Let H be a closed set in a Dieudonné complete space Y, and suppose that every real-valued continuous function defined on Y is bounded on H. Then H is compact.*

Theorem 2. *Let x_0 be a q-point of a regular Hausdorff space X, and let F be an upper semi-continuous set-valued map of X into a Dieudonné complete space Y. Then $H(x_0)$ is compact. If $F(x_0)$ is closed in the $\mathcal{G}_\delta$-topology on Y, then $K(x_0)$ coincides with $H(x_0)$.*

Proof of Lemma 9. As x_0 is a q-point of X, we can choose a decreasing sequence $(N_i)_{i=1}^{\infty}$ of open neighbourhoods of x_0, with the property that, if $x_i \in N_i$, for $i \geqq 1$, then the sequence $(x_1, x_2, \ldots)$ has a cluster point in X.

Now suppose that, for each open neighbourhood N of x_0,

$$[F(N) \setminus F(x_0)] \cap S_\alpha \neq \varnothing,$$

for infinitely many α in A, and seek a contradiction. We make an inductive choice of a decreasing sequence (L_i) of closed neighbourhoods of x_0, with

$$L_i \subset N_i, \quad i \geqq 1,$$

and a sequence (ξ_i, η_i) of points of $X \times Y$, and a sequence $\alpha(i)$ of distinct elements of A with

$$\xi_i \in L_i \quad \text{and} \quad \eta_i \in [F(\xi_i) \setminus F(x_0)] \cap S_{\alpha(i)},$$

for $i \geqq 1$, and with

$$\xi_j \notin L_i, \quad \text{for } 1 \leqq j < i.$$

As X is regular we can take L_1 to be a closed neighbourhood of x_0 that is contained in N_1. As

$$[F(L_1) \setminus F(x_0)] \cap S_\alpha \neq \varnothing,$$

for infinitely many α in A, we can choose $\alpha(1)$ to be such an α in A; then we can choose (ξ_1, η_1) with

$$\xi_1 \in L_1 \quad \text{and} \quad \eta_1 \in [F(\xi_1) \setminus F(x_0)] \cap S_{\alpha(1)}.$$

Now suppose that for some $i \geqq 1$, the required sets $L_1, L_2, \ldots, L_i$, the required points (ξ_1, η_1), (ξ_2, η_2), ..., (ξ_i, η_i), and the required indices $\alpha(1)$, $\alpha(2)$, ..., $\alpha(i)$ have been chosen. Then, for $1 \leqq j \leqq i$, we have $\eta_j \notin F(x_0)$ and $F^{-1}(\eta_j)$ is a closed set in X that contains ξ_j but does not contain x_0. As X is regular, we can now choose L_{i+1} to be a closed neighbourhood of x_0 that is contained in

$$(N_{i+1} \cap L_i) \setminus \bigcup_{j=1}^{i} F^{-1}(\eta_j).$$

This ensures that

$$\xi_j \notin L_{i+1}, \quad \text{for } 1 \leqq j < i+1.$$

As

$$[F(L_{i+1}) \setminus F(x_0)] \cap S_\alpha \neq \varnothing,$$

for infinitely many α in A, we can choose $\alpha(i+1)$ to be such an α in A, distinct from $\alpha(1)$, $\alpha(2)$, ..., $\alpha(i)$; and then we can choose ξ_{i+1}, η_{i+1} with

$$\xi_{i+1} \in L_{i+1} \quad \text{and} \quad \eta_{i+1} \in [F(\xi_{i+1}) \setminus F(x_0)] \cap S_{\alpha(i+1)}.$$

This completes the inductive construction. Note that the conditions

$$\xi_i \in L_i \quad \text{and} \quad \xi_j \notin L_i \quad \text{for } 1 \leqq j < i$$

ensure that the points ξ_i are all distinct.

As $\eta_i \in S_{\alpha(i)}$, $i \geqq 1$, and the family $\{S_\alpha\}_{\alpha \in A}$ is discrete in Y, the set

$$H = \{\eta_1, \eta_2, \ldots\}$$

is closed in Y. As H does not meet $F(x_0)$, and as F is upper semicontinuous, the set $F^{-1}(H)$ is a closed set in X that contains $\{\xi_1, \xi_2, \ldots\}$ but does not contain x_0.

As $\xi_i \in L_i \subset N_i$, for $i \geqq 1$, the choice of the N_i ensures that the sequence $(\xi_1, \xi_2, \ldots)$ of distinct points has some cluster point, say ξ_0. As L_i is closed and

$$\xi_j \in L_j \subset L_i, \quad \text{for } j \geqq i,$$

the cluster point ξ_0 belongs to L_i, for $i \geqq 1$. Hence

$$\xi_0 \in L_{i+1} \subset X \setminus F^{-1}(\eta_i),$$

for $i \geqq 1$, so that

$$\xi_0 \notin \bigcup_{i=1}^{\infty} F^{-1}(\eta_i) = F^{-1}(H).$$

As $\xi_1, \xi_2, \ldots$ all belong to the closed set $F^{-1}(H)$ this is impossible. Thus the first part of the lemma holds. Note that here we make essential use of the upper semi-continuity of F at the point ξ_0.

The second part follows immediately from the first part, by the method used in the proof of Lemma 1.

Proof of Lemma 10. Lemma 10 follows immediately from Lemma 9 by the method used to deduce Lemma 2 from Lemma 1.

Proof of Lemma 11. Suppose that η is a point in $H(x_0)$ that is not in $K(x_0)$. We seek a contradiction. We have $(x_0, \eta) \in \operatorname{cl} E(x_0)$. As $\eta \notin K(x_0)$, we conclude that $\eta \notin F(x_0)$.

As $F(x_0)$ is closed in the $\mathscr{G}_\delta$-topology of Y, the point η belongs to a $\mathscr{G}_\delta$-set in Y contained in $Y \setminus F(x_0)$. So we can choose a decreasing sequence $G_1, G_2, \ldots$ of open sets in Y, with

$$\eta \in \bigcap_{n=1}^{\infty} G_n \quad \text{and} \quad F(x_0) \cap \bigcap_{n=1}^{\infty} G_n = \varnothing.$$

As Y is a regular Hausdorff space, we can choose a decreasing sequence $H_1, H_2, \ldots$ of closed neighbourhoods of η with

$$H_n \subset G_n, \quad n \geqq 1.$$

Then η belongs to the closed set

$$H_0 = \bigcap_{n=1}^{\infty} H_n,$$

and $F(x_0) \cap H_0 = \varnothing$. By the upper semi-continuity of F, it follows that $F^{-1}(H_0)$ is a closed set in X that does not contain x_0.

The rest of the proof obtains a contradiction using a variant of the method used to prove Lemma 9.

As x_0 is a q-point, there is a decreasing sequence (N_i) of open neighbourhoods of x_0, with the property that, if $x_i \in N_i$, for $i \geqq 1$, then the sequence $(x_1, x_2, \ldots)$ has a cluster point in X. We choose a sequence (L_i) of closed neighbourhoods of x_0, with

$$L_i \subset N_i, \quad i \geqq 1,$$

and a sequence (ξ_i, η_i) of points of $E(x_0)$ by an inductive process. As X is regular, we can choose L_1 to be a closed neighbourhood of x_0 contained in the neighbourhood

$$N_1 \setminus F^{-1}(H_0)$$

of x_0. As (x_0, η) belongs to $\operatorname{cl} E(x_0)$, we can choose (ξ_1, η_1) with

$$\xi_1 \in L_1 \quad \text{and} \quad \eta_1 \in [F(\xi_1) \setminus F(x_0)] \cap H_1.$$

Suppose that, for some $i \geqq 1$, the required sets $L_1, L_2, \ldots, L_i$ and points (ξ_1, η_1), (ξ_1, η_2), $\ldots$, (ξ_i, η_i) have been chosen with

$$\xi_j \in L_j \quad \text{and} \quad \eta_j \in [F(\xi_j) \setminus F(x_0)] \cap H_j,$$

for $1 \leqq j \leqq i$. Then $\eta_j \notin F(x_0)$ and $F^{-1}(\eta_j)$ is a closed set that does not contain x_0, for $1 \leqq j \leqq i$. Take L_{i+1} to be any closed neighbourhood of x_0 contained in the neighbourhood

$$(N_{i+1} \cap L_i) \setminus \bigcup_{j=1}^{i} F^{-1}(\eta_j).$$

Then take (ξ_{i+1}, η_{i+1}) with

$$\xi_{i+1} \in L_{i+1} \quad \text{and} \quad \eta_{i+1} \in [F(\xi_{i+1}) \setminus F(x_0)] \cap H_{i+1};$$

this is possible as (x_0, η) is in the closure of $E(x_0)$. The induction is complete. Note that $\xi_j \in F^{-1}(\eta_j)$ but $\xi_i \notin F^{-1}(\eta_j)$, when $1 \leqq j < i$, so that all the points $\xi_1, \xi_2, \ldots$ are distinct.

Now any cluster point of the sequence

$$(\eta_1, \eta_2, \ldots)$$

lies in H_i, for each $i \geqq 1$, and so lies in H_0. Thus

$$H = H_0 \cup \{\eta_1, \eta_2, \ldots\}$$

is a closed set in Y that does not meet $F(x_0)$. By the upper semi-continuity of F, it follows that $F^{-1}(H)$ is a closed set in X.

As $\xi_i \in L_i \subset N_i$, for $i \geqq 1$, the choice of the N_i ensures that the sequence $(\xi_1, \xi_2, \ldots)$ of distinct points has some cluster point, say ξ_0. As $\xi_j \in L_i$, for $j \geqq i$, and L_i is closed, we have $\xi_0 \in L_i$ for $i \geqq 1$. Hence

$$\xi_0 \in L_{i+1} \subset X \setminus F^{-1}(\eta_i), \quad \text{for } i \geqq 1,$$

and also

$$\xi_0 \in X \setminus F^{-1}(H_0),$$

by the choice of L_1. Hence

$$\xi_0 \notin F^{-1}(H_0) \cup \bigcup_{i=1}^{\infty} F^{-1}(\eta_i) = F^{-1}(H).$$

As $\{\xi_1, \xi_2, \ldots\}$ is contained in the closed set $F^{-1}(H)$, this contradicts the choice of ξ_0. Note that we again make essential use of the upper semi-continuity of F at ξ_0.

Proof of Lemma 12. As the space Y is Dieudonné complete, we may suppose that Y is a closed subset of the Cartesian product

$$M = \prod_{\alpha \in A} M(\alpha),$$

of a family $\{M(\alpha): \alpha \in A\}$ of metric spaces.

For each α in A, let $\pi(\alpha)$ denote the projection map from M to $M(\alpha)$. Write

$$P(\alpha) = \mathrm{cl}\, \pi(\alpha)(H).$$

Consider any real-valued continuous function g defined initially on the closed subset $P(\alpha)$ of $M(\alpha)$. By the Tietze extension theorem, see, for example [16, p. 97], we may suppose the function g to have been extended, without change of notation, to a continuous real-valued function defined on $M(\alpha)$. Now $g \circ \pi(\alpha)$ is a continuous real-valued function defined on M. By our assumption about the subset H of Y, which is contained in M, this function $g \circ \pi(\alpha)$ is bounded on H. Hence g is bounded on

$$\pi(\alpha)(H).$$

As g is continuous on $M(\alpha)$, it is also bounded on

$$P(\alpha) = \mathrm{cl}\, \pi(\alpha)(H).$$

Thus each continuous real-valued function defined on this set $P(\alpha)$ is bounded on $P(\alpha)$. This is what is meant by saying that $P(\alpha)$ is pseudo-compact. As $P(\alpha)$ is also a metric space, it is necessarily compact, see, for example [16, p. 263].

As H is closed in Y, and Y is closed in M, the set H is a closed subset of the compact subset

$$\prod_{\alpha \in A} P(\alpha)$$

of M. Hence H is compact.

Proof of Theorem 2. By Lemma 10, each real-valued continuous function defined on Y is bounded on $H(x_0)$. Now Lemma 12 applies and shows that $H(x_0)$ is compact. When $F(x_0)$ is closed in the $\mathscr{G}_\delta$-topology on Y, Lemma 11 shows that $K(x_0)$ coincides with $H(x_0)$.

4. Locally Convex Topological Vector Spaces

If Z is a topological space, $\mathbb{R}^Z$ denotes the space of all real-valued functions defined on Z, taken with the product topology, or, equivalently, with the

topology of point-wise convergence. We use $B_p(Z)$, $B_p^1(Z)$ and $C_p(Z)$ to denote the subspaces of $\mathbb{R}^Z$ of the Baire measurable functions, of the Baire measurable functions of the first class and of the continuous functions. The Baire measurable functions of the first class are those real-valued functions on Z that are pointwise limits of sequences of continuous functions on Z.

First consider the space $B_p(Z)$, when Z is a K-analytic set. If A is any countable set in $B_p(Z)$ that is relatively countably compact, then for each point z of Z, the set of values taken at z by the functions in A is necessarily bounded. By the implications (v) implies (iv) and (v) implies (vi) in Theorem 40 of the paper [3] of Bourgain, Fremlin and Talagrand, it follows that the closure of A in $B_p(Z)$ is compact and angelic. It follows that the conditions (a) and (b), used in the proofs of Lemmas 6, 7 and 8, hold. Thus, *if Z is K-analytic, and F is a set-valued map from a Hausdorff space X to $B_p(Z)$ with F upper semi-continuous at a point x_0 of X having a countable base for its neighbourhoods, then the set $L(x_0)$ is sequentially compact (and so countably compact) and each cluster point of each countable subset of $L(x_0)$ belongs to $L(x_0)$.*

Now consider the space $B_p^1(Z)$ when Z is an analytic set, *i.e.* a Hausdorff space that is the continuous image of some Polish space. By the Corollary 3 G in [3], this space $B_p^1(Z)$ is angelic. Thus, *if Z is an analytic set, and F is a set-valued map from a Hausdorff space X to $B_p^1(Z)$, with F upper semi-continuous at a point x_0 of X having a countable base for its neighbourhoods, then the sets $L(x_0)$, $K(x_0)$ and $H(x_0)$ all coincide and are compact.*

Floret [17, p. 36], combining results of Grothendieck, Fremlin, Pryce and De Wilde, shows that, if the completely regular Hausdorff space Z can be expressed as the closure of a countable union of relatively countably compact subsets, then $C_p(Z)$ is angelic. Thus Theorem 1 again applies to this class of spaces. Of course, Lemma 2 also applies and shows that every continuous real-valued function defined on $C_p(Z)$ will be bounded on $H(x_0)$. This suggests that it might be possible to show that $H(x_0)$, regarded as a subset of $C(Z)$, should be norm bounded. In § 5, we give a simple example to show that this is not the case, even when Z is a closed interval of the real line.

The special cases when Y is a Banach with its weak topology or a dual Banach space with its weak-star topology are of special importance. We prove two theorems.

Theorem 3. *Let F be a set-valued function mapping a Hausdorff space X into a Banach space Y taken with its weak topology and suppose that F is upper semi-continuous at a point x_0 of X having a countable base for its neighbourhoods. Then the sets $L(x_0)$, $K(x_0)$ and $H(x_0)$ coincide and are weakly compact. Further, x_0 has a neighbourhood N with $F(N)\backslash F(x_0)$ a norm-bounded subset of Y.*

Theorem 4. *Let F be a set-valued function mapping a Hausdorff space X into the dual Y^* of a weakly countably determined Banach space Y, and suppose that F is upper semi-continuous from X to (Y^*, weak-star) at a point x_0 of X having a countable base for its neighbourhoods. Then the sets $L(x_0)$, $K(x_0)$ and $H(x_0)$ all coincide and are weak-star compact. Further, x_0 has a neighbourhood N with $F(N)\backslash F(x_0)$ a norm-bounded subset of Y^*.*

Note that the Banach space Y is said to be *weakly countably determined*, if (Y, weak) can be expressed as the image of some subset of the set of irrational numbers under a compact-valued upper semi-continuous map to (Y, weak). In particular, any weakly K-analytic Banach space or any weakly compactly generated Banach space is weakly countably determined.

Proof of Theorem 3. Let Z be the unit ball of the dual space Y^*, taken with its weak-star topology. Then Z is a compact Hausdorff space and the space $C_p(Z)$ is angelic, see, for example [24, p. 185]. Corresponding to each y in Y, we have a continuous real-valued function on Z, the ball of Y^*, with its weak-star topology, defined by

$$y \to \langle y, y^* \rangle, \qquad y^* \in Z.$$

This map is a homeomorphism taking (Y, weak) onto a linear subspace of $C_p(Z)$. Hence (Y, weak) is angelic and Theorem 1 applies. Thus $L(x_0)$, $K(x_0)$ and $H(x_0)$ coincide and are weakly compact.

Suppose that there were no neighbourhood N of x_0 with

$$F(N) \setminus F(x_0)$$

norm-bounded in Y. Then we could choose $\xi_1, \xi_2, \ldots$ converging to x_0 and $\eta_1, \eta_2, \ldots$ with

$$\eta_n \in F(\xi_n) \setminus F(x_0),$$

for $n \geqq 1$, and with

$$\|\eta_n\| \to \infty \qquad \text{as } n \to \infty.$$

It would follow, from Lemma 2, that each real-valued continuous function h defined on (Y, weak) would be bounded on

$$\{\eta_1, \eta_2, \eta_3, \ldots\}.$$

In particular, for each fixed y^* in the dual space Y^*, the sequence

$$\langle \eta_n, y^* \rangle, \qquad n = 1, 2, \ldots,$$

would be bounded. It would follow, by the uniform boundedness theorem, that $\|\eta_n\|$, $n = 1, 2, \ldots$ would be bounded, and we would have a contradiction. Thus x_0 must have a neighbourhood N with the set $F(N) \setminus F(x_0)$ norm-bounded in Y.

Proof of Theorem 4. Results of Talagrand [25, Theorem 3.4 and Theorem 3.6], when translated into a different language, show that the unit ball B^*, of Y^*, with the weak-star topology, is angelic, when Y is weakly countably determined. In this case tB^* is angelic for each $t > 0$. But the definition of the angelic property refers to a relatively countably compact set A and its closure. Such a relatively countably compact set A in Y^* is necessarily bounded, and so also is its weak-star closure. It follows that (Y^*, weak-star) is angelic. The rest of the proof follows the proof given for Theorem 3.

5. Some Examples of Upper Semi-Continuous Set-Valued Maps

In this section we give some simple examples of upper semi-continuous set-valued maps. In each case, we explain how the existence of the example imposes a limitation on the theory.

Example 1. (The headless man.) Take F to be the set-valued map from $[-1,1]$ to $[0,2]$ defined by

$$F(0)=[1,2].$$
$$F(x)=\{1-|x|,\ 2-|x|\}, \quad \text{for } 0<|x|\leqq 1.$$

It is easy to verify that F is upper semi-continuous and that

$$L(0)=K(0)=H(0)=\{1\},$$
$$L(x)=K(x)=H(x)=F(x), \quad \text{for } 0<|x|\leqq 1.$$

Thus the man loses his body as well as his head, his arms remain but are detached from each other, his legs remain attached to each other. Although $K(x)$ remains non-empty for all x in $[-1,1]$, this set-valued function is not upper semi-continuous at 0.

Example 2. (The fleas.) For each $n\geqq 1$, let

$$F_n=\{(x,y):\ 0\leqq x\leqq 2^{-n},\ n-1\leqq y\leqq n-1+2^{-n}\},$$
$$L_n=\{(2^{-n}+t,\ n-1+2^{-n}-t):\ 0\leqq t\leqq 2^{-n}\},$$
$$E_n=F_n\cup L_n$$

and let

$$A=\bigcup_{n=1}^{\infty} E_n.$$

Let F be the set-valued map from $\mathbb{R}$ to $\mathbb{R}$ with A for its graph. If H is any closed subset of $\mathbb{R}$, it is clear that

$$F^{-1}(H)=F^{-1}(I_n\cap H),$$

where n is the least integer, for which $I_n=[n-1,\ n-1+2^{-n}]$ meets H, if there is such an integer, and $F^{-1}(H)=\varnothing$, otherwise. In this way, it is easy to verify that F is an upper semi-continuous map from $\mathbb{R}$ to $\mathbb{R}$. The graph of the set-valued function K turns out to be the set

$$S=\bigcup_{n=1}^{\infty} L_n.$$

This example illustrates that although each point x has a neighbourhood N with

$$F(N)\backslash F(x)$$

bounded, it is possible neither to choose a neighbourhood N of 0 with $F(N\backslash\{0\})$ bounded, nor to choose such an N with $K(N\backslash\{0\})$ bounded. This example was suggested by the work of A. De Morgan [7, p. 377].

Example 3. (The aerial.) Take F to be the set-valued map from $(0,1)$ to $\mathbb{R}$ defined by

$$F(x)=\{x,x+1\} \quad \text{if } x \text{ is irrational,}$$
$$F\left(\frac{p}{q}\right)=\left[\frac{p}{q}+1-\frac{2}{q},\ \frac{p}{q}+1+\frac{2}{q}\right]\cup\left\{\frac{p}{q}\right\}, \quad \text{if } p \text{ and } q \text{ are integers with } (p,q)=1.$$

Here (p, q) denotes the highest common factor of p and q. It is easy to verify that F is upper semi-continuous and that

$$L(x) = K(x) = H(x) = \{x, x+1\} \quad \text{if } x \text{ is irrational,}$$

and that

$$L\left(\frac{p}{q}\right) = K\left(\frac{p}{q}\right) = H\left(\frac{p}{q}\right) = \left\{\frac{p}{q}\right\} \quad \text{if } p \text{ and } q \text{ are integers with } (p, q) = 1.$$

Now we have $K(x) \neq \varnothing$ for $0 < x < 1$. But

$$[1, 2] \cap K(x) = \{x+1\},$$

for the irrational values of x and

$$[1, 2] \cap K(x) = \varnothing$$

for rational values of x. While Example 1 shows that $K^{-1}(J)$ is not necessarily closed when J is closed, we now see that $K^{-1}(J)$ is not necessarily an $\mathscr{F}_\sigma$-set in the domain of K. However, $K^{-1}(J)$ must be a $\mathscr{G}_\delta$-set in $(0, 1)$ by Lemma 4 in [18].

Example 4. (Two aerials.) Take F to be the set valued map from $(0, 1)$ to $\mathbb{R}$ defined by

$$F(x) = \{x, x+2\},$$

if x and $x+\sqrt{2}$ are both irrational,

$$\left[x - \frac{1}{q},\ x + \frac{1}{q}\right] \cup \{x+2\},$$

if $x = p/q$ with p and q integers with $(p, q) = 1$, and

$$\{x\} \cup \left[x + 2 - \frac{1}{q},\ x + 2 + \frac{1}{q}\right],$$

if $x + \sqrt{2} = p/q$ with p and q integers with $(p, q) = 1$. It is easy to verify that F is upper semi-continuous with compact values. Further $H = K = L$ with

$$K(x) = \{x, x+2\},$$

if x and $x+\sqrt{2}$ are both irrational,

$$K(x) = \{x+2\},$$

if x is rational, and

$$K(x) = \{x\},$$

if $x + \sqrt{2}$ is rational.

Clearly, in this case, F has a continuous selector. The general theory, for upper semi-continuous set-valued maps with arbitrary non-empty values, between metric spaces, [18], tells us that K has a selector of the second Borel

class, we show that, in this example, *the set-valued function K has no selector of the first Borel class.*

Suppose that k were a selector of the first Borel class for K on $(0,1)$. Then the inverse images under k of the sets

$$[0,1] \quad \text{and} \quad [3-\sqrt{2},\ 5-\sqrt{2}],$$

would be $\mathscr{F}_\sigma$-sets with union $(0,1)$. By the Baire category theorem, there must be a subinterval of $(0,1)$ that is entirely covered by one of these sets. If, for example, the interval (α,β) were covered by $k^{-1}([3-\sqrt{2},\ 5-\sqrt{2}])$ we could choose a number x in (α,β) with $x+\sqrt{2}$ a rational number, and we would have a contradiction. This proves the result.

Kuratowski and Ryll-Nardzewski, [20] prove that, if F is a set-valued map with non-empty closed values from a metric space X to a Polish space Y, and $F^{-1}(U)$ is of additive class α in X, for a fixed $\alpha \geqq 1$, for each open set U in Y, then F has a selector of Borel class α. Their result shows that, if $F^{-1}(J)$ is of multiplicative class α in X, for a fixed $\alpha \geqq 0$, for each closed set J in Y, then F has a selector of Borel class $\alpha+1$, but leaves open the possibility that F might always have a selector of Borel class α, at least for $\alpha \geqq 1$. The set-valued map K, discussed above, shows that this will not always be so, in the case $\alpha = 1$.

Example 5. (The disappearing trick.) Take X to be the subset

$$\{0, 1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \ldots\}$$

of the real line. Take $Y = C_p(Z)$, with Z the closed interval $[0,4]$ of the real line. For $n = 1,2,3,\ldots$, write

$$x_n(t) = n^2 t, \qquad \text{for } 0 \leqq t \leqq \frac{1}{n},$$

$$x_n(t) = n^2\left(\frac{2}{n} - t\right), \qquad \text{for } \frac{1}{n} < t \leqq \frac{2}{n},$$

$$x_n(t) = 0, \qquad \text{for } \frac{2}{n} < t \leqq 4,$$

and

$$y_n(t) = 0, \qquad \text{for } 0 \leqq t \leqq 4 - \frac{2}{n}$$

$$y_n(t) = n^2\left(t - 4 + \frac{2}{n}\right), \qquad \text{for } 4 - \frac{2}{n} < t < 4 - \frac{1}{n},$$

$$y_n(t) = n^2(4-t), \qquad \text{for } 4 - \frac{1}{n} < t \leqq 4.$$

Take

$$F(0) = \{0, x_1, x_2, \ldots\},$$

and

$$F\left(\frac{1}{n}\right) = \{y_n, x_1 + y_n,\ x_2 + y_n, \ldots\},$$

for $n \geqq 1$. Then F is a compact-valued upper semi-continuous map from X to $C_p(Z)$. Further $K(0) = F(0)$, and $K(0)$ is compact in $C_p(Z)$ but is unbounded as a set in $C(Z)$. This contrasts with the results in Banach spaces and in dual Banach spaces where a weakly compact set or a weak-star compact set is norm-bounded.

Example 6 (Dolecki and Lechicki). In [13], Dolecki and Lechicki consider the following example. Take X to be the subset $\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ of the real line. Take Y to be the Stone-Čech compactification $\beta\mathbb{N}$ of the discrete set $\mathbb{N}$ of positive integers. Choose any point ξ of $\beta\mathbb{N} \setminus \mathbb{N}$. Take

$$F(0) = \beta\mathbb{N} \setminus (\{\xi\} \cup \mathbb{N}),$$

$$F\left(\frac{1}{n}\right) = n, \qquad n \geqq 1.$$

This set-valued function F is upper semi-continuous, but

$$H(0) = \beta\mathbb{N} \setminus \mathbb{N},$$

$$K(0) = F(0) = \beta\mathbb{N} \setminus (\{\xi\} \cup \mathbb{N}).$$

In this case $H(0)$ is compact, but is not contained in $F(0)$; while $K(0)$ is contained in $F(0)$, but is not compact. As $\beta\mathbb{N}$ is compact, it follows from Theorem 2, that $F(0)$ is not closed in the $\mathcal{G}_\delta$-topology of $\beta\mathbb{N}$.

6. Measurable Selectors

In this section we show how the results of this paper can be used to extend the range of applicability of some results of Jayne and Rogers [19]. We shall need to refer to the details of [19], which, in its turn, refers to the details of [18]. We first prove the following theorem on an upper semi-continuous map F into $c_0(\Gamma)$ with its weak topology; the result was proved in [19] only in the special case when F maps into a set that is a countable union of weakly compact subsets of $c_0(\Gamma)$.

Theorem 5. *Let F be an upper semi-continuous map, with non-empty values, from a metric space X into the Banach space $c_0(\Gamma)$ with its weak topology. Then it is possible to choose a selector f for F, and $\mathcal{F}_\sigma$-set X_1 in X and its complementary $\mathcal{G}_\delta$-set $X_2 = X \setminus X_1$, so that the restrictions of f to X_1 and X_2 are of the first Borel class, using the weak topology on $c_0(\Gamma)$. This selector f will be of the second Borel class on X, using the norm topology on $c_0(\Gamma)$. Further, if F takes only weakly compact values, then F has a selector of the first Borel class, using the weak topology on $c_0(\Gamma)$.*

In [19], Jayne and Rogers were able to obtain norm Borel measurable selectors, when F takes its values, at least locally, within some weakly σ-compact or separable subset of a Banach space, taken with its weak topology. In the following theorem we obtain a rather weaker type of measurability for a much wider class of upper semi-continuous maps.

Theorem 6. *Let Y be a Banach space that is K-analytic in its weak topology. If F is an upper semi-continuous set-valued map, with non-empty values, from a metric space X to Y with its weak topology, then F has a Baire measurable selector, using the weak topology on Y.*

Proof of Theorem 5. In the proof of Lemma 10 of [19], the assumption that F takes its values in a weakly compact set C of $c_0(\Gamma)$ is needed for the proof of Lemma 1, which asserts that $X \times Y$ is a Fréchet space. This property is needed in Lemma 2 to show that $K(x)$ is weakly closed in $Y = c_0(\Gamma)$, for each x in X. The fact that C is weakly compact then ensures that $K(x)$ is weakly compact for each x in X. By Theorem 3, above, we now know that $K(x)$ is always weakly compact in $c_0(\Gamma)$, without any assumption that the values of F lie in a weakly compact subset of $c_0(\Gamma)$. Now Theorem 5 follows by a detailed verification that the methods of [19] and [18] suffice for the proof. There are two points that deserve a few lines of explanation.

The argument in [19] depends essentially on the use of Lemma 3 in [18], and the proof of this lemma makes use of the sequential properties of the space Y (which is there a metric space) in the proof, see the bottom of page 94 and the top of page 95, in [18] that the existence of the sequence (x_i, y_i), $i \geqq 1$, with certain properties leads to a contradiction. However, when Y is a Hausdorff space, we can obtain a contradiction by the following argument. Take

$$L = \mathrm{cl}\,\{y_1, y_2, \ldots\}.$$

Then $L \subset H$. Also, by the upper semi-continuity of F, the set $F^{-1}(L)$ is closed in X and contains the points $x_1, x_2, \ldots$ converging to ξ. Hence $\xi \in F^{-1}(L)$ and there is a point, η^* say, of L in $F(\xi)$. Thus

$$(\xi, \eta^*) \in \mathrm{cl}\,\{(x_i, y_i) : i \geqq 1\} \subset \mathrm{cl}\, E(\xi),$$

and $\eta^* \in K(\xi) \cap H$ contrary to our assumptions.

To obtain the fact that f is of the second Borel class, using the norm topology on $c_0(\Gamma)$, we note that Day has defined a locally uniformly convex norm on $c_0(\Gamma)$ that is equivalent to the supremum norm, see, for example, [8, p. 94]. Now using the results of Edgar [14, Th. 1.1], as in [19], we obtain the required norm Borel measurability of f.

Proof of Theorem 6. As we have remarked in § 4, since Y is weakly K-analytic, it is weakly countably determined. Now Amir and Lindenstrauss [1] have shown that, if Y is a weakly compactly generated Banach space, then there is a set Γ and an injective linear map η of Y into $c_0(\Gamma)$ that is continuous from the weak topology of Y to the weak topology of $c_0(\Gamma)$. More recently Vašák [26] has shown that this same result also holds when Y is weakly countably determined. We use the Γ and η provided by Vašák's result.

As the set-valued map $\eta \circ F$, from X to $c_0(\Gamma)$ with its weak topology, is upper semi-continuous, we can apply Theorem 5 and obtain a selector h for $\eta \circ F$ that is Borel measurable (of the second Borel class) from X to $c_0(\Gamma)$ with its weak topology. Take $f = \eta^{-1} \circ h$. Then f is a selector for F.

Consider any weak Baire set B in Y. Then $Y\backslash B$ is also a weak Baire set. As Y is weakly K-analytic, it follows that B and $Y\backslash B$ are both weakly K-analytic. As η is continuous from (Y, weak) to ($c_0(\Gamma)$, weak) it follows that $\eta(B)$ and $\eta(Y\backslash B)$ are disjoint weakly K-analytic sets in $c_0(\Gamma)$. By the first separation theorem, see, for example, [23, p. 98], we can choose a weak Baire set C with

$$\eta(B)\subset C \quad \text{and} \quad \eta(Y\backslash B)\cap C=\varnothing.$$

Now $f^{-1}(B)=h^{-1}\circ\eta(B)=h^{-1}(C)$. As h is Borel measurable from X to ($c_0(\Gamma)$, weak), it follows that $f^{-1}(B)$ is a Borel set in X. Thus f is Baire measurable, as required.

References

1. Amir, D., Lindenstrauss, J.: The structure of weakly compact sets in Banach spaces. Ann. Math. (2) **88**, 35–46 (1968)
2. Arhangel'skii, A.V.: Functional tightness, Q-spaces and τ-embeddings. Comment. Math. Carolin. **24**, 105–120 (1983)
3. Bourgain, J., Fremlin, D.H., Talagrand, M.: Pointwise compact sets of Baire-measurable functions. Am. J. Math. **100**, 845–886 (1978)
4. Chigogidze, A.Ch.: Plotnost' prostranstva X ravnyaetsya psevdoharakteru prostranstva $C(X)$. Soobščkch. Akad. Nauk Gruz. SSR **103**, 29–32 (1981)
5. Choquet, G.: Convergences. Ann. l'Univ. Grenoble **23**, 57–112 (1947–48)
6. Corson, H.H.: The weak topology of a Banach space. Trans. Am. Math. Soc. **101**, 1–15 (1961)
7. De Morgan, A.: A Budget of Paradoxes. London: Longmans, Green and Co. 1872
8. Diestel, J.: Geometry of Banach spaces – Selected topics. Lecture Notes in Math. no. **485**. Berlin Heidelberg New York: Springer 1975
9. Dolecki, S.: Constraints, stability and moduli of semicontinuity. Preprint prepared for the 2nd IFAC Symposium, Warwick, June, 1977, but not presented in person to the Symposium and thus not published in the proceedings
10. Dolecki, S.: Remarks on semicontinuity. Bull. Acad. Polon. Sci. Sér. Sci. Math. **25**, 863–867 (1977)
11. Dolecki, S., Rolewicz, S.: Metric characterizations of upper semicontinuity. J. Math. Anal. Appl. **69**, 146–152 (1979)
12. Dolecki, S., Lechicki, A.: Semi-continuité supérieure forte et filtres adhérents. C.R. Acad. Sci. Paris, Sér. A **293**, 219–221 (1981)
13. Dolecki, S., Lechicki, A.: On structure of upper semicontinuity. J. Math. Anal. Appl. **88**, 547–554 (1982)
14. Edgar, G.A.: Measurability in a Banach space. Indiana Univ. Math. J. **26**, 663–677 (1977)
15. Edgar, G.A.: Measurability in a Banach space II. Indiana Univ. Math. J. **28**, 559–579 (1979)
16. Engelking, R.: General Topology. Warsaw: PWN 1977
17. Floret, K.: Weakly Compact Sets. Lecture Notes in Math. no. **801**. Berlin Heidelberg New York: Springer 1980
18. Jayne, J.E., Rogers, C.A.: Upper semi-continuous set-valued functions. Acta Math. **149**, 87–125 (1982), see also a correction that is under preparation
19. Jayne, J.E., Rogers, C.A.: Borel selectors for upper semi-continuous mutli-valued functions. J. Funct. Anal. **56**, 279–299 (1984)
20. Kuratowski, K., Ryll-Nardzewski, C.: A general theorem on selectors. Bull. Acad. Polon. Sci. Sér. Sci. Math. **13**, 397–403 (1965)
21. Michael, E.: A note on closed maps and compact sets. Israel J. Math. **2**, 173–176 (1964)

22. Pryce, J.D.: A device of R.J. Whiteley's applied to pointwise compactness in spaces of continuous functions. Proc. London Math. Soc., (3) **23**, 532–546 (1971)
23. Rogers, C.A., Jayne, J.E., Dellacherie, C., Topsøe, F., Hoffmann-Jørgensen, J., Martin, D.A., Kechris, A.S., Stone, A.H.: Analytic Sets. London: Academic Press 1980
24. Schaefer, H.H.: Topological Vector Spaces. Berlin Heidelberg New York: Springer 1971
25. Talagrand, M.: Espaces de Banach faiblement K-analytiques. Ann. Math. **110**, 407–438 (1978)
26. Vašák, L.: On one generalization of weakly compactly generated Banach spaces. Studia Math. **70**, 11–19 (1981)
27. Wheeler, R.F.: Weak and pointwise compactness in the space of bounded continuous functions. Trans. Am. Math. Soc. **266**, 515–530 (1981)

Received May 2, 1984

Math. Z. 189, 319-335 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Dynamics of Populations Structured by Internal Variables

G.F. Webb

Mathematics Department, Vanderbilt University, Nashville, TN 37235, USA

1. Introduction

The models of structured populations are based upon the identification of one or more fundamental characteristics of individuals such as age, size, mass, protein level, nutritional state, or other physical property. There is a considerable literature in the mathematical theory of structured population dynamics. The earliest work concerned linear models of age-dependent population dynamics and derived from the investigations of F. Sharpe and A. Lotka [43] and A. McKendrick [30]. Other researchers beginning with H. Von Foerster [49], G. Bell and E. Anderson [3], and J. Sinko and W. Streifer [46] developed structured population models of cell growth. In the past few years many researchers have contributed to the subject of structured population dynamics for both linear and nonlinear models, and we have listed some of these works in our references. The objective of this paper is to develop a general theory of linear models of populations structured by several internal variables.

The models of structured population dynamics involve first order partial differential equations with nonlocal boundary conditions. In these equations the identified characteristics of individuals are treated as independent variables. In this paper we will consider a population in which individuals are distinguished by age a and another physical characteristic b. The evolution of the population is governed by a density function $\ell(a,b,t)$, so that at time t the total population is $\int_0^\infty \int_0^\infty \ell(a,b,t)\,da\,db$. The *balance law* of the population is

$$\ell_t(a,b,t)+\ell_a(a,b,t)+\ell_b(a,b,t)=-\mu(a,b)\ell(a,b,t) \tag{1.1}$$

where μ accounts for the removal of members of the population due to such causes as mortality and migration. The *input laws* of the population are

$$\ell(a,0,t)=F(\ell(\cdot,\cdot,t))(a), \tag{1.2}$$

$$\ell(0,b,t)=G(\ell(\cdot,\cdot,t))(b). \tag{1.3}$$

Equation (1.2) accounts for the input at time t of individuals of age a with characteristic b at level 0 and (1.3) describes the input at time t of individuals of age a with characteristic b at level 0 and (1.3) describes the input at time t of individuals with age 0 and characteristic b. The initial state of the population is

$$\ell(a, b, 0) = \phi(a, b) \tag{1.4}$$

where ϕ is the prescribed *initial distribution* of the population at time 0.

To formulate precisely a model of structured population dynamics it is necessary to choose a mathematical setting. Since there is a physical interpretation attached to the integral of the density function, it is natural to choose as a setting a Banach space of integrable functions with the L^1 norm. For this setting the formulation (1.1)–(1.4) is too restrictive for a complete mathematical analysis in that the densities may be defined only almost everywhere and the partial derivatives may not exist. A more general formulation of the dynamics of structured populations is given by the integrated version of these equations. Let $L^1 = L^1((0, \infty); \mathbb{R})$, let $X = L^1((0, \infty) \times (0, \infty); \mathbb{R})$, let $L^\infty = L^\infty((0, \infty); \mathbb{R})$, let $X^\infty = L^\infty((0, \infty) \times (0, \infty); \mathbb{R})$, and let L^1_+, X_+, L^∞_+, and X^∞_+ denote the positive cones of L^1, X, L^∞, and X^∞, respectively. We suppose the following hypotheses:

$$\mu \in X^\infty_+ \quad \text{and} \quad \operatorname*{ess\text{-}inf}_{a,b>0} \mu(a, b) \equiv \boldsymbol{\mu} > 0, \tag{1.5}$$

$$F \text{ and } G \text{ are bounded linear operators from } X \text{ to } L^1, \tag{1.6}$$

$$\phi \in X. \tag{1.7}$$

Define for $a, b, t \geqq 0$, $\Pi(a, b, t) = \exp\left[-\int_0^t \mu(a-\tau, b-\tau)\, d\tau\right]$. The general formulation of structured population dynamics in the setting of the Banach space X is the functional equation

$$\ell(a, b, t) = \begin{cases} \Pi(a, b, b)\, F(\ell(\cdot, \cdot, t-b))(a-b), & 0 \leqq b \leqq t, \text{ a.e. } a > b \\ \Pi(a, b, a)\, G(\ell(\cdot, \cdot, t-a))(b-a), & 0 \leqq a \leqq t, \text{ a.e. } b > a \\ \Pi(a, b, t)\, \phi(a-t, b-t), & \text{a.e. } a > t, \text{ a.e. } b > t. \end{cases} \tag{1.8}$$

The solution of Eq. (1.8) is a continuous function $t \to \ell(\cdot, \cdot, t)$ from $[0, \infty)$ to X such that for each $t \geqq 0$, $\ell(\cdot, \cdot, t)$ satisfies (1.8). The solutions of the problem (1.8) form a strongly continuous semigroup $T(t)$, $t \geqq 0$ of bounded linear operators in the Banach space X by means of the formula $T(t)\phi = \ell(\cdot, \cdot, t)$, $t \geqq 0, \phi \in X$. We will identify the infinitesimal generator of this semigroup and analyze its spectral properties.

One of the most valuable applications of mathematical population models is the prediction of ultimate behavior. A fundamental result of mathematical demography, due to F. Sharpe and A. Lotka [43], predicts the convergence of an age-structured population to a stable age distribution. This result was first proved rigorously by W. Feller [16] in 1941. Recently, this result has been proved using semigroup methods by Song Jian, Yu Jing-Yuan, Hu Shung-Ju, and Zhu Guang-Tian [47], G. Greiner [17], and the author [51, 52]. In this

paper we will prove that populations structured by several internal variables also converge to stable distributions. The proof is based upon an analysis of the spectrum of the infinitesimal generator of the associated semigroup. This spectrum consists of essential spectrum lying in a half plane to the left of the imaginary axis, plus nonessential spectrum containing a finite number of eigenvalues of which a unique real eigenvalue is dominant.

In order to describe precisely the convergence of structured populations to stable distributions we require the following hypotheses on the input functions F and G:

$$\begin{aligned} &F\equiv 0 \text{ and for all } \phi\in X, \text{ a.e. } b>0, G(\phi)(b)=\gamma(b)\int_0^\infty\int_0^\infty \beta(\hat{a},\hat{b})\,\phi(\hat{a},\hat{b})\,d\hat{a}\,d\hat{b},\\ &\text{where } \gamma\in L^\infty_+\cap L^1,\ \gamma\not\equiv 0 \text{ and } \beta\in X^\infty_+,\ \beta(a,b)=0 \text{ for } a>b,\ \beta\not\equiv 0. \end{aligned} \tag{1.9}$$

The interpretation of these input functions is that all individuals are born (with age $a=0$) having characteristic $b\geqq 0$, and there is no input of individuals of positive age. It is assumed that b advances equally with a, so that if $\phi(a,b)=0$ for $a>b$ then $\ell(a,b,t)=0$ for $a>b$ (as may be seen from (1.8)). Define the quantities

$$\Gamma(a,b)=\beta(a,a+b)\,\Pi(a,a+b,a),\ a,b>0,$$

$$\Delta(\lambda)=\int_0^\infty\int_0^\infty e^{-\lambda a}\gamma(b)\,\Gamma(a,b)\,da\,db,\ \operatorname{Re}\lambda>-\mu,$$

$$V_\lambda(\phi)=\int_0^\infty\int_0^\infty \beta(a,b)\left[\int_0^a e^{-\lambda\sigma}\Pi(a,b,\sigma)\,\phi(a-\sigma,b-\sigma)\,d\sigma\right]da\,db,\ \operatorname{Re}\lambda>-\mu,\ \phi\in X,$$

$$M_\lambda=\int_0^\infty\int_0^\infty a e^{-\lambda a}\gamma(b)\,\Gamma(a,b)\,da\,db,\ \operatorname{Re}\lambda>-\mu.$$

The convergence of a structured population to a stable distribution is described as follows: If $\Delta(0)<1$, then $\lim_{t\to\infty}\|\ell(\cdot,\cdot,t)\|_X=0$. If $\Delta(0)\geqq 1$, then $\lim_{t\to\infty}\|e^{-\lambda_1 t}\ell(\cdot,\cdot,t)-P_0\,\phi\|_X=0$, where $P_0: X\to X$ is the projection

$$(P_0\,\phi)(a,b)=\begin{cases}0, & a>b\\ e^{-\lambda_1 a}\Pi(a,b,a)\,\gamma(b-a)\,V_{\lambda_1}(\phi)\,M_{\lambda_1}^{-1}, & \text{a.e. } b>a.\end{cases} \tag{1.10}$$

In the model of structured population dynamics we have considered it is assumed that the internal variables advance equally with age. In many applications this assumption is not valid and it is necessary to introduce a velocity function to account for the growth of an internal variable which does not increase at the same rate as age. Such a model involves the balance and input laws

$$n_t(a,c,t)+n_a(a,c,t)+(v(c)\,n(a,c,t))_c=-m(a,c)\,n(a,c,t), \tag{1.11}$$

$$n(a,0,t)=H(n(\cdot,\cdot,t))(a), \tag{1.12}$$

$$n(0,c,t)=K(n(\cdot,\cdot,t))(c), \tag{1.13}$$

$$n(a,c,0)=\psi(a,c) \tag{1.14}$$

where a represents age and c represents another internal variable with velocity function v. As we shall see, problem (1.11)-(1.14) is easily converted to the constant velocity case considered previously. Consequently, we will prove that solutions of this model of structured population dynamics also converge to stable distributions.

2. Basic Theory

The first order partial differential Eq. (1.1), together with its boundary conditions (1.2) and (1.3), and its initial condition (1.4), may be solved by the method of characteristics. The characteristic curves are the lines $a-t=\text{const}=a_0$, $b-t=\text{const}=b_0$. Consider such a characteristic curve passing through the point (a,b,t) and intersecting one of the three coordinate planes in the (a,b,t) coordinate system. Consider also the solution of (1.1) along such a characteristic curve and define $\omega_{a_0,b_0}(\tau)=\ell(a_0+\tau,b_0+\tau,\tau)$, $\tau\geqq t_0$. From (1.1) we obtain $\omega'_{a_0,b_0}(\tau)=-\mu(a_0+\tau,b_0+\tau)\omega_{a_0,b_0}(\tau)$, $\tau\geqq t_0$ so that

$$\omega_{a_0,b_0}(t)=\exp\left[-\int_{t_0}^{\tau}\mu(a_0+\sigma,b_0+\sigma)\,d\sigma\right]\omega_{a_0,b_0}(t_0),\qquad \tau\geqq t_0, \tag{2.1}$$

If the characteristic curve intersects the (a,t)-plane, then it connects the point $(a-b,0,t-b)$ to the point (a,b,t). In this case $t_0=t-b$ and (2.1) yields

$$\ell(a,b,t)=\exp\left[-\int_{t-b}^{t}\mu(a-t+\sigma,b-t+\sigma)\,d\sigma\right]\ell(a-b,0,t-b). \tag{2.2}$$

If the characteristic curve intersects the (b,t)-plane, then it connects $(0,b-a,t-a)$ to (a,b,t). In this case $t_0=t-a$ and (2.1) yields

$$\ell(a,b,t)=\exp\left[-\int_{t-a}^{t}\mu(a-t+\sigma,b-t+\sigma)\,d\sigma\right]\ell(0,b-a,t-a). \tag{2.3}$$

Lastly, if the characteristic curve intersects the (a,b)-plane, then it connects $(a-t,b-t,0)$ to (a,b,t). In this case $t_0=0$ and (2.1) yields

$$\ell(a,b,t)=\exp\left[-\int_{0}^{t}\mu(a-t+\sigma,b-t+\sigma)\,d\sigma\right]\ell(a-t,b-t,0). \tag{2.4}$$

Now use (2.2) with (1.2), (2.3) with (1.3), and (2.4) with (1.4) to obtain the Eq. (1.8). We state

Theorem 2.1. *Let* (1.5), (1.6), (1.7) *hold. There exists a unique continuous function* $t\mapsto\ell(\cdot,\cdot,t)$ *from* $[0,\infty)$ *to* X *satisfying* (1.8). *Further, for* $t\geqq 0$

$$\|\ell(\cdot,\cdot,t)\|_X\leqq\exp[t(|F|+|G|-\mu)]\,\|\phi\|_X. \tag{2.5}$$

Proof. The proof of the existence of a unique solution to (1.8) on some interval $[0,T]$ is given in [45], Proposition 3.1. That this solution exists on $[0,\infty)$

follows from Proposition 3.2 in [45], provided that we can establish (2.5). To prove (2.5) consider the estimate

$$\begin{aligned}\int_0^\infty \int_0^\infty |\ell(a,b,t)|\,dadb &\leqq \int_0^t \int_b^\infty e^{-\mu b}|F(\ell(\cdot,\cdot,t-b))(a-b)|\,dadb \\ &\quad + \int_0^t \int_a^\infty e^{-\mu a}|G(\ell(\cdot,\cdot,t-a)(b-a)|\,dbda \\ &\quad + \int_t^\infty \int_t^\infty e^{-\mu t}|\phi(a-t,b-t)|\,dadb \\ &\leqq \int_0^t e^{-\mu(t-\tau)}(|F|+|G|)\,\|\ell(\cdot,\cdot,\tau)\|_X\,d\tau + e^{-\mu t}\|\phi\|_X.\end{aligned}$$

Now use Gronwall's Lemma to obtain (2.5). □

In physical applications it is the nonnegative valued solutions of (1.8) which have significance. The proof of the following result is given in [45], Proposition 3.4:

Theorem 2.2. *Let* (1.5), (1.6), (1.7) *hold and in addition, let* $F(X_+)\subset L^1_+$, $G(X_+)\subset L^1_+$, *and* $\phi\in X_+$. *Then,* $\ell(\cdot,\cdot,t)\in X_+$ *for all* $t\geqq 0$.

3. The Semigroup and Its Infinitesimal Generator

The solutions of (1.8) give rise to a strongly continuous semigroup of bounded linear operators in the Banach space X. We state

Theorem 3.1. *Let* (1.5), (1.6), (1.7) *hold and define the family of bounded linear operators* $T(t)$, $t\geqq 0$ *in* X *by* $T(t)\phi=\ell(\cdot,\cdot,t)$, $t\geqq 0$. *Then,* $T(t)$, $t\geqq 0$ *satisfies*

$$T(t+s)=T(t)\,T(s), \qquad t,s\geqq 0, \tag{3.1}$$

$$t\mapsto T(t)\phi \quad \text{is continuous from } [0,\infty) \text{ to } X, \tag{3.2}$$

$$T(0)=I, \tag{3.3}$$

$$|T(t)|\leqq \exp[t(|F|+|G|-\mu)], \qquad t\geqq 0. \tag{3.4}$$

Proof. The semigroup property (3.1) may be shown by direct computation from (1.8) (we omit the details). The strong continuity property (3.2) follows immediately from Theorem 2.1, (3.3) follows immediately from (1.8), and (3.4) follows immediately from (2.5). □

The infinitesimal generator of the strongly continuous semigroup $T(t)$, $t\geqq 0$ is defined as the (unbounded) linear operator A in X given by $A\phi=\lim_{t\to 0^+} t^{-1}(T(t)\phi-\phi)$, $\phi\in D(A)$, $D(A)=\{\phi\in X\colon \lim_{t\to 0^+} t^{-1}(T(t)\phi-\phi) \text{ exists}\}$. We first provide a characterization of the point spectrum $P\sigma(A)$ of A.

Theorem 3.2. *Let* (1.5), (1.6), (1.7) *hold. Then,* $\phi \in D(A)$ *and* $A\phi = \lambda\phi$ *for some* $\lambda \in \mathbb{C}$ *if and only if* ϕ *satisfies*

$$\phi(a,b) = \begin{cases} e^{-\lambda b}\Pi(a,b,b)F(\phi)(a-b), & \text{a.e. } a > b \\ e^{-\lambda a}\Pi(a,b,a)G(\phi)(b-a), & \text{a.e. } b > a. \end{cases} \tag{3.5}$$

Proof. Suppose ϕ satisfies (3.5) (in the complexification of X) and define the function $t \mapsto \hat{\ell}(\cdot,\cdot,t)$ from $[0,\infty)$ to X by $\hat{\ell}(\cdot,\cdot,t) = e^{\lambda t}\phi, t \geqq 0$. We will show that $\hat{\ell}(\cdot,\cdot,t)$ satisfies (1.8). First, for $0 \leqq b \leqq t$, a.e. $a > b$,

$$\Pi(a,b,b)F(\hat{\ell}(\cdot,\cdot,t-b))(a-b) = e^{\lambda(t-b)}\Pi(a,b,b)F(\phi)(a-b) = \hat{\ell}(a,b,t)$$

and similarly, for $0 \leqq a \leqq t$, a.e. $b > a$, $\Pi(a,b,a)G(\hat{\ell}(\cdot,\cdot,t-a))(b-a) = \hat{\ell}(a,b,t)$. Next, for a.e. $a > t$, a.e. $b > t$, $a > b$

$$\begin{aligned} \Pi(a,b,t)\phi(a-t,b-t) &= e^{\lambda(t-b)}\Pi(a,b,t)\Pi(a-t,b-t,b-t)F(\phi)(a-b) \\ &= e^{\lambda(t-b)}\Pi(a,b,b)F(\phi)(a-b) = \hat{\ell}(a,b,t) \end{aligned}$$

and similarly, for a.e. $a > t$, a.e. $b > t$, $b > a$, $\Pi(a,b,t)\phi(a-t,b-t) = \hat{\ell}(a,b,t)$. Since the solution of (1.8) is unique, we must have that $T(t)\phi = e^{\lambda t}\phi$, $t \geqq 0$. Thus, $\phi \in D(A)$ and $A\phi = \lambda\phi$.

Suppose now that $\phi \in D(A)$ and $A\phi = \lambda\phi$. From standard semigroup theory we must have that $T(t)\phi = e^{\lambda t}\phi$, $t \geqq 0$. From (1.8) (in the complexification of X) we have that for $t > 0$

$$e^{\lambda t}\phi(a,b) = \begin{cases} \Pi(a,b,b)F(e^{\lambda(t-b)}\phi)(a-b), & 0 \leqq b \leqq t, \text{ a.e. } a > b \\ \Pi(a,b,a)G(e^{\lambda(t-a)}\phi)(b-a), & 0 \leqq a \leqq t, \text{ a.e. } b > a. \end{cases}$$

Now use the linearity of F and G to conclude that ϕ satisfies (3.5). □

In the case that (1.9) is satisfied the point spectrum of A can be characterized as follows:

Theorem 3.3. *Let* (1.5), (1.9) *hold and let* $\operatorname{Re}\lambda > -\mu$. *Then,* $\lambda \in P\sigma(A)$ *if and only if* $\Delta(\lambda) = 1$. *Moreover, if* $\lambda \in P\sigma(A)$, *then* $A\phi = \lambda\phi$ *if and only if there is some constant* c *such that*

$$\phi(a,b) = \begin{cases} 0, & a > b \\ ce^{-\lambda a}\Pi(a,b,a)\gamma(b-a), & \text{a.e. } b > a. \end{cases} \tag{3.6}$$

Proof. Suppose that $A\phi = \lambda\phi$ for some $\phi \neq 0$ and define $\ell(b) = G(\phi)(b)$, a.e. $b > 0$. By Theorem 3.2 ϕ must satisfy (3.5) and hence

$$\begin{aligned} G(\phi)(b) &= \gamma(b)\int_0^\infty \int_0^\infty \beta(\hat{a},\hat{b})\phi(\hat{a},\hat{b})\,d\hat{a}\,d\hat{b} \\ &= \gamma(b)\int_0^\infty \int_0^{\hat{b}} \beta(\hat{a},\hat{b})e^{-\lambda\hat{a}}\Pi(\hat{a},\hat{b},\hat{a})\ell(\hat{b}-\hat{a})\,d\hat{a}\,d\hat{b} \\ &= \gamma(b)\int_0^\infty \int_0^\infty e^{-\lambda\hat{a}}\Gamma(\hat{a},\hat{b})\ell(\hat{b})\,d\hat{b}\,d\hat{a}. \end{aligned}$$

Thus, ϕ satisfies (3.6) with $c=\int_0^\infty\int_0^\infty e^{-\lambda\hat{a}}\Gamma(\hat{a},\hat{b})\ell(\hat{b})\,d\hat{b}\,d\hat{a}$. Also, $\Delta(\lambda)=1$, since the formula for $k(b)=G(\phi)(b)$ above yields

$$\int_0^\infty\left[\int_0^\infty e^{-\lambda a}\Gamma(a,b)\,da\right]\ell(b)\,db=\Delta(\lambda)\int_0^\infty\left[\int_0^\infty e^{-\lambda\hat{a}}\Gamma(\hat{a},\hat{b})\ell(\hat{b})\,d\hat{a}\right]d\hat{b}.$$

Now suppose that $\Delta(\lambda)=1$ and define ϕ as in (3.6) with $c\neq 0$. Let $\ell(b)=c\gamma(b)$ and use the calculation above to show that $G(\phi)(b)=c\gamma(b)$. By Theorem 3.2 we conclude that $A\phi=\lambda\phi$. □

Theorem 3.4. *Let* (1.5), (1.6) *hold, let* $\lambda\in\rho(A)$ *(the resolvent set of A) such that* $\operatorname{Re}\lambda>|F|+|G|-\mu$, *let* $\psi\in X$, *and let* $\phi=(\lambda I-A)^{-1}\psi$. *Then*

$$\phi(a,b)=\begin{cases}e^{-\lambda b}\Pi(a,b,b)F(\phi)(a-b)+\int_0^b e^{-\lambda\sigma}\Pi(a,b,\sigma)\psi(a-\sigma,b-\sigma)\,d\sigma & \text{a.e. } a>b\\ e^{-\lambda a}\Pi(a,b,a)G(\phi)(b-a)+\int_0^a e^{-\lambda\sigma}\Pi(a,b,\sigma)\psi(a-\sigma,b-\sigma)\,d\sigma & \text{a.e. } b>a.\end{cases}\tag{3.7}$$

Proof. From standard results in semigroup theory (3.4) implies that $(\lambda I-A)^{-1}\psi=\int_0^\infty e^{-\lambda t}T(t)\psi\,dt$. The claim (3.7) then follows immediately from (1.8). □

Remark 3.1. Under the hypothesis (1.5) and (1.6) it can be shown that

$$A\phi(a,b)=-d/da\left[d/db\int_0^a\phi(\hat{a},b)\,d\hat{a}+\phi(a,b)\right]$$
$$-\mu(a,b)\phi(a,b),\quad \text{a.e. } a>0,\ \text{a.e. } b>0,$$

$$D(A)=\Bigg\{\phi\in X:(a,b)\mapsto\int_0^a\phi(\hat{a},b)\,d\hat{a}\text{ is absolutely continuous in }b\text{ for }a\geqq 0,$$

$$(a,b)\mapsto\left[d/db\int_0^a\phi(\hat{a},b)\,d\hat{a}+\phi(a,b)\right]\text{ is absolutely continuous in }a\text{ for a.e. }b>0,$$

$$\lim_{a\to 0^+}\left[d/db\int_0^a\phi(\hat{a},b)\,d\hat{a}+\phi(a,b)\right]=G(\phi)(b)\text{ for a.e. }b>0,$$

$$(a,b)\mapsto\int_0^b\phi(a,\hat{b})\,d\hat{b}\text{ is absolutely continuous in }a\text{ for }b\geqq 0,$$

$$(a,b)\mapsto\left[d/da\int_0^b\phi(a,\hat{b})\,d\hat{b}+\phi(a,b)\right]\text{ is absolutely continuous in }b\text{ for a.e. }a>0,$$

$$\lim_{b\to 0^+}\left[d/da\int_0^b\phi(a,\hat{b})\,d\hat{b}+\phi(a,b)\right]=F(\phi)(a)\text{ for a.e. }a>0,$$

$$\text{and } d/da\left[d/db\int_0^a\phi(\hat{a},\hat{b})\,d\hat{a}+\phi(a,b)\right]\in X\Bigg\}.$$

Further, $\phi \in D(A)$ if and only if there exists $\chi \in X$ such that for $a \geqq 0$, $b \geqq 0$

$$\int_0^a \phi(\hat{a}, b)\, d\hat{a} + \int_0^b \phi(a, \hat{b})\, d\hat{b} + \int_0^b \int_0^a \chi(\hat{a}, \hat{b})\, d\hat{a} d\hat{b} = \int_0^a F(\phi)(\hat{a})\, d\hat{a} + \int_0^b G(\phi)(\hat{b})\, d\hat{b}.$$

If $\phi \in D(A)$, then ϕ satisfies this last equation with $\chi = A\phi + \mu\phi$. If in addition ϕ is sufficiently smooth, then $A\phi = -(\phi_a + \phi_b + \mu\phi)$, $\phi(a, 0) = F(\phi)(a)$, $a \geqq 0$, $\phi(0, b) = G(\phi)(b)$, $b \geqq 0$.

Theorem 3.5. *Let* (1.5), (1.9) *hold, let* $\operatorname{Re}\lambda > |F| + |G| - \mu$, *let* $\lambda \in \rho(A)$, *and let* $\psi \in X$. *Then*

$$((\lambda I - A)^{-1}\psi)(a, b) = \begin{cases} \int_0^b e^{-\lambda\sigma} \Pi(a, b, \sigma)\, \psi(a - \sigma, b - \sigma)\, d\sigma, & a > b \\ e^{-\lambda a} \Pi(a, b, a)\, \gamma(b - a)\, V_\lambda(\psi)(1 - \Delta(\lambda))^{-1} & \\ \quad + \int_0^a e^{-\lambda\sigma} \Pi(a, b, \sigma)\, \psi(a - \sigma, b - \sigma)\, d\sigma, & \text{a.e. } b > a. \end{cases} \tag{3.8}$$

Proof. Since $\lambda \in \rho(A)$, $\Delta(\lambda) \neq 1$ by Theorem 3.3. Let $c = V_\lambda(\psi)(1 - \Delta(\lambda))^{-1}$. By Theorem 3.4 it suffices to show that if $\phi = (\lambda I - A)^{-1}\psi$ is defined as in (3.8), then $G(\phi) = c\gamma$. As in the proof of Theorem 3.3.

$$\begin{aligned} G(\phi)(b) &= \gamma(b) \int_0^\infty \int_0^\infty \beta(\hat{a}, \hat{b})\, \phi(\hat{a}, \hat{b})\, d\hat{a} d\hat{b} \\ &= \gamma(b) \int_0^\infty \int_0^{\hat{b}} \beta(\hat{a}, \hat{b}) \Big\{ e^{-\lambda\hat{a}} \Gamma(\hat{a}, \hat{b})\, c\gamma(\hat{b}) \\ &\qquad + \int_0^{\hat{a}} e^{-\lambda\sigma} \Pi(\hat{a}, \hat{b}, \sigma)\, \psi(\hat{a} - \sigma, \hat{b} - \sigma)\, d\sigma \Big\} d\hat{a} d\hat{b} \\ &= \gamma(b)\, [c\Delta(\lambda) + V_\lambda(\psi)] \end{aligned}$$

which means that $G(\phi)(b) = c\gamma(b)$. □

4. Properties of the Semigroup

The ultimate behavior of the solutions of (1.8) can be determined from the properties of the semigroup $T(t)$, $t \geqq 0$. One result in this direction is provided by Theorem 2.1, which establishes that if $\operatorname{ess\text{-}inf}_{a, b > 0} \mu(a, b) > |F| + |G|$, then the population eventually becomes extinct. Another result in this direction is established by the following theorem, which provides sufficient conditions for the portion of the population in existence at $t = 0$ to approach ultimate extinction (the proof is very similar to the proof of (2.5)):

For each $t \geqq 0$ define the operator $U(t)$ in X as follows: for $\phi \in X$

$$(U(t)\phi)(a, b) = \begin{cases} 0, 0 \leqq b \leqq t, & \text{a.e. } a > b \\ 0, 0 \leqq a \leqq t, & \text{a.e. } b > a \\ \Pi(a, b, t)\, \phi(a - t, b - t), & \text{a.e. } a > t, \text{ a.e. } b > t. \end{cases}$$

Theorem 4.1. *Let* (1.5), (1.6), (1.7) *hold. Then, for* $t \geqq 0 \| U(t)\phi\|_X \leqq e^{-\mu t}\|\phi\|_X$.

Our next objective is to determine some compactness properties of the semigroup $T(t)$, $t \geqq 0$ in the case that the input functions satisfy (1.9). The proof of the following lemma may be found is [14], Theorem 21, p. 301:

Lemma 4.1. *A closed and bounded subset* M *of* X *is compact if and only if* $\lim_{h,k\to 0} \int_0^\infty \int_0^\infty |\phi(a+h,b+k)-\phi(a,b)|\,da\,db=0$ *and* $\lim_{h,k\to\infty} \int_h^\infty \int_k^\infty |\phi(a,b)|\,da\,db=0$ *uniformly for* $\phi \in M$.

For each $t \geqq 0$ define the operator $W(t)$ in X as follows: for $\phi \in X$

$$(W(t)\phi)(a,b)=\begin{cases} 0, 0\leqq b\leqq t, \text{ a.e. } a>b \\ \Pi(a,b,a)\,G(T(t-a)\phi)(b-a), 0\leqq a\leqq t, \text{ a.e. } b>a \\ 0, \text{ a.e. } a>t, \text{ a.e. } b>t. \end{cases}$$

Theorem 4.2. *Let* (1.5), (1.9) *hold. Then, for* $t \geqq 0$ *the mapping* $W(t)$ *is compact in* X.

Proof. For $\phi \in X$, $0 \leqq a \leqq t$, a.e. $b>a$ we have that

$$\begin{aligned} G(T(t-a)\phi)(b-a) &= \gamma(b-a)\int_0^\infty\int_0^\infty \beta(\hat{a},\hat{b})\,T(t-a)\phi(\hat{a},\hat{b})\,d\hat{a}\,d\hat{b} \\ &= \gamma(b-a)\Big\{\int_{t-a}^\infty\int_{t-a}^\infty \beta(\hat{a},\hat{b})\,\Pi(\hat{a},\hat{b},t-a)\,\phi(\hat{a}-t+a,\hat{b}-t+a)\,d\hat{a}\,d\hat{b} \\ &\quad + \int_0^{t-a}\int_{\hat{a}}^\infty \beta(\hat{a},\hat{b})\,\Pi(\hat{a},\hat{b},\hat{a})\,G(T(t-a-\hat{a})\phi)(\hat{b}-\hat{a})\,d\hat{b}\,d\hat{a}\Big\} \\ &= \gamma(b-a)\Big\{\int_0^\infty\int_0^\infty \beta(\hat{a}+t-a,\hat{b}+t-a)\,\Pi(\hat{a}+t-a,\hat{b}+t-a,t-a)\,\phi(\hat{a},\hat{b})\,d\hat{a}\,d\hat{b} \\ &\quad + \int_0^{t-a}\int_0^\infty \beta(t-a-\hat{a},\hat{b}+t-a-\hat{a})\,\Pi(t-a-\hat{a},\hat{b}+t-a-\hat{a},t-a-\hat{a}) \\ &\quad \cdot G(T(\hat{a})\phi)(\hat{b})\,d\hat{b}\,d\hat{a}\Big\}. \end{aligned}$$

The compactness of $W(t)$ now follows from Lemma 4.1. □

5. Properties of the Infinitesimal Generator

In this section we develop some of the spectral properties of the infinitesimal generator A of $T(t)$, $t \geqq 0$. We require the following standard definitions: Let A be an (arbitrary) closed linear operator in the Banach space X. The *essential spectrum* of A is the subset $E\sigma(A)$ of the spectrum $\sigma(A)$ of A defined by $E\sigma(A) = \{\lambda \in \sigma(A)$: λ is not isolated in $\sigma(A)$ or the range $R(\lambda I - A)$ of $\lambda I - A$ is not closed or the generalized eigenspace $U_{k=1}^\infty N((\lambda I - A)^k)$ of λ is infinite dimensional$\}$. Let M be a bounded set in X. The *measure of noncompactness* of M is

defined as $\alpha[M]=\inf\{\varepsilon>0$: M can be covered by a finite number of sets each having diameter $\leqq\varepsilon\}$. Let T be a bounded linear operator in X. The *measure of noncompactness* of T is defined as $\alpha[T]=\inf\{\varepsilon>0$: $\alpha[T(M)]\leqq\varepsilon\alpha[M]$ for every bounded set M in $X\}$. The *spectral radius* of T is defined as $r_\sigma(T)=\sup\{|\lambda|: \lambda\in\sigma(T)\}$ and the *essential spectral radius* of T is defined as $r_{E\sigma}(T)=\sup\{|\lambda|: \lambda\in E\sigma(T)\}$. A proof of the following lemma may be found in [51], Chap. 4:

Lemma 5.1. *Let $T(t)$, $t\geqq 0$ be a strongly continuous semigroup of bounded linear operators in the Banach space X with infinitesimal generator A. The following hold:*

$$\omega_0(A)\overset{\text{def}}{=}\lim_{t\to\infty} t^{-1}\log(|T(t)|) \text{ exists,} \tag{5.1}$$

$$\omega_1(A)\overset{\text{def}}{=}\lim_{t\to\infty} t^{-1}\log(\alpha[T(t)]) \text{ exists,} \tag{5.2}$$

$$\text{If } \omega>\omega_0(A), \text{ then there exists } M_\omega\geqq 1 \text{ such that } |T(t)|\leqq M_\omega e^{\omega t},\ t\geqq 0, \tag{5.3}$$

$$\sup_{\lambda\in\sigma(A)}\operatorname{Re}\lambda\leqq\omega_0(A), \tag{5.4}$$

$$\sup_{\lambda\in E\sigma(A)}\operatorname{Re}\lambda\leqq\omega_1(A), \tag{5.5}$$

$$\omega_0(A)=\max\{\omega_1(A), \sup_{\lambda\in\sigma(A)-E\sigma(A)}\operatorname{Re}\lambda\}. \tag{5.6}$$

The determination of $\omega_0(A)$ (known as the *growth bound* of the semigroup $T(t)$, $t\geqq 0$) is of central importance in analyzing the asymptotic behavior of the semigroup trajectories. A means for determining $\omega_0(A)$ is provided by (5.6) (due to J. Prüss [36]). If $\omega_0(A)<0$, then (5.3) yields the exponential asymptotic stability of the zero equilibrium of $T(t)$, $t\geqq 0$. If $\omega_0(A)\geqq 0$, then the following lemma yields precise information concerning the asymptotic behavior (the proof may be found in [51], Chap. 4):

Lemma 5.2. *Let $T(t)$, $t\geqq 0$ be a strongly continuous semigroup of bounded linear operators in the Banach space X with infinitesimal generator A and let $\lambda_1\in\sigma(A)$ such that $\lambda_1>\omega_1(A)$, λ_1 is a real simple eigenvalue of A, (that is, the generalized eigenspace of λ_1 is $N(\lambda_1 I-A)$) and there exists a constant ω such that $\max\{\omega_1(A), \sup_{\lambda\in\sigma(A)-E\sigma(A)-\{\lambda_1\}}\operatorname{Re}\lambda\}<\omega<\lambda_1$. Then, there exists a direct sum decomposition $X=X_0\oplus X_1$ and associated projections $P_i X=X_i$, $i=0,1$, with the following properties:*

$$P_0=(2\pi i)^{-1}\int_\Gamma(\lambda I-A)^{-1}d\lambda \text{ and } P_1=I-P_0 \text{ (where } \Gamma \text{ is a positively oriented closed curve in } \mathbb{C} \text{ enclosing } \lambda_1, \text{ but no other point of } \sigma(A)), \tag{5.7}$$

$$X_0=N(\lambda_1 I-A), \tag{5.8}$$

$$P_0 T(t)=T(t)P_0=e^{t\lambda_1}P_0,\ t\geqq 0, \tag{5.9}$$

$$P_1 T(t)=T(t)P_1 \text{ and there exists } M_1\geqq 1 \text{ such that } |T(t)P_1|\leqq M_1 e^{\omega t}|P_1|,\ t\geqq 0. \tag{5.10}$$

Theorem 5.1. *Let* (1.5), (1.9) *hold. Then,*

$$\omega_1(A) \leqq -\mu, \tag{5.11}$$

For λ_0 *such that* $\operatorname{Re}\lambda_0 > -\mu$, $\lambda_0 \in \sigma(A)$ *if and only if* $\lambda_0 \in P\sigma(A)$ *if and only if* $\lambda_0 \in \sigma(A) - E\sigma(A)$ *if and only if* λ_0 *is a pole of* $(\lambda I - A)^{-1}$ *of order* m *if and only if* m *is the smallest positive integer such that the generalized eigenspace of* λ_0 *is* $N((\lambda_0 I - A)^m)$ *if and only if* λ_0 *is a pole of* $(1 - \Delta(\lambda))^{-1}$ *of order* m. (5.12)

Proof. To prove (5.11) we will use Theorems 4.1 and 4.2 together with various properties of the measure of noncompactness (see [29], Lemma 5.1, p. 17). Observe that for $t > 0$, $\alpha[T(t)] = \alpha[U(t) + W(t)] \leqq \alpha[U(t)] + \alpha[W(t)] \leqq e^{-\mu t} + 0$. The claim (5.11) now follows immediately from (5.2).

To prove (5.12) suppose that $\operatorname{Re}\lambda_0 > -\mu \geqq \omega_1(A)$. If $\lambda_0 \in \sigma(A)$, then $\lambda_0 \in \sigma(A) - E\sigma(A)$ by (5.5). If $\lambda_0 \in \sigma(A) - E\sigma(A)$, then λ_0 is a pole of $(\lambda I - A)^{-1}$ and $\lambda_0 \in P\sigma(A)$ (see [51], Proposition 4.11). Also, λ_0 is a pole of order m if and only if m is the smallest positive integer such that the generalized eigenspace of λ_0 is $N((\lambda_0 I - A)^m)$ (see [51], Proposition 4.8 and Proposition 4.11). We have shown that $\{\lambda : \operatorname{Re}\lambda > -\mu\} = \rho(A) \cup P\sigma(A)$. By Theorem 3.3 for λ such that $\operatorname{Re}\lambda > -\mu$, $\lambda \in P\sigma(A)$ if and only if $\Delta(\lambda) = 1$. By analytic continuation the formula for $(\lambda I - A)^{-1}$ in (3.8) holds for λ such that $\operatorname{Re}\lambda > -\mu$ and $\Delta(\lambda) \neq 1$. Thus, λ_0 is a pole of $(1 - \Delta(\lambda))^{-1}$ of order m if and only if λ_0 is a pole of $(\lambda I - A)^{-1}$ of order m. □

6. Convergence to Stable Distributions

We are now prepared to prove that the solutions of structured population dynamics are asymptotically convergent in time to stable distributions. We first consider the case of constant velocity in the balance law.

Theorem 6.1. *Let* (1.5), (1.9) *hold. The following are true:*

If $\Delta(0) < 1$, *then there exists constants* $M > 0$ *and* $\omega < 0$ *such that* $\|T(t)\phi\| x \leqq Me^{\omega t}\|\phi\|_X$ *for* $\phi \in X$, $t \geqq 0$. (6.1)

If $\Delta(0) \geqq 1$, *then there exists a unique real* $\lambda_1 \geqq 0$ *such that* $\Delta(\lambda_1) = 1$ *and there exist constants* $M > 0$ *and* $\omega < 0$ *such that for all* $\phi \in X$, $t \geqq 0$, $\|e^{-\lambda_1 t}T(t)\phi - P_0\phi\|_X \leqq Me^{\omega t}\|\phi\|_X$. (6.2)

Proof. Suppose $\Delta(0) < 1$. If there exists no λ such that $\operatorname{Re}\lambda > -\mu$ and $\Delta(\lambda) = 1$, then $\omega_0 \leqq -\mu$ by (5.6) and (5.12). Then, (6.1) follows immediately from (5.3). If there exists some λ_0 such that $\operatorname{Re}\lambda_0 > -\mu$ and $\Delta(\lambda_0) = 1$, then there must exist a unique real value λ_1 such that $\operatorname{Re}\lambda_0 \leqq \lambda_1 < 0$ and $\Delta(\lambda_1) = 0$. This claim follows from the fact that

$$\begin{aligned}\int_0^\infty \int_0^\infty \gamma(b)\Gamma(a,b)\,da\,db < 1 = \operatorname{Re}\int_0^\infty \int_0^\infty e^{-\lambda_0 a}\gamma(b)\Gamma(a,b)\,da\,db \\ = \int_0^\infty \int_0^\infty e^{-\operatorname{Re}\lambda_0 a}\cos(\operatorname{Im}\lambda_0 a)\,\gamma(b)\Gamma(a,b)\,da\,db \\ \leqq \int_0^\infty \int_0^\infty e^{-\operatorname{Re}\lambda_0 a}\gamma(b)\,\Gamma(a,b)\,da\,db \end{aligned} \tag{6.3}$$

which also shows that $\lambda_1 > \operatorname{Re}\lambda$ for any λ such that $\Delta(\lambda)=1$. Then, (5.6) and (5.12) yield that $\omega_0=\lambda_1$ and again (6.1) follows from (5.3).

Suppose now that $\Delta(0)\geqq 1$. A calculation similar to (6.3) shows that there must exist a unique real value λ_1 such that $\Delta(\lambda_1)=1$, $\lambda_1\geqq 0$, and $\lambda_1 > \operatorname{Re}\lambda$ for any $\lambda \neq \lambda_1$ such that $\operatorname{Re}\lambda > -\mu$ and $\Delta(\lambda)=1$. We claim that there can be at most only finitely many values λ such that $\operatorname{Re}\lambda > -\mu/2$ and $\Delta(\lambda)=1$. Assume to the contrary that there exists an infinite sequence $\{z_k\}$ such that $\operatorname{Re} z_k > -\mu/2$ and $\Delta(z_k)=1$ for $k=1,2,\ldots$. Since $1-\Delta(\lambda)$ is holomorphic for $\operatorname{Re}\lambda > -\mu/2$, its zeroes cannot accumulate in this region. Thus, $\lim_{k\to\infty} |\operatorname{Im} z_k| = \infty$. Observe that for $k=1,2,\ldots$.

$$1=\int_0^\infty \int_0^\infty e^{-\operatorname{Re} z_k a}\cos(\operatorname{Im} z_k a)\,\gamma(b)\,\Gamma(a,b)\,da\,db$$

$$<\int_0^\infty \cos(\operatorname{Im} z_k a)\left[e^{-\mu a/2}\int_0^\infty \gamma(b)\,\beta(a,b)\,db\right]da.$$

But by the Riemann-Lebesgue Theorem (see [40], p. 90), this last expression must approach 0, which is a contradiction. Thus, there exists ω such that $\sup_{\lambda\in\sigma(A)-E\sigma(A)-\{\lambda_1\}} \operatorname{Re}\lambda < \omega < \lambda_1$. The proof of (6.2) will follow from Lemma 5.2 and Theorem 5.1 if we can show that λ_1 is a simple pole of $(1-\Delta(\lambda))^{-1}$. That λ_1 is a simple pole follows from the Residue Theorem (see [42], p. 215) and the fact that the residue of $(1-\Delta(\lambda))^{-1}$ at λ_1 is $1/M_{\lambda_1}$. From Lemma 5.2 we obtain that for $\phi\in X$, $t\geqq 0$

$$\|e^{-\lambda_1 t}T(t)\phi - P_0\phi\|_X = \|e^{-\lambda_1 t}P_1 T(t)\phi\|_X \leqq e^{-(\lambda_1-\omega)t}M_1|P_1|\,\|\phi\|_X.$$

Finally, the formula P_0 in (1.10) follows immediately from (5.7), (3.8), and (5.12). □

In order to prove convergence to stable distributions of the solutions to problem (1.11)-(1.14) we require

$$\nu\colon [0,\infty)\to[0,\infty),\ \nu \text{ is continuously differentiable, } \nu' \text{ is bounded, and there exist positive constants } \underline{\nu},\ \bar{\nu} \text{ such that } \underline{\nu}\leqq\nu(c)\leqq\bar{\nu} \text{ for } c\geqq 0, \tag{6.4}$$

$$m\in X_+^\infty \text{ and } \operatorname*{ess\text{-}inf}_{a,c>0}[m(a,c)+\nu'(c)]\equiv \mathbf{m}>0, \tag{6.5}$$

$$H \text{ and } K \text{ are bounded linear operators from } X \text{ to } L^1, \tag{6.6}$$

$$\psi\in X. \tag{6.7}$$

Let $\varkappa\colon[0,\infty)\to[0,\infty)$ be the unique solution of the problem $\varkappa'(b)=\nu(\varkappa(b))$, $b\geqq 0$, $\varkappa(0)=0$. Notice that by (6.4) $\varkappa$ is strictly increasing, $\lim_{b\to\infty}\varkappa(b)=\infty$, $\varkappa^{-1}$ exists, and $\nu(c)\varkappa^{-1\prime}(c)=1$ for all $c\geqq 0$. Let J be the bounded linear operator in X defined by $(J\psi)(a,b)=\psi(a,z(b))$, $\psi\in X$, a.e. $a>0$, a.e. $b>0$. Notice that J^{-1} exists and $(J^{-1}\phi)(a,c)=\phi(a,z^{-1}(c))$, $\phi\in X$, a.e. $a>0$, a.e. $c>0$ (for a proof that J and J^{-1} are bounded linear operators in X see [2]). Define $\mu\in X_+^\infty$ by $\mu(a,b)=m(a,\varkappa(b))+\nu'(\varkappa(b))$, a.e. $a>0$, a.e. $b>0$. Lastly, define the bounded linear

operators F and G from X to L^1 by $(F\phi)(a)=(HJ^{-1}\phi)(a)$, $\phi\in X$, a.e. $a>0$, and $(G\phi)(b)=(KJ^{-1}\phi)(\varkappa(b))$, $\phi\in X$, a.e. $b>0$. This μ, F, and G satisfy the hypotheses of Theorems 2.1, 2.2, 3.1, 3.2, 3.4, 3.5, 4.1, and 4.2. Let $T(t)$, $t\geqq 0$ be the strongly continuous semigroup for this problem. Define the strongly continuous semigroup $S(t)$, $t\geqq 0$ in X by the formula $S(t)=J^{-1}T(t)J$, $t\geqq 0$. We claim that $S(t)$, $t\geqq 0$ is the strongly continuous semigroup associated with the problem (1.11)–(1.14).

Remark 6.1. It can be shown that the infinitesimal generator B of $S(t)$, $t\geqq 0$ is

$$B\psi(a,c)=-d/da\left[\nu(c)\,d/dc\int_0^a\psi(\hat{a},c)\,d\hat{a}+\psi(a,c)\right]-[m(a,c)+\nu'(c)]\,\psi(a,c),$$

a.e. $a>0$, a.e. $c>0$,

$$D(B)=\Bigg\{\psi\in X:(a,c)\mapsto\int_0^a\psi(\hat{a},c)\,d\hat{a}\text{ is absolutely continuous in }c\text{ for }a\geqq 0,$$

$$(a,c)\mapsto\left[\nu(c)\,d/dc\int_0^a\psi(\hat{a},c)\,d\hat{a}+\psi(a,c)\right]\text{ is absolutely continuous in }a$$
for a.e. $c>0$,

$$\lim_{a\to 0^+}\left[\nu(c)\,d/dc\int_0^a\psi(\hat{a},c)\,d\hat{a}+\psi(a,c)\right]=K(\psi)(c)\text{ for a.e. }c>0,$$

$$(a,c)\mapsto\int_0^c\psi(a,\hat{c})\,d\hat{c}\text{ is absolutely continuous in }a\text{ for }c\geqq 0,$$

$$(a,c)\mapsto\left[d/da\int_0^c\psi(a,\hat{c})\,\nu(\hat{c})^{-1}\,d\hat{c}+\psi(a,c)\right]\text{ is absolutely continuous in }c$$
for a.e. $a>0$,

$$\lim_{c\to 0^+}\left[d/da\int_0^c\psi(a,\hat{c})\,\nu(\hat{c})^{-1}\,d\hat{c}+\psi(a,c)\right]=H(\psi)(a)\text{ for a.e. }a>0,$$

$$\text{and }d/da\left[\nu(c)\,d/dc\int_0^a\psi(\hat{a},c)\,d\hat{a}+\psi(a,c)\right]\in X\Bigg\}.$$

If $\psi\in D(B)$ and ψ is sufficiently smooth, then $B\psi=-(\psi_a+(\nu\psi)_c+m\psi)$, $\psi(a,0)=H\psi(a)$, $a\geqq 0$, $\psi(0,c)=K\psi(c)$, $c\geqq 0$.

We require the following hypotheses on the input functions H and K:

$$H\equiv 0\text{ and for all }\psi\in X,\text{ a.e. }c>0,\ K\psi(c)=\eta(c)\int_0^\infty\int_0^\infty\zeta(\hat{a},\hat{c})\,\psi(\hat{a},\hat{c})\,d\hat{a}\,d\hat{c},$$
$$\text{where }\eta\in L^\infty_+\cap L^1,\ \eta\not\equiv 0,\text{ and }\zeta\in X^\infty_+,\ \zeta(a,c)=0\text{ for }a>z^{-1}(c),\ \zeta\not\equiv 0. \tag{6.8}$$

Define for $a,c,t\geqq 0$ $\Lambda(a,c,t)=\exp\left[-\int_0^t m(a-\tau,\varkappa(\varkappa^{-1}(c)-\tau))\,d\tau\right]\nu(\varkappa(\varkappa^{-1}(c)-t))/\nu(c)$. Notice that for $a,c,t\geqq 0$, $b=\varkappa^{-1}(c)$

$$\Lambda(a,c,t)=\exp\left[-\int_{\varkappa(\varkappa^{-1}(c)-t)}^{c} m(a-\varkappa^{-1}(c)+\varkappa^{-1}(\hat{c}),\hat{c})/\nu(\hat{c})\,d\hat{c}\right]$$
$$\exp\left[-\int_{\varkappa(\varkappa^{-1}(c)-t)}^{c} d/d\hat{c}\log(\nu(\hat{c}))\,d\hat{c}\right]$$
(by the change of variable $\tau=\varkappa^{-1}(c)-\varkappa^{-1}(\hat{c})$, $d\tau=-d\hat{c}/\nu(\hat{c})$)
$$=\exp\left[-\int_{\varkappa(b-t)}^{\varkappa(b)}(m(a-b+\varkappa^{-1}(\hat{c}),\hat{c})+\nu'(\hat{c}))/\nu(\hat{c})\,d\hat{c}\right]$$
$$=\exp\left[-\int_0^t(m(a-\tau,\varkappa(b-\tau))+\nu'(\varkappa(b-\tau))\,d\tau\right]$$
(by the change of variable $\tau=b-\varkappa^{-1}(\hat{c})$, $d\tau=-d\hat{c}/\nu(\hat{c})$)
$$=\exp\left[-\int_0^t\mu(a-\tau,b-\tau)\,d\tau\right]=\Pi(a,\varkappa^{-1}(c),t).$$

Define the quantities

$$\chi(a,c)=\zeta(a,\varkappa(a+c))\,\Lambda(a,\varkappa(a+c),a)\,\nu(z(a+c)),\ a,c\geqq 0,$$
$$\Omega(\lambda)=\int_0^\infty\int_0^\infty e^{-\lambda a}\eta(\varkappa(c))\,\chi(a,c)\,da\,dc,\ \mathrm{Re}\,\lambda>-\mathbf{m},$$
$$W_\lambda(\psi)=\int_0^\infty\int_0^\infty\zeta(a,c)\left[\int_0^a e^{-\lambda\sigma}\Lambda(a,c,\sigma)\,\psi(a-\sigma,c-\sigma)\,d\sigma\right]da\,dc,\ \mathrm{Re}\,\lambda>-\mathbf{m},\quad \psi\in X,$$
$$N_\lambda=\int_0^\infty\int_0^\infty a e^{-\lambda a}\eta(\varkappa(c))\,\chi(a,c)\,da\,dc,\ \mathrm{Re}\,\lambda>-\mathbf{m}.$$

Notice from (6.8) that for $\phi\in X$, a.e. $b>0$, $c=\varkappa(b)$,

$$(G\phi)(b)=(KJ^{-1}\phi)(c)=\eta(c)\int_0^\infty\int_0^\infty\zeta(\hat{a},\hat{c})\,\phi(\hat{a},\varkappa^{-1}(\hat{c}))\,d\hat{a}\,d\hat{c}$$
$$=\eta(c)\int_0^\infty\int_0^\infty\zeta(\hat{a},\varkappa(\hat{b}))\,\phi(\hat{a},\hat{b})\,\nu(\varkappa(\hat{b}))\,d\hat{a}\,d\hat{b}.$$

Define $\gamma(b)=\eta(\varkappa(b))$, a.e. $b>0$, and $\beta(a,b)=\zeta(a,\varkappa(b))\,\nu(\varkappa(b))$, a.e. $a>0$, a.e. $b>0$, and observe that G satisfies (1.9) for this γ and β. It is easily verified that for $\psi\in X$, $\mathrm{Re}\,\lambda>-\mathbf{m}$, $W_\lambda(\psi)=V_\lambda(J\psi)$, and for $\mathrm{Re}\,\lambda>-\mathbf{m}$, $N_\lambda=M_\lambda$. Also, from (1.10) we have that for $\psi\in X$, $a\geqq 0$, a.e. $c,\varkappa^{-1}(c)>a$

$$(J^{-1}P_0J\psi)(a,c)=(P_0J\psi)(a,\varkappa^{-1}(c))$$
$$=e^{-\lambda_1 a}\Pi(a,\varkappa^{-1}(c),a)\,\gamma(\varkappa^{-1}(c)-a)\,V_{\lambda_1}(J\psi)\,M_{\lambda_1}^{-1}$$
$$=e^{-\lambda_1 a}\Lambda(a,c,a)\,\eta(\varkappa(\varkappa^{-1}(c)-a))\,W_{\lambda_1}(\psi)\,N_{\lambda_1}^{-1}.$$

Lastly, we claim that for $\mathrm{Re}\,\lambda>-\mathbf{m}$, $\Omega(\lambda)=\Delta(\lambda)$, since

$$\Omega(\lambda)=\int_0^\infty\int_0^\infty e^{-\lambda a}\eta(\varkappa(c))\,\zeta(a,\varkappa(a+c))\,\Lambda(a,\varkappa(a+c),a)\,\upsilon(\varkappa(a+c))\,dc\,da$$

$$=\int_0^\infty\int_{\varkappa(a)}^\infty e^{-\lambda a}\eta(\varkappa(\varkappa^{-1}(\hat{c})-a))\,\zeta(a,\hat{c})\,\Lambda(a,\hat{c},a)\,d\hat{c}\,da$$

(by the change of variable $\hat{c}=\varkappa(c+a)$, $d\hat{c}=\upsilon(\varkappa(c+a))\,dc$)

$$=\int_0^\infty\int_0^{\varkappa^{-1}(\hat{c})} e^{-\lambda a}\eta(\varkappa(\varkappa^{-1}(\hat{c})-a))\,\zeta(a,\hat{c})\,\Pi(a,\varkappa^{-1}(\hat{c}),a)\,da\,d\hat{c}$$

$$=\int_0^\infty\int_0^{\hat{b}} e^{-\lambda a}\eta(\varkappa(\hat{b}-a))\,\zeta(a,\varkappa(\hat{b}))\,\Pi(a,\hat{b},a)\,\upsilon(\varkappa(\hat{b}))\,da\,d\hat{b}$$

$$=\int_0^\infty\int_0^{\hat{b}} e^{-\lambda a}\gamma(\hat{b}-a)\,\beta(a,\hat{b})\,\Pi(a,\hat{b},a)\,da\,d\hat{b}=\Delta(\lambda).$$

The characteristic equation of the problem (1.11)–(1.14) is $\Omega(\lambda)=1$. For the unique real value λ_1 such that $\Omega(\lambda_1)=\Delta(\lambda_1)=1$ define the projection $Q_0: X\to X$ by

$$(Q_0\psi)(a,c)=\begin{cases}0, & a>\varkappa^{-1}(c)\\ e^{-\lambda_1 a}\Lambda(a,c,a)\,\eta(\varkappa(\varkappa^{-1}(c)-a))\,W_{\lambda_1}(\psi)\,N_{\lambda_1}^{-1}, & \text{a.e. } c,\ \varkappa^{-1}(c)>a.\end{cases}$$

Theorem 6.8. *Let* (6.4), (6.5), (6.8) *hold. The following are true:*

If $\Omega(0)<1$, *then there exists constants* $M>0$ *and* $\omega<0$ *such that* $\|S(t)\psi\|_X\leqq Me^{\omega t}\|\psi\|_X$ *for* $\psi\in X$, $t\geqq 0$. (6.9)

If $\Omega(0)\geqq 1$, *then there exists a unique real* $\lambda_1\geqq 0$ *such that* $\Omega(\lambda_1)=1$ *and there exist constants* $M>0$ *and* $\omega<0$ *such that for all* $\psi\in X$, $t\geqq 0$, $\|e^{-\lambda_1 t}S(t)\psi-Q_0\psi\|_X\leqq Me^{\omega t}\|\psi\|_X$. (6.10)

Proof. Since $S(t)=J^{-1}T(t)J$ and $Q_0=J^{-1}P_0J$, the proof of (6.9) and (6.10) follows immediately from Theorem 6.1 and the formulas above. □

References

1. Aldenberg, T., Diekmann, O., Lauwerier, H., Metz, J.: Growth, fission, and the stable size distribution. J. Math. Biol. **18**, 135–148 (1983)
2. Badii, M., Webb, G.F.: Representation of solutions of functional differential equations with time-dependent delays. Houston J. Math. **6**, 577–594 (1980)
3. Bell, G.I., Anderson, E.C.: Cell growth and division. I.A. mathematical model with applications to cell volume distributions in mammalian suspension cultures. Biophys. J. **7**, 329–351 (1967)
4. Busenberg, S., Iannelli, M.: A class of nonlinear diffusion problems in age-dependent population dynamics. J. Nonlinear Analysis T. M. A. (to appear)
5. Chipot, M.: On the equations of age-dependent population dynamics. Arch. Rational Mech. Anal. **82**, 13–25 (1983)
6. Coale, A.J.: The Growth and Structure of Human Populations. Princeton: Princeton University Press, 1972
7. Cushing, J.M.: Model stability and instability in age structured populations. J. Theoret. Biol. **86**, 709–730 (1980)

8. Cushing, J.M.: Bifurcation of time periodic solutions of the McKendrick equations with applications to population dynamics. Internat. J. Comput. Math. Appl. **9**, 459–478 (1983)
9. Derndinger, R.: Über das Spektrum positiver Generatoren. Math. Zeit. **172**, 281–293 (1980)
10. Di Blasio, G.: Asymptotic behavior of an age-structured fish population. Internat. J. Comput. Math. Appl. **9**, 377–381 (1983)
11. Di Blasio, G., Iannelli, M., Sinestrari, E.: Approach to equilibrium in age structured populations with an increasing recruitment process. J. Math. Biol. **13**, 371–382 (1982)
12. Diekmann, O.: The stable size distribution: An example in structured population dynamics. Math. Centrum Report TW 231, Amsterdam, 1983
13. Diekmann, O., Heijmans, H., Thieme, H.: On the stability of the size distribution. Math. Centrum Report TW 242, Amsterdam, 1983
14. Dunford, N., Schwartz, J.: Linear Operators. Part I: General Theory. New York: Interscience 1958
15. Elderkin, R.H.: Seed dispersal in a patchy environment with global age dependence. J. Math. Biol. **13**, 283–303 (1982)
16. Feller, W.: On the integral equation of renewal theory. Ann. Math. Stat. **12**, 243–267 (1941)
17. Greiner, G.: A typical Perron-Frobenius theorem with application to an age-dependent population equation. Infinite-Dimensional Systems, Proceedings, Retzhof 1983 (F. Kappel and W. Schappacher, Eds.) 86–100. Berlin-Heidelberg-New York-Tokyo: Springer 1984
18. Greiner, G., Nagel, R.: On the stability of strongly continuous semigroups of positive operators on $L^2(\mu)$. Ann. Scuola Norm. Sup. Pisa, Serie IV-Vol. X, n. 2, 257–262 (1983)
19. Greiner, G., Voigt, J., Wolff, M.: On the spectral bound of the generator of semigroups of positive operators. J. Operator Theory **5**, 245–256 (1981)
20. Grippenberg, G.: Stability analysis of a distributed parameter model for the growth of microorganisms. Internat. J. Comput. Appl. Math. **9**, 431–442 (1983)
21. Groh, U., Neubrander, F.: Stabilität starkstetiger, positiver Operatorhalbgruppen auf C^*-Algebren. Math. Ann. **256**, 509–516 (1981)
22. Gurtin, M.E.: The mathematical Theory of Age-Structured Populations (To appear)
23. Gurtin, M.E., MacCamy, R.C.: Nonlinear age-dependent population dynamics. Arch. Rational Mech. Anal. **54**, 281–300 (1974)
24. Gyllenberg, M.: Nonlinear age-dependent population dynamics in continuously propagated bacterial cultures. Math. Biosci. **62**, 45–74 (1982)
25. Hadeler, K., Dietz, K.: Nonlinear hyperbolic partial differential equations for the dynamics of parasite populations. Internat. J. Comput. Math. Appl. **9**, 415–430 (1981)
26. Heijmans, H.J.A.M.: An eigenvalue problem related to cell growth. Math. Centrum Report TW 229, Amsterdam, 1982
27. Hoppensteadt, F.: Mathematical Theories of Populations: Demographics, Genetics, and Epidemics. SIAM Reg. Conf. Series in Appl. Math., 1975
28. Keyfitz, N.: Introduction to the Mathematics of Population. Reading: Addison-Wesley, 1968
29. Martin, R.H.: Nonlinear Operators and Differential Equations in Banach Space. New York: John Wiley and Sons 1976
30. McKendrick, A.G.: Applications of mathematics to medical problems. Proc. Edinburgh Math. Soc. **44**, 98–130 (1926)
31. Murphy, L.: Density dependent cellular growth in an age structured colony, Internat. J. Comput. Appl. Math. **9**, 383–392 (1983)
32. Nussbaum, R.D.: The radius of the essential spectrum. Duke Math. J. **38**, 473–478 (1970)
33. Oster, G.: Internal variables in population dynamics. Some Mathematical Questions in Biology. VII, Amer. Math. Soc., Providence. (1976), 37–68
34. Oster, G.: Lectures in population dynamics. Lectures in Applied Mathematics, American Mathematical Society **16**, 149–170 (1977)
35. Pollard, J.H.: Mathematical Models for the Growth of Human Populations. Cambridge: University Press 1973
36. Prüss, J.: Equilibrium solutions of age-specific population dynamics of several species. J. Math. Biol. **11**, 65–84 (1981)
37. Prüss, J.: On the qualitative behavior of population with age-specific interactions. Internat. J. Comput. Appl. Math. **9**, 327–340 (1983)
38. Prüss, J.: Stability analysis for equilibria in age-specific population dynamics. (To appear)

39. Prüss, J.: On the spectrum of C_0-semigroups. (To appear)
40. Royden, H.L.: Real Analysis. Second Edition. New York: Macmillan, 1968
41. Rubinov, S.: Age-structured equations in the theory of cell populations. Studies in Mathematical Biology, Part II, Math. Assoc. Amer., Washington, 1978
42. Rudin, W.: Real and Complex Analysis. New York: McGraw-Hill 1966
43. Sharpe, F.R., Lotka, A.J.: A problem in age distributions. Phil. Mag. **21**, 435–438 (1911)
44. Sinestrari, E.: Non-linear age-dependent population growth. J. Math. Biol. **128**, 1–15 (1980)
45. Sinestrari, E., Webb, G.: Hyperbolic systems with non-local boundary conditions with applications to epidemic models. (To appear)
46. Sinko, J.W., Streifer, W.: A new model for age-size structure of a population. Ecology **48**, 910–918 (1967)
47. Song, J., Yu, J.-Y., Hu S.-J., Zhu G.-T.: Spectral properties of population operators and asymptotic behavior of population semigroups. Acta Mathematica Scientia **2**, 139–148 (1982)
48. Triggiani, R.: On the stabilizability problem in Banach space. J. Math. Anal. Appl. **52**, 383–403 (1975)
49. Von Foerster, H.: Some remarks on changing populations. The Kinetics of Cellular Proliferation, 382–407. New York-London: Grune and Stratton 1959
50. Webb, G.F.: The semigroup associated with nonlinear age dependent population dynamics. Internat. J. Comput. Math. Appl. **9**, 487–498 (1983)
51. Webb, G.F.: Theory of Nonlinear Age-Dependent Population Dynamics. Monographs and Textbooks in Pure and Applied Mathematics Series, Vol. 89. New York-Basel: Marcel Dekker, Inc. 1985
52. Webb, G.F.: A Semigroup proof of the Sharpe-Lotka Theorem. Infinite-Dimensional Systems, Proceedings, Retzhof 1983 (F. Kappel and W. Schappacher, Eds.), 254–268. Berlin-Heidelberg-New York-Tokyo: Springer 1984
53. Yosida, K.: Functional Analysis. Second Edition. Berlin-Heidelberg-New York: Springer 1968

The author expresses his thanks to the referee for helpful suggestions.

Received June 5, 1984

Math. Z. 189, 337-341 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

On The Values of the Dedekind Sum

G. Myerson

Department of Mathematics, State University of New York at Buffalo, Buffalo, New York 14214, USA

1. Introduction and Statement of Main Results

The Dedekind sum $s(h, k)$ is defined for integer h and positive integer k by

$$s(h, k)=\sum_{j=0}^{k-1}((j/k))((jh/k)),$$

where $((x))=0$ if x is an integer, $((x))=x-[x]-\frac{1}{2}$ otherwise.

Historically, the Dedekind sums first arose in the study of the transformation formula for the Dedekind eta-function. Rademacher began the study by elementary means of the arithmetic properties of the Dedekind sums. More recently, the sums have found application in the study of Lehmer's linear congruential generators of pseudo-random numbers - see, for example, Knuth [3], especially pp. 78-87. The book of Rademacher and Grosswald [5] is an excellent reference for the history of the sums and for many of the known facts about them which we will need below.

We define $t(h, k)=6ks(h, k)$. It is known that $t(h, k)$ is an integer for all h and k. The problem we address is that of characterizing, for a given integer t, all pairs (h, k) such that $t(h, k)=t$.

Since $t(-h, k)=-t(h, k)$, it is sufficient to consider the cases $t\geqq 0$. It is known that, for any positive integer a, $s(ah, ak)=s(h, k)$, whence $t(ah, ak)=at(h, k)$; from here on, we assume h and k are relatively prime. With these understandings, our main results are the following two theorems.

Theorem 1. *Given integers $h, k, k>0$, let $h'=2t(h, k)-h$, $k'=(hh'-1)/k$, $Q(x, y)=ky^2+(h-h')xy-k'x^2$, and $B(u, v; x, y)=kvy+\frac{1}{2}(h-h')(uy+vx)-k'ux$. If u, v, x and y are integers, and $uy-vx=1$, then $Q(x, y)$ and $B(u, v; x, y)$ are integers; if, moreover $Q(x, y)>0$, then*

$$t(t(h, k)-B(u, v; x, y), Q(x, y))=t(h, k).$$

Theorem 2. *Given an integer t, there exists a finite, computable (and possibly empty) set Q_t of binary quadratic forms of discriminant $4t^2-4$ such that $t(h, k)$*

$=t$ if and only if there exist Q in Q_t and u, v, x, y in $\mathbf{Z}$ with $uy-vx=1$, $k=Q(x, y)$, and $h=t-B(u, v; x, y)$; here, B is the symmetric bilinear form such that $B(x, y; x, y)=Q(x, y)$.

An algorithm for computing the set Q_t of Theorem 2 is given in Sect. 3, along with applications to special cases, and a discussion of previously known results.

2. Proofs of the Main Results

We will need the most outstanding fact about the Dedekind sums, the well-known reciprocity law:

Theorem 3. $s(h, k)+s(k, h)=-\frac{1}{4}+\frac{1}{12}\left(\frac{h}{k}+\frac{1}{hk}+\frac{k}{h}\right)$.

Several proofs will be found in [5]. A simple consequence of reciprocity and of the integrality of $t(h, k)$ is

Lemma 1. *Let h, k, h' and k' be as in Theorem 1. Then k' is an integer.*

Proof. From the definitions of k' and h' we have $k'=(2ht(h, k)-(h^2+1))/k$. From the reciprocity law we obtain, on multiplying through by $12hk$, $2ht(h, k)-(h^2+1)=k^2-3hk-2kt(k, h)$. Thus $k'=k-3h-2t(k, h)$ is an integer.

Our original approach to the problem addressed in this paper involved the relation between Dedekind sums and continued fractions, as exploited by Hickerson in [2]. We present here a more succinct approach, suggested by a referee, which turns on the class-invariance of Rademacher's ψ-function.

Let $\gamma=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ belong to Γ, the group of 2-by-2 integer matrices with determinant 1, and assume $c\neq 0$. Define the matrix function ϕ by

$$c\phi(\gamma)=\operatorname{Trace}\gamma-2t(d, |c|), \tag{1}$$

and define ψ by

$$\psi(\gamma)=\phi(\gamma)-3\operatorname{sgn}(c\operatorname{Trace}\gamma). \tag{2}$$

Rademacher proved

Lemma 2. *ψ is a class-invariant.*

A proof is found in [5].

Proof of Theorem 1. Let h, k, h', k', Q, and B be as given. Let $g=\begin{pmatrix} h' & k' \\ k & h \end{pmatrix}$; then g is in Γ, by Lemma 1 and the definitions. Let $\gamma=\begin{pmatrix} u & v \\ x & y \end{pmatrix}$ be in Γ, and let $\gamma g\gamma^{-1}=\begin{pmatrix} H' & K' \\ K & H \end{pmatrix}$. A routine calculation shows that $K=Q(x, y)$ and $H=t(h, k)-B(u, v; x, y)$. The integrality of Q and B is clear, so it suffices to show that $t(H, K)=t(h, k)$.

By Lemma 2 we have $\psi(\gamma g \gamma^{-1}) = \psi(g)$. By (2), this is

$$\phi(\gamma g \gamma^{-1}) - 3\,\mathrm{sgn}(K \operatorname{trace} \gamma g \gamma^{-1}) = \phi(g) - 3\,\mathrm{sgn}(k \operatorname{Trace} g).$$

Since Trace is a class invariant, and since by assumption $\mathrm{sgn}\, k = \mathrm{sgn}\, K = 1$, we deduce $\phi(\gamma g \gamma^{-1}) = \phi(g)$. By (1), this is

$$k(\operatorname{Trace} \gamma g \gamma^{-1} - 2t(H, K)) = K(\operatorname{Trace} g - 2t(h, k)).$$

But $\operatorname{Trace} g = h' + h = 2t(h, k)$, so $2t(H, K) = \operatorname{Trace} \gamma g \gamma^{-1} = \operatorname{Trace} g = 2t(h, k)$.

Proof of Theorem 2. If there do not exist integers h, k such that $t(h, k) = t$ then the theorem is vacuously true. For each pair h, k for which $t(h, k) = t$, let $Q_{h,k}$ be the form given in Theorem 1. Each of these forms has the same discriminant, namely, $4t^2 - 4$, since $(h - h')^2 + 4kk' = (h - h')^2 + 4hh' - 4 = (h + h')^2 - 4 = 4t^2 - 4$. Choose from among these forms a maximal set Q_t of mutually inequivalent forms; Q_t is finite by the finiteness of the class number. By its construction and by Theorem 1, Q_t has the properties claimed for it in Theorem 2.

3. Applications and Discussion of Previously Known Results

We wish to apply Theorem 2 to the problem of characterizing, for given small integers t, the solutions of $t(h, k) = t$. To compute the set Q_t of Theorem 2, we apply the following theorem.

Theorem 4. *Let $Q(x, y) = ay^2 + 2bxy + cx^2$ be a binary quadratic form with discriminant $4t^2 - 4$ and $a > 0$. Then $Q(x, y)$ is equivalent to some element of Q_t if and only if $t(t - b, a) = t$.*

Proof. If $Q(x, y)$ is equivalent to some element of Q_t then by Theorem 2, $t(t - B(1, 0; 0, 1), Q(0, 1)) = t$ where $B(u, v; x, y) = avy + b(uy + vx) + cux$; thus $t(t - b, a) = t$. Conversely, assume $t(t - b, a) = t$. In Theorem 1, let $h = t - b$, $k = a$; then $h' = t + b$, $k' = (hh' - 1)/k = (t^2 - b^2 - 1)/a = -c$ (since $t^2 - 1 = b^2 - ac$), so some form equivalent to $ky^2 + (h - h')xy - k'x^2 = ay^2 - 2bxy + cx^2$ is in Q_t. But this form is equivalent to the given form $Q(x, y)$.

We now apply our results to the cases $t = 0$, 1, and 3.

Let $t = 0$. Rademacher [4] proved that $s(h, k) = 0$ (hence, $t(h, k) = 0$) if and only if k divides $h^2 + 1$. We recover this result, as follows. If $t = 0$, then $4t^2 - 4 = -4$; there is only one equivalence class of quadratic forms of discriminant -4, and we take $Q(x, y) = y^2 + x^2$ as a representative. By Theorem 4, $Q(x, y)$ is in Q_0 since $t(0, 1) = 0$. Thus, by Theorem 2, $t(h, k) = 0$ if and only if $h = -vy - ux$, $k = y^2 + x^2$, with $uy - vx = 1$. This is easily seen to be a disguised version of Rademacher's criterion.

Let $t = 1$. Salié [6] asserted that $t(h, k) = 1$ only if $k = 3y^2$ for some integer y, but it is not evident from the paper how he came to this conclusion. We note that Salié stated all of his results in terms of $D(h, k)$, where $D(h, k) = 2t(h, k)$. It follows from formulas (23) and (48) of [6] that if d and y are relatively prime then $t(3dy + 1, 3y^2) = 1$; this also follows from Theorem 5 of Dieter [1]. We characterize all solutions of $t(h, k) = 1$ in the following theorem.

Theorem 5. *$t(h, k)=1$ if and only if there exist relatively prime integers d, y such that $h=3dy+1$, $k=3y^2$.*

Proof. If $t=1$, then $4t^2-4=0$, and the binary forms of Theorem 2 are necessarily degenerate. We may assume $Q(x, y)=ay^2$ for some $a>0$. By Theorem 4, $Q(x, y)$ is in Q_1 if and only if $t(1, a)=1$. From the definitions, we calculate $s(1, a)=(a^2-3a+2)/12a$, whence $t(1, a)=(a-1)(a-2)/2$; it follows that $t(1, a)=1$ if and only if $a=3$, so Q_1 contains only the form $Q(x, y)=3y^2$. Thus $t(h, k)=1$ if and only if $h=1-3vy$, $k=3y^2$, $(v, y)=1$. Letting $d=-v$ proves the theorem.

Remark. This result clearly includes all previously known results for $t=1$.

Let $t=3$. Salié [6] proved that $t(8y+3, 8y^2-1)=3$. We characterize all solutions of $t(h, k)=3$ in the following theorem.

Theorem 6. *$t(h, k)=3$ if and only if there exist integers u, v, x and y with $uy-vx=1$ such that $h=ux-8vy+3$ and $k=8y^2-x^2>0$.*

Remark. The special case $u=0$, $v=-1$, $x=1$ yields Salié's result.

Proof. If $t=3$, then $4t^2-4=32$. An investigation of the quadratic forms of discriminant 32 shows that the only member of Q_3 must be equivalent to $Q(x, y)=8y^2-x^2$. The theorem now follows from Theorem 2.

The most general previous result on the values of Dedekind sums is the following theorem of Salié (Satz 6 in [6]), which we recover as a special case of Theorem 1.

Theorem 7. *For any integer y, $t(ky+h, [(ky+h)^2-2t(h, k)(ky+h)+1]/k)=t(h, k)$.*

Proof. In Theorem 1, let $(u, v, x, y)=(0, -1, 1, y)$; eliminating h' and k' from the resulting expression using $h'=2t(h, k)-h$ and $k'=(hh'-1)/k$ yields the result.

Salié [6] shows that $t(h, k)=t$ implies $t\equiv 0, \pm 1$, or $\pm 3 \pmod 9$. We have computed Q_t for t satisfying these congruences, $0\leqq t\leqq 15$.

Table 1

$Q_0=\{x^2+y^2\}$	$Q_1=\{3y^2\}$	$Q_3=\{4y^2-4xy-x^2\}$
$Q_6=\{5y^2-10xy-2x^2\}$	$Q_8=\{9y^2-12xy-3x^2\}$	$Q_9=\{11y^2-12xy-4x^2\}$
$Q_{10}=\{6y^2-18xy-3x^2\}$	Q_{12} is empty	$Q_{15}=\{7y^2-28xy-4x^2,$ $11y^2-26xy-5x^2\}$

Salié asserted without proof that there are precisely four values of t satisfying the congruences and $0\leqq t\leqq 150$ such that $t(h, k)=t$ is impossible, namely, $t=12$, 17, 44, and 107. We are planning a computer project to verify Salié's assertion and to extend the result to a larger range of values of t.

We would like to thank Professor D.J. Lewis for his helpful suggestions concerning the organization of this paper.

References

1. Dieter, U.: Beziehungen zwischen Dedekindschen Summen. Abh. Math. Semin. Univ. Hamburg **21**, 109–125 (1957)
2. Hickerson, D.: Continued fractions and density results for Dedekind sums. J. Reine Angew. Math. **290**, 113–116 (1977)
3. Knuth, D.: Seminumerical Algorithms, 2nd ed. Reading: Addison Wesley 1981
4. Rademacher, H.: Zur Theorie der Dedekindschen Summen. Math. Z. **63**, 445–463 (1956)
5. Rademacher, H., Grosswald, E.: Dedekind Sums. Carus Math. Monograph **16,** (1972)
6. Salié, H.: Zum Wertevorrat der Dedekindschen Summen. Math. Z. **72**, 61–75 (1959/60)

Received February 10, 1984

Math. Z. 189, 343-363 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

A Characterization of Complex Manifolds Biholomorphic to a Circular Domain

Giorgio Patrizio

Sonderforschungsbereich 40, Theoretische Mathematik, Beringstr. 4, D-5300 Bonn 1
New address: Dipartimento di Matematica II, Università di Roma, I-00173 Roma, Italy

Dedicated to Wilhelm Stoll for his sixtieth birthday

Introduction

W. Stoll [11], developing earlier ideas of Griffiths and King [5], introduced in Value Distribution Theory the concept of parabolic exhaustion. This notion has recently found important applications in the classification of domains in $\mathbb{C}^m$ and in general of complex manifolds. A *strictly parabolic exhaustion* on a non compact complex manifold M of dimension m is a C^∞ exhaustion function $\tau: M \to [0, \Delta^2), 0 < \Delta = \sup \sqrt{\tau} \leqq \infty$, such that $dd^c\tau > 0$ on M and $dd^c \log \tau \geqq 0$, $(dd^c \log \tau)^m \equiv 0$ on $M - \tau^{-1}(0)$. Using the information that τ yields and most noteworthy the existence of a Monge-Ampère foliation associated to τ (cfr. [1]), Stoll in [12] shows that there exists a biholomorphic map $h: \mathbb{B}(\Delta) \to M$, where $\mathbb{B}(\Delta) = \{Z \in \mathbb{C}^m \mid \|Z\| < \Delta\}$, such that $\tau \circ h(Z) = \|Z\|^2$ (cfr. also [4] and [14] for other proofs). The properties of the exhaustion τ on $\tau^{-1}(0)$ play a decisive role in the classification. In fact, if we relax the differentiability assumptions on $\tau^{-1}(0)$, even assuming that this set reduces to a single point (which is always the case if τ is strictly parabolic), the range of possibilities for the underlying manifold is much wider. If $G \Subset \mathbb{C}^m$ is a smooth, strictly pseudoconvex, complete circular domain and σ is its Minkowski functional squared, then σ is strictly parabolic on $G - \{0\}$ but it is smooth on the origin if and only if G is biholomorphic to the ball. On the other hand the space of biholomorphic classes of such domains is infinite dimensional (cfr. [10]). In this class of examples the Monge-Ampère foliation associated to the exhaustion is always holomorphic. This is also quite an exceptional occurrence. Let $D \Subset \mathbb{C}^m$ be a smooth, strictly convex domain and $p \in D$ be any point. There exists an exhaustion $\tau_p: \bar{D} \to [0, 1]$ such that $\tau_p^{-1}(0) = \{p\}$, $\tau | \partial D \equiv 1$ and that τ_p is strictly parabolic on $D - \{p\}$ (cfr. [6] and [9]). For a generic strictly convex domain the Monge-Ampère foliation associated to τ_p is *not* holomorphic (cfr. [9]).

In this paper we study complex manifolds M admitting an exhaustion $\tau: M \to [0, \Delta^2)$ which is strictly parabolic on $M - \tau^{-1}(0)$ and has a prescribed behavior on $\tau^{-1}(0)$. We require that there exists $p \in \tau^{-1}(0)$ such that, if $\mathbf{P}: \tilde{M} \to M$ is the blow up of M at p, then $\tau \circ \mathbf{P} \in C^\infty(\tilde{M})$ and such that, with

respect to coordinates centered at p, we have $C\|Z\|^2 \leqq \tau(Z) \leqq K\|Z\|^2$ for some positive constants C, K (cfr. Sect. 1 for detailed definitions). It is easy to see that this definition encompasses the above mentioned examples. If M is a complex manifold of dimension m equipped with such an exhaustion τ, our main result can be formulated as follows (cfr. Theorem 4.4 for the precise statement):

Theorem. *If $\Delta = \infty$, then M is biholomorphic to $\mathbb{C}^m$. If $\Delta < \infty$ and the Monge-Ampère foliation associated to τ is holomorphic, then M is biholomorphic to a bounded, strictly pseudoconvex, complete circular domain in $\mathbb{C}^m$.*

Clearly, since the assumptions are biholomorphically invariant, our Theorem characterizes, up to biholomorphic maps, $\mathbb{C}^m$ and the bounded, strictly pseudoconvex, complete circular domains in $\mathbb{C}^m$.

The methods used in this paper are a combination of Stoll's techniques ([12]) and of ideas developed in the special case of strictly convex domains ([9]). The simplifications to Stoll's original proof due to Burns [4] and Wong [14] cannot be used here because they are based on the consideration of the exponential map of the metric $dd^c\tau$ at the center point $\tau^{-1}(0)$ whereas in our situation the metric is not defined on $\tau^{-1}(0)$.

Here, once realized that, as in the case of strictly parabolic manifolds, the center $\tau^{-1}(0)$ reduces to a single point, our first step is to construct from scratch a "generalized" exponential map at the center point (Sect. 2). This map is in general only homeomorphic and in fact, it is smooth at the origin if and only if it is biholomorphic (which is the case if and only if τ is strictly parabolic on all M). The map is constructed by integrating the radial vector field $Y = \operatorname{grad}_g \sqrt{\tau}$, where g is the Riemannian metric associated to $dd^c\tau$, and by giving a precise description of its flow. The technical difficulty that one encounters is that, while it is very important to understand the flow of Y in a neighborhood of the center point, Y is not even defined at the center.

Following Stoll's original proof for the strictly parabolic manifold case, we overcome the problem by borrowing a notion from differential topology. Let $p \in M$ be the center point and D be a small ball around p. Then $M' = M - D$ is diffeomorphic to $M - \{p\}$. The orientable manifold N gotten by glueing together two copies of M' along the boundaries is called the double of M at p. Moreover there exists a differentiable map $r: N \to M$ such that r restricted to a copy of M' in N is diffeomorphic onto $M - \{p\}$ and such that $r^{-1}(p)$ is diffeomorphic to a sphere (cfr. Sect. 2 for the precise definition). The differentiability assumptions on τ guarantee that the vector field Y lifts, via r, to a smooth vector field $\tilde{Y}$ defined on *all* N (cfr. Lemma 2.3). Integrating the vector field $\tilde{Y}$ on N and pushing forward its flow to M, we get the flow of Y together with all the necessary asymptotic informations at the center (cfr. Theorem 2.6).

Next, in order to give a parametrization of the leaves of the Monge-Ampère foliation, we consider the vector field $Z = \sqrt{\tau}\, JY$, where J is the almost complex stucture of M. Since Y and Z are both tangent to the Monge-Ampère foliation, by suitably composing their flow, we get, in the case $\Delta = \sup \sqrt{\tau} < \infty$, a holomorphic covering map from a half plane onto each leaf of the foliation

of $M-\tau^{-1}(0)$ (cfr. Lemma 3.2). Then, observing that the flow of Z is periodic (cfr. Proposition 3.3), we are able to parametrize each leaf by a punctured disk in such a way that the parametrization depends smoothly on the changing of the leaf (cfr. Theorem 3.4). This will also show that our generalized exponential map transforms every disk through the origin into the closure of a leaf of the Monge-Ampère foliation. In analogy with our previous work [9], in Sect. 4 we define a "model" (G,σ) for (M,τ) where G is a circular domain and σ, up to a factor, is its Minkowski functional squared. A homeomorphism $h: G\to M$ is constructed such that $\tau\circ h=\sigma$. We show that h is biholomorphic if and only if the Monge-Ampère foliation association to τ is holomorphic. If $\Delta=\sup\sqrt{\tau}=\infty$, the proof follows the same line and, moreover, we can use the fact that, because of a theorem of Burns [4], the Monge-Ampère foliation is always holomorphic. In this case the manifold M is biholomorphic to $\mathbb{C}^m$ and τ pulls back to the Minkowski functional squared of some circular domain.

Complex manifolds carrying an exhaustion which is strictly parabolic outside its zero set and whose associated Monge-Ampère foliation is holomorphic, have been classified also by Wong [15]. His approach is based on the study of the differential geometry of the level sets of the exhaustion. His results do not overlap with ours but are clearly related.

Acknowledgements. I would like to thank here W. Stoll for sharing with me his ideas on this problem and I. Lieb for helpful discussions and remarks. I also gladly acknowledge the hospitality and the financial support of the Max-Planck-Institute für Mathematik and the SFB 40 Universität Bonn.

0. Notations

Let M be a complex manifold and let $T(M)$, $T^c(M)$ denote respectively the *tangent bundle* and the *complexified tangent bundle* of M. Then $T(M)$ is a subbundle of $T^c(M)$ and we have the splitting $T^c(M)=T^+(M)\oplus T^-(M)$ where $T^+(M)$ denote respectively the *holomorphic* and the *antiholomorphic tangent bundle* of M. The projections $\pi_+: T^c(M)\to T^+(M)$ and $\pi_-: T^c(M)\to T^-(M)$ restrict to $\mathbb{R}$-bundle isomorphisms $\pi_+: T(M)\to T^+(M)$ and $\pi_-: T(M)\to T(M)$. To the complex structure of M is canonically associated an *almost complex structure* $J: T^c(M)\to T^c(M)$ which is a bundle isomorphism such that $J^2=-\mathrm{Id}$ and such that $J_{|T^+(M)}$ is the multiplication by i and $J_{|T^-(M)}$ is the multiplication by $-i$. By an *hermitian metric* $\mathcal{k}$ on M we intend here a C^∞ function $\mathcal{k}: T^+(M)\oplus T^+(M)\to\mathbb{C}$ such that for all $p\in M$ its restriction $\mathcal{k}_p: T_p^+(M)\oplus T_p^+(M)\to\mathbb{C}$ is a positive definite hermitian form. The canonical form associated to $\mathcal{k}$ is a positive $(1,1)$ form ω defined, for all $p\in M$ and $X,Y\in T_p^+(M)$, by $\omega_p(X,\bar{Y})=(i/2\pi)\mathcal{k}_p(X,Y)$. The metric $\mathcal{k}$ is *Kähler* if and only if $d\omega=0$. Also to $\mathcal{k}$ is associated a Riemannian metric $\mathcal{g}$ defined for all $p\in M$ and $X,Y\in T(M)$, by $\mathcal{g}_p(X,Y)=\mathrm{Re}(\mathcal{k}_p(\pi_+(X),\pi_+(Y))$.

Given an exhaustion function $\tau: M\to[0,R^2)$, we shall use the following notations:

$$
\begin{aligned}
M[0] &= \{x \in M | \tau(x) = 0\} \\
M(r) &= \{x \in M | \tau(x) < r^2\} \\
M[r] &= \{x \in M | \tau(x) \leqq r^2\} \\
M\langle r \rangle &= \{x \in M | \tau(x) = r^2\} = M[r] - M(r) \\
M_* &= \{x \in M | \tau(x) \neq 0\} = M - M[0].
\end{aligned} \tag{0.1}
$$

In particular $\mathbb{C}(r)$, $\mathbb{C}[r]$ will always denote respectively the open and the closed disk of radius r in $\mathbb{C}$. For $m > 1$ we will denote the ball of radius r in $\mathbb{C}$ by $\mathbb{B}(r)$ and write $\mathbb{B} = \mathbb{B}(1)$. The unit sphere in $\mathbb{R}^n$ will be denoted by S^{n-1} or simply by S if the dimension is clear from the context. We will use the differential operators $d = \partial + \bar{\partial}$, $d^c = (i/4\pi)(\bar{\partial} - \partial)$ and $d\,d^c = (i/2\pi)\,\partial\,\bar{\partial}$. Finally, throughout the paper, upper indices will denote components and lower indices derivatives. Moreover, whenever possible, Einstein summation convention will be used. Also we shall use the Kroneker symbol $\delta_{\mu\nu}$.

1. Definition and Preliminaries

Let M be a non compact complex manifold of dimension m and $\tau: M \to [0, \Delta^2)$, $0 < \Delta = \sup\sqrt{\tau} \leqq \infty$, be an exhaustion function. We say the pair (M, τ) is a *manifold of circular type* if the following assumptions are satisfied:

(1.1) $\tau \in C^0(M) \cap C^\infty(M_*)$.

(1.2) $d\,d^c\,\tau > 0$ and $d\,d^c \log \tau \geqq 0$ on M_*.

(1.3) $(d\,d^c \log \tau)^m \equiv 0$ on M_*.

(1.4) There exists $p \in M[0]$ such that:

(i) with respect to local coordinates centered at p, if $\| \ \|$ denotes the euclidean norm, we have $C\|Z\|^2 \leqq \tau(Z) \leqq K\|Z\|$ for some positive constants C, K;

(ii) if $\mathbf{P}: \tilde{M} \to M$ is the blow up of M at p, then $\tau \circ \mathbf{P} \in C^\infty(\tilde{M})$.

In what follows we shall only use the following consequence of (1.4):

(1.4′) There exists $p \in M[0]$ and a coordinate neighbourhood U_p, centered at p, such that:

(i) there exists positive constants C, K such that $C\|Z\|^2 \leqq \tau(Z) \leqq K\|Z\|^2$ for all $Z \in U_p$;

(ii) there exists $\varepsilon_0 > 0$ so that $tZ \in U_p$ if $|t| < \varepsilon_0$ and $\|Z\| < 2$ and such that the map $h: (-\varepsilon_0, \varepsilon_0) \times \mathbb{B}(2) - \{0\} \to \mathbb{R}_+$ defined by $h(t, Z) = \tau(tZ)$ is of class C^∞.

As it will turn out from Theorem 3.4, (1.4′) is in fact equivalent to (1.4) if (1.1), (1.2), (1.3) are fulfilled. We also say that $\Delta = \sup\sqrt{\tau}$ is the *radius* of (M, τ).

It is useful to bear in mind some concrete example from the very beginning. In fact, the basic motivation for this paper was to try to generalize Example 1 according to the lines suggested by Example 2 and 3.

Example 1. Stoll defines in [12] a *strictly parabolic manifold* to be a pair (M,τ) where M is a complex manifold of dimension m and $\tau: M\to(0,\Delta^2)$, $0<\Delta=\sup\sqrt{\tau}\leqq\infty$, is a C^∞ strictly plurisubharmonic exhaustion such that $(dd^c\log\tau)^m\equiv 0$ on M_*. He proves that, up to exhaustion-preserving biholomorphic maps, the only such pairs are the balls of radius Δ in $\mathbb{C}^m$ equipped with the exhaustion $\tau_0=\|\ \|^2$ (simplified proofs of this theorem are given by Burns [4] and Wong [14]). Clearly, strictly parabolic manifolds are manifolds of circular type. This can be also easily checked directly without refering to the classification theorem (cfr. [12], Proposition 2.2).

Example 2. Let G be a strictly pseudoconvex, bounded, smooth, complete circular domain in $\mathbb{C}^m$. The *Minkowski functional* $\rho: \mathbb{C}^m\to\mathbb{R}_+$ associated to G is defined by

$$\rho(Z)=\begin{cases}0 & \text{if } Z=0\\ \inf\{1/t|t>0 \text{ and } tZ\notin G\} & \text{if } Z\neq 0.\end{cases} \tag{1.5}$$

It is not hard to check that (G,ρ^2) is a manifold of circular type with radius 1 (cfr. [3] and [10]).

Example 3. Let D be a smooth, bounded, strictly convex domain in $\mathbb{C}^m$. According to [6], for any $p\in D$ there exists a surjective C^∞ map $F:\mathbb{C}[1]\times S^{2m-1}\to\bar{D}$, which we called stationary representation of D at p in [9,], such that for all $b\in S^{2m-1}$ $F(\square,b):\mathbb{C}(1)\to D$ is a proper holomorphic embedding with $F(0,b)=p, F'(0,b)=\|F'(0,b)\|b$ and such that, if $\mathbf{K}$ denotes the Kobayashi metric of D, then $\mathbf{K}(p,b)=(\|F'(0,b)\|)^{-1}$. Also if $\lambda\in\mathbb{C}$ with $|\lambda|=1$ and $(z,b)\in\mathbb{C}[1]\times S^{2m-1}$, then $F(\lambda z,b)=F(z,\lambda b)$. An exhaustion $\tau:\bar{D}\to[0,1]$, called the *Lempert exhaustion* of D at p, is well defined by $\tau(F(z,b))=|z|^2$. Using results of Lempert [6] one can show that (D,τ) is a manifold of circular type with radius 1. Properties of the Lempert exhaustions and applications to the classification of strictly convex were studied in [9].

In the remainder of this section we give some preliminary results on manifold of circular type. First we have the following.

Theorem 1.1. *Let (M,τ) be a manifold of circular type. Then the set $M[0]$ consists exactly of one point which we call the center of M.*

Proof. Since no use of the smoothness of τ on $M[0]$ is made, a word by word repetition of the proof of Theorem 2.5 of [12] shows that $M[0]$ is connected. On the other hand, since τ satisfies (1.4), then $M[0]$ has an isolated point and therefore the claim follows. q.e.d.

Remark. One can show that given any smooth, bounded, strictly convex domain $D\subset\mathbb{C}^m$ and any $p\in D$, the only exhaustion σ of $\bar{D}$ such that (D,σ) is a manifold of circular type with center $\{p\}$ and radius 1 is the Lempert exhaustion at p. A proof of this fact based on the Bedford-Taylor minimum principle for the Monge-Ampère operator (cfr. [2]) is in [7]. Likewise given a strictly pseudoconvex, bounded, smooth, complete circular domain $G\subset\mathbb{C}^m$ then the unique exhaustion τ of $\bar{G}$ such that (G,σ) is a manifold of circular type with center at the origin and radius 1 is $\sigma=\rho^2$ where ρ is the Minkowski functional

of G. The proof of this fact goes along the same lines of the case of strictly convex domains.

Let (M,τ) be a manifold of circular type. We shall now collect some known facts which will be needed in this paper (cfr. for proofs [4] or [12] or [14]). A simple calculation shows that (1.2) and (1.3) imply that on M_*, with respect to any choice of local coordinates, the following fundamental formula holds

$$\tau_{\bar\nu}\tau^{\bar\nu\mu}\tau_\mu=\tau \tag{1.6}$$

where we set $(\tau^{\bar\nu\mu})=(\tau_{\mu\bar\nu})^{-1}$. A Kähler metric $\mathcal{k}$ on M_* is defined by the form $d\,d^c\tau>0$. A unique vector field X of type $(1,0)$, called the *complex gradient* of τ, exists which is dual to $\bar\partial\tau$ with respect to $\mathcal{k}$ i.e. such that:

$$\bar\partial\tau(\bar W)=\mathcal{k}(X,W) \tag{1.7}$$

for all vector field W of type $(1,0)$. In local coordinates we have the following expression for X:

$$X=X^\mu\frac{\partial}{\partial z^\mu}=\tau^{\bar\nu\mu}\tau_{\bar\nu}\frac{\partial}{\partial z^\mu}. \tag{1.8}$$

Because of (1.6) we have also

$$\bar\partial\tau(\bar X)=\mathcal{k}(X,X)=\tau. \tag{1.9}$$

From (1.2) and (1.3) it follows that the form $d\,d^c\log\tau$ has rank $m-1$ at every point of M_*. Then a distribution $\mathcal{A}$ of rank 1 is defined by

$$\mathcal{A}=\{W\in T^+(M_*)\,|\,d\,d^c\log\tau(W,\bar V)=0 \text{ for all } V\in T^+(M_*)\}. \tag{1.10}$$

The distribution $\mathcal{A}$ is integrable and the associated foliation of M_* in Riemann surfaces is called the *Monge-Ampère foliation* associated to τ. This foliation, first introduced by Bedford and Kalka [1], has been focus of many studies recently (see for instance [4], [10], [12], [14]). Here we shall just recall few well known properties. The distribution $\mathcal{A}$ is a trivial subbundle of $T^+(M_*)$ spanned by the vector field X i.e. for every $p\in M_*$ we have

$$\mathcal{A}_p=\mathbb{C}X_p. \tag{1.11}$$

As a consequence we get that the Monge-Ampère foliation is holomorphic – i.e., $\mathcal{A}$ is a holomorphic subbundle of $T^+(M_*)$ - if and only if X is holomorphic.

A further remarkable property of the Monge-Ampère foliation is that its leaves are totally geodesic submanifold of M_* with respect the Kähler metric $\mathcal{k}$. In fact if $\mathcal{g}$ is the Riemannian metric associated to $\mathcal{k}$ and we define a vector field Y on M_* by

$$Y=\frac{1}{\sqrt{\tau}}(X+\bar X), \tag{1.12}$$

then we have

$$Y=\operatorname{grad}_{\mathcal{g}}\sqrt{\tau} \quad\text{and}\quad \mathcal{g}(Y,Y)=1. \tag{1.13}$$

Every integral curve of Y lies on a leaf of the Monge-Ampère foliation and it is a geodesic for the metric $\mathscr{g}$.

2. The Gradient Flow

In this section we shall give a precise description of the flow of the vector field Y defined in (1.12). We shall use the same method used by Stoll in [12] for the case of strictly parabolic manifolds. We have to recall first the concept of the double at a point of a differentiable manifold. Let N be a differentiable manifold with $\dim_{\mathbb{R}} N=n$. Let $L\subset N$ be a closed submanifold with $\dim_{\mathbb{R}} L=n-1$. Let $h: N\to\mathbb{R}$ be a function of class C^∞ with $h_{|L}=0$ and $a\in L$. We say that *h vanishes of order k* at a if there exists a neighbourhood U of a and a function $g: U\to\mathbb{R}$ with $g^{-1}(0)=L\cap U$ and $dg(x)\neq 0$ if $x\in U$, such that on a neighborhood $U_1\subset U$ of a we have $h=g^k h_1$ where $h_1: U_1\to\mathbb{R}$ is a function of class C^∞ with $h_1(a)\neq 0$. Clearly h need not vanish of any order at a. If M is another differentiable manifold, with $\dim_{\mathbb{R}} M=n$, and $f: N\to M$ is a differentiable map with $f_{|L}=b\in M$, we say that *f branches of order s* at $a\in L$ if, given a local coordinates around $a\in N$ and $b\in M$, one has that

$$D=\det\left(\frac{\partial f^i}{\partial x^j}\right)_{i,j=1,\ldots,m}$$

vanishes of order $s+m-1$ at $a\in L$. This definition does not depend on the choice of local coordinates and one easily sees that $s\geqq 0$. Let $0\in M$ be any point. We say that (N,r) is the *double* of M at 0 (or connected sum of M and M at 0) (cfr. Milnor [8] and Stoll [12]) if:

(i) N is a differentiable manifold with $\dim_{\mathbb{R}} N=\dim_{\mathbb{R}} M=n$.

(ii) The map $r: N\to M$ is proper, surjective and of class C^∞.

(iii) The inverse image $r^{-1}(0)=S_0$ is a compact submanifold of N diffeomorphic to S^{m-1}.

(iv) The map r branches of order 0 at every point of S_0.

(v) There exists two open sets $N_1, N_2\subset N$ such that $N=N_1\cup N_2\cup S_0$ is a disjoint union and $r_j=r_{|N_j}: N_j\to M-\{0\}$ is a diffeomorphism.

In a neighbourhood of S_0 the double of M can be described as follows (cfr. [12], p. 113). Let U be a coordinate neighbourhood of 0 in M and let $\mathbb{B}(\varepsilon)\subset U$ be a ball of radius ε centered at 0. Let $V=r^{-1}(\mathbb{B}(\varepsilon))$. If ε is small enough, there exists a diffeomorphism $f: V\to(-\varepsilon,\varepsilon)\times S^{m-1}$ such that $f(S_0)=\{0\}\times S^{n-1}$ and $r(x)=r_0(f(x))$ for all $x\in V$ where r_0 is defined by $r_0(t,b)=t\,b$. When working near S_0, we shall identify V with $(-\varepsilon,\varepsilon)\times S^{m-1}$ so that $f=\mathrm{Id}$ and $r=r_0$ on V.

Let now (M,τ) be a manifold of circular type of dimension m and radius Δ. To describe the flow of the vector field Y defined in (1.12) we will first pull it back to a vector field $\tilde{Y}$ on the double (N,r) of M at the center $0\in M$ and study the flow of $\tilde{Y}$ on N.

Let U_0 be the coordinate neighbourhood of the center 0 of M with the properties listed in (1.4). Also the notations of Sect. 1 will be assumed. Furthermore here we write $S=S^{2m-1}$. Define $U=\{Z\in U_0 \mid Z=t\,b$ with $t\in(-\varepsilon,\varepsilon)$ and

$b \in S\}$. If (N, r) is the double of M at 0 then set $V = r^{-1}(U)$. We can choose ε small enough so that we can identily $V = (-\varepsilon, \varepsilon) \times S$ and so that $r(t, b) = t\, b$. We also set $V_1 = N_1 \cap V = (-\varepsilon, 0) \times S$ and $V_2 = N_2 \cap V = (0, \varepsilon) \times S$.

Lemma 2.1. *There exists a function $G: V - \{0\} \to \mathbb{R}_0$ of class C^∞ and constants $C_2 > C_1 > 0$, such that for all $Z \in U - \{0\}$ we have $\tau(Z) = G(Z) \|Z\|^2$ and $C_1 < G(Z) < C_2$. Furthermore there exists $c: V \to \mathbb{R}$, $H: S \to \mathbb{R}_+$, $R: V \to \mathbb{R}$ of class C^∞ such that $c^2(t, b) = \tau(r(t, b))$ for all $(t, b) \in V$ with $c_{|V_1} < 0$, $c_{|V_2} > 0$ and such that*

$$c(t, b) = t\sqrt{H(b)} + t^2 R(t, b). \tag{2.1}$$

Proof. The fact that a function G on $U - \{0\}$ exists with the above properties is an immediate consequence of (1.4′) part (i). Let $h: (-\varepsilon_0, \varepsilon_0) \times \mathbb{B}(2) - \{0\} \to \mathbb{R}_+$ be defined as in (1.4′) part (ii). Then its restriction $h: V \to \mathbb{R}_+$ is of class C^∞. There exists $G: V \to \mathbb{R}_+$ of class C^∞ and such that $\tilde{G} > 0$ on V so that $h(t, b) = \tilde{G}(t, b)\, t^2$. Then we have $\dot{h}(0, b) = 0$ and $\ddot{h}(0, b) > 0$ for all $b \in S$ where we have set $\dot{h} = \dfrac{\partial h}{\partial t}$. Define $H: S \to \mathbb{R}_+$ by

$$H(b) = \tfrac{1}{2} \ddot{h}(0, b). \tag{2.2}$$

Then there exists $T: V \to \mathbb{R}$ of class C^∞ such that if $(t, b) \in V$, then $0 \leqq h(t, b) = t^2 H(b) + t^3 T(t, b)$. In particular we have that $H(b) + t\, T(t, b) > 0$ for all $(t, b) \in V$. The function $L: V \to \mathbb{R}$ defined by $L(t, b) = (H(b) - t\, T(t, b))^{1/2} - (H(b))^{1/2}$ is of class C^∞ and $L(0, b) = 0$ for all $b \in S$. Thus there exists $R: V \to \mathbb{R}$ of class C^∞ such that $L(t, b) = t\, R(t, b)$. But then $c: V \to \mathbb{R}$ given by

$$c(t, b) = t(H(b))^{1/2} + t^2 R(t, b) = t(H(b) + t\, T(t, b)^{1/2}$$

is the required function. q.e.d.

Lemma 2.2. *Let $W = (-\varepsilon_0, \varepsilon_0) \times (\mathbb{B}(2) - \{0\})$. There exists a map $A: W \to \mathbb{C}^m$ of class C^∞ so that if X is the vector field defined in* (1.7), *$(t, Z) \in W$ and $t \neq 0$, then*

$$X(t, Z) = tZ + t^2 A(t, Z). \tag{2.3}$$

Proof. Let $(t, Z) \in W$ and $t \neq 0$, then because of Lemma 2.1 we have for $\mu, \nu = 1, \ldots, m$:

$$\begin{aligned}
&\tau(tZ) = t^2 G(tZ) \|Z\|^2 \\
&\tau_{\bar{\nu}}(tZ) = t\, G(tZ)\, z^\nu + t^2 G_{\bar{\nu}}(tZ) \|Z\|^2 \\
&\tau_{\mu\bar{\nu}}(tZ) = G(tZ)\, \delta_{\mu\nu} + t\, G_\mu(tZ)\, z^\nu + t\, G_{\bar{\nu}}(tZ)\, \bar{z}^\mu + t^2 G_{\mu\bar{\nu}}(tZ) \|Z\|^2
\end{aligned}$$

Because of (1.4′) there exists $\tilde{G}: W \to (0, \infty)$ of class C^∞ such that $h(t, Z) = \tilde{G}(t, Z) \|tZ\|^2$ and, in fact, $\tilde{G}(t, Z) = G(tZ)$ if $(t, Z) \in W$ and $t \neq 0$. Functions $\hat{G}^\mu: W \to \mathbb{C}$, $\hat{G}^{\bar{\nu}}: W \to \mathbb{C}$ and $\hat{G}^{\mu\bar{\nu}}: W \to \mathbb{C}$ of class C^∞ are defined by

$$\hat{G}^\mu(t, Z) = \begin{cases} (1/t)\, \tilde{G}_\mu(t, Z) = G_\mu(tZ) & \text{if } t \neq 0 \\ \dot{\tilde{G}}_\mu(t, Z) & \text{if } t = 0 \end{cases}$$

$$\hat{G}^{\bar{\nu}}(t,Z)=\begin{cases}(1/t)\,\tilde{G}_{\bar{\nu}}(t,Z)=G_{\bar{\nu}}(tZ) & \text{if } t\neq 0\\ \dot{\tilde{G}}_{\bar{\nu}}(t,Z) & \text{if } t=0\end{cases}$$

$$\hat{G}^{\mu\bar{\nu}}(t,Z)=\begin{cases}(1/t)\,G^{\mu}_{\bar{\nu}}(t,Z)=G_{\mu\bar{\nu}}(tZ) & \text{if } t\neq 0\\ \dot{\hat{G}}^{\mu\bar{\nu}}(t,T) & \text{if } t=0\end{cases}$$

where the dot denotes the derivative with respect to t. Then we have for $(t,Z)\in W$ and $t\neq 0$

$$\begin{aligned}\tau_{\bar{\nu}}(tZ)&=t\tilde{G}(t,Z)\,z^{\nu}+t^2\,\hat{G}^{\bar{\nu}}(t,Z)\,\|Z\|^2\\ \tau_{\mu\bar{\nu}}(tZ)&=\tilde{G}(t,Z)\,\delta_{\mu\nu}+tB^{\mu\bar{\nu}}(t,Z)\end{aligned}\tag{2.4}$$

where $B^{\mu\bar{\nu}}\colon W\to\mathbb{C}$ are functions of class C^{∞} defined by

$$B^{\mu\bar{\nu}}(t,Z)=\hat{G}^{\mu}(t,Z)\,z^{\nu}+\hat{G}^{\nu}(t,Z)\,\bar{z}^{\mu}+t\hat{G}^{\mu\bar{\nu}}(t,Z)\,\|Z\|^2.$$

There exist functions $R^{\mu\bar{\nu}}\colon W\to\mathbb{C}$ of class C^{∞} so that if $(t,Z)\in W$ and $t\neq 0$ we have

$$\tau^{\bar{\nu}\mu}(tZ)=\frac{1}{\tilde{G}(t,Z)}\,\delta_{\mu\nu}+tR^{\bar{\nu}\mu}(t,Z)$$

where $(\tau^{\bar{\nu}\mu})=(\tau_{\mu\bar{\nu}})^{-1}$. But recalling (1.8), we get at once (2.2) where $A\colon W\to\mathbb{C}^m$ is the map of class C^{∞} defined by

$$\begin{aligned}A^{\mu}(t,Z)=\sum_{\nu=1}^{m}\frac{\|Z\|^2}{\tilde{G}(t,z)}\,\delta_{\mu\nu}\,\hat{G}^{\bar{\nu}}(t,Z)+\tilde{G}(t,Z)\,R^{\mu\bar{\nu}}(t,Z)\,z^{\nu}\\ +t\,\|Z\|^2\,R^{\mu\bar{\nu}}(t,Z)\,\hat{G}^{\bar{\nu}}(t,Z).\quad\text{q.e.d.}\end{aligned}$$

In order to study vector fields on the double (N,r) of M we make the following identifications. Let $(t,b)\in V$. We identify $T_{(t,b)}(V)=\mathbb{R}\oplus T_b(S)$ where $T_b(S)$ is the tangent space of S at b. Also if $H_b(S)$ is the holomorphic tangent space to S we identify $T_{(t,b)}(V)=\mathbb{R}\oplus i\mathbb{R}\,b\oplus H_b(S)$. As seen above we can consider V as a neighbourhood of $S_0\subset N$ and set $r(t,b)=t\,b$ if $(t,b)\in V$. Also if $p\in V$ then we identify $T_p(U)=T_p^{+}(U)=T_p^{-}(U)=\mathbb{C}^m$ so that the projection π_+,π_- are the identity, the almost complex structure J is the multiplication by i and the conjugation is obtained by conjugation of the coordinates. We denote the hermitian product in $\mathbb{C}^m$ by $(|)$ so that the real scalar product $\langle\,|\,\rangle$ is given by $\langle Z|W\rangle=\mathrm{Re}(Z|W)$. Then we have the following formula:

$$dr^{-1}(t,b)(Z)=\left(\mathrm{Re}(Z|b),\frac{\mathrm{Im}(Z|b)}{t}\,i\,b+\frac{Z}{t}-\frac{(Z|b)}{t}\right)\in\mathbb{R}\oplus i\,\mathbb{R}\,b\oplus H_b(S)\tag{2.5}$$

Let J be the almost complex structure on M. On $N-S_0$ a complex structure is defined by pulling back by r the complex structure of M. The associated almost complex structure is denoted by $\tilde{J}$ and we have $\tilde{J}=dr^{-1}\circ J\circ dr$. Thus if $(t,b)\in V-S_0$ and $(u,i\,v\,b+w)\in\mathbb{R}\oplus i\mathbb{R}\oplus H_b(S)$

$$\tilde{J}_{t,b}(u,i\,v\,b+w)=\left(-t\,v,\frac{u}{t}\,i\,b+i\,w\right).\tag{2.6}$$

On $N-S_0$ we can define vector fields of class C^∞ $\tilde{X}$ and $\tilde{Y}$ by lifting via r the vector fields X, Y. Also define $\tilde{\tau}: N\to\mathbb{R}_+$ by $\tilde{\tau}=\tau\circ r$.

Lemma 2.3. *The function $\tilde{\tau}$ is of class C^∞ on N. There exists $c: N\to\mathbb{R}$ of class C^∞ such that $c^2=\tilde{\tau}$, $c_{|N_1}<0$ and $c_{|N_2}>0$. Furthermore $\tilde{X}$, $\tilde{J}\tilde{X}$ and $\tilde{Y}$ extend to vector fields of class C^∞ on N.*

Proof. We need only to check the claim on V. It is an immediate consequence of Lemma 2.1 that $\tilde{\tau}$ is of class C^∞ and that $c: N\to\mathbb{R}$ exists with the required properties. Using (2.3) and (2.5) we have for $(t,b)\in V-S_0$:

$$\begin{aligned}\tilde{X}(t,b)&=dr^{-1}(t,b)(X(r(t,b)))\\&=t(F_1(t,b),F_2(t,b)\,ib+F_3(t,b))\end{aligned}\tag{2.7}$$

where $F_1: V\to\mathbb{R}$, $F_2: V\to\mathbb{R}$, $F_3: V\to H_b(S)$ are of class C^∞ and defined by

$$\begin{cases}F_1(t,b)=1+t\,\mathrm{Re}(A(t,b)|b)\\F_2(t,b)=\mathrm{Im}(A(t,b)|b)\\F_3(t,b)=A(t,b)-(A(t,b)|b).\end{cases}\tag{2.8}$$

But then $\tilde{X}$ extends of class C^∞ on all V and thus on N. From (2.6) and (2.7) we have for $(t,b)\in V-S_0$

$$\tilde{J}\tilde{X}(t,b)=(-t^2F_2(t,b),F_1(t,b)\,ib+it\,F_2(t,b))\tag{2.9}$$

and thus also $\tilde{J}\tilde{X}$ extends of class C^∞ on all N. Finally, since if $(t,b)\in V-S_0$ we have

$$\tilde{Y}(t,b)=\frac{1}{c(t,b)}\left(\tilde{X}(t,b)+\overline{\tilde{X}(t,b)}\right),$$

$\tilde{Y}$ extends of class C^∞ on N because of (2.1) and (2.7). q.e.d.

Remark. Using (2.8) and (2.9) we also have the following useful formula for $b\in S$:

$$\tilde{J}\tilde{X}(0,b)=(0,ib).\tag{2.10}$$

Proposition 2.4. *There exists a number $\eta>0$ and a map $\Psi: (-\eta,\Delta)\times S\to N$ of class C^∞ such that*

(i) $\Psi(0,b)=(0,b)\in S_0=\{0\}\times S$ *for all* $b\in S$;
(ii) $\dot{\Psi}(t,b)=\tilde{Y}(\Psi(t,b))$ *for all* $(t,b)\in(-\eta,\Delta)\times S$;
(iii) *The map* $\Psi: [0,\Delta]\to S_0\cup N_2$ *is a diffeomorphism;*
(iv) $c(\Psi(t,b))=t$ *for all* $(t,b)\in[0,\Delta]\times S$;
(v) *If* $t\geqq 0$, *define* $\Psi_t: S\to N$ *by* $\Psi_t(b)=\Psi(t,b)$ *and* $N_t=\{c=t\}$. *Then* $\Psi_t: S\to N_t$ *is a diffeomorphism.*

Proof. Since $c: N\to\mathbb{R}$ is a function of class C^∞ such that $c^{-1}([0,\Delta))$ is non empty, non compact and connected, $c^{-1}([0,t])$ is compact for all $t\in[0,\Delta)$, $S_0=c^{-1}(\{0\}\neq\emptyset$ and $\tilde{Y}$ is a vector field of class C^∞ on N, according to Theorem 4.1 of [12] the map Ψ with the properties (i)–(v) exists provided $dc(p,\tilde{Y}(p))=1$ for all $p\in C^{-1}([0,\Delta))=S_0\cup N_2$. To show this observe that

$$dr(p, \tilde{Y}(p)) = -Y(r(p)) \quad \text{on } N_1$$
$$dr(p, \tilde{Y}(p)) = Y(r(p)) \quad \text{on } N_2.$$

But then on N_1 using (1.13) we have

$$dc(p, \tilde{Y}(p)) = -d\sqrt{\tau}(r(p), dr(p, \tilde{Y}(p)))$$
$$= -d\sqrt{\tau}(r(p), -Y(r(p))) = 1$$

In the same way one obtains $dc(p, \tilde{Y}(p)) = 1$ on N_2 and thus by continuity the claim is true on N. q.e.d.

Lemma 2.5. *Let $r_0 > 0$ small enough so that $\Psi((-\eta, r_0) \times S) \subset V$ and let $\Psi_1 : (-\eta, r_0) \times S \to \mathbb{R}$, $\Psi_2 : (-\eta, r_0) \times S \to S$ such that $\Psi = (\Psi_1, \Psi_2)$. There exist $\Psi_3 : (-\eta, r_0) \times S \to \mathbb{R}$ and $\Psi_4 : (-\eta, r_0) \times S \to \mathbb{C}^m$ of class C^∞ such that*

$$\Psi_1(t, b) = \frac{1}{\sqrt{H(b)}} t + t^2 \Psi_3(t, b), \tag{2.11}$$

$$\Psi_2(t, b) = b + t\Psi_4(t, b). \tag{2.12}$$

Proof. Because of (i) of Theorem 2.4 there exist $\Psi_4 : (-\eta, r_0) \times S \to \mathbb{C}^m$ and $\Psi_5 : (-\eta, r_0) \times S \to \mathbb{R}$ of class C^∞ such that (2.12) holds and $\Psi_1(t, b) = t\Psi_5(t, b)$. If π_+ is the isomorphism defined in Sect. 0, then

$$\lim_{t \to 0} \pi_+(\tilde{Y}(t, b)) = \lim_{t \to 0} \frac{1}{c(t, b)} \tilde{X}(t, b)$$
$$= \frac{1}{\sqrt{H(b)}}(1, F_2(0, b)\, ib + F_3(0, b))$$

where we have used (2.1), (2.7) and (2.8). Thus because of (ii) of Theorem 2.4 we have $\Psi_5(0, b) = \dot{\Psi}_1(0, b) = (H(b))^{-1/2}$. But then there exists $\Psi_3 : (-\eta, r_0) \times S \to \mathbb{R}$ of class C^∞ such that $\Psi_5(t, b) = (H(b))^{-1/2} + t\Psi_3(t, b)$ and thus the proof is completed. q.e.d.

We have now all the ingredients to describe in detail the flow on M of the vector field Y defined in (1.12). Here we identify $T_0(M) = \mathbb{C}^m$.

Theorem 2.6. *There exists a map of class C^∞*

$$\varphi : [0, \Delta) \times S \to M \tag{2.13}$$

with the following properties

(i) $\varphi(0, b) = 0$ *for all* $b \in S$;
(ii) $\dot{\varphi}(0, b) = (H(b))^{-1/2} b$ *for all* $b \in S$;
(iii) $\dot{\varphi}(t, b) = Y(\varphi(t, b))$ *for all* $(t, b) \in (0, \Delta) \times S$;
(iv) $\varphi : (0, \Delta) \times S \to M_*$ *is a diffeomorphism*;
(v) $\tau(\varphi(t, b)) = t^2$ *for all* $(t, b) \in [0, \Delta) \times S$;
(vi) *the map* $\varphi_t : S \to M\langle t \rangle$ *defined by* $\varphi_t(b) = \varphi(t, b)$ *is a diffeomorphism*;

(vii) $t\pi_+(\dot{\varphi}(t,b)) = X(\varphi(t,b))$ *for all* $(t,b)\in[0,\Delta)\times S$;

(viii) *The curve* $\varphi(\square,b):(0,\Delta)\times S\to M_*$ *is geodesic with respect the Kähler metric defined by* $dd^c\tau>0$ *for all* $b\in S$.

Proof. If $r: N\to M$ is the projection from the double, we define $\varphi = r\circ\Psi:[0,\Delta)\to M$. Then (i), (iii), (iv), (v), (vi) and (vii) are immediate from Proposition 2.4. Also, using the usual identification of the tangent space of M at the center 0 with $\mathbb{C}^m$, with respect to local coordinates around 0 we have recalling (2.11) and (2.12):

$$\begin{aligned}\varphi(t,b) &= r(\Psi(t,b))\\ &= r(\Psi_1(t,b),\Psi_2(t,b))\\ &= t(H(b))^{-1/2}b + t^2T(t,b)\end{aligned}$$

where T is of class C^∞ and thus (ii) follows. Finally, (viii) is a consequence of the well known fact recalled in Sect. 1 that the integral curves of Y are geodesics. q.e.d.

Corollary 2.7. *The map* $E:\mathbb{B}(\Delta)\to M$ *defined by*

$$E(Z) = \begin{cases} 0 & \text{if } Z=0\\ \varphi(\|Z\|, Z/\|Z\|) & \text{if } Z\neq 0\end{cases} \tag{2.14}$$

is a homeomorphism and its restriction $E:\mathbb{B}(\Delta)-\{0\}\to M_*$ *is a diffeomorphism.*

Proof. The map $E:\mathbb{B}(\Delta)-\{0\}\to M_*$ is clearly a diffeomorphism. Also $\tau(E(Z)) = \|Z\|^2$ and hence $\lim_{Z\to 0} E(Z)=0$ and $\lim_{p\to 0} E^{-1}(p)=0$ so that the proof is complete. q.e.d.

Remark. The metric ℓ defined on M_* by $dd^c\tau$ is not defined on the center 0 of M. Thus one cannot define in the usual way an exponential map at 0. The map E defined in (2.14) is a generalized exponential map at the "singular" point 0 with respect the metric ℓ. The existence of an exponential map at a singular point of a Riemannian manifold has been investigated by Stone [13] in terms of curvature conditions in a neighbourhood of the singularity. Here it follows from the behavior of the exhaustion which defines the metric. It would be interesting to understand the relationship between Stone's hypothesis and ours in this particular case. Notice also that the map E is differentiable at 0 if and only if the metric extends over 0 and this is the case if M is strictly parabolic (in which case E is biholomorphic). This will be evident from the results of the next section.

3. The Leaf Parametrization Map

In order to give a parametrization of the leaves of the Monge-Ampère foliation induced by τ we must consider also the flow of the vector field Z defined on M_* by

$$Z = i(X-\bar{X}) = J(X+\bar{X}). \tag{3.1}$$

where X is as usual the complex gradient of τ. Using (1.9) we have

$$d\tau(Z)=i\partial\tau(X)-i\bar{\partial}\tau(\bar{X})=i\tau-i\tau=0$$

and therefore any integral curve of Z is contained in some level set of τ which is compact. But this implies that Z is complete i.e. there exists a unique map $\sigma\colon \mathbb{R}\times M_*\to M_*$ such that

(3.2) $\sigma(\square, p)\colon \mathbb{R}\to M_*$ is an integral curve of Z such that $\sigma(0,p)=p$.

(3.3) For all $y\in\mathbb{R}$ the map $\sigma(y,\square)\colon M_*\to M_*$ is a diffeomorphism.

(3.4) For all $y_1, y_2\in\mathbb{R}$ and $p\in M_*$ we have $\sigma(y_1+y_2,p)=(y_1,\sigma(y_2,p))$.

Define a map $\alpha\colon [0,\Delta)\times M_*\to M$ by

$$\alpha(t,p)=\varphi(t,\varphi_{\sqrt{\tau(p)}}^{-1}(p)). \tag{3.5}$$

Clearly α is continous and of class C^∞ on $(0,\Delta)\times M_*$.

Lemma 3.1. *For all $t\in\mathbb{R}$, $s\in(0,\Delta)$ and $p\in M_*$ we have*

$$\sigma(t,\alpha(s,p))=\alpha(s,\sigma(t,p)) \tag{3.6}$$

Proof. Since α is such that for all $p\in M_*$ $\alpha(\square,p)\colon (0,\Delta)\to M_*$ is the unique integral curve of Y such that $\alpha(\sqrt{\tau(p)},p)=p$, (3.6) is equivalent to $[Y,Z]=0$ on M_*. Define $F=X+\bar{X}$ so that $Y=(\tau)^{-1/2}F$ and $Z=JF$. Then

$$[Y,Z]=\frac{1}{\sqrt{\tau}}[F,JF]-JF\left(\frac{1}{\sqrt{\tau}}\right)F.$$

Since $JF((\tau)^{-1/2})=0$ follows easily from (1.9) to show $[Y,Z]=0$ we have only to prove $[F,JF]=0$. If $[F,JF]=H^\mu\dfrac{\partial}{\partial z^\mu}+\bar{H}^\mu\dfrac{\partial}{\partial\bar{z}^\mu}$ then one computes $H^\mu=2i\bar{X}^\nu X^\mu_{\bar{\nu}}$. On the other hand

$$\bar{X}^\nu X^\mu_{\bar{\nu}}=\tau_\alpha\tau^{\bar{\nu}\alpha}(\tau_\beta\tau^{\beta\mu})_{\bar{\nu}}=\tau_\alpha\tau^{\bar{\nu}\alpha}\tau_{\beta\bar{\nu}}\tau^{\beta\mu}+\tau_\alpha\tau^{\bar{\nu}\alpha}\tau_\beta\tau^{\beta\mu}_{\bar{\nu}}.$$

Differentiating the equality $\tau_{\sigma\bar{\gamma}}\tau^{\bar{\nu}\sigma}=\delta_{\gamma\nu}$ and (1.6) we obtain the following equalities:

$$\tau_{\sigma\bar{\gamma}}\tau^{\bar{\nu}\sigma}_\beta+\tau_{\bar{\gamma}\beta\sigma}\tau^{\bar{\nu}\sigma}=0$$

$$\tau_{\beta\bar{\nu}}\tau^{\bar{\nu}\alpha}\tau_\alpha+\tau_{\bar{\nu}}\tau^{\bar{\nu}\alpha}_\beta\tau_\alpha=0.$$

But then we have

$$\begin{aligned}\tau_\alpha\tau^{\bar{\nu}\alpha}\tau_{\beta\bar{\nu}}\tau^{\beta\mu}&=-\tau_\alpha\tau^{\bar{\nu}\alpha}_\beta\tau_{\bar{\nu}}\tau^{\beta\mu}=-\tau_{\bar{\nu}}\tau^{\bar{\gamma}\alpha}\tau_{\sigma\bar{\gamma}}\tau^{\bar{\nu}\sigma}_\beta\tau_\alpha\tau^{\beta\mu}\\&=\tau_{\bar{\nu}}\tau^{\bar{\gamma}\alpha}\tau_{\sigma\bar{\gamma}\beta}\tau^{\bar{\nu}\sigma}\tau_\alpha\tau^{\beta\mu}\end{aligned}$$

and

$$\begin{aligned}\tau_\beta\tau^{\beta\mu}_{\bar{\nu}}\tau_\alpha\tau^{\bar{\nu}\alpha}&=\tau_\beta\tau^{\bar{\gamma}\mu}\tau_{\sigma\bar{\gamma}}\tau^{\beta\sigma}_{\bar{\nu}}\tau_\alpha\tau^{\bar{\nu}\alpha}=-\tau_\beta\tau^{\bar{\gamma}\mu}\tau_{\sigma\bar{\gamma}\bar{\nu}}\tau^{\beta\sigma}\tau_\alpha\tau^{\bar{\nu}\alpha}\\&=-\tau_{\bar{\nu}}\tau^{\beta\mu}\tau_{\sigma\beta\bar{\gamma}}\tau^{\bar{\nu}\sigma}\tau_\alpha\tau^{\bar{\gamma}\alpha}\end{aligned}$$

where in the last line we exchange the indices in the following way: $\beta \to \nu$, $\gamma \to \beta$, $\nu \to \gamma$. Our computation shows that $\bar{X}^{\nu} X^{\mu}_{\nu} = 0$ and thus $[F, JF] = 0$. q.e.d.

We need now some notations. For $R \in (0, \infty]$ define

$$\mathbb{H}(R) = \begin{cases} \{z \in \mathbb{C} \mid \operatorname{Re} z < \log R\} & \text{if } R < \infty \\ \mathbb{C} & \text{if } R = \infty. \end{cases} \tag{3.7}$$

Also fix $x_0 \in (0, \Delta)$. Write $q = \varphi_{x_0}$. A one parameter group of diffeomorphisms $\zeta: \mathbb{R} \times S \to S$ is defined by

$$\zeta(y, b) = q^{-1}(\sigma(y, q(b))). \tag{3.8}$$

Finally we define a surjective map $W: \mathbb{H}(\Delta) \times S \to M_*$ of class C^{∞} by

$$W(x + iy, b) = \sigma(y, \varphi(e^x, b)) \tag{3.9}$$

From (3.5), (3.6) and (3.8) one has at once

$$W(x + iy, b) = \sigma(y, \alpha(e^x, q(b)) = \alpha(e^x, \sigma(y, q(b))) = \varphi(e^x, \zeta(y, b)). \tag{3.10}$$

Lemma 3.2. *For all $b \in S$ the map $W(\square, b): \mathbb{H}(\Delta) \to M_*$ is holomorphic and $W'(z, b) = X(W(z, b))$. Furthermore if $x + iy$, $s + it \in \mathbb{H}(\Delta)$ and $b, c \in S$, then $W(x + iy, b) = W(s + it, c)$ if and only if $x = s$ and $c = \zeta(y - t, b)$.*

Proof. If we set $p = W(x + iy, b)$, then by (3.10):

$$\frac{\partial}{\partial x} W(x + iy, b) = \dot{\varphi}(e^x, \zeta(y, b))\, e^x = \sqrt{\tau(p)}\, Y(p) = X(p) + \overline{X(p)}$$

$$\frac{\partial}{\partial x} W(x + iy, b) = \dot{\sigma}(y, \varphi(e^x, b)) = Z(p) = J(X(p) + \overline{X(p)}).$$

But then $\dfrac{\partial W}{\partial x} + J \dfrac{\partial W}{\partial y} = \dfrac{\partial W}{\partial x} = \dfrac{\partial W}{\partial x} + JJ \dfrac{\partial W}{\partial x} = 0$ and $W(\square, b)$ is thus holomorphic. Moreover $W'(z, b) = \pi_+ \left(\dfrac{\partial W(z, b)}{\partial x} \right) = X(W(z, b))$. Finally if $W(x + iy, b) = W(s + it, c)$ then $\varphi(e^x, \zeta(y, b)) = \varphi(e^s, \zeta(t, b))$ and therefore $x = s$ and $\zeta(y, b) = \zeta(t, c)$ which implies $c = \zeta(y - t, b)$ since ζ is a one parameter group. In the other direction the implication is immediate. q.e.d.

We need now to compute explicitly $\zeta(y, b)$ for $y \in \mathbb{R}$ and $b \in S$. To this end we shall again use the notations introduced in Sect. 2. Let (N, r) be the double of M at the center 0, $V = (-\varepsilon, \varepsilon) \times S \subset N$ and let $U = r(V)$. If Ψ is the map defined in Proposition 2.4, let $r_0 > 0$ be such that $\Psi([0, r_0) \times S) \subset V$. Then, since $\varphi = r \circ \Psi$, we have also $\varphi([0, r_0) \times S) \subset U$. It is not restrictive to assume that, if $x_0 \in (0, \Delta)$ is the number used in the definition (3.8) of ζ, then $x_0 \in (0, r_0)$.

Proposition 3.3. *For all $y \in \mathbb{R}$ and $b \in S$ we have*

$$\zeta(y, b) = e^{iy} b. \tag{3.11}$$

Proof. Fix $b \in S$. For $t \in (0, r_0)$ and $y \in \mathbb{R}$ define $\gamma(t, y) = r^{-1}(\varphi(t, \zeta(y, b))$. Then γ is of class C^∞ on $(0, r_0) \times \mathbb{R}$ and we have

$$\gamma(t, y) = \Psi(t, \zeta(y, b)) = r^{-1}(W(\log t + iy, b)).$$

Using (2.11) and (2.12) we have

$$\begin{aligned}\gamma(t, y) &= (\Psi_1(t, \zeta(y, b)), \Psi_2(t, \zeta y, b))) \\ &= \left(\frac{t}{\sqrt{H(\zeta(y, b))}} + t^2 \Psi_3(t, \zeta(y, b)), \zeta(y, b) + t\Psi_4(t, \zeta(y, b))\right).\end{aligned}$$

Thus

$$\gamma_y(t, y) = \left(t \frac{\partial}{\partial y}\left(\frac{1}{\sqrt{H(\zeta(y, b))}} + t\Psi_3(t, \zeta(y, b)), \zeta_y(y, b) + t \frac{\partial}{\partial y} \Psi_4(t, \zeta(y, b))\right)\right).$$

The limit of $\gamma_y(t, y)$ as $t \to 0$ exists and

$$\gamma_y(0, y) = \lim_{t \to 0} \gamma_y(t, y) = (0, \zeta_y(y, b)).$$

On the other hand since

$$W_y(\log t + iy, b) = iW'(\log t + iy, b) = iX(W(\log t + iy, b)),$$

we have:

$$\begin{aligned}\gamma_y(t, b) &= dr^{-1}(W(\log t + iy, b), W_y(\log t + iy, b)) \\ &= dr^{-1}(\varphi(t, \zeta(y, b)), iX(\varphi(t, \zeta(y, b)))) \\ &= \tilde{J}\tilde{X}(\Psi(t, \zeta(y, b))).\end{aligned}$$

Thus using (2.10) we have:

$$\begin{aligned}\gamma_y(0, y) = \lim_{t \to 0} \gamma_y(t, y) &= \tilde{J}\tilde{X}(\Psi_1(0, \zeta(y, b)), \Psi_2(0, \zeta(y, b))) \\ \tilde{J}\tilde{X}(0, \zeta(y, b)) &= i\zeta(y, b).\end{aligned}$$

Comparing the two expressions that we have obtained for $\gamma_y(0, y)$ we get

$$\zeta_y(y, b) = i\zeta(y, b).$$

Integrating this equation and imposing the initial condition $\zeta(0, b) = b$ we conclude that (3.11) holds. q.e.d.

We can now state the main result of this section. As usual we shall identify the tangent space to M at the center 0 with $\mathbb{C}^m$ and we shall denote by $\| \ \|$ the euclidean norm on it.

Theorem 3.4. *There exists uniquely a surjective map, called the leaf parametrization map,*

$$F: \mathbb{C}(\Delta) \times S \to M \tag{3.12}$$

of class C^∞ and with the following properties:

(i) $F(0,b)=0$ *for all* $b\in S$;
(ii) $\tau(F(z,b))=|z|^2$ *for all* $b\in S$ *and* $z\in\mathbb{C}(\Delta)$;
(iii) $F(t,b)=\varphi(t,b)$ *for all* $b\in S$ *and* $t\in[0,\Delta)$;
(iv) $F(z,\lambda b)=F(\lambda z,b)$ *for all* $z\in\mathbb{C}(\Delta)$, $b\in S$ *and* $\lambda\in\mathbb{C}$ *with* $|\lambda|=1$;
(v) $F(\square,b)\colon\mathbb{C}(\Delta)\to M$ *is holomorphic, proper and injective*;
(vi) $X(F(z,b))=zF'(z,b)$;
(vii) $F'(0,b)=\|F'(0,b)\|\,b=\dfrac{b}{\sqrt{H(b)}}$.

Proof. Since $\exp\colon\mathbb{H}(\Delta)\to\mathbb{C}(\Delta)-\{0\}$ is the universal cover and, because of Lemma 3.2 and Proposition 3.3, we have $W(z_1,b)=W(z_2,b)$ if and only if $z_2=z_1+2k\pi i$, a unique map $F\colon\mathbb{C}(\Delta)-\{0\}\to M_*$ of class C^∞ is defined by $F(e^z,b)=W(z,b)$. Also $F(\square,b)\colon\mathbb{C}(\Delta)-\{0\}\to M_*$ is holomorphic and injective. Define $F(0,b)=0$ for all $b\in S$. Then $\tau(F(0,b))=0$. If $z\in\mathbb{C}(\Delta)-\{0\}$, then $z=e^{x+iy}$ with $x+iy\in\mathbb{H}(\Delta)$. Thus $\tau(F(z,b))=\tau(W(x+iy,b))=e^{2x}=|z|^2$ and (ii) is verified. Clearly $F\colon\mathbb{C}(\Delta)\times S\to M$ so defined is continous and, by Riemann extension theorem, $F(\square,b)\colon\mathbb{C}(\Delta)\to M$ is holomorphic for all $b\in S$. Let $U\subset M$ be a coordinate neighbourhood centered at $0\in M$. With no loss of generality we can identify it with an open subset of $\mathbb{C}^m$. Let $t_0\in(0,\Delta)$ be such that $F(\mathbb{C}(t_0)\times S)\subset U$. Let $s\in(0,t_0)$ then if $z\in\mathbb{C}(s)$ and $b\in S$ we have

$$F(z,b)=\frac{1}{2\pi i}\int_{|w|=s}\frac{F(w,b)}{w-z}\,dw$$

and hence F is of class C^∞ on $\mathbb{C}(s)\times S$ and therefore on $\mathbb{C}(\Delta)\times S$. If $K\subset M$ is compact then $K\subset M[r]$ for some $r\in(0,\Delta)$. Then $F^{-1}(K)\subset\mathbb{C}[r]\times S$ and therefore F is proper and thus so is $F(\square,b)$ for all $b\in S$. Also (iii) follows by construction and F is surjective because so is φ. As a consequence of the definition of F, Lemma 3.2 and Proposition 3.3 one obtains (iv) immediately. Again from Lemma 3.2 one gets (vi). Finally (vii) is a consequence of (iii) and of (ii) of Theorem 2.6. q.e.d.

Remark. If $E\colon\mathbb{B}(\Delta)\to M$ is the map defined in (2.14), because of (iii) and (iv) of Theorem 3.4 we have that $E(zb)=F(z,b)$ for all $(z,b)\in\mathbb{C}(\Delta)\times S$. In other words the restriction of E on any disk through the origin is holomorphic. This proves the claim made earlier that E is biholomorphic if and only if it is of class C^∞ at the origin.

Examples. Before moving on it may clarify the matter to write down the leaf parametrization map for some explicit example. Let G be a strictly pseudoconvex, bounded, smooth, complete circular domain in $\mathbb{C}^m$ and let ρ its Minkowski functional. If $\sigma=\rho^2$ one computes easily that the complex gradient vector field X of σ is given by $X(Z)=Z^\mu\dfrac{\partial}{\partial z^\mu}$. It follows then that the map $F\colon\mathbb{C}(1)\times S\to G$, relative to the manifold of circular type (G,σ), is defined by $F(z,b)=z(\rho(b))^{-1}b$. Let $D\subset\mathbb{C}^m$ be a bounded, smooth, strictly convex domain, $p\in D$ and τ be the Lempert exhaustion at p. From the results of [9] if follows that the map F for (D,τ) is just the stationary parametrization map which

defines τ. However, it is not hard to check directly that $F(\square, b)\colon \mathbb{C}(1)\to D$ is a stationary map in the sense of [6] for all $b\in S$ and thus that the map F is indeed the stationary parametrization of D.

4. The Circular Representation

Let as before (M, τ) be a manifold of circular type and Δ be its radius. We shall now define an "approximation" (G, σ) of (M, τ) where G is a circular domain and σ is, up to a factor, its Minkowski functional squared. Let $F\colon \mathbb{C}(\Delta)\to S\times M$ be the leaf parametrization map of (M, τ). As usual we identify freely $T_0(M)$ with $\mathbb{C}^m$. Define $\sigma\colon \mathbb{C}^m\to\mathbb{R}_+$ by $\sigma(0)=0$ and

$$\sigma(Z)=\left(\frac{\|Z\|}{\|F'(0, Z/\|Z\|)\|}\right)^2=H(Z/\|Z\|)\,\|Z\|^2 \tag{4.1}$$

if $Z\neq 0$ where H is defined in (2.7). One easily sees that $\sigma\in C^0(\mathbb{C}^m)\cap C^\infty(\mathbb{C}^m-\{0\})$. Also from the definition of σ and (iv) of Theorem 3.4 one has for all $Z\in\mathbb{C}^m$ and $\lambda\in\mathbb{C}$:

$$\sigma(\lambda Z)=|\lambda|^2\sigma(Z). \tag{4.2}$$

For $r\in(0, \infty]$ we define also complete circular domains $G(r)\subset\mathbb{C}^m$ by

$$G(r)=\begin{cases}\{Z\}\in\mathbb{C}^m\mid\sigma(Z)<r^2\} & \text{if } r<\infty\\ \mathbb{C}^m & \text{if } r=\infty.\end{cases} \tag{4.3}$$

Clearly the Minkowski functional of $G(r)$ is given by $\sqrt{\sigma}/r$.

Proposition 4.1. *For all $r\in(0, \infty]$ the pair $(G(r), \sigma)$ is a manifold of circular type with radius r.*

Proof. Since (1.1) and (1.4) hold by construction only (1.2) and (1.3) must be shown. But if $dd^c\sigma>0$ on $G(r)-\{0\}$ using (4.2) the other statements are obtained by simple calculations. We shall now show $dd^c\sigma>0$ on $G(r)-\{0\}$. Since $\sigma_{\mu\bar\nu}(tZ)=\sigma_{\mu\bar\nu}(Z)$ because of (4.2), it will be enough to show that for some $\varepsilon_0>0$ one has $\sigma_{\mu\bar\nu}(Z)\,V^\mu\bar V^\nu>0$ for all $Z\neq 0$ with $\|Z\|<\varepsilon_0$ and $V\in\mathbb{C}^m$ with $\|V\|=1$. In a small enough neighbourhood U of $0\in M$ we can choose coordinates so that for some $\varepsilon, \varepsilon_0>0$ if $0<t<\varepsilon$ and $Z\in\mathbb{C}^m-\{0\}$ with $\|Z\|<\varepsilon_0$, then $tZ\in U$ and we have

$$\sigma(tz)=\tau(tZ)-t^3L(t, Z)$$

where L is a function of class C^∞ on $(-\varepsilon, \varepsilon)\times(\mathbb{B}(\varepsilon_0)-\{0\})$ (cfr. Lemma 2.1). Furthermore, recalling (2.4), there are functions $\tilde G$ and $B^{\mu\bar\nu}$ of class C^∞ defined on $(-\varepsilon, \varepsilon)\times(\mathbb{B}(\varepsilon_0)-\{0\})$ such that

$$\tau_{\mu\bar\nu}(tZ)=\tilde G(t, Z)\,\delta_{\mu\bar\nu}+tB^{\mu\bar\nu}(t, Z).$$

Also there is a constant $C>0$ such that $\tilde G(t, Z)>C$ for all $(t, Z)\in(-\varepsilon, \varepsilon)\times(\mathbb{B}(\varepsilon_0)-\{0\})$. Fix $Z\in\mathbb{B}(\varepsilon_0)-\{0\}$ and define

$$M = \max_{\substack{|t| \leqq \varepsilon/2 \\ V \in S}} \| \sum_{\mu\nu} B^{\mu\bar{\nu}}(t, Z)\, V^{\mu} \bar{V}^{\nu} \|$$

$$N = \max_{\substack{|t| \leqq \varepsilon/2 \\ V \in S}} \| \sum_{\mu\nu} L_{\mu\bar{\nu}}(t, Z)\, V^{\mu} \bar{V}^{\nu} \|.$$

Then for $0 < t < \min\{\varepsilon/2, C/(M+N)\}$ we have

$$\sigma_{\mu\bar{\nu}}(Z)\, V^{\mu} \bar{V}^{\nu} = \sigma_{\mu\bar{\nu}}(tZ)\, V^{\mu} \bar{V}^{\nu} = \tau_{\mu\bar{\nu}}(tZ)\, V^{\mu} \bar{V}^{\nu} - t L_{\mu\bar{\nu}}(tZ)\, V^{\mu} \bar{V}^{\nu}$$
$$\geqq C - t(M+N) > 0 \quad \text{q.e.d.}$$

Define $\rho: \mathbb{C}^m \to \mathbb{R}_+$ by

$$\rho = \sqrt{\sigma}. \tag{4.4}$$

Then a map $h: G(\Delta) \to M$, called the *circular representation* of (M, τ), is defined by $h(0) = 0$ and

$$h(Z) = F(\rho(Z), Z/\|Z\|), \tag{4.5}$$

where F is the leaf parametrization map of (M, τ) defined in (3.12), if $Z \neq 0$.

Proposition 4.2. *The circular representation has the following properties:*

(i) $h: G(\Delta) \to M$ *is a homeomorphism with* $\tau \circ h = \sigma$;

(ii) $h: G(\Delta) - \{0\} \to M_*$ *is a diffeomorphism*;

(iii) *If* L *is any complex line through the origin in* $\mathbb{C}^m$, *then* $h_{|G(\Delta) \cap L}$ *is a holomorphic map*;

(iv) $h: G(\Delta) \to M$ *is a biholomorphic map if and only if it is of class* C^{∞} *at the origin.*

Proof. It is clear by construction that $\tau \circ h = \sigma$. Also if we define $Q: G(\Delta) \to \mathbb{B}(\Delta)$ by

$$Q(Z) = \frac{\rho(Z)}{\|Z\|} Z$$

then Q is a homeomorphism and $Q: G(\Delta) - \{0\} \to \mathbb{B}(\Delta) - \{0\}$ is a diffeomorphism. Also if $E: \mathbb{B}(\Delta) \to M$ is the map defined in (2.14) then $h = Q \circ E$. Thus (i) and (ii) follow from Corollary 2.7. Also since $\rho(\lambda Z) \subset |\lambda| \rho(Z)$ for all $\lambda \in \mathbb{C}$ and $Z \in \mathbb{C}^m$, we have $Q(\lambda Z) = \lambda Q(Z)$ for all $\lambda \in \mathbb{C}$ and $Z \in G(\Delta)$ and thus (iii) follows from the fact remarked at the end of Sect. 3 that E is holomorphic on each disk through the origin. Finally, since a smooth function on a complete circular domain which is holomorphic on each disk through the origin is holomorphic (cfr. [9], Lemma 8.1 for a detailed proof), (iii) implies (iv). q.e.d.

Remark. One can show in fact the the map h is differentiable at the origin (not necessarily of class C^1!) and that $dh(0) = \mathrm{Id}$. The proof is the same as in the case of strictly convex domains (cfr. [9], Theorem 6.2).

Proposition 4.3. *The circular representation* $h: G(\Delta) \to M$ *is biholomorphic if and only if the Monge-Ampère foliation associated to* τ *is holomorphic.*

Proof. By the definition of h and (vi) of Theorem 3.4, the disks through the origin in $G(\Delta)$ are mapped by h into the leaves of the Monge-Ampère foliation induced by τ. Thus, if h is biholomorphic, the Monge-Ampère foliation on M is holomorphic. Assume now that the Monge-Ampère foliation is holomorphic. As recalled in Sect. 1 this is equivalent to the fact that the complex gradient X of τ is holomorphic on M_*. Thus X extends holomorphically on all M with $X(0)=0$. Moreover, if we regard X as a vector function with values in $\mathbb{C}^m$, we have, for instance from (2.3), that $dX(0)=\mathbf{Id}$. There exist a neighbourhood U of $0\in M$ and a holomorphic map $Q: U\to\mathbb{C}^m$ such that on U

$$X=\mathbf{Id}+Q. \tag{4.6}$$

Define a map $A: \mathbb{C}(1)\times G(\Delta)\to M$ by $A(z, Z)=h(zZ)$. Then from (4.5) and (iv) of Theorem 3.4, for $Z\neq 0$, one has $A(z, Z)=F(z\rho(Z), Z/\|Z\|)$ so that from (vi) of Theorem 3.4 we have

$$zA'(z, Z)=X(A(z, Z)). \tag{4.7}$$

For every $j\geqq 1$, there exists $A_j: G(\Delta)\to\mathbb{C}$ such that

$$A(z, Z)=\sum_{j=1}^{\infty} A_j(Z)\, z^j.$$

Thus from (4.6) and (4.7) we obtain

$$\begin{aligned} Q(A(z, Z)) &= zA'(z, Z)-A(z, Z) \\ &= \sum_{j=2}^{\infty}(j-1)\, A_j(Z)\, z^j \end{aligned} \tag{4.8}$$

Define $P(z, Z)=\sum_{j=1}^{\infty} A_j(Z)\, z^{j-1}$. If $Q=\sum_{k=2}^{\infty} Q_k$ is the development of Q in homogeneous vector polynomials of degree k, then there exist vector polynomials Q_{kr} such that

$$Q_k(P(z, Z))=\sum_{r=0}^{\infty} Q_{kr}(A_1(Z), \ldots, A_{r+1}(Z))\, z^r.$$

Thus

$$\begin{aligned} Q(A(z, Z)) &= \sum_{k=2}^{\infty} Q_k(P(z, Z))\, z^k \\ &= \sum_{k=2}^{\infty}\sum_{r=0}^{\infty} Q_{kr}(A_1(Z), \ldots, A_{r+1}(Z))\, z^{r+k} \\ &= \sum_{j=2}^{\infty}\Big(\sum_{j=k+r} Q_{kr}(A_1(Z), \ldots, A_{r+1}(Z))\Big) z^j. \end{aligned} \tag{4.9}$$

Comparing (4.8) and (4.9), we get

$$(j-1)\, A_j(Z)=\sum_{j=k+r} Q_{kr}(A_1(Z), \ldots, A_{r+1}(Z)).$$

Since $k \geqq 2$, we have $r+1 \leqq j-1$ and hence there exists a vector polynomial R_j for all $j \geqq 2$ such that

$$(j-1)\,A_j(Z) = R_j(A_1(Z), \ldots, A_{j-1}(Z)). \tag{4.10}$$

Since $A_1 = dX(0) = \mathbf{Id}$ is holomorphic, using (4.10) it follows by induction on j that A_j is holomorphic for all $j \geqq 1$. On the other hand there exists a number $\varepsilon > 0$ such that $\mathbb{B}(2\varepsilon) \subset G(\Delta)$, and $\sum_{j=1} A_j(Z) z^j$ converges uniformly if $\|Z\| < 2\varepsilon$ and $|z| < 1/2$. If $\|Z\| < \varepsilon$ then we have

$$h(Z) = A(1/2, 2Z) = \sum_{j=1}^{\infty} A_j(2Z)(1/2)^j$$

where the series converges uniformly. Therefore h is holomorphic on a neighbourhood of the origin and hence of class C^∞ there. By Proposition 4.2 it follows that in fact h is biholomorphic. q.e.d.

We can now state our main result.

Theorem 4.4. *Let (M, τ) be a manifold of circular type of radius Δ. Then*

(i) *If $\Delta = \infty$ there exists a biholomorphic map $h: \mathbb{C}^m \to M$ such that $(\tau \circ h)^{1/2}$ is the Minkowski functional of some circular domain $G \subset \mathbb{C}^m$.*

(ii) *If $\Delta < \infty$ and the Monge-Ampère foliation associated to τ is holomorphic, then there exists a circular domain $G(\Delta) \subset \mathbb{C}^m$ and a biholomorphic map $h: G(\Delta) \to M$ such that $\Delta^{-1}(\tau \circ h)^{1/2}$* is the Minkowski functional of $G(\Delta)$.

Proof. Proposition 4.2 and 4.3 yield immediately (ii). To prove (i) we have only to recall that according to a theorem of Burns [4], when $\Delta = \infty$ the Monge-Ampère foliation associated to τ is always holomorphic. q.e.d.

References

1. Bedford, E., Kalka, M.: Foliations and complex Monge-Ampère equations. Commun. Pure Appl. Math. **30**, 510–538 (1977)
2. Bedford, E., Taylor, B.A.: The Dirichlet problem for a complex Monge-Ampère equation. Invent. Math. **37**, 1–44 (1976)
3. Bedford, E., Taylor, B.A.: Variational properties of the complex Monge-Ampère equation, II Intrinsic norms., Amer. J. Math. **101**, 1131–1166 (1979)
4. Burns, D.: Curvature of the Monge-Ampère foliations and parabolic manifolds. Ann. of Math. **115**, 349–373 (1982)
5. Griffiths, P., King, J.: Nevanlinna theory and holomorphic mappings between algebraic varieties. Acta Math. **130**, 145–220 (1973)
6. Lempert, L.: La métrique de Kobayshi et la represéntation des domaines sur la boule. Bull. Soc. Math. Fr. **109**, 427–474 (1981)
7. Lempert, L.: Solving the Degenerate Complex Monge-Ampère Equation with one Concentrated Singularity. Math. Ann. **263**, 515–532 (1983)
8. Milnor, J.: Sommes de variétés différentiables et structure différentiables des spheres. Bull. Soc. Math. Fr. **87**, 439–444 (1959)
9. Partrizio, G.: Parabolic exhaustions for strictly convex domains. Manusc. Math. **47**, 271–309 (1984)

10. Patrizio, G., Wong, P.M.: Stability of the Monge-Ampère foliation. Math. Ann. **263**, 13–29 (1983)
11. Stoll, W.: Value distribution on parabolic spaces, Lecture Notes in Math. **600**. Berlin-Heidelberg-New York: Springer 1977
12. Stoll, W.: The characterization of strictly parabolic manifolds. Ann. Scuola Norm. Super. Pisa, **7**, 81–154 (1980)
13. Stone, D.A.: The exponential map at an isolated singular point. Mem. Am. Math. Soc. **256** (1982)
14. Wong, P.M.: Geometry of the Homogeneous Complex Monge-Ampère Equation. Invent. Math., **67**, 261–274 (1982)
15. Wong, P.M.: On Umbilical Hypersurfaces and Uniformization of Circular Domains, in Complex Analysis of Several Variables, Y.T. Siu ed., Symp. Pure Math. **41**, Am. Math. Soc. Providence, 1984

Received June 8, 1984

Math. Z. 189, 365-369 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Strongly Invertible Knots Have Property *R*

Steven A. Bleiler

Department of Mathematics, University of Utah,
Salt Lake City, Utah 84112, USA

A knot K in S^3 is invertible if there exists an orientation preserving homeomorphism ξ of space which takes K to itself but reverses the orientation of K. The knot K is strongly invertible if ξ is also an involution. Given such, Waldhausen [12] shows that ξ is equivalent to a π-rotation and that K meets the axis of this rotation in precisely two points.

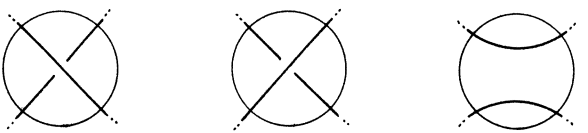
Given K and a rational number r, possibly $1/0$, let $M(K;r)$ denote the closed 3-manifold obtained by Dehn surgery of type r on K. A non-trivial knot is said to have Property R if $M(K;0)$ is not $S^2 \times S^1$. As advertised in the title, this paper shows that strongly invertible knots have Property R, thereby giving a partial solution to problem 1.16 of Kirby's list [6].

The proof is not long, and uses two quite different characterizations of the unknot: One due to Scharlemann [9] in terms of vignettes, another arising primarily from the work of Gordon [4] and Thurston [11] in terms of surgery. These are recalled in § 1. The proof is found in § 2.

§ 1. Recognizing the Unknot

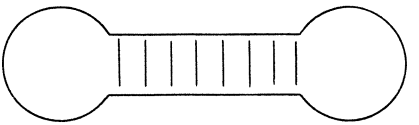
Scharlemann's solution [9] of Kirby's problems 1.1, 1.2A, and 1.3 provides a method of recognizing the unknot in terms of vignettes. It is the solution to 1.3 which is of interest here, specifically:

Proposition 1. *Let K be a knot which can be changed in a single vignette as indicated below to obtain the unknot in one case and the unlink of two unknotted components in the other:*



Then K is unknotted.

Scharlemann explicitly solves problem 1.2A and observes (along with Kirby's list) that Proposition 1 follows. Problem 1.2A concerns band sums, defined for knots K_0, K_1 imbedded in disjoint balls in the 3-sphere. If b: $I \times I \to S^3$ is an embedding such that $b^{-1}(K_i) = I \times \{i\}$ $i = 0, 1$; then the band sum of K_0 and K_1 along b is the knot resulting from joining $K_0 \backslash b(I \times \{0\})$ to $K_1 \backslash b(I \times \{1\})$ by the arcs $b(\partial I \times I)$. This knot is denoted by $K_0 \#_b K_1$. Scharlemann shows that if the band sum of K_0 and K_1 along b is the unknot then both K_0 and K_1 are unknotted and the band sum is the connected sum. In particular, the band b is isotopic to the trivial band illustrated below.



The components of the unlink lie in disjoint balls so both the unknot and the knot K in the hypothesis of Proposition 1 are there expressed as a band sum of unknots. The bands forming K and the unknot differ only by a full twist. By the above, the band yielding the unknot is isotopic to the trivial band and it follows that this twist can be achieved by isotopy. Hence K is unknotted, and we obtain Proposition 1.

The other method of recognizing the unknot arises from the general study of Dehn surgery. Thurston [11, § 5] shows that for hyperbolic knots at most finitely many Dehn surgeries yield 3-manifolds with a finite fundamental group. This result is generalized by Gordon [4, Theorem 1.11] who shows that for K a knotted, non-torus knot, $\pi_1 M(K;r)$ is finite for at most finitely many r. The Dehn surgeries on torus knots have been completely described [8]. For the purpose at hand, one only needs the observation of Seifert [10] that torus knots have 'Property P', that is, only trivial surgery on a non-trivial torus knot yields a simply connected 3-manifold. This shows:

Proposition 2. *Let K be a knot such that $M(K;r)$ is simply connected for infinitely many values of r. Then K is unknotted.*

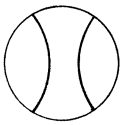
Remark. As $M(\text{unknot}; 1/n)$ is S^3 for integer values of n, Proposition 2 characterizes the unknot.

§ 2. A Proof of the Theorem

Suppose that K is a strongly invertible knot such that $M(K;0)$ is $S^2 \times S^1$. As K is strongly invertible, $M(K;r)$ double branch covers the 3-sphere [7], [1]. From [1] and [2] we have:

Theorem 3. *For a strongly invertible knot K, the base space and branch set $B(K;r)$ of the double branch covering $M(K;r) \to S^3$ decomposes as an untangle attached to a locally prime tangle T in a manner specified by r; moreover, if K is knotted, then T is in fact prime.*

Recall that a tangle is a 3-ball with a pair of properly embedded, disjoint spanning arcs, called strings. A tangle is locally prime if the strings have no local knots, prime if locally prime and does not contain a properly embedded disk separating the strings. An untangle is a homeomorph of a standard ball, arcs pair and is locally prime.

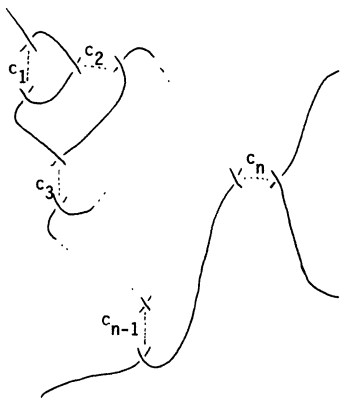


Untangle

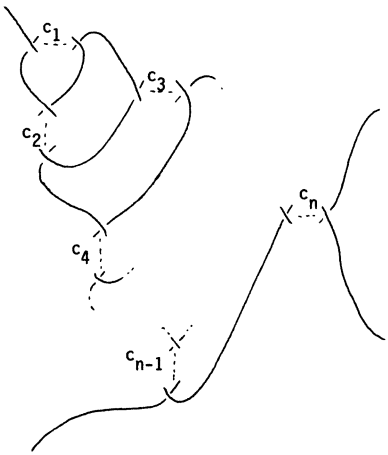


Prime Tangle

The tangle T of Theorem 3 is covered by the knot exterior while the untangle is covered by a regular neighborhood of the knot K. Via the Conway correspondence [3] certain untangle projections are associated to $Q \cup \{1/0\}$. Up to ambient isotopy fixing the boundary, Conway shows that an untangle may be put into one of the two forms illustrated below.



n even



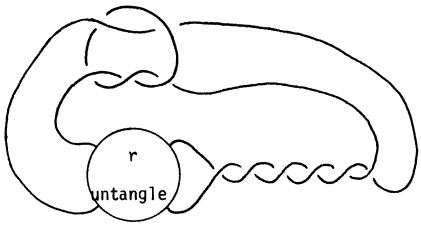
n odd

Such an untangle is associated to a rational number r, possibly 1/0, computed by the continued fraction rule $r=c_n+(c_{n-1}+(c_{n-2}\cdots+c_1^{-1})^{-1}\ldots)^{-1}$. Given such a continued fraction, the corresponding untangle projection is called an *r-untangle.* A theorem of Conway [3] says that the strings of r-untangles corresponding to different continued fraction expansions of a fixed rational r are isotopic via ambient isotopy fixing the boundary.

As 1/0 surgery on K produces S^3, and the branch covering $M(K,1/0)\to S^3$ is induced by a π-rotation, $B(K,1/0)$ is an unknotted circle. There is, therefore, the following useful corollary of Theorem 3.

Proposition 4. *For strongly invertible knot K, the base space and branch set of the double branch covering $M(K;r)\to S^3$ can be obtained by replacing a* 1/0-*untangle by an r-untangle in an appropriate projection of the unknot.*

This is illustrated below for K the right-handed trefoil. For details, see [2].

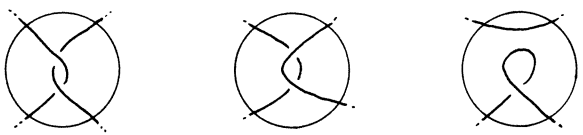


Returning to the proof, a theorem of Hodgson [5] says that whenever $S^2 \times S^1$ double branch covers S^3, the branch set is the unlink of two unknotted components. We have therefore the following vignettes of $B(K;0)$ and $B(K,1/0)$:



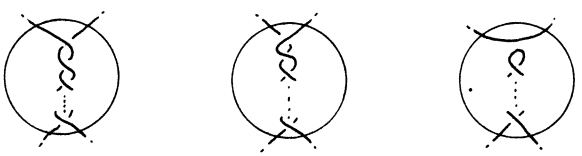
B(K;0), the unlink B(K;1/0), the unknot

Now consider the following sequence of vignettes:



B(K;1/2) B(K;1/0) the unknot B(K;0) the unlink

Changing the 'upper' crossing in $B(K;1/2)$ as illustrated above allows one to apply Proposition 1 and conclude that $B(K;1/2)$ is unknotted. Moreover, this series of vignettes can be viewed inductively:



$B(K;\frac{1}{2(n+1)})$ $B(K;\frac{1}{2n})$ B(K;0) the unlink

Induction shows that $B(K;1/2n)$ is unknotted for all $n \in \mathbb{Z}$. It follows that $M(K;1/2n)$ is S^3, and hence simply connected, for all $n \in \mathbb{Z}$. Proposition 2 now shows K to be unknotted.

References

1. Bleiler, S.A.: Knots prime on many strings. TAMS vol. **282**, 385–402 (1984)
2. Bleiler, S.A.: Prime tangles and composite knots. To appear in Proceedings Vancouver 1983. Lecture Notes in Mathematics. Berlin Heidelberg New York: Springer
3. Conway, J.H.: An enumeration of knots and links, and some of their algebraic properties. Computational problems in abstract algebra, 329–358. Oxford & New York: Pergamon Press 1969
4. Gordon, C.McA.: Dehn surgery and satellite knots. TAMS **275**, 687–708 (1983)
5. Hodgson, C.G.: Involutions and isotopies of lens spaces, Master's Thesis, University of Melbourne, 1981
6. Kirby, R.: Problems in low dimensional manifold theory. Proc. Symp. Pure Math. **XXXII**, part 2, 273–312
7. Montisenos, J.M.: Surgery on links and double branch coverings of S^3. Ann. Math. Stud. **84**, 227–259 (1977)
8. Moser, L.E.: On the impossibility of obtaining $S^2 \times S^1$ by elementary surgery along a knot. Pac. J. Math. **53**, 519–523 (1974)
9. Scharlemann, M.: Smooth spheres in $\mathbb{R}^4$ with four critical points are standard. Invent. Math. (to appear)
10. Seifert, H.: Topologie dreidimensionaler gefaserter Räume. Acta Math. **60**, 147–238 (1933)
11. Thurston, W.: The geometry and topology of 3-manifolds. Preprint, Princeton University Press
12. Waldhausen, F.: Über Involutionen der 3-Sphäre. Topology **8**, 81–91 (1969)

Received February 13, 1984, received in final form May 22, 1984

Math. Z. 189, 371-391 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

Krümmungstreue Diffeomorphismen Riemannscher und pseudo-Riemannscher Mannigfaltigkeiten

Bernhard Ruh

Département de Mathématiques, EPF-Ecublens, CH-1015 Lausanne, Suisse

Einleitung

Das von Gauß und Riemann stammende Theorema egregium besagt, daß die Krümmung eine Invariante der Metrik ist: Isometrisch aufeinander abbildbare Riemannsche oder pseudo-Riemannsche Mannigfaltigkeiten haben somit dieselbe (Gaußsche oder Riemannsche) Krümmung. Daß die Umkehrung für zweidimensionale Räume nicht gilt, kann leicht durch Beispiele gezeigt werden [2] p. 165 und [8] p. 121. In der vorliegenden Arbeit wird untersucht, inwieweit die Umkehrung für Räume der Dimension $n \geqq 3$ zutrifft, d.h., ob ein Diffeomorphismus zwischen Riemannschen oder pseudo-Riemannschen Räumen, der die Riemannsche Schnittkrümmung K invariant läßt, eine Isometrie ist, und falls nicht, wie die Räume und der Diffeomorphismus (lokal) aussehen. Der bekannte Fall konstanter Krümmung wird dabei ausgeschlossen.

Im Falle Riemannscher Räume (definite Metriken) wurde das Problem von R.S. Kulkarni 1967 weitgehend gelöst [3–5]: Falls $K \neq \text{const}$ ist, ist jeder krümmungstreue Diffeomorphismus konform und im Falle $n \geqq 4$ sogar eine Isometrie. Falls $n = 3$ ist, folgt die Isometrie, wenn die Räume konform-flach sind. Für 3-dimensionale Räume fand T. Nasu 1974 eine Relation, die im Falle eines krümmungstreuen Diffeomorphismus' erfüllt sein muß [7]. Anschließend konstruierte S.T. Yau Beispiele von 3-dimensionalen Riemannschen Räumen, zwischen denen nicht-isometrische krümmungstreue Diffeomorphismen existieren [11]. Kulkarni bewies 1977, daß auch im Falle pseudo-Riemannscher Räume (indefinite Metriken) krümmungstreue Diffeomorphismen konform sein müssen [6].

In dieser Arbeit werden nun alle Riemannschen und pseudo-Riemannschen Räume bestimmt, welche nicht-triviale krümmungstreue Diffeomorphismen gestatten. Dabei zeigt sich folgendes: Für $n \geqq 4$ gibt es genau eine Klasse von Räumen mit indefiniter Metrik, welche nicht-triviale krümmungstreue Diffeomorphismen erlauben. Es handelt sich dabei um die konform-flachen Räume mit rekurrenter Krümmung im Sinne von Walker [9]. (Die im Gegensatz dazustehende Behauptung in [6] ist nur unter der zusätzlichen Annahme, daß

der Gradient des Konformitätsfaktors nicht verschwindende Norm hat, richtig.) Für $n=3$ wird bewiesen, daß die von Yau gefundenen Beispiele im definiten Fall die einzigen sind. Bei Zulassung indefiniter Metriken ergeben sich drei Klassen von Beispielen, je nachdem, ob die Eigenwerte des Ricci-Tensors reell und verschieden sind, ob komplexe Eigenwerte auftreten, oder ob der Ricci-Tensor einen dreifachen Eigenwert hat, aber nicht diagonalisierbar ist. Im letzten Fall findet man drei verschiedene Typen von Beispielen. In den beiden ersten Fällen gelingt es, die Metrik vollständig durch gewisse nur vom Konformitätsfaktor abhängige Invarianten auszudrücken.

In Kapitel I wird zunächst ein einfacher Beweis (unabhängig von der Signatur der Metriken) gegeben, daß nicht-triviale krümmungstreue Diffeomorphismen für $n \geqq 3$ konform sind. Anschließend werden die Beziehungen zwischen konform äquivalenten bzw. krümmungsgleichen Metriken zusammengestellt. Kapitel II enthält die Diskussion für $n \geqq 4$ und $n=3$.

Die vorliegende Arbeit ist ein etwas überarbeiteter Auszug aus meiner Dissertation an der ETH Zürich. Ich möchte mich an dieser Stelle bei meinem Lehrer, Prof. Dr. K. Voß, für seine Unterstützung und seine wichtigen Anregungen herzlich bedanken. Ebenfalls gebührt mein Dank der Universität Zürich, vor allem Prof. Dr. H.H. Keller, für ihre Gastfreundschaft.

Kapitel I. Krümmungsgleichheit und Konformität

Seien (M, g) und $(\bar{M}, \bar{g})$ zwei Riemannsche oder pseudo-Riemannsche Mannigfaltigkeiten (Riem. oder ps.-Riem. Mgf.) der Dimension $n \geqq 3$ (der Differenzierbarkeitsklasse C^r, $r \geqq 4$) mit Krümmungstensoren R bzw. $\bar{R}$. Sei (G_2, π, M) das Graßmannbündel der Flächenelemente (Ebenen) auf M. Ein Element A in G_2 heißt *ausgeartet* (resp. *nicht-ausgeartet*), falls $g_{|A}$ ausgeartet (resp. nicht-ausgeartet) ist. Sodann ist die Riemannsche Krümmung (Schnittkrümmung) definiert auf den nicht-ausgearteten Elementen von G_2 durch

$$K(A) := \frac{R(X, Y, X, Y)}{G(X, Y)},$$

wobei (X, Y) eine Basis von A ist, und der Tensor G definiert ist durch $G(X, Y) := G(X, Y, X, Y)$, $G(X, Y, V, W) := \langle X, V\rangle\langle Y, W\rangle - \langle X, W\rangle\langle Y, V\rangle$. Dabei schreiben wir – wie es häufig geschehen wird – $\langle\ ,\ \rangle$ für $g(\ ,\ )$.

Definition. Zwei Riem. oder ps.-Riem. Mgf. (M, g) und $(\bar{M}, \bar{g})$ heißen *krümmungsgleich*, falls ein Diffeomorphismus $f\colon M \to \bar{M}$ existiert, so daß

$$K(A) = \bar{K}(f_* A)$$

für alle nicht-ausgearteten Flächenelemente $A \in G_2$, für welche das Flächenelement $f_* A$ nicht-ausgeartet ist. Die Abbildung f heißt dann ein *krümmungstreuer Diffeomorphismus*.

Bemerkung. In Abweichung von [6] fordern wir nicht, daß f_* ausgeartete Flächenelemente in solche überführen muß.

Definition. Ein Punkt $p \in M$ heißt *isotroper Punkt von* (M, g), falls $K_{|\pi^{-1}(p)} = \text{const}$ ist. Anderenfalls heißt p ein *nicht-isotroper Punkt von* (M, g).

§ 1. Krümmungsgleiche Räume sind konform äquivalent

Wir beweisen das

Theorem I.1. *Sei* $f: (M, g) \to (\bar{M}, \bar{g})$ *ein krümmungstreuer Diffeomorphismus zwischen zwei Riem. oder ps.-Riem. Mgf. der Dimension* $n \geqq 3$. *Die Menge* I *der nicht-isotropen Punkte sei dicht in* M. *Dann ist* f *ein konformer Diffeomorphismus, d.h. es ist* $f^* \bar{g} = \lambda g$ *mit einer Funktion* $\lambda: M \to \mathbb{R}^* := \mathbb{R} \setminus \{0\}$.

Dieses Theorem ist "punktuell"; es folgt nämlich sofort aus dem

Satz I.2. *Sei* $f: (M, g) \to (\bar{M}, \bar{g})$ *ein krümmungstreuer Diffeomorphismus. Sei* $p \in M$ *ein nicht-isotroper Punkt. Dann ist der Vektorraumisomorphismus* $T_p f: T_p M \to T_{f(p)} \bar{M}$ *eine Homothetie.*

Theorem I.1 wurde für Riem. Mgf. in [5] mit Hilfe einer speziellen Basis bewiesen, deren Existenz relativ schwierig zu zeigen war. Für ps.-Riem. Mgf. beruht der Beweis in [6] auf der Tatsache, daß jeder Diffeomorphismus, der ausgeartete Flächenelemente in solche überführt, konform sein muß.

Unser Beweis ist unabhängig von der Signatur der Metrik und rein algebraisch. Er benötigt drei Hilfssätze. Sei im folgenden $p \in M$ fest. Ferner identifizieren wir M und $\bar{M}$ via f, was uns eine Mgf. M mit zwei Metriken g und $\bar{g}$ liefert.

Hilfssatz 1. *Sei* $K(A) = \bar{K}(A)$ *für alle (bez.* g *und* $\bar{g}$*) nicht-ausgearteten Flächenelemente* A *in* $T_p M$. *Dann gilt für alle* $X, Y \in T_p M$: $\bar{G}(X, Y) R(X, Y, X, Y) = \bar{R}(X, Y, X, Y) G(X, Y)$.

Beweis. Die reellwertige Funktion

$$(X, Y) \mapsto \bar{G}(X, Y) R(X, Y, X, Y) - \bar{R}(X, Y, X, Y) G(X, Y)$$

ist nach Voraussetzung gleich Null auf der in $T_p M \times T_p M$ überall dichten Menge $\{(X, Y) \mid G(X, Y) \bar{G}(X, Y) \neq 0\}$. Aus Stetigkeitsgründen folgt die Behauptung. □

Die beiden folgenden Hilfssätze sind von rein algebraischer Natur.

Hilfssatz 2. *Sei* $(e_1, \ldots, e_n)$ *eine Vektorraumbasis von* $T_p M$. *Seien* $g_{ij} := g(e_i, e_j)$, $1 \leqq i, j \leqq n$, *die Komponenten von* $g_{|T_p M}$. *Dann ist das Polynom*

$$G(x^1, \ldots, x^n, y^1, \ldots, y^n) := \left(\sum_{i,j} g_{ij} x^i x^j\right)\left(\sum_{i,j} g_{ij} y^i y^j\right) - \left(\sum_{i,j} g_{ij} x^i y^j\right)^2$$

für $n \geqq 3$ *irreduzibel.*

Beweis. Sei o.B.d.A. $(e_1, \ldots, e_n)$ eine bezüglich g orthonormierte Basis, d.h. es sei $g_{ij} = \varepsilon_i \delta_{ij}$, $\varepsilon_i := \pm 1$. Dann ist

$$G(x^1, \ldots, x^n, y^1, \ldots, y^n) = a(x^n)^2 + 2b x^n y^n + c(y^n)^2 + d$$

mit

$$a=\varepsilon_n\left(\sum_{i=1}^{n-1}\varepsilon_i(y^i)^2\right);\quad b=-\varepsilon_n\left(\sum_{i=1}^{n-1}\varepsilon_i x^i y^i\right);\quad c=\varepsilon_n\left(\sum_{i=1}^{n-1}\varepsilon_i(x^i)^2\right);$$
$$d=G(x^1,\ldots,x^{n-1},y^1,\ldots,y^{n-1}).$$

Wir fassen nun G als Polynom in den Variablen x^n, y^n auf und schreiben es als Summe zweier homogener Polynome, d.h. wir setzen $G=G^2+G^0$, wo $G^2:=a(x^n)^2+2bx^ny^n+c(y^n)^2$ und $G^0:=d$ ist. Sei G zerlegbar, sei also $G=PQ$ für zwei Polynome P und Q. Dann zerlegen wir P und Q in homogene Teile $P=P^2+P^1+P^0$ und $Q=Q^2+Q^1+Q^0$. Durch Ausmultiplizieren und Vergleich folgt zunächst $P^2Q^2=0$, $P^2Q^1+P^1Q^2=0$. Sei z.B. $P^2=0$; daraus folgt $P^1=0$ oder $Q^2=0$. Ist $Q^2=0$, so folgt durch Vergleich $P^1Q^1=G^2=a(x^n)^2+2bx^ny^n+c(y^n)^2$; $P^1Q^0+P^0Q^1=0$; $P^0Q^0=G^0=d$. Aus den letzten beiden Gleichungen folgt $P^1=\psi Q^1$ für ein $\psi=\psi(x^1,\ldots,x^{n-1},y^1,\ldots,y^{n-1})$ und damit aus der ersten $ac-b^2=0$, was im Widerspruch steht zu

$$ac-b^2=\tfrac{1}{2}\sum_{i,j=1}^{n-1}\varepsilon_i\varepsilon_j(x^iy^j-x^jy^i)^2\neq 0\qquad\text{für } n\geqq 3.$$

Es ist also $P^1=0$, und somit teilt $P=P^0$ insbesondere die Koeffizienten a und c; daraus folgt $P=\text{const}$. □

Hilfssatz 3. *Sei $\bar{G}(X,Y)=cG(X,Y)$, $(\forall(X,Y)\in T_pM\times T_pM)$, für ein $c\in\mathbb{R}^*$. Dann ist $\bar{g}=\lambda g$ für ein $\lambda\in\mathbb{R}^*$.*

Beweis. Wie im Beweis von Hilfssatz 2 sei $(e_1,\ldots,e_n)$ eine Basis mit $g_{ij}=\varepsilon_i\delta_{ij}$ für $1\leqq i,j\leqq n$. Sei $a_{pq}:=\bar{g}(e_p,e_q)$, $1\leqq p,q\leqq n$. Für feste i,j,k, $1\leqq i\leqq j\leqq k\leqq n$, setzen wir für $\alpha,\beta=1,2,3$

$$(A_{\alpha\beta})=\begin{pmatrix}a_{ii}&a_{ij}&a_{ik}\\a_{ji}&a_{jj}&a_{jk}\\a_{ki}&a_{kj}&a_{kk}\end{pmatrix},\qquad(G_{\alpha\beta})=\begin{pmatrix}\varepsilon_i&0&0\\0&\varepsilon_j&0\\0&0&\varepsilon_k\end{pmatrix}.$$

Es sei $(A^*_{\alpha\beta})$ bzw. $(G^*_{\alpha\beta})$ die Matrix der algebraischen Komplemente von $(A_{\alpha\beta})$ bzw. $(G_{\alpha\beta})$. Nach Voraussetzung ist nun für beliebige $(x,y)\in\mathbb{R}^2$

$$\begin{aligned}\bar{G}(xe_i+ye_j,e_k)&=A^*_{22}x^2+A^*_{11}y^2-2A^*_{12}xy=cG(xe_i+ye_j,e_k)\\&=c(\varepsilon_i\varepsilon_kx^2+\varepsilon_j\varepsilon_ky^2)=c(G^*_{22}x^2+G^*_{11}y^2-2G^*_{12}xy)\end{aligned}$$

und analog für zyklisch permutierte (i,j,k) bzw. $(1,2,3)$. Man hat also $(A^*_{\alpha\beta})=c(G^*_{\alpha\beta})$. Damit existiert die zu $(A_{\alpha\beta})$ inverse Matrix; da diese proportional zu $(A^*_{\alpha\beta})$ ist, folgt $(A_{\alpha\beta})=\lambda(G_{\alpha\beta})$. Da eine entsprechende Gleichung für jedes Tripel i,j,k gilt, ist λ unabhängig von i,j,k. □

Beweis von Satz I.2. Nach Identifikation von M und $\bar{M}$ via f folgt aus der Krümmungsgleichheit die Polynomgleichung von Hilfssatz 1. Da $G(X,Y)$ nach Hilfssatz 2 irreduzibel und nicht Teiler von $R(X,Y,X,Y)$ ist, folgt $\bar{G}(X,Y)=cG(X,Y)$. Daraus erhält man die Behauptung mit Hilfssatz 3. □

Bemerkung 1. Der Konformitätsfaktor λ ist eine differenzierbare Funktion ohne Nullstellen. Er ist somit überall negativ oder überall positiv. Wir nehmen im folgenden an, daß $\lambda(p)>0$ ist für alle $p\in M$. Die Metriken g und $\bar{g}$ haben dann dieselbe Signatur. Falls $\lambda(p)<0$ ist für alle $p\in M$, ergeben sich nur unwesentliche Änderungen.

Bemerkung 2. Die Voraussetzung in Theorem I.1, daß die Menge der nicht-isotropen Punkte dicht in M liegt, ist insbesondere dann erfüllt, wenn g analytisch und $K\neq\text{const}$ ist. Beweis: Falls es nämlich eine nicht-leere offene und zusammenhängende Menge U von isotropen Punkten gäbe, so wäre nach dem Satz von Schur $K_{|U}=\text{const}$, also wegen der Analytizität K überall konstant.

Bemerkung 3. Aus Theorem I.1 folgt insbesondere:

Sei $f\colon (M, g)\to(\bar{M}, \bar{g})$ ein krümmungstreuer Diffeomorphismus, und sei die Menge der nicht-isotropen Punkte dicht in M. Dann ist ein Flächenelement A von M genau dann ausgeartet, wenn das Flächenelement $f_* A$ von $\bar{M}$ ausgeartet ist.

§ 2. Konform äquivalente Metriken

Sei M eine Mgf. der Dimension $n\geqq 3$, versehen mit zwei Riem. oder ps.-Riem. konform äquivalenten Metriken g und $\bar{g}=\lambda g$, $\lambda\colon M\to\mathbb{R}^+$.

Seien $R^h{}_{ijk}$ die Komponenten des Riemannschen Krümmungstensors, $R_{ij}=R^h{}_{ihj}$ die Komponenten des Ricci-Tensors und sei $R=g^{st}R_{st}$ die skalare Krümmung. Die kovariante Ableitung wird in der invarianten Schreibweise mit D, in der Komponentendarstellung mit „ , “ bezeichnet.

Wir definieren die symmetrischen, 2-fach kovarianten Tensorfelder Q und P durch ihre Komponenten

$$Q_{ij}:=\frac{1}{n-2}\left(R_{ij}-\frac{1}{2(n-1)}R\,g_{ij}\right), \tag{1}$$

$$P_{ij}:=Q_{ij}-\beta g_{ij};\qquad \beta:=\frac{1}{n}g^{st}Q_{st}. \tag{2}$$

Der konforme Weyl-Tensor C und der Weylsche 3-Index Tensor S sind gegeben durch

$$C^h{}_{ijk}:=R^h{}_{ijk}-(Q_{ik}\delta^h{}_j-Q_{ij}\delta^h{}_k+Q^h{}_j g_{ik}-Q^h{}_k g_{ij}), \tag{3}$$

$$S_{ijk}:=Q_{ij'k}-Q_{ik'j}. \tag{4}$$

Wir setzen $\mu:=\frac{1}{2}\log\lambda$, also $\lambda=e^{2\mu}$. Ferner sei $G:=\operatorname{grad}\mu$ der kontravariante Gradient, den wir den *Konformitätsvektor* nennen wollen. Sind $(u^1, \ldots, u^n)$ lokale Koordinaten von M, so definieren wir $\mu_i:=\dfrac{\partial\mu}{\partial u^i}$, $1\leqq i\leqq n$. Das symmetrische Tensorfeld ρ sei definiert durch

$$\rho_{ij}:=\mu_{ij}-\mu_i\mu_j+\tfrac{1}{2}\mu^s\mu_s g_{ij}, \tag{5}$$

dabei ist $\mu_{ij} := \mu_{i'j}$. Man bemerkt, daß wegen $g^{ij}R_{ik'j} = \frac{1}{2}R_{'k}$ gilt:

$$g^{ij}S_{ijk} = 0. \tag{6}$$

Satz I.3. *Bei konformer Änderung der Metrik gelten die folgenden Formeln:*

(a) $\bar{\Gamma}^h_{ij} - \Gamma^h_{ij} = \mu_i \delta^h{}_j + \mu_j \delta^h{}_i - \mu^h g_{ij} =: T^h_{ij}$;

(a') $\bar{D}_X Y - D_X Y = X\mu\, Y + Y\mu\, X - \langle X, Y\rangle G =: T(X, Y)$;

(b) $\bar{R}^h{}_{ijk} - R^h{}_{ijk} = \rho_{ij}\delta^h{}_k + \rho^h{}_k g_{ij} - \rho_{ik}\delta^h{}_j - \rho^h{}_j g_{ik}$;

(b') $\bar{R}(X, Y)Z - R(X, Y)Z$
$= (D_X T)(Y, Z) - (D_Y T)(X, Z) + T(X, T(Y, Z)) - T(Y, T(X, Z))$;

(c) $\bar{R}_{ik} - R_{ik} = \rho_{ik}(2-n) - \rho^s{}_s g_{ik}$;

(d) $\lambda\bar{R} - R = 2(1-n)\rho^s{}_s$;

(e) $Q_{ik} - \bar{Q}_{ik} = \rho_{ik}$;

(f) $C^h{}_{ijk} = \bar{C}^h{}_{ijk}$;

(g) $S_{ijk} - \bar{S}_{ijk} = C^h{}_{ijk}\mu_h$;

(h) $C^h{}_{ijk'h} = (3-n)S_{ijk}$.

Für die Beweise siehe z.B. [1].

Die beiden folgenden Sätze stammen von H. Weyl [10]:

Satz I.4. *Sei $n \geqq 4$. Dann ist $C = 0$ genau dann, wenn g lokal zur Euklidischen oder zu einer pseudo-Euklidischen Metrik konform äquivalent ist. In diesem Fall ist $S = \bar{S} = 0$.*

Satz I.5. *Sei $n = 3$. Dann gilt:*

(a) *$C = 0$ und $S = \bar{S}$.*

(b) *$S = 0$ genau dann, wenn g lokal zur Euklidischen oder zu einer pseudo-Euklidischen Metrik konform äquivalent ist.*

§ 3. Beziehungen zwischen konform äquivalenten, krümmungsgleichen Räumen

Satz I.6. *Für eine Mgf. M der Dimension $n \geqq 3$, versehen mit zwei krümmungsgleichen Riem. oder ps.-Riem. Metriken g und $\bar{g} = \lambda g$, gelten folgende Formeln (mit den Bezeichnungen von § 2)*

(a) $\bar{R}^h{}_{ijk} = \lambda R^h{}_{ijk}$;

(b) $\bar{R}_{ik} = \lambda R_{ik}$;

(c) $\bar{R} = R$;

(d) $\bar{Q}_{ij} = \lambda Q_{ij}$;

(e) $\bar{C}^h{}_{ijk} = \lambda C^h{}_{ijk}$;

(f) $\frac{1}{\lambda}\bar{S}_{ijk} - S_{ijk} = P_{ij}\mu_k - P_{ik}\mu_j$;

(g) $Q_{ij}\mu^j = \beta\mu_i$ *d.h.* $P_{ij}\mu^j = 0$.

Beweis. Wegen $\bar{g}=\lambda g$ und $\bar{K}=K$ gilt $\bar{R}(X,Y,X,Y)=\lambda^2 R(X,Y,X,Y)$ zunächst für alle X, Y mit $G(X,Y)\neq 0$, aus Stetigkeitsgründen folgt diese Gleichheit aber für alle X, Y. Bekanntlich ist der Krümmungstensor durch die Werte der Form $R(X,Y,X,Y)$ bestimmt. Daraus folgt (a) und damit auch die Formeln (b) bis (e). Um (f) und (g) zu verifizieren, differenzieren wir die Beziehung (d) $\bar{Q}(Y,Z)=\lambda Q(Y,Z)$ kovariant in Richtung von X. Man erhält

$$(\bar{D}_X\bar{Q})(Y,Z)+\bar{Q}(\bar{D}_X Y,Z)+\bar{Q}(Y,\bar{D}_X Z)$$
$$=(X\lambda)Q(Y,Z)+\lambda((D_X Q)(Y,Z)+Q(D_X Y,Z)+Q(Y,D_X Z)).$$

Mit Satz I.3 (a′) folgt

$$(\bar{D}_X\bar{Q})(Y,Z)+\lambda Q(T(X,Y),Z)+\lambda Q(Y,T(X,Z))$$
$$=(X\lambda)Q(Y,Z)+\lambda(D_X Q)(Y,Z).$$

Man vertausche nun X und Y und subtrahiere die beiden Gleichungen. Nach Division durch λ ergibt sich, wenn man noch $T(X,Y)$ nach Satz I.3 (a′) einsetzt und berücksichtigt, daß $X\lambda=2\lambda\langle X,G\rangle$ ist,

$$\frac{1}{\lambda}\bar{S}(Z,Y,X)-S(Z,Y,X)=Q(Y,G)\langle X,Z\rangle+Q(Y,Z)\langle X,G\rangle$$
$$-Q(X,G)\langle Y,Z\rangle-Q(X,Z)\langle Y,G\rangle,$$

bzw. für die Komponenten

$$\frac{1}{\lambda}\bar{S}_{ijk}-S_{ijk}=Q_{js}\mu^s g_{ik}+Q_{ij}\mu_k-Q_{ks}\mu^s g_{ij}-Q_{ik}\mu_j.$$

Multipliziert man mit $g^{ij}=\lambda\bar{g}^{ij}$, so erhält man wegen (6) die Formel (g). Durch Einsetzen von (g) in die obige Gleichung findet man (f). □

Bemerkung. Liegen die nicht-isotropen Punkte dicht in M, so folgt wegen Theorem I.1 die Bedingung $\bar{g}=\lambda g$ aus der Krümmungsgleichheit.

Kapitel II. Klassifikation der krümmungsgleichen Räume

Nach Theorem I.1 ist ein krümmungstreuer Diffeomorphismus $f\colon (M,g)\to(\bar{M},\bar{g})$ zwischen Riem. oder ps.-Riem. Mgf. der Dimension $n\geqq 3$ konform (d.h. es ist $f^*\bar{g}=\lambda g$), falls die nicht-isotropen Punkte dicht in M liegen. Wir bestimmen nun (in speziellen Koordinaten) alle Metriken g und die jeweils dazugehörenden Konformitätsfaktoren λ, $\lambda\not\equiv 1$, für welche die Räume (M,g) und $(\bar{M},f_*(\lambda g))$ krümmungsgleich, aber nicht von konstanter Krümmung sind. Je nach Art der Metrik g unterscheiden wir verschiedene Fälle.

Aus den Formeln von Kapitel I ergeben sich vorerst die beiden folgenden Hilfssätze:

Hilfssatz 1. *Sei M eine Mgf. versehen mit zwei konform äquivalenten Metriken g und $\bar{g}=\lambda g$. Dann sind folgende Aussagen* (i) *und* (ii) *äquivalent:*

(i) $K=\bar{K}$,

(ii) $\bar{R}_{hijk}=\lambda^2 R_{hijk}$.

Gilt zusätzlich $C=0$, *so ist jede dieser Aussagen gleichbedeutend mit*

(iii) $Q_{ij}(1-\lambda)=\rho_{ij}$.

Beweis. Die Äquivalenz von (i) und (ii) folgt aus Satz I.6(a). Sei also $C=0$. Wegen $\bar{Q}_{ij}=\lambda Q_{ij}$ und $Q_{ij}-\bar{Q}_{ij}=\rho_{ij}$ folgt (iii) aus (ii). Die Aussage (ii) ergibt sich aus (iii) wie folgt: Aus $C=0$ folgt $\bar{C}=0$. Aus $Q_{ij}(1-\lambda)=\rho_{ij}$ und $Q_{ij}-\bar{Q}_{ij}=\rho_{ij}$ folgt $\bar{Q}_{ij}=\lambda Q_{ij}$ und damit – mit Hilfe der Definition von C – (ii). □

Hilfssatz 2. *Sei* $(M, g, \bar{g})$ *eine Mgf. versehen mit zwei Metriken. Sei* $K=\bar{K}$ *und sei die Menge der nicht-isotropen Punkte dicht, folglich* $\bar{g}=\lambda g$. *Dann folgt aus* $\lambda=\text{const}$, *daß* $\lambda=1$ *ist.*

Beweis. Mit Satz I.3(b) und Satz I.6(a) folgt $(1-\lambda)R_{hijk}=0$ und damit, wegen der Dichtheit der nicht-isotropen Punkte, $\lambda=1$. □

§4. Der Fall $C\neq 0$ (Der nicht konform-flache Fall für $n\geqq 4$)

In diesem Fall ist im Gegensatz zu den folgenden Fällen die Definitheit bzw. die Indefinitheit der Metriken nicht von Bedeutung.

Wir bemerken, daß in einem isotropen Punkt $C=0$ ist; in einem solchen Punkt $p\in M$ ist nämlich $R^h{}_{ijk}=K(g_{ik}\delta^h{}_j-g_{ij}\delta^h{}_k)$ für eine Zahl K. Folglich ist $R_{ik}=(n-1)Kg_{ik}$, $R=n(n-1)K$ und $Q_{ik}=\frac{1}{2}Kg_{ik}$ $(n=\dim M)$. Hieraus folgt nach Definition von C, daß (in p) $C=0$ ist.

Theorem II.1. *Seien* (M, g) *und* $(\bar{M}, \bar{g})$ *zwei Riem. oder ps.-Riem. Mgf. der Dimension* $n\geqq 4$ *und sei* $f: M\to\bar{M}$ *ein krümmungstreuer Diffeomorphismus. Sei* (M, g) *nirgends konform-flach. Dann ist* f *eine Isometrie.*

Beweis. Wir identifizieren M und $\bar{M}$ via f. Da C in keiner offenen Menge verschwindet, sind die isotropen Punkte – nach obiger Bemerkung – nirgends dicht. Nach Theorem I.1 ist also $\bar{g}=\lambda g$, somit einerseits nach Satz I.3 $C=\bar{C}$ und andererseits nach Satz I.6 $\bar{C}=\lambda C$; also gilt $\lambda=1$ in den Punkten mit $C\neq 0$. □

§5. Der Fall $C=S=0$ (Der konform-flache Fall)

Wir betrachten auf der Mgf. M zwei Metriken g und $\bar{g}$ mit $K=\bar{K}$, wobei die nicht-isotropen Punkte in M dicht liegen. Wir setzen nun zusätzlich voraus: $C=0$ und $S=0$ (für $n\geqq 4$ folgt $S=0$ aus $C=0$, vgl. Satz I.4). Wegen $\bar{g}=\lambda g$ (Theorem I) gilt nun auch $\bar{C}=0$ und $\bar{S}=0$ (Satz I.3(f) (g)).

Aus Satz I.6(f) schließen wir $P_{ij}\mu_k-P_{ik}\mu_j=0$, durch Multiplikation mit μ^k also, wegen Satz I.6(g)

$$P_{ij}\mu^k\mu_k=0. \tag{7}$$

§ 5.1. Der Fall $\langle G, G\rangle \neq 0$ (Definiter Fall)

In jedem Punkt M sei $G=(\mu^i)$ ein nicht-isotroper Vektor (d.h. entweder ist $G=0$ oder $\langle G, G\rangle \neq 0$). In einem Punkt p mit $G\neq 0$ ist also wegen (7) $P_{ij}=\frac{1}{n-2}\left(R_{ij}-\frac{R}{n}g_{ij}\right)=0$, folglich, wegen $C=0$, $K_{|\pi^{-1}(p)}=\text{const}$, d.h. p ist isotroper Punkt. Da nach Voraussetzung die Menge der nicht-isotropen Punkte dicht in M liegt, so ist aus Stetigkeitsgründen $G=0$, folglich $\lambda=e^{2\mu}=\text{const}$, also hat man wegen Hilfssatz 2, daß $\lambda=1$ ist.

Damit folgt das

Theorem II.2. *Seien (M, g) und $(\bar{M}, \bar{g})$ zwei Riem. Mgf. der Dimension $n\geqq 3$ und sei $f: M\to\bar{M}$ ein krümmungstreuer Diffeomorphismus. Sei (M, g) konform flach und sei die Menge der nicht-isotropen Punkte dicht in M. Dann ist f eine Isometrie. Sind (M, g) und $(\bar{M}, \bar{g})$ ps.-Riem. Mgf. und ist der Konformitätsvektor nirgends isotrop, so ist f eine Isometrie.*

Bemerkung. Theorem II.1 und Theorem II.2 sind schon in [4] und [5] bewiesen worden und werden hier mehr vollständigkeitshalber aufgeführt.

§ 5.2. Der Fall $\langle G, G\rangle = 0$ (Indefiniter Fall)

In diesem Abschnitt sei $\langle G, G\rangle = 0$. Wir setzen weiter voraus, daß die Menge $\{p\in M \mid \lambda(p)\neq 1\}$ dicht in M liegt. Die Punkte p mit $G(p)=0$ sind dann nirgends dicht (Hilfssatz 2). Die folgenden Gleichungen, die zunächst nur auf der Menge $\{p\in M \mid (\lambda(p)-1)G(p)\neq 0\}$ gelten, sind aus Stetigkeitsgründen auf ganz M gültig.

Aus $Q_{ik}(1-\lambda)=\rho_{ik}$ (vgl. Hilfssatz 1) und $Q_{ik}\mu^i=\beta\mu_k$ (vgl. Satz I.6) folgt, wegen $0=(\mu^i\mu_i)_k=2\mu_{ik}\mu^i$,

$$\beta\mu_k(1-\lambda)=\rho_{ik}\mu^i=\mu_{ik}\mu^i-\mu_i\mu_k\mu^i=0,$$

d.h. es ist notwendigerweise

$$\beta=0, \qquad R=0.$$

Aus Satz I.6 (f) folgt somit $Q_{ik}\mu_j-Q_{ij}\mu_k=0$, was bedeutet, daß gilt:

$$Q_{ij}=\phi\,\mu_i\mu_j, \tag{8}$$

für eine Funktion $\phi: M\to\mathbb{R}$. Durch Differenzieren dieser Gleichung erhält man unter Verwendung von Hilfssatz 1, der Formel (8) und der Definition von ρ_{ij}

$$Q_{ij'k}=v_kQ_{ij}, \qquad v_k=\frac{\phi_k}{\phi}+2\mu_k(1+\phi(1-\lambda)) \tag{9}$$

in Punkten p mit $\phi(p)\neq 0$ (dies sind genau die nicht-isotropen Punkte). Wegen $C=0$ gilt also $R_{hijk'l}=v_lR_{hijk}$, d.h. (M, g) ist ein rekurrenter oder symmetrischer Raum. Aus $S=0$ und (9) folgt weiter

$$\mu_k=\kappa v_k \tag{10}$$

falls $(v_k) \neq 0$. Wegen $C=0$ und (8) hat man auch

$$\sum_{\substack{j,k,l \\ \text{zykl.}}} R_{hijk}\mu_l = 0. \tag{11}$$

Wir werden damit auf Räume geführt, die von Walker in [9] untersucht worden sind:

Definition (Walker). Ein nicht-flacher Riem. oder ps.-Riem. Raum (M, g) der Dimension n heißt ein K_n^*-*Raum*, falls eine der folgenden Bedingungen (1), (2) erfüllt ist:

(1) (M, g) ist rekurrent, d.h. es ist $R_{hijk'l} = v_l R_{hijk}$ für einen Vektor $(v_i) \neq 0$.

(2) (M, g) ist symmetrisch (d.h. es ist $R_{hijk'l} = 0$) und es gilt $\sum_{\substack{j,k,l \\ \text{zykl.}}} R_{hijk} v_l = 0$ für einen Vektor $(v_i) \neq 0$.

Walker bewies, daß der Rekurrenzvektor (v_i) als Gradientenvektor $(v_i) = \operatorname{grad} v$ angenommen werden kann. Wir nennen v eine *Rekurrenzfunktion.*

Satz (Walker). *Sei (M, g) ein konform-flacher K_n^*-Raum mit $n \geqq 3$.*

(a) *Dann existieren Koordinaten $(u^1, \ldots, u^n)$, so daß die Metrik die Form hat*

$$ds^2 = \psi (du^1)^2 + 2\, du^1\, du^2 + k_{\alpha\beta}\, du^\alpha\, du^\beta,$$

wobei $\psi = \theta(u^1) k_{\alpha\beta} u^\alpha u^\beta + b_\alpha(u^1) u^\alpha$ und $(k_{\alpha\beta})$, $\alpha, \beta = 3, \ldots, n$, eine konstante, symmetrische, nicht-ausgeartete $(n-2) \times (n-2)$-Matrix ist. Die Rekurrenzfunktion v ist eine Funktion von u^1.

(b) *Für die inverse Matrix gilt dann*

$$g^{1i} = \delta_2^i, \quad g^{22} = -\psi, \quad g^{\alpha\beta} = k^{\alpha\beta}, \quad g^{2\alpha} = 0,$$

mit $1 \leqq i \leqq n$, $3 \leqq \alpha, \beta \leqq n$.

Die nichtverschwindenden Größen sind

$$\Gamma_{11'1} = \tfrac{1}{2}(\theta' k_{\alpha\beta} u^\alpha u^\beta + b'_\alpha u^\alpha); \quad \Gamma_{11'\alpha} = -\Gamma_{1\alpha'1} = -\theta k_{\alpha\beta} u^\beta - \tfrac{1}{2} b_\alpha;$$
$$\Gamma_{11}^2 = \Gamma_{11'1}; \quad \Gamma_{1\alpha}^2 = -\Gamma_{11'\alpha}; \quad \Gamma_{11}^\alpha = -\theta u^\alpha - \tfrac{1}{2} k^{\alpha\beta} b_\beta;$$
$$R_{1\alpha1\beta} = -\theta k_{\alpha\beta}; \quad R_{11} = -(n-2)\theta; \quad Q_{11} = -\theta;$$
$$R_{1\alpha1\beta'1} = -\theta' k_{\alpha\beta}.$$

Dabei bedeutet „$'$“ die Ableitung nach u^1.

(c) *Es gelten die Äquivalenzen:* (M, g) *symmetrisch* $\Leftrightarrow \theta = \text{const}$; $K = \text{const} \Leftrightarrow \theta = 0$.

Erlaubt also ein konform-flacher Raum (in dem die nicht-isotropen Punkte dicht liegen) einen nicht-trivialen krümmungstreuen Diffeomorphismus, so ist der Raum notwendigerweise ein K_n^*-Raum. Der Konformitätsvektor ist dabei proportional zum Rekurrenzvektor (vgl. (10), (11), Satz von Walker).

Wir beweisen nun das

Theorem II.3. *Die konform-flachen Räume der Dimension $n \geqq 3$, welche nicht-triviale krümmungstreue Diffeomorphismen erlauben, sind genau die konform-*

flachen K_n^*-Räume. Die Diffeomorphismen sind konform und der Konformitätsfaktor ist eine Funktion der Rekurrenzfunktion.

Beweis. Die Notwendigkeit wurde oben bewiesen. Sei also (M,g) ein konformflacher K_n^*-Raum, $n\geqq 3$, mit Rekurrenzfunktion v. Wir wählen die speziellen Koordinaten nach dem Satz von Walker und setzen $\mu=\mu(u^1)$. Es ist also $(\mu_i)=(\mu',0,\ldots,0)$ und $\mu_{ij}=\mu''\delta^1{}_i\delta^1{}_j$. Die Gleichungen $Q_{ij}(1-\lambda)=\rho_{ij}$ (vgl. Hilfssatz 1) reduzieren sich auf $Q_{11}(1-\lambda)=\rho_{11}$, d.h. auf $-\theta(u^1)(1-e^{2\mu(u^1)})=\mu''-(\mu')^2$. Für gegebenes θ (mit nirgends dichten Nullstellen) besitzt diese Gleichung eine 2 parametrige Schar von Lösungen. □

Ein explizites Beispiel:

$$\theta(u^1)=\frac{e^{2u^1}}{e^{2u^1}-1};\qquad \lambda(u^1)=e^{2\mu(u^1)}=\theta(u^1)-1=\frac{1}{e^{2u^1}-1}.$$

§6. Der Fall der Dimension drei

In diesem Paragraphen wird nun der Fall der Dimension drei vollständig behandelt.

Sei $(M,g,\bar{g})$ eine Mgf. der Dimension $n=3$, versehen mit zwei krümmungsgleichen Riem. oder ps.-Riem. Metriken, wobei die Menge der nicht-isotropen Punkte dicht in M ist. Aus Theorem I.1 folgt die Konformität: $\bar{g}=\lambda g$.

Ist (M,g) konform-flach (d.h. $S=0$), so wissen wir aus §5, daß bei definiten Metriken nur der triviale Fall $\bar{g}=g$ möglich ist; im pseudo-Riemannschen Fall kann $\bar{g}\neq g$ sein: (M,g) ist dann ein konform-flacher K_3^*-Raum und $\bar{g}$ eine gewisse zu g passende konform äquivalente Metrik. Diese Beispiele von §5.2 werden in §6.3 nochmals auftreten.

Sei die Menge $\{p\in M\mid\lambda(p)\neq 1\}$ dicht in M; die Punkte p mit $G(p)=0$ sind folglich nirgends dicht. Ferner sei M zusammenhängend. Wir setzen (mit den Bezeichnungen von §2)

$$\alpha:=\tfrac{1}{2}\langle G,G\rangle=\tfrac{1}{2}\mu^s\mu_s;\qquad \beta:=\tfrac{1}{3}g^{st}Q_{st};\qquad \gamma:=\tfrac{1}{2}P_{ij}P^{ij}.$$

Im folgenden erweist sich die invariante Schreibweise als nützlich. Den zu einem (0,2)-Tensor T gehörigen (1,1)-Tensor bezeichnen wir wieder mit T: $T(X,Y)=\langle T(X),Y\rangle$; entsprechend schreiben wir bei einem (0,3)-Tensor S:

$$S(Z,X,Y)=\langle S(X,Y),Z\rangle.$$

Wir benötigen nun die folgenden grundlegenden Formeln:

$$D_XG=\langle X,G\rangle G-\alpha X+(1-\lambda)Q(X) \tag{12}$$

$$D_XG=\langle X,G\rangle G-(\alpha-(1-\lambda)\beta)X+(1-\lambda)P(X) \tag{12'}$$

$$\operatorname{Sp}P=0 \tag{13}$$

$$P(G)=0 \tag{14}$$

$$(D_X P)(G) = -P(D_X G) = (\alpha - (1-\lambda)\beta)P(X) - (1-\lambda)P^2(X) \tag{15}$$

$$\frac{\lambda-1}{\lambda} S(Y,X) = \langle Y, G\rangle P(X) - \langle X, G\rangle P(Y)$$

$$= \frac{\lambda-1}{\lambda}((D_X P)(Y) - (D_Y P)(X) - \rho(\langle Y, G\rangle X - \langle X, G\rangle Y)), \tag{16}$$

$$\rho = \begin{cases} -\frac{1}{2}(1-\lambda)\frac{\gamma}{\alpha} & \text{falls } \alpha \neq 0 \\ 0 & \text{falls } \alpha \equiv 0. \end{cases}$$

(12) ist die invariante Schreibweise von $Q_{ij}(1-\lambda) = \rho_{ij}$ (Hilfssatz 1, Kap. II); (12′) ergibt sich, wenn man Q durch P ersetzt. (13) und (14) folgen aus der Definition von P und aus Satz I.6 (g). Differenziert man (14) in Richtung X und benützt (12′), so erhält man (15). Die erste Gleichheit von (16) folgt wegen $S = \bar{S}$ aus Satz I.6 (f), die zweite ergibt sich wie folgt: Multipliziert man (15) skalar mit Y, vertauscht X und Y und subtrahiert, so erhält man $(D_X P)(G, Y) - (D_Y P)(G, X) = 0$. Mit $P = Q - \beta g$ folgt also $S(G, Y, X) - X\beta\langle G, Y\rangle + Y\beta\langle G, X\rangle = 0$. Da aber $S(G, X, Y) = 0$ ist (Satz I.6 (f), (g)), ist obige Formel äquivalent zu

$$X\beta = \rho\langle X, G\rangle \tag{17}$$

für eine noch zu bestimmende Funktion ρ. Nach Definition von S ist $S(Y, X) = (D_X Q)(Y) - (D_Y Q)(X)$. Ersetzt man hier Q durch P und benützt (17), so folgt die zweite Gleichheit von (16). Zur Berechnung von ρ bilden wir in (15) die Spur. Aus der Formel (6) folgt $\mathrm{Sp}(X \mapsto (D_X Q)(Y)) = 3Y\beta$. Wegen $P = Q - \beta g$ erhält man also auf der linken Seite von (15) $\mathrm{Sp}(X \mapsto (D_X P)(G)) = 3G\beta - G\beta = 4\alpha\rho$. Andererseits erhält man für die Spur der rechten Seite $-(1-\lambda)2\gamma$. Durch Vergleich folgt

$$2\alpha\rho = -(1-\lambda)\gamma \tag{18}$$

und daraus (16), falls $\alpha \neq 0$ ist. Ist $\alpha \equiv 0$, so hat man auch $D_G G \equiv 0$. Setzt man in (12′) $X = G$, so folgt $(1-\lambda)\beta G \equiv 0$, also $\beta \equiv 0$. Mit (17) hat man folglich $\rho \equiv 0$.

Satz II.4. *Die Funktionen* $\alpha = \frac{1}{2}\langle G, G\rangle$, $\beta = \frac{1}{3}\mathrm{Sp}(Q)$, $\gamma = \frac{1}{2}\mathrm{Sp}(P^2)$ *sind Funktionen von* μ. *Bezeichnet* „′“ *die Ableitung nach* μ, *so erfüllen sie im Falle* $\alpha \neq 0$ *die Differentialgleichungen (in Punkten* p *mit* $\lambda(p) \neq 1$)

$$\alpha' = \alpha + (1-\lambda)\beta \tag{19}$$

$$\beta' = -\frac{1}{2}(1-\lambda)\frac{\gamma}{\alpha} \tag{20}$$

$$\gamma' = \frac{1+\lambda}{1-\lambda}\gamma - (1-\lambda)\frac{\beta\gamma}{\alpha}. \tag{21}$$

Falls $\alpha \equiv 0$ *ist, so ist* $\beta \equiv 0$ *und* $\gamma \equiv 0$.

Beweis. Für eine Funktion $\phi = \phi(\mu)$ ist $X\phi = \phi' X\mu$. Wegen (12′) und (14) gilt $X\alpha = \langle D_X G, G\rangle = (\alpha + (1-\lambda)\beta)X\mu$. Daraus folgt $\alpha = \alpha(\mu)$, (19) und $\beta = \beta(\mu)$. Aus

(17) schließt man $\beta'=\rho$; zusammen mit (16) ergibt dies (20) und $\gamma=\gamma(\mu)$. Es genügt nun noch, $G\gamma=\mathrm{Sp}(P\circ D_G P)$ zu berechnen. Setzt man in (16) $X=G$, so findet man

$$(D_G P)(Y)=\frac{\lambda}{1-\lambda}2\alpha P(Y)+(D_Y P)(G)+\rho(\langle Y,G\rangle G-2\alpha Y),$$

bzw. wegen (15)

$$(D_G P)(Y)=\left(\frac{1+\lambda}{1-\lambda}\alpha-(1-\lambda)\beta\right)P(Y)-(1-\lambda)P^2(Y)+\rho(\langle Y,G\rangle G-2\alpha Y).$$

Anwendung von P auf diese Gleichung liefert wegen $P(G)=0$

$$P\circ D_G P=\left(\frac{1+\lambda}{1-\lambda}\alpha-(1-\lambda)\beta\right)P^2-(1-\lambda)P^3-2\alpha\rho P.$$

Durch Spurbildung folgt wegen (13)

$$\mathrm{Sp}(P\circ D_G P)=\left(\frac{1+\lambda}{1-\lambda}\alpha-(1-\lambda)\beta\right)2\gamma-(1-\lambda)\mathrm{Sp}(P^3).$$

Da ein Eigenwert von P gleich Null ist und die zwei anderen entgegengesetzt sind, gilt $\mathrm{Sp}(P^3)=0$. Somit erhält man schließlich

$$G\gamma=2\alpha\gamma'=\left(\frac{1+\lambda}{1-\lambda}\alpha-(1-\lambda)\beta\right)2\gamma$$

und daraus die Gleichung (21). Ist $\alpha\equiv 0$, so folgt aus (19), daß $\beta\equiv 0$ ist; aus (18) folgt $\gamma\equiv 0$. □

Bemerkung. Aus obigen Differentialgleichungen folgt durch Differenzieren sofort das Lemma von Nasu, das besagt, daß gilt:

$$\left(\frac{1}{\lambda}-1\right)\alpha\gamma=\text{const.} \tag{22}$$

Lemma II.5. *Die Orthogonaltrajektoren der Flächen* $\mu=\text{const}$ *sind geodätische Linien.*

Beweis. Aus (12) folgt $D_G G=\alpha' G$. □

Aus (22) und aus Satz II.4 folgt, daß die Menge der Punkte, in denen γ verschwindet, offen in M liegt. Ist also M zusammenhängend, so hat γ keine Nullstellen, außer γ verschwinde identisch. Wir unterscheiden daher folgende drei Fälle: 1) $\gamma>0$ auf M. 2) $\gamma<0$ auf M. 3) $\gamma=0$ auf M.

§ 6.1. Der Fall $\gamma>0$

Sei $\gamma(p)>0$ für alle $p\in M$. Auf M existieren somit keine isotropen Punkte. Aus (22) und (18) folgt, daß $\lambda(p)\neq 1$ und $\alpha(p)\neq 0$ für alle $p\in M$.

Die Eigenwerte von P sind 0, ω, $-\omega$, mit $\omega(\mu)=\sqrt{\gamma}>0$. Die dazu gehörenden Eigenvektoren G, V, W sind linear unabhängig, paarweise orthogonal und nicht-isotrop.

Satz II.6. *Sind G, V, W, $\langle V,V\rangle=\varepsilon_1$, $\langle W,W\rangle=\varepsilon_2$, $\varepsilon_1,\varepsilon_2=\pm 1$, normierte Eigenvektoren von P zu den Eigenwerten 0, ω, $-\omega$, so gilt für deren kovariante Ableitungen:*

$$D_G G=(\alpha+(1-\lambda)\beta)G; \quad D_V G=c_1 V; \quad D_W G=c_2 W;$$

$$D_G V=0; \quad D_V V=-\frac{1}{2\alpha}\varepsilon_1 c_1 G; \quad D_W V=0;$$

$$D_G W=0; \quad D_V W=0; \quad D_W W=-\frac{1}{2\alpha}\varepsilon_2 c_2 G.$$

Dabei ist $c_1=(1-\lambda)(\omega+\beta)-\alpha$; $c_2=(1-\lambda)(-\omega+\beta)-\alpha$.

Beweis. Man hat

$$P(G)=0; \quad P(V)=\omega V; \quad P(W)=-\omega W \tag{23}$$

$$\langle G,V\rangle=\langle G,W\rangle=\langle V,W\rangle=0; \quad \langle V,V\rangle=\varepsilon_1; \quad \langle W,W\rangle=\varepsilon_2. \tag{24}$$

Die Formeln der Ableitungen von G, d.h. für $D_G G$, $D_V G$ und $D_W G$ folgen mit (12′). Als erstes zeigen wir das Verschwinden von $D_G V$. Aus der zweiten Gleichung von (23) erhalten wir durch Differenzieren in Richtung von G

$$(D_G P)(V)+P(D_G V)=(G\omega)V+\omega D_G V. \tag{25}$$

Nun setzen wir in (16) $X=G$ und $Y=V$ und erhalten

$$(D_G P)(V)-(D_V P)(G)+2\rho\alpha V=-\frac{\lambda}{\lambda-1}2\alpha\omega V. \tag{26}$$

Aus (15) folgt ferner $(D_V P)(G)=-c_1\omega V$ und damit aus (26)

$$(D_G P)(V)=\hat{c}V; \quad \hat{c}=-2\alpha\omega\frac{\lambda}{\lambda-1}-2\rho\alpha-c_1\omega. \tag{27}$$

Setzt man nun den Ansatz $D_G V=aV+bW+cG$ in (25) ein, so liefert (27) nach einem Koeffizientenvergleich $b=c=0$, also $D_G V=aV$. Aus $\langle V,V\rangle=\varepsilon_1$ folgt aber durch Differenzieren in Richtung von G: $0=\langle D_G V,V\rangle=\pm a$. Somit gilt $D_G V=0$. Ganz analog folgt $D_G W=0$.

Als nächstes zeigen wir das Verschwinden von $D_V W$ und $D_W V$. Zunächst folgt aus den zwei letzten Gleichungen von (24), daß $D_W V=a_1 W+a_2 G$ bzw. $D_V W=b_1 V+b_2 G$ ist. Differenziert man nun $\langle G,V\rangle=0$ in Richtung von W bzw. $\langle G,W\rangle=0$ in Richtung von V, so folgt mit Hilfe der Formeln für $D_W G$ bzw. $D_V G$, daß $a_2=b_2=0$ ist. Jetzt differenziert man noch die zweite Gleichung von (23) in Richtung von W bzw. die dritte in Richtung von V; man erhält wegen $\omega=\omega(\mu)$

$$(D_W P)(V)+P(D_W V)=\omega D_W V, \quad (D_V P)(W)+P(D_V W)=-\omega D_V W.$$

Subtraktion dieser Gleichungen und Anwendung von (16) liefert $P([W,V])$ $=\omega D_W V+\omega D_V W$. Durch Einsetzen von $D_W V=a_1 W$ bzw. $D_V W=b_1 V$ folgt a_1 $=b_1=0$.

Es bleiben nun noch die Formeln für $D_V V$ und $D_W W$ zu verifizieren. Aus $\langle V,V\rangle=\varepsilon_1$, $\langle W,W\rangle=\varepsilon_2$ und $\langle V,W\rangle=0$ folgt durch jeweilige Differentiation in Richtung nach V bzw. W mit Hilfe der eben bewiesenen Gleichungen, daß $D_V V$ bzw. $D_W W$ proportional zu G sind. Die entsprechenden Koeffizienten bestimmen sich leicht durch Differenzieren von $\langle V,G\rangle=0$ bzw. $\langle W,G\rangle=0$ in Richtung von V bzw. W mit Hilfe der Formeln für $D_V G$ bzw. $D_W G$. □

Wir versuchen nun, die Eigenvektoren G, V, W in Satz I.6 so mit Faktoren zu multiplizieren, daß die Lie-Klammern von je zweien verschwinden; wir setzen an: $X^*=\phi_1(\mu)V$, $Y^*=\phi_2(\mu)W$. Dann folgt $[X^*,G]=\phi_1[V,G]-(G\phi_1)V$ $=(c_1(\mu)\phi_1-G\phi_1)V$. Wir wählen also für ϕ_1 eine (nicht-verschwindende) Lösung der Differentialgleichung $G\phi=c_1\phi$. Analog hat man $[Y^*,G]=(c_2(\mu)\phi_2$ $-G\phi_2)W$, und man nimmt für ϕ_2 eine Lösung von $G\phi=c_2\phi$. Da ϕ_1, ϕ_2 Funktionen von μ sind, ist dann auch $[X^*,Y^*]=0$.

Da die Lie-Klammern verschwinden, gibt es lokale Koordinaten u, v, w so, daß $\frac{1}{2\alpha}G$, X^*, Y^* die Tangentialvektoren der Koordinatenlinien sind, d.h. man hat

$$ds^2=\frac{1}{2\alpha}du^2+\varepsilon_1(\phi_1(u))^2\,dv^2+\varepsilon_2(\phi_2(u))^2\,dw^2,\qquad \mu=u, \tag{28}$$

wobei

$$\begin{aligned}&\phi_1=\exp\left(\int\frac{c_1}{2\alpha}du\right),\qquad &&\phi_2=\exp\left(\int\frac{c_2}{2\alpha}du\right),\\ &c_1=(1-\lambda)(\sqrt{\gamma}+\beta)-\alpha,\qquad &&c_2=(1-\lambda)(-\sqrt{\gamma}+\beta)-\alpha.\end{aligned} \tag{29}$$

Dabei sind α, β, γ Lösungen der Differentialgleichungen (19)–(21). Die in (28), (29) bestimmten Metriken g und $\bar{g}=e^{2\mu}g$ sind i.A. noch nicht krümmungsgleich. Wie wir sehen werden, tritt noch eine Forderung an die Anfangsbedingungen der Funktionen α, β, γ auf.

Hilfssatz II.7. *Für die in* (28), (29) *bestimmten Metriken* g *und* $\bar{g}=e^{2\mu}g$ *sind die nicht-verschwindenden Größen des Riemann-Christoffel-Tensors gegeben durch*

$$R(X^*,G,X^*,G)=2\alpha(2\beta+\omega)\varepsilon_1\phi_1^2;\qquad R(Y^*,G,Y^*,G)=2\alpha(2\beta-\omega)\varepsilon_2\phi_2^2;$$

$$R(X^*,Y^*,X^*,Y^*)=-\frac{1}{2\alpha}\varepsilon_1\varepsilon_2\phi_1^2\phi_1^2c_1c_2;$$

$$\bar{R}(X^*,G,X^*,G)=\lambda^2R(X^*,G,X^*,G);\qquad \bar{R}(Y^*,G,Y^*,G)=\lambda^2R(Y^*,G,Y^*,G);$$

$$\bar{R}(X^*,Y^*,X^*,Y^*)=\lambda^2R(X^*,Y^*,X^*,Y^*)-\lambda(1-\lambda)\varepsilon_1\varepsilon_2\phi_1^2\phi_2^2\frac{1}{2\alpha}(c_1c_2+4\alpha\beta).$$

Beweis. Aus Satz II.6 erhält man leicht die kovarianten Ableitungen für die Vektoren G, X^*, Y^* und damit – unter Benützung von (19)–(21) – die Formeln für den Krümmungstensor von g. Die Werte des Krümmungstensors von $\bar{g}$ folgen mit der Beziehung (b) von Satz I.3, wobei die ρ_{ij} gemäß (5) aus (28), (29) zu berechnen sind. □

Die Bedingungen $\bar{R}(A,B,C,D)=\lambda^2 R(A,B,C,D)$ (vgl. Hilfssatz 1, Kap. II) für Krümmungsgleichheit reduzieren sich also auf

$$\delta := (1-\lambda)^2(\beta^2-\gamma)+2(1+\lambda)\alpha\beta+\alpha^2=0. \tag{30}$$

Wir bemerken, daß auf Grund von Satz II.4 gilt: $\delta'=2\delta$. Wählt man für ein $\mu_0 \neq 0$ die Anfangswerte $\alpha(\mu_0)$, $\beta(\mu_0)$, $\gamma(\mu_0)$ so, daß $\delta(\mu_0)=0$ ist, so ist also $\delta \equiv 0$. Wir haben somit das

Theorem II.8. *Zwei dreidimensionale Metriken ds^2, $d\bar{s}^2$ mit $\gamma>0$ haben dann und nur dann dieselbe Riemannsche Krümmung, wenn sich ds^2 durch eine lokale Koordinatentransformation auf die Form* (28), (29) *bringen läßt. Dabei sind α, β, γ Lösungen der Differentialgleichungen* (19)–(21) *mit Anfangsbedingungen $\alpha(\mu_0)\neq 0$, $\gamma(\mu_0)>0$, $\delta(\mu_0)=0$ für ein $\mu_0 \neq 0$, und es ist $d\bar{s}^2=e^{2\mu}ds^2$.*

Bemerkungen. 1) Durch Lösen der Gleichung (30) für β, findet man wegen der Wahl obiger Anfangsbedingungen, daß der Anfangswert $\beta(\mu_0)$ reell ist.

2) Sind $\phi_1(u)$, $\phi_2(u)$ Lösungen von $G\phi=c_1\phi$, $G\phi=c_2\phi$, so auch $\phi_1(u)\psi_1(v,w)$, $\phi_2(u)\psi_2(v,w)$. Wegen der Forderung $[X^*,Y^*]=0$ folgt $\psi_1=\psi_1(v)$, $\psi_2=\psi_2(w)$. Diese Funktionen lassen sich aber wegtransformieren.

3) Die ersten Beispiele nicht-trivialer krümmungstreuer Diffeomorphismen – sie wurden von Yau [11] gefunden – waren genau von der Art (28) mit $\varepsilon_1=\varepsilon_2=1$ und zunächst beliebigen $\phi_1(u)$, $\phi_2(u)$, $\lambda(u)$. Yau ging mit einem solchen Ansatz in die Gleichungen $\bar{R}_{hijk}=\lambda^2 R_{hijk}$ und zeigte die Existenz von Lösungen. Anstelle der 4 Differentialgleichungen für ϕ_1, ϕ_2, λ treten bei uns die 4 Gleichungen (19)–(21) und $\delta'=2\delta$ für die drei Invarianten α, β, γ auf. Nach Theorem II.8 sind also die Beispiele von Yau im Riemannschen Fall die einzigen.

§ 6.2. Der Fall $\gamma<0$

Sei $\gamma<0$ auf M. Auf M existieren somit keine isotropen Punkte, und wir haben wieder wegen (22) und (18) $\lambda(p)\neq 1$, $\alpha(p)\neq 0$ für alle $p\in M$. Neben dem Eigenwert 0 besitzt P jetzt die beiden rein imaginären Eigenwerte ω, $-\omega$, mit $\omega := \sqrt{\gamma}$, deren dazugehörige Eigenvektoren V, W konjugiert komplex sind, d.h. es gilt $W=\bar{V}$. Die Eigenvektoren G, V, W sind linear unabhängig, paarweise orthogonal und nicht-isotrop (bezüglich der komplexen Fortsetzung der Bilinearform $g(X,Y)$). Wir normieren $\langle V,V\rangle=\langle W,W\rangle=2$.

Da die Gleichungen (12)–(16) auch für komplexe Vektoren X, Y ihre Gültigkeit behalten, gelten die für den Fall $\gamma>0$ durchgeführten Rechnungen wörtlich auch für den Fall $\gamma<0$. Die Eigenvektoren $V^*=\phi_1 V$, $W^*=\phi_2 W=\bar{V}^*$, deren Lie-Klammern verschwinden, sind jetzt komplex, und es ist $\phi_1=\psi_1 e^{i\psi_2}$, $\phi_2=\bar{\phi}_1$, wobei

$$\begin{aligned} &\psi_1=\exp\left(\int\frac{d_1}{2\alpha}d\mu\right); \qquad \psi_2=\int\frac{d_2}{2\alpha}d\mu; \\ &d_1=-\alpha+(1-\lambda)\beta; \qquad d_2=(1-\lambda)\sqrt{-\gamma}. \end{aligned} \tag{31}$$

Die durch die Skalarprodukte dieser Vektorfelder als Basisfelder bestimmten komplexifizierten Metriken g und $e^{2\mu}g$ sind genau dann krümmungsgleich, wenn $\delta=0$ ist. Als reelle Vektorfelder nehmen wir nun neben G noch $X^*=\mathrm{Re}\,V^*$, $Y^*=\mathrm{Im}\,V^*$. Nun verschwinden auch die Lie-Klammern von G, X^*, Y^*. Wegen

$$\langle X^*, X^*\rangle = \tfrac{1}{4}\langle V^*+\bar{V}^*, V^*+\bar{V}^*\rangle = \tfrac{1}{2}(\psi_1^2 e^{2i\psi_2}+\psi_1^2 e^{-2i\psi_2}) = \psi_1^2\cos 2\psi_2,$$

$$\langle X^*, Y^*\rangle = \frac{1}{4i}\langle V^*+\bar{V}^*, V^*-\bar{V}^*\rangle = \frac{1}{2i}(\psi_1^2 e^{2i\psi_2}-\psi_1^2 e^{-2i\psi_2}) = \psi_1^2\sin 2\psi_2$$

und $\langle Y^*, Y^*\rangle = -\psi_1^2\cos 2\psi_2$ erhalten wir das

Theorem II.9. *Erlaubt eine ps.-Riem. Mgf. (M,g) der Dimension drei einen nichttrivialen krümmungstreuen Diffeomorphismus mit $\gamma<0$, so ist der Diffeomorphismus konform, und es existieren Koordinaten (u,v,w), in denen die Metrik g und der Konformitätsfaktor $e^{2\mu}$ die folgende Form haben*

$$ds^2 = \frac{1}{2\alpha(u)}du^2+(\psi_1(u))^2\cos 2\psi_2(u)(dv^2-dw^2)$$
$$-2(\psi_1(u))^2\sin 2\psi_2(u)\,dv\,dw; \qquad \mu=u.$$

Dabei sind ψ_1, ψ_2 gegeben durch die Gleichungen (31), *und α, β, γ sind Lösungen von* (19)–(21) *mit den Anfangsbedingungen $\alpha(\mu_0)\neq 0$, $\gamma(\mu_0)<0$, $\delta(\mu_0)=0$, $4e^{2\mu_0}\alpha^2(\mu_0)+(1-e^{2\mu_0})^4\gamma(\mu_0)\geqq 0$, für ein $\mu_0\neq 0$. Umgekehrt ist jedes derartige Paar von Räumen (M,g) und $(M,\bar{g})$, $\bar{g}=e^{2\mu}g$, krümmungsgleich.*

Bemerkung. Die obige Ungleichung für die Anfangswerte garantiert, daß der Wert $\beta(\mu_0)$ reell ist.

§ 6.3. Der Fall $\gamma=0$

Sei $\gamma=0$ auf M, aber $P\neq 0$; somit sind alle Eigenwerte von P gleich Null; die Menge der nicht-isotropen Punkte sei dicht in M. Es zeigt sich, daß man im wesentlichen zwei Unterfälle unterscheiden muß: (A) $\alpha=0$ auf M; (B) $\alpha\neq 0$ auf M. Im Falle (B) nehmen wir an, daß die Menge $\{p\in M\,|\,\alpha(p)\neq 0\}$ dicht in M ist. Die beiden Unterfälle lassen sich aber doch in großen Teilen parallel behandeln.

Hilfssatz II.10. *Unter obigen Voraussetzungen ist $\beta=\mathrm{const}$. Ist $\alpha\equiv 0$, so ist sogar $\beta\equiv 0$.*

Beweis. Falls $\alpha\neq 0$ ist, folgt aus (20), daß $\beta=\mathrm{const}$ ist. Ist $\alpha\equiv 0$ auf M, so ergibt sich die Behauptung aus (19). □

Satz II.11. *Es existieren Koordinaten (u,v,w), so daß die Metrik und der Konformitätsfaktor auf folgende Form gebracht werden können:*

$$ds^2 = A\,du^2+2\,du\,dv+2B\,du\,dw+dw^2, \qquad \mu=\mu(u,w).$$

Ist $\alpha\equiv 0$, so kann $\mu=u$ gewählt werden.

Beweis. (A) Sei $\alpha\equiv 0$. Wir führen „isotrope" Gaußsche Koordinaten (u^1,u^2,u^3) ein: Die Differentialgleichung $g^{ij}\mu_i\frac{\partial\eta}{\partial u^j}=0$ hat zwei unabhängige Lösungen μ, η. Von der Differentialgleichung $g^{ij}\mu_i\frac{\partial\theta}{\partial u^j}=1$ wählen wir eine von μ und η unabhängige Lösung θ und setzen $u^1=\mu$, $u^2=\theta$, $u^3=\eta$. Es folgt $g^{11}=0$, $g^{12}=1$, $g^{13}=0$, $\mu=u_1$; also

$$(g_{ij})=\begin{pmatrix} A & 1 & B\\ 1 & 0 & 0\\ B & 0 & C\end{pmatrix},\qquad \begin{aligned} g^{22}&=\frac{B^2}{C}-A,\\ g^{23}&=-\frac{B}{C},\quad g^{33}=\frac{1}{C}.\end{aligned}$$

Aus Hilfssatz II.10 folgt $\beta=0$. Damit gilt $Q_{ij}\mu^j=Q_{2i}=0$, also $\beta=\frac{1}{3}g^{33}Q_{33}=0$, und somit $Q_{33}=0$. Die Bedingung $Q_{33}(1-\lambda)=\rho_{33}$ liefert nun: $0=\rho_{33}=\mu_{33}-(\mu_3)^2=\mu_{33}=\frac{1}{2}(\partial C/\partial u^2)$; es ist also $C=C(u^1,u^3)$. Die Transformation $u=u^1$, $v=u^2$, $w=h(u^1,u^3)$, $\frac{\partial h}{\partial u^3}=\sqrt{C}$, ergibt nun die gewünschte Form von (g_{ij}). (Bemerkung: Die obige Annahme $C>0$ besagt, daß wir uns – der Einfachheit halber – auf den Fall $\det(g_{ij})=-1$ beschränken.)

(B) Sei $\alpha\neq 0$. Dann besitzt P noch einen von G linear unabhängigen Eigenvektor zum Eigenwert 0. (P_{ij}) ist folglich vom Rang 1; m.a.W.: Es existiert ein Vektor $Z=(z^1,z^2,z^3)$, so daß $(P_{ij})=\pm(z_iz_j)$ ist. Wegen $P^i{}_i=0$ ist $\langle Z,Z\rangle=0$. Wir wollen zeigen, daß Z ein Vielfaches eines Gradientenvektors ist. Dazu genügt es zu beweisen, daß die zu Z orthogonale Distribution $\{Z,G\}$ integrabel ist. Aus $\langle Z,Z\rangle=0$ folgt aber durch Differenzieren, daß $D_GZ\in\{Z,G\}$ ist; andererseits hat man aus (12′), daß $D_ZG=d_1Z$ ist. Wir wählen nun wieder „isotrope" Gaußsche Koordinaten (u^1,u^2,u^3), setzen aber $u^1=z$, wobei $\operatorname{grad}z\sim Z$ ist, $u^2=\theta$ mit $g^{ij}z_i\frac{\partial\theta}{\partial u^j}=1$ und $u^3=\mu$. Wir erhalten dieselbe Metrik wie unter (A), zusätzlich noch $P_{ij}=P_{11}\delta_{i1}\delta_{j1}$ und $\mu=u^3$. Wegen $\gamma=0$ erhält man aus (16) $P_{ij'k}-P_{ik'j}=\frac{\lambda}{1-\lambda}(P_{ij}\mu_k-P_{ik}\mu_j)$ und daraus $0=P_{31'3}=-P_{11}\Gamma_{33'2}=\frac{1}{2}P_{11}(\partial C/\partial u^2)$. Die gleiche Transformation wie unter (A) ergibt das Gewünschte. □

Im folgenden benützen wir für die partiellen Ableitungen die Indexschreibweise, also z.B. $\frac{\partial B}{\partial v}=B_v$.

Durch eine Routine-Rechnung erhält man folgenden

Hilfssatz II.12. *Für die Metrik in Satz* II.11 *gilt:*

(a) $\Gamma_{11'1}=\frac{1}{2}A_u$; $-\Gamma_{12'1}=\Gamma_{11'2}=-\frac{1}{2}A_v$; $\Gamma_{11'3}=B_u-\frac{1}{2}A_w$;
$\Gamma_{12'3}=-\Gamma_{13'2}=\Gamma_{23'1}=\frac{1}{2}B_v$; $\Gamma_{13'1}=\frac{1}{2}A_w$; $\Gamma_{33'1}=B_w$.

(b) $R_{1212}=:D=-\frac{1}{2}A_{vv}+\frac{1}{4}B_v^2$;
$R_{1313}=:E=B_{uw}-\frac{1}{2}A_{ww}-\frac{1}{2}A_wB_v+\frac{1}{2}A_vB_w+\frac{1}{4}(B^2-A)B_v^2$;
$R_{2323}=0$;

$R_{1213} =: F = \frac{1}{2} B_{uv} - \frac{1}{2} A_{vw} + \frac{1}{4} B B_v^2$;
$R_{2123} = -\frac{1}{2} B_{vv}$;
$R_{3132} =: H = \frac{1}{2} B_{vw} - \frac{1}{4} B_v^2$.

(c) *Es ist* $Q_{23} = -R_{2123}$ *und* $\rho_{23} = \mu_{23} = -\Gamma_{23}^s \mu_s = 0$. *Aus* $Q_{23}(1-\lambda) = \rho_{23}$ *folgt somit* $R_{2123} = 0$.

(d) *Unter Berücksichtigung von* (c) *gilt:*
$Q_{11} = (B^2 - \frac{1}{2} A) D - 2BF + E - AH$; $Q_{12} = -\frac{1}{2} D$;
$Q_{13} = \frac{1}{2} BD - F$; $Q_{22} = Q_{23} = 0$; $Q_{33} = \frac{1}{2} D + H$;
$\beta = -\frac{1}{6} D + \frac{1}{3} H$; $\gamma = \frac{1}{3}(D+H)^2$.

Mit dem Ansatz von Satz II.11 gehen wir nun in die Gleichungen $Q_{ij}(1-\lambda) = \rho_{ij}$. Aus (c) und (b) des obigen Satzes folgt zunächst $B = k(u,w) v + \hat{k}(u,w)$. Wegen der Forderung $\gamma = 0$ haben wir $D + H = 0$, d.h. $A_{vv} = k_w$, was gleichbedeutend ist zu $A = \frac{1}{2} k_w v^2 + g(u,w) v + h(u,w)$. Durch die Transformation $\bar{u} = u$, $\bar{v} = v + \psi(u,w)$, $\bar{w} = w$, $\psi_w = -\psi k + \hat{k}$, erhält die Metrik die Gestalt

$$\begin{aligned} ds^2 &= (\tfrac{1}{2} k_w v^2 + g v + h) du^2 + 2 du dv + 2 k v du dw + dw^2, \\ k_v &= g_v = h_v = 0. \end{aligned} \tag{32}$$

Somit ergibt sich

$$\beta = \tfrac{1}{4}(k_w - \tfrac{1}{2} k^2) = \text{const}, \tag{33}$$

und daraus

$$k_{uw} = k k_u; \qquad k_{ww} = k k_w. \tag{34}$$

Für die μ_{ij} erhält man nun

$$\mu_{11} = -\tfrac{1}{4} k k_w v^2 + \tfrac{1}{2}(k_w \mu_1 - g k \mu_3 - 2 k_u \mu_3 + g_w \mu_3) v + \mu_{uu} + \tfrac{1}{2} g \mu_1 + \tfrac{1}{2} h_w \mu_3;$$

$$\mu_{12} = -\tfrac{1}{2} k \mu_3; \qquad \mu_{13} = -\tfrac{1}{2} k^2 \mu_3 v + \mu_{uw} + \tfrac{1}{2} k \mu_1; \qquad \mu_{22} = \mu_{23} = 0; \qquad \mu_{33} = \mu_{ww}.$$

Die linken und rechten Seiten der Bedingungen

(i) $Q_{11}(1-\lambda) = \rho_{11}$;
(ii) $Q_{12}(1-\lambda) = \rho_{12}$;
(iii) $Q_{13}(1-\lambda) = \rho_{13}$;
(iv) $Q_{33}(1-\lambda) = \rho_{33}$,

sind Polynome in der Variablen v. Durch Koeffizientenvergleich findet man unter Verwendung von (33) & (34)

(ii) $\frac{1}{2} k_w - \frac{1}{4} k^2 = \dfrac{1}{1-\lambda}((\mu_3)^2 - k \mu_3) = 2\beta$;

(iii) $-\frac{1}{2} k_u + \frac{1}{2} g_w = \dfrac{1}{1-\lambda}(\mu_{uw} + \frac{1}{2} k \mu_1 - \mu_1 \mu_3)$;

(iv) $\frac{1}{2} k_w - \frac{1}{4} k^2 = \dfrac{2}{1-\lambda}(\mu_{ww} - \frac{1}{2}(\mu_3)^2)$;

(i) $-\frac{1}{2} g_{ww} + \frac{1}{2} k g_w = \dfrac{1}{1-\lambda}(\frac{1}{2} g_w \mu_3 + \frac{1}{2} k_w \mu_1 - k_u \mu_3)$;

$-\frac{1}{2} h_{ww} - \frac{1}{2} h_w k - \frac{1}{4} h(k_w + \frac{1}{2} k^2)$
$= \dfrac{1}{1-\lambda}(\frac{1}{2} h_w \mu_3 + \frac{1}{2} h(\mu_3)^2 + \frac{1}{2} g \mu_1 + \mu_{uu} - (\mu_1)^2)$.

Wir unterscheiden jetzt die folgenden Fälle:

(A) Sei $\alpha=0$, d.h. $\mu_3=0$. Wir setzen gemäß Satz II.11 $\mu_1=1$.

(A1) $k=0$. Das System (i)–(iv) reduziert sich auf

$$k=\mu_3=0; \quad g_w=0; \quad h_{ww}=\frac{1}{1-\lambda}(2-g).$$

Aus Hilfssatz II.12 folgt $Q_{12}=Q_{13}=Q_{22}=Q_{23}=Q_{33}=0$. Also ist (wegen Satz I.6) $S=0$. Nach Paragraph 5.2 muß die Metrik rekurrent und der Konformitätsvektor isotrop sein. Wie man leicht sieht, wird dies durch die obigen Gleichungen bestätigt.

(A2) $k\neq 0$. Das System (i)–(iv) reduziert sich dann auf

$$\begin{aligned} &k_w-\tfrac{1}{2}k^2=0; \quad g_w=k_u+\frac{1}{1-\lambda}k; \\ &h_{ww}+h_w k+\tfrac{1}{2}hk^2=\frac{1}{1-\lambda}(2-g). \end{aligned} \tag{35}$$

Die erste Gleichung von (i) kann weggelassen werden, da sie durch Differenzieren der Gleichung (iii) folgt. Wegen $Q_{12}=-\frac{1}{2}D=0$ und $Q_{13}=-F\neq 0$ ist die Metrik für jede Lösung keine Einsteinmetrik und somit nicht von konstanter Krümmung. Aus Satz I.6 folgt $S_{1\,3\,1}\neq 0$. Es liegen also Beispiele mit $\alpha=0$ und nicht konform-flacher Metrik vor. (Bemerkung: Die ersten beiden Gleichungen lassen sich leicht integrieren. Die Funktionen k, g, h sind jeweils nur bis auf Funktionen von u bestimmt.)

(B) Sei $\alpha\neq 0$. Dann ist $\mu_3=\sqrt{2\alpha}$, wegen des rechten Teils von (ii) also $k=k(\mu)=\frac{\sqrt{2\alpha}}{\alpha}(\alpha-(1-\lambda)\beta)$. Dieses k erfüllt wegen (19) tatsächlich (ii). Wegen

$$\mu_{uw}+\tfrac{1}{2}k\mu_1-\mu_1\mu_3=(\alpha'\mu_1/\sqrt{2\alpha})+\tfrac{1}{2}\frac{\sqrt{2\alpha}}{\alpha}(\alpha-(1-\lambda)\beta)\mu_1-\sqrt{2\alpha}\mu_1=0$$

folgt aus (iii) $g_w=k_u$, d.h. $g_{ww}=kk_u$. Die erste Gleichung von (i) ist also eine Folgerung aus (ii) und (iii). Ferner folgt (iv) aus (ii) wegen

$$2\mu_{ww}-(\mu_3)^2=2(\alpha'-\alpha)=2(1-\lambda)\beta=(\mu_3)^2-k\mu_3.$$

Das System reduziert sich also auf

$$\begin{aligned} &\alpha'=\alpha+(1-\lambda)\beta \quad \text{d.h. } \alpha=\alpha_0\sqrt{\lambda}-(1+\lambda)\beta; \quad \mu_3=\sqrt{2\alpha}; \\ &k=\frac{\sqrt{2\alpha}}{\alpha}(\alpha-(1-\lambda)\beta); \quad g_w=k'\mu_1; \\ &h_{ww}+h_w k+\tfrac{1}{2}h(4\beta+k^2)=\frac{1}{1-\lambda}(-2\mu_{uu}+2(\mu_1)^2-g\mu_1-h_w\sqrt{2\alpha}-2\alpha h). \end{aligned} \tag{36}$$

Man findet:

$$K\neq\text{const} \Leftrightarrow \mu_{uu}-(\mu_1)^2+\tfrac{1}{2}g\mu_1+\tfrac{1}{2}h_w\sqrt{2\alpha}+\tfrac{1}{2}hk\sqrt{2\alpha}\neq 0. \tag{37}$$

Wir wählen also Lösungen des Systems, die in einem Punkt dieser Bedingung genügt.

Beispiele:

(B1) $\beta=0$. Dann ist $\alpha=\alpha_0 e^{\mu}$, $k=\sqrt{2\alpha}$, $\mu=-2\ln[-\frac{1}{2}\sqrt{2\alpha_0}\,w+\mu_0(u)]$.

(B2) Sei $\beta<0$. Dann ist eine Lösung gegeben durch

$$k^2=-8\beta, \quad \alpha=-\beta(1+e^{\mu})^2, \quad \mu=\ln[e^{\sqrt{-2\beta}w}\mu_0(u)/(1-e^{\sqrt{-2\beta}w}\mu_0(u))]$$

$$g=g(u).$$

Wir fassen die Ergebnisse zusammen im

Theorem II.13. *Sei f: $(M,g)\to(\bar{M},\bar{g})$ ein krümmungstreuer Diffeomorphismus zwischen zwei ps.-Riem. Mgf. der Dimension $n=3$. Sei die Menge der nicht-isotropen Punkte dicht in M. Sei $\gamma=\frac{1}{2}\mathrm{Sp}(P^2)=0$. Dann ist $f^*\bar{g}=\lambda g$, wobei für g und λ einer der folgenden (nicht-trivialen) Fälle auftritt:*

(A1) *Die Metrik g ist konform-flach. Dann ist (M,g) ein K^*_3-Raum und der Konformitätsvektor grad λ ist isotrop (vgl. Theorem* II.3).

(A2) *Die Metrik g ist nicht konform-flach und der Konformitätsvektor ist isotrop. Dann existieren Koordinaten (u,v,w), so daß die Metrik – bis auf eine nirgends dichte Menge – von der Form* (32) *und $\mu=u$ ist. Die auftretenden Funktionen sind Lösungen von* (35).

(B) *Die Metrik g ist nicht konform-flach und der Konformitätsvektor nicht isotrop. Dann existieren Koordinaten (u,v,w), so daß die Metrik – bis auf eine nirgends dichte Menge – von der Form* (32) *und $\mu=\mu(u,w)$ ist. Die auftretenden Funktionen sind Lösungen von* (36), *so daß in einem Punkt* (37) *gilt. Dabei kann für β eine beliebige Konstante gewählt werden.*

Literatur

1. Eisenhart, L.P.: Riemannian Geometry. Princeton: University Press 1949
2. Kreyszig, E.: Introduction to differential geometry and Riemannian geometry. University of Toronto Press, 1968
3. Kulkarni, R.S.: Une réciproque du ‹theorema egregium›. C.R. Acad. Sci. **264**, série A, 944–946 (1967); C.R. Acad. Sci. **265**, série A, 388–390 (1967)
4. Kulkarni, R.S.: Curvature and metric. Ann. of Math. **91**, 311–331 (1970)
5. Kulkarni, R.S.: On the Bianchi-Identities. Math. Ann. **199**, 175–204 (1972)
6. Kulkarni, R.S.: Maps of Manifolds with indefinite metric preserving certain geometrical entities. Int. J. Math. Math. Sci. **1**, 137–140 (1978)
7. Nasu, T.: Curvature and metric in Riemannian 3-manifolds. J. Math. Soc. Japan, **27**, 194–206 (1975)
8. Strubecker, K.: Differentialgeometrie III. Berlin: de Gruyter 1969
9. Walker, A.G.: On Ruse's spaces of recurrent curvature. Proc. Lond. Math. Soc., II. Sér. **52**, 36–64 (1951)
10. Weyl, H.: Zur Infinitesimalgeometrie. Göttinger Nachr., 1921, 99–112
11. Yau, S.T.: Curvature preserving diffeomorphisms. Ann. of Math. **100**, 121–130 (1974)

Eingegangen am 15. April 1983; in endgültiger Fassung am 31. August 1984

Math. Z. 189, 393-413 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Cohomology of Bimodules Over Enveloping Algebras

Kenneth A. Brown[1], and Thierry Levasseur[2]

[1] Department of Mathematics, University of Glasgow, Glasgow G12 8QW, U.K.
[2] Dept. de Mathématiques, Université de Paris VI, Tour 45-46-5e, 4, place Jussieu, Paris, France

Introduction

Let $\mathbf{g}$ be a finite dimensional Lie algebra over $\mathbb{C}$ with enveloping algebra U. Let M be a $(U-U)$-bimodule, finitely generated on both sides. For $i \geqq 0$, $\mathrm{Ext}^i_U(M, U)$, evaluated for the *left* modules M and U, is a $(U-U)$-bimodule via the right action of U on M and on U. These bimodules are the principal focus of interest in this paper. The central result, (5.5), concerns the case when $M = U/P$, where P is a prime ideal of height t of U, and $\mathbf{g}$ is solvable. *Under these circumstances,* $\mathrm{Ext}^i_U(U/P, U) = 0$ *for* $i < t$; *and* $\mathrm{Ext}^t_U(U/P, U)$ *is isomorphic, as a left module, to an non-zero left ideal of* U/P, *and, as a right module, to a right ideal of* $U/\tau_\lambda(P)$, *where* τ_λ *is the winding automorphism of* U *obtained using a certain element* λ *of* $(\mathbf{g}/[\mathbf{g}, \mathbf{g}])^*$ *determined by* P. (*For example, if* P *is co-artinian, then* $\lambda = -\mathrm{trace}\ \mathrm{ad}\ \mathbf{g}$.)

The vanishing of $\mathrm{Ext}^i_U(U/P, U)$ for $i < t$ and the fact that these groups are "small" – in the sense of G.K.-dimension – for $i > t$ are easy consequences of earlier work, principally of Björk and Roos. A question of Roos', motivated by well-known open problems concerning the finitistic global dimension of Artin algebras, led to our work described here. Furthermore, recent important work of O. Gabber [9], making use of the modules $\mathrm{Ext}^i_U(\mathrm{Ext}^i_U(X, U), U)$, suggests that it may be worthwhile to gain a better understanding of the *right* modules $\mathrm{Ext}^i_U(X, U)$, where X is a finitely generated left module.

If $\mathbf{g}$ is abelian, so that U is a polynomial algebra (and so Cohen-Macaulay), the P contains a U-sequence of length t generating the ideal I, say, and $\mathrm{Ext}^t_U(U/P, U)$ is just $\mathrm{Hom}_U(U/P, U/I)$. That $\mathrm{Ext}^t_U(U/P, U)$ is a (i) torsion-free (ii) uniform U/P-module then correspond to the facts that U (and so U/I) are (i) Cohen-Macaulay and (ii) Gorenstein. So 5.5 can be viewed as a generalisation of these aspects of the commutative theory.

We now discuss the contents of the paper in detail. In § 1 we state some isomorphisms arising from spectral sequences in bimodule categories over K-algebras, (K a field). This material is implicit in work of Grothendieck [10]; but we have chosen to give a self-contained account, with the aim of stimulating further applications in other contexts.

The main result of §1 is used in §2 to describe the bimodules $\mathrm{Ext}^i_R(R/I, R)$, where R is a ring and I is an ideal generated by a regular normalising sequence (as defined in (2.1)). Of course, even when $\mathbf{g}$ is abelian, not all prime ideals P of U are generated by regular normalising sequences. However, when $\mathbf{g}$ is solvable there is a subset $\mathscr{S}(P)$ of $U \setminus P$ which is an Ore set, such that the localised ring $U_{\mathscr{S}(P)} = R$ has many desirable properties [6]. For our purposes, the most important are that R/PR is the division ring of fractions of U/P, PR has a regular normalising set of t generators, and R has global dimension t, where t is the height of P. Now, as bimodules, $\mathrm{Ext}^i_U(U/P, U) \otimes_U R \cong \mathrm{Ext}^i_R(R/PR, R)$, so the main result of §2 yields information about $\mathrm{Ext}^i_U(U/P, U)$.

To obtain the result stated in the first paragraph, then, two things must be shown: The bimodule $\mathrm{Ext}^i_U(U/P, U)$ has to be left and right $\mathscr{S}(P)$-torsion free; and it has to be finitely generated on each side. The first of these facts is proved in §5, using a result on module extensions which is obtained in §3. The second is a special case of an unpublished result of I.N. Bernstein (4.4): *If M is a Harish-Chandra module for an enveloping algebra U, then $\mathrm{Ext}^i_U(M, U)$ is also a Harish-Chandra module; in particular it is finitely generated on the right and the left.*

In §6 three applications of 5.5 are given. We give a new proof of a result of M.-P. Malliavin [14] on the minimal injective resolution of $U(\mathbf{g})$, where $\mathbf{g}$ is solvable (6.1); we discuss the projective dimensions, as U-modules, of certain prime factor rings of U (again for solvable $\mathbf{g}$) (6.2); and for $\mathbf{g}$ semisimple and M a Verma module for $U(\mathbf{g})$, we give a description, (implicit in the work of Duflo [8]), of the right modules $\mathrm{Ext}^*_{U(\mathbf{g})}(M, U(\mathbf{g}))$.

In §7, (5.5) is generalised in two directions: For $\mathbf{g}$ solvable, the structure of the modules $\mathrm{Ext}^i_U(M, U)$ is examined for M a torsion free module over a prime factor ring of U (7.1); and for $M = U/I$, I an ideal of U (7.2). For a co-Artinian ideal I, (7.2) is refined in (7.3).

Acknowledgements. Much of this work was done while the first author was visiting Université Paris VI. We would like to thank Marie-Paule Malliavin for arranging this visit. We are also indebted to her, and to Colette Moeglin, for many helpful observations; and (most of all) to I.N. Bernstein for explaining (4.4) to us and for allowing us to include it here. We also thank the London Mathematical Society, A.W. Goldie, and J.C. McConnell for supporting and organising the conference on enveloping algebras at Durham, 1983, when some of this research was done.

Index of Notation: All modules in this paper are left modules, except where otherwise described. The category of left [resp. right] R-modules, for a ring R, is denoted R-Mod [resp. Mod-R]. The first occurrences of some of the most frequently used notations are as follows.

$l(\ )$	3.1	$r(\ )$	3.1
$d(\ )$	5.1	${}^{\sigma}M$	2.1
τ_λ	5.2	$\mathscr{C}(P)$	3.1
$\mathscr{S}(P)$	3.1	$\mathscr{E}(P)$	5.2
$O(P)$	5.2	$G(\ )$	5.1
$\mathbf{k}$	4.1	U^2	4.1
$\mathscr{H}$	4.1	$\mathscr{H}_f$	4.1
Φ_V	4.2		

1. Spectral Sequences of Bimodules

In this section K will be a fixed commutative field.

1.1. Let Γ, Σ be two K-algebras. A $(\Gamma-\Sigma)$-bimodule, ${}_\Gamma C_\Sigma$, is a left Γ-module and a right Σ-module such that $\gamma(c\sigma)=(\gamma c)\sigma$, $\alpha c = c\alpha$, for every $\gamma\in\Gamma$, $\sigma\in\Sigma$, $\alpha\in K$, $c\in C$. Clearly the category of $(\Gamma-\Sigma)$-bimodules can be identified with mod-$(\Gamma^0 \otimes_K \Sigma)$. The following result is well-known [7, chap. XVI, §4]:

Proposition. *Let C be injective [resp. projective] in* Mod-$(\Gamma^0\otimes_K\Sigma)$. *Then C is injective [resp. projective] in* Mod-Σ *and* Γ-mod.

Consequently any projective [resp. injective] resolution of $C\in$Mod-$(\Gamma^0 \otimes_K \Sigma)$ will be a projective [resp. injective] resolution in Mod-Σ and Γ-mod.

1.2. Let $\Gamma, \Sigma, \Lambda, \Delta$ be K-algebras. We set $\mathscr{C}_1 = \text{Mod-}(\Gamma^0\otimes_K\Delta)$, $\mathscr{C}_2 = \text{Mod-}(\Gamma^0\otimes_K\Sigma)$ and $\mathscr{C}_3 = \text{Mod-}(\Delta^0\otimes_K\Sigma)$. Define a bifunctor T from $\mathscr{C}_1\times\mathscr{C}_2$ to $\mathscr{C}_3$ by $T(A,B)=\text{Hom}_\Gamma(A,B)$. Let $P(A)$ be a projective resolution of A in $\mathscr{C}_1$ and $I(B)$ an injective resolution of B in $\mathscr{C}_2$. For each $i\geqq 0$ set

$$R_1^i T(A,B)=H^i(T(P(A),B))=H^i(\text{Hom}_\Gamma(P(A),B))$$

and

$$R_2^i T(A,B)=H^i(T(A,I(B))=H^i(\text{Hom}_\Gamma(A,I(B))).$$

All these objects are in $\mathscr{C}_3$. Denote the derived functors of T by $R^iT(A,B)$, these also being objects in $\mathscr{C}_3$.

Note that if B is injective in $\mathscr{C}_2$ then the functor $T(-,B)$ from $\mathscr{C}_1$ to $\mathscr{C}_3$ is exact, since B is injective in Γ-mod by 1.1. In the same way if A is projective in $\mathscr{C}_1$ then $T(A,-)$ is an exact functor from $\mathscr{C}_2$ to $\mathscr{C}_3$ since A is projective in Γ-mod. Thus by [10, §2.3] we can make the natural identifications in $\mathscr{C}_3$:

$$R_1^i T(A,B)=R_2^i T(A,B)=R^i T(A,B),$$

for all $i\geqq 0$. In particular the $(\Delta-\Sigma)$-bimodule $R^iT(A,B)$ is equal to $\text{Ext}_\Gamma^i(A,B)$, as one can see by writing $R_1^iT(A,B)=H^i(\text{Hom}_\Gamma(P(A),B))=\text{Ext}_\Gamma^i(A,B)$, since $P(A)$ is a projective resolution of A in Γ-mod, by 1.1.

Similarly, consider the bifunctor $S:\mathscr{C}_1\times\mathscr{C}_3\to\mathscr{C}_2$ given by $S(A,B)=A\otimes_\Delta B$. If $P(A)$, [resp. $P(B)$] is a projective resolution of A [resp. B] in $\mathscr{C}_1$ [resp. $\mathscr{C}_3$], then we have the natural identifications in $\mathscr{C}_2$:

$$H_i(P(A)\otimes_\Delta B)=H_i(A\otimes_\Delta P(B))=\text{Tor}_i^\Delta(A,B).$$

1.3. Let $B\in\text{Mod}(\Gamma^0\otimes_K\Delta)$, $A\in\text{Mod}(\Lambda^0\otimes_K\Gamma)$, $C\in\text{Mod}(\Lambda^0\otimes_K\Sigma)$, with Λ, Γ, Δ, and Σ as in 1.2. There is a natural isomorphism [7, Chap. II, §5]

$$s:\text{Hom}_\Gamma(B,\text{Hom}_\Lambda(A,C))\to\text{Hom}_\Lambda(A\otimes_\Gamma B,C)$$

where $s(f)(a\otimes b)=f(b)(a)$, which is an isomorphism of $(\Delta-\Sigma)$-bimodules, as one can easily verify.

Fix a projective resolution of B in Mod$(\Gamma^0\otimes_K\Delta)$,

$$P(B):\ldots\to P_1\xrightarrow{d_1}P_0\to B\to 0;$$

and an injective resolution of C in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$,

$$E(C): 0 \to C \to E^0 \xrightarrow[d_{II}]{} E^1 \to E^2 \to \ldots .$$

Denote by $\mathscr{K}$ the double complex with modules $K^{p,q} = \mathrm{Hom}_\Gamma(P_p, \mathrm{Hom}_\Lambda(A, E^q))$ and differentials $d_I^{p,q}: K^{p,q} \to K^{p+1,q}$ and $d_{II}^{p,q}: K^{p,q} \to K^{p,q+1}$ induced by d_I and d_{II}. Note that $\mathscr{K}$ is a double complex in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$. Furthermore, as observed above, $K^{p,q} \cong \mathrm{Hom}_\Lambda(A \otimes_\Gamma P_p, E^q)$ in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$. Denote by $H^*(\mathscr{K})$ the total cohomology of $\mathscr{K}$; this is also an object in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$. With these notations we can restate the assertion of [7, p. 345] as follows. (We shall supply a proof for completeness.)

Proposition. *There are two convergent spectral sequences in* $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$, *given by*

$$'E_2^{p,q} = \mathrm{Ext}_\Gamma^p(B, \mathrm{Ext}_\Lambda^q(A, C)) \Rightarrow H^*(\mathscr{K})$$

and

$$''E_2^{p,q} = \mathrm{Ext}_\Lambda^q(\mathrm{Tor}_p^\Gamma(A, B), C) \Rightarrow H^*(\mathscr{K}).$$

Proof. The two filtrations F_I and F_{II} of the single complex associated to $\mathscr{K}$ yield two spectral sequences, namely $'E_2^{p,q} = H_I^p H_{II}^q(\mathscr{K})$ and $''E_2^{p,q} = H_{II}^q H_I^p(\mathscr{K})$. Since $K^{p,q} = 0$ if $p < 0$ or $q < 0$, both sequences converge to $H^*(\mathscr{K})$. The cohomology of the complex $(K^{p,*}, d_{II})$ is given by

$$H_{II}^q(K^{p,*}) = H_{II}^q(\mathrm{Hom}_\Gamma(P_p, \mathrm{Hom}_\Lambda(A, E^*(C)))).$$

Since P_p is Γ-projective, and since $H^q(\mathrm{Hom}_\Lambda(A, E^*(C))) = \mathrm{Ext}_\Lambda^q(A, C)$ in $\mathrm{Mod}(\Gamma^0 \otimes_K \Sigma)$ by 1.2, the following equalities hold in $\mathrm{Mod}(\Gamma^0 \otimes_K \Sigma)$:

$$H_{II}^q(K^{p,*}) = \mathrm{Hom}_\Gamma(P_p, H^q(\mathrm{Hom}_\Lambda(A, E^*(C)))) = \mathrm{Hom}_\Gamma(P_p, \mathrm{Ext}_\Lambda^q(A, C)).$$

From 1.2 it also follows that, in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$,

$$'E_2^{p,q} = H_I^p(H_{II}^q(\mathscr{K}), d_I) = H_I^p(\mathrm{Hom}_\Gamma(P(B), \mathrm{Ext}_\Lambda^q(A, C))) = \mathrm{Ext}_\Gamma^p(B, \mathrm{Ext}_\Lambda^q(A, C)).$$

Finally, using the last part of 1.2 and the isomorphism $K^{p,q} \cong \mathrm{Hom}_\Lambda(A \otimes_\Gamma P_p, E^q)$, one sees that, in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$,

$$''E_2^{p,q} = \mathrm{Ext}_\Lambda^q(\mathrm{Tor}_p^\Gamma(A, B), C).$$

Remark 1.4. All the morphisms which can be constructed from the morphisms d_I and d_{II} between the terms of the spectral sequence are morphisms in $\mathrm{Mod}(\Lambda^0 \otimes_K \Sigma)$ – that is, they are maps of $(\Lambda - \Sigma)$-bimodules.

1.5. Let R, S, T be three K-algebras and $z \in R$ a normal element with $\bar{R} = R/zR$ (see 2.1). Let ${}_R M_T$ and ${}_R X_S$ be bimodules. Suppose that z is a non-zero divisor in R and on the left R-module M.

Proposition. *For all* $n \geqq 1$ *there is an isomorphism of* $(S - T)$*-bimodules,*

$$\mathrm{Ext}_R^{n-1}(X, \mathrm{Ext}_R^1(\bar{R}, M)) \cong \mathrm{Ext}_R^n(X, M).$$

(*Here,* $\mathrm{Ext}_R^1(\bar{R}, M)$ *is computed using the left* R*-modules* R *and* M.)

Proof. Apply Proposition 1.3 with $\Gamma = \bar{R}$, $\Lambda = R$, $\Delta = S$, $\Sigma = T$ and ${}_\Gamma B_\Delta = {}_{\bar{R}}X_S$. We obtain two spectral sequences in $\mathrm{Mod}(S^0 \otimes_K R)$ convergent to the same $H^*(\mathscr{X})$; namely

$$'E_2^{p,q} = \mathrm{Ext}_R^p(X, \mathrm{Ext}_{\bar{R}}^q(\bar{R}, M)),$$

and

$$''E_2^{p,q} = \mathrm{Ext}_{\bar{R}}^q(\mathrm{Tor}_p^R(\bar{R}, X), M).$$

Since $''E_2^{p,q} = 0$ for $p \neq 0$, the second sequence collapses to isomorphisms: – For $n \geqq 0$,

$$H^n(\mathscr{X}) \cong ''E_2^{0,n} = \mathrm{Ext}_{\bar{R}}^n(X, M).$$

The projective dimension of $\bar{R}$ as a left R-module is one, and moreover $\mathrm{Ext}_R^0(\bar{R}, M) = \mathrm{Hom}_R(\bar{R}, M) = 0$, by the hypotheses on z. Hence the first spectral sequence collapses to isomorphisms: – For $n \geqq 1$

$$'E_2^{n-1,1} = \mathrm{Ext}_R^{n-1}(X, \mathrm{Ext}_R^1(\bar{R}, M)) \cong \mathrm{Ext}_R^n(X, M) \cong H^n(\mathscr{X}).$$

By Remark 1.4, these maps are all $(S-T)$-bimodule isomorphisms.

1.6. We shall also need the following well-known result:

Proposition. *Let R and S be rings and $R \to S$ a ring homomorphism such that S is flat as a left and right R-module. Let ${}_R X_R$ be an R-bimodule such that the $(R-S)$-bimodule $X \otimes_R S$ is an S-bimodule. Then for every finitely generated left R-module M and all $i \geqq 0$, there are isomorphisms of right S-modules,*

$$\mathrm{Ext}_R^i(M, X) \otimes_R S \cong \mathrm{Ext}_S^i(S \otimes_R M, X \otimes_R S).$$

Moreover, if M is an R-bimodule these isomorphisms are isomorphisms of (R, S)-bimodules.

Proof. Let ${}_R M$ be a finitely generated free R-module and consider the map $\theta: \mathrm{Hom}_R(M, X) \otimes_R S \to \mathrm{Hom}_S(S \otimes_R M, X \otimes_R S)$ defined by $\theta(f \otimes s)(t \otimes m) = t[f(m) \otimes s]$ for $f \in \mathrm{Hom}_R(M, X)$, $m \in M$, $t, s \in S$. It's easy to show that θ is a well-defined isomorphism of right S-modules. Using the facts that R is left Noetherian and S is right and left R-flat one can then obtain the isomorphisms of the first statement. The second statement is an immediate consequence of the structure as $(R-S)$-bimodules of $\mathrm{Ext}_R^i(M, X) \otimes_R S$ and $\mathrm{Ext}_S^i(S \otimes_R M, X \otimes_R S)$.

2. Ext Groups and Regular Normalising Sequences

2.1. Let R be a ring. Recall that an element $z \in R$ is *normal* if $Rz = zR$, and *regular* if it is not a zero divisor on the right or the left.

Definition. The elements $z_1, \ldots, z_t$ of R form a *regular normalising sequence* if, (setting $z_0 = 0$), for all i, $0 \leqq i \leqq t-1$, z_{i+1} maps to a regular normal element in the ring $R / \sum_{j<i+1} z_j R$.

Definition. Let σ be an automorphism of R and let T be a ring. Let M be an $(R-T)$-bimodule. The $(R-T)$-bimodule ${}^{\sigma}M$ is defined by: ${}^{\sigma}M=M$ as a right T-module and, as a left R-module, $r\in R$ acts on $m\in{}^{\sigma}M$ by $r.\ m=\sigma(r)m$.

Note that, if z is a regular normal element of R, the rule $r^{\sigma}z=zr$ for $r\in R$ defines an automorphism σ of R.

Lemma. *Let z be a regular normalising element of R and σ the corresponding automorphism as defined above. Let M be an $(R-T)$-bimodule.*

(i) *If z is not a zero divisor on M then* $\mathrm{Hom}_R(R/Rz, M)=0$.

(ii) *There is an isomorphism of $(R-T)$-bimodules, (in fact of $(R/Rz-T)$-bimodules),*

$$\mathrm{Ext}^1_R(R/Rz, M)\cong {}^{\sigma}M/z\,{}^{\sigma}M.$$

Proof. (i) Clear.

(ii) Let $a\in R$. The right multiplications by a on Rz and R, both denoted by r_a, yield a commutative diagram of left R-modules,

$$\begin{array}{ccccccc} 0\to & Rz & \xrightarrow{d} & R & \longrightarrow & R/Rz\to 0 \\ & \downarrow{\scriptstyle r_a} & & \downarrow{\scriptstyle r_a} & & \downarrow & \\ 0\to & Rz & \xrightarrow{d} & R & \longrightarrow & R/Rz\to 0. \end{array}$$

Applying $\mathrm{Hom}_R(-, M)$ yields the commutative diagram of right T-modules

$$\begin{array}{ccccc} 0\to & \mathrm{Hom}_R(R, M) & \xrightarrow{d^*} & \mathrm{Hom}_R(Rz, M) \\ & \downarrow{\scriptstyle r_a^*} & & \downarrow{\scriptstyle r_a^*} \\ 0\to & \mathrm{Hom}_R(R, M) & \xrightarrow{d^*} & \mathrm{Hom}_R(Rz, M). \end{array}$$

Thus the left R-module structure of $\mathrm{Ext}^1_R(R/Rz, M)$ is given by the left action of R induced by right multiplication on Rz. In other words, there is an isomorphism of $(R-T)$-bimodules, $\mathrm{Ext}^1_R(R/Rz, M)\cong \mathrm{Hom}_R(Rz, M)/\mathrm{Im}\, d^*$. It is easily seen that $\mathrm{Im}\, d^*=\{f\in\mathrm{Hom}_R(Rz, M): f(z)\in zM\}$, so the above isomorphism is actually of $(R/Rz-T)$-bimodules. Define now a map $\varphi\colon \mathrm{Hom}_R(Rz, M)\to {}^{\sigma}M$ by $\varphi(f)=f(z)$. Clearly, φ is right T-linear and the following calculation shows that φ is left R-linear: For all $r\in R$ and $f\in\mathrm{Hom}_R(Rz, M)$,

$$\varphi(rf)=(rf)(z)=f(zr)=f(r^{\sigma}z)=r^{\sigma}f(z)=r\cdot\varphi(f).$$

It is now clear that the induced map $\overline{\varphi}\colon \mathrm{Hom}_R(Rz, M)/\mathrm{Im}\, d^*\to {}^{\sigma}M/z\,{}^{\sigma}M$ is an $(R/Rz-T)$-bimodule isomorphism.

2.2. All the rings appearing in this paragraph are algebras over a (commutative) field K. Let I be an ideal of a ring R with I generated by a regular normalising sequence $\{z_1, \ldots, z_t\}$. Thus, for $0\leqq i\leqq t-1$, z_{i+1} defines an automorphism $\overline{\sigma_{i+1}}$ of the ring $R_i:=R\Big/\sum_{j=0}^{i} z_jR$, (setting $z_0=0$). We shall assume

that each $\bar{\sigma}_i$ is induced by an automorphism σ_i of R, and we shall write $r^{\sigma_i} z_i \equiv z_i r (\mathrm{mod} \sum_{j<i} z_j R)$ for $r \in R$, (setting $\sigma_0 = 1$).

Let T be another K-algebra and M an $(R-T)$-bimodule. For each $i=0, \ldots, n$ define $M_i := {}^{\sigma_i}M_{i-1}/z_i {}^{\sigma_i}M_{i-1}$, $M_0 := M$, $(R_i - T)$-bimodules. We shall need an additional hypothesis:

(*) For $0 \leqq i \leqq n-1$, z_{i+1} is not a zero divisor on the left R_i-module M_i.

Lemma. (i) *There exists an isomorphism of $(R_i - T)$-bimodules between M_i and* ${}^{\sigma_1 \ldots \sigma_i}M \Big/ \sum_{l=0}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M$ *given by*

$$\varphi(x + z_i {}^{\sigma_i}M_{i-1}) = m + \sum_{l=1}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M,$$

for every $x \in {}^{\sigma_i}M_{i-1}$ represented by an element m of M.

(ii) *Condition (*) is equivalent to: (**) For $0 \leqq i \leqq n-1$, z_{i+1} is not a zero divisor on the left R_i-module* ${}^{\sigma_1 \ldots \sigma_i}M \Big/ \sum_{l=0}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M$.

Proof. (i) Note first that if $[x_{i-1} + z_i {}^{\sigma_i}M_{i-1}] \in M_i$ with $x_{i-1} \in {}^{\sigma_i}M_{i-1}$, then we can represent x_{i-1} by an element x_{i-2} of ${}^{\sigma_{i-1}}M_{i-2}$, and so on, until we reach x_1 represented by x_0 in ${}^{\sigma_1}M_0 = M$. The map φ is simply $\varphi(x_{i-1} + z_i {}^{\sigma_i}M_{i-1}) = x_0 + \sum_{l=1}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M$. For each i, consider the map $\varepsilon_{i-1} : {}^{\sigma_i}M_{i-1} \to M_{i-1}$ defined by forgetting the twisting on the left structure; that's to say, $\varepsilon_{i-1}(m) = m$ for $m \in {}^{\sigma_i}M_{i-1}$. It is clearly an isomorphism of right T-modules which identifies the right T-modules M_i and $M_{i-1}/z_i M_{i-1}$. Thus, by an easy induction, the right T-modules M_i and $\bar{M} = M \Big/ \sum_{l=0}^{i} z_l M$ are isomorphic, via the map ε, where $\varepsilon(x_{i-1} + z_i {}^{\sigma_i}M_{i-1}) = \left[x_0 + \sum_{l=1}^{i} z_l M\right]$. In the same way ${}^{\sigma_1 \ldots \sigma_i}M \Big/ \sum_{l=1}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M$ and $\bar{M}$ are isomorphic as T-modules by the map $\rho: \left[m + \sum_{l=1}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M\right] \to \left[m + \sum_{l=1}^{i} z_l M\right]$. Clearly, there is a commutative diagram:

$$\begin{array}{ccc} M_i & \xrightarrow{\varepsilon} & \bar{M} \\ & {\scriptstyle\varphi}\searrow \quad \swarrow{\scriptstyle\rho} & \\ & {}^{\sigma_1 \ldots \sigma_i}M \Big/ \sum_{l=1}^{i} z_l {}^{\sigma_1 \ldots \sigma_i}M. & \end{array}$$

This shows that ρ is an isomorphism of right T-modules.

Consider now the left action by R_i. Let $r \in R$ and $[x_{i-1} + z_i {}^{\sigma_i}M_{i-1}] \in M_i$ with $x_{i-1} \in {}^{\sigma_i}M_{i-1}$. As above, we define the sequence $x_{i-2}, \ldots, x_0$. Then

$$\varphi(r[x_{i-1} + z_i {}^{\sigma_i}M_{i-1}]) = \varphi([\sigma_i(r) x_{i-1} + z_i {}^{\sigma_i}M_{i-1}]),$$

where $\sigma_i(r)x_{i-1} \in M_{i-1}$ is represented by $\sigma_i(r)x_{i-2}$ in ${}^{\sigma_{i-1}}M_{i-2}$, which is equal to $\sigma_{i-1}\sigma_i(r)x_{i-2}$ in M_{i-2}. Using this remark at each step one gets

$$\begin{aligned}
\varphi(r[x_{i-1}+z_i{}^{\sigma_i}M_{i-1}]) &= \varphi([\sigma_{i-1}\sigma_i(r)x_{i-2}+z_i{}^{\sigma_i}M_{i-1}]) \\
&= \dots \\
&= \varphi\left(\left[\sigma_1\dots\sigma_i(r)x_0+\sum_{l=0}^{i} z_l{}^{\sigma_i}M_{i-1}\right]\right) \\
&= \left[\sigma_1\dots\sigma_i(r)x_0+\sum_{l=0}^{i} z_l{}^{\sigma_1\dots\sigma_i}M\right] \\
&= r\cdot\left[x_0+\sum_{l=0}^{i} z_l{}^{\sigma_1\dots\sigma_i}M\right] \\
&= r\cdot\varphi([x_{i-1}+z_i{}^{\sigma_i}M_{i-1}]).
\end{aligned}$$

That is, φ is left R_i-linear.

(ii) This is an immediate consequence of (i).

Corollary 2.3. *Let R and I be as in 2.2. Then the $(R-R)$-bimodule R satisfies (*).*

Proof. By 2.2(ii) it's enough to prove that R satisfies (**). Let $r\in R$ and $1\leqq i\leqq n-1$, and suppose that $z_{i+1}\cdot r\in\sum_{l=1}^{i} z_l{}^{\sigma_1\dots\sigma_i}R$. Then there exist $a_1,\dots,a_i\in R$ such that, in the ring R,

$$\sigma_1\dots\sigma_i(z_{i+1})r=\sum_{l=1}^{i}(\sigma_1\dots\sigma_i)(z_l)a_l.$$

Apply $\tau=\sigma_i^{-1}\dots\sigma_1^{-1}$ to this identity to obtain

$$z_{i+1}\tau(r)=\sum_{l=1}^{i} z_l\tau(a_l).$$

But z_{i+1} is not a zero divisor on the left module R_i, so there exist $b_1,\dots,b_i\in R$ such that $\tau(r)=\sum_{l=1}^{i} z_l b_l$.

Therefore,

$$r=\sum_{l=1}^{i}\tau^{-1}(z_l)\tau^{-1}(b_l)=\sum_{l=1}^{i}(\sigma_1\dots\sigma_i)(z_l)\tau^{-1}(b_l)\in\sum_{l=1}^{i} z_l{}^{\sigma_1\dots\sigma_i}R.$$

Theorem 2.4. *Let R, T and I be as in 2.1 and 2.2. Let M be an $(R-T)$-bimodule satisfying (*). Put $\sigma=\sigma_1\sigma_2\dots\sigma_t$. There is an isomorphism of $(R-T)$-bimodules, (indeed, of (R_t-T)-bimodules):*

$$\mathrm{Ext}^i_R(R/I,M)\cong\begin{cases}0 & \text{if } i\neq t;\\ {}^{\sigma}M\Big/\sum_{l=1}^{t} z_l{}^{\sigma}M & \text{if } i=t.\end{cases}$$

Proof. By induction on t. When $t=1$, apply Lemma 2.1. When $t>1$ and $i=0$, $\mathrm{Hom}_R(R/I,M)=0$. Now assume that $t>1$ and $i\geqq 1$, and that the result is

known for smaller values of t. Put $S=R_t$, $T=T$, $\bar{R}=R_1$ and $X=R/I$ in Proposition 1.5. This yields the (R_t-T)-bimodule isomorphism

$$\mathrm{Ext}^{i-1}_{R_1}(R/I, \mathrm{Ext}^1_R(R_1, M)) \cong \mathrm{Ext}^i_R(R/I, M). \tag{1}$$

By Lemma 2.1, $\mathrm{Ext}^1_R(R_1, M) \cong {}^{\sigma_1}M/z_1{}^{\sigma_1}M = M_1$, so (1) becomes, for each $i \geqq 1$,

$$\mathrm{Ext}^{i-1}_{R_1}(R/I, M_1) \cong \mathrm{Ext}^i_R(R/I, M). \tag{2}$$

Condition (*) on M ensures that the (R_1-T)-bimodule M_1 satisfies (*) with respect to the ring R_1 and its ideal $\langle \overline{z_2}, \ldots, \overline{z_t} \rangle$. Therefore, by induction,

$$\mathrm{Ext}^{i-1}_{R_1}(R/I, M_1) \cong \begin{cases} 0 & \text{if } i \neq t; \\ {}^{\sigma_2 \ldots \sigma_t}M_1 \Big/ \sum_{l=2}^{t} z_l{}^{\sigma_2 \ldots \sigma_t}M_1 & \text{if } i = t. \end{cases} \tag{3}$$

As in the proof of Lemma 2.2(i), one easily confirms that the (R_t-T)-bimodule ${}^{\sigma_2 \ldots \sigma_t}M_1 \Big/ \sum_{l=2}^{t} z_l{}^{\sigma_2 \ldots \sigma_t}M_1$ is isomorphic to ${}^{\sigma_1 \ldots \sigma_t}M \Big/ \sum_{l=1}^{t} z_l{}^{\sigma_1 \ldots \sigma_t}M$. Thus the induction step follows from (2) and (3).

Corollary 2.5. *Let R and I be as above. Put $\sigma = \sigma_1 \ldots \sigma_t$. As (R_t-R)-bimodules,*

$$\mathrm{Ext}^i_R(R/I, R) \cong \begin{cases} 0 & \text{if } i \neq t; \\ {}^{\sigma}R \Big/ \sum_{i=1}^{t} z_i{}^{\sigma}R & \text{if } i = t. \end{cases}$$

Thus, as right R-modules, $\mathrm{Ext}^t_R(R/I, R) \cong R/\sigma(I)$.

3. Essential Extensions and Ore Sets

3.1. Let $\mathbf{g}$ be a solvable Lie algebra, finite dimensional over $\mathbb{C}$, and let P be a prime ideal of $U=U(\mathbf{g})$. Recall from [6] that there is a unique largest Ore subset, $\mathscr{S}(P)$, of $\mathscr{C}(P)$, where $\mathscr{C}(P)=\{r \in U : r+P \text{ is regular in } U/P\}$, and $U_{\mathscr{S}(P)}/PU_{\mathscr{S}(P)}$ is the ring of fractions of U/P.

If X is a subset of a left module over a ring R, $l(X)=\{r \in R : rX=0 \text{ for all } x \in X\}$. For X in a right R-module, $r(X)=\{r \in R : Xr=0 \text{ for all } x \in X\}$.

Theorem. *Let $\mathbf{g}$ be a solvable Lie algebra, finite dimensional over $\mathbb{C}$, and let P be a prime ideal of $U=U(\mathbf{g})$. Let*

$$0 \to V \to M \to A \to 0 \tag{1}$$

be an exact sequence of finitely generated left U-modules with V $\mathscr{S}(P)$-torsion and A isomorphic to a non-zero left ideal of U/P. Then V is not essential in M.

Proof. There is a finite chain $0=V_0 \subset V_1 \subset \ldots \subset V_t = V$ of submodules of V such that, for $1 \leqq i \leqq t$, $l(V_i/V_{i-1})$ is prime and equals the annihilator of every non-zero submodule of V_i/V_{i-1}. Arguing by induction on t, one may pass to the case $t=1$, so that there is a prime ideal Q of U such that

$$l(X)=Q \; \forall 0 \neq X \subseteq V. \tag{2}$$

Case 1. Suppose that $V \subsetneqq \{m \in M : Qm = 0\} =: M'$. There is then no harm in replacing M by M', so that $QM = 0$. Now $Q \subseteq P$; if $Q = P$ then V cannot be essential in M since A contains a non-zero free U/Q-module.

Suppose that $Q \subsetneqq P$, and that V is essential in M. Thus, by (2),

$$l(X) \subseteq Q \; \forall 0 \neq X \subseteq M. \tag{3}$$

Choose $\gamma \in P \setminus Q$ such that $\gamma + Q$ is normal in $\bar{U} = U/Q$. Let $m \in M \setminus V$. There exists $d \in \mathscr{S}(P)$ with $d(\gamma m) = 0$. Since $Qm = 0$ and there exists $d' \in U \setminus Q$ with $d\gamma + Q = \gamma d' + Q$, $\gamma d' m = 0$. Therefore $\gamma(U d' m) \subseteq U \gamma d' m = 0$, and hence $U d' m = 0$ by (3). Since $m \notin V$, $d' \in P$. It follows that $\gamma \bar{U}/\gamma \bar{P}$ is a free right U/P-module which is left $\mathscr{S}(P)$-torsion. Since $\mathscr{S}(P)$ is an Ore subset of $\mathscr{C}(P)$, this contradicts [5, Lemma 1.4].

Case 2. Suppose that $V = \{m \in M : Qm = 0\}$. Suppose for a contradiction that V is essential in M. Thus (3) holds for M. Let $I = l(M)$. Replacing M by a submodule if necessary, and using the Noetherian condition on U, we may assume that

$$l(M') = I \; \forall M' \subseteq M \quad \text{with } M' \not\subseteq V. \tag{4}$$

Now $QP \subseteq I \subseteq Q \cap P$. We claim that

$$I \subsetneqq Q \cap P. \tag{5}$$

Suppose that $I = Q \cap P$; then $P(QM) \subseteq (P \cap Q) M = 0$. But $QM \neq 0$ by hypothesis, so $P \subseteq Q$ by (3). Thus $I = P$ and V cannot be essential since A contains a non-zero free U/I-module. Thus (5) holds.

Put $\bar{I} = \{u \in U : du \in I, \; d \in \mathscr{C}(Q \cap P)\}$, so $\bar{I}$ is an ideal, since certainly $\bar{I} \subset Q \cap P$, $Q \cap P/I$ is a $U/Q \cap P$-module, and the image of $\mathscr{C}(Q \cap P)$ in $U/Q \cap P$ is an Ore set. We claim that

$$I = \bar{I}. \tag{6}$$

Suppose not; since U is right Noetherian there exists $d \in \mathscr{C}(Q \cap P)$ with $d\bar{I} \subseteq I$ and since $\bar{I} M \neq 0$ we arrive at a contradiction to (3).

Now choose $\gamma \in (P \cap Q) \setminus I$ such that $\gamma + I$ is normal in U/I. Let $m \in M$. There exists $d \in \mathscr{S}(P)$ with $d\gamma m = 0$. Let $d' \in U$ satisfy $d\gamma + I = \gamma d' + I$. By [5, Lemma 5.3], $Q \rightsquigarrow P$ and so $\mathscr{S}(P) \subseteq \mathscr{C}(Q)$ by [5, Lemma 1.4]. In particular, $d \in \mathscr{C}(Q)$. Thus (6) shows that $d\gamma \notin I$, so $\gamma d' \notin I$ and therefore $d' \notin P$, since $QP \subseteq I$. Hence, $d' m \notin V$. But $\gamma \in l(U d' m)$, so $I \subsetneqq l(U d' m)$. This is a contradiction to (4), and so the result is proved.

Remarks 3.2. (i) It is tempting to propose the following generalisation of 3.1: Every uniform $U(\mathbf{g})$-module is either $\mathscr{S}(P)$-torsion or $\mathscr{S}(P)$ torsion free. This is false. Take $\mathbf{g} = \mathbb{C}x \oplus \mathbb{C}y$ with $[x, y] = x$ and let M be the module $U(\mathbf{g})/U(\mathbf{g})(y-1)(x-1)$. This is uniform [15, Lemma 4]; if one takes P to be the augmentation ideal, so $\mathscr{S}(P) = \bigcap_{n \in \mathbb{Z}} \mathscr{C}(\langle x, y - n \rangle)$, then M is neither torsion nor torsion-free with respect to $\mathscr{S}(P)$.

(ii) Case 2 of the proof of 3.1 can also be deduced from results of A.V. Jategaonkar [11, § 3].

4. Ext Groups for k-finite Modules

4.1. Let K be a (commutative) field and $\mathbf{g}$ a Lie algebra, finite dimensional over K. In this section we shall adopt the notations of [1]. In particular, $U=U(\mathbf{g})$, $\mathbf{g}^2=\mathbf{g}\oplus\mathbf{g}^0$, where $\mathbf{g}^0$ is the opposite Lie algebra to $\mathbf{g}$, $\mathbf{k}=\{(x,-x): x\in\mathbf{g}\}\subseteq\mathbf{g}^2$. Further $U^2=U(\mathbf{g}\oplus\mathbf{g}^0)=U(\mathbf{g})\otimes_K U(\mathbf{g}^0)=U(\mathbf{g})\otimes_K U(\mathbf{g})^0$; and $\{U_n\}_{n\in\mathbb{N}}$ denotes the canonical filtration on U, with $\operatorname{gr}U=S(\mathbf{g})$ the associated graded ring. The U^2-module M is $\mathbf{k}$-finite if $\dim_K U(\mathbf{k})m<\infty$ for every $m\in M$. Denote by $\mathscr{H}$ the category of $\mathbf{k}$-finite (left) U^2-modules, and by $\mathscr{H}_f$ the subcategory of $\mathscr{H}$ consisting of finitely generated U^2-modules.

For example, $U/I\in\mathscr{H}_f$ when I is a two-sided ideal of U.

4.2. Let V be a finite dimensional U-module. Denote by Φ_V the U^2-module $V\otimes_K U$, where the actions are defined as follows, see [1], 2.2: - For $x\in\mathbf{g}$, $v\in V$ and $u\in U$, $x(v\otimes u)=xv\otimes u+v\otimes xu$; $(v\otimes u)x=v\otimes ux$. The K-dual of V, $\operatorname{Hom}_K(V,K)$, is a U-module via the contragradient action, denoted V^*. For completeness we record the following results, taken from various parts of [1].

Proposition. *(i) $\mathscr{H}$ and $\mathscr{H}_f$ are closed under taking subquotients.*

(ii) Let V be a finite dimensional U-module. Then $\Phi_V\in\mathscr{H}_f$, and Φ_V is finitely generated and free as a right and as a left U-module.

(iii) The U^2-module $\operatorname{Hom}_U(\Phi_V,U)$ is isomorphic to Φ_{V^*}, and so belongs to $\mathscr{H}_f$. (Recall that $\operatorname{Hom}_U(-,-)$ is computed on left modules.)

Proof. (i) Obvious.

(ii) Let $u\in U_n$, $y=(x,-x)\in\mathbf{k}$, and $v\in V$. Then $y(v\otimes u)=xv\otimes u+v\otimes[x,u]\in V\otimes U_n$, so $\Phi_V\in\mathscr{H}$. Clearly, if $\{e_i: 1\leqq i\leqq n\}$ is a K-basis for V, Φ_V is free as a right U-module on $\{e_i\otimes 1: 1\leqq i\leqq n\}$.

Define $\varphi: U\otimes_K V\to V\otimes_K U$ by $\varphi(u\otimes v)=u(v\otimes 1)$ for $u\in U$, $v\in V$. The map φ is well defined: If $x=\sum_i u_i\otimes v_i=0$ in $U\otimes_K V$, write x as $\sum_{i,j}\lambda_{ij}u_i\otimes e_j$, so that $\sum_i\lambda_{ij}u_j=0$ for each j. Now

$$\sum_i u_i(v_i\otimes 1)=\sum_i u_i\Big(\sum_j\lambda_{ij}e_j\otimes 1\Big)=\sum_j\Big(\sum_i\lambda_{ij}u_i\Big)(e_j\otimes 1);$$

and thus $\varphi(x)=0$. Secondly, φ is left U-linear: Let $x\in\mathbf{g}$, $u\in U$ and $v\in V$. Then

$$\varphi(x(u\otimes v))=\varphi(xu\otimes v)=xu(v\otimes 1)=x(u(v\otimes 1))=x\varphi(u\otimes v).$$

We claim that φ is an isomorphism. Filter $U\otimes_K V$ by the subspaces $\{U_n\otimes V: n\in\mathbb{N}\}$; the $\operatorname{gr}U$-module $\operatorname{gr}(U\otimes_K V)$ is clearly $\operatorname{gr}U\otimes_K V$. Note that $\{V\otimes_K U_n: n\in\mathbb{N}\}$ is a filtration of the left U-module $V\otimes\operatorname{gr}U$, and that φ induces a homomorphism $\operatorname{gr}\varphi$ of graded modules, $\operatorname{gr}\varphi:\operatorname{gr}U\otimes V\to V\otimes\operatorname{gr}U: s\otimes v\to v\otimes s$. Since $\operatorname{gr}\varphi$ is an isomorphism of $\operatorname{gr}U$-modules, φ is an isomorphism of U-modules. Thus Φ_V is free as a left U-module, with basis $\{e_i\otimes 1: 1\leqq i\leqq n\}$.

(iii) Define $\Theta:\operatorname{Hom}_U(\Phi_V,U)\to V^*\otimes_K U$ as follows. Let $\{u_j: j\geqq 0\}$ be a K-basis for U. Let $\varphi\in\operatorname{Hom}_U(V\otimes_K U,U)$, and write

$$\varphi(e_i\otimes 1)=\sum_j\lambda_{ij}u_j. \tag{1}$$

Denote by $\Psi_j^{\varphi} \in V^*$ the map defined by

$$\Psi_j^{\varphi}(e_i) = \lambda_{ij}, \tag{2}$$

and set

$$\Theta(\varphi) = \sum_j \Psi_j^{\varphi} \otimes u_j.$$

Let $\{\varepsilon_i : 1 \leqq i \leqq n\}$ be the dual basis of $\{e_i : 1 \leqq i \leqq n\}$. Thus

$$\Psi_j^{\varphi} = \sum_i \lambda_{ij} \varepsilon_i. \tag{3}$$

To define the inverse map Θ', let $f \in V^* \otimes_K U$, so there exist (uniquely determined) $\Psi_j \in V^*$ such that $f = \sum_j \Psi_j \otimes u_j$. Since Φ_V is a free left U-module, $\Theta'(f) \in \mathrm{Hom}_U(\Phi_V, U)$ is well-defined by setting

$$\Theta'(f)(e_i \otimes 1) = \sum_j \Psi_j(e_i)\, u_j,$$

for $1 \leqq i \leqq n$. From the definitions, $\Theta\Theta' = \Theta'\Theta = 1$. We now prove that

$$\Theta \text{ is right } U\text{-linear.} \tag{4}$$

Let $\varphi \in \mathrm{Hom}_U(\Phi_V, U)$ and $x \in \mathbf{g}$. By definition, $(\varphi x)(e_i \otimes 1) = \varphi(e_i \otimes 1)\, x = \sum_j \lambda_{ij} u_j x$. Writing $u_j x = \sum_r \mu_{rj} u_r$, we get

$$(\varphi x)(e_i \otimes 1) = \sum_r (\sum_j \lambda_{ij} \mu_{rj})\, u_r. \tag{5}$$

By (3) and (5),

$$\Psi_r^{\varphi x} = \sum_i (\sum_j \lambda_{ij} \mu_{rj})\, \varepsilon_i = \sum_j \mu_{rj} (\sum_i \lambda_{ij} \varepsilon_i) = \sum_j \mu_{rj}\, \Psi_j^{\varphi}.$$

Therefore,

$$\Theta(\varphi x) = \sum_r \Psi_r^{\varphi x} \otimes u_r = \sum_r \sum_j \mu_{rj}\, \Psi_j^{\varphi} \otimes u_r.$$

On the other hand,

$$\Theta(\varphi)\, x = \sum_j \Psi_j^{\varphi} \otimes u_j x = \sum_j (\Psi_j^{\varphi} \otimes \sum_r \mu_{rj} u_r),$$

and so $\Theta(\varphi x) = \Theta(\varphi)\, x$, proving (4).

Finally, we prove that

$$\Theta \text{ is left } U\text{-linear.} \tag{6}$$

Let $\varphi \in \mathrm{Hom}_U(\Phi_V, U)$ and $x \in \mathbf{g}$. Then $x\, \varphi(e_i \otimes 1) = \varphi(e_i \otimes x) = \varphi(x(e_i \otimes 1) - x\, e_i \otimes 1)$, so by linearity of φ,

$$(x\, \varphi)(e_i \otimes 1) = x(\varphi(e_i \otimes 1)) + \varphi(-x\, e_i \otimes 1).$$

Write $x\, e_i = \sum_t \gamma_{it} e_t$, so that by (1) and (2),

$$\begin{aligned} x\varphi(e_i\otimes 1) &= x(\sum_j \Psi_j^{\varphi}(e_i)u_j) - \sum_t \gamma_{it}\varphi(e_t\otimes 1) \\ &= \sum_j \Psi_j^{\varphi}(e_i)xu_j - \sum_t \gamma_{it}\sum_j \Psi_j^{\varphi}(e_t)u_j \\ &= \sum_j \Psi_j^{\varphi}(e_i)xu_j - \sum_j (\sum_t \gamma_{it}\Psi_j^{\varphi}(e_t))u_j \\ &= \sum_j \Psi_j^{\varphi}(e_i)xu_j + \sum_j \Psi_j^{\varphi}(-xe_i)u_j \\ &= \sum_j \Psi_j^{\varphi}(e_i)xu_j + \sum_j (x\Psi_j^{\varphi})(e_i)u_j. \end{aligned}$$

Thus,

$$\Theta(x\varphi) = \sum_j \Psi_j^{\varphi}\otimes xu_j + \sum_j x\Psi_j^{\varphi}\otimes u_j. \tag{7}$$

However,

$$x\Theta(\varphi) = x(\sum_j \Psi_j^{\varphi}\otimes u_j) = \sum_j x\Psi_j^{\varphi}\otimes u_j + \sum_j \Psi_j^{\varphi}\otimes xu_j, \tag{8}$$

so (6) follows from (7) and (8).

4.3. Proposition. *Let $M\in\mathscr{H}_f$. Then there is a finite dimensional U-module V such that M is an image of Φ_V. In particular, M is finitely generated as a right and left U-module.*

Proof. Let $\{m_1,\ldots,m_t\}$ generate M as a U^2-module. By definition of $\mathscr{H}$, there is a finite dimensional $\mathbf{k}$-submodule V of M with $\{m_1,\ldots,m_t\}\subseteq V$. Since $\mathbf{k}\cong\mathbf{g}$, V is a finite dimensional U-module, so there is a U^2-homomorphism π from Φ_V onto M, with $\pi(v\otimes u)=vu$. The final statement of the propostion now follows from 4.2 *(ii)*.

4.4. Theorem *(Bernstein). Let $M\in\mathscr{H}_f$. Then $\mathrm{Ext}_U^i(M,U)\in\mathscr{H}_f$ for every $i\geqq 0$. In particular, $\mathrm{Ext}_U^i(M,U)$ is a finitely generated right and left U-module.*

Peoof. By 4.3 there is a resolution of M in $\mathscr{H}_f$ of the form

$$\ldots\to\Phi_{V_i}\to\Phi_{V_{i-1}}\to\ldots\to\Phi_{V_0}\to M\to 0,$$

with each V_i a finite dimensional U-module. Applying $\mathrm{Hom}_U(-,U)$, we get, in view of 4.2*(iii)*, the complex in $\mathscr{H}_f$

$$0\to\Phi_{V_0^*}\to\Phi_{V_1^*}\to\ldots\to\Phi_{V_i^*}\to\ldots.$$

By 4.2 *(ii)*, the cohomology of this complex is $\mathrm{Ext}_U^*(M,U)$, so the result follows from 4.2*(i)*.

4.5. Corollary. *Let I be an ideal of U. For every $i\geqq 0$, $\mathrm{Ext}^i(U/I,U)$ is finitely generated as a right and a left U-module.*

5. The Structure of Certain Ext Groups

5.1. In this paragraph, U will denote the enveloping algebra of a complex Lie algebra $\mathbf{g}$ with $\dim_{\mathbf{C}}\mathbf{g}=n<\infty$, and $d(X)$ will denote the Gelfand-Kirillov di-

mension of the U-module X. (See [3] for the basic properties of $d(\ )$.) We shall say that X is *d-homogeneous* (of dimension t) if $d(Y)=t$ for all non-zero submodules Y of X. Although our main interest lies in the bimodules $\mathrm{Ext}^i_U(U/I, U)$ for I an ideal of U, it pays to consider somewhat more general entries in the left argument.

Let X be a finitely generated U-module. The *grade* of X, $G(X)$, is $\inf\{i: \mathrm{Ext}^i_U(X, U)\neq 0\}$. We shall need the formula [2, 2.7.2]

$$G(X)+d(X)=n. \tag{1}$$

Our first result is hardly more than a reformulation of special cases of earlier work.

Proposition. *Let $i\in\mathbb{N}$, let X be a $(U-U)$-bimodule finitely generated on left and right, and write E_i for $\mathrm{Ext}^i_U(X, U)$. Let $d(X)=m$ and set $t=n-m$. Let $I=r(X)$ and set $r=\inf\{d(U/J): I\subseteq J\subsetneqq U\}$.*

(i) $E_i=0$ if $i<t$ or $i>n-r$, while $E_t\neq 0$.

(ii) If $t<i\leqq n-r$ then $d(E_i)<m$ and $I\subsetneqq l(E_i)$.

(iii) Suppose $m=r$. Then $E_i\neq 0$ if and only if $i=t$. In particular, $\mathrm{pr\,dim}(X)=t$. *(Note that $m=r$ if $X=U/I$ for a maximal ideal I of U, for example.)*

Proof. (i) By (1), $E_i=0$ for $i<t$. Suppose $i>n-r$. Now $E_{i_{|U}}$ is finitely generated, so by [13, Lemma 2] and [2, 2.7.3],

$$d({}_U|E_i)\leqq d(E_{i_{|U}})\leqq n-i<r. \tag{2}$$

Let $E_i=\sum_{j=1}^{s} x_j U$, so $I\subseteq l(E_i)=\bigcap_j l(x_j)$. By (2) $d(U/l(x_j))<r$ for $1\leqq j\leqq s$, so $d(U/l(E_i))<r$. The definition of r shows that $l(E_i)=U$; that is, $E_i=0$.

(ii) follows by the same proof as *(i)*.

(iii) For $i\neq t$, $E_i=0$ by *(i)*. Since there must exist i with $E_i\neq 0$, we must have $E_t\neq 0$.

5.2. We shall now study the groups E_i in the case where X is a prime factor ring of an enveloping algebra U of a *solvable* Lie algebra $\mathbf{g}$. (We continue to assume $\dim_{\mathbb{C}}\mathbf{g}=n$.) The notation developed here will be fixed for the rest of § 5.

Let P be a prime ideal of U and let $\mathscr{S}(P)$ be the unique largest Ore set in $\mathscr{C}(P)$, as in 3.1. Put $R=U_{\mathscr{S}(P)}$. We need the following facts from [6]. The ideal PR has a regular normalising set of generators $x_1, \ldots, x_t$ (see 2.1), where t = height P. For each i, $x_i+\sum_{j<i} x_j R$ is a $\mathbf{g}$-eigen-vector under the action of $\mathbf{g}$ on $R/\sum_{j<i} x_j R$ induced by the adjoint action on U, with eigenvalue $m_i\lambda_i$, where m_i is a non-negative integer and λ_i is an eigenvalue for $\mathrm{ad}\,\mathbf{g}$ acting on $\mathbf{g}$. Set $\mathscr{E}(P)=\{m_i\lambda_i: 1\leqq i\leqq t\}$ and $O(P)=\langle\mathscr{E}(P)\rangle$, a finitely generated subgroup of $\mathbf{g}^*$. Then $\mathscr{S}(P)=\bigcap_{\lambda\in O(P)}\mathscr{C}(\tau_\lambda(P))$, where τ_λ is the winding automorphism of U given by $x\to x+\lambda(x)$, for $x\in\mathbf{g}$. The primitive ideals of R are $\tau_\lambda(P)R$, for $\lambda\in O(P)$. If Q is such a primitive ideal then $R/Q\cong D$, the division ring of fractions of U/P. The global dimension of R is t.

Recall also that by [17, Théorème 2.6]

$$t+d(U/P)=n. \tag{3}$$

We put $r=\inf\{d(U/J)\colon P\subseteqq J\subsetneqq U\}$.

We begin with an addendum to 5.1.

Lemma 5.3. *If $i>t$ then $\mathrm{Ext}^i_U(U/P,U)$ is right and left $\mathscr{S}(P)$-torsion. If $i<t$ or $i>n-r$ then $\mathrm{Ext}^i_U(U/P,U)=0$.*

Proof. Apply Proposition 1.6 with $S=U_{\mathscr{S}(P)}$, noting that S has global dimension t, to show that $E_{i|U}$ is right torsion. By [2, 2.7.3], [13, Lemma 2] and (3), $d({}_U|E_i)\leqq n-i<n-t=d(U/P)$. Since $E_{i|U}$ is finitely generated, $P\subsetneqq l(E_i)$ as in the proof of 5.1. Hence, $l(E_i)\cap\mathscr{S}(P)\neq\varnothing$. The other assertions are given by 5.1.

The remainder of § 5 is devoted to $E_t=:\mathrm{Ext}^t_U(U/P,U)$.

Lemma 5.4. *The left $\mathscr{S}(P)$-torsion submodule T of E_t is zero.*

Proof. Since $E_{t|U}$ is finitely generated, the bimodule T is finitely generated on the right. Hence there exists $c\in\mathscr{S}(P)$ with $cT=0$, and so $P\subsetneqq J=l(T)$. Choose $x\in J\setminus P$ with $x+P$ normal in U/P. The exact sequence

$$0\to U/P\xrightarrow{\times x} U/P\to U/P+Ux\to 0$$

yields

$$\mathrm{Ext}^t(U/P+Ux,U)\to E_t\xrightarrow{\times x}E_t, \tag{4}$$

where the second arrow is afforded by *left* multiplication by x. Since $d(U/P+Ux)<d(U/P)=n-t$ by [3, Satz 3.4] and (3), $\mathrm{Ext}^t(U/P+Ux,U)=0$ by 5.1*(i)*. Thus (4) show that left multiplication by x is a monomorphism on E_t. Therefore $T=0$.

5.5. We can now state our central result. The notation remains that given in 5.2.

Theorem. *(i) As left modules E_t is isomorphic to a nonzero left ideal of U/P.*

(ii) As right modules E_t is isomorphic to a non-zero right ideal of $U/\tau_\lambda(P)$, where $\lambda=-\sum_{\mu\in\mathscr{E}(P)}\mu$.

Proof. Set $R=U_{\mathscr{S}(P)}$. Note that if $\mu,\nu\in(\mathbf{g}/[\mathbf{g},\mathbf{g}])^*$ then $\tau_\mu\tau_\nu=\tau_{\mu+\nu}$. Thus, by Corollary 2.6 $\tilde{E}_t=:\mathrm{Ext}^t_R(R/PR,R)$ is isomorphic as left R-module to R/PR and as right R-module to $R/\tau_\lambda(P)R$. Let W be the right $\mathscr{S}(P)$-torsion submodule of E_t. By 1.6,

$$0\neq E_t/W\subseteq\tilde{E}_t, \tag{5}$$

the embedding being as $(U-U)$-bimodules. Since E_t/W is finitely generated as a right and as a *left* module, by 4.5, E_t/W has the structure attributed to E_t in the statement of the theorem. Thus 5.5 will follow from

Lemma 5.6. *The right $\mathscr{S}(P)$-torsion submodule W of E_t is zero.*

Proof. Since $\lambda\in O(P)$, $\mathscr{S}(P)=\mathscr{S}(\tau_\lambda(P))$. Now $(E_t/W)_{|U}\subseteq U/\tau_\lambda(P)$, so by 3.1 W is not an essential right submodule of E_t. Fix a non-zero right submodule Y of E_t with $Y\cap W=0$, so that $Y\tau_\lambda(P)=0$.

By [2, 2.7.3], $d(E_{t|U}) \leqq n-t$. Suppose E_t has a non-zero right submodule X with $d(X) < n-t$. Then without loss of generality X is a bimodule, so that $d({}_U|X) < n-t$. Since $PE_t = 0$ and $d(U/P) = n-t$, we must have $P \subsetneqq l(x)$ for all $x \in X$. But R/PR is the full division ring of fractions of U/P, so that $l(x) \cap \mathscr{S}(P) \neq \varnothing$ for all $x \in X$. By 5.4, $X = 0$. That is,

$$E_t \text{ is } d\text{-homogeneous of dimension } n-t. \tag{6}$$

The exact sequence of finitely generated right U-modules

$$0 \to Y \oplus W \xrightarrow{i} E_t \to Q \to 0$$

yields an exact sequence of finitely generated left modules

$$\mathrm{Ext}^t_U(E_t, U) \xrightarrow{i^*} \mathrm{Ext}^t_U(Y \oplus W, U) \to \mathrm{Ext}^{t+1}_U(Q, U). \tag{7}$$

By [2, 2.7.3],

$$d(\mathrm{Ext}^{t+1}_U(Q, U)) \leqq n-(t+1). \tag{8}$$

Suppose that

$$W \neq 0. \tag{9}$$

Then $\mathrm{Ext}^t_U(Y \oplus W, U) \cong \mathrm{Ext}^t_U(Y, U) \oplus \mathrm{Ext}^t_U(W, U)$, where

$$d(\mathrm{Ext}^t_U(Y, U)) = d(\mathrm{Ext}^t_U(W, U) = n-t, \tag{10}$$

by (6) and [9, 3.3.4]. By (8) and (10),

$$n-t = d(\mathrm{im}\, i^* \cap \mathrm{Ext}^t_U(Y, U)) = d(\mathrm{im}\, i^* \cap \mathrm{Ext}^t_U(W, U)). \tag{11}$$

But $\mathrm{Ext}^t_U(E_t, U)$ is an extension of U/P by a module of dimension less than $n-t$, by [2, 2.4.15, 2.7.5, 2.7.8 and 2.7.11]. In particular, its d-homogeneous homomorphic images of dimension $n-t$ are all uniform. This contradicts (10) and (11), so the assumption (9) must be false.

6. Some Applications

This paragraph contains 3 applications of the result of § 5.

6.1. From parts of 5.1 and 5.5 one can read off information about the minimal injective resolution of U, (where $U = U(\mathbf{g})$ and $\mathbf{g}$ is solvable).

The ring U is Noetherian, so every injective U-module is a direct sum of indecomposable injective modules [16, V.4.5]. An indecomposable injective U-module has the form $E(X)$, the injective hull of a d-critical U-module X with $l(X) = P$ prime, (since indecomposable injectives are uniform and every non-zero module contains a d-critical submodule with prime annihilator). Now X is either torsion free as U/P-module, (in which case X is (isomorphic to) a left ideal of the domain U/P and $E(X) = E(U/P)$), or X is (U/P)-torsion. The upshot of these remarks is: An injective U-module E has the form $\bigoplus_{P \in \mathrm{Spec}(U)} c(P)\, E(U/P) \oplus E'$, where the $c(P)$ are cardinals and E' is the hull of a

sum of d-critical modules each of which is torsion modulo its annihilator. In particular, this applies to the terms $I_i (0 \leqq i \leqq n)$ of a minimal injective resolution of U. Following [14] we write $\mu_i(P, U)$ for the cardinal $c(P)$ given by the decomposition of I_i.

The result we obtain is due to M.-P. Malliavin [14]. Her proof is completely different from ours.

Theorem. *Let P be a prime ideal of height t.*
Then

$$\mu_i(P, U) = \begin{cases} 0 & \text{if } i \neq t; \\ 1 & \text{if } i = t. \end{cases}$$

Proof. Since $\mathrm{Hom}_U\left(U/P, \bigoplus_{j=1}^{r} U/P\right) \cong \bigoplus_{j=1}^{r} U/P$ as left U-modules, it's easy to see – in fact, it's essentially the definition of Ext in terms of injective resolutions – that $\mu_i(P, U)$ is precisely the maximal uniform rank of a (U/P)-torsion free submodule of the *left* module $\mathrm{Ext}^i_U(U/P, U)$. Thus the theorem follows from 5.3 and 5.5.

6.2. Let P be a prime ideal of the enveloping algebra U of the n-dimensional Lie algebra $\mathbf{g}$, with $d(U/P) = m$. Generalising a concept from the commutative theory [12, p. 126], we shall say that P is *perfect* if the projective dimension of U/P, $\mathrm{pr\,dim}(U/P)$, is exactly $n - m =: t$; that is, P is perfect if $G(U/P) = \mathrm{pr\,dim}(U/P)$, by 5.1(1). By definition, $G(U/P) \leqq \mathrm{pr\,dim}(U/P)$. By Proposition 5.1 *(iii) maximal ideals are perfect.*

If $\mathbf{g}$ is abelian and P has height $n-1$, then $\mathrm{pr\,dim}(U/P) = n-1$ [12, Theorem 176]; that is, P is perfect. This result generalises as follows.

Proposition. *Let P be a prime ideal of the enveloping algebra U of the solvable Lie algebra $\mathbf{g}$ with $\dim_{\mathbb{C}} \mathbf{g} = n$. Let $d(U/P) = m$ and suppose that every prime ideal strictly containing P is maximal. Then* $\mathrm{pr\,dim}(U/P) = n - m = \mathit{height}(P)$.

Proof. By (3) of 5.2 and 5.5 only $\mathrm{pr\,dim}(U/P) \leqq n - m$ remains to be proved. Choose $x \in U \setminus P$ with $x + P$ normal in U/P. Put $t = n - m$. The exact sequence

$$0 \to U/P \xrightarrow{\times x} U/P \to U/P + Ux \to 0$$

yields, for $i \geqq 1$

$$\mathrm{Ext}^{t+i}(U/P, U) \xrightarrow{x^* \times} \mathrm{Ext}^{t+i}(U/P, U) \to \mathrm{Ext}^{t+i+1}(U/P + Ux, U). \tag{1}$$

Now every prime ideal Q of U containing $P + Ux$ is maximal, by hypothesis, and has $d(U/Q) = m - 1$ by [9, 4.3.1]. By 5.1 *(iii)*, $\mathrm{pr\,dim}(U/P + Ux) = n - (m-1) = t + 1$. Hence the right hand term of (1) is 0 and so left multiplication of $\mathrm{Ext}^{t+i}(U/P, U)$ by x is onto. Choose x in $l(\mathrm{Ext}^{t+i}(U/P, U)) \setminus P$ (as we may do by 5.1 *(ii)*) to deduce that $\mathrm{Ext}^{t+i}(U/P, U) = 0$.

6.3. Let $\mathbf{p} \subset \mathbf{g}$ be Lie algebras, with $\mathbf{p}$ solvable. Let P be a prime ideal of $U(\mathbf{p})$ of height t. By 1.6,

$$\mathrm{Ext}^i_{U(\mathbf{p})}(U(\mathbf{p})/P, U(\mathbf{p})) \otimes_{U(\mathbf{p})} U(\mathbf{g}) \cong \mathrm{Ext}^i_{U(\mathbf{g})}(U(\mathbf{g})/U(\mathbf{g})P, U(\mathbf{g})), \tag{2}$$

as right $U(\mathbf{g})$-modules, for all $i \geqq 0$. Thus information about the cohomology of induced modules of the form $U(\mathbf{g})/U(\mathbf{g})P$ can be read off from the results of § 5.

As an example, we consider Verma modules. Let $\mathbf{g}$ be semisimple and $\mathbf{p}=\mathbf{b}$ a Borel subalgebra containing a Cartan subalgebra $\mathbf{h}$, so $\mathbf{b}=\mathbf{h}\oplus\mathbf{n}^+$. Let $\dim_{\mathbb{C}}\mathbf{b}=t$. Let $\lambda\in\mathbf{h}^*$ and set $M(\lambda)=U(\mathbf{g})\otimes_{U(\mathbf{b})}\mathbb{C}_{\lambda-\rho}$, where ρ is half the sum of the positive roots and $\mathbb{C}_\mu$ is the 1-dimensional $U(\mathbf{b})$-module of weight μ, for $\mu\in\mathbf{h}^*$.

Theorem. (Duflo [8]) *With the above notation,* $\mathrm{Ext}^i_{U(\mathbf{g})}(M(\lambda), U(\mathbf{g}))$ *is* 0 *for* $i\neq t$, *while* $\mathrm{Ext}^t_{U(\mathbf{g})}(M(\lambda), U(\mathbf{g}))\cong\mathbb{C}_{\lambda+\rho}\otimes_{U(\mathbf{b})}U(\mathbf{g})$ *as right* $U(\mathbf{g})$*-modules. Moreover,*

$$l(M(\lambda))=r(\mathrm{Ext}^t_{U(\mathbf{g})}(M(\lambda), U(\mathbf{g})).$$

Proof. The first claim follows from (2) and 5.1. Let $\mathbb{C}_{\lambda-\rho}=U(\mathbf{g})/P$. Then $\sum_{\mu\in\mathscr{E}(P)}\mu=\mathrm{tr}\,\mathrm{ad}\,\mathbf{b}_{|\mathbf{b}}=2\rho$. Thus by 5.5 *(ii)* there is an isomorphism of right modules

$$\mathrm{Ext}^t_{U(\mathbf{g})}(\mathbb{C}_{\lambda-\rho}, U(\mathbf{g}))\cong\mathbb{C}_{(\lambda-\rho)+2\rho}.$$

Coupled with (2), this yields the result. The final equality now follows from [8, Théorème 2].

7. Non-prime Ideals

Throughout this paragraph, $\mathbf{g}$ is solvable and $U=U(\mathbf{g})$. We shall examine $\mathrm{Ext}^i(U/I, U)$ for an arbitrary ideal I of U.

7.1. We must first extend parts of § 5 to modules, rather than bimodules.

Proposition. *Let* P *be a prime ideal of* U *of height* t, *and let* M *be a finitely generated torsion free left* U/P*-module of uniform rank* r, *(where* $r\geqq 1$*).*

Let $\lambda=-\sum_{\mu\in\mathscr{E}(P)}\mu$. *Then* $\mathrm{Ext}^i(M, U)$ *is (i)* 0, *for* $i<t$; *(ii) a non-zero submodule of a finitely generated free right* $U/\tau_\lambda(P)$*-module of rank* r, *for* $i=t$; *(iii)* $\mathscr{S}(P)$*-torsion, for* $i>t$.

Note. A module M over a prime ring R is "torsion free" if it is torsion free with respect to $\mathscr{C}_R(0)$.

Proof. (i) [2, 2.7.1] and Tauvel's identity (3) of § 5.

(ii) Let F be the ring of quotients of U/P, so $M\subseteq F\otimes_U M\cong F^{(r)}$ by hypothesis. Since every finitely generated U-submodule of F is contained in a free U/P-module of rank 1, we easily see that M is a submodule of a free U/P-module B of rank r with B/M $\mathscr{C}(P)$-torsion. Also, since M is essential in $F^{(r)}$, we can find a submodule C of M which is U/P-free of rank r; again, M/C is $\mathscr{C}(P)$-torsion. Put $m=d(U/P)$, so $m=n-t$ by (3) of § 5, and we have exact sequences

$$0\to M\to B\to B/M\to 0 \tag{1}$$

and

$$0\to C\to M\to M/C\to 0 \tag{2}$$

with $d(B/M)<m$, $d(M/C)<m$.

By [2, 2.7.1],

$$\mathrm{Ext}^t(B/M, U)=0=\mathrm{Ext}^t(M/C, U),$$

so (1) and (2) yield exact sequences

$$0\to\mathrm{Ext}^t(B, U)\to\mathrm{Ext}^t(M, U) \tag{3}$$

and

$$0\to\mathrm{Ext}^t(M, U)\to\mathrm{Ext}^t(C, U). \tag{4}$$

Thus *(ii)* follows from (3), (4) and 5.5 *(ii)*.

(iii) Since $\mathrm{gl\,dim}\, U_{\mathscr{S}(P)}=t$ (see 5.2), this follows from 1.6.

7.2. Let I be an ideal of U with $d(U/I)=m$. Put $t=n-m$, where $n=\dim_{\mathbb{C}}\mathfrak{g}$. For $i\neq t$, information about $\mathrm{Ext}^i_U(U/I, U)$ is given by Proposition 5.1. For $i=t$, we offer

Theorem. *Let the notation be as above, and let $P_1, \ldots, P_s$ be the prime ideals of U with $I\subseteq P_j$ and $d(U/P_j)=m$ for $1\leqq j\leqq s$. For each j, let τ_j be the winding automorphism defined by $\tau_j=\tau_{\lambda_j}$, where $\lambda_j=-\sum_{\mu\in\mathscr{E}(P_j)}\mu$. Put $P_j'=\tau_j(P_j)$, $1\leqq j\leqq s$.*

(i) $\mathrm{Ext}^t(U/I, U)$ *is d-homogeneous of dimension m.*

(ii) There exists a finite product Y of the ideals P_j' with $\mathrm{Ext}^t(U/I, U)\,Y=0$.

Proof. Let J/I be the sum of all the submodules of U/I of dimension less than m, so $J\subseteq\bigcap_j P_j$ and U/J is d-homogeneous with $d(U/J)=m$. The two outer terms of the exact sequence

$$\mathrm{Ext}^{t-1}(J/I, U)\to\mathrm{Ext}^t(U/J, U)\to\mathrm{Ext}^t(U/I, U)\to\mathrm{Ext}^t(J/I, U)$$

are zero by [2, 2.7.1], so without loss of generality $I=J$. There is a chain of left ideals of U/I

$$0=W_0\subset W_1\subset\ldots\subset W_k=U/I$$

with each factor W_p/W_{p-1} isomorphic to a left ideal of one of the rings U/P_j, $1\leqq j\leqq s$, by [4, 5.4 *(ii)* and 4.4]. The result now follows from Proposition 7.1 and an easy induction on k.

7.3. For certain co-Artinian ideals I we can give a complete description of $\mathrm{Ext}^n_U(U/I, U)$.

Theorem. *Let I be a co-Artinian ideal of U. Let J/I be the nilpotent radical of U/I.*

(i) As left modules, $\mathrm{Ext}^n_U(U/I, U)$ *is the U/I-injective hull of U/J. Let M be a finite dimensional left U-module with $l(M)=I$.*

(ii) $\dim_{\mathbb{C}}\mathrm{Ext}^n_U(M, U)=\dim_{\mathbb{C}}M$.

(iii) $\mathrm{gr}\,\mathrm{Ext}^n_U(M, U)\cong\mathrm{Ext}^n_{\mathrm{gr}\,U}(\mathrm{gr}\,M, \mathrm{gr}\,U)$.

(iv) If I can be expressed as $\bigcap_{j=1}^{t} I_j$, where $I_i+(\bigcap_{j\neq i} I_j)=U$ for all $i=1, \ldots, t$ and each ideal I_j is a product of maximal ideals, then $\mathrm{Ext}^n_U(U/I, U)_{|U}$ *is the right*

$U/\tau_\lambda(I)$-injective hull of $U/\tau_\lambda(J)$, where $\lambda = -\operatorname{tr}\operatorname{ad}\mathbf{g}$. In particular, $r(\operatorname{Ext}^n_U(U/I, U)) = \tau_\lambda(I)$.

Proof. (i) Let $0 \to U \to I_0 \to \ldots \to I_{n-1} \to I_n \to 0$ by the minimal injective resolution of U. Now $\operatorname{Hom}_U(U/I, I_{n-1}) = 0$ by Theorem 6.1, so as left modules $\operatorname{Ext}^n_U(U/I, U)$ is isomorphic to $\operatorname{Hom}_U(U/I, I_n)$. Since each irreducible U/I-module occurs in the socle of I_n with multiplicity one, by Theorem 6.1 again, *(i)* follows.

(ii) This is an easy induction on $\dim_{\mathbb{C}} M$, using the long exact sequence for $\operatorname{Ext}^*_U(-, U)$. It is a special case of [2, 2.7.14].

(iii) By [2, 2.6.10], $\operatorname{gr}\operatorname{Ext}^n_U(M, U)$ is a subfactor of $\operatorname{Ext}^n_{\operatorname{gr} U}(\operatorname{gr} M, \operatorname{gr} U)$. But

$$\begin{aligned}\dim_{\mathbb{C}} \operatorname{gr}\operatorname{Ext}^n_U(M, U) &= \dim_{\mathbb{C}} \operatorname{Ext}^n(M, U) = \dim_{\mathbb{C}} M = \dim_{\mathbb{C}} \operatorname{gr} M \\ &= \dim_{\mathbb{C}} \operatorname{gr}\operatorname{Ext}^n_{\operatorname{gr} U}(\operatorname{gr} M, \operatorname{gr} U),\end{aligned}$$

by *(ii)* applied first in U, then in $\operatorname{gr} U$.

(iv) Since $U/I \cong \sum_{j=1}^{t} \oplus U/I_j$, it's enough to prove *(iv)* for $I = P_1 P_2 \ldots P_s$, a product of s not necessarily distinct maximal ideals. We induct on s – Theorem 5.5 gives the case $s = 1$. Assume $s > 1$, and set $T = P_2 \ldots P_s$ and $P = P_1$, so $I = PT$. The exact sequence

$$0 \to T/I \to U/I \to U/T \to 0$$

yields

$$0 \to \operatorname{Ext}^n_U(U/T, U) \to \operatorname{Ext}^n_U(U/I, U) \to \operatorname{Ext}^n_U(T/I, U) \to 0.$$

Since $r(\operatorname{Ext}(T/I, U)) = \tau_\lambda(P)$ by the case $s = 1$ and $r(\operatorname{Ext}^n(U/T, U)) = \tau_\lambda(T)$ by induction,

$$K := r(\operatorname{Ext}^n_U(U/I, U) \supseteq \tau_\lambda(P)\tau_\lambda(T) = \tau_\lambda(I). \tag{5}$$

However, by *(iii)*, and by *(i)* applied in $\operatorname{gr} U$,

$$\operatorname{gr} K \subseteq r(\operatorname{gr}\operatorname{Ext}^n_U(U/I, U)) = r(\operatorname{Ext}^n_{\operatorname{gr} U}(\operatorname{gr} U/\operatorname{gr} I, \operatorname{gr} U)) = \operatorname{gr} I. \tag{6}$$

Since $\operatorname{gr}\tau_\lambda(I) = \operatorname{gr} I$, (5) and (6) show that $K = \tau_\lambda(I)$. Now $\operatorname{Ext}^n_U(U/I, U)_{|U}$ contains the completely reducible submodule $\operatorname{Ext}^n_U(U/J, U)_{|U} := S$. Suppose there is another irreducible submodule V, with $V \oplus \operatorname{Ext}^n_U(U/J, U) \subseteq \operatorname{Ext}^n(U/I, U)$. Applying $\operatorname{Ext}^n_U(-, U)$ to this inclusion gives, in view of [2, p. 76],

$$U/I = \operatorname{Ext}^n_U(\operatorname{Ext}^n_U(U/I, U)) \to \operatorname{Ext}^n_U(V \oplus S, U) \to 0.$$

Since $\operatorname{Ext}^n_U(V \oplus S, U)$ is completely reducible of dimension $\dim_{\mathbb{C}}(U/J) + 1$, this is a contradiction to the definition of J. Hence $\operatorname{Ext}^n(U/I, U)_{|U}$ is a submodule of the right $U/\tau_\lambda(I)$-injective hull E of $U/\tau_\lambda(J)$.

By right and left handed use of *(i)* and *(ii)*, $\dim_{\mathbb{C}} E = \dim_{\mathbb{C}} U/\tau_\lambda(I) = \dim_{\mathbb{C}} U/I = \dim_{\mathbb{C}} \operatorname{Ext}^n_U(U/I, U)$. Therefore $E = \operatorname{Ext}^n_U(U/I, U)_{|U}$, as required.

7.4. Naturally, one conjectures that 7.3 *(iv)* is true for all co-Artinian ideals I. Let $K = r(\operatorname{Ext}^n(U/I, U))$. One can show-in fact, we've essentially done so above-that $\operatorname{gr} I = \operatorname{gr} K$ and $\operatorname{Ext}^n_U(U/I, U)_{|U}$ is the right U/K-injective hull of $U/\tau_\lambda(J)$. But the proof that $K = \tau_\lambda(I)$ has eluded us.

References

1. Bernstein, I.N., Gel'fand, S.I.: Tensor products of finite and infinite dimensional representations of semisimple Lie algebras. Comp. Math. **41**, 245–285 (1980)
2. Björk, J.-E.: Rings of Differential Operators. Amsterdam: North-Holland 1979
3. Borho, W., Kraft, H.: Über die Gelfand-Kirillov dimension. Math. Ann. **220**, 1–26 (1976)
4. Brown, K.A.: Module estensions over Noetherian rings. J. Algebra **69**, 247–260 (1981)
5. Brown, K.A.: Localisation, bimodules and injective modules for enveloping algebras of solvable Lie algebras. Bull. Sci. Math. **107**, 225–251 (1983)
6. Brown, K.A.: Ore sets in enveloping algebras. Comp. Math. **53**, 347–367 (1984)
7. Cartan, H., Eilenberg, S.: Homological Algebra. Princeton: Univ. Press 1956
8. Duflo, M.: Sur les idéaux induits dans les algèbres enveloppantes, Invent. Math. **67**, 385–393 (1982)
9. Gabber, O.: Equidimensionalité de la variété caracteristique. Unpublished manuscript by T. Levasseur
10. Grothendieck, A.: Sur quelques points d'algèbre homologique. Tôhoku Math. J. **9**, 119–221 (1957)
11. Jategaonkar, A.V.: Solvable Lie algebras, polycyclic-by-finite groups, and bimodule Krull dimension. Commun. Algebra **10**, 19–69 (1982)
12. Kaplansky, I.: Commutative Rings. Boston: Allyn and Bacon 1970
13. Lenagan, T.H.: Gelfand-Kirillov dimension and affine *PI*-rings. Commun Algebra **10**, 87–92 (1982)
14. Malliavin, M.-P.: Modules sans torsion et modules injectifs sur les algèbres de Lie résolubles. J. Algebra **83**, 126–156 (1983)
15. Musson, I.N.: Some examples of modules over Noetherian rings. Glasgow Math. J. **23**, 9–13 (1982)
16. Stenström, B.: Rings of Quotients. Berlin, Heidelberg, New York: Springer 1975
17. Tauvel, P.: Sur les quotients premiers de l'algèbre enveloppante d'une algèbre de Lie résoluble. Bull. Soc. Math. Fr. **106**, 177–205 (1978)

Received May 17, 1984

Math. Z. 189, 415–419 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Equipartitioning Common Domains of Non-Atomic Measures *

Theodore P. Hill

Department of Mathematics, University of Hawaii, Honolulu, HI 96822, USA

§ 1. Introduction

Let $h(a_1,\ldots,a_n)$ denote the *harmonic mean* $n(a_1^{-1}+\ldots+a_n^{-1})^{-1}$ of positive numbers $a_1,\ldots,a_n$, and extend this definition (by continuity) to the non-negative case by setting $h(a_1,\ldots,a_n)=0$ if $a_i=0$ for some $i\leqq n$. Recall that a measure μ is *non-atomic* if every measurable set A of strictly positive measure has a measurable subset $B\subseteq A$ satisfying $\mu(A)>\mu(B)>0$. A *measurable partition* of a measure space $(X,\mathscr{F})$ is a collection of disjoint subsets $A_1,\ldots,A_n$ of X satisfying $A_i\in\mathscr{F}$ for all i, and $\bigcup_{i=1}^{n} A_i=X$. Throughout this paper, "measure" means "countably additive, non-negative measure", and $\|\mu\|$ is the total mass (or total variation) of the measure μ; the main purpose is to prove the following theorem.

Theorem 1.1. *Let $\mu_1,\mu_2,\ldots,\mu_n$ be non-atomic finite measures on a measure space $(X,\mathscr{F})$. Then there is a measurable partition $A_1,\ldots,A_n$ of X satisfying*

$$\mu_i(A_i)\geqq n^{-1}h(\|\mu_1\|,\ldots,\|\mu_n\|) \quad \text{for all } i, \tag{1}$$

and this bound is best possible. Moreover, if $\|\mu_i\|>0$ for all i, then:

(2a) *if $\mu_i/\|\mu_i\|=\mu_j/\|\mu_j\|$ for all $i,j\leqq n$, then every partition satisfying* (1) *holds with equality for all i; and*

(2b) *otherwise, the $\{A_i\}$ may be chosen so that strict inequality holds in* (1) *for all i.*

Theorem 1.1 has an interesting "distribution of value" or "fair-division" interpretation. Suppose an object, such as a cake or piece of land, is to be divided among n people whose values of the object may differ. What is the largest c possible so that each person may receive a piece he values at least c, *according to his own values*? Theorem 1.1 essentially says that if the values vary continuously (e.g., pieces of zero volume are worth nothing), then c is at least $1/n$ times the harmonic mean of the total values assigned by each person, and that in general it is not possible to do better.

* Research partially supported by NSF Grant DMS-84-01604

Corollary 1.2 (Dubins and Spanier [2]). *Let $\mu_1, \ldots, \mu_n$ be non-atomic probability measures on a measure space $(X, \mathscr{F})$. Then there is a measurable partition $A_1, \ldots, A_n$ of X satisfying*

$$\mu_i(A_i) \geqq 1/n \quad \text{for all } i, \tag{3}$$

and if $\mu_i \neq \mu_j$ for some $i \neq j$, then strict inequality in (3) *for all i may be attained.*

(Actually, Dubins and Spanier showed more: for example, $\mu_i(A_i) \geqq p_i$ may be attained for *any* $p_i \geqq 0$, $\sum_{i=1}^{n} p_i = 1$.)

The question underlying (3) dates back to Steinhaus [7], and the reader is referred to [2] for a number of interesting related results many of which are based on Lyapounov's convexity theorem. A proof of Theorem 1.1 can also be given using the convexity theorem, but the argument below is based on a recent $L_1(0,1)$ inequality of Dor.

§ 2. Proof of Main Theorem

For measures $m_1, m_2, \ldots, m_n$ on $(X, \mathscr{F})$, let $\bigvee_{i=1}^{n} m_i$ denote the smallest measure which dominates each of the $\{m_i\}$, that is, for each $A \in \mathscr{F}$,

$$\left(\bigvee_{i=1}^{n} m_i\right)(A) = \sup\left\{\sum_{i=1}^{n} m_i(B_i) \colon B_1, \ldots, B_n \text{ is a measurable partition of } A\right\}.$$

The next proposition provides the key step in the proof; it is a standard generalization, via the non-atomic separable-measure-algebra isomorphism theorem (e.g. [4] Theorem C, p. 173), of a result of Dor ([1], p. 472) who stated it only for $L_1(0,1)$.

Proposition 2.1. *Let $\mu_1, \mu_2, \ldots, \mu_n$ be non-atomic finite measures on a measure space $(X, \mathscr{F})$, and let $c \geqq 0$. If*

$$\left\| \bigvee_{i=1}^{n} \alpha_i \mu_i \right\| \geqq c \sum_{i=1}^{n} \alpha_i \quad \text{for all } \alpha_1, \ldots, \alpha_n \geqq 0, \tag{4}$$

then there are disjoint measurable sets $A_1, \ldots, A_n \in \mathscr{F}$ satisfying

$$\mu_i(A_i) \geqq c \quad \text{for all } i \leqq n. \tag{5}$$

The proof of Theorem 1.1 will also use three short lemmas; the first lemma must be known, but as no reference is known to the author, a proof is given.

For each positive integer n, let $\prod_n$ denote the space of probabilities on n points, that is $\prod_n = \{\vec{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n \colon p_i \geqq 0$ for all i, and $\sum_{i=1}^{n} p_i = 1\}$.

Lemma 2.2. *If $a_1, \ldots, a_n$ are strictly positive, then*

$$\inf_{\vec{p}\in\prod_n} \max_{i\leqq n} \{p_i a_i\} = n^{-1} h(a_1, \ldots, a_n) = \sup_{\vec{p}\in\prod_n} \min_{i\leqq n} \{p_i a_i\}, \tag{6}$$

and both equalities are attained only for $\vec{p}=(p_1^*, \ldots, p_n^*)$ where $p_i^* = a_i^{-1} n^{-1} h(a_1, \ldots, a_n)$.

Proof. By the compactness of $\prod_n$ and the continuity in $\vec{p}$ of both $\min_{i\leqq n}\{p_i a_i\}$ and $\max_{i\leqq n}\{p_i a_i\}$, the inf and sup in (6) are both attained. An easy redistribution of mass argument then shows the inf and sup are each attained only when $p_1 a_1 = \ldots = p_n a_n$; substitution into the equation $\sum_{i=1}^{n} p_i = 1$ yields the desired conclusion. □

Lemma 2.3. *Let $\mu_1, \ldots, \mu_n$ be finite measures. Then*

$$\left\| \bigvee_{i=1}^{n} p_i \mu_i \right\| \geqq \max_{i\leqq n} \|p_i \mu_i\| \geqq n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$$
$$\text{for all } (p_1, \ldots, p_n) \in \prod_n. \tag{7}$$

Proof. The first inequality in (7) is elementary, and the second follows by Lemma 2.2 with $a_i = \|\mu_i\|$. □

Lemma 2.4. *Suppose $\|\mu_i\| > 0$ for all $i \leqq n$. If $\left\| \bigvee_{i=1}^{n} p_i \mu_i \right\| = n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$ for some $(p_1, \ldots, p_n) \in \prod_n$, then $p_i = \|\mu_i\|^{-1} n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$ for all $i \leqq n$, and $p_i \mu_i = p_j \mu_j$ for all $i, j \leqq n$.*

Proof. By (7), $\max_{i\leqq n} \|p_i \mu_i\| = n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$, and by Lemma 2.2 this implies $p_i = \|\mu_i\|^{-1} n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$ for all $i \leqq n$. Hence

$$\left\| \bigvee_{i=1}^{n} p_i \mu_i \right\| = \|p_1 \mu_1\| = \ldots = \|p_n \mu_n\|,$$

which in turn implies $p_i \mu_i = p_j \mu_j$ for all $i, j \leqq n$. □

Proof of Theorem 1.1. If $\|\mu_i\| = 0$ for some i, then $h(\|\mu_1\|, \ldots, \|\mu_n\|) = 0$, and (1) follows immediately; that this bound is best possible in this case is clear.

For the remainder of the proof, suppose $\|\mu_i\| > 0$ for all $i = 1, \ldots, n$. Lemma 2.3 implies that (4) is satisfied with $c = n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$, (the case $\alpha_1 = \ldots = \alpha_n = 0$ being trivial), and (1) follows from Proposition 2.1; that this bound is best possible will follow from (2a).

Suppose $\mu_i/\|\mu_i\| = \mu_j/\|\mu_j\|$ for all $i, j \leqq n$, and let $A_1, \ldots, A_n$ be a measurable partition of X satisfying (1). Then for all $i \leqq n$,

$$\begin{aligned}\mu_i(A_i) &= \sum_{j=1}^{n} \mu_i(A_j) - \sum_{j\neq i}^{n} \mu_i(A_j) \\ &= \|\mu_i\| - \sum_{j\neq i}^{n} \|\mu_i\| \, \|\mu_j\|^{-1} \mu_j(A_j) \\ &\leqq \|\mu_i\| (1 - n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|) \sum_{j\neq i}^{n} \|\mu_j\|^{-1}) \\ &= n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|),\end{aligned}$$

so by (1) equality holds throughout, which establishes (2a).

Finally, suppose $\mu_i/\|\mu_i\| \neq \mu_j/\|\mu_j\|$ for some $i \neq j$. By Lemmas 2.3 and 2.4, $\left\| \bigvee_{i=1}^{n} p_i \mu_i \right\| > n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$ for all $(p_1, \ldots, p_n) \in \prod_n$. The compactness of $\prod_n$ and continuity of $\left\| \bigvee_{i=1}^{n} p_i \mu_i \right\|$ in $\vec{p}$ then imply the existence of a $c > n^{-1} h(\|\mu_1\|, \ldots, \|\mu_n\|)$ satisfying $\left\| \bigvee_{i=1}^{n} p_i \mu_i \right\| > c$ for all $(p_1, \ldots, p_n) \in \prod_n$, which by Proposition 2.1 completes the proof of (2b). □

§ 3. Structured Domains

When $\mu_1, \ldots, \mu_n$ are all probability measures, applications of inequalities of the type in Corollary 1.2 have been made to statistical hypotheses testing (see for example [3, 6]). If the domain in question has a particular structure (e.g., as in [5]), then one may ask whether there is a measurable partition $\{A_i\}$ *with the same structure* satisfying (1) or (3). The following theorem is a typical result along these lines; its proof uses an idea in [5] and is given here primarily as an application of the strict inequality conclusion (2b) of Theorem 1.1.

Theorem 3.1. *Suppose $f_1, \ldots, f_n$ are continuous non-negative functions defined on a polygon $P \subset \mathbb{R}^2$. Then there is a polygonal decomposition $P_1, \ldots, P_n$ of P satisfying*

$$\int_{P_i} f_i \geqq [(\int_P f_1)^{-1} + \ldots + (\int_P f_n)^{-1}]^{-1} \quad \text{for all } i \leqq n, \tag{8}$$

and this bound is best possible.

Sketch of Proof. If $f_i \equiv 0$ for some i, any polygonal decomposition of P will suffice, so assume $\int_P f_i > 0$ for all $i \leqq n$. Suppose first that $f_i/\int_P f_i = f_j/\int_P f_j$ for all $i, j \leqq n$. Take as P_1 any polygon contained in P for which $P \setminus P_1$ is also a polygon and $\int_{P_1} f_1 = [(\int_P f_1)^{-1} + \ldots + (\int_P f_n)^{-1}]^{-1}$. Continue with P_2 in $P \setminus P_1$, and so on; since the functions are proportional, (8) is seen to be satisfied with equality throughout. On the other hand, if $f_i/\int_P f_i \neq f_j/\int_P f_j$ for some $i \neq j$, then use

Theorem 1.1 (2b) to infer the existence of a *Borel* partition $A_1, \ldots, A_n$ of P satisfying

$$\int_{A_i} f_i > [(\int_P f_1)^{-1} + \ldots + (\int_P f_n)^{-1}]^{-1} \quad \text{for all } i \leqq n.$$

Using this strict inequality, approximate the $\{A_i\}$ simultaneously by disjoint polygons $\{P_i\}$ in P in such a way as to retain the inequality. □

§ 4. Remarks

Clearly the conclusion of Theorem 1.1 may fail if the measures have atoms (the worst case being when each μ_i is the dirac delta measure on the same point $x \in X$), and as the next example shows, the conclusion (1) may also fail if the requirement of finiteness of the measures is dropped (where h is defined using $0^{-1} = \infty$ and $\infty^{-1} = 0$).

Example 4.1. Let μ_1 be Lebesgue measure on $[0, 1]$, and let μ_2 be given by $\mu_2(A) = \int_A f d\mu_1$ where $f(x) = 1/x$ for all $x \in (0, 1)$. Then $\|\mu_1\| = 1$, $\|\mu_2\| = \infty$, and $(\|\mu_1\|^{-1} + \|\mu_2\|^{-1})^{-1} = 1$, but for any measurable set A, either $\mu_1(A) < 1$ or $\mu_2(A^c) = 0$, so (1) fails.

If the requirement of non-negativity of the measures is dropped, it is also easy to see that inequality (1) may fail, whether $\|\mu\|$ is interpreted either as the total variation or as the total mass of μ.

Acknowledgements. The author is grateful to Stephen Demko for bringing the Dor result to his attention, to Carl Spruill for several conversations, and to the referee for several suggestions.

Bibliography

1. Dor, L.: On projections in L_1. Ann. Math. **102**, 463-474 (1975)
2. Dubins, L., Spanier, E.: How to cut a cake fairly. Am. Math. Mon. **68**, 1-17 (1961)
3. Feller, W.: Note on regions similar to the sample space. Statistical Research Memoirs, Univ. of London, 116-125 (1938)
4. Halmos, P.: Measure Theory. New York: Van Nostrand 1950
5. Hill, T.: Determining a fair border. Am. Math. Mon. **90**, 438-442 (1983)
6. Neyman, J., Pearson, E.: On the problem of the most efficient tests of statistical hypotheses. Philos. Trans. R. Soc. Lond., Ser. A **231**, 289-337 (1933)
7. Steinhaus, H.: Sur la division pragmatique. Econometrica (Supplement) **17**, 315-319 (1949)

Received April 13, 1984; in revised form August 27, 1984

Math. Z. 189, 421–447 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

Division et extension dans l'algèbre $A^\infty(\Omega)$ d'un ouvert pseudo-convexe à bord lisse de $\mathbb{C}^n$

Roger Gay et Ahmed Sebbar

Université de Bordeaux I, U.E.R. de mathématiques et informatique,
Laboratoire Associe au CNRS n° 226, 351, Cours de la Libération, F-33405 Talence, France

I. Introduction

Le but de ce travail est l'étude, dans le cadre des ouverts pseudo-convexes a bord lisse, bornés ou non bornés, des deux questions suivantes qui ont reçu, récemment, un regain d'attention ([1, 2, 4, 9, 16, 17, 20]) surtout dans le cas strictement pseudo-convexe.

Soit Ω un ouvert pseudo-convexe à bord lisse de classe C^∞ de $\mathbb{C}^n$ et Z un espace analytique défini par p fonctions holomorphes $h_1, \ldots, h_p$ au voisinage de $\bar{\Omega}$.

a) Toute fonction holomorphe f dans Ω et C^∞ dans $\bar{\Omega}$ (ie: $f \in A^\infty(\Omega)$) et nulle avec multiplicité sur $\bar{Z} \cap \Omega$ s'écrit-elle $f = \sum_{1 \leqq i \leqq p} \varphi_1 h_1$ avec $\varphi_1 \in A^\infty(\Omega)$ $(1 \leqq i \leqq p)$?

b) Toute fonction C^∞ dans $Z \cap \bar{\Omega}$ et "holomorphe" dans l'intérieur "relatif" de $Z \cap \bar{\Omega}$ dans Z est-elle la restriction d'une fonction de $A^\infty(\Omega)$ à $Z \cap \bar{\Omega}$?

Dans [2] a) et b) reçoivent une réponse affirmative pour Ω pseudo-convexe borné, $h_1, \ldots, h_p \in A^\infty(\Omega)$ et Z intersection complète globale de la forme $Z = \tilde{Z} \cap \Omega$, avec $\tilde{Z}$ variété C^∞ dans un voisinage de $\bar{\Omega}$ ayant un ordre de contact borné avec Ω (la preuve vaut, en fait dans le cas de $\bar{\tilde{Z}} \setminus \bar{\Omega}$ et $\bar{\Omega}$ régulièrement situés).

Dans [4] les auteurs répondent à a) et à b) dans le cas Ω strictement pseudo-convexe, Z variété au voisinage de $\partial\Omega$, Z et $\partial\Omega$ régulièrement situés et les fonctions $h_1, \ldots, h_p$ étant choisies, en fonctions de Z, de façon à engendrer, en chaque point z d'un voisinage Ω' de Ω, la fibre en z de l'idéal des fonctions holomorphes nulles sur Z. Lorsqu'on donne d'abord les fonctions $h_1, \ldots, h_p$ et que Z est l'espace analytique défini par $h_1 = \ldots = h_p = 0$ le résultat de factorisation d'une fonction nulle sur Z (sans multiplicité) peut ne pas être exact sous l'hypothèse que Z est une variété au voisinage de $\partial\Omega$. Voici un exemple:

On prend $n = 3$, Ω la boule unité de $\mathbb{C}^3$ et $h_1(x_1, x_2, x_3) = x_1 x_3$, $h_2(x_1, x_2, x_3) = x_1(x_1 - \varepsilon)^2$ $(0 < \varepsilon < 1)$, $h_3(x_1, x_2, x_3) = x_1 - x_2$. L'ensemble Z est $Z = \{x_1 = x_2$

$=0\}\cup\{(\varepsilon,\varepsilon,0)\}$. La fonction $x_1(x_1-\varepsilon)$ s'annule sur Z et n'appartient pas à l'idéal (h_1,h_2,h_3). Cependent Z est une variété au voisinage de $\partial\Omega$.

Notons que des résultats de division ont été prouvés par Nagel dans [16, 17] et qu'on peut se ramener aux idées de cet auteur pour obtenir les résultats cités ci-dessus.

Dans [20] l'auteur étend certains des résultats de [4] au cas des ouverts à bord lisse C^∞, strictement pseudo-convexes, Ω, non nécessairement bornés en construisant une exhaustion $(\bar{\Omega}_j)_{j\geq 1}$ de $\bar{\Omega}$ par des ouverts Ω_j bornés strictement pseudo-convexes à bord lisse.

Notre étude est organisée en cinq parties:

La première partie (§ II) est consacrée à l'exhaustion de l'adhérence d'un ouvert pseudo-convexe à bord lisse de classe C^∞ par les adhérences d'ouverts de ce type mais bornés.

La seconde partie (§ III) montre un résultat de densité de $A^\infty(\Omega)$ dans $A^\infty(\Omega_1)$ lorsque Ω et Ω_1 sont pseudo-convexes bornés à bord lisse de classe C^∞ et Ω_1 obtenu à partir de Ω par le procédé décrit dans la première partie.

Dans le cas où l'ouvert Ω est non borné la troisième partie (§ IV) est consacrée à prouver l'exactitude du complexe de Dolbeault pour les formes C^∞ jusqu'au bord.

Dans la quatrième partie (§ V) nous répondons par l'affirmative à a) et b) dans le cas des ouverts pseudo-convexes à bord lisse C^∞ avec Z espace analytique au voisinage de $\bar{\Omega}$. Pour avoir a) une condition de fermeture, portant sur l'idéal $(h_1,\ldots,h_p)A^\infty(\Omega)$, est faite comme dans [17] dont on retrouve les résultats en les étendant au cas pseudo-convexe; pour b) on suppose $\overline{Z\setminus\bar{\Omega}}$ et $\bar{\Omega}$ régulièrement situés.

Enfin, dans la cinquième partie (§ VI), nous étudions a) et b) dans le cas d'un ouvert Ω non borné et de Z intersection complète globale de la forme $Z=\tilde{Z}\cap\Omega$ avec $\tilde{Z}$ variété C^∞ dans un voisinage de $\partial\Omega$ ce qui prolonge à ce cas les résultats de [2].

Nous sommes heureux de remercier E. Amar, J.E. Fornaess et C.O. Kiselmann pour les utiles conversations que nous avons eu avec eux.

Notations. Si Ω est un ouvert à bord lisse de classe C^∞ dans C^n nous désignons par:

$\mathscr{E}_\Omega$ le faisceau sur $\bar{\Omega}$ des germes de fonctions C^∞ jusqu'au bord de $\bar{\Omega}$.

A^∞ le faisceau sur $\bar{\Omega}$ des germes de fonctions holomorphes dans Ω et C^∞ jusqu'au bord de $\bar{\Omega}$.

On désigne par $\mathscr{E}$ (resp: $\mathscr{O}$) le faisceau de germes de fonctions C^∞ (resp: holomorphes) dans $\mathbb{C}^n$ ou, plus généralement, sur une variété, enveloppe d'holomorphie d'un ouvert U de $\mathbb{C}^n$, et par $\mathscr{P}_\Omega$ le sous-faisceau de $\mathscr{E}_\Omega$ des germes de fonctions C^∞ infiniment plates sur $\bar{\Omega}$.

Si $\mathscr{F}$ est un faisceau de germes de fonctions ou distributions, on désigne par $\mathscr{F}^{p,q}$ le faisceau des germes de (p,q)-formes à coefficients dans $\mathscr{F}$.

L'espace des sections de $\mathscr{E}_\Omega$ sur $\bar{\Omega}$ sera noté $\mathscr{E}_\Omega(\bar{\Omega})$ ou $\Gamma(\bar{\Omega},\mathscr{E}_\Omega)$ et de même pour les autres faisceaux.

On désigne par $BO(\ ,\ )$ (resp. $BF(\ ,\ )$) une boule ouverte (resp. fermée).

II. Exhaustion des ouverts pseudo-convexes à bord lisse de classe C^∞ non bornés

2.1 Lemme. *Soit Ω un ouvert pseudo-convexe à bord lisse défini par $\hat\rho: \mathbb{C}^n \to \mathbb{R}$ de classe C^∞ ($\Omega = \{z \in \mathbb{C}^n : \hat\rho(z) < 0\}$, $d\hat\rho(z) \neq 0$ si $\hat\rho(z) = 0$ ($\hat\rho(z) = -d(z, \partial\Omega)$ dans un voisinage de $\partial\Omega$). Pour tout $R > 1$ il existe $L > 0$, un entier $m \geqq 1$ tel que*

$$-(-\hat\rho e^{-L|z|^2})^{1/2m}$$

soit strictement plurisousharmonique dans $U = U_1 \cap BO(0, R+r) \cap \Omega$ où U_1 est un voisinage ouvert de $\partial\Omega$ et $r > 0$.

Preuve. On reprend les calculs de [8]. Soient

$$\hat\rho_1 = -(-\hat\rho e^{-L|z|^2})^{1/2m}$$

et $\varphi(z) = |z|^2$. On a

$$\mathscr{L}(\hat\rho_1, z, t) = \frac{1}{2m}(-\hat\rho)^{1/2m-2} \exp\left(\frac{L}{2m}\varphi(z)\right)\{A + B + C\}$$

où

$$A = L \cdot (\hat\rho)^2 \{\mathscr{L}(\varphi, z, t) - \frac{L}{2m}|\langle \partial\varphi, t\rangle|^2\}$$

$$B = (-\hat\rho)\{\mathscr{L}(\hat\rho, z, t) - \frac{L}{m}\operatorname{Re}\langle \partial\hat\rho, t\rangle \overline{\langle \partial\varphi, t\rangle}\}$$

$$C = \left(1 - \frac{1}{2m}\right)|\langle \partial\hat\rho, t\rangle|^2$$

et $\mathscr{L}(f, ., .)$ désigne la forme de Lévi d'une fonction f.

On minore A, B, C comme suit:

a) Puisque $\mathscr{L}(\varphi, z, t) = |t|^2$ on a $A \geqq L(\hat\rho)^2\left(1 - \frac{L}{2m}|\partial\varphi|^2\right)|t|^2$ et $\frac{L}{2m}|\partial\varphi(z)|^2 \leqq \frac{L}{2m}\sup_{|z| \leqq R+1}|\partial\varphi|^2 \leqq \frac{1}{2}$ donc $A \geqq \frac{L(\hat\rho)^2}{2}|t|^2$ si $m \geqq m_0(L)$.

b) si $m \geqq 1$ $\left(1 - \frac{1}{2m}\right)|\langle \partial\hat\rho, t\rangle|^2 \geqq \frac{1}{2}|\langle \partial\hat\rho, t\rangle|^2$.

c) $\mathscr{L}(\hat\rho, z, t) - \frac{L}{m}\operatorname{Re}\langle \partial\hat\rho, t\rangle\overline{\langle \partial\varphi, t\rangle} \geqq \mathscr{L}(\hat\rho, z, t) - \frac{L}{m}|\langle \partial\hat\rho, t\rangle|\,|\partial\varphi|\,|t|$. On peut supposer qu'il existe $C = C(R, r) > 0$ tel que, pour $z \in U = U_1 \cap BO(0, R+1) \cap \Omega$

$$\mathscr{L}(\hat\rho, z, t) \geqq -C|\langle \partial\hat\rho, t\rangle|\,|t|$$

où U_1 est un voisinage de $\partial\Omega$ ([8]). Alors

$$\mathscr{L}(\hat\rho, z, t) - \frac{L}{m}\operatorname{Re}\langle \partial\hat\rho, t\rangle\overline{\langle \partial\varphi, t\rangle} \geqq -C_2|\langle \partial\hat\rho, t\rangle|\,|t|$$

avec

$$C_2 = C + \frac{L}{m} \sup_{|z| \leqq R+r} |\partial\varphi(z)|.$$

Ainsi, si $m \geqq m_0(L)$ et $m \geqq 1$ on a, dans U

$$A+B+C \geqq \tfrac{1}{2} L(\hat{\rho})^2 |t|^2 + C_2 \hat{\rho} |\langle \partial\hat{\rho}, t\rangle|\, |t| + \tfrac{1}{2} |\langle \partial\hat{\rho}, t\rangle|^2$$

$$A+B+C \geqq \tfrac{1}{2}\{L(\hat{\rho})^2 |t|^2 + 2C_2 \hat{\rho} |\langle \partial\hat{\rho}, t\rangle|\, |t| + |\langle \partial\hat{\rho}, t\rangle|^2\}$$

$$A+B+C \geqq \frac{L}{2}\left(\hat{\rho}|t| + \frac{C_2}{L}|\langle \partial\hat{\rho}, t\rangle|\right)^2 + \frac{1}{2}\left(1 - \frac{C_2^2}{L}\right)|\langle \partial\hat{\rho}, t\rangle|^2.$$

On fixe alors L pour que $\dfrac{C^2}{L} < \dfrac{1}{6}$ puis $m \geqq 1$ et $m \geqq m_0(L)$ pour que:

$$\frac{L}{m}\left(\sup_{|z| \leqq R+1} |\partial\varphi(z)|\right)^2 + \frac{2C}{m} \sup_{|z| \leqq R+1} |\partial\varphi(z)| < \tfrac{2}{6}.$$

Alors $A+B+C > 0$ dans U car si

$$\frac{L}{2}\left(\hat{\rho}|t| + \frac{C_2}{L}|\langle \partial\hat{\rho}, t\rangle|\right)^2 + \frac{1}{2}\left(1 - \frac{C_2^2}{L}\right)|\langle \partial\hat{\rho}, t\rangle|^2 = 0$$

en un $z \in U$ pour $t \neq 0$ on aurait $\hat{\rho}(z) = 0$ ce qui est impossible.

Soient Ω et $\hat{\rho}$ comme dans le lemme 2.1. On suppose que $0 \notin \bar{\Omega}$. Soit $r: \partial\Omega \to]0, \infty[$ la fonction $z \to r(z) = |z|$. Soient $]a, b[$ un intervalle contenu dans $r(\partial\Omega)$, $L > 0$ et m untier $\geqq 1$ tel que $-(-\hat{\rho} e^{-L|z|^2})^{1/2m}$ soit strictement plurisousharmonique dans $U = U_1 \cap BO(0, b) \cap \Omega$ comme dans 2.1. Soit $\rho = \hat{\rho} e^{-L|z|^2}$ de telle sorte que $-(-\rho)^{1/2m} = -(-\hat{\rho} e^{-L|z|^2})^{1/2m}$. On peut supposer que U est tel qu'il existe $\alpha_0 > 0$ pour lequel $z \in \Omega \setminus U$ si et seulement si $-(-\rho)^{1/2m} \leqq -2\alpha_0$. Soit $\chi: \mathbb{R} \to \mathbb{R}$ une fonction convexe croissante de classe C^∞ telle que $\chi(s) = -\beta_0$ (avec $\beta_0 > \alpha_0$) pour $s \leqq -\frac{3}{2}\alpha_0$ et $\chi(s) = s$ pour $s \geqq -\alpha_0$. Soit

$$\tilde{\rho}(z) = -[\chi(-(-\rho)^{1/2m})]^{2m}.$$

Si $-(-\rho)^{1/2m} \geqq -\alpha_0$ on a $\tilde{\rho} = \rho$ ce qui assure que $\tilde{\rho}$ est de classe C^∞. Par ailleurs, pour $z \in \Omega$, on a

$$[-\tilde{\rho}(z)]^{1/2m} = -\chi(-(-\rho(z))^{1/2m})$$

car $\chi(-(-\rho(z))^{1/2m}) < 0$. Ainsi:

$$-[-\tilde{\rho}(z)]^{1/2m} = \chi(-(-\rho(z))^{1/2m})$$

est plurisousharmonique dans U. (Elle est constante si $-(-\rho)^{1/2m} \leqq \frac{3}{2}\alpha_0$ et psh dans $-(-\rho)^{1/2m} > -2\alpha_0$.

Le théorème de Sard assure l'existence de $r_1 \in]a, b[$ tel que la sphère de centre 0 et de rayon r_1 et $\partial\Omega$ soient transverses aux points de leur intersection. En particulier $d\rho(z)$ et z ne peuvent être colinéaires en ces points donc aussi dans un voisinage V.

Soit φ une fonction convexe croissante de classe C^∞ nulle pour $x \leqq r_1^2$ et >0 pour $x > r_1^2$. On pose:

$$F_{r_1, K}(z) = -(-\hat{\rho}(z))^{1/2m} + K\varphi(|z|^2)$$

($K>0$). La fonction $F_{r_1, K}$ est plurisousharmonique dans $BO(0, b) \cap \Omega$ et l'ouvert

$$\omega_{r_1, K} = \{z \in \Omega \cap BO(0, b) : F_{r_1, K}(z) < 0\}$$

est pseudo-convexe; il est aussi à bord lisse (et d'adhérence contenue dans $\bar{\Omega} \cap BF(0, \lambda)$ avec $r_1 < \lambda < b$ si K est assez grand). En effet $\partial\omega_{r_1, K}$ est aussi l'ensemble

$$\{z \in \Omega \cap BO(0, b) : \tilde{\rho}(z) + K^{2m}\varphi^{2m}(|z|^2) = 0\}$$

car $-(-\tilde{\rho}(z))^{1/2m} + K\varphi(|z|^2) = 0$ si et seulement si $\tilde{\rho}(z) + K^{2m}\varphi^{2m}(|z|^2) = 0$. De plus le système d'équation

$$\tilde{\rho}(z) + K^{2m}\varphi^{2m}(|z|^2) = 0$$

$$\frac{\partial \tilde{\rho}}{\partial z_j}(z) + K^{2m}\varphi^{2m-1}(|z|^2)\,\varphi'(|z|^2)\,\bar{z}_j = 0 \qquad 1 \leqq j \leqq n$$

ne peut avoir de solution dans $U \cap V$ car $\tilde{\rho}(z) = \rho(z)$, et, si K est assez grand, on a $z \in V$ pour $z \in U$ tel que $\rho(z) + K^{2m}\varphi^{2m}(|z|^2) = 0$ ce qui interdit à $\operatorname{grad}\rho(z)$ et $\bar{z}$ d'être colinéaires.

Ce système ne peut avoir non plus de solutions si $-(-\rho(z))^{1/2m} \leqq -2\alpha_0$ car alors $\dfrac{\partial \tilde{\rho}}{\partial z_j}(z) = 0$ (puisque $\tilde{\rho}$ est constante) ce qui implique $\varphi^{2m}(|z|^2) \neq 0$ et $\varphi'(|z|^2) = 0$ ce qui est impossible.

Ainsi $\omega_{r, K}$ est un ouvert pseudo-convexe à bord lisse, contenu dans $\Omega \cap BO(0, \lambda)$ pour $r_1 < \lambda < b$ pour K assez grand et λ assez proche de r_1.

On a ainsi prouvé le

2.1. Théorème. *Soit Ω un ouvert borné ou non borné, pseudo-convexe à bord lisse, défini par $\tilde{\rho}: \mathbb{C}^n \to \mathbb{R}$ de classe C^∞: $\Omega = \{z \in \mathbb{C}^n : \hat{\rho}(z) < 0\}$; $d\hat{\rho}(z) \neq 0$ si $\hat{\rho}(z) = 0$; $\hat{\rho}(z) = \pm d(z, \partial\Omega)$ dans un voisinage de $\partial\Omega$. On suppose que $0 \notin \bar{\Omega}$. Soit $r: \partial\Omega \to]0, \infty[$ la fonction $r(z) = |z|$. Pour tout intervalle $]a, b[$ contenu dans $r(\partial\Omega)$ il existe $r_1, r_2, \lambda, K_0 (r_1 < r_2 < \lambda, K_0 > 0)$ tel que, pour $r \in [r_1, r_2[$ et $K \geqq K_0$ l'ouvert $\omega_{r, K}$ soit un ouvert pseudo-convexe à bord lisse.*

2.3 Remarque. Dans une première version nous avions indiqué une autre manière de procéder pour obtenir cette exhaustion. Nous la donnons rapidement, (bien qu'elle exige que $\hat{\rho}$ soit psh dans Ω) pour les services qu'elle peut rendre, soit dans le cas d'une exhaustion d'un ouvert convexe, soit d'un ouvert strictement pseudo-convexe (cf. [20] dans ce dernier cas).

Pour $0 < \delta$ soit ψ_δ une fonction C^∞ paire, convexe égale à x pour $x \geqq \delta$. Pour $(x, y) \in \mathbb{R}^2$ on pose $h_\delta(x, y) = \Psi_\delta\left(\dfrac{x-y}{2}\right) + \dfrac{x+y}{2}$. Soit $\rho_1(z) = |z|^2 - r_1^2$, on pose $H_\delta(z) = h_\delta(\hat{\rho}(z), \rho_1(z))$ et

$$\Omega_\delta = \{z \in \Omega : H_\delta(z) < 0\}.$$

Si $\hat{\rho}$ est psh dans Ω alors Ω_δ est pseudo-convexe et à bord lisse si δ est assez petit.

2.4 Remarque. En corollaire on obtient aussi le résultat de [2] sur l'existence d'ouverts admissibles (cf. 5 pour leur définition) arbitrairement petits.

Une suite croissante $(K_j)_{j\geq 1}$ de compacts de $\bar{\Omega}$ est dite une suite exaustive (cf. [20]) si les conditions suivantes sont remplies:

(1) $\bar{\Omega}=\bigcup_{j\geq 1} K_j$

(2) $\partial K_j \cap \partial K_k \subset \partial\Omega$ pour $j\neq k\ j,k\geq 1$

(3) Pour tout $j\geq 1$ l'intérieur de $K_j\cap\partial\Omega$ dans $\partial\Omega$ est contenu dans $K_{j+1}\cap\partial\Omega$.

(4) Pour tout $j\geq 1$ l'intérieur de K_{j+1} dans $\bar{\Omega}$ contient K_j.

Le théorème 2.2 nous permet d'énoncer

2.4 Proposition. *Soit Ω un ouvert pseudo-convexe de $\mathbb{C}^n$ non borné à bord lisse de classe C^∞.*

Il existe une suite d'ouverts pseudo-convexes bornés à bord lisse de classe C^∞: $(\Omega_j)_{j\geq 1}$ tels que, si $K_j=\bar{\Omega}_j$, la suite $(K_j)_{j\geq 1}$ soit une suite exhaustive de compacts pour $\bar{\Omega}$.

Preuve. L'ensemble $r(\partial\Omega)$ est un ensemble non borné contenant des points intérieurs aussi grands qu'on le désire.

III. Un résultat de densité

Nous nous donnons, dans cette partie, un ouvert Ω non borné à bord lisse de classe C^∞ défini par ρ, C^∞ dans $\mathbb{C}^n$ et nous allons montrer que, pour un ouvert borné Ω_1 construit par le procédé de 2.2, le sous-espace de $A^\infty(\Omega_1)$ des restrictions à Ω_1 des fonctions de $A^\infty(\Omega)$ est dense dans $A^\infty(\Omega_1)$. Nous pouvons considérer que Ω_1 est le premier élément d'une suite $(\Omega_j)_{j\geq 1}$ d'ouverts pseudo-convexes tels que $(\bar{\Omega}_j)_{j\geq 1}$ soit une suite exhaustive de compacts pour $\bar{\Omega}$. Un procédé classique à la Mittag-Leffler nous permet de nous ramener à prouver que $A^\infty(\Omega_k)$ est dense dans $A^\infty(\Omega_1)$.

Si Ω_0 est un ouvert pseudo-convexe à bord lisse de classe C^∞ borné et $P^0(\Omega)$ (resp. $P^\infty(\Omega)$) est la classe des fonctions psh dans Ω et continue dans $\bar{\Omega}$ (resp. psh dans Ω et C^∞ dans $\bar{\Omega}$) on sait d'après [7] et [11], l'égalité des trois enveloppes $K, \hat{K}_{p^0}, \hat{K}_{p^\infty}$ pour un compact K de $\bar{\Omega}$

$$\hat{K}=\{z\in\bar{\Omega}: |f(z)|\leq \sup_{\zeta\in K}|f(\zeta)|, f\in A^\infty(\Omega)\}$$

$$\hat{K}_{p^0}=\{z\in\bar{\Omega}: \varphi(z)\leq \sup_{\zeta\in K}\varphi(\zeta), \varphi\in P^0(\Omega)\}$$

$$\hat{K}_{p^\infty}=\{z\in\bar{\Omega}: \varphi(z)\leq \sup_{\zeta\in K}\varphi(\zeta), \varphi\in P^\infty(\Omega)\}.$$

Pour un ouvert U de $\mathbb{C}^n$ et une forme ω de bidegré (p,q) $\omega=\sum_{I,J}\omega_{I,J}dz_I\wedge d\bar{z}_J$ on introduit les normes de Sobolev $\|\omega\|_{m,U}$ $(m\geq 0)$

$$\|\omega\|^2_{m,U}=\sum_{I,J}\sum_{|\alpha|\leq m}\int_U |D^\alpha\omega_{I,J}(z)|^2\, d\lambda(z)$$

où λ est la mesure de Lebesgue dans $\mathbb{C}^n$ et

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\partial y_1^{\alpha_2}\dots\partial x_n^{\alpha_{2n-1}}\partial y_n^{\alpha_{2n}}} \qquad z=(z_1,\dots,z_n)\ z_j=x_j+iy_j$$

$\alpha=(\alpha_1,\dots,\alpha_{2n})\in\mathbb{N}^{2n}$. On note $H_m^{pq}(U)$ les espaces de Sobolev correspondants. On a alors le résultat suivant, dû a Catlin [7]:

3.1 Théorème (Catlin [7]). Soit Ω un ouvert borné pseudo-convexe à bord lisse de classe C^∞ de $\mathbb{C}^n$. On suppose que K est un compact de $\bar\Omega$ tel que $K=\hat K_{p^\infty}$. Soit G un voisinage de K pour la topologie de $\bar\Omega$. Soit f holomorphe dans $\mathring{G}\cap\Omega$ ($\mathring{G}$ est l'intérieur de G dans $\mathbb{C}^n$) telle que $\|f\|_{m,G}<+\infty$. Alors il existe une suite de fonctions $f_k\in A^\infty(\Omega)$ et un ouvert G' tel que $K\Subset G'\Subset G$ et $\lim_{k\to\infty}\|f_k-f\|_{m,G'}=0$.

Nous démontrons d'abord le

3.2 Théorème. *Soient Ω un ouvert pseudo-convexe borné à bord lisse de classe C^∞ et Ω_1 un ouvert de Ω construit par le procédé de 2.2 (avec $r_1<\sup_{\zeta\in\Omega} r(\zeta)$ et $0\notin\bar\Omega$). Le sous-espace de $A^\infty(\Omega_1)$ des restrictions à Ω_1 des fonctions de $A^\infty(\Omega)$ est dense dans $A^\infty(\Omega_1)$, la topologie de $A^\infty(\Omega_1)$ étant définie par les semi normes $p_q(f)=\sup_{\substack{\zeta\in\bar\Omega_1\\|\alpha|\leqq q}}|D_z^\alpha(f)(\zeta)|$ pour lesquelles $A^\infty(\Omega_1)$ est une algèbre de Fréchet.*

Preuve. L'ensemble compact $\bar\Omega_1=\bar\omega_{r_1,K}=\{z\in\bar\Omega: F_{r_1,K}\leqq 0\}$ est $A^\infty(\Omega)$-convexe puisque $F_{r_1,K}$ est psh dans Ω et continue dans $\bar\Omega$.

Soient $s\geqq 1$ et $\varepsilon>0$ fixés. Soit $f\in A^\infty(\Omega_1)$ prolongée dans $\mathbb{C}^n$ en $\tilde f\in\mathscr{E}(\mathbb{C}^n)$ grace au théorème d'extension de Whitney.

On peut trouver $\eta>0$ tel que, si r_1, K sont les nombres à partir desquels on a construit Ω_1, et si $r_1\leqq r_1'\leqq r_1+\eta<\sup_{\zeta\in\Omega} r(\zeta)$ alors l'ouvert Ω_1' construit par le procédé 2.2 avec r_1' et K soit tel que $\bar\Omega_1$ soit un compact de l'intérieur de $\bar\Omega_1'$ dans Ω.

D'après Kohn [13], pour tout $m\geqq 0$, il existe une solution U_m dans $H_m(\Omega_1')$ de l'équation $\bar\partial U=\bar\partial\tilde f$ telle que

$$\|U_m\|_{s,\Omega_1'}\leqq C(m,r_1')\,\|\bar\partial\tilde f\|_{s,\Omega_1'} \qquad (s\leqq m) \tag{$*$}$$

où la constante $C(m,r_1')$ reste bornée par $M_m>0$ au voisinage de r_1 pour chaque m fixé.

A partir de ces solutions Kohn [14], selon une idée de Hormander, construit une solution v de classe C^∞ de $\bar\partial v=\bar\partial\tilde f$ dans Ω_1' de la façon suivante: On choisit, par récurrence, pour chaque $j\geqq 1$, une solution $U_j\in H_j(\Omega_1')$ telle que

$$\|U_{j+1}-U_j\|_{j,\Omega_1'}<\frac{1}{2^j}$$

et on prend $v=U_k+(U_{k+1}-U_k)+\dots$ qui est une solution C^∞. Dans cette construction U_k est une modification d'une solution $\tilde U_k$ vérifiant ($*$) donc; à priori, U_k ne vérifie pas cette inégalité.

Mais, si l'on commence la construction à partir d'un indice m avec $U_m \in H_m(\Omega_1')$ vérifiant (*) puis en construisant, comme Hormander, des solutions $(U_j)_{j \geqq m+1}$ de telle sorte que $\|U_{j+1} - U_j\|_{j, \Omega_1'} \leqq \frac{1}{2^j}$ pour $j \geqq m+1$ la fonction

$$v = U_m + (U_{m+1} - U_m) + \ldots$$

est une solution C^∞ de $\bar{\partial}\tilde{v} = \partial f$ et

$$\|v\|_{m, \Omega_1'} \leqq \|U_m\|_{m, \Omega_1'} + \|U_{m+1} - U_m\|_{m, \Omega_1'} + \ldots + \|U_{j+1} - U_j\|_{m, \Omega_1'} + \ldots$$

donc

$$\|v\|_{m, \Omega_1'} \leqq \|U_m\|_{m, \Omega_1'} + \sum_{j=m}^{\infty} \frac{1}{2^j}$$

soit

$$\|v\|_{m, \Omega_1'} \leqq C(m, r_1') \|\bar{\partial}\tilde{f}\|_{m, \Omega_1'} + \frac{1}{2^{m-1}}.$$

On va donc choisir $m \geqq s$ tel que $\frac{1}{2^{m-1}} \leqq \frac{\varepsilon}{4}$ puis choisir $r_1' > r_1$ assez proche de r_1 pour que $\|\bar{\partial}\tilde{f}\|_{m, \Omega_1'} < \frac{\varepsilon}{4 K_m}$.

En raisonnant comme ci-dessus dans le Ω_1' qu'on vient de fixer on trouve une solution de classe C^∞ v dans $\bar{\Omega}_1'$ telle que

$$\|v\|_{s, \Omega_1'} \leqq \|v\|_{m, \Omega_1'} \leqq \frac{\varepsilon}{2}.$$

Soit alors $h = \tilde{f} - v$. On a $\bar{\partial} h = 0$ dans Ω_1' donc $h \in A^\infty(\Omega_1')$ et, de plus:

$$(f-h)|_{\Omega_1} = (f - (\tilde{f} - v))|_{\Omega_1} = v|_{\Omega_1}.$$

Ainsi:

$$\|f - h\|_{s, \Omega_1} \leqq \frac{\varepsilon}{2}.$$

Posons $G = \bar{\Omega}_1'$. D'après le résultat de Catlin, puisque h est holomorphe dans $\mathring{G} \cap \Omega$ et $\|h\|_{s, G} < \infty$, il existe: un ouvert G' tel que $\bar{\Omega}_1 \subset G'$ et G' relativement compact dans l'intérieur dans $\bar{\Omega}$ de $\bar{\Omega}_1'$ et il existe une suite $f_k \in A^\infty(\Omega)$ telle que $\lim_{k \to \infty} \|f_k - h\|_{s, G'} = 0$.

On peut donc choisir $g \in A^\infty(\Omega)$ telle que $\|h - g\|_{s, \Omega_1} \leqq \|h - g\|_{s, G'} \leqq \frac{\varepsilon}{2}$. Dans ces conditions $\|f - g\|_{s, \Omega_1} \leqq \varepsilon$.

Puisque $\mathscr{E}(\bar{\Omega}_1) = \bigcap_{s \geqq 1} H_s(\Omega_1)$ (topologiquement) si $(f_j)_{j \geqq 1}$ est une suite de $A^\infty(\Omega)$ telle que $\|f - f_j\|_{j, \Omega_1} \leqq \frac{1}{j}$ on a $\lim_{j \to \infty} f_j = f$ dans chaque $H_s(\Omega_1)$ donc dans $A^\infty(\Omega_1)$. D'où le théorème.

3.3 Corollaire. *Soient Ω un ouvert pseudo-convexe, non borné à bord lisse de classe C^∞, et Ω_1 un ouvert de Ω construit par le procédé 2.2. Le sous-espace de $A^\infty(\Omega_1)$ constitué des restrictions à $\bar{\Omega}_1$ des fonctions de $A^\infty(\Omega)$ est dense dans $A^\infty(\Omega_1)$.*

Preuve. Soit $(\Omega_j)_{j\geq 1}$ une suite d'ouverts de Ω dont le premier élément est le Ω_1 de l'énoncé construit par 2.2 et tels qu'en posant $K_j=\bar{\Omega}_j$ la suite $(K_j)_{j\geq 1}$ soit une suite exhaustive de compacts pour $\bar{\Omega}$. La topologie de $A^\infty(\Omega)$ est décrite par la famille de semi-normes

$$p_{q,k}(f)=\sup_{\substack{|\alpha|\leq q\\ \zeta\in K_k}}|(D^\alpha f)(\zeta)|.$$

Soient q un entier ≥ 1 et $\varepsilon>0$ fixés. Soit $(q_k)_{k\geq 1}$ une suite strictement croissante d'entiers tels que $q_1\geq q$. Pour tout $k\geq 1$, on peut trouver $f_k\in A^\infty(\Omega_k)\cap\mathscr{E}(\mathbb{C}^n)$ telle que $p_{q_{k-1},k-1}(f_{k-1}-f_k)\leq\frac{\varepsilon}{2^{k+1}}$, par récurrence, en posant $f=f_1$ et en utilisant 3, 2 pour chaque couple Ω_k, Ω_{k-1} ainsi que le théorème de Withney pour $k\geq 2$.

La suite $(f_j)_{j\geq 1}$ est de Cauchy dans l'espace de Fréchet $A^\infty(\Omega)$. *En effet, pour* $j_2>j_1\geq k$

$$p_{q_k,k}(f_{j_2}-f_{j_1})\leq\sum_{j_1\leq l\leq j_2-1}p_{q_k,k}(f_{k+1}-f_l)\leq\sum_{l\geq j_1}\frac{\varepsilon}{2^{l+1}}\leq\frac{\varepsilon}{2^{j_1}}.$$

Soit g la limite de la suite f_j dans $A^\infty(\Omega)$.

Il existe N tel que $p_{q_1,1}(f_N-g)\leq\frac{\varepsilon}{2}$. Alors

$$p_{q,1}(f-g)\leq p_{q_1,1}(f-g)\leq p_{q_1,1}(f_1-f_2)+p_{q_2,2}(f_2-f_3)+\ldots+p_{q_{N-1},N-1}(f_{N-1}-f_N)+p_{q_1,1}(f_N-g).$$

Donc $p_{q,1}(f-g)\leq\varepsilon$. D'où le corollaire.

IV. Exactitude du complexe de Dolbeault

4.1 Théorème. *Soit Ω un ouvert pseudo-convexe non borné à bord lisse de classe C^∞. Pour tout p,q $(q\geq 1)$ et toute $\omega\in\mathscr{E}^{p,q}_\Omega(\Omega)$, $\bar{\partial}$-fermée, il existe $\alpha\in\mathscr{E}^{p,q-1}_\Omega(\bar{\Omega})$ telle que $\omega=\bar{\partial}\alpha$. En particulier le complexe:*

$$0\to A^\infty(\Omega)\to\mathscr{E}^{0,0}_\Omega(\bar{\Omega})\xrightarrow{\bar{\partial}}\mathscr{E}^{0,1}_\Omega(\bar{\Omega})\xrightarrow{\bar{\partial}}\ldots\xrightarrow{\bar{\partial}}\mathscr{E}^{0,n}_\Omega(\bar{\Omega})\to 0$$

est exact. Autrement dit $H^q(\bar{\Omega},A^\infty)=0$ pour $q\geq 1$.

Preuve. On peut supposer $p=0$. On introduit une suite d'ouverts $(\Omega_j)_{j\geq 1}$ construits par le procédé 2.2 tels qu'en posant $K_j=\bar{\Omega}_j$ la suite $(K_j)_{j\geq 1}$ soit une suite exhaustive de compacts pour $\bar{\Omega}$. Comme usuellement nous distinguons deux cas. Nous répétons les preuves, traditionnelles, pour la commodité du lecteur.

1) Le cas de $q=1$

D'après Kohn il existe $f_1\in\mathscr{E}(\mathbb{C}^n)$ telle que $\bar\partial f_1=\omega$ dans $\bar\Omega_1$, puis $f_2\in\mathscr{E}(\mathbb{C}^n)$ telle que $\bar\partial f_2=\omega$ dans $\bar\Omega_2$. Puisque $f_1-f_2\in A^\infty(\Omega_1)\cap\mathscr{E}(\mathbb{C}^n)$ on peut, grace à 3.3, modifier f_2 en lui ajoutant un élément de $A^\infty(\Omega)\cap\mathscr{E}(\mathbb{C}^n)$ de telle sorte que $p_{1,1}(f_1-f_2)\leqq 1$. (Avec les notations de 3.3). Par le même raisonnement on peut construire, par récurrence, une suite $(f_N)_{N\geqq 1}$, $f_N\in\mathscr{E}_\Omega(\bar\Omega)$ telle que $\bar\partial f_N=\omega$ dans $\bar\Omega_N$ et $P_{N,N}(f_N-f_{N+1})\leqq\frac{1}{2^N}$. Alors la suite $(f_N)_{N\geqq 1}$ est de Cauchy dans $\mathscr{E}_\Omega(\bar\Omega)$ et, si $f=\lim_{N\to\infty} f_N$, on a $\bar\partial f=\omega$.

2) Le cas de $q\geqq 2$

Toujours grace au théorème de Kohn, il existe $\alpha_1\in\mathscr{E}^{p,q-1}(\mathbb{C}^n)$ telle que $\bar\partial\alpha_1=\omega$ dans $\bar\Omega_1$ et $\vartheta\in\mathscr{E}^{p,q-1}(\mathbb{C}^n)$ telle que $\bar\partial\vartheta=\omega$ dans $\bar\Omega_2$. Puisque $\bar\partial(\vartheta-\alpha_1)=0$ dans $\bar\Omega_1$ il existe $\beta\in\mathscr{E}^{p,q-2}(\mathbb{C}^n)$ telle que $\alpha_1-\vartheta=\bar\partial\beta$ dans $\bar\Omega_1$. On pose $\alpha_2=\vartheta+\bar\partial\beta\in\mathscr{E}^{p,q-1}(\mathbb{C}^n)$ pour laquelle $\bar\partial\alpha_2=\omega$ dans $\bar\Omega_2$ et $\alpha_2=\alpha_1$ dans $\bar\Omega_1$. Par récurrence, on construit une suite $(\alpha_N)_{N\geqq 1}$ d'éléments de $\mathscr{E}^{p,q-1}(\mathbb{C}^n)$ telles que $\bar\partial\alpha_N=\omega$ dans $\bar\Omega_N$ et $\alpha_{N+1}-\alpha_N$ dans $\bar\Omega_N$. Il est alors clair que $\alpha=\lim_{N\to\infty}\alpha_N$ existe dans $\mathscr{E}^{p,q-1}_\Omega(\bar\Omega)$ et que $\bar\partial\alpha=\omega$. D'où le théorème.

Pour tout entiers $1\leqq p\leqq n$ et $1\leqq q\leqq n$ on a les suites exactes d'espaces Fréchet-Nucléaires (F, N) et d'applications linéaires continues.

$$0\to\Gamma(\mathbb{C}^n,\mathscr{D}^{p,q}_\Omega)\to\Gamma(\mathbb{C}^n,\mathscr{E}^{p,q})\to\Gamma(\mathbb{C}^n,\mathscr{E}^{p,q}_\Omega)\to 0$$

d'où l'identification du dual de l'espace F, N $\mathscr{E}^{p,q}_\Omega(\bar\Omega)=\Gamma(\mathbb{C}^n,\mathscr{E}^{p,q}_\Omega)$ avec $'\mathscr{D}^{n-p,n-q}_c(\bar\Omega)$ espace des $(n-p,\ n-q)$ courants à supports compacts dans $\bar\Omega$. En transposant la suite exacte d'espace FN et d'applications linéaires continues

$$0\to A^{\infty p,0}(\Omega)\to\mathscr{E}^{p,0}_\Omega(\bar\Omega)\xrightarrow{\bar\partial}\mathscr{E}^{p,1}_\Omega(\bar\Omega)\xrightarrow{\bar\partial}\dots\xrightarrow{\bar\partial}\mathscr{E}^{p,n}_\Omega(\bar\Omega)\to 0$$

on obtient

4.2 Corollaire. *Soit Ω un ouvert pseudo-convexe à bord lisse de classe C^∞. La suite d'espaces D.F.N et d'applications linéaires continues*

$$0\to'\mathscr{D}^{n-p,0}_c(\bar\Omega)\xrightarrow{\bar\partial}'\mathscr{D}^{n-p,1}_c(\bar\Omega)\xrightarrow{\bar\partial}\dots$$
$$\xrightarrow{\bar\partial}'\mathscr{D}^{n-p,n-1}_c(\bar\Omega)\xrightarrow{\bar\partial}'\mathscr{D}^{n-p,n}_c(\bar\Omega)\to(A^{\infty p,0}_\cdot(\Omega))'\to 0$$

ext exact. En particulier:

$$(A^\infty(\bar\Omega))'=\frac{'\mathscr{D}^{n,n}_c(\bar\Omega)}{\bar\partial\,'\mathscr{D}^{n,n-1}_c(\bar\Omega)}.$$

V. Division et extension. Case de Z espace analytique au voisinage de $\bar\Omega$

Nous renvoyons à [15, 21] pour les définitions et les principaux faits relatifs à la notion de platitude dans le contexte des idéaux de fonctions différentiables.

Dans [16, 17] et dans le cas d'un ouvert Ω de $\mathbb{C}^n$ borné à bord lisse de $\mathbb{C}^n$ vérifiant:

a) Pour tout $\zeta \in \partial\Omega$ la suite

$$0 \to A^\infty_\zeta \to \mathscr{E}^{0,0}_{\Omega,\zeta} \xrightarrow{\bar\partial} \mathscr{E}^{0,1}_{\Omega,\zeta} \xrightarrow{\bar\partial} \dots \xrightarrow{\bar\partial} \mathscr{E}^{0,n}_{\Omega,\zeta} \to 0$$

est exacte.

b) Il existe une suite de domaines d'holomorphie $\Omega_i \subset \mathbb{C}^n$ tel que $\bar\Omega = \bigcap_{i \geqq 1} \Omega_i$.

c) $H^q(\bar\Omega, A^\infty) = 0$ pour tout $q > 0$.

Nagel obtient les résultats suivants:

A) Pour chaque $\zeta \in \bar\Omega$ $\mathscr{E}_{\Omega,\zeta}$ et A^∞_ζ sont des $\mathscr{O}_\zeta$-modules (fidélement) plats.

B) Si $\mathscr{F}$ est un $\mathscr{O}$-module cohérent défini dans un voisinage U de $\bar\Omega$

$$H^q(\bar\Omega, \mathscr{F} \otimes_{\mathscr{O}} A^\infty) = 0 \qquad (q \geqq 1).$$

La démonstration de A) n'utilise que a) et les propriétés classiques des fonctions différentiables.

Selon [2], un ouvert U de $\bar\Omega$ est dit admissible si $U \cap \Omega$ est un ouvert pseudo-convexe à bord lisse C^∞ et si $\overline{U \cap \Omega}$ possède, dans $\bar\Omega$, une base de voisinages pseudo-convexes à bord lisse C^∞. D'après cet auteur, pour un ouvert pseudo-convexe à bord lisse C^∞, il existe des recouvrement arbitrairement fins de $\bar\Omega$ constitués d'ouverts admissibles. Ceci, joint au théorème de Kohn assure que a) est vrai et donc aussi A) pour les pseudo-convexes à bord lisse. Dans la preuve de B) l'hypothèse b) intervient seulement pour assurer que $\mathscr{F}$ admet une résolution libre de longueur $2n+1$ au moins dans un voisinage de $\bar\Omega$. Ainsi B) sera valable pour tout $\mathscr{O}$-module cohérent au voisinage de $\bar\Omega$ pour lequel on sera capable de construire une telle suite.

Nous allons étendre les résultats de division de Nagel [17] au cas pseudo-convexe borné à bord lisse et démontrer un théorème d'extension.

5.1 Proposition. *Soient Ω un ouvert pseudo-convexe borné à bord lisse de classe C^∞ et $h_1, \dots, h_p$ des fonctions holomorphes au voisinage U de $\bar\Omega$. Alors $h_1, \dots, h_p$ engendrent $\Gamma(\bar\Omega, I \otimes_{\mathscr{O}} A^\infty) = \Gamma(\bar\Omega, IA^\infty)$ comme $A^\infty(\Omega)$-module avec $I = (h_1, \dots, h_p)\mathscr{O}$. Autrement dit $\Gamma(\bar\Omega, IA^\infty) = \Gamma(\bar\Omega, I)A^\infty(\Omega)$.*

Preuve. On peut supposer Ω connexe. Soit h l'homomorphisme de $\mathscr{O}$-modules défini dans l'enveloppe d'holomorphie $\tilde\Omega$ de U supposé aussi connexe qui est alors une variété de Stein.

On peut trouver un voisinage U_1 de $\bar\Omega$ dans $\tilde\Omega$ et, dans ce voisinage, une résolution libre

$$\dots \to \mathscr{O}^{p_{2n+2}} \to \mathscr{O}^{p_{2n+1}} \to \dots \to \mathscr{O}^{p_2} \to \mathscr{O}^{p_1} \xrightarrow{\mu} \mathscr{O}^p \xrightarrow{h} \mathscr{O} \to 0$$

pour h. Le faisceau $G = \operatorname{Ker} h$ est cohérent et admet, dans U_1, une résolution libre de longueur $2n+1$ au moins.

On a la suite exacte de $\mathscr{O}$-modules cohérents dans U_1

$$0 \to \operatorname{Ker} h \to \mathscr{O}^p \xrightarrow{h} I \to 0.$$

En tensorisant par $\otimes_{\mathcal{O}} A^\infty$ il vient la suite exacte de faisceaux sur $\bar{\Omega}$

$$0 \to \operatorname{Ker} h \otimes_{\mathcal{O}} A^\infty \to (A^\infty)^p \xrightarrow{h} I \otimes_{\mathcal{O}} A^\infty \to 0$$

car A^∞ est $\mathcal{O}$-plat ce qui assure aussi $I \otimes_{\mathcal{O}} A^\infty = IA^\infty$.

L'extension à ce cas du résultat B) de Nagel donne la suite exacte

$$0 \to \Gamma(\bar{\Omega}, \operatorname{Ker} h \otimes_{\mathcal{O}} A^\infty) \to (A^\infty(\Omega))^p \xrightarrow{h} \Gamma(\bar{\Omega}, IA^\infty) \to 0$$

d'où la proposition.

5.2 Proposition. *Soient Ω un ouvert pseudo-convexe borné à bord lisse de classe C^∞, $h_1, \ldots, h_p$ des fonctions holomorphes au voisinage de $\bar{\Omega}$ et $I = (h_1, \ldots, h_p)\mathcal{O}$. Le complexe:*

$$0 \to \Gamma(\bar{\Omega}, IA^\infty) \to \Gamma(\bar{\Omega}, I\mathcal{E}^{0,0}_\Omega) \xrightarrow{\bar{\partial}} \Gamma(\bar{\Omega}, I\mathcal{E}^{0,1}_\Omega) \xrightarrow{\bar{\partial}} \ldots \xrightarrow{\bar{\partial}} \Gamma(\bar{\Omega}, I\mathcal{E}^{0,n}_\Omega) \to 0$$

est exact. En particulier:

$$\Gamma(\bar{\Omega}, IA^\infty) = \Gamma(\bar{\Omega}, I)\, A^\infty(\Omega) = \Gamma(\bar{\Omega}, I\mathcal{E}^{0,0}_\Omega) \cap A^\infty(\Omega).$$

Cette dernière égalité est, dans dans le cas considéré par Nagel, son théorème 3.2 ([17]).

Preuve. Pour tout $\zeta \in \bar{\Omega}$ la suite de $\mathcal{O}_\zeta$-modules

$$0 \to A^\infty_\zeta \to \mathcal{E}^{0,0}_{\Omega,\zeta} \xrightarrow{\bar{\partial}} \mathcal{E}^{0,1}_{\Omega,\zeta} \xrightarrow{\bar{\partial}} \ldots \xrightarrow{\bar{\partial}} \mathcal{E}^{0,n}_{\Omega,\zeta} \to 0$$

est exacte. On la découpe en courtes suites exactes

$$\begin{aligned}
&0 \to A^\infty_\zeta \to \mathcal{E}^{0,0}_{\Omega,\zeta} \xrightarrow{\bar{\partial}} K_{1,\zeta} \to 0\\
&0 \to K_{1,\zeta} \to \mathcal{E}^{0,1}_{\Omega,\zeta} \xrightarrow{\bar{\partial}} K_{2,\zeta} \to 0\\
&\cdots\\
&0 \to K_{n-1,\zeta} \to \mathcal{E}^{0,n-1}_{\Omega,\zeta} \xrightarrow{\bar{\partial}} \mathcal{E}^{0,n}_{\Omega,\zeta} \to 0
\end{aligned}$$

où $K_j = \operatorname{Ker}[\mathcal{E}^{0,j}_{\Omega,\zeta} \xrightarrow{\bar{\partial}} \mathcal{E}^{0,j+1}_{\Omega,\zeta}]$.

En tensorisant par $I_\zeta \otimes_{\mathcal{O}_\zeta}$ il vient les longues suites exactes:

$$\begin{aligned}
&\ldots \to \operatorname{Tor}_1^{\mathcal{O}_\zeta}(I_\zeta, K_{1,\zeta}) \to I_\zeta \otimes_{\mathcal{O}_\zeta} A^\infty_\zeta \to I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,0}_{\Omega,\zeta} \to I_\zeta \otimes_{\mathcal{O}_\zeta} K_{1,\zeta} \to 0\\
&\ldots \to \operatorname{Tor}_1^{\mathcal{O}_\zeta}(I_\zeta, K_{2,\zeta}) \to I_\zeta \otimes_{\mathcal{O}_\zeta} K_{1,\zeta} \to I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,1}_{\Omega,\zeta} \to I_\zeta \otimes_{\mathcal{O}_\zeta} K_{2,\zeta} \to 0\\
&\ldots \to \operatorname{Tor}_1^{\mathcal{O}_\zeta}(I_\zeta, \mathcal{E}^{0,n}_{\Omega,\zeta}) \to I_\zeta \otimes_{\mathcal{O}_\zeta} K_{n-1,\zeta} \to I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,n-1}_{\Omega,\zeta} \to I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,n}_{\Omega,\zeta} \to 0.
\end{aligned}$$

Mais les $\mathcal{O}_\zeta$-modules $K_{1,\zeta}, K_{2,\zeta}, \ldots, K_{n-1,\zeta}$ et $\mathcal{E}^{0,n}_{\Omega,\zeta}$ sont $\mathcal{O}_\zeta$-plats; par suite:

$$\operatorname{Tor}_1^{\mathcal{O}_\zeta}(I_\zeta, K_{1,\zeta}) = \ldots = \operatorname{Tor}_1^{\mathcal{O}_\zeta}(I_\zeta, K_{n-1,\zeta}) = \operatorname{Tor}_1^{\mathcal{O}_\zeta}(I_\zeta, \mathcal{E}^{0,n}_{\Omega,\zeta}) = 0$$

et l'on obtient la longue suite exacte de $\mathcal{O}_\zeta$-modules:

$$0 \to I \otimes_{\mathcal{O}_\zeta} A^\infty_\zeta \to I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,0}_{\Omega,\zeta} \xrightarrow{1\otimes\bar{\partial}} I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,1}_{\Omega,\zeta} \xrightarrow{1\otimes\bar{\partial}} \ldots \xrightarrow{1\otimes\bar{\partial}} I_\zeta \otimes_{\mathcal{O}_\zeta} \mathcal{E}^{0,n}_{\Omega,\zeta} \to 0$$

qui induit la résolution suivante de $I\otimes_{\mathcal{O}}A^\infty$ par des $\mathcal{O}$-modules:

$$0\to I\otimes_{\mathcal{O}}A^\infty\to I\otimes_{\mathcal{O}}\mathscr{E}^{0,0}_{\bar\Omega}\xrightarrow{1\otimes\bar\partial} I\otimes_{\mathcal{O}}\mathscr{E}^{0,1}_{\bar\Omega}\xrightarrow{1\otimes\bar\partial}\dots\xrightarrow{1\otimes\bar\partial} I\otimes_{\mathcal{O}}\mathscr{E}^{0,n}_{\bar\Omega}\to 0$$

qui est une résolution fine de $I\otimes_{\mathcal{O}}A^\infty$ sur $\bar\Omega$, ce qui assure que la suite de ses sections sur $\bar\Omega$ a pour cohomologie les $H^q(\bar\Omega, I\otimes_{\mathcal{O}}A^\infty)=0$.

D'où la proposition.

5.3 Remarque. En utilisant une partition de l'unité on voit aussi que

$$\Gamma(\bar\Omega, I\cdot\mathscr{E}^{p,q}_{\bar\Omega})=\Gamma(\bar\Omega, I)\cdot\Gamma(\bar\Omega,\mathscr{E}^{p,q}).$$

Dans [17] Nagel démontre le théorème suivant où Ω est un ouvert borné à bord lisse de classe C^∞ de $\mathbb{C}^n$ (théorème 4.2) et $h_1,\dots,h_p$ des fonctions holomorphes dans un voisinage U de $\bar\Omega$.

5.4 Théorème ([17]). *Si* $\Gamma(\bar\Omega, I)\cdot\Gamma(\bar\Omega,\mathscr{E}^{0,0}_{\bar\Omega})$ *est fermé dans* $\mathscr{E}^{0,0}_{\bar\Omega}(\bar\Omega)$ *une* $g\in A^\infty(\Omega)$ *s'écrit* $g=\sum_{1\le j\le p} g_j h_j$ *avec* $g_j\in A^\infty(\Omega)$ *si et seulement si, pour chaque* $\zeta\in\bar\Omega$, *la série formelle de Taylor* $T_\zeta(g)$ *de* g *en* ζ *appartient à l'idéal engendré par* $T_\zeta(h_1),\dots,T_\zeta(h_p)$ *dans l'anneau des séries formelles en* ζ.

Lorsque la condition sur les séries formelles qui apparait dans cet énoncé sera réalisée nous dirons que g s'annulle avec multiplicité sur $Z\cap\bar\Omega$ où $Z=\{\zeta\in U: h_1(\zeta)=\dots=h_p(\zeta)=0\}$.

Il revient au même *de dire que, pour tout* $\zeta\in Z\cap\bar\Omega$, et toute distribution $T=\sum_{\alpha\in\mathbb{N}^{2n}} a_\alpha\delta^\alpha_\zeta$ à support dans $\{\zeta\}$ telle que $\langle T,u\rangle=0$ pour tout $u\in I_\zeta\mathscr{E}^{0,0}_{\bar\Omega,\zeta}$ on a $\langle T,g\rangle=0$. (La somme est finie,

$$\delta^\alpha_\zeta=\frac{\partial^{\alpha_1+\dots+\alpha_{2n}}\delta_\zeta}{\partial^{\alpha_1}_{z_1}\dots\partial^{\alpha_n}_{z_n}\partial\bar z_1^{\alpha_{n+1}}\dots\partial\bar z_n^{\alpha_{2n}}}$$

où δ_ζ est la distribution de Dirac en ζ et les séries formelles sont en $(z_1-\zeta_1,\dots,z_n-\zeta_n,\bar z_1-\bar\zeta_1,\dots,\bar z_n-\bar\zeta_n)$).

Toujours dans [17] Nagel montre si Ω est strictement pseudo-convexe et l'hypothèse de fermeture de $\Gamma(\bar\Omega, I)\cdot\Gamma(\bar\Omega,\mathscr{E}^{00}_{\bar\Omega})$ est réalisée, alors le résultat précédent de division dans $A^\infty(\Omega)$ d'une fonction s'annulant avec multiplicité sur $Z\cap\bar\Omega$ est exact. Il donne ensuite le cas d'un ouvert strictement pseudo-convexe à frontière réelle analytique "par morceaux" comme exemple dans lequel l'hypothèse de fermeture est réalisée.

Ces résultats s'étendent maintenant au cas de Ω pseudo-convexe.

5.5 Proposition. *Soient* Ω *un ouvert pseudo-convexe à bord lisse de classe* C^∞ *borné et* $h_1,\dots,h_p$ *des fonctions holomorphes dans un voisinage* U *de* $\bar\Omega$. *Si* $\Gamma(\bar\Omega, I\mathscr{E}^{00}_{\bar\Omega})$ *est fermé dans* $\Gamma(\bar\Omega,\mathscr{E}^{00}_{\bar\Omega})$, *une* $g\in A^\infty(\Omega)$ *s'écrit* $g=\sum_{1\le j\le p} g_j h_j$ *avec* $g_j\in A^\infty(\Omega)$ *si et seulement si* g *s'annule avec multiplicité sur* $Z\cap\bar\Omega$.

Preuve. Si g s'annule avec multiplicité le théorème de synthèse spectrale de Whitney ([21]) montre que $g\in\Gamma(\bar\Omega, I)\cdot\Gamma(\bar\Omega,\mathscr{E}^{00}_{\bar\Omega})=\Gamma(\bar\Omega, I\mathscr{E}^{00}_{\bar\Omega})$. D'où la proposition d'après 5.2.

5.6 Remarque. On remarquera qu'il n'est pas vraiment nécessaire que $\Gamma(\bar{\Omega}, I\mathscr{O}_\Omega^{00})$ soit fermé pour utiliser 5.2 et conclure à la factorisation de g. Il suffit de savoir que $g \in \Gamma(\bar{\Omega}, I\mathscr{O}_\Omega^{00})$. On peut alors démontrer

5.7 Proposition. *Soient Ω un ouvert pseudo-convexe à bord lisse de classe C^∞, borné et $h_1, \ldots, h_p$ des fonctions holomorphes dans un voisinage U de $\bar{\Omega}$.* Soit $Z = \{z \in U : h_1(z) = \ldots = h_p(z) = 0\}$

1) *si $n \geqq 2$ et $n - p \geqq 1$ et $\partial h_1(z) \wedge \ldots \wedge \partial h_p(z) \neq 0$ pour $z \in \partial\Omega$ alors toute $f \in A^\infty(\Omega)$ nulle sur $Z \cap \bar{\Omega}$ s'écrit $f = \sum_{1 \leqq j \leqq p} g_j h_j$ avec $g_j \in A^\infty(\Omega)$ si $Z = \overline{Z \setminus \partial\Omega}$ et Z et $\partial\Omega$ sont régulièrement situés.*

2) *si Z est une variété au voisinage de $\partial\Omega$, $Z = \overline{Z \setminus \partial\Omega}$ et Z et $\partial\Omega$ sont régulièrement situés toute $f \in A^\infty(\Omega)$ nulle sur $Z \cap \bar{\Omega}$ avec multiplicité s'écrit $f = \sum_{1 \leqq j \leqq p} g_j h_j$ avec $g_j \in A^\infty(\Omega)$.*

Preuve. Considérons d'abord le premier point. D'après la remarque 5.6 il suffit de prouver que $f \in \Gamma(\bar{\Omega}, I\mathscr{O}_\Omega^{00})$. Cette question est locale et a une réponse positive, dans le voisinage d'un point intérieur, grace à l'énoncé suivant ([12]) que l'on peut soit démontrer comme corollaire du théorème d'extension des classes de cohomologie de G. Scheja [3], soit en utilisant les formules d'interpolations que l'on trouve dans [5] comme nous l'a indiqué B.A. Taylor [19].

5.8 Proposition. *Soient n et p deux entiers positifs tels que $n \geqq 2$ et $n - p \geqq 1$, Ω la boule unité de $\mathbb{C}^n$ et $h_1, \ldots, h_p$ des fonctions holomorphes dans Ω dont on désigne par X l'ensemble des zéros communs: $X = \{z \in \Omega : h_1(z) = \ldots = h_p(z) = 0\}$. On suppose qu'il existe $0 < \varepsilon < 1$ tel que $\partial h_1(z) \wedge \ldots \wedge \partial h_p(z) \neq 0$ si $\varepsilon < |z| < 1$, $z \in X$. Si f est une fonction holomorphe dans Ω et nulle sur X, elle s'écrit $f = \sum_{1 \leqq j \leqq p} \varphi_j h_j$ avec φ_j holomorphe dans Ω.*

En un point du bord il suffit, après un difféomorphisme local, de démontrer le lemme suivant:

5.9 Lemme. *Soient Ω la boule unité ouverte de $\mathbb{C}^n$, $\rho : \Omega \to \mathbb{R}$ une fonction de classe C^∞ telle que $\rho(0) = 0$ et $d\rho(z) \neq 0$ si $\rho(z) = 0$,*

$$U = \{z \in \Omega : \rho(z) < 0\}, \qquad Z = \{z \in \Omega : z_1 = z_2 = \ldots = z_p = 0\}.$$

On suppose que $\Omega \setminus \partial U$ a deux composantes connexes U et $\Omega \setminus \bar{U}$, que $Z = \overline{Z \setminus \partial U}$ et que Z et ∂U sont régulièrement situés. Alors si $I = (z_1, \ldots, z_p)\mathscr{O}$ et $g \in A^\infty(\bar{U})$ s'annule sur $Z \cap \bar{U}$ on a $g \in \Gamma(\bar{U}, I\mathscr{O}_U^{00})$. ($\bar{U}$ est l'adhérence de U dans Ω).

Preuve. On montre d'abord le lemme suivant dont l'énoncé nous a été proposé par E. Amar.

5.10 Lemme. *Soient X de Y deux fermés non vides de W ouvert de $\mathbb{R}^n$ tels que $\overline{X \setminus Y} = X$ et X et Y soient régulièrement situés. Il existe un voisinage V de $X \setminus Y$ contenu dans $W \setminus Y$ tel que*

1) $V \subset X \setminus Y$
2) $\bar{V} \cap Y = X \cap Y$
3) *$\bar{V}$ et Y soient régulièrement situés.*

Preuve. Posons $V=\{z\in W: d(z,Y)>d(z,X)\}$ où d est la distance dans $\mathbb{R}^n$ déduite de la norme euclidienne $|\ |$. On a, successivement:

1) $V\subset W\setminus Y$ car, si $z\in Y$, $d(z,Y)=0$ ce qui interdit $z\in V$ par l'inégalité stricte.

2) $z\in X\setminus Y$ implique $d(z,X)=0$ et $d(z,Y)>0$ donc $z\in V$.

3) $\bar{V}\cap Y\subset X\cap Y$ car, si $z\in\bar{V}\cap Y$ $d(z,Y)=0$ et il existe $z_k\in V$ tel que $d(z_k,z)\to 0$ si $k\to+\infty$ et $d(z_k,Y)>d(z_k,X)$ donc $d(z_k,Y)\to 0$ et ainsi $d(z_k,X)\to d(z,X)=0$.

4) $X\cap Y\subset\bar{V}\cap Y$ car, si $z\in X\cap Y=\overline{X\setminus Y}\cap Y$ il existe $z_k\in X\setminus Y$ rel que $d(z_k,z)\to 0$ avec $k\to+\infty$. D'après 2) $z_k\in V\subset\bar{V}$ donc $z\in\bar{V}\cap Y$.

5) Montrons que $\bar{V}$ et Y sont régulièrement situés. Puisque X et Y sont régulièrement situés par hypothèse il existe, pour tout $z_0\in X\cap Y=\bar{V}\cap Y$ un voisinage W_1 de z_0 et des constantes $C>0$ et $\alpha\geqq 0$ tels que, pour tout $z\in W_1\cap Y$: $d(z,X)\geqq C(d(z,X\cap Y))^\alpha$.

Mais

$$d(z,X)\leqq d(z,\bar{V})+d(\zeta,X)$$

où ζ est un point de $\bar{V}$ tel que $d(z,\bar{V})=|\zeta-z|$.

Par définition de V on a $d(\zeta,X)\leqq d(\zeta,Y)$. Ainsi

$$d(z,X)\leqq d(z,\bar{V})+d(\zeta,Y).$$

Comme $z\in Y$ on a aussi:

$$d(\zeta,Y)=\inf_{z'\in Y}|\zeta-z'|=d(z,\bar{V}).$$

Par suite

$$d(z,X)\leqq 2\,d(z,\bar{V})$$

et, pour $z\in W_1\cap Y$

$$d(z,\bar{V})\geqq\frac{C}{2}(d(z,X\cap Y))^\alpha=\frac{C}{2}(d(z,\bar{V}\cap Y))^\alpha.$$

D'où le lemme.

5.11 Remarque. Le fait que X et Y soient régulièrement situés n'implique pas que $\overline{X\setminus Y}$ et Y le soient. Par exemple $X=[0,1/2]\cup G$ avec $G=\{(x,e^{-1/x^2}): 0\leqq x\leqq 1/2\}$ et $Y=[0,1]$ dans $\mathbb{C}$.

Revenant à 5.9 nous posons $X=Z\setminus\partial U=Z$ et $Y=\partial U$. Nous avons $\overline{X\setminus Y}=\overline{Z\setminus\partial U\setminus\partial U}=\overline{Z\setminus\partial U}=Z$ car si $z\in X$ il existe $z_k\in Z\setminus\partial U\subset X$ donc $z_k\in X\setminus\partial U$ tel que $z_k\to z$. Ainsi $X\subset X\setminus Y$. L'autre inclusion est évidente.

D'après le lemme 5.10 il existe un voisinage V de $Z\setminus\partial U=\overline{Z\setminus\partial U}\setminus\partial U$ contenu dans $\Omega\setminus\partial U$ tel que:

1) $\overline{Z\setminus\partial U}\setminus\partial U\subset V=Z\setminus\partial U$

2) $\bar{V}\cap\partial U=\overline{Z\setminus\partial U}\cap\partial U=Z\setminus\partial U$

3) $\bar{V}$ et ∂U soient régulièrement situés.

On écrit $V=V_1\cup V_2$ avec $V_1=V\cap U$ et $V_2=V\cap(\Omega\setminus\bar{U})$.

Soit V_3 l'ensemble des $z=(z_1,\ldots,z_p,z_{p+1},\ldots,z_n)\in V_1$ tels que

$$(0,\ldots,0,z_{p+1},\ldots,z_n)\in\overline{Z\setminus\partial U}=Z\setminus\partial U$$

et le segment d'origine $(0,\ldots,z_{p+1},\ldots,z_n)$ et d'extrémité z soit contenu dans V_1. L'ensemble V_3 est un voisinage ouvert de $Z\setminus\partial U$ et, pour $z\in V_3$ on a

$$g(z)=\sum_{1\leq j\leq p} z_j\int_0^1\frac{\partial g}{\partial z_j}(t\,z_1,\ldots,t\,z_p,z_{p+1},\ldots,z_n)\,dt$$

car g est holomorphe dans U.

Les fonctions $\varphi_j(z)=\int_0^1\frac{\partial\tilde{g}}{\partial z_j}(t\,z_1,\ldots,t\,z_p,z_{p+1},z_n)\,dt$ sont définies et C^∞ dans Ω en appelant $\tilde{g}$ un prolongement à Ω de g.

On considère la fonction de Whitney sur $\bar{U}\cap\bar{V}_2$, définie par le recollement de g sur $\bar{U}$ et de $\sum_{1\leq j\leq p} z_j\varphi_j$ sur $\bar{V}_2$, qui est bien définie car $\bar{U}$ et $\bar{V}_2$ sont régulièrement situés et les deux champs de Whitney considérés coincident, par construction, sur $\bar{U}\cap\bar{V}_2$. Cette fonction est restriction à $\bar{U}\cap\bar{V}_2$ d'une fonction $G\in\mathscr{E}^{00}(\Omega)$ qui est telle que, pour tout $\zeta\in\Omega$, la série formelle $T_\zeta(G)$ soit dans l'idéal engendré par $z_1,\ldots,z_p$ dans l'anneau des séries formelles en $(z_1-\zeta_1,\ldots,\bar{z}_n-\bar{\zeta}_n)$. Ainsi, d'après [21] G s'écrit $G=\sum_{1\leq j\leq p} z_j G_j$ avec $G_j\in\mathscr{E}^{00}(\Omega)$ et

$$g=G_{|\bar{U}}=\sum_{1\leq j\leq p} z_j\,G_{j|\bar{U}}\in\Gamma(\bar{U},I\mathscr{E}^{00}_{\bar{U}}).$$

D'où le lemme 5.9.

Pour le second point la factorisation aux points intérieurs est acquise par hypothèse et, pout les points ζ du bord, on se ramène au lemme 5.9 en extrayant, localement, un système $h_{i_1},\ldots,h_{i_k}$ tel que $\partial h_{i_1}(\zeta)\wedge\ldots\wedge\partial h_{i_k}(\zeta)\neq 0$ avec $k.=n-\dim_\zeta Z$. D'où la proposition.

En utilisant les mêmes idées nous prouvons maintenant un théorème d'extension.

5.12 Théorème. *Soient Ω un ouvert pseudo-convexe borné à bord lisse de classe C^∞ et $h_1,\ldots,h_p$ des fonctions holomorphes dans un voisinage U de $\bar{\Omega}$. Soit $Z=\{z\in U: h_1(z)=\ldots=h_p(z)=0$.*

Soit f une fonction C^∞ sur $Z\cap\bar{\Omega}$, restriction à $Z\cap\bar{\Omega}$ d'une fonction $F\in\Gamma(\bar{\Omega},\mathscr{E}^{00}_\Omega)$ telle que $\bar{\partial}F$ soit infiniment plate sur $Z\cap\bar{\Omega}$.

Si $X=\overline{Z\setminus\bar{\Omega}}$ est régulièrement situé par rapport à $Y=\bar{\Omega}$ dans U il existe $G\in A^\infty(\Omega)$ telle que $G-F\in\Gamma(\bar{\Omega},I\mathscr{E}^{00}_\Omega)$.

Preuve. Les coefficients $\frac{\partial F}{\partial\bar{z}_j}$ de $\bar{\partial}F$ sont infiniment plats sur $Z\cap\bar{\Omega}$. En appliquant le lemme 5.10 à X et Y on trouve un voisinage V de $X\setminus\bar{\Omega}$ contenu dans $U\setminus\bar{\Omega}$ tel que

1) $X\setminus\bar{\Omega}\subset V$
2) $\bar{V}\cap\bar{\Omega}=X\cap\bar{\Omega}$
3) $\bar{V}$ et $\bar{\Omega}$ soient régulièrement situés.

On peut alors trouver $\omega_j \in \Gamma(U, \mathscr{E}^{00})$ telle que $\omega_{j|\Omega} = \dfrac{\partial F}{\partial \bar{z}_j}$ et $\omega_{j|V} = 0$ de telle sorte que $\omega_j \in \Gamma(U, I\mathscr{E}^{00})$ et $\sum_{1 \leqq j \leqq n} \omega_j d\bar{z}_j = \omega \in \Gamma(U, I\mathscr{E}^{0,1})$ ce qui assure que $\bar{\partial} F = \omega_{|\Omega} \in \Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{0,1})$.

Puisque $\bar{\partial}(\omega_{|\Omega}) = 0$ il existe, d'après 5.2, $g \in \Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00})$ telle que $\bar{\partial} g = \bar{\partial} F$. Alors $\bar{\partial}(g - F) = 0$ donc $G = F - g \in A^\infty(\Omega)$ et $G - F = -g \in \Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00})$. D'où le théorème.

5.13 Proposition (cf. [10]). *Soient Ω un ouvert pseudo-convexe borné à bord lisse de classe C^∞, Z un espace analytique défini par l'annulation des fonctions $h_1, \ldots, h_p$ holomorphes dans un voisinage U de $\bar{\Omega}$, $I = (h_1, \ldots, h_p)\mathscr{O}$. On suppose que $\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00})$ est fermé dans $\Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{00})$. L'orthogonal de $\Gamma(\bar{\Omega}, IA^\infty)$ dans*

$$(A^\infty(\Omega))' = \frac{{}'\mathscr{D}_c^{n,n}(\bar{\Omega})}{\bar{\partial}\,{}'\mathscr{D}_c^{n,n-1}(\bar{\Omega})} \quad \textit{s'identifie à} \quad \frac{{}'\mathscr{D}_c^{n,n}(\bar{\Omega} \cap Z)}{\bar{\partial}\,{}'\mathscr{D}_c^{n,n-1}(\bar{\Omega} \cap Z)}$$

où ${}'\mathscr{D}_c^{p,q}(\bar{\Omega} \cap Z)$ désigne l'ensemble des courants de bidegré (p,q) à support compact contenu dans $\bar{\Omega} \cap Z$ tels que $g(T_{|w})$ soit le courant nul pour toute $g \in \Gamma(w \cap \bar{\Omega}, IA^\infty)$ et tout w ouvert de U.

Preuve. D'après 5.2 puisque $\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00})$ est fermé dans $\Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{00})$ l'idéal $\Gamma(\bar{\Omega}, IA^\infty)$ est fermé dans $A^\infty(\Omega)$.

Considérons le double complexe d'espaces FN. (Les espaces $\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{pq})$ sont fermés dans les espaces $\Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{pq})$ d'après l'hypothèse)

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \Gamma(\bar{\Omega}, IA^\infty) & \xrightarrow{\bar{\partial}} & \Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00}) & \xrightarrow{\bar{\partial}} & \Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{01}) \xrightarrow{\bar{\partial}} \cdots \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \Gamma(\bar{\Omega}, A^\infty) & \longrightarrow & \Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{00}) & \xrightarrow{\bar{\partial}} & \Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{01}) \xrightarrow{\bar{\partial}} \cdots \\
 & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & \dfrac{\Gamma(\bar{\Omega}, A^\infty)}{\Gamma(\bar{\Omega}, IA^\infty)} & \longrightarrow & \dfrac{\Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{00})}{\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00})} & \xrightarrow{\bar{\partial}_1} & \dfrac{\Gamma(\bar{\Omega}, \mathscr{E}_\Omega^{01})}{\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{01})} \xrightarrow{\bar{\partial}} \cdots \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & 0 & & 0 & & 0
\end{array}$$

où $\bar{\partial}_1$ est $\bar{\partial}$ passé aux quotients.

Les deux premières lignes étant exactes la troisième l'est et le lemme de dualité de Serre [18] donne, par transposition (au signe près ${}^t\bar{\partial}_1 = \bar{\partial}$) la suite exacte

$$0 \to (\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{0,n}))^\perp \xrightarrow{\bar{\partial}} \cdots \xrightarrow{\bar{\partial}} (\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{0,1}))^\perp \xrightarrow{\bar{\partial}} (\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00}))^\perp \to (\Gamma(\bar{\Omega}, IA^\infty))^\perp \to$$

En particulier:

$$(\Gamma(\bar{\Omega}, IA^\infty))^\perp = \frac{(\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{00}))^\perp}{\bar{\partial}(\Gamma(\bar{\Omega}, I\mathscr{E}_\Omega^{0,1}))^\perp}.$$

Pour conclure il reste à prouver que $(\Gamma(\bar{\Omega}, I\mathscr{E}^{p,q}))^{\perp}$ s'identifie à $\mathscr{D}_c'^{n-p,n-q}(\bar{\Omega} \cap Z)$. Si $T \in \mathscr{D}_c'^{n-p,n-q}(\bar{\Omega} \cap Z)$ et $\varphi \in \Gamma(\bar{\Omega}, I\mathscr{E}_{\Omega}^{pq})$

$$\varphi = \sum_{1 \leqq j \leqq p} h_j \sum_{\substack{|I|=p \\ |J|=q}} \varphi_{j,I,J}\, dZ_I \wedge d\bar{Z}_J$$

on a

$$\langle T, \varphi \rangle = \sum_{I,J,j} \langle h_j T, \varphi_{j,I,J}\, dZ_I \wedge d\bar{Z}_J \rangle = 0$$

car $h_j T = 0$. Ainsi $T \in (\Gamma(\bar{\Omega}, I\mathscr{E}_{\Omega}^{pq}))^{\perp}$.

Réciproquement soit $T \in (\Gamma(\bar{\Omega}, I\mathscr{E}_{\Omega}^{pq}))^{\perp}$. Alors, pour toute φ comme ci-dessus, $\langle T, \varphi \rangle = 0$; on en déduit aisément que $h_j T_{|w} = 0$ pour tout ouvert w de U. D'où la proposition.

Dans la suite de ce paragraphe nous allons étendre au cas non borné les résultats précédents.

5.14 Proposition. *Soient Ω un ouvert pseudo-convexe non borné à bord lisse de classe C^{∞}, $h_1, \ldots, h_p$ des fonctions holomorphes dans un voisinage U de $\bar{\Omega}$, $I = (h_1, \ldots, h_p)\mathscr{O}$. On suppose que $\Gamma(\bar{\Omega}, I\mathscr{E}_{\Omega}^{00}))$ est fermé dans $\Gamma(\bar{\Omega}, \mathscr{E}_{\Omega}^{00})$. Alors*

$$H^q(\bar{\Omega}, I \otimes_{\mathscr{O}} A^{\infty}) = 0 \qquad (q \geqq 1).$$

Preuve. Soit $(\bar{\Omega}_j)_{j \geqq 1}$ une exhaustion de $\bar{\Omega}$ obtenue par le procédé 2.2. Pour tout $j \geqq 1$ et $q \geqq 1$ on a donc $H^q(\bar{\Omega}_j, I \otimes_{\mathscr{O}} A^{\infty}) = 0$ la propriété B, de Nagel étant exacte ici.

Pour $q \geqq 2$ nous utilisons le lemme classique suivant:

5.15 Lemme ([3]). *Soient X un espace topologique paracompact $(X_j)_{j \geqq 1}$ une suite de fermés de X tels que $X_j \subset \mathring{X}_{j+1}$ pour tout $j \geqq 1$, $X = \bigcup_{j \geqq 1} X_j$ et $\mathscr{F}$ un faisceau de groupes abéliens sur X. Alors l'application canonique*

$$H^p(X, \mathscr{F}) \to \varprojlim_j H^p(X_j; \mathscr{F})$$

est surjective et bijective si l'on suppose de plus que l'application $H^{p-1}(X_{j+1}, \mathscr{F}) \to H^{p-1}(X_j, \mathscr{F})$ est surjective pour tout $j \geqq 1$.

Dans notre cas, pour $q \geqq 2$ $H^{q-1}(\bar{\Omega}_{j+1}, I \otimes_{\mathscr{O}} A^{\infty}) = H^{q-1}(\bar{\Omega}_j, I \otimes_{\mathscr{O}} A^{\infty}) = 0$. D'où la conclusion dans ce cas.

Examinons à présent le cas $q = 1$. Nous utiliserons les semi-normes

$$p_{m,k}(f) = \sup_{\substack{|\alpha| \leqq m \\ \zeta \in \bar{\Omega}_k}} |D_z^{\alpha} f(\zeta)|$$

5.16 Lemme. *Pour tout $k \geqq 1$ les espaces $\Gamma(\bar{\Omega}_k, I\mathscr{E}_{\Omega}^{00})$ sont fermés dans $\Gamma(\bar{\Omega}_k, \mathscr{E}_{\Omega}^{00})$.*

Preuve. Soit $f = \lim_{j \to \infty} f_j$ avec $f_j \in \Gamma(\bar{\Omega}_k, I\mathscr{E}_{\Omega}^{00})$ la limite ayant lieu dans $\Gamma(\bar{\Omega}_k, \mathscr{E}_{\Omega}^{00})$.

Pour voir que $f \in \Gamma(\bar{\Omega}_k, I\mathscr{E}_{\Omega}^{00})$ il suffit de s'assurer que, pour tout $\zeta \in \bar{\Omega}_k$ le germe f_{ζ} de f en ζ appartient à $I_{\zeta}\mathscr{E}_{\Omega}^{00}$.

a) Si $\zeta \notin Z \cap \bar{\Omega}_k$ ceci est clair

b) Si $\zeta \in Z \cap \bar{\Omega}_k \cap \Omega$ on a $f_\zeta \in I_\zeta \mathscr{E}^{00}_{\Omega,\zeta} = I_\zeta \mathscr{E}^{00}_\zeta$ et on sait que cet idéal est fermé dans $\mathscr{E}^{00}_\zeta$ et que $f_\zeta = \lim_{j\to\infty} f_{j,\zeta}$ dans $\mathscr{E}^{00}_\zeta$.

c) Si $\zeta \in Z \cap \bar{\Omega}_k \cap \partial\Omega$ et si $\chi \in \mathscr{D}(\mathbb{C}^n)$ est à support compact dans un voisinage de ζ assez petit et est égale à 1 dans un voisinage de ζ la suite χf_j converge dans $\Gamma(\bar{\Omega}, \mathscr{E}^{00}_\Omega)$ vers χf et, puisque $\chi f_j \in \Gamma(\bar{\Omega}, I\mathscr{E}^{00}_\Omega)$, on a $\chi f \in \Gamma(\bar{\Omega}, I\mathscr{E}^{00}_\Omega)$ d'après l'hypothèse et $f_\zeta \in I_\zeta \mathscr{E}^{00}_\Omega$. D'où le lemme.

Puisque

$$0 \to I \otimes_\mathscr{O} A^\infty \to I \otimes_\mathscr{O} \mathscr{E}^{00}_\Omega \xrightarrow{1\otimes\bar\partial} \dots \xrightarrow{1\otimes\bar\partial} I \otimes_\mathscr{O} \mathscr{E}^{0,n}_\Omega \to 0$$

est une résolution fine de $I \otimes A^\infty$ on a

$$H^1(\bar{\Omega}, I \otimes_\mathscr{O} A^\infty) = \frac{\operatorname{Ker}[\Gamma(\bar{\Omega}, I\mathscr{E}^{01}_\Omega) \xrightarrow{\bar\partial} \Gamma(\bar{\Omega}, I\mathscr{E}^{0,2}_\Omega)]}{\operatorname{Im}[\Gamma(\bar{\Omega}, I\mathscr{E}^{00}_\Omega) \xrightarrow{\bar\partial} \Gamma(\bar{\Omega}, I\mathscr{E}^{0,1}_\Omega)]}.$$

Soit donc $\omega \in \Gamma(\bar{\Omega}, I\mathscr{E}^{0,1})$ telle que $\bar\partial\omega = 0$. Pour tout $k \geqq 1$, il existe $\alpha_k \in \Gamma(\bar{\Omega}_k, I\mathscr{E}^{00}_\Omega)$ telle que $\bar\partial\alpha_k = \omega$ (par 5.2) et, alors, $\alpha_{k+1} - \alpha_k \in \Gamma(\bar{\Omega}_k, IA^\infty)$. Mais, puisque $\Gamma(\bar{\Omega}_k, IA^\infty) = \Gamma(\bar{\Omega}_k, I) A^\infty(\Omega_k)$ et puisque $A^\infty(\Omega_{k+1})$ est dense dans $A^\infty(\Omega_k)$ (par 3.2), l'espace $\Gamma(\bar{\Omega}_{k+1}, IA^\infty)$ est dense dans $\Gamma(\bar{\Omega}_k, IA^\infty)$. On peut donc modifier α_{k+1} en lui ajoutant $\gamma_k \in \Gamma(\bar{\Omega}_{k+1}, IA^\infty)$ de telle sorte que $p_{k,k}(\alpha_{k+1} - \alpha_k) \leqq \frac{1}{2^k}$.

La série:

$$(\alpha_{k+1} - \alpha_k) + (\alpha_{k+2} - \alpha_{k+1}) + \dots$$

converge dans $A^\infty(\bar{\Omega}_k)$. Soit s_k sa limite. La fonction $s_k \in \Gamma(\bar{\Omega}_k, IA^\infty)$ d'après le lemme 5.16. Alors, la fonction $g_k = \alpha_k + s_k$ appartient à $\Gamma(\bar{\Omega}_k, I\mathscr{E}^{00}_\Omega)$, vérifie $\bar\partial g_k = \bar\partial\alpha_k$ dans $\bar{\Omega}_k$ et $g_{k+1|\Omega_k} = g_k = \lim_{N\to\infty} \alpha_N$.

La fonction $g = \lim_{N\to\infty} \alpha_N$ est donc élément de $\Gamma(\Omega, I\mathscr{E}^{00}_\Omega)$ et est telle que $\bar\partial g = \omega$. D'où la proposition.

5.17 Corollaire. *Avec les données et hypothèse de* 5.14 *la suite d'espaces F-N*

$$0 \to \Gamma(\bar{\Omega}, IA^\infty) \to \Gamma(\bar{\Omega}, I\mathscr{E}^{00}_\Omega) \xrightarrow{\bar\partial} \dots \xrightarrow{\bar\partial} \Gamma(\bar{\Omega}, I\mathscr{E}^{0,n}_\Omega) \to 0$$

est exacte.

On peut alors énoncer un théorème d'extension:

5.18 Théorème. *Soient Ω un ouvert pseudo-convexe non borné à bord lisse de classe C^∞, $h_1, \dots, h_p$ des fonctions holomorphes dans un voisinage U de $\bar{\Omega}$. Soit $Z = \{z \in U : h_1(z) = \dots = h_p(z) = 0\}$. Soit $I = (h_1, \dots, h_p)\mathscr{O}$. On suppose que $\Gamma(\bar{\Omega}, I\mathscr{E}^{00}_\Omega)$ est fermé dans $\Gamma(\bar{\Omega}, \mathscr{E}^{00}_\Omega)$.*

Soit f une fonction C^∞ sur $Z \cap \bar{\Omega}$, restriction à $Z \cap \bar{\Omega}$ d'une fonction $F \in \Gamma(\bar{\Omega}, \mathscr{E}^{00}_\Omega)$ telle que $\bar\partial F$ soit infiniment plate sur $Z \cap \bar{\Omega}$.

Si $X=Z\setminus\bar{\Omega}$ est régulièrement situé par rapport à $Y=\bar{\Omega}$ dans U il existe $G\in A^\infty(\Omega)$ telle que $G-F\in\Gamma(\bar{\Omega}, I\mathscr{E}^{00}_\Omega)$.

Preuve. Analogue à 5.12 en utilisant 5.17.

Examinons si l'on peut avoir $H^q(\bar{\Omega}, \mathscr{R}(h_1,\dots,h_p; A^\infty))=0$ $q\geqq 1$ où $\mathscr{R}(h_1,\dots,h_p, A^\infty)$ désigne le faisceau des relations entre les h_j sur $\bar{\Omega}$, ce qui conduirait à un résultat de division.

Soit $\tilde{U}$ l'enveloppe d'holomorphie de U. Le faisceau $\mathscr{R}(h_1,\dots,h_p;\mathscr{O})$ est cohérent sur $\tilde{U}$; donc, au voisinage de $\bar{\Omega}_{k+1}$, on peut trouver une résolution:

$$\dots\to\mathscr{O}^{p_{2n+1}}\to\mathscr{O}^{p_{2n}}\to\dots\to\mathscr{O}^{p_2}\to\mathscr{O}^{p_1}\xrightarrow{\mu}\mathscr{R}(h_1,\dots,h_p;\mathscr{O})\to 0$$

avec $\mu=(\mu_1,\dots,\mu_p)$, $\mu_j\in\Gamma(V,\mathscr{R}(h_1,\dots,h_p:\mathscr{O}))$, où V est un voisinage du compact $\widehat{\bar{\Omega}_{k+1}}$ enveloppe holomorphiquement convexe du compact $\bar{\Omega}_{k+1}$ de la variété de Stein $\tilde{U}$.

En tensorisant par $\otimes_{\mathscr{O}} A^\infty$ la suite exacte de faisceaux cohérents au voisinage de $\bar{\Omega}_{k+1}$ dans $\mathbb{C}^n$

$$0\to\operatorname{Ker}\mu\to\mathscr{O}^{p_1}\xrightarrow{\mu}\mathscr{R}(h_1,\dots,h_p;\mathscr{O})\to 0$$

on obtient, par platitude, la suite exacte

$$0\to\operatorname{Ker}\mu\otimes_{\mathscr{O}}A^\infty\to(A^\infty)^{p_1}\to\mathscr{R}(h_1,\dots,h_p;A^\infty)\to 0$$

car $\mathscr{O}^{p_1}\otimes_{\mathscr{O}}A^\infty=(A^\infty)^{p_1}$ et

$$\mathscr{R}(h_1,\dots,h_p;\mathscr{O})\otimes_{\mathscr{O}}A^\infty=\mathscr{R}(h_1,\dots,h_p;\mathscr{O})\,A^\infty=\mathscr{R}(h_1,\dots,h_p;A^\infty)$$

la dernière égalité venant de la platitude de A^∞ sur $\mathscr{O}$ et l'avant dernière de (Bourbaki Algèbre linéaire-chapitre 2–3, n^0 7, Prop. 7).

L'extension au cas pseudo-convexe de la propriété b)
$H^q(\bar{\Omega}_\alpha, \operatorname{Ker}\mu\otimes_{\mathscr{O}}A^\infty)=0$ $q\geqq 1$, $\alpha=k$ ou $k+1$ donne les suites exactes

$$0\to\Gamma(\bar{\Omega}_{k+1},\operatorname{Ker}\mu\otimes_{\mathscr{O}}A^\infty)\to(A^\infty(\Omega_{k+1}))^{p_1}\xrightarrow{\mu}\Gamma(\bar{\Omega}_{k+1},\mathscr{R}(h,A^\infty))\to 0$$
$$0\to\Gamma(\bar{\Omega}_{k},\operatorname{Ker}\mu\otimes_{\mathscr{O}}A^\infty)\to(A^\infty(\Omega_{k}))^{p_1}\xrightarrow{\mu}\Gamma(\bar{\Omega}_{k},\mathscr{R}(h,A^\infty))\to 0$$

ce qui assure que $\mu_1,\dots,\mu_p$ engendrent $\Gamma(\Omega_\alpha,\mathscr{R}(h,A^\infty))$ $\alpha\geqq 1$. Puisque $A^\infty(\Omega_{k+1})$ est dense dans $A^\infty(\Omega_k)$ on en déduit que $\Gamma(\bar{\Omega}_{k+1},\mathscr{R}(h,A^\infty))$ est dense dans $\Gamma(\bar{\Omega}_k,\mathscr{R}(h,A^\infty))$, comme sous-espace de $(A^\infty(\Omega_k))$.

En tensorisant par $\otimes_{\mathscr{O}}A^\infty$ la suite exacte de faisceaux cohérents dans un voisinage de $\bar{\Omega}_{k+1}$ dans $\mathbb{C}^n$

$$0\to\mathscr{R}(h;\mathscr{O})\to\mathscr{O}^p\xrightarrow{h}I\to 0$$

on obtient, toujours par platitude, la suite exacte

$$0\to\mathscr{R}(h;A^\infty)\to(A^\infty)^p\xrightarrow{h}IA^\infty\to 0$$

d'où l'on déduit une suite exacte

(car, par la propriété *B*), $H^q(\bar{\Omega}_k,\mathscr{R}(h;\mathscr{O})\otimes_{\mathscr{O}}A^\infty)=0$ $\quad(q\geqq 1)$
$$0\to\Gamma(\bar{\Omega}_k,\mathscr{R}(h;A^\infty))\to(A^\infty(\Omega_k))^p\xrightarrow{h}\Gamma(\bar{\Omega}_k,IA^\infty)\to 0.$$

Comme h est continue on obtient que $\Gamma(\bar{\Omega}_k, \mathscr{R}(h; A^\infty))$ est un sous-espace fermé de $(A^\infty(\Omega_k))^p$, donc un espace de Fréchet.

Montrons alors que $H^1(\bar{\Omega}; \mathscr{R}(h, A^\infty)) = 0$.

Soit

$$0 \to \mathscr{R}(h, A^\infty) \to \mathscr{F}^0 \xrightarrow{d_0} \mathscr{F}^1 \xrightarrow{d_1} \mathscr{F}^2 \to \dots$$

la résolution flasque canonique de $\mathscr{R}(h, A^\infty)$ de telle sorte que:

$$H^1(\bar{\Omega}, \mathscr{R}(h; A^\infty)) = \frac{\operatorname{Ker}[\Gamma(\bar{\Omega}, \mathscr{F}^1) \to \Gamma(\bar{\Omega}, \mathscr{F}^2)]}{\operatorname{Im}[\Gamma(\bar{\Omega}, \mathscr{F}^0) \to \Gamma(\bar{\Omega}, \mathscr{F}^1)]}.$$

Soit $\alpha \in \Gamma(\bar{\Omega}, \mathscr{F}^1)$ tel que $d_1 \alpha = 0$. Pour tout $i \geqq 1$ il existe $\beta_i \in \Gamma(\bar{\Omega}_i, \mathscr{F}^0)$ tel que $d_0(\beta_i) = \alpha$ sur $\bar{\Omega}_i$ puisque $H^1(\bar{\Omega}_i, \mathscr{R}(h; A^\infty)) = 0$.

On a donc $d_0(\beta_{i+1|\Omega_i} - \beta_i) = 0$ sur $\bar{\Omega}_i$, ce qui signifie que $\beta_{i+1|\Omega_i} - \beta_i \in \Gamma(\bar{\Omega}_i, \mathscr{R}(h; A^\infty))$. D'après la densité de $\Gamma(\bar{\Omega}_{i+1}, \mathscr{R}(h; A^\infty))$ dans $\Gamma(\bar{\Omega}_i, \mathscr{R}(h; A^\infty))$, il existe $\gamma_i \in \Gamma(\bar{\Omega}_{i+1}, \mathscr{R}(h, A^\infty))$ tel que

$$d_i(\beta_{i+1} - \beta_i, \gamma_i) < \frac{1}{2^i}$$

où d_i est une distance sur $\Gamma(\bar{\Omega}_i; \mathscr{R}(h; A^\infty))$ invariante par translation définissant la topologie de cet espace de Fréchet. On peut supposer $d_i \leqq d_{i+1}$.

En remplacant β_{i+1} par $\beta_{i+1} - \gamma_i$ on peut supposer que $d_i(\beta_{i+1} - \beta_i, 0) < \frac{1}{2^i}$.

Pour tout k soit $s_k \in \Gamma(\bar{\Omega}_k, \mathscr{R}(h; A^\infty))$ la somme de la série convergente

$$\sum_{l \geqq k} (\beta_{l+1} - \beta_l)_{|\Omega_k}$$

et soit $\omega_k = \beta_k + s_k \in \Gamma(\bar{\Omega}_k, \mathscr{F}^0)$.

On a $d_0(\omega_k) = d_0(\beta_k) = \alpha$ sur $\bar{\Omega}_k$ car $s_k \in \Gamma(\bar{\Omega}, \mathscr{R}(h, A^\infty))$ et

$$0 \to \Gamma(\bar{\Omega}, \mathscr{R}(h, A^\infty)) \to \Gamma(\bar{\Omega}, \mathscr{F}^0) \xrightarrow{d_0} \Gamma(\bar{\Omega}, \mathscr{F}^1)$$

est exacte.

On a $\omega_{k+1|\bar{\Omega}_k} = \omega_k$. Ainsi les ω_k définissent une $\omega \in \Gamma(\bar{\Omega}, \mathscr{F}^0)$ telle que $d_0 \omega = \alpha$.

Enfin $H^q(\bar{\Omega}\, \mathscr{R}(h; A^\infty)) = 0$ pour $q \geqq 2$ car $H^{q-1}(\bar{\Omega}, \mathscr{R}(h; A^\infty)) = H^{q-1}(\bar{\Omega}_{k+1}, \mathscr{R}(h; A^\infty)) = 0$ pour tout k (d'après 5.15).

Nous avons ainsi prouvé le

5.19 Théorème. *Soient Ω un ouvert pseudo-convexe non borné à bord lisse de classe C^∞, $h_1, \dots, h_p$ des fonctions holomorphes dans un voisinage U de $\bar{\Omega}$, $I = (h_1, \dots, h_p)\mathscr{O}$. On a*

$$H^q(\Omega, \mathscr{R}(h_1, \dots, h_p; A^\infty)) = 0 \quad (q \geqq 1).$$

5.20 Corollaire. *Sous les mêmes hypothèses $h_1, \dots, h_p$ engendrent $\Gamma(\bar{\Omega}, IA^\infty)$ comme $A^\infty(\Omega)$-module.*

Preuve. Soit $h: \mathcal{O}^p \to \mathcal{O}$ l'homomorphisme de $\mathcal{O}$-modules dans la variété de Stein $\tilde{U}$ défini par les h_j.

La suite de $\mathcal{O}$-modules cohérents

$$0 \to \operatorname{Ker} h \to \mathcal{O}^p \xrightarrow{h} I \to 0$$

au voisinage de $\bar{\Omega}$ donne, par tensorisation par $\otimes_{\mathcal{O}} A^\infty$, grâce à la platitude de A^∞ sur $\mathcal{O}$, la suite exacte:

$$0 \to \mathcal{R}(h, A^\infty) \to (A^\infty)^p \xrightarrow{h} IA^\infty \to 0$$

et le théorème 5.19 donne alors la suite exacte:

$$0 \to \Gamma(\bar{\Omega}, \mathcal{R}(h; A^\infty)) \to (A^\infty(\Omega))^p \xrightarrow{h} \Gamma(\bar{\Omega}, IA^\infty) \to 0.$$

D'où le corollaire.

5.21 Remarque. $\Gamma(\bar{\Omega}, \mathcal{R}(h; A^\infty))$ est alors un espace de Fréchet.

5.22 Corollaire. Sous les mêmes hypothèses, une $g \in A^\infty(\Omega)$ s'écrit $g = \sum_{1 \leqq j \leqq p} g_j h_j$, avec $g_j \in A^\infty(\Omega)$, si seulement si g s'annule avec multiplicité sur $Z \cap \bar{\Omega}$.

5.23 Corollaire. *Sous les mêmes hypothèses l'énoncé* 5.13 *est valable.*

VI. Extension: cas de Z ne débordant pas

Nous nous proposons d'étendre au cas non borné les résultats d'extension, obtenus par E. Amar [2], dans le cas d'un ouvert pseudoconvexe borné à bord lisse de classe C^∞ et de Z de la forme $Z = \tilde{Z} \cap \Omega$ avec $\tilde{Z}$ variété C^∞ dans un voisinage de $\partial\Omega$, intersection complète globale, sous l'hypothèse que $\overline{\tilde{Z} \setminus \bar{\Omega}}$ et $\bar{\Omega}$ sont régulièrement situés.

Nous reprenons certains des résultats de [2] en utilisant les idées des paragraphes précédents.

6.1 Proposition. *Soient X un espace topologique paracompact de dimension cohomologique l_0 finie, A un faisceau d'anneaux sur X et $\mathcal{F}$ un faisceau de A-modules. Soit O un ouvert de X. On suppose que $H^p(O, A) = H^p(O, \mathcal{F}) = H^p(X, A) = 0$ pour $p \geqq 1$. Si $\mathcal{F}$ possède sur un voisinage U de $X \setminus O$ une résolution libre;*

$$\ldots \to A^{p_n} \to A^{p_{n-1}} \to \ldots \to A^{p_1} \xrightarrow{u} \mathcal{F} \to 0$$

alors $H^p(X; \mathcal{F}) = 0$ pour $p \geqq 1$.

Preuve. Nous utiliserons la cohomologie à support dans une partie fermée (voir [3] pour cette notion).

6.2 Lemme. $H^p_{X \setminus O}(X, A) = H^p_{X \setminus O}(U, A) = 0$ *pour* $p \geqq 2$.

Preuve. L'égalité provient du théorème d'excision. Par ailleurs on a la longue suite de cohomologie à support dans $X \setminus O$ pour le faisceau A aur X

$$\begin{array}{ccccccccc} 0\to \Gamma_{X\setminus O}(X,A) & \to \Gamma(O,A) & \to H^1_{X\setminus O}(X,A) & \to & H^1(X,A) & \to H^1(O,A) & \to & & \\ & & & & \| & \| & & & \\ & & & & 0 & 0 & & & \\ \ldots\to H^2_{X\setminus O}(X,A) & \to H^2(X,A) & \to & H^2(O,A) & \to H^3(X\setminus O,A) & \to H^3(X,A) & \to \ldots & & \\ & \| & & \| & \| & \| & & & \\ & 0 & & 0 & 0 & 0 & & & \end{array}$$

qui assure la conclusion.

Sur U découpons la résolution libre donnée en courtes suites exactes:

$$\begin{array}{l} 0\to L_1\to A^{p_1}\to \mathscr{F}=L_0\to 0 \\ 0\to L_2\to A^{p_2}\to L_1\to 0 \\ \cdots \end{array}$$

6.3 Lemme. *$H^p_{X\setminus O}(U,L_k)=H^{p+1}_{X\setminus O}(U,L_{k+1})$ pour tout $k\geqq 0$ et $p\geqq 2$ et $H^p_{X\setminus O}(X,L_k)=H^{p+1}_{X\setminus O}(X,L_{k+1})$ dans les mêmes conditions.*

Preuve. La deuxième affirmation vient de la première et du théorème d'excision. Montrons la première.

Ecrivons d'abord la longue suite de cohomologie à support dans $X\setminus O$ pour la courte suite $0\to L_1\to A^{p_1}\to L_0\to 0$. Il vient:

$$\begin{array}{l} 0\to \Gamma_{X\setminus O}(U,L_1)\to \Gamma_{X\setminus O}(U,A^{p_1})\to \Gamma_{X\setminus O}(U,L_0)\to \\ \to H^1_{X\setminus O}(U,L_1)\to H^1_{X\setminus O}(U,A^{p_1})\to H^1_{X\setminus O}(U,L_0)\to \\ \to H^2_{X\setminus O}(U,L_1)\to H^2_{X\setminus O}(U,A^{p_1})\to H^2_{X\setminus O}(U,L_0)\to \\ \qquad\qquad\qquad\qquad \| \\ \qquad\qquad\qquad\qquad 0 \\ \to H^3_{X\setminus O}(U,L_1)\to H^3_{X\setminus O}(U,A^{p_1})\to \qquad \ldots \\ \qquad\qquad\qquad\qquad \| \\ \qquad\qquad\qquad\qquad 0 \end{array}$$

et il en résulte $H^p_{X\setminus O}(U,L_0)=H^{p+1}_{X\setminus O}(U,L_1)$ $(p\geqq 2)$.

Le raisonnement pour les autres suites est identique.

6.4 Lemme. *Pour $l\geqq l_0+1$ et tout faisceau G de A-modules sur X on a $H_{X\setminus O}(X,G)=0$.*

Preuve. La longue suite de cohomologie à support dans $X\setminus O$ donne

$$\begin{array}{ccccc} \ldots\to H^{l-1}(O,G) & \to H^l_{X\setminus O}(X,G) & \to H^l(X,G) & \to \ldots & \\ \| & & \| & & \\ 0 & & 0 & & \end{array}$$

pour $l\geqq l_0+1$.

Ainsi $H^{p+k}_{X\setminus O}(X,L_k)=0$ pour $p+k\geqq l_0+1$ et par suite

6.5 Lemme. *$H^p_{X\setminus O}(X,\mathscr{R})=H^{p+k}_{X\setminus O}(X,L_k)=0$.*

6.6 Lemme. *$H^p(X,\mathscr{F})=0$ pour $p\geqq 2$.*

Preuve. On écrit la longue suite de cohomologie à support dans $X\setminus O$ pour le faisceau $\mathscr{F}$ sur X

$$\begin{array}{c} 0\to \Gamma_{X\setminus O}(X,\mathscr{F}) \to \Gamma(X,\mathscr{F}) \to \Gamma(O,\mathscr{F}) \to \\ \to H^1_{X\setminus O}(x,\mathscr{F}) \to H^1(X,\mathscr{F}) \to H^1(O,\mathscr{F}) \to \\ \qquad\qquad\qquad\qquad\qquad\qquad \| \\ \qquad\qquad\qquad\qquad\qquad\qquad 0 \\ \to H^2_{X\setminus O}(X,\mathscr{F}) \to H^2(X,\mathscr{F}) \to H^2(O,\mathscr{F}) \to \\ \qquad \| \qquad\qquad\qquad \| \\ \qquad 0 \qquad\qquad\qquad 0 \end{array}$$

Il nous reste à prouver que $H^1(X,\mathscr{F})=0$. Partons de la suite

$$0\to L_1\to A^{p_1}\xrightarrow{u}\mathscr{F}$$

où $\ker u=L_1$. On peut écrire le diagramme commutatif (car $0\to L_1\to A^{p_1}\xrightarrow{u}\mathscr{F}\to 0$ est exacte sur U):

$$\begin{array}{ccccccccc} & & \downarrow & & \downarrow & & \downarrow & & \\ 0\to\Gamma_{X\setminus O}(X,A^{p_1}) & \to & \Gamma(X,A^{p_1}) & \to & \Gamma(O,A^{p_1}) & \xrightarrow{\alpha} & H^1_{X\setminus O}(X,A^{p_1}) & \longrightarrow & H^1(X,A^{p_1})\longrightarrow\dots \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \| \\ 0\to\Gamma_{X\setminus O}(X,\mathscr{F}) & \to & \Gamma(X,\mathscr{F}) & \to & \Gamma(O,\mathscr{F}) & \xrightarrow{\beta} & H^1_{X\setminus O}(X,\mathscr{F}) & \xrightarrow{\varepsilon} & H^1(X,\mathscr{F}) \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \| \\ & & & & & & H^2_{X\setminus O}(X,L_1)=0 & & 0 \\ & & & & & & \downarrow & & \end{array}$$

ou ^{a} et β sont des morphismes de connexion et γ, δ sont induites par u ce qui assure que δ est surjective. Soit alors $x\in H^1_{X\setminus O}(X,\mathscr{F})$. Il existe $y\in H^1_{X\setminus O}(X,A^{p_1})$ tel que $x=\delta(y)$. Puisque $H^1(X,A^{p_1})=0$ ^{a} est surjective donc il existe $z\in\Gamma(\mathscr{O},A^{p_1})$ tel que $y={}^a(z)$. Alors $x=\beta(\gamma(z))$. Donc β est surjective, donc $\varepsilon=0$. Comme $H^1_{X\setminus O}(O,\mathscr{F})=0$ on a $\varphi=0$, donc $\operatorname{Ker}\varphi=\operatorname{Im}\varepsilon=H^1(X,\mathscr{F})=0$. D'où la proposition.

6.7 Definition ([2]). Soit Ω un ouvert à bord lisse de classe C^∞ de $\mathbb{C}^n$. Un faisceau $\mathscr{F}$ sur $\bar\Omega$ de A^∞-modules sera dit globalement cohérent si $\mathscr{F}_{|\Omega}$ est O-cohérent et s'il existe un voisinage V de $\partial\Omega$ et, sur $U=V\cap\bar\Omega$, une résolution libre

$$\to(A^\infty)^{p_{j+1}}\to(A^\infty)^{p_j}\to\dots\to(A^\infty)^{p_1}\xrightarrow{u}\mathscr{F}\to 0.$$

Ceci assure, au moins que sur Ω, le faisceau $\mathscr{F}$ est globalement de type fini, cohérent. En particulier, pour tout ouvert pseudo-convexe $O\subset\Omega$ on aura $H^p(O,\mathscr{F})=0$. On a alors la

6.8 Proposition. *Si Ω est un ouvert pseudo-convexe à bord lisse de classe C^∞ et $\mathscr{F}$ est un faisceau globalement cohérent sur $\bar\Omega$ alors $H^p(\bar\Omega,\mathscr{F})=0$ pour tout $p\geqq 1$.*

Preuve. $\bar\Omega$ est paracompact de dimension cohomologique $2n$ et l'on a $H^p(\Omega,A^\infty)=H^p(\Omega,O)=H^p(\Omega,\mathscr{F})=H(\bar\Omega,A^\infty)=0$ pour $p\geqq 1$. La conclusion résulte de 6.1.

Soient maintenant Ω un ouvert pseudo-convexe à bord lisse de classe C^∞ et $U_1, \ldots, U_p \in A^\infty(\Omega)$ telles que:

$$X = \{z \in \bar{\Omega} : U_1(z) = \ldots = U_p(z) = 0\}$$

soit tel que, pour tout $\zeta \subset \partial\Omega \cap X$, il existe un difféomorphisme φ_ζ de classe C^∞ d'un voisinage V_ζ de ζ dans $\mathbb{C}^n$ sur un voisinage U_ζ de 0 dans C^n tel que

a) $\varphi_{\zeta|V_\zeta \cap \Omega}$ soit biholomorphe

b) Les fonctions $h_j = U_j \circ \varphi_\zeta^{-1}$ soient égales à z_j dans U_ζ.

c) L'ensemble $Z_\zeta = \{z \in U_\zeta : z_1 = \ldots = z_p = 0\}$ est tel que $\overline{Z_\zeta \setminus \overline{\varphi_\zeta(V_\zeta \cap \Omega)}}$ est régulièrement situé par rapport à $\overline{\varphi_\zeta(V_\zeta \cap \Omega)}$.

Soit I le faisceau des germes de fonctions f such $\bar{\Omega}$ telles que $f_\zeta \in A^\infty_\zeta$ pour tout $\zeta \in \bar{\Omega}$, $f_\zeta \in I_\zeta$ si $\zeta \in \Omega \cap X$, avec $I_\zeta = (U_{1,\zeta}, \ldots, U_{p,\zeta})\mathcal{O}_\zeta$, et, si $\zeta \in X \cap \partial\Omega$ et f est définie dans $V_\zeta^1 \subset V_\zeta \cap \bar{\Omega}$, alors $f \circ \varphi_\zeta^{-1}$ s'annule sur $Z_\zeta \cap \varphi_\zeta(V_\zeta^1)$ dans $\varphi_\zeta(V_\zeta^1)$).

Alors:

6.9 Théorème. *Le faisceau I est globalement cohérent.*

Preuve. On remarque d'abord que $I_{|\Omega}$ est cohérent puisque de type fini engendré sur $\mathcal{O}_{|\Omega}$ par $U_1, \ldots, U_p$ dans Ω.

Si $\zeta \in \partial\Omega \cap X$ le complexe de Koszul:

$$0 \to \Lambda^p(\mathcal{O}_0^p) \xrightarrow{p_z} \Lambda^{p-1}(\mathcal{O}_0^p) \xrightarrow{p_z} \Lambda^{p-2}(\mathcal{O}_0^p) \xrightarrow{p_z} \ldots \xrightarrow{p_z} \Lambda^1(\mathcal{O}_p^0) \to J_0 \to 0$$

est une résolution de J_0 avec $J = \varphi_*(I) = (z_1, \ldots, z_p)\mathcal{O}$ dans U_ζ et $\Lambda(\mathcal{O}_0^p)$ est l'algèbre extérieure sur $\mathcal{O}_0^p$:

Si $(e_i)_{1 \leqq i \leqq p}$ est une base de $\Lambda^1(\mathcal{O}_0^p) \simeq \mathcal{O}_0^p$ toute $f \in \Lambda^k(\mathcal{O}_0^p)$ s'écrit:

$$f = \sum_{1 \leqq i_1 < \ldots < j_{k-1} \leqq p} f_{i_1 \ldots i_k} e_{i_1} \wedge \ldots \wedge e_{i_k}$$

et

$$P_z(f) = \sum_{1 \leqq j_1 < \ldots < j_{k-1} \leqq p} (P_z(f))_{j_1 \ldots j_{k-1}} e_{j_1} \wedge \ldots \wedge e_{j_{k-1}}$$

avec

$$(P_z(f))_{j_1 \ldots j_{k-1}} = \sum_{\substack{1 \leqq j \leqq p \\ j \notin \{j_1 \ldots j_{k-1}\}}} f_{j_1 \ldots j_{k-1}, j} z_j.$$

Soit alors A_0^∞ l'anneau des germes de fonctions holomorphes dans $\varphi_\zeta(V_\zeta \cap \Omega)$ et de classe C^∞ dans $\overline{\varphi_\zeta(V_\zeta \cap \Omega)}$. Comme $\varphi_\zeta(V_\zeta \cap \Omega)$ est pseudo-convexe à bord lisse de classe C^∞ au voisinage de 0 l'anneau A_0^∞ est $\mathcal{O}_0$-plat et, en tensorisant par $\otimes_{\mathcal{O}_0} A_0^\infty$ la résolution de Koszul de J_0, on obtient la résolution suivante de $J_0 \otimes_{\mathcal{O}_0} A_0^\infty = J_0 A_0^\infty$:

$$0 \to \Lambda^p((A_0^\infty)^p) \xrightarrow{p_z} \ldots \xrightarrow{p_z} \Lambda^1((A_0^\infty)^p) \to J_0 A_0^\infty \to 0.$$

En transformant par φ_ζ on obtient la résolution:

$$0 \to \Lambda^p((A_\zeta^\infty)^p) \xrightarrow{p_u} \ldots \xrightarrow{p_u} \Lambda^1((A_\zeta^\infty)^p) \to I_\zeta \to 0$$

car, si $f \in I_\zeta$, alors $f \circ \varphi_\zeta^{-1} \in J_0 A_0^\infty$ car $f \circ \varphi_\zeta^{-1}$ s'annule sur z_ζ (d'après 5.9).

Cette résolution est globale dans U, où $U=V\cap\bar{\Omega}$ et V est un voisinage de $\partial\Omega$ car, si $\zeta\in\partial\Omega\cap X$, l'un au moins des U_i est une unité. Par suite I est globalement cohérent.

6.10 Théorème. *Avec les données et hypothèses précédentes la suite*

$$0\to\Gamma(\bar{\Omega},I)\to\Gamma(\bar{\Omega},I\mathscr{E}^{00}_{\Omega})\xrightarrow{\bar\partial}\dots\xrightarrow{\bar\partial}\Gamma(\bar{\Omega},I\mathscr{E}^{0,n}_{\Omega})\to 0$$

est exacte.

Preuve. En tout point $\zeta\in\bar{\Omega}$ on a la résolution

$$0\to I_\zeta\to I_\zeta\mathscr{E}^{00}_{\Omega,\zeta}\xrightarrow{\bar\partial}\dots\xrightarrow{\bar\partial}I_\zeta\mathscr{E}^{0,n}_{\Omega,\zeta}\to 0. \qquad (*)$$

Ceci est clair si $\zeta\in\Omega$ ou $\zeta\in\bar{\Omega}\setminus X$ et, si $\zeta\in X\cap\partial\Omega$, on a la suite exacte

$$0\to J_0A^\infty_0\to J_0\mathscr{E}^{00}_0\xrightarrow{\bar\partial}\dots\xrightarrow{\bar\partial}J_0\mathscr{E}^{00}_0\to 0$$

où $\mathscr{E}_0$ est l'anneau des germes de fonctions C^∞ sur $\overline{\varphi_\zeta(V_\zeta\cap\Omega)}$.

On utilise enfin que $(*)$ est une résolution fine et I globalement cohérent.

6.11 Théorème. *Avec les données et hypothèses précédentes si f est une fonction C^∞ sur $X\cap\bar{\Omega}$ restriction à $X\cap\bar{\Omega}$ d'une F de classe C^∞ dans $\bar{\Omega}$ telle que $\bar\partial F$ soit infiniment plate sur $X\cap\bar{\Omega}$ il existe $G\in A^\infty(\bar{\Omega})$ telle que $G-F\in\Gamma(\bar{\Omega},I\mathscr{E}^{00}_{\Omega})$.*

Preuve. Analogue à celle de 5.12.

Note added in proof. Les énoncés 5.7, 5.8, 5.22 supposent $\overline{Z\cap\Omega}=Z\cap\bar{\Omega}$ et 5.9 suppose $\overline{Z\cap U}=Z\cap\bar{U}$.

Bibliographie

1. Adachi, K.: Continuation of A^∞-functions from submanifolds to strictly pseudo-convex domains. Math. Soc. Japan **32**, 331–341 (1980)
2. Amar, E.: Cohomologie complexe et application. A paraitre
3. Banica, C., Stanasila, O.: Méthodes algébriques dans la théorie globale des espaces complexes. Paris: Gauthier-Villars 1977
4. De Bartolomeis, P., Tomassini, G.: Finitely generated ideals in $A^\infty(\bar{D})$. Adv. Math. **46**, 162–170 (1982)
5. Berenstein, C.A., Taylor, B.A.: Interpolation problems in $\mathbb{C}^n$ with applications to harmonic analysis. J. Anal. Math. **38**, 188–254 (1980)
6. Bruna, J., Ortega, J.M.: Closed finitely generated ideals in algebras of holomorphic functions and smooth to the boundary in strickly pseudo convex domains. Math. Ann. **268**, 137–157 (1984)
7. Catlin, D.W.: Boundary behavior of holomorphic functions on weakly pseudo convex domains. Dissertation. Princeton University (1978)
8. Diederich, K., Fornaess, J.: Pseudoconvex domains: Bounded strictly plurisubharmonic exhaustion functions. Invent. Math. **39**, 129–141 (1977)
9. Elgueta, M.: Extensions strictly pseudoconvex domains of functions holomorphic in a submanifold in general position and C^∞ up to boundary. Ill. J. Math. **24**, 1–17 (1980)
10. Gay, R.: Division des fonctionnelles analytiques. Applications aux fonctions moyennes périodiques. CRAS Paris **283** série A, 835–838 (1976)
11. Hakim, M., Sibony, N.: Spectre de $A(\bar{\Omega})$ pour les domaines bornés faiblement pseudo-convexes réguliers. J. Funct. Anal. **37**, 127–135 (1980)
12. Henaut, A.: Communication orale

13. Kohn, J.J.: Global regularity for $\bar{\partial}$ on weakly pseudoconvex manifolds. T.A.M.S. **181**, 273–392 (1973)
14. Kohn, J.J.: Methods of Partial differential equations in complex analysis. Proceedings of symposia in pure mathematics. Vol. **30**, 215–237 (1977)
15. Matsumura, H.: Commutative algebra. New York: W.A. Benjamin C.O. 1970
16. Nagel, A.: Flatness criteria for modules of holomorphic functions over $\mathcal{O}_n$. Duke. Math. J. **40**, 433–448 (1973)
17. Nagel, A.: On algebras of holomorphic functions zith C^∞ boundary values. Duke. Math. J. **41**, 527–535 (1974)
18. Serre, J.P.: Un théorème de dualité. Comm. Math. Helv. **29**, 9–26 (1955)
19. Taylor, B.A.: Communication orale
20. Tomassini, G.: Sur les algèbres $A^0(\bar{D})$ et $A^\infty(\bar{D})$ d'un domaine pseudo-convexe non borné. Ann. Scuola Norm. Pisa. Serie IV, Vol. X, **2**, 243–256 (1983)
21. Tougeron, J.C.: Idéaux de fonctions différentiables. Berlin Heidelberg New York: Springer 1972

Reçu le 14 Mai 1984

Math. Z. 189, 449–456 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Characterizations of the Unit Ball B^n in Complex Euclidean Space

Ian Graham[1]* and H. Wu[2]**

[1] University of Toronto, Toronto, Canada M5S 1A1
[2] University of California at Berkeley, Berkeley, CA 94720, USA

The purpose of this note is to present some characterizations of the unit ball $B^n \subset \mathbb{C}^n$ in terms of the standard invariant norms. For the first theorem we need to recall two norms introduced in Eisenman [4]. Let M be an n-dimensional complex manifold, and let $\{w_1, \ldots, w_n\}$ be complex coordinates around a point $p \in M$. Then the norms E_n and C_n on $\Lambda^n TM$ (n-th exterior power of the holomorphic tangent bundle TM of M) are defined by:

$$E_n\left(p, \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right) = \inf\{|Jf(0)|^{-2}: \tag{1}$$
there exists a holomorphic map $f\colon B^n \to M$ satisfying $f(0)=p\},$

$$C_n\left(p, \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right) = \sup\{|Jf(p)|^2: \tag{2}$$
there exists a holomorphic map $f\colon M \to B^n$ satisfying $f(p)=0\},$

where Jf denotes the Jacobian determinant of f. These are the analogues of the Kobayashi-Royden and Carathéodory-Reiffen metrics, respectively. Since $\Lambda^n TM$ has a 1-dimensional fibre at each point of M, (1) and (2) uniquely define $E_n(p)$ and $C_n(p)$ at p for each $p \in M$. E_n is studied in detail in [5]. It is elementary that $C_n \leqq E_n$, but this inequality can be considerably sharpened: the following theorem shows that, if $C_n \not\equiv 0$, then either $C_n \equiv E_n$ (in which case M is biholomorphic to B^n) or the strict inequality $C_n < E_n$ prevails wherever $C_n \neq 0$.

By contrast the Carathéodory and Kobayashi metrics are identically equal on all convex domains in $\mathbb{C}^n$ [11].

Theorem 1. *Let M be an n-dimensional complex manifold. Suppose for some $p \in M$,*

$$E_n(p) = C_n(p) \neq 0. \tag{3}$$

Then M is biholomorphic to B^n.

* Partially supported by the Natural Sciences and Engineering Research Council of Canada
** Partially supported by the National Science Foundation

We emphasize that in Theorem 1 there are no assumptions on M except that of (3) at one point. When M is in addition a complete hyperbolic domain in $\mathbb{C}^n$, Theorem 1 was first proved by Bun Wong [13]. The assumption in Wong's theorem was subsequently relaxed to M being a bounded domain in $\mathbb{C}^n$ by Rosay [10] and later yet to M being a hyperbolic manifold by Dektyarev [3]. The starting point of our work was the discovery of the simple but decisive fact that a C^∞ map $f: M_1 \to M_2$ between C^∞ manifolds of the same dimension has a one-sided C^∞ inverse $g: M_2 \to M_1$ iff it has a two-sided C^∞ inverse, i.e. iff f is itself diffeomorphic. (This is the precise counterpart in differential topology of the well known fact in linear algebra that a linear map between vector spaces of the same finite dimension has a one-sided linear inverse iff it is an isomorphism.) This striking fact should be in every introductory text on manifolds because its proof is immediate. The proof of Theorem 1 is guided by this simple observation, although for technical reasons a more elaborate version (Lemma 1 below) is actually needed. This more elaborate version, it should be noted, is strictly for holomorphic maps.

It is worth pointing out that the reasoning leading to the proof of Theorem 1 yields at the same time a more general theorem:

Let M, N be complex manifolds of the same dimension n and let M be taut (in the sense of [14]*). Let $a \in M$ and $b \in N$ and consider the set of all holomorphic maps $f: M \to N$ and $g: N \to M$ such that $f(a)=b$ and $g(b)=a$. Suppose*

$$\sup (gf)^* E_n(a) = E_n(a)$$

when the supremum is taken over all such f and g. Then M and N are biholomorphic.

This theorem complements a similar theorem in Dektyarev [3] where he assumes instead that M and N are both hyperbolic. Since this theorem is only incidental to our purpose, we will not discuss it any further.

For the statement of the second theorem, it is necessary to recall the definition of the Kobayashi-Royden metric K and the Carathéodory-Reiffen metric C in order to fix the normalizing constants. Let $X \in T_p M$ (holomorphic tangent space to M at p). Then

$$K(p, X) = \inf \left\{ |\alpha| : \text{there exists a holomorphic mapping } f: D \to M \text{ such that } f(0)=p \text{ and } f_* \left(\alpha \frac{d}{dz}\right) = X \right\},$$

$$C(p, X) = \sup \left\{ |\alpha| : \text{there exists a holomorphic mapping } f: M \to D \text{ such that } f(p)=0 \text{ and } f_*(X) = \alpha \frac{d}{dz} \right\},$$

where D denotes the unit disc in $\mathbb{C}$. The respective *indicatrices* $I(p)$ and $J(p)$ are now defined by:

$$I(p)=\{X\in T_pM\colon K(p,X)<1\},$$
$$J(p)=\{X\in T_pM\colon C(p,X)<1\}.$$

Holomorphic mappings carry indicatrix into indicatrix, and further, we have normalized things so that the indicatrix of B^n at the origin 0 with respect to K or C is just B^n itself (the unit ball of the tangent space to B^n at 0 is canonically identified with B^n itself).

Theorem 2. *Let Ω be a (bounded) strictly convex domain in $\mathbb{C}^n$ in the sense that if $p, q\in\bar{\Omega}$, then the interior of the line segment joining p to q lies entirely in Ω. Suppose Ω satisfies one of the following conditions: for some $p\in\Omega$,*

(4) *there exists a holomorphic map $\tau\colon B^n\to\Omega$ such that $\tau(0)=p$ and $\tau_*(B^n)=I(p)$;*

(5) *there exists a holomorphic map $\sigma\colon\Omega\to B^n$ such that $\sigma(p)=0$ and $\sigma_*(J(p))=B^n$.*

Then Ω is biholomorphic to B^n.

In §4 we give a reformulation of Theorem 2 in terms of a relationship between the intrinsic norm E_n (respectively C_n) and the volume of $I(p)$ (respectively $J(p)$).

The proof of Theorem 2 relies on the recent results of Lempert [7] and Royden-Wong [11]. (See the note added in proof.)

We should also mention that C. Stanton [12] and G. Patrizio [9] have given characterizations of B^n which are in spirit related to Theorems 1 and 2.

§1. We Now Prove the Lemma Basic to the Proof of Theorem 1

Lemma 1. *Let M, N be complex manifolds of dimension n and let $f_i\colon M\to N$ and $g\colon N\to M$ be holomorphic mappings $(i=1,2,\ldots)$ such that for some points $a\in M$, $b\in N$, $f_i(a)=b$ for all i. Furthermore, suppose gf_i converges uniformly on compact sets to a (holomorphic) automorphism $\tau\colon M\to M$. Then g is biholomorphic.*

Proof. First observe that $g\colon N\to M$ must be surjective. Indeed letting $y\in M$, we shall produce a $p\in N$ such that $g(p)=y$. Let $\tau(x)=y$. Since gf_i converges uniformly to τ near x and since τ is open, Hurwitz's theorem implies that for all i sufficiently large, $y=gf_i(p_i)$ for some p_i near x. Fix such an i and let $p=f_i(p_i)$. Then $g(p)=y$.

It remains to show g is injective (injective holomorphic maps between manifolds of the same dimension have nowhere zero differential). The crux of the matter lies in proving:

(6) given $s\in N$, there exists a compact set $K\subset M$ such that for some $i_0\in\mathbb{Z}^+$, $s\in f_i(K)$ for all $i\geqq i_0$.

To prove this, let ξ be a continuous curve in N joining b to s. Let $\zeta=g(\xi)$ and let $c=g(b)$, $t=g(s)$. Thus ζ is a curve in M joining c to t. Choose an open set

W in M such that $\overline{W}$ is compact and $\zeta \subset W$. Finally let $V = \tau^{-1}(W)$. Then V is an open set in M and $\overline{V}$ is compact.

We claim that letting $K = \overline{V}$ would satisfy the assertion in (6). Suppose not. We define two positive constants ε_1 and ε_2 as follows. Let $\varepsilon_1 = \operatorname{dist}(\zeta, \partial W)$ and let ε_2 be a constant so small that any holomorphic map $h: M \to M$ which is pointwise ε_2-close to τ on $\overline{V}$ is necessarily biholomorphic on $\overline{V}$. The latter can be achieved because (1) a sufficiently small neighborhood of an imbedding in the C^1-topology consists of only imbeddings, and (2) uniform convergence of holomorphic mappings implies uniform convergence of all derivatives. Now let $\varepsilon = \min\{\frac{1}{4}\varepsilon_1, \varepsilon_2\}$. Now take i_0 so large that for all $i \geqq i_0$, gf_i is ε-close to τ on $\overline{V}$. Thus gf_i is a biholomorphic map from $\overline{V}$ to $gf_i(\overline{V})$. In particular, $g: f_i(\overline{V}) \to gf_i(\overline{V})$ is biholomorphic. By our assumption that (6) is false, we may pick i to be both bigger than i_0 and $s \notin f_i(V)$. However, $f_i(a) = b$, so the curve ξ joining b to s has the property that $\xi \cap f_i(V)$ does not have compact closure in $f_i(V)$. On the other hand $\zeta = g(\xi)$ by definition, and since the boundary of $gf_i(V)$ is within ε of $\partial W (\varepsilon \leqq \frac{1}{4}\varepsilon_1)$, ζ is a compact subset of $gf_i(V)$. Thus the biholomorphic map $g: f_i(V) \to gf_i(V)$ carries $\xi \cap f_i(V)$ onto a subset of ζ which has therefore compact closure in $gf_i(V)$. Contradiction.

It is now easy to prove that g is injective. Given $s_1, s_2 \in N$, (6) implies that there is a compact set $K \subset M$ and an $i_0 \in \mathbb{Z}^+$ such that $s_1, s_2 \in f_i(K)$ for all $i \geqq i_0$. We may assume i_0 is so big that $i \geqq i_0$ implies gf_i is sufficiently close to τ on K. For such an i, let $p_1, p_2 \in K$ such that $f_i(p_1) = s_1$, $f_i(p_2) = s_2$. Then $\tau(p_1) \neq \tau(p_2)$ implies that $g(s_1) \neq g(s_2)$. Q.E.D.

§ 2. Proof of Theorem 1

By the Definitions (1) and (2) of E_n and C_n, there exist a sequence of holomorphic maps $f_i: B^n \to M$ and a sequence of holomorphic maps $g_i: M \to B^n$ such that:

$$
\begin{aligned}
& f_i(0) = p \quad \text{for all } i \\
& |Jf_i(0)|^{-2} \to E_n\left(p, \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right) \\
& g_i(p) = 0 \quad \text{for all } i \\
& |Jg_i(0)|^2 \to C_n\left(p, \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right).
\end{aligned} \tag{7}
$$

Now B^n is complete hyperbolic, so a standard normal family argument shows that in fact the g_i's converge uniformly on compact sets to a holomorphic mapping $g: M \to B^n$ such that:

$$
\begin{aligned}
& g(p) = 0 \\
& |Jg(0)|^2 = C_n\left(p, \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right).
\end{aligned} \tag{8}
$$

By the hypothesis in (3) and using (7) and (8), we see that the sequence $gf_i: B^n \to B^n$ has the property that $gf_i(0)=0$ and $|J(gf_i)(0)| \to 1$ as $i \to \infty$. The same normal family argument again shows the existence of a holomorphic map $\tau: B^n \to B^n$ to which gf_i converges uniformly on compact sets; τ necessarily satisfies $\tau(0)=0$ and $|J\tau(0)|=1$. By the classical Cartan-Carathéodory theorem (see Bochner-Martin [2, Chap. III]), τ is an automorphism of B^n. Since $f_i(0)=p$ for all i, Lemma 1 shows that g is a biholomorphic map between M and B^n. Q.E.D.

§ 3. Proof of Theorem 2

For a strictly convex bounded domain Ω in $\mathbb{C}^n$ and $p \in \Omega$, Lempert [7] defined a $1-1$ onto map $\phi_p: \Omega \to B^n$. We begin by recalling some of the basic properties of this map.

For each nonzero $X \in T_p\Omega$ there exists a unique (holomorphic) extremal map $g_X: D \to \Omega$ such that $g_X(0)=p$ and $(g_X)_*\left(\alpha \frac{d}{dz}\right)=X$ with $\alpha=F(p, X)$. The mapping g_X has the additional property that it realizes the Kobayashi distance (in Ω) between any two points in its image (cf. (ii) below), and also the Kobayashi-Royden length of any vector tangent to its image. We let $S(X)$ denote the range of g_X; Lempert showed that $S(X) \cap S(Y)=\{p\}$ if $X, Y \in T_p\Omega$ are linearly independent over $\mathbb{C}$. The Lempert map $\phi_p: \Omega \to B^n$ is constructed by inverting extremal maps and has the following properties:

(i) ϕ_p is $1-1$ and onto, and $\phi_p(p)=0$;

(ii) If $0 \neq X \in T_p\Omega$ then we may regard X as a vector in T_0B^n by parallel translation. Let $D(X)$ denote the disc of radius one centered at $0 \in B^n$ and tangent to X, i.e.

$$D(X)=\left\{z \frac{X}{\|X\|}: z \in \mathbb{C}, |z|<1\right\}.$$

Then $\phi_p|_{S(X)}$ maps $S(X)$ to $D(X)$, is holomorphic (when so restricted), and is an isometry relative to the Kobayashi distance of Ω (equivalently: of $S(X)$) and of B^n (equivalently: of $D(X)$).

(iii) If $0 \neq X \in T_p\Omega$ then $(\phi_p)_*(X)=cX$ where $c>0$ (identifying $T_p\Omega$ and T_0B^n by parallel translation in $\mathbb{C}^n$; in fact $c=F(p, X)$).

We now proceed with the proof of Theorem 2 under the assumption (4). First we claim that the mapping $\tau: B^n \to \Omega$ realizes the Kobayashi-Royden metric at p in all directions. By this we mean the following: let $0 \neq X \in T_p\Omega$, and let $v \in T_0B^n$ be such that $\tau_*(v)=X$. Then $\|v\|=F(p, X)$ and furthermore $\tau|_{D(v)}$ may be regarded as a mapping from the unit disc $D \subset \mathbb{C}$ to Ω which realizes the Kobayashi length of X. This follows immediately from the assumption that $\tau_*(B^n)=I(p)$. More precisely, define the mapping $i_v: D \to B^n$ by $i_v(z)=z \frac{v}{\|v\|}$.

Then $\tau\circ i_v: D\to\Omega$ realizes the Kobayashi length of X and $(\tau\circ i_v)_*\left(\frac{d}{dz}\right)$ is a positive multiple of X, hence according to Lempert [7], $\tau\circ i_v$ must be the unique extremal map with these two properties. It follows that $\tau|_{D(v)}: D(v)\to S(X)$ is $1-1$ and onto. Since τ_* is nonsingular at p it follows that if $v_1, v_2\in T_0B^n$ are linearly independent on $\mathbb{C}$ then $X_1=\tau_*(v_1)$ and $X_2=\tau_*(v_2)$ are likewise linearly independent over $\mathbb{C}$. Since $S(X_1)\cap S(X_2)=\{p\}$ in this case we see that τ is globally $1-1$. Finally since τ_* is nonsingular at p and since $\bigcup_{x\in T_p\Omega} S(X)=\Omega$ it follows that τ is onto. Hence τ is a biholomorphic map.

We next prove Theorem 2 under assumption (5). If $0\neq v\in T_0B^n$ we denote by π_v the orthogonal projection of B^n onto $D(v)$. We first observe that σ must realize the Carathéodory length of any $X\in T_p\Omega$ in the sense that $\pi_{\sigma_*(X)}\circ\sigma$ may be regarded as a holomorphic mapping from Ω to D which realizes $C(p,X)$ and $C(p,X)=\|\sigma_*(X)\|$. This follows from the assumption that $\sigma_*(J(p))=B^n$.

Fixing $X\in T_p\Omega$, $X\neq 0$, we shall abbreviate $\pi_{\sigma_*(X)}$ to π. We next show

Lemma 2. *$\pi\circ\sigma|_{S(X)}$ maps $S(X)$ $1-1$ and onto $D(\sigma_*(X))$.*

Proof. The extremal mapping $g_X: D\to\Omega$ realizes the Kobayashi-Royden length of X. The mapping $\pi\circ\sigma: \Omega\to D(\sigma_*(X))$ realizes the Carathéodory length of X. Identifying $D(\sigma_*(X))$ with D, and making use of the recent result of Royden-P.M. Wong [11] that $K(p,X)=C(p,X)$ on a convex domain, we see that the composed mapping $\pi\circ\sigma\circ g_X: D\to D$ satisfies $\pi\circ\sigma\circ g_X(0)=0$ and $|(\pi\circ\sigma\circ g_X)'(0)|=1$. From the Schwarz lemma it follows that $\pi\circ\sigma\circ g_X(z)=\lambda z$ with $|\lambda|=1$. In particular $\pi\circ\sigma\circ g_X: D\to D$ is $1-1$ and onto. Since the same is true of $g_X: D\to S(X)$, the statement of the lemma follows.

Continuing with the proof of the theorem, we next show that $\sigma|_{S(X)}$ maps $S(X)$ $1-1$ and onto $D(\sigma_*(X))$ (without composing with π). This will imply that σ (unrestricted) is $1-1$ and onto. This time we shall assume that $D(\sigma_*(X))$ is the unit disc in the z_1 direction - this may be achieved via a unitary change of coordinates. We may suppose that $\sigma\circ g_X: D\to B^n$ is given in coordinates by $\sigma\circ g_X(z)=(f_1(z), \ldots, f_n(z))$. It suffices to show that $f_2\equiv\ldots\equiv f_n\equiv 0$. If this is not the case, then we may suppose without loss of generality that $f_2\not\equiv 0$. By the maximum principle there exist a sequence of points $\{\xi_j\}_{j=1}^{\infty}\subset D$ such that $|\xi_j|\to 1$ and a number $\varepsilon>0$ such that $|f_2(\xi_j)|\geqq\varepsilon$. However we know that $|f_1(\xi_j)|=|\xi_j|\to 1$, so for sufficiently large j this will contradict the fact that $|f_1(\xi_j)|^2+|f_2(\xi_j)|^2<1$. This completes the proof.

§ 4. Remarks

(i) Theorem 2 can be reformulated in terms of a relationship between the intrinsic norm E_n (respectively C_n) and the volume of $I(p)$ (respectively $J(p)$). First we note the following:

Lemma 3. *Let M be an n-dimensional complex manifold, let $p \in M$, and let $w_1, \ldots, w_n$ be local coordinates at p. Then*

$$E_n\left(p; \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right) \geqq \frac{\mathrm{vol}(B^n)}{\mathrm{vol}(I(p))}, \tag{9}$$

$$C_n\left(p; \frac{\partial}{\partial w_1} \wedge \ldots \wedge \frac{\partial}{\partial w_n}(p)\right) \leqq \frac{\mathrm{vol}(B_n)}{\mathrm{vol}(J(p))}. \tag{10}$$

(All volumes are Euclidean volumes.)

Proof. If $\tau: B^n \to M$ is a holomorphic map such that $\tau(0)=p$ then $\tau_*(B^n) \subset I(p)$. Hence $|J\tau(0)|^2 \,\mathrm{vol}(B^n) \leqq \mathrm{vol}(I(p))$. Equation (9) now follows on taking the appropriate infimum. Similarly if $\sigma: M \to B^n$ is a holomorphic such that $\sigma(p)=0$ then $\sigma_*(J(p)) \subset B^n$. This implies $|J\sigma(p)|^2 \,\mathrm{vol}(J(p)) \leqq \mathrm{vol}(B^n)$, from which (10) follows on taking the appropriate supremum.

The reformulated version of Theorem 2 is the following:

Theorem 2′. *Let Ω be a bounded strictly convex domain in $\mathbb{C}^n$ which satisfies one of the following conditions: for some $p \in \Omega$,*

(11) *equality holds in* (9)

(12) *equality holds in* (10).

Then Ω is biholomorphic to B^n.

It is easily seen that Theorems 2 and 2′ are equivalent. The mapping τ which occurs in the assumption (4) must satisfy $|J\tau(0)|^2 \,\mathrm{vol}(B^n) = \mathrm{vol}(I(p))$. Hence τ realizes the infimum in the definition of E_n and equality holds in (9). A similar argument applies to the mapping σ in assumption (5). Conversely if equality holds in (9) then the mapping τ which realizes the infimum in the definition of E_n (which exists because a convex domain containing no complex lines is taut, and even complete hyperbolic, by the result of Barth [1]) must satisfy $\mathrm{vol}(\tau_*(B_n)) = \mathrm{vol}(I(p))$, and hence from the fact that $\tau_*(B^n) \subset I(p)$ and the continuity of $K(p, X)$ on a taut domain, $\tau_*(B^n) = I(p)$. Similarly if equality holds in (10) then the mapping σ which realizes the supremum in the definition of C_n must satisfy assumption (5).

(ii) With reference to the assumptions (11) and (12), we make the following observation: suppose we are given a hermitian metric ds^2 on an n-dimensional complex manifold M. Let $p \in M$ and let $H(p)$ denote the indicatrix of ds^2 at p. Then in terms of local coordinates $w_1, \ldots, w_n$ at p the volume element at p is precisely

$$\frac{\mathrm{vol}(B^n)}{\mathrm{vol}(H(p))} \cdot \left(\frac{i}{2}\right)^n dw_1 \wedge d\bar{w}_1 \wedge \ldots \wedge dw_n \wedge d\bar{w}_n. \tag{13}$$

Equation (13) applies for example to the volume element of the Bergman metric of M (if M admits it). From Eq. (13) for the Bergman metric, Eq. (10), and the result of K.T. Hahn [6] and K.H. Look [8] to the effect that the

Carathéodory metric is majorized by the Bergman metric if M admits them, we can deduce that the Carathéodory volume form is majorized by the volume form of the Bergman metric.

(iii) We'd like to mention an open problem. Let Ω be a convex domain in $\mathbb{C}^n$, and let K^{2n} and C^{2n} be the differential volume forms constructed out of the $2n$-dimensional Hausdorff measures associated with the Kobayashi and Carathéodory distances, resp. In general,

$$E_n dV \geqq K^{2n} \geqq C^{2n} \geqq C_n dV$$

where dV denotes the euclidean volume form. Royden and Wong showed that $K^{2n} \equiv C^{2n}$ on such an Ω [11]. The theorem of B. Wong [13] already shows that if $E_n = C_n$ *at one point*, then Ω is necessarily biholomorphic to the ball. The natural question is if $E_n dV = K^{2n}$ or $C^{2n} = C_n dV$ (either at one point, or everywhere, or anything in between), what can be said about Ω?

References

1. Barth, T.J.: Convex domains and Kobayashi hyperbolicity. Proc. Am. Math. Soc. **79**, 556–558 (1980)
2. Bochner, S., Martin, W.T.: Several complex variables. Princeton, N.J.: Princeton Univ. Press 1948
3. Dektyarev, I.M.: Tests for equivalence of hyperbolic manifolds. Funkts. Anal. i Prilozh. **15**, 73–74 (1981), (English translation: Funct. Anal. Appl. **15**, 292–293 (1981) (1982)
4. Eisenman, D.A.: Intrinsic measures on complex manifolds and holomorphic mappings. Mem. Am. Math. Soc., No. 96, Am. Math. Soc., Providence 1970
5. Graham, I., Wu, H.: Some remarks on the intrinsic measures of Eisenman. To appear
6. Hahn, K.T.: Inequality between the Bergman metric and the Carathéodory differential metric. Proc. Am. Math. Soc. **68**, 193–194 (1978)
7. Lempert, L.: La métrique de Kobayashi et la representation des domaines sur la boule. Bull. Soc. Math. Fr. **109**, 427–474 (1981)
8. Look, K.H.: Schwarz lemma and analytic invariants. Sci. Sin. **7**, 435–504 (1958)
9. Patrizio, G.: Parabolic exhaustions for strictly convex domains, Thesis, University of Notre Dame, 1983
10. Rosay, J.-P.: Sur une caractérization de la boule parmi les domaines de $\mathbb{C}^n$ par son groupe d'automorphismes. Ann. Inst. Fourier **29**, 91–97 (1979)
11. Royden, H.L., Wong, P.M.: Carathéodory and Kobayashi metric on convex domains. To appear
12. Stanton, C.: A characterization of the ball by its intrinsic metrics. Math. Ann. **264**, 271–275 (1983)
13. Wong, B.: Characterization of the unit ball in $\mathbb{C}^n$ by its automorphism group. Invent. Math. **41**, 253–257 (1977)
14. Wu, H.: Normal families of holomorphic mappings. Acta Math. **119**, 193–233 (1967)

Received July 7, 1984; received in final form December 12, 1984

Added in Proof:

The equality of the Carathédory and Kobayashi metrics on a strictly convex domain, which is used in the proof of Theorem 2, was first proved in Lempert, L.: Holomorphic retracts and intrinsic metrics in convex domains. Anal. Math. **8**, 257–261 (1982).

The open problem mentioned above has been solved by J. Bland and the first author (to appear in Ann. Scuola Norm. Sup. Pisa).

Math. Z. 189, 457-463 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Framing Sphere Bundles over Spheres, the Smith Pairing, and Three-Fold Toda Brackets

George Al-Sabti

Department of Mathematics, College of Science, University of Basrah, Basrah, Iraq

Let M be an oriented stably parallelizable m-sphere bundle over S^n, $m<n$. In this paper, we compute products in $\tilde{\pi}_s^0(M)$, the zero-dimensional component of the stable cohomotopy ring of M. There is a relation between such products and three-fold Toda brackets (Theorem 1). We give three applications, the first of which is concerned with framings on sphere bundles over spheres. The second gives a formula for the deviation from linearity in the second variable of the Smith pairing defined in [6]. The final application is the representation of certain three-fold Toda brackets on framed sphere bundles over spheres. The middle element of some of these brackets does not belong to $\operatorname{Im} J$, the image of the classical J-homomorphism, and so cannot be gotten by the method of [5].

I would like to thank Professor L. Smith for his interest and encouragemant to write this work. I also would like to thank the Deutscher Akademischer Austauschdienst for a grant to spend some time in Göttingen where this paper was completed.

1. Preliminaries

Given any framed manifold (M, V) of dimension n, and any element $\alpha \in \tilde{K}O^{-1}(M)$, one can twist the framing V by α and the difference element is given by:

$$(M, V)^{\alpha} - (M, V) = V^* J(-\alpha) \tag{1}$$

where $J: \tilde{K}O^{-1}(M) \to \tilde{\pi}_s^0(M)$ is the generalized J-map, and $V^*: \tilde{\pi}_s^0(M) \to \pi_n^s$ is a homomorphism given by composition with the reduced fundamental class $\tilde{\sigma}_V: S^{n+k} \to S^k \wedge M$ (k is large). The J-map is natural and satisfies a quadratic formula:

$$J(\alpha+\beta) = J(\alpha) + J(\beta) + J(\alpha) J(\beta). \tag{2}$$

For details we refer to [2] and [9]. We shall think of elements of $\tilde{\pi}_s^0(M)$ as stable maps $M \to S^0$, and if $A, B \in \tilde{\pi}_s^0(M)$, then the product AB is represented by the stable composition,

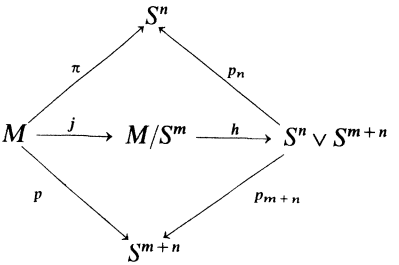
$$M \xrightarrow{\Delta} M \wedge M \xrightarrow{A \wedge B} S^0 \wedge S^0 \xrightarrow{\wedge} S^0 \tag{3}$$

where Δ is the diagonal map, and $\wedge$ is the smash product of the sphere spectrum.

Let M be an oriented stably parallelizable m-sphere bundle over S^n, $m<n$. Let $f: S^{n-1} \to SO(m+1)$ be a classifying map for M and $q: SO(m+1) \to S^m$ be the natural projection. Then M has the following cellular structure:

$$M \simeq S^m \bigcup_{qf} e^n \cup e^{m+n}. \tag{4}$$

The top cell of M splits off stably, and according to [3], there is a homotopy commutative diagram



(5)

where π is the sphere bundle projection, j is a map which collapses a fibre, p is a projection of M onto its top cell, h is a homotopy equivalence and p_n, p_{m+n} are the obvious projections.

2. Products in $\tilde{\pi}_s^0(M)$

The following theorem computes products in $\tilde{\pi}_s^0(M)$ and shows their relation to three-fold Toda brackets.

Theorem 1. *Let M be an oriented stably parallelizable m-sphere bundle over $S^n, m<n$. Let $i: S^m \to M$ be the inclusion of a fibre, and $\lambda_M = \sum^\infty qf$ as in* (1.4). *Then for any $A, B \in \tilde{\pi}_s^0(M)$, the products AB is given by:*

$$AB = X + Y$$

where X and Y are described by the following stable compositions:

$$X: M \xrightarrow{p} S^{m+n} \xrightarrow{T} S^0,$$

T is an element in the Toda bracket $\langle Ai, \lambda_M, Bi \rangle$.

$$Y: M \xrightarrow{\pi} S^n \xrightarrow{Y'} S^{2m} \xrightarrow{(Ai)(Bi)} S^0, \text{ for some } Y' \in \pi_{n-2m}^s.$$

Proof. Let $C_\theta = S^m \bigcup_\theta e^n$ be the mapping cone of $\theta = qf$, and let $a = A|C_\theta$, $b = B|C_\theta$. Then AB is represented by the composition of maps in the top row of the following commutative diagram:

$$\begin{array}{ccccccc}
M & \xrightarrow{\Delta} & M \wedge M & \xrightarrow{A \wedge B} & S^0 \wedge S^0 & \xrightarrow{\wedge} & S^0 \\
\downarrow{\scriptstyle j} & & \uparrow{\scriptstyle i \wedge i} & & \uparrow{\scriptstyle 1 \wedge b} & & \\
M/S^m & & & & & & \\
\downarrow{\scriptstyle h} & & & & & & \\
S^n \vee S^{m+n} & \xrightarrow{(\Delta_n, \Delta_{m+n})} & C_\theta \wedge C_\theta & \xrightarrow{a \wedge 1} & S^0 \wedge C_\theta & &
\end{array}$$

Since $m < n$, dimensional considerations imply that Δ factors up to homotopy through the projection j and the inclusion i of C_θ in M. Replacing M/S^m by the homotopy equivalent space $S^n \vee S^{m+n}$ produces well defined maps Δ_n and Δ_{m+n}. Therefore we have $AB = X + Y$, where X and Y are given by the stable compositions:

$$X\colon M \xrightarrow{p} S^{m+n} \xrightarrow{\Delta_{m+n}} C_\theta \wedge C_\theta \xrightarrow{a \wedge 1} S^0 \wedge C_\theta \xrightarrow{1 \wedge b} S^0 \wedge S^0 \xrightarrow{\wedge} S^0$$

$$Y\colon M \xrightarrow{\pi} S^n \xrightarrow{\Delta_n} C_\theta \wedge C_\theta \xrightarrow{a \wedge b} S^0 \wedge S^0 \xrightarrow{\wedge} S^0,$$

where we have replaced $p_{m+n} hj$ and $p_n hj$ by p and π respectively, according to (1.5).

Since $m < n$, Δ_n factors through the inclusion $i\colon S^m \to M$, to give the desired description for Y. It remains to show that the stable composition:

$$S^{m+n} \xrightarrow{\Delta_{m+n}} C_\theta \wedge C_\theta \xrightarrow{a \wedge 1} S^0 \wedge C_\theta \xrightarrow{1 \wedge b} S^0 \wedge S^0 \xrightarrow{\wedge} S^0,$$

can be written as a coextension of Ai composed with an extension of Bi to C_θ, so as to lie in the Toda bracket $\langle Ai, \lambda_M, Bi \rangle$.

Let $k\colon C_\theta \to S^n$ be the projection onto S^n, and consider the following diagram:

$$\begin{array}{ccccccc}
S^{m+n} & \xrightarrow{\Delta_{m+n}} & C_\theta \wedge C_\theta & \xrightarrow{a \wedge 1} & S^0 \wedge C_\theta & \xrightarrow{\wedge(1 \wedge b)} & S^0 \\
\downarrow{\scriptstyle \bar{\Delta}_{m+n}} & \cong & \downarrow{\scriptstyle 1 \wedge k} & = & \downarrow{\scriptstyle 1 \wedge k} & & \\
S^m \wedge S^n & \xrightarrow{i \wedge 1} & C_\theta \wedge S^n & \xrightarrow{a \wedge 1} & S^0 \wedge S^n & &
\end{array}$$

The left square commutes up to homotopy for dimensional reasons. Since Δ_{m+n} was induced from Δ, which induces the cup product in M, and since $H^*(M, \mathbb{Z})$

is an exterior algebra on two generators in dimensions m and n $(m<n)$, it follows that $\bar{\Delta}_{m+n}$ must be a map of degree ± 1. Therefore the composition $(a \wedge 1)\ \Delta_{m+n}$ can be taken as a coextension of $Ai=ai$. It is obvious that $\wedge(1 \wedge b)$ is an extension of $Bi=bi$. Remembering that $\lambda_M=\sum^{\infty}\theta$, gives the required result.

3. Applications

(3.1) Reframing sphere bundles over spheres.

We prove some results about reframing sphere bundles over spheres. We start with the following Lemma which is probably known but not documented.

Lemma 1. *Let (W,V) be a framed manifold of dimension n. Let $\alpha \in \tilde{K}O^{-1}(W)$ be an element that factors through S^k with $k>\frac{n}{2}$. Then*

$$(W,V)^{\alpha}=(W,V) \quad \textit{modulo } \operatorname{Im} J.$$

Proof. Suppose $\alpha: W \xrightarrow{\beta} S^k \xrightarrow{\gamma} SO$, $k>\frac{n}{2}$. According to (1.1), $(W,V)^{\alpha}-(W,V)$ $=J(-\alpha)\tilde{\sigma}_V$. Now $J(-\alpha)\tilde{\sigma}_V$ is represented by the stable composition:

$$S^n \xrightarrow{\tilde{\sigma}_V} W \xrightarrow{\beta} S^k \xrightarrow{J(-\gamma)} S^0,$$

which is a composition of an element in the image of J, namely $J(-\gamma)$, with $\beta\tilde{\sigma}_V \in \pi^s_{n-k}$. Since $k>\frac{n}{2}$, it follows from (7.6) of [4], that the above composition lies in $\operatorname{Im} J$.

Now let M be an oriented stably parallelizable m-sphere bundle over S^n with $m<n$. Let $i: S^m \to M$ be the inclusion of a fibre and $j: M \to M/S^m$ as before.

Let V be any framing of M, then Theorem 1 tells us that in general, the composition:

$$\tilde{K}O^{-1}(M) \xrightarrow{J} \tilde{\pi}^0_s(M) \xrightarrow{V^*} \pi^s_{m+n}$$

need not be a homomorphism. However, we have

Lemma 2. *The composition:*

$$\Phi: \tilde{K}O^{-1}(M/S^m) \xrightarrow{j^*} \tilde{K}O^{-1}(M) \xrightarrow{J} \tilde{\pi}^0_s(M) \xrightarrow{V^*} \pi^s_{m+n}$$

is a homomorphism. Moreover, $\operatorname{Im}\Phi=\operatorname{Im} J$.

As a consequence we see that $(M,V)^{\alpha}=(M,V)$ modulo $\operatorname{Im} J$, whenever the restriction of α to the fibre is trivial.

Proof. Since $\tilde{K}O^{-1}(M/S^m)=\tilde{K}O^{-1}(S^n)\oplus\tilde{K}O^{-1}(S^{m+n})$, then let $\alpha\in\tilde{K}O^{-1}(S^n)$, $\beta\in\tilde{K}O^{-1}(S^{m+n})$. From the definition, (1.2) and (1.5), we get

$$\begin{aligned}\Phi(\alpha+\beta)&=V^*J(\alpha\pi+\beta p)\\&=V^*J(\alpha\pi)+V^*J(\beta p)+V^*(J(\alpha\pi)J(\beta p)).\end{aligned}$$

Now since J is natural, and since the restrictions of $\alpha\pi$ and βp to S^m are trivial, it follows from Theorem 1, that the product $J(\alpha\pi)J(\beta p)$ is equal to zero and hence Φ is a homomorphism. Applying Lemma 1 gives $\operatorname{Im}\Phi=\operatorname{Im}J$.

We immediately have the following.

Corollary. *Let M be as above and suppose that M admits a framing V such that $(M,V)=0$. If the dimension m of the fibre equals 2, 4, 5, 6 mod 8, then $(M,F)\in\operatorname{Im}J$, for any framing F on M.*

We end our discussion by reproving Proposition 2.1 of [6] using our method.

Lemma 3. *Let (M,V) be as above and let $i: S^m\to M$ be the inclusion of a fibre. Then if $\alpha_1,\alpha_2\in\tilde{K}O^{-1}(M)$ are such that $\alpha_1 i=\alpha_2 i$, we get*

$$(M,V)^{\alpha_1}=(M,V)^{\alpha_2}\quad \textit{modulo } \operatorname{Im}J$$

Proof. Write $\alpha_1=\alpha_2+\beta$, for some $\beta\in\tilde{K}O^{-1}(M/S^m)$, and observe $V^*J(-\alpha_2)\cdot J(-\beta)$ is a product of two elements in $\operatorname{Im}J$ in different stems (remember $m<n$), and hence lies in $\operatorname{Im}J$. The result follows from Lemma 2.

(3.2) The Smith pairing

In [6], a pairing

$$\mathscr{S}\colon \pi_n(V^{m+1}(\mathbb{R}^\infty))\times\pi_m(SO)\to\pi^s_{m+n}/\operatorname{Im}J$$

was defined for $n-m\neq 2$, as follows:

$$\mathscr{S}(\xi,\alpha)=(S(\xi),\Phi)^{\hat{\alpha}}$$

where ξ is an $(m+1)$-dimensional oriented vector bundle over S^n, with a stable trivialization which induces a framing Φ on the sphere bundle $S(\xi)$ such that $(S(\xi),\Phi)=0$, and $\hat{\alpha}$ is any extension of α to $S(\xi)$. The above pairing is not linear in the second variable (see [7], Appendix 1), and we are going to describe the deviation from linearity in the second variable.

Let $D_\xi(\alpha,\beta)=\mathscr{S}(\xi,\alpha+\beta)-\mathscr{S}(\xi,\alpha)-\mathscr{S}(\xi,\beta)$

Theorem 2.

$$D_\xi(\alpha,\beta)\in\begin{cases}\pm\langle J(-\alpha),\lambda_\xi,J(-\beta)\rangle, & m\neq 1,3,7\\ \pm\langle J(-\alpha),\lambda_\xi,J(-\beta)\rangle+D_m, & m=1,3,7\end{cases}$$

where D_m is an element in $\eta^3\circ\pi^s_{n-2}$, $\nu^3\circ\pi^s_{n-6}$, $\sigma^3\circ\pi^s_{n-14}$, for $m=1,3,7$ respectively, and $\lambda_\xi=\sum^\infty\theta_\xi$, $\theta_\xi\colon S^{n-1}\xrightarrow{\xi}SO(m+1)\xrightarrow{q}S^m$.

Proof. From (1.1) and (1.2), we get

$$D_\xi(\alpha,\beta)=\Phi^*(J(-\hat{\alpha})J(-\hat{\beta})),$$

so Theorem 1 and naturality of J imply that,

$$\begin{aligned} D_\xi(\alpha,\beta)&=\Phi^*(X+Y)\\ &=X\tilde{\sigma}_\Phi+Y\tilde{\sigma}_\Phi \end{aligned}$$

Since $p\tilde{\sigma}_\Phi=\pm 1$, $\pi\tilde{\sigma}_\Phi$ is an element in π^s_m and products of elements in $\operatorname{Im}J\subset\pi^s_m$ vanish except when $m=1,3,7$; The result follows.

(3.3) Representing some Toda brackets on sphere bundles over spheres.

Theorem 1 and the results of (3.1) enable us to represent the following Toda brackets on framed sphere bundles over spheres.

(i) $\bar{\nu}=\langle\nu,\eta,\nu\rangle\in\pi^s_8=\underset{(\bar{\nu})}{\mathbf{Z}_2}\oplus\underset{(\varepsilon)}{\mathbf{Z}_2}$

(ii) $\beta_1=\langle\nu,2\nu,\nu\rangle\in\pi^s_{10}=\underset{(\eta\mu)}{\mathbf{Z}_2}\oplus\underset{(\beta_1)}{\mathbf{Z}_3}$

(iii) $\eta\kappa=\langle\nu,\varepsilon,\nu\rangle\in\pi^s_{15}=\underset{(\rho)}{\mathbf{Z}_{480}}\oplus\underset{(\eta\kappa)}{\mathbf{Z}_2}$

(iv) $\nu^*=-\langle\sigma,\nu,\sigma\rangle\in\pi^s_{18}=\underset{(\eta\bar{\mu})}{\mathbf{Z}_2}\oplus\underset{(\nu^*)}{\mathbf{Z}_8}$

(v) $\langle\sigma,\bar{\kappa},\sigma\rangle\in{}_2\pi^s_{35}=\underset{(J_{35})}{\mathbf{Z}_8}\oplus\underset{(\langle\sigma,\bar{\kappa},\sigma\rangle)}{\mathbf{Z}_2}\oplus\underset{(\eta\kappa\bar{\kappa})}{\mathbf{Z}_2}$

We have used the notation of [8] and [1], for ${}_2\pi^s_{35}$.

Remark. Kosinski's method [5], fails to represent the brackets $\langle\nu,\varepsilon,\nu\rangle$ and $\langle\sigma,\bar{\kappa},\sigma\rangle$, because the middle elements do not lie in $\operatorname{Im}J$. Moreover, although $\nu\in\operatorname{Im}J$; the bracket $\langle\sigma,\nu,\sigma\rangle$ does not fulfill the conditions of [5].

Our idea to represent the above brackets is to choose a sphere bundle where the fibre has the dimension of the first element in the bracket. The bundle is classified by a map $S^{n-1}\xrightarrow{f}SO(m+1)$ such that qf equals the desuspension of the middle map of the bracket. The bounding framing is then twisted by an iteration of a certain map and Theorem 1 gives the Toda bracket.

To illustrate the method we present the details for the case of $\eta\kappa$. Let M be a 3-sphere bundle over S^{12}, classified by the composition, $S^{11}\xrightarrow{\varepsilon_3}S^3\xrightarrow{s}SO(4)$, where s is a map given by $s(x)y=xy$ (quaternionic multiplication). The bundle structure is stably trivial and hence M admits a framing V such that $(M,V)=0$, [6]. Let $\alpha:S^3\to SO$ be such that $J(-\alpha)=\nu$ and let $\hat{\alpha}$ be an extension of α to M.

We have $(M,V)^{\hat{\alpha}}=m\rho+n\eta\kappa$, for some integers m,n. Then (1.1), (1.2) and Theorem 1, give

$$(M,V)^{2\hat{\alpha}}=2m\rho+\langle\nu,\varepsilon,\nu\rangle$$

Since $2m\rho\in\operatorname{Im}J$, the results of (3.1) say that there exists $\beta\in\tilde{K}O^{-1}(M)$ such that

$$(M,V)^{2\hat{\alpha}+\beta}=\langle\nu,\varepsilon,\nu\rangle.$$

References

1. Barratt, M.G., Mahowald, M., Tangora, M.: Some Differentials in the Adams Spectral Sequence II. Topology **9**, 309–316 (1970)
2. Gershenson, H.H.: On Framed Cobordism Classes Representable on a Fixed Manifold. Math. Z. **122**, 189–202 (1971)
3. James, I.M., Whitehead, J.H.C.: The Homotopy Theory of Sphere Bundles over Spheres I, II. Proc. Lond. Math. Soc. **4**, 196–218 (1954); **5**, 148–166 (1955)
4. Kosinski, A.: Inertia Groups of π-Manifolds. Am. J. Math. **89**, 227–248 (1967)
5. Kosinski, A.: Toda Brackets in Differential Topology. Comment. Math. Helv. **46**, 113–123 (1971)
6. Smith, L.: Framings of Sphere Bundles over Spheres, the Plumbing Pairing, and the Framed Bordism Classes of Rank Two Simple Lie Groups. Topology **13**, 401–415 (1974)
7. Steer, B.: Orbits and the homotopy class of a Classical Map. Topology **15**, 383–393 (1976)
8. Toda, H.: Composition Methods in Homotopy Groups of Spheres. Princeton; University Press 1962
9. Wood, R.M.W.: Framing the Exceptional Lie Group G_2. Topology **15**, 303–320 (1976)

Received June 20, 1984, received in final form October 10, 1984

Math. Z. 189, 465–474 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

Uniform Estimate for the $\bar{\partial}$-Problem in a Class of Branched Covering of Strictly Pseudoconvex Domains

Carla Parrini and Angela Selvaggi Primicerio
Istituto Matematico "U. Dini", Viale Morgagni 67/a, I-50134 Firenze, Italy

Introduction

It is well known that if D is a strictly pseudo-convex domain in $\mathbb{C}^n$ with smooth boundary and if g is an uniformly bounded $\bar{\partial}$-closed $C^1(0,1)$-form on D, then there exists an uniformly bounded function α on D such that $\bar{\partial}\alpha = g$.

In general this is not true for a pseudoconvex domain, see Sibony's counter-example [14].

In this paper we solve the $\bar{\partial}$-problem with an uniform estimate in a class of bounded pseudoconvex domain D which can be mapped by a proper holomorphic mapping onto a strictly pseudoconvex domain $\Omega \subset \mathbb{C}^n$. We construct an integral representation formula for $(0,q)$ forms on D and we prove that if $g \in C^1_{(0,1)}(\bar{D})$ and is $\bar{\partial}$-closed in D, then the equation $\bar{\partial}\alpha = g$ has a solution $\alpha(z)$ with the estimate

$$\|\alpha(z)\|_{C(D)} \leqq c(D)\|g\|$$

where $\| \ \|_{C(D)}$ is the suprenum norm, $\|g\| = \max_i \|g_i\|_{C(D)}$ and g_i are the coefficients of the form g.

The representation formula is obtained in section 1, using an appropriate pull back of the Henkin function of Ω. In Section 2 we study the boundary behaviour of the covering mapping f, and we find a fundamental inequality (2.4) between the values of the map $f(z)-f(\zeta)$ and its zero set. This boundary regularity allows us to estimate the solution of the $\bar{\partial}$-equation (Theorem 3.1).

1. Integral Representation Formula for $(0,q)$ Forms on Branched Covering of Strictly Pseudoconvex Domains

Let Ω and D be bounded domains of $\mathbb{C}^n$ such that:

$$\Omega = \{w \in \mathbb{C}^n : \tilde{\lambda}(w) < 0\} \tag{1.1}$$

where $\tilde{\lambda}$ is a real valued function which is C^∞ and strictly plurisubharmonic in a neighborhood of the boundary, $b\Omega$, of Ω, $d\tilde{\lambda}_{|b\Omega} \neq 0$, and

$$D = \{z \in \mathbb{C}^n : \lambda(z) = \tilde{\lambda} \circ f(z) < 0\} \tag{1.2}$$

with $d\lambda_{|bD} \neq 0$ and $f : D \to \Omega$ is a proper holomorphic mapping $C^\infty(\bar{D})$.

We assume $\bar{D}\in\mathscr{F}$[1] where $\mathscr{F}$ is defined as follows

Definition 1.1. Let $\mathscr{F}$ be the family of compact sets $K\subset\mathbb{C}^n$ with the following properties:

i) $K=\bigcap_j \Omega_j$, Ω_j pseudoconvex domains of $\mathbb{C}^n$,

ii) there exists a constant $p\geqq 1$ such that, for every $0<\varepsilon<\frac{1}{2}$ it is possible to find Ω_j with

$$\{z\in\mathbb{C}^n: d(z,K)<\varepsilon^p\}\subset\Omega_j\subset\{z\in\mathbb{C}^n: d(z,K)<\varepsilon\}.$$

$d(z,A)$ is the euclidean distance of a point z from a set A.

Examples of domains whose clousure belongs to $\mathscr{F}$ are convex sets, uniformly H-convex domains and hence [5] C^∞ regular pseudoconvex domains.

Note that D is pseudoconvex and if $u=\det(f')$ is the determinant of the jacobian matrix of f, then $\{z\in\mathbb{C}^n: u(z)=0\}\cap bD$ is the set in which there is not strict pseudoconvexity.

We construct an integral representation formula for $(0,q)$ forms in D, using a pullback of the Henkin function of Ω. This pullback was used in [4] to get an integral representation formula for functions in $A(D)=H(D)\cap C(\bar{D})$, where D is as in (1.2) but the function f is holomorphic in a neighborhood of D.

For reader's convenience we summarize in the following lemma some properties of a strictly pseudoconvex domain Ω.

Let

$$V_\delta=\{w\in\mathbb{C}^n: |\lambda(w)|<\delta\},\qquad \Omega_\delta=V_\delta\cup\Omega,$$

$$U_{\varepsilon,\delta}=\{(v,w)\in\mathbb{C}^n\times\mathbb{C}^n: v\in V_\delta,\ w\in\Omega_\delta \text{ and } |v-w|<\varepsilon\}$$

$C^\infty(V,H(\Omega))$ the space of C^∞-functions on V with values in the space $H(\Omega)$ of holomorphic functions in Ω.

Lemma 1.2 ([7], [10]). *There exist positive constants γ, ε, δ, a function $\Phi\in C^\infty\cdot(V_\delta,H(\Omega_\delta))$ and functions F, $H\in C^\infty(U_{\varepsilon,\delta})$ such that*

a) $\Phi=F\cdot H$ *on* $U_{\varepsilon,\delta}$; $|H|>\gamma$ *on* $U_{\varepsilon,\delta}$; $|\Phi|>\gamma$ *outside* $U_{\varepsilon,\delta}$;

b) $\operatorname{Re} F(v,w)>(\lambda(v)-\lambda(w))+\gamma|v-w|^2$ *on* $U_{\varepsilon,\delta}$;

c) $d_v F(v,w)_{|v=w}=-d_w F(v,w)_{|v=w}=2\partial_w\lambda(w)$;

d) $\Phi(v,w)=\sum_{i=1}^{n}(v_i-w_i)P_i(v,w)$, $\qquad P_i\in C^\infty(V_\delta,H(_\delta))$.

Let $f=(f_1,f_2,\ldots,f_n)$. Any $f_i\in A^\infty(D)=C^\infty(\bar{D})\cap H(D)$. As $\bar{D}\in\mathscr{F}$ it is possible to factorize any function belonging to $A^\infty(D)$, according to the following

Lemma 1.3 ([9]). *Let $D\subset\mathbb{C}^n$ be a bounded pseudoconvex domain with regular C^2-boundary. Assume $\bar{D}\in\mathscr{F}$. Then for any $h\in A^\infty(D)$, there exist $h_i\in A^\infty(D\times D)$, $i=1,2,\ldots,n$, such that*

$$h(\zeta)-h(z)=\sum_{i=1}^{n}(\zeta_i-z_i)\,h_i(\zeta,z)\qquad \zeta,z\in\bar{D}.$$

[1] For the possibility of releasing this assumption see the final acknowledgement.

Thus $f_i(\zeta)-f_i(z)=\sum_{j=1}^{n}(\zeta_j-z_j)f_{ij}(\zeta,z)$ $i=1,\ldots,n$ $(\zeta,z)\in\bar{D}\times\bar{D}$, $f_{ij}\in A^\infty(D\times D)$.
Let us consider

$$P(\zeta,z)=\Phi(f(\zeta),f(z))=\sum_{j=1}^{n}(\zeta_j-z_j)\,p_j(\zeta,z) \tag{1.3}$$

where

$$p_j(\zeta,z)=\sum_{i=1}^{n}f_{ij}(\zeta,z)\,p_i(f(\zeta),f(z)).$$

The application $\left(\frac{p_1(\zeta,z)}{P(\zeta,z)},\ldots,\frac{p_n(\zeta,z)}{P(\zeta,z)}\right)$ is a C^∞ map from a suitable neighborhood W of $(bD\times\bar{D})\setminus\{(\zeta,z)\in\bar{D}\times\bar{D}\colon f(\zeta)=f(z)\}$ in $\mathbb{C}^n\times\mathbb{C}^n$.

Let

$$\Delta=\{t=(t_1,t_2)\in\mathbb{R}^2\colon t_1+t_2=1,\ t_1+t_2\geqq 0\}$$

with orientation so that $b\Delta=\Delta_1-\Delta_2$, $\Delta_i=\{t\in\Delta\colon t_i=1\}$, and

$$\mu_i=t_1\frac{\bar{\zeta}_i-\bar{z}_i}{|\zeta-z|^2}+t_2\frac{p_i(\zeta,z)}{P(\zeta,z)}\qquad i=1,2,\ldots,n,\ t\in\Delta.$$

For $0\leqq q\leqq n-1$ it is possible to define in a standard way the double differential form

$$\Omega_{n,q}(t_1,t_2,\zeta,z)=(-1)^q\binom{n-1}{q}\det\Big(\mu,\underbrace{\bar{\partial}_z\mu,\ldots,\bar{\partial}_z\mu}_{q\text{-times}},\underbrace{\bar{\partial}_{\zeta,t}\mu,\ldots,\bar{\partial}_{\zeta,t}\mu}_{(n-q-1)\text{-times}}\Big)\wedge d\zeta$$

where $\mu=(\mu_1,\ldots,\mu_n)$, $d\zeta=d\zeta_1\wedge d\zeta_2\wedge\ldots\wedge d\zeta_n$, $\bar{\partial}_{\zeta,t}=\bar{\partial}_\zeta+d_t$, and set $\Omega_{n,n}=\Omega_{n,-1}=0$.

Note that

i) $\Omega_{n,q}(1,0,\zeta,z)=q$-th Bochner-Martinelli-Koppelman Kernel

ii) $\Omega_{n,q}=0$ on $bD\times\Delta_2$, if $q>0$.

Following Range-Siu [13] one can see that

iii) $\bar{\partial}_{\zeta,t}\Omega_{n,q}=(-1)^q\bar{\partial}_z\Omega_{n,q-1}$.

Then apply Stokes' theorem to the differential form $d_{\zeta,t}(g\wedge\Omega_{n,q})$ on $bD\times\Delta$, $g\in C^1(\bar{D})$, and use iii) and the Bochner-Martinelli-Koppelman formula to define operators

$$T_0\colon C_{(0,0)}(bD)\to H(D)$$
$$T_q\colon C_{(0,q)}(\bar{D})\to C_{(0,q-1)}(D)\qquad 1\leqq q\leqq n$$

as follows

$$T_0g=c_n\int_{bD\times\Delta_2}g\Omega_{n,0},\qquad c_n=(-1)^n\frac{(n-1)/2}{(2\pi i)^n}$$
$$T_qg=-c_n\Big(\int_{bD\times\Delta}g\wedge\Omega_{n,q-1}+\int_{D\times\Delta_1}g\wedge\Omega_{n,q-1}\Big).$$

Thus we have

Theorem 1.4. *Let D be as in (1.2) a branched cover of a bounded strictly pseudoconvex domain in $\mathbb{C}^n$. Assume $\bar{D} \in \mathcal{F}$. Then*

$$\begin{aligned} g &= (\bar{\partial} T_q + T_{q+1} \bar{\partial})\, g \quad \text{for } g \in C^1_{(0,q)}(\bar{D}), \quad 1 \leqq q \leqq n \\ g &= (T_0 + T_1 \bar{\partial})\, g \quad \text{for } g \in C^1_{(0,0)}(\bar{D}). \end{aligned} \tag{1.4}$$

2. A Fundamental Inequality for Proper Holomorphic Mapping

First, we introduce some notations. We follow ([15]), appendix V) and we refer to it for more details.

Definition 2.1. The symmetric m^{th}-power, $(\mathbb{C}^n)^m_{\text{sym}}$, of $\mathbb{C}^n$ is the space of equivalence classes in $\mathbb{C}^n \dots \mathbb{C}^n$, m-times, under the equivalence relation

$$(x^1, \dots, x^m) = (y^1, \dots, y^m) \leftrightarrow y^i = x^{p(i)}$$
$$\text{for some permutation } p \text{ of } (1, \dots, m).$$

An element of $(\mathbb{C}^n)^m_{\text{sym}}$ will be denoted by $\zeta = \langle \zeta^1, \dots, \zeta^m \rangle$, $\zeta^i = (\zeta^i_1, \dots, \zeta^i_n) \in \mathbb{C}^n$. $(\mathbb{C}^n)^m_{\text{sym}}$ is a metric space with distance given by

$$\operatorname{dist}(\langle x^1, \dots, x^m \rangle, \langle y^1, \dots, y^m \rangle) = \inf_p \sup_i \text{ euclidean dist } (x^i, y^{p(i)}),$$

and there exists a homeomorphism $\Psi^{(m)} : (\mathbb{C}^n)^m_{\text{sym}} \to \mathbb{C}^N$, $N = \binom{n+m}{n} - 1$, whose components $\Psi^{(m)}_\eta$, (η multiindex $\eta = \eta_1, \dots, \eta_n$, $|\eta| \geqq 1$), are the coefficients of the polynomial in the $n+1$ variables z, $u_i \in \mathbb{C}$, $i = 1, \dots, n$,

$$\begin{aligned} P_\zeta(u, z) &= (z - u \cdot \zeta^1) \dots (z - u \cdot \zeta^m) \\ &= \sum_{k=0}^{m} (-1)^k z^{m-k} \sum_{|\eta| = k} \Psi^{(m)}_\eta(\zeta)\, u_1^{\eta_1} \dots u_n^{\eta_n} \end{aligned} \tag{2.1}$$

where $\zeta \in (\mathbb{C}^n)^m_{\text{sym}}$, $u = (u_1, \dots, u_n)$ is a covector in $\mathbb{C}^n$ and $u \cdot \zeta^i$ means $u_1 \zeta^i_1 + \dots + u_n \zeta^i_n$.

In other words

$$\Psi^{(m)}_\eta(\zeta) = \sum_{\lambda, \gamma} \zeta^{\lambda_1}_{\gamma_1} \dots \zeta^{\lambda_k}_{\gamma_k} \qquad (|\eta| = k)$$

the sum being over all multiindex λ, γ such that

i) $1 \leqq \lambda_1 < \lambda_2 < \dots < \lambda_k \leqq m$

ii) for each $j (1 \leqq j \leqq n)$ the number j occurs exactly η_j times in the expression $(\gamma_1, \dots, \gamma_k)$.

For example, in the case $(\mathbb{C}^2)^3_{\text{sym}}$ with $\zeta = \langle (\zeta^1_1, \zeta^1_2), (\zeta^2_1, \zeta^2_2), (\zeta^3_1, \zeta^3_2) \rangle$ we have

$$\begin{aligned} &\Psi^{(3)}_{10} = \zeta^1_1 + \zeta^2_1 + \zeta^3_1, \ \Psi^{(3)}_{01} = \zeta^1_2 + \zeta^2_2 + \zeta^3_2, \\ &\Psi^{(3)}_{20} = \zeta^1_1 \zeta^2_1 + \zeta^1_1 \zeta^3_1 + \zeta^2_1 \zeta^3_1, \ \Psi^{(3)}_{11} = \zeta^1_1 \zeta^2_2 + \zeta^1_1 \zeta^3_2 + \zeta^2_1 \zeta^3_2 \\ &\qquad + \zeta^2_1 \zeta^1_2 + \zeta^3_1 \zeta^2_2 + \zeta^3_1 \zeta^1_2, \text{ etc.} \dots \end{aligned}$$

For these coefficients we can prove the following relation

Lemma 2.2. *There exists a positive constant c such that, for any $\zeta = \langle \zeta^1, \dots, \zeta^m \rangle \in (\mathbb{C}^n)^m_{\mathrm{sym}}$ one has*

$$|\zeta^1|^2 \dots |\zeta^m|^2 \leqq c \sum_{|\eta|=m} |\Psi^{(m)}_\eta(\zeta)|^2. \tag{2.2}$$

Proof. With no loss of generality, we prove the Lemma in $(\mathbb{C}^2)^m_{\mathrm{sym}}$ by induction on $m \geqq 2$.

If $m=2$, we have

$$\begin{aligned} |\zeta^1|^2 |\zeta^2|^2 &\leqq 2[|\zeta^1_1 \zeta^2_1|^2 + |\zeta^1_2 \zeta^2_2|^2 + |\zeta^1_1 \zeta^2_2 + \zeta^1_2 \zeta^2_1|^2] \\ &= 2[|\Psi^{(2)}_{20}(\zeta)|^2 + |\Psi^{(2)}_{02}(\zeta)|^2 + |\Psi^{(2)}_{11}(\zeta)|^2] \end{aligned}$$

because

$$\begin{aligned} |\zeta^1_1 \zeta^2_2|^2 + |\zeta^1_2 \zeta^2_1|^2 &\leqq |\zeta^1_1 \zeta^2_1|^2 + |\zeta^1_2 \zeta^2_2|^2 + |\zeta^1_1 \zeta^2_2 + \zeta^1_2 \zeta^2_1|^2 \\ &= |\zeta^1_1 \zeta^2_2|^2 + |\zeta^1_2 \zeta^2_1|^2 + |\zeta^1_1 \bar{\zeta}^2_1 + \zeta^1_2 \bar{\zeta}^2_2|^2. \end{aligned}$$

Fix $\zeta = \langle \zeta^1, \dots, \zeta^m \rangle$ and consider a rotation σ in $\mathbb{C}^2$, given by a matrix $M = (a_{ij})$, $a_{ij} \in \mathbb{C}$, such that $\sigma(\zeta^m) = \xi^m = (0, \xi^m_2)$ where ξ_1, ξ_2 are the new coordinates in $\mathbb{C}^2$. Let $\xi = \langle \xi^1, \dots, \xi^m \rangle$ and $\hat{\xi} = \langle \xi^1, \dots, \xi^{m-1} \rangle$ where $\xi^i = M(\zeta^i)$, $i = 1, \dots, m$.

First it is easy to see that

$$\Psi^{(m)}_{m0}(\xi) = 0, \qquad \Psi^{(m)}_{i(m-1)}(\xi) = \xi^m_2 \Psi^{(m-1)}_{i(m-i-1)}(\hat{\xi}) \quad i = 1, \dots, m-1. \tag{2.3}$$

Moreover, the rotation induces a transformation $\tilde{\sigma}$ on $\mathbb{C}^N$. As

$$P_\zeta(uM, z) = P_\xi(u, z)$$

i.e.

$$\begin{aligned} &\sum_{k=0}^{m} (-1)^k z^{m-k} \sum_{|\lambda|=k} \Psi^{(m)}_\lambda(\zeta) (uM)^{\lambda_1}_1 (uM)^{\lambda_2}_2 \\ &\quad = \sum_{k=0}^{m} (-1)^k z^{m-k} \sum_{|\mu|=k} \Psi^{(m)}_\mu(\xi) u^{\mu_1}_1 u^{\mu_2}_2 \end{aligned}$$

the identity in z gives

$$\sum_{|\mu|=k} \Psi^{(m)}_\mu(\xi) u^{\mu_1}_1 u^{\mu_2}_2 = \sum_{|\lambda|=k} \Psi^{(m)}_\lambda(\zeta) \left(\sum_{i=1}^{2} a_{i1} u_i \right)^{\lambda_1} \left(\sum_{i=1}^{2} a_{i2} u_i \right)^{\lambda_2} \qquad k = 1, \dots, m$$

and the comparison of the coefficients of the above polynomials shows that $\tilde{\sigma}$ is a linear map, whose norm is bounded by a constant independent of a_{ij} as M is an unitary matrix and thus $|a_{ij}| \leqq 1$.

Using the induction hypothesis of step $m-1$, (2.3), and the boundedness of $\|\tilde{\sigma}\|$ respectively, we have

$$\begin{aligned} |\zeta^1|^2 \dots |\zeta^m|^2 = |\xi^m_2|^2 |\xi^1|^2 \dots |\xi^{m-1}|^2 &\leqq c |\xi^m_2|^2 \sum_{|\mu|=m-1} |\Psi^{(m-1)}_\mu(\hat{\xi})|^2 \\ = c \sum_{|\mu|=m} |\Psi^{(m)}_\mu(\xi)|^2 &\leqq c \sum_{|\mu|=m} |\Psi^{(m)}_\mu(\zeta)|^2. \end{aligned}$$

Here and from now on, we denote different constants by the same symbol c.

We are now able to prove the following

Theorem 2.3. *Let $D\in\mathbb{C}^n$ be a bounded pseudoconvex m-sheeted covering of a strictly pseudoconvex bounded domain $\Omega\in\mathbb{C}^n$. Let $f: D\to\Omega$ be the holomorphic covering mapping. Assume $f\in C^\infty(\bar{D})$. Then, for any $\zeta\in\bar{D}$*

$$|f(\zeta)-f(z)|^2\geqq c\Big(\prod_{i=1,\dots,k}|\zeta^i-z|^{2\beta_i}\Big),\qquad z\in\bar{D} \tag{2.4}$$

where $\{\zeta^1=\zeta,\zeta^2,\dots,\zeta^k\}$, $1\leqq k\leqq m$ is the zero set of the mapping $f(x)-f(\zeta)$, β_i is the branching order of ζ^i and $c=c(f,\Omega)$.

Before proving the theorem we state some properties of the holomorphic map f.

Lemma 2.4. *Let D, Ω, f be as in the above Theorem.*

a) *Let $h\in A^\infty(D)$. Let $F_1, F_2,\dots,F_m$ be the m local inverses to f defined on $\Omega\setminus\{f(z): u(z)=0\}$. Then, every symmetric function of $h\circ F_1,\dots,h\circ F_m$ is $A^\infty(\Omega)$.*

b) *The function u vanishes at most to finite order at any boundary point of D.*

c) *If $m=\#(w)$ is the number of points in $f^{-1}(w)$, $w\in\Omega\setminus\{f(z): u(z)=0\}$, then $\#(w^0)\leqq m$ for any $w^0\in b\Omega$.*

Proof. Let S_r be the r-symmetric function of $h\circ F_1,\dots,h\circ F_m$ and $H_r=\sum_{k=1}^{m}(h\circ F_k)^r$. By the Riemann removable singularity theorem H_r extends to a holomorphic function in Ω. From Newton's identity

$$H_r+H_{r-1}S_1+\dots+H_1S_{r-1}+rS_r=0$$

is then enough to prove that $H_r\in C^\infty(\bar{\Omega})$. As Ω is strongly pseudoconvex this result follows from Remark B of [2] applied to our H_r. This proves a) and hence, in particular it follows that every symmetric function of $F_1,\dots,F_m$ is $A^\infty(\Omega)$.

The proof af b) can be found in [3] or [6].

To prove c) we can use a slight variation of proposition 1 of [1]. Let $w^0\in b\Omega$. Assume that $\#(w^0)=m+1$ and $w^0=f(z^k)$ $k=1,\dots,m+1$, $z^k\in bD$. Let Q be an holomorphic polynomial on $\mathbb{C}^n$ which separates $z^1,\dots,z^{m+1}$. The polynomial in the indeterminated X

$$P(X,w)=\prod_{k=1}^{m}(X-Q\circ F_k(w))=X^m+S_1(w)X^{m-1}+\dots+S_m(w)$$

is such that: i) $S_i\in A^\infty(\Omega)$, $i=1,\dots,m$ by part a); ii) $P(Q(z^j), w^0)=0$ for $j=1,\dots,m+1$. This produces a contradiction because P is a polynomial of degree m.

Proof of the Theorem. We prove the inequality for every z. With no loss of generality we can assume $z=0$. Let $f(0)=w$ and $f(\zeta)=v$.

Consider the mapping $\chi\colon \bar{\Omega}\to(\mathbb{C}^n)^m_{\mathrm{sym}}$ given by

$$\chi(v)=\begin{cases}\langle F_1(v)=\zeta^1,\dots,F_m(v)=\zeta^m\rangle & \text{if } v\in\bar{\Omega}\setminus\{f(z)\colon u(z)=0\},\ \text{and}\\ \{\zeta^1,\zeta^2,\dots,\zeta^m\}=f^{-1}(v).\\ \langle\zeta^1,\dots,\zeta^k\rangle \quad \text{where } \zeta^i \text{ is counted } \beta_i \text{ times, and}\\ \{\zeta^1,\dots,\zeta^k\}=f^{-1}(v), \quad v\in\{f(z)\colon u(z)=0\}.\end{cases}$$

Then, the mapping $\tilde{\chi}=\Psi\circ\chi\colon \bar{\Omega}\to\mathbb{C}^N$ whose components are the elementary symmetric functions of the local inverses $F_1,\dots,F_m$ of f, is $C^\infty(\bar{\Omega})$ by Lemma 2.4 a). Thus,

$$\begin{aligned}|z-\zeta^1|^2\dots|z-\zeta^m|^2 &\leqq c\sum_{i=1}^{N}|\chi_i(w)-\chi_i(v)|^2\\ &\leqq c|w-v|^2=c|f(\zeta)-f(z)|^2\end{aligned}$$

where the first inequality is given by Lemma 2.2, the second by the regularity of $\tilde{\chi}$ and the constants c are independent of z.

Remark. The proof of Theorem 2.3 can also be obtained using the techniques introduced by Diederich and Fornaess in their investigation of the polynomial neighborhoods of the branching locus of f [5].

Indeed, estimate (7) of Proposition 1 of [5] holds true also if ρ is defined as follows

$$\rho(w,v)=\frac{|z-\zeta^1|\dots|z-\zeta^m|}{d^{m-1}}$$

where $d=$ diameter of D, $\{\zeta^1,\dots,\zeta^m\}=f^{-1}(v)$ and $f^{-1}(w)=z$.

Moreover, the function h defined in the proof of the same Proposition belongs, in our case, to $A^\infty(\Omega)$ because of Lemma 2.4 a).

3. Uniform Estimate for the $\bar{\partial}$-Problem on D

Let us define

$$U'_\varepsilon(\zeta,z)=\{(\zeta,z)\bar{D}\times\bar{D}\colon|f(\zeta)-f(z)|<\varepsilon\} \tag{3.1}$$

where ε is as in Lemma 1.2. Thus, in analogy with the same lemma we have

$$\begin{gathered}P=(F\circ f)(H\circ f) \quad \text{on} \quad U'_\varepsilon;\quad |H\circ f|>\gamma' \quad \text{on} \quad U'_\varepsilon;\\ |P|>\gamma' \text{ outside } U'_\varepsilon\end{gathered} \tag{3.2}$$

$$\begin{gathered}\mathrm{Re}(F\circ f)(\zeta,z)>\lambda(\zeta)-\lambda(z)+\gamma'(|\zeta-z|^{2\beta_1}\dots\\ \dots|\zeta-z^k|^{2\beta_k}) \text{ on } U'_\varepsilon.\end{gathered} \tag{3.3}$$

The last inequality follows from Theorem 2.3.

Theorem 3.1. *Let D be as in* (1.2) *a branched cover of a bounded strictly pseudoconvex domain of $\mathbb{C}^n$. Assume $\bar{D}\in\mathscr{F}$. Then, for any $\bar{\partial}$-closed form*

$g \in C^1_{(0,1)}(\bar{D})$ there exists a solution α of $\bar{\partial}\alpha = g$ satisfying

$$\|\alpha\|_{C(D)} \leqq c(D)\|g\|$$

where $\| \ \|_{C(D)}$ is the supremum norm, $\|g\| = \max_i \|g_i\|_{D(D)}$ and g_i are the coefficients of the form g.

Proof. For sake of simplicity we confine ourselves to $\mathbb{C}^2$. Theorem 1.4 gives a solution $\alpha = T_1 g$ of $\bar{\partial}\alpha = g$. We prove now its boundedness.

$$T_1 g(z) = -c\Big(\int_{D \times \Delta_1} g(\zeta)\, \Omega_{2,0}(\zeta, z) + \int_{bD \times \Delta} g(\zeta)\, \Omega_{2,0}(\zeta, z)\Big) = J_1(z) + J_2(z), \quad z \in \bar{D}.$$

After the integration in t, we get

$$J_1(z) = c \int_D \frac{g_1(\zeta)(\bar{\zeta}_1 - \bar{z}_1) + g_2(\zeta)(\bar{\zeta}_2 - \bar{z}_2)}{|\zeta - z|^4}\, d\bar{\zeta} \wedge d\zeta$$

$$J_2(z) = c \int_{bD} g(\zeta) \frac{p_1(\zeta)(\bar{\zeta}_2 - \bar{z}_2) - p_2(\zeta)(\bar{\zeta}_1 - \bar{z}_1)}{|P(\zeta, z)||\zeta - z|^2}\, d\zeta.$$

Using spherical coordinates centred in $\zeta = z$ in J_1 and Cauchy's inequality in J_2 we have

$$|J_1(z)| \leqq c(D)\|g\|$$

$$|J_2(z)| \leqq c\|g\| \int_{bD} \frac{d\sigma(\zeta)}{|P(\zeta, z)||\zeta - z|}.$$

Thus, by (3.2) it follows

$$\int_{bD} \frac{d\sigma(\zeta)}{|P(\zeta, z)||\zeta - z|} \leqq \int_{bD \setminus U'_\varepsilon} \frac{d\sigma(\zeta)}{|P(\zeta, z)||\zeta - z|} + c \int_{U'_\varepsilon \cap bD} \frac{d\sigma(\zeta)}{|F(f(\zeta), f(z))||\zeta - z|}$$

and moreover the boundedness of the first integral in the right hand side.

Let us observe that, for any fixed z, $U'_\varepsilon = \bigcup_i^k B_i$, $1 \leqq k \leqq m$, B_i open neighborhood of the inverse images $z = z^1, \ldots, z^k$ of $f(z)$. These neighborhoods may be mutuallydisjoint or not. In any case, for ε sufficiently small, we can suppose $\dfrac{\partial \lambda}{\partial \zeta_j}(Q) \neq 0$, $\forall Q \in B_i$, $j = 1$ or 2, and that $B_1, \ldots, B_s$ be such that i) $B_i \cap B_j = \varnothing$ $1 \leqq i, j \leqq s$, ii) $\bigcup_i^s B_i = U'_\varepsilon$.

Let $F(f(\zeta), f(z)) = \hat{F}(\zeta, z)$. Then,

$$\int_{bD \cap U'_\varepsilon} \frac{d\sigma(\zeta)}{|\hat{F}(\zeta, z)||\zeta - z|} = \sum_{i=1}^{s} \int_{B_i \cap bD} \frac{d\sigma(\zeta)}{|\hat{F}(\zeta, z)||\zeta - z|}$$

We confine ourselves to study $\int_{B_1 \cap bD} \frac{d\sigma(\zeta)}{|\hat{F}(\zeta, z)||\zeta - z|}$, where B_1 contains z and other inverse images of $f(z)$. With no loss of generality, let z and z^2 belong to B_1 and let $\frac{\partial \lambda}{\partial \zeta_1}(Q) \neq 0$, $Q \in B_1$. Estimating the other integrals is easier and it can be done analogously.

Since

$$bD \cap B_1 = \{\zeta \in bD : |\zeta - z| < |\zeta - z^2|\} \cup \{\zeta \in bD : |\zeta - z| \geqq |\zeta - z^2|\} = A_1 \cup A_2$$

then, using (3.3) we get

$$\int_{bD \cap B_1} \frac{d\sigma(\zeta)}{|\hat{F}(\zeta, z)||\zeta - z|} = \sum_{i=1}^{2} \int_{A_i} \frac{d\sigma(\zeta)}{|\zeta - z|(|\operatorname{Im} \hat{F}(\zeta, z)|^2 + |\operatorname{Re} \hat{F}(\zeta, z)|^2)^{\frac{1}{2}}}$$

$$\leqq \int_{A_1} \frac{d\sigma(\zeta)}{|\zeta - z|(|\operatorname{Im} \hat{F}(\zeta, z)|^2 + (|\lambda(z)| + c|\zeta - z|^{2\beta})^2)^{\frac{1}{2}}} + \int_{A_2} \frac{d\sigma(\zeta)}{|\zeta - z^2|(|\operatorname{Im} \hat{F}(\zeta, z^2)| + (|\lambda(z^2)| + c|\zeta - z^2|^{2\beta})^2)^{\frac{1}{2}}}$$

where $\beta = k \max_{i=1,\ldots,k} \beta_i$.

In order to estimate the first integral we introduce local real coordinates as follows

$$\tau_z(\zeta) = \begin{cases} t_1 = \lambda(\zeta) - \lambda(z) \\ t_2 = \operatorname{Im} \hat{F}(\zeta, z) \\ t_3 = \operatorname{Re}(\zeta_2 - z_2) \\ t_4 = \operatorname{Im}(\zeta_2 - z_2) \end{cases}$$

The jacobian of this change of coordinates is, for $\zeta = z$

$$-\left(\frac{\partial \lambda}{\partial \zeta_1}(z) \cdot \overline{\frac{\partial \hat{F}}{\partial \zeta_1}(z, z)} + \overline{\frac{\partial \lambda}{\partial \zeta_1}(z)} \cdot \frac{\partial \hat{F}}{\partial \zeta_1}(z, z)\right).$$

As $\frac{\partial \hat{F}}{\partial \zeta_1}(z, z) = -\frac{\partial \lambda}{\partial \zeta_1}(z)$, the jacobian is equal to $2\left|\frac{\partial \lambda}{\partial \zeta_1}(z)\right|^2 \neq 0$.

Thus the change is valid in a neighborhood of A_1. Note that A_1 is contained in $t_1 = 0$, $t_2^2 + t_3^2 + t_4^2 < R^2$ with R independent of z. Hence it is possible to proceed as in [8]. Using spherical coordinates $t_2 = r \cos \phi$, $t_2^2 + t_3^2 + t_4^2 = r^2$ we get

$$\int_{A_1} \frac{d\sigma(\zeta)}{|\hat{F}||\zeta - z|} \leqq c 2\pi \int_0^R dr \int_0^\pi \frac{r^2 \operatorname{sen} \phi \, d\phi}{r(r^2 \cos^2 \phi + (|\lambda(z)| + r^{2\beta})^2)^{\frac{1}{2}}}$$

$$= c 2\pi \int_0^R \log \frac{(\sqrt{r^2 + (|\lambda(z)| + r^{2\beta})^2} + r)^2}{(|\lambda(z)| + r^{2\beta})^2}$$

$$< c(D).$$

The same is true for the integral on A_2 using the change of coordinates $\tau_{z^2}(\zeta)$.

Acknowledgement. We want to express our gratitude to Prof. K. Diederich for his interest in this paper and for useful discussion. He communicated to us that in a joint paper with Fornaess, in preparation, he can prove that if D, $\Omega \Subset \mathbb{C}^n$ are C^∞-smooth pseudoconvex domains, Ω is regular and f: $D \to \Omega$ is a proper holomorphic mapping, $f \in C^\infty(\bar{D})$, then $\bar{D}$ is uniform H-convex. This property will assure automatically that the clousure of our domain D belongs to $\mathscr{F}$.

References

1. Alexander, H.: Proper holomorphic mappings in $\mathbb{C}^n$. Indiana Univ. Math. J. **26**, 137–146 (1977)
2. Bell, S.: Proper holomorphic mappings and the Bergman projection. Duke Math. J. **48**, 167–175 (1981)
3. Bell, S., Catlin, D.: Boundary regularity of proper holomorphic mappings. Duke Math. J. **49**, 385–396 (1982)
4. Bonami, A., Lohoue', N.: Projecteurs de Bergman et Szegö pour une classe de domaines faiblement pseudo convexes et estimations L^p. Compos. Math. **46**, 159–226 (1982)
5. Diederich, K., Fornaess, J.E.: Pseudoconvex domains: existence of Stein neighborhoods. Duke Math. J. **44**, 641–662 (1977)
6. Diederich, K., Fornaess, J.E.: Boundary regularity of proper holomorphic mapping. Invent. Math. **67**, 363–384 (1982)
7. Henkin, G.M.: Integral representations of functions holomorphic in strictly pseudoconvex domains and some applications. Math. USSR, Sb. **7**, 597–616 (1969)
8. Henkin, G.M.: Integral representations of functions holomorphic in strictly pseudoconvex domains and applications to $\bar{\partial}$-problem. Math. USSR, Sb. **II**, 273–281 (1970)
9. Parrini, C., Selvaggi Primicerio, A.: A decomposition theorem for $A^\infty(D)$ in pseudoconvex domains. Boll. Unione Mat. Ital. (6) 3-B (1984)
10. Øvrelid, N.: Integral representation formulas and L^p-estimates for $\bar{\partial}$-equation. Math. Scand. **29**, 137–160 (1971)
11. Range, R.M.: On Hölder estimates for $\bar{\partial}u = f$ on weakly pseudoconvex domains. In: Proceedings of International Conferences Cortona, Italy 1976–1977
12. Range, R.M.: The Caratheodory metric and holomorphic maps on a class of weakly pseudoconvex domains. Pacific J. Math. **78**, 173–189 (1978)
13. Range, R.M., Siu, Y.T.: Uniform estimates for the $\bar{\partial}$-equation on domains with piecewise smooth strictly pseudoconvex boundaries. Math. Ann. **206**, 325–354 (1973)
14. Sibony, N.: Un exemple de domain pseudoconvexe regulier on l'equation $\bar{\partial}u = f$ n'admet pas de solution bornée pour f bornée. Invent. Math. **62**, 235–242 (1980)
15. Whitney, H.: Complex Analytic Varieties. Addison-Wesley Series in Mathematics, 1972

Received December 5, 1983

Math. Z. 189, 475–486 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Actions of Compact Lie Groups on Homogeneous Spaces

Volker Hauschild

Mathematisches Institut der Universität, Bunsenstr. 3-4, D-3400 Göttingen, Federal Republic of Germany

0. Introduction

Let G be a compact connected Lie group, $H \subset G$ a closed connected subgroup, then G acts on the homogeneous space $X = G/H$ in a natural way by left translations. We call this action of G on X the natural action or the linear action. One might ask, given a G and a homogeneous space G/H, if, beside the linear action, there exist essentially different actions of G on G/H. This is a classical problem in transformation group theory and the present situation seems to be far from a complete and satisfying solution of it. If $H = \{e\}$ is the trivial subgroup we are in the special case of actions of G on itself. In this case we know of at least two actions, namely the natural action and the adjoint action by conjugation, and in some cases (see e.g. [4], p. 151) it has been proved that these are indeed all actions up to an equivariant equivalence of G on itself. Nevertheless even this question remains totally open.

In the more general case $H \neq \{e\}$ some new results have been obtained, so for example for the actions of $SU(n)$ on its corresponding Stiefel manifolds, by W.Y. Hsiang and P. Tomter in [5], where they show, that up to conjugacy, any differentiable action of $SU(n)$ on $W_{n,k} = SU(n)/SU(k)$, $k > \frac{1}{2}n$, is the natural action.

The case $X = G/NT$, G a compact Lie group, NT the normalizor of the maximal torus T of G, has been completely settled by the author, in fact with the same answer. (See [3], Satz 3.1, ii.)

In this paper we are concerned with the case of the flag manifolds G/T. (This case can not be treated by the lifting methods in [3].) Here our main theorem is:

Theorem 1. *Let X be a simply connected compact closed oriented manifold with $H^*(X; \mathbb{R}) \cong H^*(G/T; \mathbb{R})$, $\mathbb{R}$ the reals, G a simple compact Lie group, $T \subset G$ its maximal torus. If G acts on X with finitely many orbit types, then $X = G/T$ and every G-action on X is conjugate to the natural action.*

As a corollary we obtain:

Theorem 1'. *Let G be a compact simple Lie group, $T \subset G$ a fixed maximal torus, let G act on $X = G/T$ with finitely many orbit types. Then the action is conjugate to the natural action.*

One might conjecture that similar results hold at least in all cases G/H, where H is a closed subgroup of maximal rank, see e.g. [4], p. 144.

I conclude these remarks with an outline of the proof. If $H \subset G$ is a closed subgroup of maximal rank, then $H^*(B_H)$ is a free finitely generated module over $H^*(B_G)$ if the coefficient field of the cohomology has characteristic zero. Let Γ be a compact Lie group and let $\rho: \Gamma \to G$ be a homomorphism. Then this gives an action of Γ on G and all its homogeneous spaces G/H by left translations: $\varphi_\rho: \Gamma \times G/H \to G/H$, $\varphi_\rho(\gamma, gH) = \rho(\gamma)\, gH$. In the case $\operatorname{rk} G = \operatorname{rk} H$ it is now easy to compute the equivariant cohomology $H^*_\Gamma(G/H) = H^*_\Gamma(G/H; k)$, $\operatorname{char} k = 0$. We look on the obvious pullback diagram

$$\begin{array}{ccc} E_\Gamma \times_\Gamma X & \longrightarrow & B_H \\ \downarrow & & \downarrow \\ B_\Gamma & \longrightarrow & B_G \end{array}$$

By the methods of the Eilenberg-Moore spectral sequence we obtain for $H^*_\Gamma(X) = H^*(E_\Gamma \times_\Gamma X)$

$$H^*_\Gamma(X) \cong R_\Gamma \otimes_{R_G} R_H.$$

(See e.g. [4], p. 38.) Here is $R_K = H^*(B_K)$ for a group K. In particular, if G acts on G/H by left translations, we obtain

$$H^*_G(X) \cong R_H$$

or, if $T \subset G$ is a maximal torus

$$H^*_T(X) \cong R_T \otimes_{R_G} R_H.$$

Conversely, it can be shown: If X equals G/H, $\operatorname{rk} G = \operatorname{rk} H$, and G acts on X such that $H^*_G(X) \cong R_H$, then the action is linear. So, our main problem in proving the above theorems, is to show that a G-action on G/H has a "linear" equivariant cohomology. In the case G/T, G simple, this can be done in a surprisingly simple way using solely elementary facts from commutative algebra. For example it is shown, that any graded endomorphism of the k-algebra $H^*(G/T)$ is either trivial or an isomorphism, a result which in an other context has also been proved by A. Liulevicius.

1. Localization and Other Preliminary Facts

We begin with some commutative algebra. To fix notations let k be a field of characteristic zero. Denote by $k[x_1, \ldots, x_n]$ the graded commutative polynomial algebra in the indeterminants $x_1, \ldots, x_n$ of positive degrees. Note

$\deg x_i \equiv 0(2)$, $i=1, \ldots, n$. If B is a positively graded k-algebra let $B_+ \subset B$ denote the augmentation ideal.

Theorem 1.1. *Let $R=k[t_1, \ldots, t_r]$, $P=k[x_1, \ldots, x_n]$ and $I_0 \subset P_+$ the ideal generated by the homogeneous elements $f_1, \ldots, f_m$. Set $A_0=P/I_0$ and assume $\dim_k A_0 < \infty$. If A is a homogeneous R-algebra, free as an R-module, such that $A \otimes_R k \cong A_0$ in the graded sense, then $A \cong P_R/I$, where $P_R=R \otimes_k P$ and $I \subset P_R$ is an ideal generated by homogeneous elements of the form $1 \otimes f_j - r_j$, $j=1, \ldots, m$, where the r_j are elements of $R_+ P_R$. In particular A is a finitely generated R-module.*

Proof. From $A \otimes_R k = A_0$ it follows $\mathrm{rk}_R A = \dim_k A_0$, i.e., A is finitely generated.

Let $p: A \to A_0$ denote the canonical projection, let $\bar{X}_1, \ldots, \bar{X}_n \in A$ be homogeneous elements with $p(\bar{X}_i)=\bar{x}_i$, where $\bar{x}_i$ is the remainder of $x_i \in P$ modulo I_0. Then A_0 is generated as a k-vector space by 1 and some polynomials $p_j(\bar{x}_1, \ldots, \bar{x}_n)$, $j=1, \ldots, N-1$, $N=\dim_k A_0$. So, it follows by Nakayama, that A as a R-module is generated by 1 and the polynomials $p_j(\bar{X}_1, \ldots, \bar{X}_n)$, $j=1, \ldots, N-1$. We write $P_R=R[X_1, \ldots, X_n]$, $X_i=1 \otimes x_i$. So the homomorphism $P_R \to A$, $X_i \to \bar{X}_i$ is surjective, i.e., we can write $A=P_R/I$, where I is a certain ideal. Let $F_j=1 \otimes f_j \in P_R$ and $\bar{F}_j$ its image in A, then $p(\bar{F}_j)=0 \in A_0$. This shows $\bar{F}_j \in R_+ A$. Let $r_j \in R_+ P_R$ such that $\bar{F}_j - \bar{r}_j \in 0$ in A. Then $F_j - r_j \in I$, $j=1, \ldots, m$. Let $I' \subset I$ be the ideal generated by the $F_j - r_j$, $j=1, \ldots, m$. Let $\varphi: P_R/I' \to P_R/I$ denote the obvious quotient map and let $K=\ker \varphi$. Then by the freeness of A the exact sequence

$$0 \to K \to P_R/I' \to P_R/I \to 0$$

induces an exact sequence of k-vector spaces. In particular the induced homomorphism $\bar{\varphi}: (P_R/I) \otimes_R k \to A_0$ becomes an isomorphism. In other words $\bar{K}=K \otimes_R k=0$. If we have chosen the r_j homogeneous such that I' is homogeneously generated, the above sequence is an exact sequence of graded modules over the graded polynomial ring R. So $\bar{K}=0$ implies $K=0$ (see e.g. Hilton-Stammbach, p. 252, Prop. 7.2). This shows φ to be an isomorphism, *i.e.*, $I=I'$. □

Let now X be a topological space, let G be a group having a universal fibration $E_G \to B_G$, let G act on X as a group of homeomorphisms. We then consider the Borel-fibration

$$X \to E_G \times_G X \to B_G.$$

We set $X_G=E_G \times_G X$ and $H_G^*(X) \cong H^*(X_G)$, where we use Cech-cohomology with compact supports in the coefficient field k. We denote the projection of the above fibre bundle by π. If $R_G=H^*(B_G)$, then $H_G^*(X)$ is a R_G-algebra via π^* and the cup-product. Obviously, the above construction is functorial with respect to equivariant maps. The functor obtained in this way is called the ordinary equivariant cohomology with coefficients in the constant sheaf $\underline{k}$ given by the ground field. In the following we assume always, that G is acting with finitely many orbit types.

Let X be a space such that $H^*(X)$ is multiplicatively generated by finitely many elements of even degree, i.e.,

$$H^*(X) \cong k[x_1, \ldots, x_n]/I_0,$$
$$\deg x_i \equiv 0 \pmod 2,$$
$$I_0 = (f_1, \ldots, f_m), \quad f_j \text{ homogeneous.}$$

Let G be a compact Lie group acting on X, then the Serre-spectral sequence of $X \to X_G \to B_G$ degenerates at the 2-term. This shows $H_G^*(X)$ to be a free graded finitely generated R_G-module. Let $R_G^+ \subset R_G$ the maximal ideal generated by the elements of strictly positive degree, then $R_G/R_G^+ = k$. The homomorphism $i^*: H_G^*(X) \to H^*(X)$ induced by the inclusion $X \subset X_G$ is then given by the projection map

$$H_G^*(X) \to H_G^*(X)/R_G^+ H_G^*(X) \cong H^*(X).$$

(See e.g. [1], Th. 14.2) Therefore by Theorem 1.1 we get

$$H_G^*(X) = R_G[X_1, \ldots, X_n]/I,$$

where the defining ideal I is generated by elements of the form $F_j - r_j$, $r_j \in R_G^+ \cdot R_G[X_1, \ldots, X_n]$, $F_j = 1 \otimes f_j$, $j = 1, \ldots, m$. Obviously, the polynomials r_j depend from the action. Therefore we have to investigate what kind of polynomials r_j can actually occur from a G-action. A partial answer to this question is given by the localization theorem in ordinary equivariant cohomology theory. Let $S \subset R_G$ a multiplicatively closed subset, let $X^S \subset X$ be the closed G-equivariant subspace defined as follows: If $x \in X$ is a point, $G_x \subset G$ the closed isotropy subgroup of x, let $P_x \subset R_G$ be the kernel of the natural restriction homomorphism $R_G \to R_{G_x}$. Then X^S is the subset of all points $y \in X$, s.t., $P_y \cap S = \varnothing$.

Theorem 1.2 (Localization theorem) [4]. *Let X be a compact or paracompact Hausdorff space of finite k-cohomology dimension. Let the compact Lie group G act on X with finitely many orbit types. If $S \subset R_G$ is a multiplicatively closed subset, then the inclusion $X^S \subset X$ induces an isomorphism of graded $S^{-1}R_G$-algebras $S^{-1}H_G^*(X) \cong S^{-1}H_G^*(X^S)$.*

As an example, we consider the special case $S = R_G - \{0\}$, G a torus. Then X^S is the fixed set $F = F(G; X)$ and $S^{-1}H_G^*(X) \cong S^{-1}R_G \otimes_k H^*(F)$.

Theorem 1.3 (Special case of Th. IV.1 in [4]). *Let G be a torus and X a compact space or a paracompact space with finite k-cohomology dimension such that $H^*(X)$ is generated as a k-algebra by elements of even degree $X_1, \ldots, X_n$. Let $P_G = R_G[X_1, \ldots, X_n]$ and $H_G^*(X) = P_G/I$. If $S = R_G - \{0\}$, we write $J = S^{-1}I$ as an ideal in $S^{-1}P_G$.*

(i) *If $J = \bigcap_{i \in \Delta} J_i$ is the reduced primary decomposition of J, then there is a natural correspondence between the sets $\{J_i\}_{i \in \Delta}$ and $\pi_0 F(T; X)$ of connected components F_i of the fixed set such that $S^{-1}R_G \otimes_k H^*(F_i) \cong S^{-1}P_G/J_i$ and the localization isomorphism is given by the canonical homomorphism*

$$S^{-1}P_G/J \to \prod_{i \in \Delta} S^{-1}P_G/J_i.$$

(ii) *The set of associated prime ideals of J is given by a family $\{\underline{m}_i\}_{i\in\Delta}$ of maximal ideals $\underline{m}_i=(X_1-\omega_{i1},\ldots,X_n-\omega_{in})$, where the ω_{ij} are homogeneous elements of R_G with $\deg\omega_{ij}=\deg X_j$.*

As a simple consequence it follows:

Lemma 1.1. *Under the above assumptions the set of associated prime ideals $\mathrm{Ass}_{P_G}H_G^*(X)$ is in a natural way bijective with the set of connected components $\pi_0F(G;X)$. Moreover any $P\in\mathrm{Ass}_{P_G}H_G^*(X)$ is of the form*

$$P_i=(X_1-\omega_{i1},\ldots,X_n-\omega_{in}),$$

i.e., $H_G^*(X)/P_iH_G^*(X)\cong R_G$.

Proof. The first statement follows from (i) of the above theorem. The second is a consequence of Bourbaki, Alg. comm., Ch IV, 1.2, Prop. 5.

2. A Special Case

As a very important special case we are considering spaces X, such that $H^*(X)$ as a graded k-algebra is generated by finitely many elements $x_1,\ldots,x_n$ of degree two. Then $H_G^*(X)$ is a free R_G-module and the long exact sequence of the couple (X_G,F_G) degenerates into the short exact sequence

$$(*)\qquad 0\to H_G^*(X)\to H_G^*(F)\to H_G^{\mathrm{odd}}(X,F)\to 0.$$

(By the localization theorem $H_G^*(X,F)$ is a torsion module over R_G. By the freeness of $H_G^*(X)$ the second arrow must be injective.)

Theorem 2.1. *Let G be a torus and X a finitistic space such that $H^*(X)$ is generated as a k-algebra only by elements of degree two. Let $H^*(X)=P/I_0$, $H_G^*(X)=P_G/I$, then we have:*

(i) *The set of associated prime ideals $\mathrm{Ass}_{P_G}H_G^*(X)=\{P_i\}_{i\in\Delta}$, is in a natural correspondence with the connected components of the fixed set; moreover any P_i is of the form $P_i=(X_1-\omega_{i1},\ldots,X_n-\omega_{in})$, $\omega_{ij}\in R_G$, where the homomorphism $H_G^*(X)\to H_G^*(\{s_i\}\times B_G)$, $s_i\in F_i$, is given by $X_j\to\omega_{ij}$.*

(ii) *Let $I=\bigcap_{i\in\Delta}I_i$ the corresponding reduced primary decomposition, then the sequence (*) is equivalent to the canonical sequence*

$$0\to P_G/I\to\prod_{i\in\Delta}P_G/I_i\to\mathrm{coker}\to 0.$$

Proof. The first statement is an immediate consequence of Lemma 1.1. For the second statement we observe that the homomorphism $r\colon H_G^*(X)\to H_G^*(F')$ induced by the inclusion of a connected fixed space component $F'\subset X$ is given by $X_j\to Y_j+\omega_j$ with $Y_j\in H^2(F')$, $\omega_j\in H^2(B_G)$. By the localization theorem the homomorphism $S^{-1}r\colon S^{-1}H_G^*(X)\to S^{-1}H_G^*(F')$ is an epimorphism. From the isomorphism $S^{-1}H_G(F')\cong S^{-1}R_G\otimes_kH^*(F')$ it follows that $Y_1,\ldots,Y_n\in H^2(F')$ is a set of generators of the k-algebra $H^*(F')$. This shows $\bar{r}=r\otimes 1\colon H_G^*(X)$

$\otimes_{R_G} k \to H_G^*(F') \otimes_{R_G} k \cong H^*(F')$ to be an epimorphism. Let $C = \operatorname{coker} r$, then $\bar{C} = C \otimes_{R_G} k = 0$. Because C is finitely generated, the R_G-module C must be the zero module. So, the map r is an epimorphism too. We conclude that the restriction map $i_i^*: H_G^*(X) \to H_G^*(F_i)$ is an epimorphism for all i. This shows i_i^* to be a projection map $P_G/I \to P_G/I_i$ for an ideal $I_i \subset P_G$ with $I \subset I_i$. Let $\{I_i\}_{i \in \Delta}$ be the family of ideals obtained in this way. We define $I' = \bigcap_{i \in \Delta} I_i$. Then we consider the following commutative diagram:

$$\begin{array}{ccc} H_G^*(X) & \xrightarrow{i^*} & \prod_{i \in \Delta} H_G^*(F_i) \\ \downarrow{\scriptstyle p} & & \downarrow{\scriptstyle \cong} \\ P_G/I' & \xrightarrow[\mu]{} & \prod_{i \in \Delta} P_G/I_i \end{array}$$

Here, the first horizontal arrow is induced by the inclusion of the fixed set, whereas μ is the obvious canonical map. The map p is the projection induced by $I \subset I'$. By the injectivity of i^* it follows that $I' \subset I$, i.e., $I = I'$. So, p is an isomorphism and we are done.

As a corollary we obtain:

Corollary. *Let X be a compact closed oriented manifold, $H^*(X) = P/I_0$, $P = k[x_1, \ldots, x_n]$, $\deg x_i = 2$, let G a torus acting continuously and nontrivially and with finitely many orbit types on X. Then there exists a connected component F of the fixed set $F(T; X)$ such that the restriction homomorphism $H_G^*(X) \to H_G^*(\{q\}) \cong R_G$, $q \in F$, is nontrivial, i.e., the prime ideal P of the component F is given as $P = (X_1 - \omega_1, \ldots, X_n - \omega_n)$, $\omega_i \neq 0$, for at least one index i, $1 \leqq i \leqq n$.*

Proof. If all the associated prime ideals P of the P_G-module $H_G^*(X)$ are trivial, then $\operatorname{Ass}_{P_G} H_G^*(X) = \{P\}$, $P = (X_1, \ldots, X_n)$. Then, by the theorem, we have $H_G^*(X) = P_G/I$, $\operatorname{Rad}(I) = P$. Therefore $H_G^*(X) \cong R_G \otimes_k H^*(F)$, F being the single connected component. It follows $H^*(X) \cong H^*(F)$. This implies $\dim F \geqq \dim_c F = \dim_c X$, where $\dim_c(-)$ denotes the k-cohomology dimension. Now, because X is a closed compact oriented manifold, it follows $\dim_c X = \dim X$. This shows $\dim F \geqq \dim X$, i.e., the action is trivial, which is a contradiction.

3. Equivariant Cohomology of Linear Actions

Let G be a compact Lie group, $H \subset G$ a closed subgroup, let Γ be a compact connected Lie group and let $\rho: \Gamma \to G$ be a homomorphism. Then ρ induces an action of Γ on $X = G/H$ in the following way:

$$\begin{aligned} \varphi_\rho: \Gamma \times G/H &\to G/H \\ (\gamma, gH) &\to \rho(\gamma)\, gH. \end{aligned}$$

In the following such actions will be called linear or homomorphic. If ρ is injective, then the induced action φ_ρ is effective. We assume, that ρ is always

injective, so Γ can be considered as a subgroup of G. If $H \subset G$ is a closed connected subgroup of maximal rank, it is not very difficult to compute the equivariant cohomology $H^*_\rho(G/H) = H^*_\Gamma(G/H)$, $\operatorname{char} k = 0$, of a linear action. Let $i: H \subset G$ denote the inclusion homomorphism. We consider the obvious fiber square

$$\begin{array}{ccc} E_\Gamma \times_\Gamma G/H & \longrightarrow & B_H \\ \downarrow & & \downarrow{\scriptstyle Bi} \\ B_\Gamma & \xrightarrow[B\rho]{} & B_G \end{array}$$

Let $R_G = H^*(B_G)$, $R_H = H^*(B_H)$ and $R_\Gamma = H^*(B_\Gamma)$, then R_H and R_Γ are R_G-modules by the induced mappings $(Bi)^*$ and $(B\rho)^*$. In particular, R_H is a free R_G-module. The E_2-term of the corresponding Eilenberg-Moore spectral sequence is then given by:

$$\begin{aligned} E_2 &\cong \operatorname{Tor}_{R_G}(R_\Gamma, R_H) \\ &\cong R_\Gamma \otimes_{R_G} R_H. \\ E_2^{*,q} &= 0, \quad q > 0. \end{aligned}$$

Therefore the spectral sequence degenerates at the 2-term and we get:

$$H^*_\Gamma(G/H) \cong R_\Gamma \otimes_{R_G} R_H,$$

where the isomorphism is induced by the natural map $R_\Gamma \otimes_{R_G} R_H \to H^*_\Gamma(G/H)$.

Lemma 3.1. *Let G be a compact connected Lie group and $H \subset G$ a closed connected subgroup of maximal rank, let $X = G/H$, let $\varphi_\rho: \Gamma \times X \to X$ be a linear action of a compact connected Lie group Γ, then the canonical homomorphism $R_\Gamma \otimes_{R_G} R_H \to H^*_\Gamma(X)$ is an isomorphism of graded R_Γ-algebras.*

As an example take $H = T$, a fixed maximal torus, let $\rho: \Gamma \to T$ be an isomorphism, then $H^*_\Gamma(G/T) \cong R_T \otimes_{R_G} R_T$.

In the following we want to investigate in greater detail the structure of $\operatorname{Spec} R_T \otimes_{R_G} R_T$. Let $\underline{g}$ be the Lie algebra of G, $\underline{h} \subset \underline{g}$ the Lie algebra of T (both real). Let $\mathcal{O}_{\underline{h}} = k[\underline{h}]$ denote the ring of k-valued regular functions on $\underline{h}$. Then the Weyl group $W = W_G = NT/T$ acts in a natural way on $\underline{h}$ and $\mathcal{O}_{\underline{h}}$. Let $\mathcal{O}^W_{\underline{h}} \subset \mathcal{O}_{\underline{h}}$ be the corresponding invariant ring. Because W acts on $\underline{h}$ as a real reflection group, the ring $\mathcal{O}^W_{\underline{h}}$ is a polynomial ring [2]. It is a well known fact in fiber bundle theory that the natural homomorphism $\pi^*: R_G \to R_T$ induced by the projection map $\pi: B_T \to B_G$ can be identified with the inclusion $\mathcal{O}^W_{\underline{h}} \subset \mathcal{O}_{\underline{h}}$ of invariants [1]. The push-out diagram

$$\begin{array}{ccc} R_T & \longrightarrow & R_T \otimes_{R_G} R_T \\ \uparrow & & \uparrow \\ R_G & \longrightarrow & R_T \end{array}$$

corresponds to the pullback-diagram

$$\begin{array}{ccc} \underline{h} \times_{\underline{h}/W} \underline{h} & \longrightarrow & \underline{h} \\ \downarrow & & \downarrow p \\ \underline{h} & \xrightarrow[p]{} & \underline{h}/W, \end{array}$$

where $p: \underline{h} \to \underline{h}/W$ is the obvious quotient morphism. Because of the $(\underline{h}/W)$-automorphisms induced by the elements $\sigma \in W$,

$$\begin{array}{ccc} \underline{h} & \xrightarrow{\sigma} & \underline{h} \\ & \searrow \quad \swarrow & \\ & \underline{h}/W, & \end{array}$$

we obtain $|W|$-many sections $s_\sigma: \underline{h} \to \underline{h} \times_{\underline{h}/W} \underline{h}$, $\sigma \in W$, i.e., there exist $|W|$-many different surjective $\mathcal{O}_{\underline{h}}$-homomorphisms

$$\varphi_\sigma: \mathcal{O}_{\underline{h}} \otimes_{\mathcal{O}_{\underline{h}/W}} \mathcal{O}_{\underline{h}} \to \mathcal{O}_{\underline{h}}$$

or

$$\varphi_\sigma: R_T \otimes_{R_G} R_T \to R_T.$$

Let

$$R_T^l = k[t_1, \ldots, t_n],$$
$$R_T^r = k[x_1, \ldots, x_n],$$

then we have

$$R_T \otimes_{R_G} R_T = \frac{R_T^l \otimes_k R_T^r}{I} \cong \frac{R_T[x_1, \ldots, x_n]}{I},$$

where the ideal I is generated by the elements $g_i(x) - g_i(t)$, all i. Here $g_1, \ldots, g_n$ is a system of algebraically independent homogeneous invariants in R_T, such that $R_G = k[g_1, \ldots, g_n]$. So we see that the kernel of φ_σ as an ideal is generated by the "diagonal" elements $x_i - t_i^\sigma$, $\sigma \in W$. Let $P^\sigma = (x_1 - t_1^\sigma, \ldots, x_n - t_n^\sigma) \subset R_T^l \otimes_k R_T^r$, then P^σ is an associated prime ideal of I. Conversely, the following lemma shows, that any associated prime ideal of I is given by a P^σ.

Lemma 3.2. *Let* $P^\sigma = (x_1 - t_1^\sigma, \ldots, x_n - t_n^\sigma) \subset R_T[x]$, $\sigma \in W$, *then* $I = \bigcap_{\sigma \in W} P^\sigma$, *in particular, the action of* W *on* R_T *induces a transitive and free action on the set of associated prime ideals* $\mathrm{Ass}_{R_T[x]}(R_T[x]/I)$.

Proof. Let $A = R_T[x]/I$, let $I' = \bigcap_{\sigma \in W} P^\sigma$ and $B = R_T[x]/I'$. Obviously, we have $I \subset P^\sigma$, all $\sigma \in W$, i.e., $I \subset I'$. The inclusion $I \subset I'$ induces an epimorphism $\varphi: A \to B$ of graded R_T-modules. Now, A is a free graded R_T-module, therefore $\dim_k A \otimes_{R_T} k = |W|$, implies $\mathrm{rk}_{R_T} A = |W|$. So B is a finitely generated R_T-module. Let $S = R_T - \{0\}$, let $Q = S^{-1} R_T$, then $\dim_Q S^{-1} B = |W|$. By the semicontinuity of the dimension function, the generic dimension of B is less or equal to the closed dimension, i.e.,

$$\dim_Q S^{-1} B \leqq \dim_k B \otimes_{R_T} k.$$

Moreover, the existence of the surjection $\varphi: A \to B$ implies

$$\dim_k B \otimes_{R_T} k \leqq \dim_k A \otimes_{R_T} k.$$

So we conclude

$$\dim_k B \otimes_{R_T} k = \dim_Q S^{-1} B = |W|.$$

This implies B to be a free R_T-module with $\mathrm{rk}_{R_T} B = \mathrm{rk}_{R_T} A$, i.e., φ is an isomorphism, and so it follows $I = I'$. □

4. On a Rigidity Property of Flag Manifolds

Let k be a field of characteristic zero, let $\mathscr{A}_k$ the category whose objects are the graded commutative k-algebras A with $A = P/J$, $P = k[x_1, \ldots, x_n]$, $\deg x_i = 2$, $J = (g_1, \ldots, g_m)$, g_j homogeneous with respect to the above grading and $\deg g_1 < \deg g_2 \leqq \ldots \leqq \deg g_m$, $\deg g_1 = 4$. Moreover, let g_1 be the quadratic form of a definite symmetric bilinear form. As morphisms we take the gradation preserving homomorphisms. So, in the defining ideal J there exists precisely one definite quadratic form $q = g_1$.

Lemma 4.1. *Let $\varphi: A \to B$ be a morphism in $\mathscr{A}_k$, then either $\mathrm{Im}\,\varphi = B^0 \subset B$ or φ is an epimorphism.*

Proof. Let $A = P/J^A$, $B = S/J^B$, $P = k[x_1, \ldots, x_n]$, $S = k[y_1, \ldots, y_m]$, $J^A \subset P$, $J^B \subset S$, the defining ideals. We can assume φ to be induced by a gradation preserving homomorphism Φ, such that the diagram

$$\begin{array}{ccc} J^A & \xrightarrow{\Phi|_{JA}} & J^B \\ \downarrow & & \downarrow \\ P & \xrightarrow{\Phi} & S \\ \downarrow & & \downarrow \\ A & \xrightarrow{\varphi} & B \end{array}$$

is commutative. Let $g^A \subset J^A$ and $g^B \subset J^B$ be the corresponding quadratic forms. Then we have by degree reasons: $\Phi(g^A) = a g^B$, $a \in k$. Now, we have

$$\Phi(g^A(x_1, \ldots, x_n)) = g^A(\Phi(x_1), \ldots, \Phi(x_n)) = a g^B(y_1, \ldots, y_m).$$

If $a = 0$, then it follows $g^A(\Phi(x_1), \ldots, \Phi(x_n)) = 0$, and therefore by definiteness $\Phi(x_1) = \Phi(x_2) = \ldots = \Phi(x_n) = 0$, i.e., $\mathrm{Im}\,\varphi = B^0$.

Let $a \neq 0$, then by degree reasons again, we have $\Phi(x_i) = \omega_i$, $\omega_i = \sum_{j=1}^{m} \omega_{ij} y_j$, $\omega_{ij} \in k$. Therefore, we have $g^A(\omega_1, \ldots, \omega_n) = a g^B(y_1, \ldots, y_m)$.

Let $\omega: k^m \to k^n$ the linear map given by the (m, n)-matrix $\omega = (\omega_{ij})$. Then, for a $v \in k^m$, we have $g^B(v) = \frac{1}{a}(g^A \omega)(v)$. Let $\mathrm{rk}\,\omega < m$, then there exists a nonvanish-

ing v such that $\omega(v)=0$. So $g^B(v)=0$ in contradiction to the definiteness of g^B. This shows $\operatorname{rk}\omega=m$, i.e., the elements ω_i, $i=1,\ldots,n$, span the subspace B^2. So Φ and φ must be epimorphisms.

Lemma 4.2. *Let G be a compact simple connected Lie group, let $T\subset G$ a maximal torus. Then $H^*(G/T;\mathbb{R})$ is an object in the category $\mathscr{A}_{\mathbb{R}}$.*

Proof. We have $H^*(G/T;\mathbb{R})\cong H^*(B_T;\mathbb{R})/(\pi^* H^*_+(B_G;\mathbb{R}))$, $H^*(B_G;\mathbb{R})\cong\mathbb{R}[g_1,\ldots,g_n]$, g_j homogeneous, $n=\operatorname{rk}G$. The degrees $d_j=\deg g_j$ are given by the exponents e_j of the Weylgroup $W_G=NT/T$ as follows: $d_j=2e_j+2$, $j=1,\ldots,n$. See e.g. [6], p. 112. Therefore it exists precisely one generator, say g_1, with $d_1=4$, in other terms, g_1 is a quadratic form. By degree reasons any WG-invariant quadratic polynomial $q\in H^*(B_T;\mathbb{R})$ must be proportional to g_1. Let q the restriction to the Lie algebra LT of T of the quadratic polynomial which stems from the Killing form κ: $LG\times LG\to\mathbb{R}$. By the compactness of G the form q must be negatively definite and WG-invariant. So, g_1 is also definite.

Theorem 4.1. *Let X be a compact closed oriented manifold, such that $H^*(X;\mathbb{R})\in|\mathscr{A}_{\mathbb{R}}|$. Let G be a compact connected simple Lie group acting on X with finitely many orbit types. Then there exists a surjective graded R_G-algebra homomorphism $\varphi: H^*_G(X)\to R_T$ and therefore a surjective graded $\mathbb{R}$-algebra homomorphism $\bar{\varphi}: H^*(X)\to H^*(G/T)$.*

Corollary. *Let X be a compact closed oriented manifold such that $H^*(X;\mathbb{R})\in|\mathscr{A}_{\mathbb{R}}|$. Let G be a compact connected simple Lie group acting on X with finitely many orbit types. Then we have the following inequality of Euler characteristics:*

$$\chi(X)\geqq\chi(G/T).$$

Proof. The commutative diagram (see e.g. [4], p. 38)

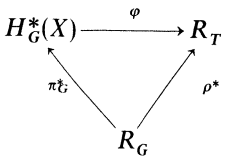
$$\begin{array}{ccc} X_T & \xrightarrow{\tilde{\rho}} & X_G \\ {\scriptstyle\pi_T}\downarrow & & \downarrow{\scriptstyle\pi_G} \\ B_T & \xrightarrow[\rho]{} & B_G \end{array}$$

induces a commutative diagram

$$\begin{array}{ccc} H^*_G(X) & \xrightarrow{\tilde{\rho}^*} & H^*_T(X) \\ {\scriptstyle\pi_G^*}\uparrow & & \uparrow{\scriptstyle\pi_T^*} \\ R_G & \xrightarrow[\rho^*]{} & R_T \end{array}.$$

Here we have the isomorphism $H^*_T(X)\cong H^*_G(X)\otimes_{R_G}R_T$, which follows from [4], p. 38, proposition 1 (ii). First, we remark, that the choice of a fixed point $q\in F(T,X)$ induces a section s in the fibration $X_T\to B_T$ in the following way:
If $b\in B_T$, the choose an element $e\in E_T$ which projects to b. Then set $s(b)=[e,q]\in E_T\times_T X$. It follows $\pi_T\circ s=id_{B_T}$, i.e., s is a section. In particular, we

have $s^* \circ \pi_T^* = id_{R_T}$. Define φ to be the composition $\varphi = s^* \circ \tilde{\rho}^*$. We then have a commutative triangle



as can be shown by a simple diagram chasing argument. In other terms, φ is a graded R_G-algebra homomorphism. We want to describe the homomorphism φ a little bit more closely: Let

$$H_G^*(X) = \frac{R_G[X_1, \ldots, X_n]}{I}, \quad \text{(Th. 1.1, ii)}$$

let $J \subset R_T[X_1, \ldots, X_n]$ be the ideal generated by the elements of I, using the canonical inclusion $R_G[X_1, \ldots, X_n] \subset R_T[X_1, \ldots, X_n]$, then we have

$$H_T^*(X) = \frac{R_T[X_1, \ldots, X_n]}{J}.$$

Now, we write

$$R_T = \frac{R_G[t_1, \ldots, t_n]}{K},$$

where the ideal K is generated by $g_k - g_k(t_1, \ldots, t_n)$, $k = 1, \ldots, n$. Here we put $R_G = \mathbb{R}[g_1, \ldots, g_n]$. Now, s^* is given by $X_j \to w_j$, $j = 1, \ldots, n$, where the w_j are linear polynomials in the t's. So, it follows $\varphi(X_j) = w_j$, $j = 1, \ldots, n$. By Th. 2.1, Cor., the point q can be chosen in such a way, that at least one of the w_j does not vanish. Therefore φ induces nontrivial graded homomorphism of $\mathbb{R}$-algebras:

$$\bar{\varphi}: H^*(X) \cong H_G^*(X) \otimes_{R_G} \mathbb{R} \to R_T \otimes_{R_G} \mathbb{R} \cong H^*(G/T).$$

By the Lemmata 4.1 and 4.2 the homomorphism $\bar{\varphi}$ is epi. Because we are considering graded modules over the graded polynomial ring R_G, the map φ must be an epimorphism too.

As a second corollary of Th. 4.1 we obtain:

Lemma 4.3. *Let X be compact closed oriented manifold with $H^*(X; \mathbb{R}) \in |\mathcal{A}_{\mathbb{R}}|$ and $\chi(X) = |WG|$. Let the compact connected simple Lie group G set on X with finitely many orbit types. Then $H_G^*(X)$ is isomorphic to R_T as a graded R_G-algebra.*

Now, we come to the proof of our main theorems:

Theorem 1. *Let X be a simply connected closed oriented compact manifold with $H^*(X; \mathbb{R}) \cong H^*(G/T; \mathbb{R})$, G a simple compact Lie group, $T \subset G$ its maximal torus. If G acts on X with finitely many orbit types, then $X = G/T$ and every G-action on X is conjugate to the natural action.*

As a corollary we get:

Theorem 1'. *Let G be a compact simple Lie group, $T \subset G$ a fixed maximal torus, let G act on $X = G/T$ with finitely many orbit types. Then the action is conjugate to the natural action.*

Proof of Theorem 1. By the lemma 4.3 the R_G-algebra $H_G^*(X)$ is isomorphic to R_T. It follows that $H_T^*(X)$ is isomorphic to $R_T \otimes_{R_G} R_T$. The lemma 3.2 says that the Weyl group WG acts freely and transitively on the set of associated prime ideals of the $R_T \otimes_{\mathbb{R}} R_T$-module $H_T^*(X)$. This implies that the action of WG on the set $\pi_0 F(T; X)$ of connected components of the fixed set $F(T; X)$ is free and transitive. Let F be a connected component of $F(T; X)$ and let $x \in F$. Denote by G_x the isotropy group of x, let $G_x^0 \lhd G_x$ its one-connected component. Then it follows that the Weyl group WG_x^0 of G_x^0 is trivial, i.e., $G_x^0 = T$. This shows $\dim G(x) = \dim G/T = \dim X$, i.e., $G(x) = X$. Because X is simply connected we have $G_x = G_x^0 = T$. □

References

1. Borel, A.: Topics in the homology theory of fiber bundles, LNM N° 36. Berlin-Heidelberg-New York: Springer 1967
2. Chevalley, C.: Invariants of finite groups generated by reflections. Am. J. Math. **77**, 778–782 (1955)
3. Hauschild, V.: Über den Symmetriegrad der rational azyklischen kompakten homogenen Räume. Manuscr. Math. **32**, 365–379 (1980)
4. Hsiang, W.Y.: Cohomology theory of topological transformation groups, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 85. Berlin-Heidelberg-New York: Springer 1975
5. Hsiang, W.Y., Tomter, P.: Transformation groups on complex Stiefel manifolds. Acta math. **152**:1-2, 107–126 (1984)
6. Slodowy, P.: Simply singularities and simple algebraic groups. Lecture Notes in Math. 815. Berlin-Heidelberg-New York: Springer 1980

Received April 1, 1984; received in final form December 12, 1984

Math. Z. 189, 487–505 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

The Global Cauchy Problem for the Non Linear Klein-Gordon Equation

J. Ginibre [1] and G. Velo [2]

[1] Laboratoire de Physique Théorique et Hautes Energies*, Université de Paris-Sud, F-91405 Orsay Cedex, France

[2] Dipartimento di Fisica, Università di Bologna and INFN, Sezione di Bologna, Italy

1. Introduction

A large amount of work has been devoted to the study of the Cauchy problem for the non linear Klein-Gordon (NLKG) equation

$$\ddot{\varphi} - \Delta\varphi + f(\varphi) = 0, \tag{1.1}$$

where φ is a complex valued function defined in space time $\mathbb{R}^{n+1}$, the upper dot denotes the time derivative, Δ is the Laplace operator in $\mathbb{R}^n$ and f is a non linear complex valued function [3–5, 7, 8, 10–14, 16, 18–24, 28, 29] (Some of those papers are concerned with more general semilinear hyperbolic equations and/or contain additional references to previous work). That problem has been studied mostly by compactness or by contraction methods. In both cases the proof of existence of global solutions makes use at some stage of the conservation of the energy, thereby yielding solutions which are functions of time with values in a space not larger than the energy space X_e (see definition in Sect. 2). The assumptions on the interaction are best visualized on the example of the sum of two powers, namely

$$f(\varphi) = \sum_{j=1,2} \lambda_j |\varphi|^{p_j - 1}\varphi \tag{1.2}$$

with $1 \leqq p_1 \leqq p_2 < \infty$ and $\lambda_j \geqq 0$. All the uniqueness proofs are of a perturbative nature and an elementary power counting argument shows that they require $p_2 < 1 + 4/(n-2)$. The available results include the existence of weak global solutions in the energy space with no upper bound on p_2 [11, 19, 23], their uniqueness for $p_2 \leqq 1 + 4n/[(n+1)(n-2)]$ [7], the existence of unique global strongly continuous solutions in the energy space under the assumption $p_2 \leqq 1 + 2/(n-2)$ [5, 8, 10, 20, 28, 29], and the existence of unique global strongly continuous solutions under the assumption $p_2 < 1 + 4/(n-2)$, but in spaces

* Laboratoire associé au Centre National de la Recherche Scientifique

smaller than X_e and with limitations on n, namely for $n \leqq 3$ and φ not only in X_e, but also bounded and continuous [10, 22], and for $n \leqq 9$ and $\varphi \in H^k(\mathbb{R}^n)$, $k \geqq n/2$ [3, 4, 13, 14]. Also, global classical solutions are obtained in the last cases.

The main purpose of the present paper is to prove the uniqueness of weak solutions and the existence and uniqueness of global strongly continuous solutions in the energy space under the natural assumption $p_2 < 1 + 4/(n-2)$ and for arbitrary n. The corresponding results are stated below in precise form in Proposition 3.1 and Proposition 3.2 respectively.

We rely for the existence result on the treatment given in [11], which we recall briefly in Sect. 2 with some minor additions. The proof of uniqueness follows the same pattern as that of the corresponding property for the non linear Schrödinger (NLS) equation [6] and combines a contraction method on bounded sets of the energy space with suitable space-time integrability estimates of the solutions of (1.1). The proof is given in Sect. 3.

We conclude this introduction by giving the main notation used in this paper. We restrict our attention to $n \geqq 2$ since the simpler special case $n = 1$ would require slightly modified statements. We use the notation $2^* = 2n/(n-2)$. We denote by $\| \cdot \|_r$ the norm in $L^r \equiv L^r(\mathbb{R}^n)$. Pairs of conjugate indices are written as $r, \bar{r}$, where $2 \leqq r \leqq \infty$ and $r^{-1} + \bar{r}^{-1} = 1$. For any integer k, we denote by $H^k \equiv H^k(\mathbb{R}^n)$ the usual Sobolev spaces. For any interval I of $\mathbb{R}$, we denote by $\bar{I}$ the closure of I. For any Banach space B, we denote by $\mathscr{C}(I, B)$ the space of strongly continuous functions from I to B, by $\mathscr{C}_w(I, B)$ the space of weakly continuous functions from I to B, and by $\mathscr{C}^\alpha(I, B)$, $0 < \alpha < 1$, (resp. by $\mathscr{C}^L(I, B)$) the space of functions from I to B that are strongly Hölder continuous with exponent α (resp. strongly Lipschitz continuous), uniformly in compact subintervals of I. For any q, $1 \leqq q \leqq \infty$, we denote by $L^q(I, B)$ (resp. $L^q_{\mathrm{loc}}(I, B)$) the space of measurable functions φ from I to B such that $\|\varphi(\cdot)\|_B \in L^q(I)$ (resp. $\|\varphi(\cdot)\|_B \in L^q_{\mathrm{loc}}(I)$). If I is open we denote by $\mathscr{D}'(I, B)$ the space of vector valued distributions from I to B [11].

We shall need the operators $\omega = (-\Delta)^{1/2}$, $K(t) = \omega^{-1} \sin \omega t$ and $\dot{K}(t) = \cos \omega t$. The operators $K(t)$ and $\dot{K}(t)$ are bounded and strongly continuous with respect to t in H^k for any k.

2. Existence of Solutions

In this section we briefly recall the existence result of weak global solutions of (1.1) in the energy space that can be obtained by the compactness method. We state the result without proof and refer to [11] for details and to [6] for a similar treatment for the NLS equation. We introduce the following assumptions on f:

(H1) $f \in \mathscr{C}(\mathbb{C}, \mathbb{C})$ and for some p, $1 \leqq p < \infty$ and for all $z \in \mathbb{C}$,

$$|f(z)| \leqq C(|z| + |z|^p). \tag{2.1}$$

(H2) There exists a function $V \in \mathscr{C}^1(\mathbb{C}, \mathbb{R})$ such that $V(0)=0$, $V(z)=V(|z|)$ for all $z \in \mathbb{C}$ and $f(z)=\partial V/\partial \bar{z}$. For all $\rho \in \mathbb{R}^+$, V satisfies the estimate

$$V(\rho) \geqq -a^2 \rho^2 \tag{2.2}$$

for some $a \geqq 0$.

For $(\varphi, \psi) \in H^1 \oplus L^2$ and such that $V(\varphi) \in L^1$, the energy is defined by

$$E(\varphi, \psi) = \|\psi\|_2^2 + \|\nabla \varphi\|_2^2 + \int dx\, V(\varphi(x)). \tag{2.3}$$

The assumption (H2) formally implies the conservation of the energy for the equation (1.1).

We introduce the Banach space $X = H^1 \cap L^{p+1}$. The dual space is $X' = H^{-1} + \overline{L^{p+1}}$, the duality being realized through the scalar product $\langle \varphi_1, \varphi_2 \rangle$ in L^2, linear in φ_2 and antilinear in φ_1. We define the energy space $X_e = X \oplus L^2$. We first state some elementary properties of the solutions of the equation (1.1).

Lemma 2.1. *Let f satisfy* (H1), *let I be a bounded open interval of $\mathbb{R}$ and let $\varphi \in L^\infty(I, X)$. Then*

(1) $f(\varphi) \in L^\infty(I, L^2 + \overline{L^{p+1}})$.

Let in addition φ satisfy the equation (1.1) *in $\mathscr{D}'(I, X')$. Then*

(2) $\ddot{\varphi} \in L^\infty(I, X')$ *and* $\dot{\varphi} \in \mathscr{C}^L(I, X')$. *If in addition* $\dot{\varphi} \in L^\infty(I, L^2)$, *then* $\dot{\varphi} \in \mathscr{C}_w(I, L^2)$ *and*

$$\varphi \in \mathscr{C}_w(\bar{I}, X) \cap \mathscr{C}^L(\bar{I}, L^2) \cap \bigcap_{2 < r < \mathrm{Max}(p+1, 2^*)} \mathscr{C}^{\alpha(r)}(I, L^r)$$

where

$$\alpha(r) = 1 - \delta(r)\, \mathrm{Min}\{1, \delta(p+1)^{-1}\} \tag{2.4}$$

with $\delta(r) = n/2 - n/r$.

(3) *For any $t, s \in \bar{I}$, φ satisfies the equation*

$$\varphi(t) = \dot{K}(t-s)\, \varphi(s) + K(t-s)\, \dot{\varphi}(s) - \int_s^t d\tau\, K(t-\tau)\, f(\varphi(\tau)) \tag{2.5}$$

where the integral is a Bochner integral in H^{-k}, $k \geqq \mathrm{Max}\{0, \delta(p+1)-1\}$.

The proof is similar to that of Lemma 2.1 of [6] and will be omitted.

Remark 2.1. If φ is a solution of (1.1) in $L^\infty(I, X)$, then $\ddot{\varphi} \in L^\infty(I, X')$ and, by an elementary computation, one obtains, at least formally, for any $a, b \in I$,

$$\begin{aligned} &\int_{a \leqq s \leqq t \leqq b} ds\, dt\, \|\dot{\varphi}(s) - \dot{\varphi}(t)\|_2^2 \\ &= \int_a^b d\tau\, \langle \ddot{\varphi}(\tau), (b-\tau)\, \varphi(a) + (\tau - a)\, \varphi(b) - (b-a)\, \varphi(\tau) \rangle \\ &\leqq 2\, |b-a|^2\, \|\ddot{\varphi}\|_{L^\infty([a,b], X')}\, \|\varphi\|_{L^\infty([a,b], X)}. \end{aligned}$$

However that estimate is not sufficient to prove that $\dot{\varphi} \in L^\infty(I, L^2)$ and we assume that last fact explicitly.

We now state the basic existence result.

Proposition 2.1. *Let f satisfy* (H1) *and* (H2). *If $p+1>2^*$ assume in addition that*

$$V(\rho) \geqq -a^2\rho^2 + c\rho^{p+1} \tag{2.6}$$

for some $c>0$ and all $\rho \in \mathbb{R}^+$. Let $t_0 \in \mathbb{R}$ and let $(\varphi_0, \psi_0) \in X \oplus L^2$. Then the equation (1.1) *has a solution*

$$\varphi \in L^\infty_{\mathrm{loc}}(\mathbb{R}, X) \cap \mathscr{C}_w(\mathbb{R}, X) \cap \mathscr{C}^L(\mathbb{R}, L^2) \cap \bigcap_{2<r<\mathrm{Max}(p+1, 2^*)} \mathscr{C}^{\alpha(r)}(\mathbb{R}, L^r)$$

(*where $\alpha(r)$ is defined by* (2.4)) *with*

$$\dot{\varphi} \in L^\infty_{\mathrm{loc}}(\mathbb{R}, L^2) \cap \mathscr{C}_w(\mathbb{R}, L^2) \cap \mathscr{C}^L(\mathbb{R}, X')$$

and with $\varphi(t_0)=\varphi_0$, $\dot{\varphi}(t_0)=\psi_0$. For any $t \in \mathbb{R}$, φ satisfies the estimates

$$\|\varphi(t)\|_2 \leqq e(t-t_0) \tag{2.7}$$

$$\|\dot{\varphi}(t)\|_2^2 + \|\nabla\varphi(t)\|_2^2 + c\|\varphi(t)\|_{p+1}^{p+1} \leqq \dot{e}(t-t_0)^2 \tag{2.8}$$

where

$$e(\tau) \equiv \|\varphi_0\|_2 \cosh(a\tau) + (E(\varphi_0, \psi_0) + a^2\|\varphi_0\|_2^2)^{1/2} a^{-1} \sinh(a|\tau|). \tag{2.9}$$

Assume in addition that the map $\varphi \mapsto \int dx\, V(\varphi(x))$ is weakly lower semi-continuous from X to $\mathbb{R}$ on the bounded sets of X. Then, for all $t \in \mathbb{R}$, φ satisfies the energy inequality

$$E(\varphi(t), \dot{\varphi}(t)) \leqq E(\varphi_0, \psi_0). \tag{2.10}$$

The proof is given in [11]. We simply remark that the finite dimensional approximants to the solution can be shown to converge pointwise in time weakly in X_e, thereby yielding the energy inequality (2.10) for each t. The basic estimates (2.7) and (2.8) are proved first for the finite dimensional approximants by an elementary computation: formally the conservation of the energy and the condition (2.6) imply

$$\begin{aligned} E \equiv E(\varphi_0, \psi_0) &\geqq \|\dot{\varphi}\|_2^2 + \|\nabla\varphi\|_2^2 - a^2\|\varphi\|_2^2 + c\|\varphi\|_{p+1}^{p+1} \\ &\geqq \|\dot{\varphi}\|_2^2 - a^2\|\varphi\|_2^2 \end{aligned} \tag{2.11}$$

so that the quantity $y = a\|\varphi\|_2$ satisfies the inequality .

$$|\dot{y}| \leqq a(E+y^2)^{1/2}. \tag{2.12}$$

Integrating (2.12) yields (2.7), and substituting (2.7) into the first inequality in (2.11) yields (2.8).

The NLKG equation differs from the NLS equation most notably by the lack of conservation of the L^2-norm. As a consequence we are not able to prove that the approximants to the solution φ tend to φ strongly in L^2 and we need a stronger semi-continuity (in fact convexity) condition on V in order to

establish the energy inequality. This stronger condition will be removed in Sect. 3 under the assumptions that ensure the uniqueness of the solution at the cost of an additional approximation procedure.

3. Uniqueness of Solutions

In this section, we prove that under stronger assumptions on the interaction, the solutions obtained in Sect. 2 are unique. We replace (H1) by the following stronger assumption.

(H3) $f\in\mathscr{C}^1(\mathbb{C},\mathbb{C})$, $f(0)=0$ and, for $n\geqq 2$,

$$|f'(z)|\equiv \mathrm{Max}\,\{|\partial f/\partial z|, |\partial f/\partial \bar{z}|\}\leqq C(1+|z|^{p-1}) \tag{3.1}$$

where

$$0\leqq p-1<4/(n-2). \tag{3.2}$$

Clearly (H3) implies (H1) and the condition (3.2) implies that $H^1\subset L^{p+1}$, so that $X=H^1$ and $X_e=H^1\oplus L^2$.

The proof of uniqueness proceeds by a contraction method. In dimension $n=1, 2$, under the assumptions (H2) and (H3), one can solve the equation (1.1) by a contraction method in $\mathscr{C}(\mathbb{R}, X_e)$, and uniqueness in $L^\infty_{\mathrm{loc}}(\mathbb{R}, X_e)$ follows immediately from the same estimates. In dimensions $n\geqq 3$, the proof of uniqueness requires some space time integrability properties of the solutions, which we derive below. We shall use the homogeneous Besov and Sobolev spaces of arbitrary order and the associated Sobolev inequalities. For the convenience of the reader, we briefly recall in the Appendix the definitions and some basic properties of those spaces [1, 27]. We use notations such as $\dot{B}^\sigma_{p,q}$ and $\dot{H}^\sigma_p$ for those spaces, and for the legibility of the formulas, we denote by $\|\cdot;B\|$ and $\|\cdot;L^q(I,B)\|$ the norms in B and in $L^q(I,B)$ when B is such a space. With each r, $1\leqq r\leqq\infty$, it is convenient to associate the variables $\beta(r)$, $\gamma(r)$ and $\delta(r)$ defined by

$$2\beta(r)/(n+1)=\gamma(r)/(n-1)=\delta(r)/n=1/2-1/r. \tag{3.3}$$

With this notation the basic estimate used in this section can be written as:

$$\|\exp(i\omega t)\varphi;\dot{B}^{-\beta(r)}_{r,2}\|=C_r|t|^{-\gamma(r)}\|\varphi;\dot{B}^{\beta(r)}_{r',2}\| \tag{3.4}$$

for all $t\in\mathbb{R}\setminus\{0\}$, $2\leqq r\leqq\infty$, $\varphi\in\dot{B}^{\beta(r)}_{r',2}$. A proof of (3.4) can be found in [2, 13].

We now recall some space time integrability properties of the solutions of the free equation [21, 25, 15]. We restrict our attention to the case $n\geqq 3$.

Lemma 3.1. (1) *Let r and q satisfy*

$$0\leqq 2/q=\gamma(r)<1. \tag{3.5}$$

Then, for any $\varphi\in L^2$, $\exp(i\omega t)\varphi$ belongs to $L^q(\mathbb{R},\dot{B}^{-\beta(r)}_{r,2})$ and satisfies the estimate

$$\|\exp(i\omega t)\varphi;L^q(\mathbb{R},\dot{B}^{-\beta(r)}_{r,2})\|\leqq C\|\varphi\|_2. \tag{3.6}$$

(2) *Let r and ρ satisfy*

$$\begin{aligned} &0 \leqq \delta(r) \leqq n/2 \\ &0 \leqq \rho + \delta(r) - 1 < 1/2 \\ &\rho \leqq 1 - \beta(r), \end{aligned} \tag{3.7}$$

and let q satisfy

$$1/q = \rho + \delta(r) - 1. \tag{3.8}$$

Then, for any $(\varphi, \psi) \in \dot{H}^1_2 \oplus L^2$, $K\psi + \dot{K}\varphi \in L^q(\mathbb{R}, \dot{B}^\rho_{r,2})$ and the following estimate holds

$$\|K\psi + \dot{K}\varphi; L^q(\mathbb{R}, \dot{B}^\rho_{r,2})\| \leqq C(\|\psi\|_2 + \|\nabla\varphi\|_2). \tag{3.9}$$

Proof. (1) Let $U(t) = \exp(i\omega t)$ and $\beta = \beta(r)$.

By density, it is sufficient to prove (3.6) for $\varphi \in \mathscr{S}(\mathbb{R}^n)$, and by duality and density again, it is sufficient to prove that

$$|\langle \theta, U(\cdot)\varphi \rangle_{n+1}| \leqq C\|\varphi\|_2 \|\theta; L^{\bar{q}}(\mathbb{R}, \dot{B}^\beta_{\bar{r},2})\| \tag{3.10}$$

for any $\theta \in \mathscr{Z}^n_{n+1}$, where

$$\mathscr{Z}^n_{n+1} = \{u \in \mathscr{S}(\mathbb{R}^{n+1}) : D^\alpha_\xi \hat{u}(\xi_0, \xi) = 0 \text{ for any multiindex } \alpha\}$$

and $\langle \cdot, \cdot \rangle_{n+1}$ denotes the L^2-duality in $n+1$ variables. Now

$$\begin{aligned} |\langle \theta, U(\cdot)\varphi \rangle_{n+1}| &= |\langle \int d\tau\, U(-\tau)\theta(\tau, \cdot), \varphi \rangle_n| \\ &\leqq \|\varphi\|_2 \|\int d\tau\, U(-\tau)\theta(\tau, \cdot)\|_2. \end{aligned} \tag{3.11}$$

The last term in (3.11) is estimated by

$$\begin{aligned} \|\int d\tau\, U(-\tau)\theta(\tau, \cdot)\|_2^2 &= \int dt \langle \theta(t, \cdot), \int d\tau\, U(t-\tau)\theta(\tau, \cdot) \rangle_n \\ &\leqq \|\theta; L^{\bar{q}}(\mathbb{R}, \dot{B}^\beta_{\bar{r},2})\| \, \|\int d\tau\, U(\cdot - \tau)\theta(\tau); L^q(\mathbb{R}, \dot{B}^{-\beta}_{r,2})\|. \end{aligned} \tag{3.12}$$

Using (3.4) and the Hardy-Littlewood-Sobolev inequality in the time variable ([9], p. 117) we can continue (3.12) as

$$\ldots \leqq C\|\theta; L^{\bar{q}}(\mathbb{R}, \dot{B}^\beta_{\bar{r},2})\|^2$$

which implies (3.10) and thereby completes the proof.

(2) By the definition of K and $\dot{K}$ it is sufficient to prove that

$$\|\exp(i\omega t)\psi; L^q(\mathbb{R}, \dot{B}^{\rho-1}_{r,2})\| \leqq C\|\psi\|_2. \tag{3.13}$$

Under the assumptions (3.7) and (3.8), (3.13) follows from (3.6) through the Sobolev inequalities (A.12). Q.E.D.

Note that $K(t)\psi + \dot{K}(t)\varphi$ is the solution of the free equation with initial data (φ, ψ) at $t = 0$ and that, for $(\varphi, \psi) \in H^1 \oplus L^2$, $K(t)\psi + \dot{K}(t)\varphi \in L^\infty_{\mathrm{loc}}(\mathbb{R}, L^2)$ with $\|K(t)\psi + \dot{K}(t)\varphi\|_2 \leqq t\|\psi\|_2 + \|\varphi\|_2$.

We shall need the following estimate on the interaction term in the equation (1.1). That result is an improvement and a simplification of Lemma 3.2 in [6].

Lemma 3.2. *Let $f \in \mathscr{C}^1(\mathbb{C}, \mathbb{C})$ with $|f'(z)| = C_0|z|^{p-1}$ for some p, $1 \leqq p < \infty$. Let $0 < \lambda < 1$, let $1 \leqq \bar{l} \leqq k \leqq \infty$, $1 \leqq m \leqq \infty$ and $1/s = 1/\bar{l} - 1/k$. Then the following in-*

equality holds

$$\|f(\varphi); \dot{B}^{\lambda}_{l,m}\| \leqq C \|\varphi; \dot{B}^{\lambda}_{k,m}\| \, \| |\varphi|^{p-1} \|_s \tag{3.14}$$

for all φ *such that* $\varphi \in L^q$ *for some* q, $1 \leqq q < \infty$, *and such that the norms in the right hand side are finite. The same result holds for* $\lambda = 1$, $1 < l \leqq 2 = k = m$.

Proof. We use the fact that for $0 < \lambda < 1$, the Besov spaces $\dot{B}^{\lambda}_{l,m}$ can be defined through the equivalent norm [1]

$$\|v; \dot{B}^{\lambda}_{l,m}\|_{\#} = \left\{ \int_0^{\infty} \frac{dt}{t} \{ \operatorname*{Sup}_{|y| \leqq t} t^{-\lambda} \|\tau_y v - v\|_l \}^m \right\}^{1/m} \tag{3.15}$$

where τ_y denotes the space translation by $y \in \mathbb{R}^n$. In order to use (3.15) with $v = f(\varphi)$, we note that

$$|\tau_y f(\varphi) - f(\varphi)| \leqq |\tau_y \varphi - \varphi| \int_0^1 d\alpha \, |f'(\alpha \tau_y \varphi + (1-\alpha)\varphi)| \tag{3.16}$$

and therefore by the Hölder inequality

$$\|\tau_y f(\varphi) - f(\varphi)\|_l \leqq C_0 \|\tau_y \varphi - \varphi\|_k \, \| |\varphi|^{p-1} \|_s \tag{3.17}$$

from which (3.14) follows immediately.

For $\lambda = 1$, we use the fact that for $m = 2$, $\dot{B}^1_{l,2} \supset \dot{F}^1_{l,2} \equiv \dot{H}^1_l$ by (A.7), so that

$$\|f(\varphi); \dot{B}^1_{l,2}\| \leqq C \|\nabla f(\varphi)\|_l \leqq C \|\varphi; \dot{B}^1_{2,2}\| \, \| |\varphi|^{p-1} \|_s. \tag{3.18}$$

Q.E.D.

We are now in a position to derive the relevant integrability properties of the solutions of the equation (1.1).

Lemma 3.3. *Let* $n \geqq 3$. *Let* f *satisfy* (H3). *Let* I *be an open interval, let* $t_0 \in I$, *let* $(\varphi_0, \psi_0) \in H^1 \oplus L^2$ *and let* $\varphi \in L^{\infty}_{\mathrm{loc}}(I, H^1)$ *be a solution of the equation* (1.1) *with* $\varphi(t_0) = \varphi_0$ *and* $\dot{\varphi}(t_0) = \psi_0$. *Let* ρ, r *satisfy* (3.7) *and let* q *satisfy* $1/q \geqq \rho + \delta(r) - 1$. *Then* $\varphi \in L^q_{\mathrm{loc}}(I, \dot{B}^{\rho}_{r,2})$ *and, for any compact interval* $J \subset I$, *with* $t_0 \in J$, φ *is estimated in* $L^q(J, \dot{B}^{\rho}_{r,2})$ *in terms of the norm of* φ *in* $L^{\infty}(J, H^1)$ *and of the norm of* (φ_0, ψ_0) *in* $H^1 \oplus L^2$. *If in addition* $r < \infty$, *the same results hold with* $\dot{B}^{\rho}_{r,2}$ *replaced by* $\dot{H}^{\rho}_r$.

Proof. The proof is based on the integral equation (2.5) or

$$\varphi(t) = \dot{K}(t - t_0)\varphi_0 + K(t - t_0)\psi_0 - \int_{t_0}^{t} d\tau \, K(t - \tau) f(\varphi(\tau)) \tag{3.19}$$

and on estimates on its integrand. In addition to the variables ρ and r considered in Lemma 3.1, it is convenient to use also the variable $\sigma = \rho + \delta(r) - 1$. The second condition in (3.7) then becomes $0 \leqq \sigma < 1/2$. We first take r with $1 \leqq \delta(r) \leqq n/2$, and σ', σ'' such that $0 \leqq \sigma' < \sigma'' \leqq \sigma' + \varepsilon < 1/2$ with ε sufficiently small, to be chosen later, and we prove that for any $\varphi \in H^1 \cap \dot{B}^{\rho'}_{r,2}$ and for any $t \neq 0$, $K(t) f(\varphi) \in \dot{B}^{\rho''}_{r,2}$ and $K(t) f(\varphi)$ satisfies the estimate

$$\|K(t) f(\varphi); \dot{B}^{\rho''}_{r,2}\| \leqq M(1 + |t|^{-\gamma}(1 + \|\varphi; \dot{B}^{\rho'}_{r,2}\|^{\nu})) \tag{3.20}$$

for some γ, ν such that $0 \leqq \gamma < 1$, $0 \leqq \nu \leqq 1$, and with M depending only on γ, ν and ε, but not otherwise on ρ' and ρ'', and depending on φ through $\|\varphi\|_{H^1}$

only. By (3.1), we can decompose f as $f=f_1+f_2$ with $|f_1'(z)|\leqq C$ and $|f_2'(z)|\leqq C|z|^{p-1}$ and we estimate separately the contributions of f_1 and f_2 to (3.20). The contribution of f_1 is estimated by the use of (A.12) as

$$\|K(t)f_1(\varphi);\dot{B}^{\rho''}_{r,2}\|\leqq C\|K(t)f_1(\varphi);\dot{B}^{1+\sigma''}_{2,2}\|\leqq C\|\varphi;\dot{H}^{\sigma''}_2\| \tag{3.21}$$

thereby yielding the first term in the right hand side of (3.20). The contribution of f_2 is estimated by (A.12) and (3.4) as

$$\begin{aligned}\|K(t)f_2(\varphi);\dot{B}^{\rho''}_{r,2}\|&\leqq C\|K(t)f_2(\varphi);\dot{B}^{1+\lambda-2\beta(l)}_{l,2}\|\\&\leqq C|t|^{-\gamma(l)}\|f_2(\varphi);\dot{B}^{\lambda}_{l,2}\|\end{aligned} \tag{3.22}$$

provided $r\geqq l$ and

$$\lambda=\sigma''+\delta(l)/n. \tag{3.23}$$

The latter implies $0<\lambda<1$ for $0\leqq\sigma''<1/2$ and $2<r\leqq\infty$. We estimate the last norm in (3.22) by Lemma 3.2 as

$$\|f_2(\varphi);\dot{B}^{\lambda}_{l,2}\|\leqq C\|\varphi;\dot{B}^{\lambda}_{k,2}\|\,\||\varphi|^{p-1}\|_s \tag{3.24}$$

with

$$n/s=\delta(l)+\delta(k). \tag{3.25}$$

We estimate the last two norms in (3.24) by interpolation between the norms in L^2, in $\dot{B}^1_{2,2}\equiv\dot{H}^1_2$ and in $\dot{B}^{\rho'}_{r,2}$ and by the Sobolev inequalities through the use of Lemmas (A.1) and (A.2). Since $\delta(r)\geqq 1$, we have $\lambda\geqq\sigma''\geqq\sigma'\geqq\rho'$. The interpolation is then possible provided

$$0\leqq\delta(k)\leqq\delta(r)(1-\lambda)/(1-\rho') \tag{3.26}$$

and

$$0\leqq\delta((p-1)s)\leqq 1+\mathrm{Min}\left\{\sigma',\frac{\sigma'}{1-\rho'}\right\}\equiv 1+\frac{\sigma'}{1+\rho'_-} \tag{3.27}$$

where $\rho'_\pm=\mathrm{Max}(0,\pm\rho')$, or equivalently

$$(p-1)(n/2-1-\sigma'/(1+\rho'_-))\leqq n/s\leqq(p-1)n/2. \tag{3.28}$$

After elimination of s through (3.25), we see that the interpolation conditions are satisfied, namely there exists k such that (3.26) and (3.27) hold, provided $(p-1)n/2\geqq\delta(l)$ and

$$(p-1)(n/2-1-\sigma'/(1+\rho'_-))\leqq\delta(l)+\delta(r)(1-\lambda)/(1-\rho'). \tag{3.29}$$

The lower interpolation condition is easily satisfied within the admissible range of p, possibly after replacing the original p by a larger one, and the upper condition (3.29) will be considered below.

When the interpolation is possible, and since s is a decreasing function of k by (3.25), one can always arrange for $\lambda+\delta(k)-1$ and $\delta((p-1)s)-1$ to have the same sign. If both are negative, one can estimate both norms in the right hand side of (3.24) (modulo Sobolev inequalities) in terms of the norms of φ in L^2 and $\dot{H}^1_2$, thereby obtaining an estimate of (3.24) in terms of the norm of φ in

H^1. If both are positive, one can estimate both norms in the right hand side of (3.24) in terms of the norms of φ in $\dot{H}^1_2$ and $\dot{B}^{\rho'}_{r,2}$, thereby obtaining

$$\|f_2(\varphi); \dot{B}^{\lambda}_{l,2}\| \leqq C \|\nabla\varphi\|_2^{p-\nu} \|\varphi; \dot{B}^{\rho'}_{r,2}\|^{\nu} \tag{3.30}$$

with ν simply given by homogeneity, namely

$$(p-1)(n/2-1)=\delta(l)+1-\lambda+\nu\sigma'. \tag{3.31}$$

Substituting (3.30) into (3.22) and using the previous estimate of the term with f_1 yields (3.20) with $\gamma=\gamma(l)$ provided p satisfies (3.29). The condition $\nu\leqq 1$ becomes after elimination of λ through (3.23)

$$(p-1)(n/2-1)\leqq 1+\gamma(l)-\sigma''+\sigma' \tag{3.32}$$

and is satisfied for $\gamma(l)=1-\varepsilon$, $\sigma''\leqq\sigma'+\varepsilon$, provided

$$p-1\leqq 4(1-\varepsilon)/(n-2) \tag{3.33}$$

which can be ensured for $\varepsilon>0$, ε sufficiently small, under the assumption (H3).

We now consider the interpolation condition (3.29). By (3.23) that condition holds for $\sigma''\leqq\sigma'+\varepsilon$ provided

$$\begin{aligned}&(p-1)(n/2-1-\sigma'/(1+\rho'_-))\\ &\quad\leqq 1+\gamma(l)-\varepsilon-[\sigma'/(1-\rho')][\delta(r)-1+\varepsilon+\delta(l)/n].\end{aligned} \tag{3.34}$$

The latter condition reduces to

$$p-1\leqq 2(1+\gamma(l)-\varepsilon)/(n-2) \tag{3.35}$$

for $\sigma'=0$ and is then satisfied for $\gamma(l)=1-\varepsilon$ under the condition (3.33). It is implied by (3.35) for all $\sigma'\geqq 0$ such that

$$(n-2)(\delta(r)-1+\varepsilon+\delta(l)/n)\leqq 2(1+\gamma(l)-\varepsilon)(1-\rho'_+). \tag{3.36}$$

For $\rho'\leqq 0$, (3.36) reduces to

$$(n-2)\gamma(r)\leqq(n-1)(1-\varepsilon)+\gamma(l) \tag{3.37}$$

and is satisfied for $\gamma(l)=1-\varepsilon$ and $\gamma(r)\leqq(1-\varepsilon)n/(n-2)$, and in particular for $\gamma(r)=\gamma(l)=1-\varepsilon$. For $\rho'\geqq 0$ (3.36) reduces to

$$\delta(r)(1-A)\leqq 1-\varepsilon-\delta(l)/n-A\sigma' \tag{3.38}$$

with $A=2(1-\varepsilon+\gamma(l))/(n-2)$. That condition is easily seen to hold for all σ', $0\leqq\sigma'\leqq 1/2$, and $\gamma(r)=\gamma(l)=1-\varepsilon$ for ε sufficiently small (actually $\varepsilon\leqq 3/(2n)$).

This completes the proof of (3.20) with $\gamma(r)=\gamma(l)=\gamma=1-\varepsilon$, $\nu\leqq 1$ and $\varepsilon>0$, ε sufficiently small.

We now use (3.20) to complete the proof of Lemma 3.3. Let $\varphi\in L^{\infty}_{\mathrm{loc}}(I,H^1)$ be solution of the equation (1.1). We first estimate φ in $L^q_{\mathrm{loc}}(I,\dot{B}^{\rho}_{r,2})$ for $\gamma(r)=1-\varepsilon$ with ε sufficiently small and for values of ρ that increase from $\rho=-(1-n\varepsilon)/(n-1)$, corresponding to $\sigma=0$, to $\rho=1/2-(1-(n+1)\varepsilon/2)/(n-1)$ corresponding to

the upper limit allowed by the third condition in (3.7). For that purpose, we estimate iteratively the right hand side of the integral equation (3.19) in $L^q_{\mathrm{loc}}(I, \dot{B}^{\rho''}_{r,2})$ for successive values of ρ'' increasing in the previous range in steps of ε and for $1/q=\sigma''$. Using (3.20), we obtain at each step

$$\|\varphi(t);\dot{B}^{\rho''}_{r,2}\| \leqq \|\dot{K}(t-t_0)\varphi_0+K(t-t_0)\psi_0;\dot{B}^{\rho''}_{r,2}\| + M\int_{t_0}^{t} d\tau\{1+|t-\tau|^{\varepsilon-1}(1+\|\varphi;\dot{B}^{\rho'}_{r,2}\|^{\nu})\}. \tag{3.39}$$

The norm of the free term in (3.39) is in $L^q(\mathbb{R})$ by Lemma 3.1. Applying the Young inequality to the time integral we obtain for any compact interval J containing t_0

$$\|\varphi;L^q(J,\dot{B}^{\rho''}_{r,2})\| \leqq \|\dot{K}(t-t_0)\varphi_0+K(t-t_0)\psi_0;L^q(\mathbb{R},\dot{B}^{\rho''}_{r,2})\| + MC(J)(1+\|\varphi;L^{vq}(J,\dot{B}^{\rho'}_{r,2})\|^{\nu}) \tag{3.40}$$

where M depends only on the norm of φ in $L^\infty(J,H^1)$ and $C(J)$ depends only on $|J|$. The last norm in (3.40) is controlled from the previous step of the iteration since $vq\leqq q=1/\sigma''\leqq 1/\sigma'$. After a finite number of steps, we obtain that $\varphi\in L^q(J,\dot{B}^{\rho}_{r,2})$ for all ρ in the previous range and $1/q=\sigma$. Furthermore, the norm of φ in $L^q(J,\dot{B}^{\rho}_{r,2})$ depends on φ only through the constant M in (3.20) and through the norm of the free term, which are estimated in terms of the norm of φ in $L^\infty(J,H^1)$ and of the norm of (φ_0,ψ_0) in $H^1\oplus L^2$. This completes the proof of the result for (ρ,r) in the subset $\gamma(r)=1-\varepsilon$ of the set defined by (3.7), with ε small and positive. The result for general (ρ,r) in that set follows then from that special case by the Sobolev inequality (A.12) in the region $\gamma(r)\geqq 1$ and by interpolation with local boundedness in time for $\sigma=0$, which follows from local boundedness in time of the norm in H^1, in the region $0\leqq\gamma(r)<1$.

The last statement of Lemma 3.3 follows from the embedding (A.8). Q.E.D.

We are now in a position to prove the uniqueness of the solutions of (1.1) under the assumption (H3).

Proposition 3.1. *Let f satisfy* (H3). *Let I be an open interval, let $t_0\in\bar{I}$ and let $(\varphi_0,\psi_0)\in H^1\oplus L^2$. Then there exists at most one $\varphi\in L^\infty(I,H^1)$ which satisfies the equation* (1.1) *in $\mathscr{D}'(I,H^1)$ and the initial condition $\varphi(t_0)=\varphi_0$, $\dot{\varphi}(t_0)=\psi_0$.*

Proof. We restrict our attention to $n\geqq 3$. Let φ_1 and φ_2 in $L^\infty(I,H^1)$ be two solutions of the equation (1.1) with the same initial data at time t_0. From Lemma 2.1, part (3), we obtain

$$\varphi_1(t)-\varphi_2(t)=-\int_{t_0}^{t} d\tau\, K(t-\tau)(f(\varphi_1(\tau))-f(\varphi_2(\tau))) \tag{3.42}$$

where the integral is taken in L^2. We estimate (3.42) in $L^{r'}$ for some r' satisfying $0\leqq\gamma'\equiv\gamma(r')\leqq(n-1)/(n+1)$, to be chosen more precisely below. We decompose f as $f=f_1+f_2$ as in the proof of Lemma 3.3 and we treat separately the contributions of f_1 and f_2. The term with f_1 is estimated by

$$\|K(t-\tau)(f_1(\varphi_1(\tau))-f_1(\varphi_2(\tau)))\|_{r'} \leqq C|t-\tau|\,\|\varphi_1(\tau)-\varphi_2(\tau)\|_{r'} \tag{3.43}$$

where we have used the estimate [17]

$$\|K(t)\psi\|_{r'} \leqq C|t|\,\|\psi\|_{r'} \tag{3.44}$$

(which holds in the larger range $|\gamma(r')| \leqq 1$).

The term with f_2 is estimated by the use of (3.4) and (A.8)(A.1) as

$$\begin{aligned}
&\|K(t-\tau)(f_2(\varphi_1(\tau))-f_2(\varphi_2(\tau)))\|_{r'} \\
&\quad \leqq C|t-\tau|^{-\gamma'}\,\|f_2(\varphi_1(\tau))-f_2(\varphi_2(\tau))\|_{s'} \\
&\quad \leqq C|t-\tau|^{-\gamma'}\,\|\varphi_1(\tau)-\varphi_2(\tau)\|_{r'} \sum_{i=1,2} \|\varphi_i(\tau)\|^{p-1}_{(p-1)l'}
\end{aligned} \tag{3.45}$$

with $n/s' = 1/r' + (n+1)/2$ and $n/l' = 1+\gamma'$. Substituting (3.43) and (3.45) into (3.42) and estimating the time integral by the Hölder inequality, we obtain

$$\begin{aligned}
\|\varphi_1(t)-\varphi_2(t)\|_{r'} \leqq C \operatorname*{Sup}_{t_0 \leqq \tau \leqq t} &\|\varphi_1(\tau)-\varphi_2(\tau)\|_{r'} \\
&\cdot \{|t-t_0|^2 + |t-t_0|^{1-\gamma'-1/m'} \sum_{i=1,2} |\varphi_i|^{p-1}_{(p-1)l',(p-1)m',[t_0,t]}\}
\end{aligned} \tag{3.46}$$

with

$$1/m' < 1-\gamma'. \tag{3.47}$$

We estimate the last norms in (3.46) by using the boundedness of φ in H^1 if $(p-1)l' \leqq 2^*$ and Lemma 3.3 with $\rho = 0$ if $(p-1)l' \geqq 2^*$. Those norms are then estimated in terms of t and of the norms of φ_i in $L^\infty(I, H^1)$ provided there exists r such that $1 \leqq \delta(r) < 3/2$ and

$$2 \leqq (p-1)l' \leqq r, \tag{3.48}$$

$$[(p-1)m']^{-1} \geqq \delta((p-1)l') - 1. \tag{3.49}$$

The condition $(p-1)l' \geqq 2$ can be easily satisfied within (H 3), possibly by replacing p by some higher value, and (3.48) and (3.49) reduce to the conditions

$$p-1 \leqq 2(1+\gamma')[n-2\delta(r)]^{-1} \tag{3.50}$$

and

$$1/m' \geqq (p-1)(n/2-1) - (1+\gamma'). \tag{3.51}$$

The existence of m' satisfying (3.47) and (3.51) is obviously ensured by (H 3). The condition (3.50) also follows from (H 3) provided

$$(1+\gamma')(n-2) \geqq 2(n-2\delta(r))$$

a condition which is easily satisfied for γ' and $\delta(r)$ in the allowed range, for instance by $\gamma' = (n-1)/(n+1)$ and $\delta(r) = 3n[2(n+1)]^{-1}$. Taking in (3.46) the Supremum over t in a sufficiently small interval containing t_0, one obtains a linear inequality which implies that $\varphi_1 = \varphi_2$ in that interval. Iterating the process yields $\varphi_1 = \varphi_2$ everywhere in I. Q.E.D.

Under the assumptions of both Propositions 2.1 and 3.1, we obtain the existence and uniqueness of the solution of the Cauchy problem for the equation (1.1). In addition the solutions satisfy the conservation of the energy, even without the lower semi-continuity assumption made on V in the last statement of Proposition 2.1, and that conservation implies a stronger regularity in time than previously, namely strong continuity in the energy space.

Proposition 3.2. *Let f satisfy* (H 2) *and* (H 3), *let $t_0 \in \mathbb{R}$ and let $(\varphi_0, \psi_0) \in H^1 \oplus L^2$. Then the equation* (1.1) *has a unique solution in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1)$ with $\varphi(t_0) = \varphi_0$, $\dot\varphi(t_0) = \psi_0$. In addition, $\varphi \in \mathscr{C}(\mathbb{R}, H^1) \cap \mathscr{C}^1(\mathbb{R}, L^2)$, $\dot\varphi \in \mathscr{C}(\mathbb{R}, L^2) \cap \mathscr{C}^L(\mathbb{R}, H^{-1})$, and φ satisfies the estimates*

$$\|\varphi(t)\|_2 \leqq e(t - t_0) \tag{2.7}$$

$$\|\dot\varphi(t)\|_2^2 + \|\nabla \varphi(t)\|_2^2 \leqq \dot e(t - t_0)^2 \tag{3.52}$$

with $e(\cdot)$ given by (2.9), *and the conservation of the energy*

$$E(\varphi(t), \dot\varphi(t)) = E(\varphi_0, \psi_0) \tag{3.53}$$

for all $t \in \mathbb{R}$.

Proof. Propositions 2.1 and 3.1 assert the existence and uniqueness of the solution φ of the equation (1.1) in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1)$, the fact that

$$\varphi \in \mathscr{C}_w(\mathbb{R}, H^1) \cap \mathscr{C}^L(\mathbb{R}, L^2) \cap \bigcap_{2 < r < 2^*} \mathscr{C}^{\alpha(r)}(\mathbb{R}, L^r)$$

and

$$\dot\varphi \in L^\infty_{\mathrm{loc}}(\mathbb{R}, L^2) \cap \mathscr{C}_w(\mathbb{R}, L^2) \cap \mathscr{C}^L(\mathbb{R}, H^{-1}),$$

and the estimates (2.7) and (2.8), the second of which reduces to (3.52) in the present case. We postpone the proof of the conservation of the energy since it requires an independent approximation argument if, as is the case here, one does not assume V to satisfy the lower semi-continuity property contained in the last statement of Proposition 2.1, and we continue the proof of Proposition 3.2 under the assumption that energy conservation holds. Since $\varphi \in \bigcap_{2 \leqq r < 2^*} \mathscr{C}(\mathbb{R}, L^r)$, it follows from (3.53) that the function

$$\|\dot\varphi(t)\|_2^2 + \|\nabla \varphi(t)\|_2^2 = E(\varphi_0, \psi_0) - \int dx\, V(\varphi(x))$$

and therefore the norm of $(\varphi, \dot\varphi)$ in $H^1 \oplus L^2$, is a continuous function of time. The strong continuity of $(\varphi, \dot\varphi)$ in $H^1 \oplus L^2$ now follows from the weak continuity and the continuity of the norm.

We now come back to the proof of energy conservation. For that purpose, we approximate φ by smooth solutions of a regularized equation. We choose an even non-negative function $h_1 \in \mathscr{C}_0^\infty(\mathbb{R}^n)$ such that $\|h_1\|_1 = 1$. For any positive integer j, we define

$$h_j(x) = j^n h_1(jx),$$

$$f_j(\varphi) = h_j * f(h_j * \varphi)$$

and correspondingly

$$E_j(\varphi, \psi) = \|\psi\|_2^2 + \|\nabla \varphi\|_2^2 + \int dx\, V(h_j * \varphi), \tag{3.54}$$

where $*$ denotes convolution in $\mathbb{R}^n$. We consider the regularized equation

$$\varphi(t)=\dot{K}(t-t_0)(h_j*\varphi_0)+K(t-t_0)(h_j*\psi_0)-\int_{t_0}^{t}d\tau\,K(t-\tau)\,f_j(\varphi(\tau)). \tag{3.55}$$

For all $j\in\mathbb{Z}^+$, the equation (3.55) has a unique solution $\varphi_j\in\mathscr{C}(\mathbb{R},H^1)$ [20]. Furthermore, $(\varphi_j,\dot{\varphi}_j)\in\mathscr{C}^1(\mathbb{R},H^{k+1}\oplus H^k)$ for any positive integer k, and φ_j satisfies the differential equation

$$\ddot{\varphi}_j-\Delta\varphi_j+f_j(\varphi_j)=0 \tag{3.56}$$

in H^{k-1}. In addition φ_j satisfies the conservation of the energy

$$E_j(\varphi_j(t),\dot{\varphi}_j(t))=E_j(h_j*\varphi_0,h_j*\psi_0) \tag{3.57}$$

and the estimates

$$\|\varphi_j(t)\|_2\leqq e_j(t-t_0)\leqq\bar{e}(t-t_0) \tag{3.58}$$

$$\|\dot{\varphi}_j(t)\|_2^2+\|\nabla\varphi_j(t)\|_2^2\leqq\dot{e}_j(t-t_0)^2\leqq\dot{\bar{e}}(t-t_0)^2, \tag{3.59}$$

where the functions e_j and $\bar{e}$ are obtained from e, as defined in (2.9), by replacing $E(\varphi_0,\psi_0)$ respectively by $E_j(h_j*\varphi_0,h_j*\psi_0)$ and by

$$\bar{E}\equiv\operatorname*{Sup}_j E_j(h_j*\varphi_0,h_j*\psi_0)<\infty.$$

Furthermore, by Lemma 3.3, $\varphi_j\in L^q_{\mathrm{loc}}(\mathbb{R},\dot{H}^\rho_r)$ for ρ, r satisfying (3.7) and $1/q=\rho+\delta(r)-1$ and, for any compact interval I, φ_j is estimated in $L^q(I,\dot{H}^\rho_r)$ uniformly with respect to j. The last two facts follow from the uniform estimates (3.58) and (3.59) and from the fact that all the estimates in the proof of Lemmas 3.2 and 3.3 still hold with f replaced by f_j, since convolution with h_j is a contraction in $\dot{B}^\rho_{r,2}$ for all ρ and r.

We next prove that φ_j tends to φ when $j\to\infty$ in $\mathscr{C}(I,L^r)$ for any r, $2\leqq r<2^*$ and for any compact interval I (that result is stronger than needed for the proof of energy conservation, for which pointwise strong convergence in L^r, $2\leqq r<2^*$, would be sufficient). We first prove convergence of φ_j to φ in $\mathscr{C}(I,L^{r'})$ for r' chosen as in the proof of Proposition 3.1 (for instance with $\gamma(r')=(n-1)/(n+1)$). The proof is essentially a repetition of that of Proposition 3.1 and will only be sketched briefly. Comparing the equation (3.19) with the equation (3.55) for φ_j, we estimate

$$\begin{aligned}\|\varphi_j(t)-\varphi(t)\|_{r'}&\leqq C\,|t-t_0|^{1-\delta(r')}\,\|h_j*\psi_0-\psi_0\|_2\\&+C\,\|\omega^{\delta(r')}(h_j*\varphi_0-\varphi_0)\|_2+\int_{t_0}^{t}d\tau\,\|K(t-\tau)(h_j*f(\varphi(\tau))-f(\varphi(\tau))\|_{r'}\\&+\int_{t_0}^{t}d\tau\,M_j(t,\tau)(\|h_j*\varphi(\tau)-\varphi(\tau)\|_{r'}+\|\varphi_j(\tau)-\varphi(\tau)\|_{r'}),\end{aligned} \tag{3.60}$$

where

$$M_j(t,\tau)=C\{|t-\tau|+|t-\tau|^{-\gamma(r')}(\|\varphi(\tau)\|^{p-1}_{(p-1)l'}+\|\varphi_j(\tau)\|^{p-1}_{(p-1)l'})\} \tag{3.61}$$

and $n/l'=1+\gamma(r')$. Here we have estimated the contribution of the free term by (A.13), we have used the identity

$$h_j * f(h_j * \varphi_j) - f(\varphi) = h_j * f(\varphi) - f(\varphi) + h_j * \left\{ \int_0^1 d\alpha f'(\alpha h_j * \varphi_j + (1-\alpha)\varphi)(h_j * \varphi - \varphi + h_j * (\varphi_j - \varphi)) \right\} \tag{3.62}$$

and the fact that convolution with h_j is a contraction in L^r for all r, we have separated again $f = f_1 + f_2$ and estimated the contributions of the two terms as in the proof of Proposition 3.1. We now remark that, since $\varphi \in \mathscr{C}(\mathbb{R}, L^{r'})$, the sequence of functions $\|h_j * \varphi - \varphi\|_{r'}$ is equicontinuous in time and therefore tends to zero uniformly in compact intervals. On the other hand, by Lemma 3.3 and its analogue for φ_j and by the same arguments as in the proof of Proposition 3.1, the functions $M_j(t,\tau)$ are locally integrable in τ and the integral $\int_{t_0}^{t} d\tau M_j(t,\tau)$ is estimated uniformly in j by a quantity which tends to zero when t tends to t_0. The norm

$$\|K(t-\tau)(h_j * f(\varphi(\tau)) - f(\varphi(\tau)))\|_{r'}$$

is bounded by $2\|K(t-\tau) f(\varphi(\tau))\|_{r'}$, which can be shown to be integrable in τ by similar arguments, and to tend to zero pointwise almost everywhere in τ. It follows then from the Lebesgue theorem that the first integral in (3.60) tends to zero as $j \to \infty$ for fixed t, and by an additional elementary argument taking into account the singularity at $\tau = t$, that the convergence is actually uniform for t in compact intervals. Taking in (3.60) the Supremum over t in a sufficiently small interval containing t_0 and using the previous remarks, one obtains an inequality which implies that $\varphi_j(t)$ tends to $\varphi(t)$ in L^r uniformly for t in that interval. Iterating the process yields convergence of φ_j to φ in $\mathscr{C}(I, L^r)$ for any compact interval I. By interpolation between the convergence of φ_j to φ in $\mathscr{C}(I, L^r)$ and the uniform boundedness of φ_j in $L^\infty_{\mathrm{loc}}(\mathbb{R}, H^1)$ expressed by (3.58) (3.59), we obtain the convergence of φ_j to φ in $\mathscr{C}(I, L^r)$ for any r, $2 < r < 2^*$.

We next prove that φ_j tends to φ in $\mathscr{C}(I, L^2)$ for any compact I, by a similar argument using the convergence in $\mathscr{C}(I, L^r)$ for $2 < r < 2^*$. Comparing again (3.19) and (3.55) for φ_j, we obtain in the same way as above

$$\begin{aligned} \|\varphi_j(t) - \varphi(t)\|_2 \leqq{} & |t - t_0| \, \|h_j * \psi_0 - \psi_0\|_2 + \|h_j * \varphi_0 - \varphi_0\|_2 \\ & + \int_{t_0}^{t} d\tau \{ \|K(t-\tau)(h_j * f(\varphi(\tau)) - f(\varphi(\tau)))\|_2 \\ & + C|t-\tau|(\|\varphi_j(\tau) - \varphi(\tau)\|_2 + \|h_j * \varphi(\tau) - \varphi(\tau)\|_2) \\ & + C(\|\varphi(\tau)\|^{p-1}_{(p-1)l} + \|\varphi_j(\tau)\|^{p-1}_{(p-1)l}) \\ & \cdot (\|h_j * \varphi(\tau) - \varphi(\tau)\|_r + \|\varphi_j(\tau) - \varphi(\tau)\|_r) \} \end{aligned} \tag{3.63}$$

with $1/l + 1/r = 1/2 + 1/n$. That estimate differs from the special case of (3.60) with r' replaced by 2 by the fact that the contribution of f_2' from (3.62) is estimated in terms of the previously controlled norms of $\varphi_j - \varphi$ in L^r instead of the norms in L^2. In fact, one can choose $r < 2^*$ such that φ, $\varphi_j \in L^{p-1}_{\mathrm{loc}}(\mathbb{R}, L^{(p-1)l})$ according to Lemma 3.3, by choosing l such that

$$\mathrm{Max}\{(p-1)(n-3)/2, (p-1)(n/2-1)-1\} < n/l < 2,$$

a condition which is satisfied under (H 3) if

$$\mathrm{Max}\{2-2/(n-2),1\}\leqq n/l<2.$$

Using the previously proved convergence of φ_j to φ in $\mathscr{C}(I,L^r)$ for compact I, one obtains the convergence in $\mathscr{C}(I,L^2)$ by an argument similar to that used above for the proof of convergence in $\mathscr{C}(I,L^r)$.

From the uniform estimates (3.58) and (3.59) on φ_j and the convergence of φ_j to φ in $\mathscr{C}(I,L^r)$ for some r, $2\leqq r<2^*$, for instance $r=2$, several additional convergences of φ_j to φ can be derived by the same compactness arguments as used in the proof of Proposition 2.1 to prove the convergence of the finite dimensional approximants to the solution. In particular, for any compact interval I, φ_j tends to φ in the weak-* topology of $L^\infty(I,H^1)$ and $\dot\varphi_j$ tends to $\dot\varphi$ in the weak-* topology of $L^\infty(I,L^2)$ ([11], but here no extraction of subsequence is needed). In addition, for each t, uniform boundedness of $\varphi_j(t)$ in H^1 and convergence of $\varphi_j(t)$ to $\varphi(t)$ in L^r implies that $\varphi_j(t)$ converges to $\varphi(t)$ weakly in H^1. Finally we prove that $\dot\varphi_j(t)$ converges to $\dot\varphi(t)$ weakly in L^2 for each t. In fact from the equation (3.56) it follows that $\dot\varphi_j$ is Lipschitz continuous in H^{-1} uniformly in j. On the other hand by (3.59), $\dot\varphi_j(t)$ is uniformly bounded and therefore weakly relatively compact in L^2, so that it is sufficient to prove that $\dot\varphi_j(t)$ can have no other weak accumulation point than $\dot\varphi(t)$ in L^2. Assume now that a subsequence, still called $\dot\varphi_j(t)$ for simplicity, converges weakly to χ in L^2. Then, for any $\theta>0$ and any $u\in H^1$,

$$\begin{aligned}\langle u,\dot\varphi(t)-\chi\rangle=(2\theta)^{-1}\int_{t-\theta}^{t+\theta}d\tau\langle u,(\dot\varphi(t)-\dot\varphi(\tau))+(\dot\varphi(\tau)-\dot\varphi_j(\tau))\\+(\dot\varphi_j(\tau)-\dot\varphi_j(t))+(\dot\varphi_j(t)-\chi)\rangle.\end{aligned}\tag{3.64}$$

The contributions of the first and third differences in the right hand side of (3.64) tend to zero with θ uniformly in j by the uniform Lipschitz continuity of $\dot\varphi$ and $\dot\varphi_j$ in H^{-1}, the contribution of the second difference tends to zero for fixed θ by the weak-* convergence of $\dot\varphi_j$ to $\dot\varphi$ in $L^\infty(I,L^2)$ and the contribution of the last difference tends to zero by the assumed weak convergence of $\dot\varphi_j(t)$ to χ in L^2. Therefore $\langle u,\dot\varphi(t)-\chi\rangle=0$ for any $u\in H^1$ so that $\chi=\dot\varphi(t)$. This proves that $\dot\varphi_j(t)$ tends to $\dot\varphi(t)$ weakly in L^2 for each t.

We can now complete the proof of energy conservation. We take the limit $j\to\infty$ in (3.57). The right hand side tends to $E(\varphi_0,\psi_0)$ since convolution with h_j tends to the identity strongly in L^r for $1\leqq r<\infty$ and since the map $\varphi\mapsto V(\varphi)$ is continuous from $L^2\cap L^{p+1}$ to L^1. In the left hand side, the term with V also converges to the obvious value, since in addition $\varphi_j(t)$ tends to $\varphi(t)$ strongly in $L^2\cap L^{p+1}$ as just proved. Therefore the following limit exists:

$$\lim_{j\to\infty}(\|\dot\varphi_j(t)\|_2^2+\|\nabla\varphi_j(t)\|_2^2)=E(\varphi_0,\psi_0)-\int dx\,V(\varphi(t)).\tag{3.65}$$

Together with weak convergence of $\dot\varphi_j(t)$ to $\dot\varphi(t)$ in L^2 and of $\varphi_j(t)$ to $\varphi(t)$ in H^1, (3.65) implies that

$$E(\varphi(t),\dot\varphi(t))\leqq E(\varphi_0,\psi_0)\tag{3.66}$$

for all t. By the uniqueness of the solution φ, we can apply time reversal to the equation (1.1) and therefore to the energy inequality (3.66), thereby completing the proof of the energy conservation (3.53). Q.E.D.

Acknowledgments. We are grateful to Professor L. Streit for the warm hospitality extended to us at the Zentrum für Interdisziplinäre Forschung in Bielefeld, where part of this work was done.

Appendix

In this appendix, we collect the definition of and some basic facts about the homogeneous Besov and Sobolev spaces of arbitrary order. The reader is referred to [27], p. 237–244, and to [1], p. 146–154 for a general framework and a sketch of the main results.

We denote by $\mathscr{F}$ the Fourier transform, and we write $\hat{v}=\mathscr{F}v$ for any $v\in\mathscr{S}'\equiv\mathscr{S}'(\mathbb{R}^n)$. We define the subspaces $\mathscr{Z}\equiv\mathscr{Z}(\mathbb{R}^n)$ and $\hat{\mathscr{D}}_0$ of $\mathscr{S}\equiv\mathscr{S}(\mathbb{R}^n)$ by

$$\mathscr{Z}=\{u\in\mathscr{S}: D^\alpha\hat{u}(0)=0 \text{ for any multindex } \alpha\} \tag{A.1}$$

and

$$\hat{\mathscr{D}}_0=\{u\in\mathscr{S}: \hat{u}\in\mathscr{C}_0^\infty(\mathbb{R}^n\setminus\{0\})\}. \tag{A.2}$$

Clearly $\mathscr{Z}$ is a closed subspace of $\mathscr{S}$, while $\hat{\mathscr{D}}_0$ is easily seen to be dense in $\mathscr{Z}$. For any $\sigma\in\mathbb{R}$, we define the operator ω^σ by $\widehat{\omega^\sigma u}(\xi)=|\xi|^\sigma\hat{u}(\xi)$. Then ω^σ is an homeomorphism of $\mathscr{Z}$ onto itself which leaves $\hat{\mathscr{D}}_0$ invariant and satisfies the group property $\omega^{\sigma_1}\omega^{\sigma_2}=\omega^{\sigma_1+\sigma_2}$. The inclusion $\mathscr{Z}\subset\mathscr{S}$ induces a map π: $\mathscr{S}'\mapsto\mathscr{Z}'$ of the dual spaces, defined by $\langle\pi v,u\rangle=\langle v,u\rangle$ for any $v\in\mathscr{S}'$ and $u\in\mathscr{Z}$. The nullspace of π consists of those $v\in\mathscr{S}'$ with $\operatorname{Supp}\hat{v}=\{0\}$, namely of the polynomials, and π is surjective by the Hahn-Banach theorem, so that $\mathscr{Z}'=\mathscr{S}'/\mathscr{P}$. By duality, the map ω^σ induces an homeomorphism of $\mathscr{Z}'$ onto itself, still denoted by ω^σ.

For any p, $1\leqq p\leqq\infty$, the space L^p is continuously embedded in $\mathscr{S}'$, so that $\pi(L^p)$ is continuously embedded in $\mathscr{Z}'$. For $1\leqq p<\infty$, any $v\in\pi(L^p)$ is the image of only one element of L^p, since L^p contains no polynomial, so that L^p is canonically embedded in $\mathscr{Z}'$. For any $\sigma\in\mathbb{R}$ and any p, $1\leqq p<\infty$, we define the homogeneous Sobolev space $\dot{H}_p^\sigma$ by

$$\dot{H}_p^\sigma=\omega^{-\sigma}L^p. \tag{A.3}$$

The space $\dot{H}_p^\sigma$ is a Banach space with the norm

$$\|v;\dot{H}_p^\sigma\|=\|\omega^\sigma v\|_p, \tag{A.4}$$

and for any σ_1, $\sigma_2\in\mathbb{R}$, ω^σ is an isometry of $\dot{H}_p^{\sigma_1}$ onto $\dot{H}_p^{\sigma_2}$ for $\sigma=\sigma_1-\sigma_2$. For $1<p<\infty$, an easy duality argument shows that $\hat{\mathscr{D}}_0$ and $\mathscr{Z}$ are dense in L^p and therefore in $\dot{H}_p^\sigma$.

One can give an alternative definition of the spaces $\dot{H}_p^\sigma$ by the use of a dyadic decomposition. Let $\hat{\psi}\in\mathscr{C}_0^\infty(\mathbb{R}^n)$ with $0\leqq\hat{\psi}\leqq1$, $\hat{\psi}(\xi)=1$ for $|\xi|\leqq1$ and $\hat{\psi}(\xi)=0$ for $|\xi|\geqq2$, and define for any $j\in\mathbb{Z}$

$$\hat{\varphi}_j(\xi)=\hat{\psi}(2^{-j}\xi)-\hat{\psi}(2^{-(j+1)}\xi)$$

so that

$$\mathrm{Supp}\,\hat{\varphi}_j \subset \{\xi : 2^{j-1} \leqq |\xi| \leqq 2^{j+1}\}$$

and for any $\xi \neq 0$, $\sum_j \hat{\varphi}_j(\xi) = 1$, with at most two nonvanishing terms in the sum. With any $v \in \mathscr{S}'$ one can associate the sequence of $\mathscr{C}^\infty$ functions $\Phi(v) = \{v_j \equiv \varphi_j * v\}$, $j \in \mathbb{Z}$. Clearly $\Phi(v)$ depends only on πv so that Φ is actually defined on $\mathscr{Z}'$. For any $\sigma \in \mathbb{R}$ and any p, q with $1 \leqq p, q \leqq \infty$, one defines the spaces $\dot{F}^\sigma_{pq}$ and $\dot{B}^\sigma_{pq}$ by

$$\dot{F}^\sigma_{pq} = \{v \in \mathscr{Z}' : \| \{\sum_j |2^{j\sigma}(\varphi_j * v)|^q\}^{1/q} \|_p \equiv \|v; \dot{F}^\sigma_{p,q}\| < \infty\} \tag{A.5}$$

and

$$\dot{B}^\sigma_{pq} = \{v \in \mathscr{Z}' : \{\sum_j 2^{j\sigma q} \|\varphi_j * v\|_p^q\}^{1/q} \equiv \|v; \dot{B}^\sigma_{p,q}\| < \infty\} \tag{A.6}$$

with obvious modifications if $q = \infty$. One can then prove that the spaces $\dot{F}^\sigma_{pq}$ and $\dot{B}^\sigma_{pq}$ are Banach spaces when equipped with the norms given in the definitions (A.5), (A.6), and that for $1 < p, q < \infty$, $\hat{\mathscr{D}}_0$ and $\mathscr{Z}$ are dense in $\dot{F}^\sigma_{pq}$ and $\dot{B}^\sigma_{pq}$. The spaces $\dot{B}^\sigma_{pq}$ are the homogeneous Besov spaces, and it follows immediately from the Minkowski inequality that for any p, q with $1 \leqq p \leqq q \leqq \infty$ (resp. $1 \leqq q \leqq p \leqq \infty$)

$$\dot{F}^\sigma_{pq} \subset \dot{B}^\sigma_{pq} \quad (\text{resp. } \dot{B}^\sigma_{pq} \subset \dot{F}^\sigma_{pq}). \tag{A.7}$$

Using the vector valued version [26] of the Mikhlin-Hörmander theorem ([9], Theorem 7.9.5, p. 243) one can prove that for any $\sigma \in \mathbb{R}$ and any p, $1 < p < \infty$, $\dot{F}^\sigma_{p2} = \dot{H}^\sigma_p$ with equivalent norms. Moreover, the norms given in (A.5) and (A.4) are uniformly equivalent for σ in a bounded subset of $\mathbb{R}$ and for p in a compact subset of $(1, \infty)$. From that identification and from (A.7), one obtains immediately the embeddings

$$\dot{B}^\sigma_{p2} \subset \dot{H}^\sigma_p \subset \dot{B}^\sigma_{pp} \quad , \quad \dot{B}^\sigma_{p\bar{p}} \subset \dot{H}^\sigma_{\bar{p}} \subset \dot{B}^\sigma_{\bar{p}2} \tag{A.8}$$

for $2 \leqq p < \infty$. On the other hand, from the same identification and from the Hölder inequality, one obtains useful convexity properties which we state as follows.

Lemma A.1. *Let $0 \leqq \mu_i \leqq 1$, $1 \leqq i \leqq N$, with $\sum_i \mu_i = 1$. Let $\sigma_i \in \mathbb{R}$ and $\sigma = \sum_i \mu_i \sigma_i$. Let $1 \leqq p_i \leqq \infty$ and $1/p = \sum_i \mu_i / p_i$.*

(1) *Let in addition $1 \leqq q_i \leqq \infty$ and $1/q = \sum_i \mu_i / q_i$. Then $\bigcap_i \dot{B}^{\sigma_i}_{p_i,q_i} \subset \dot{B}^\sigma_{p,q}$ and for any $v \in \bigcap_i \dot{B}^{\sigma_i}_{p_i,q_i}$*

$$\|v; \dot{B}^\sigma_{p,q}\| \leqq \prod_{1 \leqq i \leqq N} \|v; \dot{B}^{\sigma_i}_{p_i,q_i}\|^{\mu_i}. \tag{A.9}$$

The same property holds for the spaces $\dot{F}$.

(2) *Let in addition $1 < p_i < \infty$. Then $\bigcap_i \dot{H}^{\sigma_i}_{p_i} \subset \dot{H}^\sigma_p$ and for any $v \in \bigcap_i \dot{H}^{\sigma_i}_{p_i}$,*

$$\|v; \dot{H}^\sigma_p\| \leqq C \prod_{1 \leqq i \leqq N} \|v; \dot{H}^{\sigma_i}_{p_i}\|^{\mu_i}. \tag{A.10}$$

The constant C is uniform in σ_i, p_i, μ_i for σ_i in a bounded subset of $\mathbb{R}$ and p_i in a compact subset of $(1, \infty)$.

The Sobolev inequalities in the (homogeneous) Besov spaces are a consequence of the Hölder inequality. In the Sobolev spaces, they follow from the fact that in a suitable range of σ, namely for $0<\sigma<n$, the operator $\omega^{-\sigma}$ acting on sufficiently regular v's can be represented by the convolution with the Riesz potential $I_\sigma = \mathscr{F}^{-1}|\cdot|^{-\sigma}$, namely $I_\sigma(x) = C_\sigma |x|^{-n+\sigma}$. The Hardy-Littlewood-Sobolev inequality ([9], Theorem 4.5.3 p. 117) can be stated as

$$\|I_\sigma * v\|_r \leqq C \|v\|_p \tag{A.11}$$

with $1<p<r<\infty$ and $\sigma/n = 1/p - 1/r$. One then obtains the following results.

Lemma A.2. *Let $1 \leqq p \leqq r \leqq \infty$, let $\rho, \sigma \in \mathbb{R}$ and $1/p - 1/r = (\sigma - \rho)/n$.*

(1) *Let in addition $1 \leqq q \leqq \infty$. Then $\dot{B}^\sigma_{p,q} \subset \dot{B}^\rho_{r,q}$ and for any $v \in \dot{B}^\sigma_{p,q}$,*

$$\|v; \dot{B}^\rho_{r,q}\| \leqq C \|v; \dot{B}^\sigma_{p,q}\| \tag{A.12}$$

for some constant C independent of p, r, q, ρ and σ.

(2) *Let in addition $1<p \leqq r<\infty$. Then $\dot{H}^\sigma_p \subset \dot{H}^\rho_r$ and for any $v \in \dot{H}^\sigma_p$,*

$$\|v; \dot{H}^\rho_r\| \leqq C \|v; \dot{H}^\sigma_p\|. \tag{A.13}$$

The constant C in (A.13) *is uniform in p, r, σ and ρ for p, r in a compact subset of $(1, \infty)$.*

In the applications, we use in general Lemmas A.1 and A.2 put together in an obvious way. The inequalities (A.9) and (A.12) are understood to hold in appropriate subspaces of $\mathscr{Z}'$, and are used eventually to obtain estimates in spaces L^l, considered as subspaces of $\mathscr{Z}'$, of equivalence classes modulo polynomials of certain functions v. On the other hand, in the applications, those functions are known from other sources to belong to L^s for some s, $1<s<\infty$ (in general $s=2$), so that the estimates in L^l hold actually in the ordinary sense for the functions v themselves.

References

1. Bergh, J., Löfström, J.: Interpolation spaces. Berlin-Heidelberg-New York: Springer 1976
2. Brenner, P.: On $L_p - L_{p'}$ estimates for the wave equation. Math. Z. **145**, 251–254 (1975)
3. Brenner, P.: On the existence of global smooth solutions of certain semilinear hyperbolic equations. Math. Z. **167**, 99–135 (1979)
4. Brenner, P., von Wahl, W.: Global classical solutions of non-linear wave equations. Math. Z. **176**, 87–121 (1981)
5. Browder, F.E.: On nonlinear wave equations. Math. Z. **80**, 249–264 (1962)
6. Ginibre, J., Velo, G.: The global Cauchy problem for the nonlinear Schrödinger equation revisited. Ann. Inst. Henri Poincaré (in press)
7. Glassey, R., Tsutsumi, M.: On uniqueness of weak solutions to semi linear wave equations. Commun. Partial Differ. Equations **7**, 153–195 (1982)
8. Heinz, E., von Wahl, W.: Zu einem Satz von F.E. Browder über nichtlineare Wellengleichungen. Math. Z. **141**, 33–45 (1975)

9. Hörmander, L.: The analysis of linear partial differential operators I. Berlin-Heidelberg-New York: Springer 1983
10. Jörgens, K.: Das Anfangswertproblem im Grossen für eine Klasse nichtlinearer Wellengleichungen. Math. Z. **77**, 295–308 (1961)
11. Lions, J.L.: Quelques méthodes de résolution des problèmes aux limites non linéaires. Paris: Dunod and Gauthier-Villars 1969
12. Parenti, C., Strocchi, F., Velo, G.: A local approach to some nonlinear evolution equations of hyperbolic type. Ann. Sc. Norm. Super. Pisa, Cl. Sci., IV. Ser. **3**, 443–500 (1976)
13. Pecher, H.: L^p-Abschätzungen und klassische Lösungen für nichtlineare Wellengleichungen I. Math. Z. **150**, 159–183 (1976)
14. Pecher, H.: Ein nichtlinearer Interpolationssatz und seine Anwendung auf nichtlineare Wellengleichungen. Math. Z. **161**, 9–40 (1978)
15. Pecher, H.: Nonlinear small data scattering for the wave and Klein-Gordon equation. Math. Z. **185**, 261–270 (1984)
16. Pecher, H.: Low energy scattering for nonlinear Klein-Gordon equations. Preprint
17. Peral, J.C.: L^p-estimates for the wave equation. J. Funct. Anal. **36**, 114–145 (1980)
18. Reed, M.: Abstract nonlinear wave equations. Berlin-Heidelberg-New York: Springer 1976
19. Segal, I.E.: The global Cauchy Problem for a relativistic scalar field with power interaction. Bull. Soc. Math. Fr. **91**, 129–135 (1963)
20. Segal, I.E.: Nonlinear semigroups. Ann. Math. **78**, 339–364 (1963)
21. Segal, I.E.: Space-time decay for solutions of wave equations. Adv. Math. **22**, 305–311 (1976)
22. Strauss, W.: Decay and asymptotics for $\Box u = f(u)$. J. Funct. Anal. **2**, 409–457 (1968)
23. Strauss, W.: On weak solutions of semilinear hyperbolic equations. An. Acad. Bras. Cienc. **42**, 645–651 (1970)
24. Strauss, W.: Nonlinear scattering theory at low energy. J. Funct. Anal. **41**, 110–133 (1981). Ibid. **43**, 281–293 (1981)
25. Strichartz, R.: Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations. Duke Math. J. **44**, 705–714 (1977)
26. Triebel, H.: Spaces of distributions of Besov type on Euclidean n-space. Duality, interpolation. Ark. Mat. **11**, 13–64 (1973)
27. Triebel, H.: Theory of function spaces. Basel: Birkhäuser 1983
28. von Wahl, W.: Klassische Lösungen nichtlinearer Wellengleichungen im Grossen. Math. Z. **112**, 241–279 (1969)
29. von Wahl, W.: Über nichtlineare Wellengleichungen mit zeitabhängigem elliptischem Hauptteil. Math. Z. **142**, 105–120 (1975)

Received September 10, 1984; in final form December 11, 1984

Math. Z. 189, 507–513 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Finiteness Results for Algebraic K3 Surfaces

Hans Sterk

Mathematisch Instituut, Katholieke Universiteit, Toernooiveld, NL-6525 ED Nijmegen, The Netherlands

0. Introduction

The aim of this paper is to prove the following theorem:

0.1 Theorem. *Let* X *be an algebraic* K3 *surface over the complex numbers. Then*

a) *the group* Aut(X) *of (biholomorphic) automorphisms is finitely generated*

b) *for every even integer* $d \geqq -2$, *the number of* Aut(X)*-orbits in the collection of complete linear systems which contain an irreducible curve of selfintersection* d *is finite.*

In [5] A. Bremner states a conjecture which is related to the following problem: Is the number of Aut(X)-orbits in the set of nodal classes finite? Here X is a (algebraic) K3 surface over the complex numbers. Bremner's own version is more specific than this problem which was formulated by W. Barth. Theorem (0.1) gives an answer to a somewhat generalized form of the problem. The proof of (0.1 b) we present here was outlined by E. Looyenga, to whom the author wishes to express his thanks. The result in (0.1 a) is implicit in [7] § 7 theorem 1 up to a quotation from [4]. We include it here for convenience of the reader.

The cup-product endows the second cohomology group $H^2(X, \mathbb{Z})$ of a K3 surface over the complex numbers with the structure of an even unimodular lattice of signature (3, 19).

An automorphism of X induces an automorphism of $H^2(X; \mathbb{Z})$ which preserves the lattice structure, i.e. an isometry; it is known that this representation of Aut(X) on $H^2(X; \mathbb{Z})$ is faithful. If X is algebraic, its Néron-Séveri group NS is a nondegenerate primitive sublattice of $H^2(X; \mathbb{Z})$ of signature $(1, \rho - 1)$ where $\rho = \text{rank}(NS)$. Each element of NS is represented by a unique divisor class. The set $\{x \in NS \otimes_{\mathbb{Z}} \mathbb{R} : \langle x, x \rangle > 0\}$ consists of two disjoint connected components, only one of which contains ample classes: the positive cone. Using the Torelli-theorem for (algebraic) K3 surfaces we can relate Aut(X) to certain arithmetic subgroups of the orthogonal group of $NS \otimes_{\mathbb{Z}} \mathbb{R}$ acting on the

positive cone. This enables one to apply results from reduction theory for arithmetic groups. Our main reference for this last part will be the work of A. Ash in [1]. In fact, implicit in Ash's work is a method for constructing polyhedral (almost strict) fundamental domains for such group actions. These fundamental domains satisfy a certain finiteness property: the so-called Siegel-property. This allows one to obtain the desired finiteness results.

1. Preliminaries

In this paragraph we briefly review the most important notions for the proof of theorem (0.1) leaving out most of the details and proofs. We also fix most of the notation here. A good reference for more details is [2].

1.1 K3 *Surfaces*

Let X be an algebraic K3 surface over the complex numbers (i.e. X is a smooth connected algebraic surface with irregularity $q=0$ and which has trivial canonical class) and let $\langle\ ,\ \rangle$ denote the scalar product on $H^2(X;\mathbb{Z})$. The short exact sequence of sheaves on X

$$0\to\mathbb{Z}\to\mathcal{O}\to\mathcal{O}^*\to 1$$

gives rise to a connecting homomorphism in cohomology: $H^1(X;\mathcal{O}^*)\to H^2(X;\mathbb{Z})$, which is injective because $H^1(X;\mathcal{O})$ vanishes.

We fix the following notations:

$NS:=$ image $(H^1(X,\mathcal{O}^*)\to H^2(X;\mathbb{Z}))$, the *Néron-Séveri group* of X. It is a primitive nondegenerate sublattice of $H^2(X;\mathbb{Z})$ of signature $(1,\rho-1)$ where $\rho=\operatorname{rank} NS$. By the Lefschetz-theorem on $(1,1)$-classes we have $NS=H^2(X;\mathbb{Z})\cap H^{1,1}_{\delta}(X)$. Here $H^{1,1}_{\delta}(X)$ denotes cohomology of type $(1,1)$ in the Hodge-decomposition of $H^2(X;\mathbb{C})$.

$T:=NS^{\perp}$ (in $H^2(X;\mathbb{Z})$), the *transcendental lattice* of X. It can alternatively be characterized as the smallest primitive sublattice of $H^2(X;\mathbb{Z})$ such that its complexification contains $H^{2,0}_{\delta}(X)\cong H^0(X;\Omega^2)$. Clearly $S\cap T=\{0\}$, and $S+T$ is of finite index in $H^2(X;\mathbb{Z})$.

$\Delta:=\{\delta\in NS:\langle\delta,\delta\rangle=-2\}$.

$\Delta^+:=\{\delta\in\Delta:\delta$ is effective$\}$.

A *nodal class* is the class of a smooth rational curve on X. By the genus formula such a curve has selfintersection -2.

$B:=\{b\in\Delta^+:b$ is a nodal class$\}$.

$\mathscr{C}(\subset NS\otimes_{\mathbb{Z}}\mathbb{R}):=$ connected component of $\{x\in NS\otimes_{\mathbb{Z}}\mathbb{R}:\langle x,x\rangle>0\}$ which contains ample classes. We call it the *positive cone.*

$\mathscr{C}_+:=$ convex hull of $\bar{\mathscr{C}}\cap NS\otimes_{\mathbb{Z}}\mathbb{Q}$.

$K(\subset\mathscr{C}):=\{x\in\mathscr{C}:\langle x,b\rangle>0$ for all $b\in B\}$. We call this set the *ample cone.* The intersection $K\cap NS$ consists precisely of the ample classes in NS.

$O(NS) := $ group of isometries of the lattice NS.

$O^{\uparrow}(NS) := \{g \in O(NS) : g(\mathscr{C}) = \mathscr{C}\}$.

We do not introduce a separate notation for the $\mathbb{R}$- (or $\mathbb{Q}$-)linear extension of an element of $O(NS)$.

Every $\delta \in \Delta$ induces an isometry s_δ of $H^2(X;\mathbb{Z})$ (and NS) namely $s_\delta : x \to x + \langle x, \delta\rangle \delta$. Its $\mathbb{R}$-linear extension is a reflection in the hyperplane

$$\{x \in H^2(X;\mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{R} \text{ (or } NS \otimes_{\mathbb{Z}} \mathbb{R}) : \langle x, \delta\rangle = 0\}.$$

We will write H_δ for the hyperplane

$$\{x \in NS \otimes_{\mathbb{Z}} \mathbb{R} : \langle x, \delta\rangle = 0\}.$$

If $b \in B$, then $H_b \cap \bar{K} \cap \mathscr{C}$ contains elements which are not contained in any other $H_c (c \in B, c \neq b)$. Namely if $a \in K$, then $a + \frac{1}{2}\langle a, b\rangle\, b$ is such an element. So none of the hyperplanes H_b $(b \in B)$ is superfluous in defining K.

$W (\lhd O^{\uparrow}(NS)) := $ group generated by $\{s_b\}_{b \in B}$. It can be shown that the pair $(W, \{s\}_{b \in B})$ is a Coxeter-system. The $\mathbb{R}$-linear extension of an element of W induces an isometry of $NS \otimes_{\mathbb{Z}} \mathbb{R}$ which preserves $\mathscr{C}$. The closure $\bar{K} \cap \mathscr{C}$ of K in $\mathscr{C}$ is a (strict) fundamental domain for the W-action on $\mathscr{C}$ in the sense that every W-orbit in $\mathscr{C}$ meets $\bar{K} \cap \mathscr{C}$ in precisely one point. We can even say a little bit more.

Let $D := \{x \in NS \otimes_{\mathbb{Z}} \mathbb{R} : \langle x, b\rangle > 0 \text{ for all } b \in B\}$. It is easily checked that the arguments in [3] Ch. V §4.4 and §4.5 can be used to prove the analogues of the statements (P_n) and (Q_n) ([3] p. 93), for W, $\{s_b : b \in B\}$ and D.

1.2 Lemma. *Every W-orbit meets $\bar{D}$ in at most one point.*

Proof. Let $x \in \bar{D}$ and suppose $w \in W$ is such that $w(x) \in \bar{D}$. Then we must show that $x = w(x)$. We proceed by induction on the length $l(w)$ of w ($l(w)$ is the smallest integer $q \geqq 0$ such that w can be written as $w = s_{b_1} \ldots s_{b_q}$ for elements $b_1, \ldots, b_q \in B$). If $l(w) = 0$, the result is clear. Now assume $l(w) = n+1$ and the result is proven for $w' \in W$ with $l(w') \leqq n$. Write $w = s_b w'$ with $b \in B$ and $l(w') = n$. Define $A_b := \{x \in NS \otimes_{\mathbb{Z}} \mathbb{R} : \langle s, b\rangle > 0\}$. Using (P_n) of [3] p. 93 we find $w(A_b) \subset s_b A_b$ and therefore $\langle w(x), b\rangle \leqq 0$. As $w(x) \in \bar{D}$ we also have $\langle w(x), b\rangle \geqq 0$, so we conclude $w(x) \in H_b$ and $w'(x) = s_b w(x) = w(x) \in \bar{D}$. As $l(w') = n$, the induction hypothesis implies $x = w'(x)$. Therefore $x = w'(x) = w(x)$.

The nondegenerate lattice structure on NS makes it possible to view $NS^* (= \mathrm{Hom}(NS, \mathbb{Z}))$ as a sublattice of $NS \otimes_{\mathbb{Z}} \mathbb{R}$ which contains NS as a sublattice of finite index.

$\Gamma := \{g \in O^{\uparrow}(NS) : g \text{ induces identity on } NS^*/NS\}$. It is a subgroup containing W and it is of finite index in $O^{\uparrow}(NS)$. The subgroup Γ consists precisely of those elements of $O^{\uparrow}(NS)$ which extend to $H^2(X;\mathbb{Z})$ inducing identity on T.

$\Gamma_B := \{g \in \Gamma : g(B) = B\}$. As elements of Γ_B preserve nodal classes they also preserve the ample cone. The Torelli-theorem for (algebraic) K3 surfaces implies that Γ_B can be identified with the subgroup of Γ consisting of isometries induced by automorphism of X acting as identity on $H^0(X, \Omega^2)$.

$O_e^{\uparrow}(NS) := \{g \in O^{\uparrow}(NS) : g \text{ preserves effective classes}\}$.

An argument as in the proof of [7] §6 theorem 1 yields the following semi-direct product decomposition $O^{\uparrow}(NS) = O_e^{\uparrow}(NS) \cdot W$.

In [8] §1 we find the semi-direct product decomposition $\Gamma = \Gamma_B \cdot W$.

1.3 Convex Cones, Polyhedral Cones

The material presented here is based on Chap. II of [1]. Let V be a finite-dimensional $\mathbb{R}$-vectorspace with $\mathbb{Q}$-structure $V_{\mathbb{Q}}$ and let C be an open convex non-degenerate cone in V.

$C^* := \{\xi \in V^* (= \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})) : \xi_{|C} \geqq 0\}$. The interior $\mathring{C}^*$ is then again a strictly convex open cone, called the *dual cone.*

$C_+ :=$ convex hull of $\bar{C} \cap V_{\mathbb{Q}}$

Define for any subset $A \subset V$:
semi-hull $(A) :=$ closed convex hull of $\{\lambda a : \lambda \in \mathbb{R},\ \lambda \geqq 1 \text{ and } a \in A\}$.

A *polyhedral cone* $\Pi \in V$ is a closed subset of V which is the intersection of a finite number of closed halfspaces

$$\Pi = \{x \in V : \xi_1(x) \geqq 0, \ldots, \xi_l(x) \geqq 0 \quad \text{for } \xi_1, \ldots, \xi_l \in V^*\}.$$

Π is called *rational* if $\xi_1, \ldots, \xi_l$ can be chosen in $V_{\mathbb{Q}}^*$. We call a subset B of C_+ *rationally locally polyhedral* if for any rational polyhedral cone $\Pi \subset C_+$, the intersection $\Pi \cap B$ is described within Π by a finite number of inequalities

$$\xi_1 \geqq 1, \ldots, \xi_l \geqq 1 \quad \text{where } \xi_1, \ldots, \xi_l \in V_{\mathbb{Q}}^*.$$

Note that the positive cone introduced in 1.1 is a homogeneous selfdual cone, and that $O^{\uparrow}(NS)$ is an arithmetic group. This means that the results in Ch. II §§ 4, 5 of [1] can be applied in our case.

2. Finiteness Results

Let X be an algebraic K3 surface (over the complex numbers), and let ω be a nowhere vanishing holomorphic 2-form on X. Define

$\mathrm{Aut}^\circ(X) = \{h \in \mathrm{Aut}(X) : h^* \omega = \omega\}$;
then $\mathrm{Aut}^\circ(X)$ is clearly a normal subgroup of $\mathrm{Aut}(X)$.

2.1 Lemma. *The quotient* $\mathrm{Aut}(X)/\mathrm{Aut}^\circ(X)$ *is a finite cyclic group.*

Proof. The signature of the bilinear form restricted to the realification $T \otimes_{\mathbb{Z}} \mathbb{R}$ is $(2, 20-\rho)$. A positive definite 2-plane is $(H_{\delta}^{2,0} \oplus H_{\delta}^{0,2})_{\mathbb{R}}$ (the real elements in $H_{\delta}^{2,0} \oplus H_{\delta}^{0,2}$). Together with its negative definite orthogonal complement this yields an orthogonal decomposition of $T \otimes_{\mathbb{Z}} \mathbb{R}$. The special choice of the decomposition implies that every $\sigma \in \mathrm{Aut}(T \otimes_{\mathbb{Z}} \mathbb{R})$ which preserves $H_{\delta}^{2,0}$ also preserves this decomposition. So all the eigenvalues of σ have absolute value 1. If, moreover, $\sigma \in \mathrm{Aut}(T)$, then these eigenvalues are algebraic integers and we conclude that they are in fact roots of unity ([6] proof of theorem 38). Now let $f \in \mathrm{Aut}(X)$, then f induces $f^* : T \to T$ and $f^* : H_{\delta}^{2,0} \to H_{\delta}^{2,0}$. The above statements apply and we find that $f^* : H_{\delta}^{2,0} \to H_{\delta}^{2,0}$ is multiplication by some root of unity. Its minimumpolynomial is a cyclotomic polynomial Φ_n (for some n) of degree $\phi(n)$. As Φ_n necessarily divides the characteristic polynomial of $f^* : T \to T$ we

must have: $\phi(n) \leqq \operatorname{rank} T$. The set $\{n \in \mathbb{N}: \phi(n) \leqq \operatorname{rank} T\}$ is finite, so we conclude that the representation of $\operatorname{Aut}(X)$ on $H^{2,0}_{\delta}$ assumes values in a finite group of roots of unity (after identifying $H^{2,0}_{\delta}$ with $\mathbb{C}$). But then $\operatorname{Aut}(X)/\operatorname{Aut}^{\circ}(X)$ is identified with a finite subgroup of $\mathbb{C}^*$ and must therefore be finite.

Remark. This result is actually implicit in [7] § 7 theorem 1 and can also be derived easily using the two semi-direct products mentioned in (1.1).

2.2 Proposition. $\operatorname{Aut}(X)$ *is a finitely generated group.*

Proof. By (2.1) it suffices to prove that $\operatorname{Aut}^{\circ}(X)$ is finitely generated. As $\operatorname{Aut}^{\circ}(X) \cong \Gamma_B$ we work with the latter group. Now Γ is an arithmetic subgroup of $O(NS \otimes_{\mathbb{Z}} \mathbb{R})(\cong O(1, \rho - 1))$, so it follows from [4] theorem 6.12 that Γ is finitely generated. As Γ_B is a quotient of Γ, we conclude that Γ_B is a finitely generated group.

Next we seek a rational polyhedral fundamental domain for the Γ-action on $\mathscr{C}_+$, which will also serve as a fundamental domain for the Γ_B-action on $\bar{K} \cap \mathscr{C}_+$.

Choose $\xi \in \mathring{\mathscr{C}}^* \cap NS^*$, then ξ is an extreme point of $A := \text{semi-hull}(\Gamma \cdot \xi + \mathring{\mathscr{C}}^*)$ (for at least one point of $\Gamma \cdot \xi$ is an extreme point and A is Γ-invariant.). According to [1] prop. 11 p. 142, A is locally rationally polyhedral, so we can choose a rational polyhedral cone $\Pi' \subset \mathscr{C}^*_+ (= (\mathring{\mathscr{C}}^*)_+)$ such that $\xi \in \mathring{\Pi}'$ and such that $\Pi' \cap A$ is defined within Π' by a finite number of (rational) inequalities $x_1 \geqq 1, \ldots, x_l \geqq 1$ (where every $x_i \in NS \otimes_{\mathbb{Z}} \mathbb{Q}$ and where we make the identification $(NS \otimes_{\mathbb{Z}} \mathbb{Q})^{**} \cong NS \otimes_{\mathbb{Z}} \mathbb{Q}$). We may suppose that no inequality is superfluous and that $x_1(\xi) = \ldots = x_l(\xi) = 1$ $(i = 1, \ldots, l)$; (if necessary take a smaller Π'). The convexity of A implies $x_{i|A} \geqq 1$ $(i = 1, \ldots, l)$ and the inclusion $\xi + \mathring{\mathscr{C}}^* \subset A$ yields $x_{i|\mathscr{C}^*} \geqq 0$, so $x_i \in \mathscr{C} \setminus \{0\} \cap NS \otimes_{\mathbb{Z}} \mathbb{Q} \subset \mathscr{C}_+$. Define $\Pi := \mathbb{R}_{\geqq 0}\{x_1, \ldots, x_l\} (\subset \mathscr{C}_+)$.

2.3 Lemma. $\Pi = \{x \in \mathscr{C}_+ : \xi\gamma(x) \geqq \xi(x) \; \forall \gamma \in \Gamma\}$.

Proof. Let $x \neq 0$ be an element of the right hand side.

In order to show that $x \in \Pi$, we may assume $x(\xi) = 1$, for both sets are cones. The inequalities $\xi(\gamma(x)) \geqq \xi(x) \; \forall \gamma \in \Gamma$ and $x_{|\mathscr{C}^*} \geqq 0$ imply $x_{|A} \geqq 1$. But then we conclude that there exist nonnegative scalars $\lambda_1, \ldots, \lambda_l$ with $x = \lambda_1 x_1 + \ldots + \lambda_l x_l$, i.e. $x \in \Pi$.

To prove the other inclusion, it is sufficient to show that $x_1, \ldots, x_l$ are contained in the set on the righthand side. But $x_i(\xi) = 1$ $(i = 1, \ldots, l)$ and $x_{i|A} \geqq 1$ $(i = 1, \ldots, l)$ so this is obviously true.

For every positive scalar N, the set $\{x \in \bar{\mathscr{C}} \mid \xi(x) \leqq N\}$ is clearly compact. It is easy to see that the Γ-orbit of an element of $\mathscr{C}_+$ meets such a set in at most a finite number of points. So the linear function ξ restricted to a Γ-orbit in $\mathscr{C}_+$ assumes its minimum (in at most a finite number of points). From the description of Π given in (2.3) it is then clear that Π is a fundamental domain for the Γ-action on $\mathscr{C}_+$: from every Γ-orbit in $\mathscr{C}_+$, Π picks up the points (finite in number) where the linear function ξ restricted to this orbit assumes its minimum. Now Π is a rational polyhedral cone, so according to [1] for every rational polyhedral cone $\hat{\Pi} \subset \mathscr{C}_+$ the set $\{\gamma \in \Gamma: \gamma(\hat{\Pi}) \cap \Pi \neq \emptyset\}$ is finite. Now we assume that $\xi = \langle p, - \rangle$ where $p \in K \cap NS$ i.e. an ample class.

2.4 Lemma. $\Gamma_B \cdot \Pi = \bar{K} \cap \mathscr{C}_+$.

Proof. First we show that $\Pi \subset \bar{K}$. Let $x \in \Pi$ and let $b \in B$. Then $\xi s_b(x) \geqq \xi(x)$, so $\langle s_b(x), p\rangle \geqq \langle x, p\rangle$ from which we obtain $\langle x, b\rangle \geqq 0$. We find therefore: $x \in \bar{K}$. As Γ_B leaves B invariant, it also leaves $\bar{K}$ invariant. But then we may conclude that $\Gamma_B \cdot \Pi \subset \bar{K} \cap \mathscr{C}_+$. Now let $x \in \bar{K} \cap \mathscr{C}_+$, then there exist $w \in W$, $\gamma \in \Gamma_B$ such that $w\gamma(x) \in \Pi$. We have $\gamma(x) \in \Gamma_B \cdot \bar{K} \subset \bar{K}$ and $w\gamma(x) \in \Pi \subset \bar{K}$. This implies that $\gamma(x) = w\gamma(x) \in \Pi$ (1.2).

2.5 Proposition. *The number of* $\mathrm{Aut}(X)$*-orbits in the set* B *of nodal classes is finite.*

Proof. Let $b \in B$. The wall $H_b \cap \mathscr{C}$ meets $\bar{K} \cap \mathscr{C}$, so there exist $x \in \Pi$ and $\gamma \in \Gamma_B$ such that $\gamma(x) \in \bar{K} \cap \mathscr{C} \cap H_b$ (see (2.4)). If $\gamma^{-1}(b) = c (\in B)$ then γ^{-1} transforms the wall $\bar{K} \cap H_b$ into the wall $\bar{K} \cap H_c$. By the Siegelproperty $\Pi \cap \mathscr{C}$ meets only finitely many walls H_e ($e \in B$). We conclude that the Γ_B-orbit of b contains an element of the finite set $\{c \in B : (\Pi \cap \mathscr{C}) \cap H_c \neq \emptyset\}$. Using (2.1) we obtain the proposition.

Let x, y be two linear independent elements of Π, then $\langle x, y\rangle > 0$ and $\langle x, x\rangle \geqq 0$, $\langle y, y\rangle \geqq 0$. As $\Pi \cap NS$ is a finitely generated semi-group, we find that $\Pi \cap NS$ contains only a finite number of elements of a given norm $d > 0$, and only a finite number of primitive elements of norm $d = 0$.

2.6 Proposition. *Let* d *be an even integer* $\geqq -2$. *Then* $\mathrm{Aut}(X)$ *has a finite number of orbits in the collection of complete linear systems which contain an irreducible curve of selfintersection* d.

Proof. The case $d = -2$ has been dealt with in (2.4). Let C be an irreducible curve on X with $\langle [C], [C]\rangle \geqq 0$, then $[C]$ is necessarily primitive because of [7] §3 theorem 1. There exists a $\gamma \in \Gamma_B$ such that $\gamma([C]) \in \Pi$ because $[C] \in \bar{K} \cap \mathscr{C}_+$. Now apply the remarks above.

2.7 Corollary. *Up to automorphisms of* X, *there are only a finite number of elliptic pencils on* X.

Proof. An elliptic pencil is determined by the class of a smooth elliptic curve on X, which has selfintersection 0. So proposition (2.6) applies.

Remark. Corollary (2.7) answers a question raised by U. Persson.

Combining the propositions (2.2), (2.5) and (2.6) gives theorem (0.1) stated in the introduction.

Bibliography

1. Ash, A., Mumford, D., Rapoport, M., Tai, Y.: Smooth compactification of locally symmetric varieties, Lie groups: history frontiers and applications volume IV. Brookline Massachusetts: Math. Sci. Press 1975
2. Barth, W., Peters, C., Van de Ven, A.: Compact Complex Surfaces. Berlin Heidelberg New York Tokyo: Springer 1984
3. Bourbaki, N.: Groupes et Algébres de Lie, Ch. 4, 5 et 6. Paris: Masson 1981

4. Borel, A., Harish-Chandra: Arithmetic subgroups of algebraic groups. Ann. Math. **75**, 485–535 (1962)
5. Bremner, A.: A geometric approach to equal sums of sixth powers. Proc. Lond. Math. Soc. **43**, 544–581 (1981)
6. Marcus, D.A.: Number Fields. Berlin Heidelberg New York: Springer 1977
7. Piateckii-Shapiro, I., Shafarevic, I.R.: A Torelli theorem for algebraic surfaces of type K-3. Isv. Akad. Nauk. **35**, 530–572 (1971)
8. Vinberg, È.B.: The two most algebraic K3 surfaces. Math. Ann. **265**, 1–21 (1983)

Received November 22, 1984

Math. Z. 189, 515–527 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Filtrations of Principal Indecomposable Modules of Frobenius Kernels of Reductive Groups

Ye Jia-chen*

Department of Mathematics, East China Normal University, Shanghai 200062, People's Republic of China

Let G be a simply connected, semi-simple algebraic group over an algebraically closed field K of characteristic $p>0$, and $\mathscr{U}_n$ the hyperalgebra of the n-th Frobenius kernel (It is a $p^{n\dim G}$-dimensional subalgebra of the hyperalgebra U_K of G). Let T be a maximal torus of G, let X be the character group of T, and W the Weyl group. To each dominant weight $\mu\in X^+\subset X$ associate the Weyl module $V(\mu)$ with highest weight μ. To each "restricted" dominant weight $\lambda\in X_n\subset X$ there corresponds a simple $\mathscr{U}_n$-module $M(\lambda,n)$ and a principal indecomposable $\mathscr{U}_n-T$-module $Q(\lambda,n)$ with $M(\lambda,n)$ as its unique top and bottom composition factor. J. Ballard proved for $p\geqq 3h-3$ (where h is the Coxeter number of G) that each $Q(\lambda,n)$ with $\lambda\in X_n$ admits a G-structure (cf. [2]). This bound was improved to $p\geqq 2h-2$ by J. Jantzen in [10], 4.5. He also proved in [10], 5.9 that each $Q(\lambda,n)$ admits a Weyl filtration, i.e. a G-filtration with subquotients of the form $V(\mu)$, $\mu\in X^+$. More generally such filtrations exist also for all $Q(\lambda,n)\otimes V(\nu)^{(p^n)}$ with $\nu\in X^+$, where the exponent (p^n) denotes a Frobenius twist of a module.

The purpose of this paper is to deal with the multiplicity of $V(\mu)$ occuring in these Weyl filtrations. This problem is related to the question of multiplicities of the simple $\mathscr{U}_n-T$-modules $\hat{M}(\mu,n)$, $\mu\in X$ in composition series of the universal $\mathscr{U}_n-T$-modules $\hat{Z}(\lambda,n)$ with highest weight $\lambda\in X$.

In Sect. 1 we introduce the affine Weyl group $W_{a,p}$ of G, discuss its alcove geometry and a partial order on $W_{a,p}$ which induces an order $\uparrow\uparrow$ on X.

In Sect. 2 we mainly discuss the Weyl filtration of $Q(\lambda,n)$. First we define for each $\lambda\in X_n$ the set $\mathscr{S}_{\lambda,n}=\{\mu\in X^+\mid\mu\uparrow\uparrow\lambda^0-\delta\}$ where $\lambda^0=\sigma_0\lambda-\delta+p^n\delta$. (Here $\sigma_0\in W$ is the longest element, and δ is half the sum of all positive roots.) For each $\mu\in\mathscr{S}_{\lambda,n}$ set b_μ^λ equal to the multiplicity of $\hat{M}(\lambda,n)$ as composition factor of $\hat{Z}(\mu+p^n\delta,n)$. We then prove for $p\geqq 2h-2$, for all $\sigma\in W$, $\lambda\in X_n$, $\mu\in\mathscr{S}_{\lambda,n}$, $\nu\in X^+$ with $\sigma\cdot\mu+p^n(\nu+\delta)\in X^+$ that the multiplicity of $V(\sigma\cdot\mu+p^n(\nu+\delta))$ in a Weyl filtration of $Q(\lambda,n)\otimes V(\nu)^{(p^n)}$ is equal to

$$\sum_{\tau\in W}\det(\tau)[\hat{Z}(\sigma\cdot\mu+p^n(\nu+\delta),n):\hat{M}(\lambda+p^n\tau\cdot\nu,n)].$$

(Here we use the dot notation $\sigma\cdot\mu=\sigma(\mu+\delta)-\delta$.)

* *Current address:* Department of Applied Mathematics, Tong Ji University, Shanghai 200092, People's Republic of China

This is applied in Sect. 3 in the case $n=1$ to prove the following main theorem:

Suppose $p \geqq 2h-2$. *Let* $\lambda \in X_1$ *be p-regular. Then*

(1) $b^{\lambda}_{\lambda^0 - \delta} = 1$;

(2) *If* $\mu, \mu' \in \mathscr{S}_{\lambda,1}$ *with* $\mu' \uparrow\uparrow \mu$, *then* $b^{\lambda}_{\mu} \leqq b^{\lambda}_{\mu'}$.

As a corollary we get $b^{\lambda}_{\mu} > 0$ for all $\mu \in \mathscr{S}_{\lambda,1}$.

The author welcomes this opportunity to express his deepest gratitude to Professor Cao Xi-hua for his assistance and encouragement throughout this work. The author also wishes to give his thanks to Professor Jens C. Jantzen for his comments on the organization of this paper, and for many conversations on this and related subjects during his stay at East China Normal University in August 1984.

1. The Affine Weyl Group $W_{a,p}$

Let G be a simply connected, semi-simple algebraic group over an algebraically closed field K of characteristic $p>0$, let T be a maximal torus in G and B a Borel subgroup of G containing T. Let Φ be the root system of (G, T), and let $\Delta = \{\alpha_1, \ldots, \alpha_l\}$ be the set of all simple roots such that B corresponds to the set Φ^+ of positive roots. Set $\Phi^- = -\Phi^+$. We assume that Φ is irreducible and that G, B, T are defined and split over the prime field $\mathbb{F}_p$.

Let W be the Weyl group of G and X the character group of T. Then W acts on X and on $E = X \otimes \mathbb{R}$, and there is on E a positive defined inner product $(\ ,\ )$ invariant under W (and unique up to a scalar multiple with this property). For each root α define the dual root $\alpha^{\vee}$ as $\alpha^{\vee} = \dfrac{2\alpha}{(\alpha, \alpha)}$. Define the fundamental weights ω_{α_i} by $(\omega_{\alpha_i}, \alpha_j^{\vee}) = \delta_{ij}$, $1 \leqq i, j \leqq l$. Then we have $X = \sum_{i=1}^{l} \mathbb{Z}\omega_{\alpha_i}$, we call $X^+ = \sum_{i=1}^{l} \mathbb{Z}^+ \omega_{\alpha_i}$ the set of dominant weights and $X_{rt} = \sum_{i=1}^{l} \mathbb{Z}\alpha_i$ the root lattice of G. Denote the index of connection by $f = [X : X_{rt}]$ (=determinant of the Cartan matrix of Φ), and write $\delta = \sum_{i=1}^{l} \omega_{\alpha_i}$ (=half the sum of positive roots).

For each root β let σ_β be the reflection with $\sigma_\beta(\chi) = \chi - (\chi, \beta^{\vee})\beta$ for all $\chi \in E$, and let $P_\beta \subset E$ be the reflecting hyperplane, i.e. $P_\beta = \{\chi \in E \mid (\chi, \beta^{\vee}) = 0\} = \{\chi \in E \mid \sigma_\beta(\chi) = \chi\}$. Set $\sigma_i = \sigma_{\alpha_i}$. Then W is a Coxeter group with generators $\{\sigma_1, \ldots, \sigma_l\}$. The connected components of $E - \bigcup_{\beta \in \Phi^+} P_\beta$ are called Weyl chambers. The group W acts simply transitively on the set of Weyl chambers. The closure of any Weyl chamber is a fundamental region for the action of W on E. The open region $E^+ = \{\chi \in E \mid (\chi, \alpha_i^{\vee}) > 0 \text{ for } 1 \leqq i \leqq l\}$ is a Weyl chamber which we call the fundamental Weyl chamber. Let $\sigma_0 \in W$ be the longest element of W and $\alpha_0 \in \Phi^+$ be the highest short root. Denote by h the Coxeter number, $h = (\delta, \alpha_0^{\vee}) + 1$.

Let $W_{a,p}$ be the affine Weyl group of G relative to p, i.e. the group generated by all reflections $S_{\beta,m}$ ($\beta \in \Phi^+$, $m \in \mathbb{Z}$) where $S_{\beta,m}(\chi) = \chi - ((\chi, \beta^{\vee})$

$-mp)\beta$ (clearly $S_{\beta,0}=\sigma_\beta$). Then $W_{a,p}$ is a Coxeter group on generators $\{S_0, S_1, \ldots, S_l\}$ where $S_i=\sigma_i$, $1\leqq i\leqq l$ and $S_0=S_{\alpha_0,1}$. Denote by T_v for each v the translation on E sending χ to $\chi+v$. Then $W_{a,p}$ is also equal to the semidirect product of W with the normal subgroup (isomorphic to X_{rt}) consisting of all translation T_{pv}, $v\in X_{rt}$. The following properties of the reflections in $W_{a,p}$ are well known and easily proved.

(1.1) **Lemma.** *Suppose* $\beta, v\in\Phi^+$, $m,n\in\mathbb{Z}$, $\sigma\in W$ *and* $v\in X$. *Then*

(1) $S_{\beta,m}^2=1$

(2) $S_{\beta,m}=S_{-\beta,-m}$

(3) $\sigma S_{\beta,m}\sigma^{-1}=S_{\sigma(\beta),m}$

(4) $T_{pv}S_{\beta,m}T_{-pv}=S_{\beta,m+(v,\beta^\vee)}$

(5) $S_{\gamma,n}S_{\beta,m}S_{\gamma,n}=S_{\sigma_\gamma(\beta),m-n(\gamma,\beta^\vee)}$.

Define the "dot" action of W and $W_{a,p}$ on E as follows:

$$\sigma\cdot\chi=\sigma(\chi+\delta)-\delta \quad\text{and}\quad w\cdot\chi=w(\chi+\delta)-\delta \qquad \forall\,\chi\in E,\ \sigma\in W,\ w\in W_{a,p}.$$

Let $H_{\beta,m}$ ($\beta\in\Phi^+$, $m\in\mathbb{Z}$) denote the reflecting hyperplane with respect to $S_{\beta,m}$ under the "dot" action. Then the connected components of $E-\bigcup_{\substack{\beta\in\Phi^+\\ m\in\mathbb{Z}}} H_{\beta,m}$ are called alcoves. In particular, the open region C_0 bounded by the $H_{\alpha_i,0}$ ($1\leqq i\leqq l$) and $H_{\alpha_0,1}$ is an alcove which we call the lowest dominant alcove. The closure of any alcove (e.g. $\bar{C}_0$) is a fundamental region for the "dot" action of $W_{a,p}$ on E and $W_{a,p}$ operates simply transitively on the set $\mathscr{C}$ of all alcoves. By sending w to $w\cdot C_0$ we get therefore a $1-1$ correspondence between $W_{a,p}$ and the set $\mathscr{C}$. If $C=w\cdot C_0$ for some $w\in W_{a,p}$ and $\lambda\in C_0\cap X$, then we denote by $\lambda_c=w\cdot\lambda$ the mirror image of λ in C.

Any $w\in W_{a,p}$ can be written in the form $w=S_{i_1}\ldots S_{i_t}$ with $0\leqq i_1,\ldots,i_t\leqq l$. If t is minimal, we call this a reduced decomposition of w and set $l(w)=t$. This length is equal to the number of reflecting hyperplanes that separate C_0 from $w\cdot C_0$. The restriction of $l(\ )$ to W is the usual length function on W.

Let $W_{a,p}^+=\{w\in W_{a,p}\mid l(S_iw)>l(w), 1\leqq i\leqq l\}$. This is a set of distinguished left coset representatives for $W_{a,p}/W$. Let $\mathscr{C}^+=\{C\in\mathscr{C}\mid(\chi+\delta,\alpha_i^\vee)>0,\ 1\leqq i\leqq l,\ \forall\,\chi\in C\}$ be the set of all alcoves which lie in $E^+-\delta$. Then the restriction of the $1-1$ correspondence $w\mapsto w\cdot C_0$ to $W_{a,p}^+$ is a $1-1$ correspondence between $W_{a,p}^+$ and $\mathscr{C}^+$.

Let C be an alcove. Then (for each $\beta\in\Phi^+$) the number $n_\beta=\left[\dfrac{(\chi+\delta,\beta^\vee)}{p}\right]$ is independent of the choice of χ inside C. Therefore

$$d(C)=\sum_{\beta\in\Phi^+} n_\beta \quad\text{and}\quad \bar{d}(C)=\sum_{\beta\in\Phi}|n_\beta|$$

are well defined numbers. For $C\in\mathscr{C}^+$ obviously $d(C)=\bar{d}(C)$. For any $C=w\cdot C_0$ with $w\in W_{a,p}$ one has $\bar{d}(C)=l(w)$ and $\bar{d}(C)$ is equal to the number of reflecting hyperplanes separating C and C_0.

Following Verma (cf. [13]) we introduce partial orders on X, $\mathscr{C}$ and $W_{a,p}$. For $\lambda\in X$ and $\beta\in\Phi^+$, $m\in\mathbb{Z}$ we set $S_{\beta,m}\cdot\lambda\uparrow\lambda$ if $(\lambda+\delta,\beta^\vee)\geqq mp$ and $\lambda\uparrow S_{\beta,m}\cdot\lambda$ if

$(\lambda+\delta, \beta^\nu) \leqq mp$. Let $\uparrow$ be the partial order on X generated by this, i.e. for $\lambda, \mu \in X$ the relation $\lambda \uparrow \mu$ holds if and only if there is a sequence $\lambda_0 = \lambda$, $\lambda_1, \ldots, \lambda_k = \mu$ in X and reflections $S_{\beta_i, m_i} \in W_{a,p}$ with $\lambda_{i-1} = S_{\beta_i, m_i} \cdot \lambda_i \uparrow \lambda_i$ for $1 \leqq i \leqq k$. Obviously $\lambda \uparrow \mu$ implies $\lambda \leqq \mu$ and $\lambda \in W_{a,p} \cdot \mu$.

For any alcove $C \in \mathscr{C}$ and any reflection $S_{\beta,m} \in W_{a,p}$ (with $\beta \in \Phi^+$, $m \in \mathbb{Z}$) we set $S_{\beta,m} \cdot C \underset{\overline{A}}{\leqq} C$ if $(\chi+\delta, \beta^\nu) > mp$ for all $\chi \in C$ and $C \underset{\overline{A}}{\leqq} S_{\beta,m} \cdot C$ otherwise. Let $\underset{\overline{A}}{\leqq}$ be the partial order on $\mathscr{C}$ generated by this, i.e. for $C, C' \in \mathscr{C}$ we have $C \underset{\overline{A}}{\leqq} C'$ if and only if there is a sequence $C_0 = C, C_1, \ldots, C_k = C'$ in $\mathscr{C}$ and reflections $S_{\beta_i, m_i} \in W_{a,p}$ with $C_{i-1} = S_{\beta_i, m_i} \cdot C_i \underset{\overline{A}}{\leqq} C_i$ for $1 \leqq i \leqq k$. It is easy to show that this coincides with Lusztig's definition in [12].

For $w, w' \in W_{a,p}$ we set $w \underset{\overline{A}}{\leqq} w'$ if and only if $w \cdot C_0 \underset{\overline{A}}{\leqq} w' \cdot C_0$.

Let $C \in \mathscr{C}$, $\lambda \in \bar{C} \cap X$ and $w \in W_{a,p}$. The definitions imply easily: If $C \underset{\overline{A}}{\leqq} w \cdot C$, then $\lambda \uparrow w \cdot \lambda$. If $\lambda \uparrow w \cdot \lambda$, then $C' \underset{\overline{A}}{\leqq} w \cdot C$ for some alcove C' with $\lambda \in \bar{C}'$. If λ belongs to the upper (resp. lower) closure of C (i.e. if $(\lambda+\delta, \beta^\nu) = mp$ with $\beta \in \Phi^+$ and $m \in \mathbb{Z}$ implies $(\chi+\delta, \beta^\nu) < mp$ resp. $(\chi+\delta, \beta^\nu) > mp$ for all $\chi \in C$), then $C \underset{\overline{A}}{\leqq} C'$ (resp. $C' \underset{\overline{A}}{\leqq} C$) for any C' with $\lambda \in \bar{C}'$ by [7], Lemma 6 (ii). Thus we get

(1.2) **Lemma.** *If $\lambda \in X$ belongs to the upper (resp. lower) closure of $C \in \mathscr{C}$, then for all $w \in W_{a,p}$*

$$\lambda \uparrow w \cdot \lambda \Leftrightarrow C \underset{\overline{A}}{\leqq} w \cdot C$$

$$(\textit{resp. } w \cdot \lambda \uparrow \lambda \Leftrightarrow w \cdot C \underset{\overline{A}}{\leqq} C).$$

The following properties of $\uparrow$ and $\underset{\overline{A}}{\leqq}$ are well known:

(1.3) **Lemma.** *Suppose $C, C' \in \mathscr{C}$ and $v \in X$, $\beta \in \Phi^+$. Then we have*

(1) $C' \underset{\overline{A}}{\leqq} C \Leftrightarrow T_{pv} C' \underset{\overline{A}}{\leqq} T_{pv} C$

(2) $C' \underset{\overline{A}}{\leqq} C \Leftrightarrow \sigma_0 \cdot C \underset{\overline{A}}{\leqq} \sigma_0 \cdot C'$

(3) $C \underset{\overline{A}}{\leqq} T_{p\beta} C$

(4) *If $C \in \mathscr{C}^+$, then $\sigma_0 \cdot C \underset{\overline{A}}{\leqq} \sigma \cdot C \underset{\overline{A}}{\leqq} C$ for all $\sigma \in W$.*

(1.4) **Lemma.** *Suppose $\lambda, \mu \in X$ and $v \in X$, $\beta \in \Phi^+$. Then we have*

(1) $\lambda \uparrow \mu \Leftrightarrow T_{pv} \lambda \uparrow T_{pv} \mu$

(2) $\lambda \uparrow \mu \Leftrightarrow \sigma_0 \cdot \mu \uparrow \sigma_0 \cdot \lambda$

(3) $\mu \uparrow T_{p\beta} \mu$

(4) *If $\mu \in X^+ - \delta$, then $\sigma_0 \cdot \mu \uparrow \sigma \cdot \mu \uparrow \mu$ for all $\sigma \in W$.*

The following properties of $\underset{\overline{A}}{\leqq}$ were proved in [7], Lemma 4 and Corollary 1 to Satz 9.

(1.5) **Lemma.** *Suppose* $C \in \mathscr{C}$.

(1) *If* $S_{\beta,m} \in W_{a,p}$ *is a reflection, then* $S_{\beta,m} \cdot C \underset{\bar{A}}{\leqq} C \Leftrightarrow d(S_{\beta,m} \cdot C) \leqq d(C)$;

(2) *If* $C' \in \mathscr{C}$ *with* $C' \underset{\bar{A}}{\leqq} C$, *then* $d(C') \leqq d(C)$ *and there exist* $C_i \in \mathscr{C}$ *with* $C' = C_1 \underset{\bar{A}}{\leqq} C_2 \underset{\bar{A}}{\leqq} \ldots \underset{\bar{A}}{\leqq} C_k \underset{\bar{A}}{\leqq} C_{k+1} = C$ *and* $d(C_i) = d(C) + i - (k+1)$ *for* $1 \leqq i \leqq k+1$.

It was noted by H. Andersen in [1] that (2) can be sharpened to the following (Remember [13] 1.6)

(1.6) **Lemma.** *If* $C, C' \in \mathscr{C}^+$ *with* $C' \underset{\bar{A}}{\leqq} C$, *then we can choose* $\{C_i\}_{1 \leqq i \leqq k+1}$ *in* (1.5(2)) *such that all* C_i *belong to* $\mathscr{C}^+$.

For each $\lambda \in X$ there is a unique $C \in \mathscr{C}$ such that λ belongs to its lower closure. If $\lambda \in X^+ - \delta$, then $C \in \mathscr{C}^+$. Therefore (1.6) implies

(1.7) **Lemma.** *If* $\lambda, \mu \in X^+ - \delta$ *with* $\lambda \uparrow \mu$, *then there is a sequence* $\lambda = \mu_1 \uparrow \mu_2 \uparrow \ldots \uparrow \mu_k \uparrow \mu_{k+1} = \mu$ *and reflections* $S_{\beta_i, m_i} \in W_{a,p}$ *with* $\beta_i \in \Phi^+$ *and* $m_i > 0$ *such that all* $\mu_i \in X^+ - \delta$ *and* $\mu_{i+1} = S_{\beta_i, m_i} \cdot \mu_i$.

(Note that for $m_i \leqq 0$ we can not have $\mu_i, \mu_{i+1} \in X^+ - \delta$ at the same time except when $\mu_{i+1} = \mu_i$ where we can delete μ_{i+1}.)

Let us consider the following modification of $\uparrow$ and $\underset{\bar{A}}{\leqq}$. For $\lambda \in X$, $\beta \in \Phi^+$ and $m \geqq 0$ we set $S_{\beta,m} \cdot \lambda \uparrow\uparrow \lambda$ if $(\lambda + \delta, \beta^\vee) \geqq mp$ and $\lambda \uparrow\uparrow S_{\beta,m} \cdot \lambda$ if $mp \leqq (\lambda + \delta, \beta^\vee)$, and define $\uparrow\uparrow$ in general as the order relation generated by this. Similarly we define $C \uparrow\uparrow C'$. Obviously $\lambda \uparrow\uparrow \mu$ implies $\lambda \uparrow \mu$ and $C \uparrow\uparrow C'$ implies $C \underset{\bar{A}}{\leqq} C'$. From (1.6) and (1.7) we get the converse in a special case.

(1.8) **Lemma.** (1) *For* $C', C \in \mathscr{C}^+$ *we have* $C' \underset{\bar{A}}{\leqq} C \Leftrightarrow C' \uparrow\uparrow C$.

(2) *For* $\lambda, \mu \in X^+ - \delta$ *we have* $\lambda \uparrow \mu \Leftrightarrow \lambda \uparrow\uparrow \mu$.

The statement in (1.4(3)) has the following counterpart

(1.9) **Lemma.** *Let* $\lambda \in X$ *and* $\beta \in \Phi^+$. *Suppose* $n \in \mathbb{Z}^+$ *satisfies* $np < (\lambda + \delta, \beta^\vee) \leqq (n+1)p$. *Then* $\lambda - mp\beta \uparrow\uparrow \lambda$ *for all* $m \in \mathbb{Z}$ *with* $0 \leqq m \leqq n$.

Proof. For $m > 0$ we have $\lambda - mp\beta = (S_{\beta,n-m} S_{\beta,n}) \cdot \lambda \uparrow\uparrow S_{\beta,n} \cdot \lambda \uparrow\uparrow \lambda$ as

$$(S_{\beta,n} \cdot \lambda + \delta, \beta^\vee) = (\sigma_\beta(\lambda + \delta) + np\beta, \beta^\vee) = 2np - (\lambda + \delta, \beta^\vee)$$
$$\geqq (n-1)p \geqq (n-m)p \geqq 0.$$

Finally we need

(1.10) **Lemma.** (1) *Let* $C, C' \in \mathscr{C}$ *with* $C \uparrow\uparrow C'$. *Let* $\sigma, \sigma' \in W$ *be the unique elements with* $\sigma \cdot C, \sigma' \cdot C' \in \mathscr{C}^+$. *Then* $\sigma \cdot C \uparrow\uparrow \sigma' \cdot C'$.

(2) *Let* $\lambda, \mu \in X$ *with* $\lambda \uparrow\uparrow \mu$. *There are* $\sigma, \sigma' \in W$ *with* $\sigma \cdot \lambda$, $\sigma' \cdot \mu \in X^+ - \delta$. *Then* $\sigma \cdot \lambda \uparrow\uparrow \sigma' \cdot \mu$.

Proof. (1) We may assume $C = S_{\beta,m} \cdot C'$ with $\beta \in \Phi^+$ and $m \geqq 0$. If $m = 0$, then $\sigma \cdot C = \sigma' \cdot C'$ and nothing has to be shown. Suppose $m > 0$. Then $\sigma' \cdot C$

$=\sigma' S_{\beta,m}\sigma'^{-1}\cdot(\sigma'\cdot C')=S_{\sigma'(\beta),m}\cdot(\sigma'\cdot C')$. For any $\chi\in C'$ we have $0<mp<(\chi+\delta,\beta^\vee)=(\sigma'(\chi+\delta),\sigma'(\beta)^\vee)$, hence $\sigma'(\beta)\in\Phi^+$. Therefore [7] Satz 9 implies $\sigma\cdot C=\sigma\sigma'^{-1}\cdot(\sigma'\cdot C)\uparrow\sigma'\cdot C'$ and by (1.8(1)) we get $\sigma\cdot C\uparrow\uparrow\sigma'\cdot C'$.

(2) *Can be proved similarly.*

2. The Weyl Filtration of $Q(\lambda, n)$

Let $\mathfrak{g}$ be the restructed Lie algebra of G, and $\mathscr{U}$ its restricted universal enveloping algebra. Let U_K be the hyperalgebra of G, and $\mathscr{U}_n$ the hyperalgebra of the n-th Frobenius kernel of G. This is a $p^{n\dim G}$-dimensional subalgebra of U_K, we have $\mathscr{U}_1\simeq\mathscr{U}$ and $\mathscr{U}_1\subset\mathscr{U}_2\subset\ldots\subset\mathscr{U}_n\subset\ldots\subset U_K$. The algebras $\mathscr{U}_n$ are Hopf algebras and symmetric Frobenius algebras.

Let $X_n=\{\lambda\in X^+|(\lambda,\alpha_i^\vee)<p^n,\ 1\leqq i\leqq l\}$ be the set of "restricted" dominant weights of G. For each $\lambda\in X_n$ there is an irreducible $\mathscr{U}_n$-module $M(\lambda,n)$ with "highest" weight λ and each irreducible $\mathscr{U}_n$-module is isomorphic to exactly one $M(\lambda,n)$ with $\lambda\in X_n$. Therefore X_n may also be seen as an index set for distinct $\mathscr{U}_n$-PIMs (principal indecomposable module). Denote by $Q(\lambda,n)$ the PIM with $M(\lambda,n)$ as its unique top and bottom composition factor. For each $\lambda\in X_n$ set $Z(\lambda,n)=\mathscr{U}_n\otimes_{\mathscr{E}_n}K_\lambda$ where $\mathscr{E}_n$ is the hyperalgebra of the n-th Frobenius kernel of B and K_λ is the 1-dimensional $\mathscr{E}_n$-module corresponding to $\lambda\in X\simeq X(B)$. The $\mathscr{U}_n$-module $Z(\lambda,n)$ is the "universal highest weight module" with "highest" weight λ. It is indecomposable and has unique irreducible quotient $M(\lambda,n)$. In [8] J. Jantzen proved that $Q(\lambda,n)$ has a $\mathscr{U}_n$-filtration with subquotients $Z(\mu,n)$, the Z-filtration, and that the number of times $Z(\mu,n)$ occurs as a subquotient in such a filtration of $Q(\lambda,n)$ is equal to the multiplicity of $M(\lambda,n)$ as a $\mathscr{U}_n$-composition factor of $Z(\mu,n)$, i.e.

$$(Q(\lambda,n):Z(\mu,n))=[Z(\mu,n):M(\lambda,n)].$$

For $\lambda\in X^+$ let $V(\lambda)$ be the Weyl module with highest weight λ. This is an indecomposable G-module having a unique irreducible quotient denoted by $M(\lambda)$, and it is a universal highest weight module. We write $\mathrm{ch}(\lambda)$ (resp. $p-\mathrm{ch}(\lambda)$) for the formal character of $V(\lambda)$ (resp. $M(\lambda)$). For $\mu\in X^+-\delta$, $\mu\notin X^+$ we set $V(\mu)=0$ and $\mathrm{ch}(\mu)=0$. We extend the definition of $\mathrm{ch}(\mu)$ to all $\mu\in X$: There is $\sigma\in W$ with $\sigma\cdot\mu\in X^+-\delta$. Then set $\mathrm{ch}(\mu)=\det(\sigma)\,\mathrm{ch}(\sigma\cdot\mu)$. (If $\mathrm{ch}(\sigma\cdot\mu)\neq 0$, then σ is unique.)

Let $F_n: G\to G$ be the n-th power of the Frobenius endomorphism (Remember that G is defined over $\mathbb{F}_p$). For each G-module V we introduce a new G-module $V^{(p^n)}$ which we call the n-th Frobenius twist of V. It has the same underlying vector space as V with a new action of G given by $g\circ v=F_n(g)v$ for every $g\in G$ and $v\in V$. Any $\lambda\in X^+$ has a p-adic expansion $\lambda=\lambda_0+p\lambda_1+\ldots+p^k\lambda_k$ where $\lambda_0,\lambda_1,\ldots,\lambda_k\in X_1$. By Steinberg's tensor product theorem we have then

$$M(\lambda)\simeq M(\lambda_0)\otimes M(\lambda_1)^{(p)}\otimes\ldots\otimes M(\lambda_k)^{(p^k)}.$$

For $\lambda\in X_n$ we have $M(\lambda)\simeq M(\lambda,n)$ as a $\mathscr{U}_n$-module, for $\mu\in X_n$ and $\nu\in X^+$ we have $M(\mu+p^n\nu)\simeq M(\mu)^{\dim M(\nu)}$ as a $\mathscr{U}_n$-module. For $\lambda=(p^n-1)\delta$ we have $V((p^n$

$-1)\delta)=M((p^n-1)\delta)$ and this G-module is as $\mathcal{U}_n$-module isomorphic to $Q((p^n-1)\delta, n)=Z((p^n-1)\delta, n)=M((p^n-1)\delta, n)$. We call this module the n-th Steinberg module and denote it by St_n.

There is a bijective map $\lambda\mapsto\lambda^0$ of X_n onto itself given by $\lambda^0=\sigma_0\lambda+(p^n-1)\delta=\sigma_0\cdot(\lambda-\delta)+p^n\delta$ for all $\lambda\in X_n$. Obviously $(\lambda^0)^0=\lambda$ for all λ.

J. Ballard proved for $p\geqq 3h-3$ that each $Q(\lambda, n)$ with $\lambda\in X_n$ admits a (unique by [8]) G-module structure (cf. [2]). This bound was improved to $p\geqq 2h-2$ in [10], 4, 5. In both cases $Q(\lambda, n)$ was constructed as a direct summand of $M(\lambda^0)\otimes St_n$. It is however clear that we can replace $M(\lambda^0)$ by $V(\lambda^0)$ or any G-module such that all weights μ satisfy $\mu\leqq\lambda^0$ and such that the λ^0 weight space has dimension 1 (see [5]).

It is proved in [10], 5.9 that $Q(\lambda, n)$ admits a Weyl filtration, i.e. a G-filtration with subquotients of the form $V(\mu)$ with $\mu\in X^+$. All μ occuring satisfy $(\mu, \alpha_0^\vee)<2p^n(h-1)$ and for such μ the Weyl module $V(\mu)$ occurs $[V(\mu): M(\lambda)]$ times as a subquotient in such a filtration. The existence of the filtration can be proved more easily than in [10], 5.6 using [14], Theorem B and Proposition 3.4 together with the realization of $Q(\lambda, n)$ as a direct summand of $V(\lambda^0)\otimes St_n$.

More generally one has

(2.1) **Proposition.** *Assume $p\geqq 2h-2$. For all $\lambda\in X_n$ and $\nu\in X^+$ the tensor product $Q(\lambda, n)\otimes V(\nu)^{(p^n)}$ has highest weight $\lambda^0-\delta+p^n(\nu+\delta)$. It is a direct G-summand of $V(\lambda^0)\otimes St_n\otimes V(\nu)^{(p^n)}$ and admits a Weyl filtration which starts with a submodule isomorphic to $V(\lambda^0-\delta+p^n(\nu+\delta))$.*

Proof. The previous remarks show that $Q(\lambda, n)\otimes V(\nu)^{(p^n)}$ is a direct summand of $V(\lambda^0)\otimes St_n\otimes V(\nu)^{(p^n)}$ which is isomorphic to $V(\lambda^0)\otimes V(p^n(\nu+\delta)-\delta)$ by [10], 5.5. Again using [14] Theorem B and Proposition 3.4 gives the Weyl filtration. (It was first proved to exist in [10], 5.5.) For the statement about the highest weight see e.g. [10], 6.2. Whenever we have a Weyl filtration of a G-module V then we can take any maximal weight μ of V and get a submodule $V(\mu)$ at the beginning of the filtration, cf. [10], 5.2(2).

Let us use the category of $\mathcal{U}_n-T$-modules (the definition in [8] or [10]). For each $\lambda\in X$ there are a universal highest weight module $\hat{Z}(\lambda, n)=\mathcal{U}_n\otimes_{\mathfrak{b}_n}K_\lambda$ of weight λ, its unique irreducible quotient $\hat{M}(\lambda, n)$ and the projective cover $\hat{Q}(\lambda, n)$ of $\hat{M}(\lambda, n)$. For $\lambda\in X_n$ there are isomorphisms of $\mathcal{U}_n$-modules $\hat{M}(\lambda, n)\simeq M(\lambda, n)$, $\hat{Z}(\lambda, n)\simeq Z(\lambda, n)$, $\hat{Q}(\lambda, n)\simeq Q(\lambda, n)$. For all $\lambda, \nu\in X$ there are isomorphisms of $\mathcal{U}_n-T$-modules $\hat{M}(\lambda+p^n\nu, n)\simeq\hat{M}(\lambda, n)\otimes p^n\nu$, $\hat{Z}(\lambda+p^n\nu, n)\simeq\hat{Z}(\lambda, n)\otimes p^n\nu$, $\hat{Q}(\lambda+p^n\nu, n)\simeq\hat{Q}(\lambda, n)\otimes p^n\nu$. For $\lambda\in X_n$ the G-module $M(\lambda)$ is isomorphic to $M(\lambda, n)$ as a $\mathcal{U}_n-T$-module, if $Q(\lambda, n)$ admits a G-module structure, this G-structure induces the $\mathcal{U}_n-T$-module structure of $\hat{Q}(\lambda, n)$.

Let $e(\mu)$, $\mu\in X$ be the canonical basis of $\mathbb{Z}[X]$. Set for any $\mu\in X$

$$S(\mu)=\sum_{\nu\in W_\mu}e(\nu)=\sum_{\sigma\in W}\frac{1}{|\mathrm{Stab}_W\mu|}e(\sigma\mu).$$

By [11], 2.5(2) we have for all $\lambda\in X_n$

$$\mathrm{ch}\,\hat{Q}(\lambda, n)=\sum_{\mu\in X^+}[\hat{Z}(\mu-\delta, n):\hat{M}(\lambda-p^n\delta, n)]\,S(\mu)\,\mathrm{ch}\,St_n,$$

hence for all $v \in X^+$

$$\mathrm{ch}\,\hat{Q}(\lambda, n) \otimes V(v)^{(p^n)} = \sum_{\mu \in X^+} [\hat{Z}(\mu + (p^n - 1)\delta, n) : \hat{M}(\lambda, n)]\, S(\mu)\, \mathrm{ch}(St_n \otimes V(v)^{(p^n)}).$$

As $St_n \otimes V(v)^{(p^n)} \simeq V(p^n(v+\delta) - \delta)$ by [10], 5.5 we get

(2.2) **Proposition.** *For all $\lambda \in X_n$, $v \in X^+$ we have*

$$\mathrm{ch}\,\hat{Q}(\lambda, n) \otimes V(v)^{(p^n)} = \sum_{\mu \in X^+} b_\mu^\lambda \sum_{\sigma \in W} \frac{1}{|\mathrm{Stab}_W \mu|}\, \mathrm{ch}(\sigma\mu + p^n(v+\delta) - \delta)$$

where $b_\mu^\lambda = [\hat{Z}(\mu + (p^n - 1)\delta, n) : \hat{M}(\lambda, n)]$.

Set $\mathscr{S}_{\lambda,n} = \{\mu \in X^+ \mid \mu \uparrow\uparrow \lambda^0 - \delta\}$. For each $\mu \in \mathscr{S}_{\lambda,n}$ we have $\mu \leqq \lambda^0$, hence μ is a Weight of $V(\lambda^0)$ by [3], §13.2 Lemma B and §13.4 Lemma B. Thus $\mathscr{S}_{\lambda,n}$ is a finite set with $\lambda^0 - \delta$ as its largest element.

(2.3) **Lemma.** *Let $\lambda \in X_n$, $\mu \in X^+$. If $b_\mu^\lambda \neq 0$, then $\mu \in \mathscr{S}_{\lambda,n}$.*

Proof. By definition $0 \neq b_\mu^\lambda = [\hat{Z}(\mu + p^n\delta, n) : \hat{M}(\lambda, n)]$. Now [10], 3.3 implies $\mu + p^n\delta \uparrow \sigma_0 \cdot \lambda + 2p^n\delta = \lambda^0 - \delta + p^n\delta$, hence $\mu \uparrow \lambda^0 - \delta$ by (1.4(1)), hence $\mu \uparrow\uparrow \lambda^0 - \delta$ by (1.8(2)).

(2.4) **Proposition.** *Let $\lambda \in X_n$ and $\mu \in \mathscr{S}_{\lambda,n}$. Then for all $\sigma \in W$ we have*

$$\lambda \uparrow \sigma \cdot \mu + p^n\delta \uparrow\uparrow \lambda^0 - \delta + p^n\delta.$$

Proof. By (1.4(4)) we have $\sigma_0 \cdot \mu \uparrow\uparrow \sigma \cdot \mu \uparrow\uparrow \mu$ where we can go step by step with reflections of the form $S_{\beta,0}$, $\beta \in \Phi^+$ only. Then we get

$$\sigma_0 \cdot \mu + p^n\delta \uparrow\uparrow \sigma \cdot \mu + p^n\delta \uparrow\uparrow \mu + p^n\delta$$

as we are using reflections $S_{\beta,m}$ with $m = p^{n-1}(\delta, \beta^\vee) > 0$ only. From $\mu \uparrow\uparrow \lambda^0 - \delta$ we get $\mu + p^n\delta \uparrow\uparrow \lambda^0 - \delta + p^n\delta$, hence $\sigma \cdot \mu + p^n\delta \uparrow\uparrow \lambda^0 - \delta + p^n\delta$. Using (1.4(2)) we get $\sigma_0 \cdot (\lambda^0 - \delta + p^n\delta) \uparrow \sigma_0 \cdot (\mu + p^n\delta)$, hence $\lambda - 2p^n\delta \uparrow \sigma_0 \cdot \mu - p^n\delta$ and $\lambda \uparrow \sigma_0 \cdot \mu + p^n\delta \uparrow \sigma \cdot \mu + p^n\delta$.

(2.5) **Proposition.** *Let $\lambda \in X_n$ and $\mu \in X^+$ with $\mu \leqq \lambda^0$. If $\lambda \uparrow \sigma_0 \cdot \mu + p^n\delta$ or $\mu + p^n\delta \uparrow\uparrow \lambda^0 - \delta + p^n\delta$, then $\mu \in \mathscr{S}_{\lambda,n}$.*

Proof. Certainly $\mu + p^n\delta \uparrow\uparrow \lambda^0 - \delta + p^n\delta$ implies $\mu \uparrow \lambda^0 - \delta$, hence $\mu \in \mathscr{S}_{\lambda,n}$ using (1.8(2)). If $\lambda \uparrow \sigma_0 \cdot \mu + p^n\delta$, apply σ_0 to both sides and then argue as before.

(2.6) **Proposition.** *Let $\lambda \in X_n$, $\mu \in \mathscr{S}_{\lambda,n}$, $\sigma \in W$ and $v \in X^+$. There is $\tau \in W$ with $\tau \cdot (\sigma \cdot \mu + p^n(v+\delta)) \in X^+ - \delta$, and there are a unique $\mu' \in X^+$ and a $\sigma' \in W$ with $\tau \cdot (\sigma \cdot \mu + p^n(v+\delta)) = \sigma' \cdot \mu' + p^n(v+\delta)$. Then $\mu' \in \mathscr{S}_{\lambda,n}$ and $\mu' \uparrow\uparrow \mu$.*

Proof. The existence of τ, μ', σ' and the uniqueness of μ' are clear. We have only to prove the last statements. Let us use induction over the number of $\beta \in \Phi^+$ with $(\sigma \cdot \mu + p^n(v+\delta) + \delta, \beta^\vee) < 0$. If this number is 0, then $\sigma \cdot \mu + p^n(v + \delta) \in X^+ - \delta$ and $\mu' = \mu$.

If not, then there is a simple root α_i with $(\sigma \cdot \mu + p^n(v+\delta) + \delta, \alpha_i^\vee) < 0$. We can write $\sigma_i \cdot (\sigma \cdot \mu + p^n(v+\delta))$ in the form $\sigma' \cdot \mu' + p^n(v+\delta)$ with $\mu' \in X^+$ and $\sigma' \in W$.

We want to prove $\mu' \in \mathscr{S}_{\lambda,n}$ and $\mu' \uparrow\uparrow \mu$. Then we can apply the induction hypothesis to $\sigma' \cdot \mu' + p^n(\nu+\delta)$ and get the proposition.

We have

$$\begin{aligned}\sigma' \cdot \mu' + p^n(\nu+\delta) &= \sigma_i \cdot (\sigma \cdot \mu + p^n(\nu+\delta)) = \sigma_i\sigma \cdot \mu + p^n \sigma_i(\nu+\delta) \\ &= \sigma_i\sigma \cdot \mu + p^n(\nu+\delta) - p^n(\nu+\delta, \alpha_i^\vee)\alpha_i,\end{aligned}$$

hence $\sigma' \cdot \mu' = \sigma_i\sigma \cdot \mu - p^n(\nu+\delta, \alpha_i^\vee)\alpha_i$. By our choice of α_i we have $(\sigma \cdot \mu + \delta, \alpha_i^\vee) < -p^n(\nu+\delta, \alpha_i^\vee)$, hence $(\sigma_i\sigma \cdot \mu + \delta, \alpha_i^\vee) > p^n(\nu+\delta, \alpha_i^\vee) > 0$. Therefore (1.9) implies $\sigma' \cdot \mu' \uparrow\uparrow \sigma_i\sigma \cdot \mu$ and (1.10(2)) implies $\mu' \uparrow\uparrow \mu$ and therefore also $\mu' \in \mathscr{S}_{\lambda,n}$. This was what we had to show.

(2.7) **Corollary.** *Let $\lambda \in X_n$ and $\nu \in X^+$. Then $\operatorname{ch} \hat{Q}(\lambda, n) \otimes V(\nu)^{(p^n)}$ is a linear combination of the $\operatorname{ch}(\sigma \cdot \mu + p^n(\nu+\delta))$ with $\mu \in \mathscr{S}_{\lambda,n}$ and $\sigma \in W$ such that $\sigma \cdot \mu + p^n(\nu+\delta) \in X^+$. The coefficient of $\operatorname{ch}(\sigma \cdot \mu + p^n(\nu+\delta))$ for such μ, σ is equal to*

$$\sum_{\tau \in W} \det(\tau)[\hat{Z}(\sigma \cdot \mu + p^n(\nu+\delta), n) : \hat{M}(\lambda + p^n\tau \cdot \nu, n)].$$

Proof. By (2.2) and (2.3) we know that $\operatorname{ch} \hat{Q}(\lambda, n) \otimes V(\nu)^{(p^n)}$ is a linear combination of all $\operatorname{ch}(\sigma \cdot \mu + p^n(\nu+\delta))$ with $\mu \in \mathscr{S}_{\lambda,n}$ and $\sigma \in W$. If $\sigma \cdot \mu + p^n(\nu+\delta) \notin X^+$ and $\operatorname{ch}(\sigma \cdot \mu + p^n(\nu+\delta)) \neq 0$, then there is a unique $\tau \in W$ with $\tau \cdot (\sigma \cdot \mu + p^n(\nu+\delta)) \in X^+$. Then $\operatorname{ch}(\sigma \cdot \mu + p^n(\nu+\delta)) = \det(\tau) \operatorname{ch}(\tau \cdot (\sigma \cdot \mu + p^n(\nu+\delta)))$ and $\tau \cdot (\sigma \cdot \mu + p^n(\nu+\delta))$ has by (2.6) the required form. This proves that $\operatorname{ch} \hat{Q}(\lambda, n) \otimes V(\nu)^{(p^n)}$ is a linear combination as resired.

In (2.2) the coefficient of $\operatorname{ch}(\sigma \cdot \mu + p^n(\nu+\delta))$ is equal to

$$\begin{aligned}[\hat{Z}(\mu + p^n\delta, n) : \hat{M}(\lambda, n)] &= [Z(\sigma \cdot \mu + p^n\delta, n) : \hat{M}(\lambda, n)] \\ &= [\hat{Z}(\sigma \cdot \mu + p^n(\nu+\delta), n) : \hat{M}(\lambda + p^n\nu, n)].\end{aligned}$$

If $\sigma \cdot \mu + p^n(\nu+\delta) = \tau^{-1} \cdot (\sigma' \cdot \mu' + p^n(\nu+\delta))$ for $\mu' \in \mathscr{S}_{\lambda,n}$, $\tau, \sigma' \in W$, then this number is by [10], 3.2(2) equal to $[\hat{Z}(\sigma' \cdot \mu' + p^n(\nu+\delta), n) : \hat{M}(\lambda + p^n\tau \cdot \nu, n)]$. From this the claim about the coefficients in the corollary follows easily.

(2.8) **Proposition.** *Let ν be a dominant weight such that $(\nu + 2\delta, \alpha_0^\vee) \leqq p$. For $\lambda \in X_n$, $\mu \in \mathscr{S}_{\lambda,n}$ and $\sigma \in W$ with $\sigma \cdot \mu + p^n(\nu+\delta) \in X^+$ we have*

$$\begin{aligned}&[V(\sigma \cdot \mu + p^n(\nu+\delta)) : M(\lambda + p^n\nu)] \\ &\quad = \sum_{\tau \in W} \det(\tau)[\hat{Z}(\sigma \cdot \mu + p^n(\nu+\delta), n) : \hat{M}(\lambda + p^n\tau \cdot \nu, n)].\end{aligned}$$

Proof. By [10] 3.1(5) we have

$$\begin{aligned}&[V(\sigma \cdot \mu + p^n(\nu+\delta)) : M(\lambda + p^n\nu)] \\ &\quad = \sum_{\nu_1 \in X^+} \sum_{\tau \in W} \det(\tau)[\hat{Z}(\sigma \cdot \mu + p^n(\nu+\delta), n) : \hat{M}(\lambda + p^n\tau \cdot \nu_1, n)][V(\nu_1) : M(\nu)].\end{aligned}$$

We have to show that only $\nu_1 = \nu$ contributes to this sum. If we get something from $\nu_1 \neq \nu$, then $[V(\nu_1) : M(\nu)] \neq 0$, hence $\nu_1 = w \cdot \nu$ for some $w \in W_{a,p}$, $w \neq 1$. As $\nu \in \bar{C}_0$, we get $S_{\alpha_{0,1}} \cdot \nu \uparrow\uparrow \nu_1$, hence

$$p^n(S_{\alpha_{0,1}} \cdot \nu, \alpha_0^\vee) \leqq p^n(\nu_1, \alpha_0^\vee) \leqq (\lambda + p^n \nu_1, \alpha_0^\vee) \leqq (\tau^{-1} \cdot (\sigma \cdot \mu + p^n(\nu + \delta)), \alpha_0^\vee)$$
$$\leqq (\sigma \cdot \mu + p^n(\nu + \delta), \alpha_0^\vee) \leqq (\mu + p^n(\nu + \delta), \alpha_0^\vee)$$
$$\leqq (p^n(\nu + 2\delta) - \delta, \alpha_0^\vee) < p^n(\nu + 2\delta, \alpha_0^\vee).$$

This implies

$$(\nu + 2\delta, \alpha_0^\vee) > (S_{\alpha_{0,1}} \cdot \nu, \alpha_0^\vee) = (\sigma_{\alpha_0}(\nu + \delta) - \delta + p\alpha_0, \alpha_0^\vee) = 2p - (\nu + 2\delta, \alpha_0^\vee)$$

and $2p < 2(\nu + 2\delta, \alpha_0^\vee) \leqq 2p$ a contradiction.

(2.9) *Remarks.* (1) Comparing (2.8) and (2.7) we get a reciprocity for the Weyl filtration of $\hat{Q}(\lambda, n) \otimes V(\nu)^{(p^n)}$ generalizing that in [10], 5.9. (Note that the existence of ν as in (2.8) implies $p \geqq 2h - 2$.)

(2) We call the formula in (2.2) the pattern of a Weyl filtration of $Q(\lambda, n) \otimes V(\nu)^{(p^n)}$, briefly the filtration pattern of $Q(\lambda, n) \otimes V(\nu)^{(p^n)}$. It is obvious that the filtration pattern of $Q(\lambda, n) \otimes V(\nu)^{(p^n)}$ is only the translation of the filtration pattern of $Q(\lambda, n)$ by $p^n \nu$. The filtration pattern of $Q(\lambda, n)$ does not depend on the chosen λ but only on the facet to which λ belongs because of Jantzen's translation principle (cf. [6, 9]). For regular λ there are exactly $\frac{|W|}{f} p^{2(n-1)}$ filtration patterns.

(3) Because of (1) the filtration pattern is dual to Jantzen's generic decomposition pattern for Weyl modules (cf. [6]).

3. The Case $n = 1$

(3.1) **Lemma.** *For any $\beta \in \Phi^-$ there exist $\sigma \in W$ and $\alpha_k \in \Delta$ with $\sigma(\alpha_k) = \beta$ and $\sigma(\alpha_i) \in \Phi^+$ for all $i \neq k$, $1 \leqq i \leqq l$.*

Proof. Let $-\beta = \sum_{i=1}^{l} a_i \alpha_i$, where $a_i \in \mathbb{Z}^+$, $1 \leqq i \leqq l$. We use induction on $ht(-\beta) = \sum_{i=1}^{l} a_i$.

If $ht(-\beta) = 1$, then $\beta = -\alpha_k$ for some simple root α_k and we can take $\sigma = \sigma_k$. Suppose $ht(-\beta) > 1$. There exists $\alpha_j \in \Delta$ with $(\beta, \alpha_j^\vee) < 0$. Then $\beta' = \sigma_j(\beta) = \beta - (\beta, \alpha_j^\vee)\alpha_j \in \Phi^-$ and $ht(-\beta') < ht(-\beta)$. By the induction hypothesis there exist $\sigma' \in W$, $\alpha_k \in \Delta$ such that $\sigma'(\alpha_k) = \beta'$, $\sigma'(\alpha_i) \in \Phi^+$ for all $i \neq k$, $1 \leqq i \leqq l$. We take now $\sigma = \sigma_j \sigma'$, thus $\sigma(\alpha_k) = \sigma_j \sigma(\alpha_k) = \sigma_j(\beta') = \beta$. Consider $\alpha_i \in \Delta - \{\alpha_k\}$. In order to prove that $\sigma(\alpha_i) = \sigma_j \sigma'(\alpha_i) \in \Phi^+$, we have to prove that $\sigma'(\alpha_i) \neq \alpha_j$. If $\sigma'(\alpha_i) = \alpha_j$, then

$$\alpha_k = (\sigma')^{-1}(\beta') = (\sigma')^{-1}(\sigma_j(\beta)) = (\sigma')^{-1}(\beta - (\beta, \alpha_j^\vee)\alpha_j)$$
$$= (\sigma')^{-1}(\beta) - (\beta, \alpha_j^\vee)(\sigma')^{-1}(\alpha_j) = (\sigma')^{-1}(\beta) - (\beta, \alpha_j^\vee)\alpha_i$$

and $(\sigma')^{-1}(\beta) = \alpha_k + (\beta, \alpha_j^\vee)\alpha_j$. As $(\beta, \alpha_j^\vee) < 0$, this is a contradiction to the definition of a basis of a root system. Hence $\sigma'(\alpha_i) \neq \alpha_j$.

We call a weight $\mu' \in X^+$ maximal with respect to a weight $\mu \in X^+$ if $\mu' \uparrow\uparrow \mu$, $\mu \neq \mu'$, and if there is no $\mu'' \in X^+$ with $\mu' \uparrow\uparrow \mu'' \uparrow\uparrow \mu$ and $\mu'' \neq \mu, \mu'$. If $\mu \in C$, $\mu' \in C'$

for $C, C' \in \mathscr{C}$, then μ' is maximal with respect to μ if and only if $\mu' \uparrow\uparrow \mu$ and $d(C') = d(C) - 1$ (by [7], Corollary 2 to Satz 9).

We shall from now on restrict ourselves to the case of p-regular $\lambda \in X_1$, i.e. to λ inside an alcove. Then also $\lambda^0 - \delta$ and all μ with $\mu \in \mathscr{S}_{\lambda,1}$ are p-regular.

(3.2) **Lemma.** *Suppose $\lambda \in X_1$ p-regular and $\mu \in \mathscr{S}_{\lambda,1}$. Let C be the alcove with $\mu \in C$.*

(1) *If $C = C_0$ then $\sigma \cdot \mu + p(v+\delta) \in X^+$ for all $\sigma \in W$ and $v \in X^+$.*

(2) *Let $C \neq C_0$ and let $\mu' \in \mathscr{S}_{\lambda,1}$ be maximal with respect to μ. Then there exist $\sigma \in W$, $v \in X^+$ and a simple reflection σ_k such that $\sigma_k \cdot (\sigma \cdot \mu + p(v+\delta)) = \sigma \cdot \mu' + p(v+\delta) \in X^+$ and such that for all $\tau \in W$, $\tau \neq 1$, σ_k the weight $\tau \cdot (\sigma \cdot \mu' + p(v+\delta))$ is not of the form $\sigma'' \cdot \mu'' + p(v+\delta)$ with $\mu'' \in \mathscr{S}_{\lambda,1}$ and $\sigma'' \in W$.*

Proof. (1) As $\mu \in C_0$ we have $0 \leqq (\mu+\delta, \beta^\vee) \leqq (\mu+\delta, \alpha_0^\vee) < p$ for all $\beta \in \Phi^+$, hence $(\sigma \cdot \mu + \delta, \beta^\vee) > -p$ for all $\sigma \in W$ and

$$(\sigma \cdot \mu + p(v+\delta) + \delta, \beta^\vee) > p(\delta, \beta^\vee) - p \geqq 0.$$

This shows $\sigma \cdot \mu + p(v+\delta) \in X^+$.

(2) As μ' is maximal with respect to μ there are $\beta \in \Phi^+$ and $m \in \mathbb{Z}^+$ with $\mu' = S_{\beta,m} \cdot \mu$. Furthermore there has to be $a_\beta \in \mathbb{Z}$ with $0 < a_\beta < p$ and $(\mu+\delta, \beta^\vee) = mp + a_\beta$, hence $(\mu'+\delta, \beta^\vee) = mp - a_\beta$ and $\mu' = \mu - a_\beta \beta$. By (3.1) there exist $\sigma \in W$ and $\alpha_k \in \Delta$ such that $\sigma(\beta) = -\alpha_k$ and $\sigma^{-1}(\alpha_i) \in \Phi^+$ for all $i \neq k$, $1 \leqq i \leqq l$. Then $(\sigma \cdot \mu + \delta, \alpha_k^\vee) = -(\mu+\delta, \beta^\vee) = -(mp + a_\beta)$ and $(\sigma \cdot \mu' + \delta, \alpha_k^\vee) = -mp + a_\beta$ whereas $(\sigma \cdot \mu + \delta, \alpha_i^\vee) \geqq 0$ and $(\sigma \cdot \mu' + \delta, \alpha_i^\vee) \geqq 0$ for $i \neq k$, $1 \leqq i \leqq l$. Set $v = (m-1)\omega_{\alpha_k} + \sum_{i \neq k} m_i \omega_{\alpha_i}$ with suitable $m_i \in \mathbb{Z}^+$ to be determined later. Then $(\sigma \cdot \mu + p(v+\delta) + \delta, \alpha_k^\vee) = -a_\beta$ and

$$\begin{aligned}\sigma_k \cdot (\sigma \cdot \mu + p(v+\delta)) &= \sigma \cdot \mu + p(v+\delta) - (\sigma \cdot \mu + p(v+\delta) + \delta, \alpha_k^\vee)\alpha_k \\ &= \sigma \cdot \mu + p(v+\delta) + a_\beta \alpha_k = \sigma \cdot (\mu - a_\beta \beta) + p(v+\delta) \\ &= \sigma \cdot \mu' + p(v+\delta).\end{aligned}$$

For all $i \neq k$ we have $(\sigma \cdot \mu' + p(v+\delta) + \delta, \alpha_i^\vee) \geqq (\sigma \cdot \mu' + \delta, \alpha_i^\vee) \geqq 0$, hence $\sigma \cdot \mu' + p(v+\delta) \in X^+$. The set $\{\sigma'' \cdot \mu'' \mid \mu'' \in \mathscr{S}_{\lambda,1}, \sigma'' \in W\}$ is finite. Therefore we can choose the m_i in the definition of v such that $(\sigma'' \cdot \mu'' + p(v+\delta) + \delta, \gamma^\vee) > 0$ for all $\gamma \in \Phi^+$, $\gamma \neq \alpha_k$ and all these $\sigma'' \cdot \mu''$. Then either $\sigma'' \cdot \mu'' + p(v+\delta) \in X^+$ or $\sigma_k \cdot (\sigma'' \cdot \mu'' + p(v+\delta)) \in X^+$. From this the last claim of the lemma follows.

(3.3) **Theorem.** *Suppose $p \geqq 2h-2$. Let $\lambda \in X_1$ be p-regular. Then we have*

(1) $b^\lambda_{\lambda^0 - \delta} = 1$.

(2) *If $\mu, \mu' \in \mathscr{S}_{\lambda,1}$ with $\mu' \uparrow\uparrow \mu$, then $b^\lambda_\mu \leqq b^\lambda_{\mu'}$.*

Proof. (1) By definition

$$b^\lambda_{\lambda^0 - \delta} = [\hat{Z}(\lambda^0 + (p^n - 1)\delta, 1) : \hat{M}(\lambda, 1)] = [\hat{Z}(\sigma_0 \cdot \lambda + 2p^n \delta, 1) : \hat{M}(\lambda, 1)].$$

This number is equal to 1 by [10], Bemerkung 1 in 3.3.

(2) We may assume that μ' is maximal with respect to μ. Choose v, σ as in (3.2(2)). We can express $\mathrm{ch}\, Q(\lambda, 1) \otimes V(v)^{(p)}$ as linear combination of the $\mathrm{ch}(\mu_1)$

with $\mu_1 \in X^+$. The coefficient of $\mathrm{ch}(\sigma \cdot \mu' + p(\nu+\delta))$ is by (3.2(2)) and by (2.7) equal to $b^{\lambda}_{\mu'} - b^{\lambda}_{\mu}$. As $p \geqq 2h-2$ we know by (2.1) that $Q(\lambda, 1) \equiv V(\nu)^{(p)}$ has a Weyl filtration. Therefore the coefficient of each $\mathrm{ch}(\mu_1)$ has to be non-negative. This implies $b^{\lambda}_{\mu'} - b^{\lambda}_{\mu} \geqq 0$.

Remark. Observe that (1) holds not only for $n=1$ but for arbitrary n.

(3.4) **Corollary.** *Suppose $p \geqq 2h-2$. Let $\lambda \in X_1$ be p-regular. Then $b^{\lambda}_{\mu} \geqq 1$ for all $\mu \in \mathscr{S}_{\lambda,1}$.*

This is obvious as $\mu \uparrow\uparrow \lambda^0 - \delta$.

(3.5) **Corollary.** *Suppose $p \geqq 2h-2$. Let $\lambda \in X_1$ be p-regular. Then*

$$\{\lambda' \in X \mid [\hat{Z}(\lambda', 1) : \hat{M}(\lambda, 1)] \neq 0\} = \{\sigma \cdot \mu + p\delta \mid \mu \in \mathscr{S}_{\lambda,1}, \sigma \in W\}.$$

This is now clear by [10], 3.2, by the definition of b^{λ}_{μ} and by (2.3).

(3.6) *Remarks.* (1) For $n>1$ Corollary 3.4 does not hold in general. For G of type A_2, for $n=2$ and $p=5>2h-2=4$ one has

$$\mathscr{S}_{\lambda,2} = \{(1.12), (3.8), (6.2), (2.7), (3.4), (1.2)\} \quad \text{for } \lambda = (12.23).$$

One can show $b^{\lambda}_{(3.4)} = b^{\lambda}_{(1.2)} = 0$.

(2) Let $\lambda = \lambda_0 + p\lambda_1 + \ldots + p^{n-1}\lambda_{n-1}$ be the p-adic expansion of some $\lambda \in X_n$. Then we have an isomorphism of G-modules

$$Q(\lambda, n) \simeq Q(\lambda_0, 1) \otimes Q(\lambda_1, 1)^{(p)} \otimes \ldots \otimes Q(\lambda_{n-1}, 1)^{(p^{n-1})}.$$

We can obtain the filtration pattern of $Q(\lambda, n)$ from those of the $Q(\lambda_i, 1)$ by decomposing tensor products.

(3.7) **Proposition.** *Let $\lambda \in X_n$ be p-regular, let $C \in \mathscr{C}$ be the alcove with $\lambda \in C$. Let $s \in W_{a,p}$ be a reflection with respect to a facet of the alcove C such that $C \underset{\bar{A}}{\leqq} s \cdot C$. Let $\mu \in \mathscr{S}_{\lambda,n}$. Then there is $w \in W_{a,p}$ with $\mu + p^n\delta = w \cdot \lambda$, set $\mu' = ws \cdot \lambda - p^n\delta$. Then we have that if $\mu' \in X^+$, then $b^{\lambda}_{\mu'} = b^{\lambda}_{\mu}$.*

Proof. We have to show

$$[\hat{Z}(\mu + p^n\delta, n) : \hat{M}(\lambda, n)] = [\hat{Z}(\mu' + p^n\delta, n) : \hat{M}(\lambda, n)],$$

i.e.

$$[\hat{Z}(w \cdot \lambda, n) : \hat{M}(\lambda, n)] = [\hat{Z}(ws \cdot \lambda, n) : \hat{M}(\lambda, n)].$$

Applying the translations functors for the category of $\mathscr{U}_n - T$-modules (cf. [10] 5.2) this follows in the same way as the corresponding statement for Weyl modules in, say [9], § 3, Theorem (b).

(3.8) **Proposition.** *Let $\lambda \in X_n$ be p-regular. Suppose $\mu \in \mathscr{S}_{\lambda,n}$ is maximal with respect to $\lambda^0 - \delta$. Then $b^{\lambda}_{\mu} = 1$.*

Proof. We have to show $[\hat{Z}(\mu + p^n\delta, n) : \hat{M}(\lambda, n) = 1$ or, equivalently, by [10], 3.2, $[\hat{Z}(\sigma_0 \cdot \mu + p^n\delta, n) : \hat{M}(\lambda, n)] = 1$ or, equivalently, for some $\nu \in X$

$$[\hat{Z}(\sigma_0 \cdot \mu + p^n(\nu+\delta), n) : \hat{M}(\lambda + p^n\nu, n)] = 1.$$

As $\mu \uparrow\uparrow \lambda^0 - \delta$ and $\mu \neq \lambda^0 - \delta$ is maximal with this property we have also $\sigma_0 \cdot (\lambda^0 - \delta) \uparrow \sigma_0 \cdot \mu$ and maximality of $\sigma_0 \cdot (\lambda^0 - \delta) = \lambda - p^n \delta$ for this property. Hence we see for all $\nu \in X$ that $\lambda + p^n \nu \uparrow \sigma_0 \cdot \mu + p^n(\nu + \delta)$, that $\lambda + p^n \nu \neq \sigma_0 \cdot \mu + p^n(\nu + \delta)$ and that $\lambda + p^n \nu$ is maximal with this property. If we choose $\nu \in X^+$ sufficiently large such that $[\hat{Z}(\sigma_0 \cdot \mu + p^n(\nu + \delta)), n) : \hat{M}(\lambda', n)] \neq 0$ implies $\lambda' \in X^+$, then $\sigma_0 \cdot \mu + p^n(\nu + \delta) \in X^+$ and $[V(\sigma_0 \cdot \mu + p^n(\nu + \delta)) : M(\lambda + p^n \nu)] = 1$ by [7], Satz 10. Using [10], 3.1 (5) and Satz 3.3 one gets the desired equality.

Remark. The following results about $\mathrm{ch}\, Q(\lambda, n)$ are well known, and can now easily be proved. Set $C^0 = \sigma_0 \cdot C_0 + p\delta$. Then

(1) If $\lambda \in C^0 \cap X_1$, then $\mathrm{ch}\, Q(\lambda, 1) = \sum_{\sigma \in W} \mathrm{ch}(\sigma \cdot (\lambda^0 - \delta) + p\delta)$.

(2) *If* $\lambda \in S_{\alpha_0, h-2} \cdot C^0 \cap X_1$, then

$$\mathrm{ch}\, Q(\lambda, 1) = \sum_{\sigma \in W_1} (\mathrm{ch}(\sigma \cdot (\lambda^0 - \delta) + p\delta) + \mathrm{ch}(\sigma \cdot \mu + p\delta))$$

where μ is the unique element in $\mathscr{S}_{\lambda, 1}$ with $\mu \neq \lambda^0 - \delta$. Here $W_1 = \{\sigma \in W \mid \sigma(\alpha_0) \notin -\Delta\}$.

References

1. Andersen, H.H.: The strong linkage principle. J. Reine Angew. Math. **315**, 53–59 (1980)
2. Ballard, J.W.: Injective modules for restricted enveloping algebras. Math. Z. **163**, 57–63 (1978)
3. Humphreys, J.E.: Introduction to Lie algebras and representation theory, GTM 9. Berlin-Heidelberg-New York: Springer 1972
4. Humphreys, J.E.: Linear algebraic groups, GTM 21. Berlin-Heidelberg-New York: Springer 1975
5. Humphreys, J.E.: Ordinary and modular representation of Chevalley groups. Lecture Notes in Math. **528**. Berlin-Heidelberg-New York: Springer 1976
6. Jantzen, J.C.: Über das Dekompositionsverhalten gewisser modularer Darstellungen halbeinfacher Gruppen und ihrer Lie-Algebren. J. Algebra **49**, 441–469 (1977)
7. Jantzen, J.C.: Darstellungen halbeinfacher Gruppen und Kontravariante Formen. J. Reine Angew. Math. **290**, 117–141 (1977)
8. Jantzen, J.C.: Über Darstellungen höherer Frobenius-Kerne halbeinfacher algebraischer Gruppen. Math. Z. **164**, 271–292 (1979)
9. Jantzen, J.C.: Weyl modules for group of Lie type, in: Finite simple groups II, pp. 291–300. London: Academic Press 1980
10. Jantzen, J.C.: Darstellungen halbeinfacher Gruppen und ihrer Frobenius-Kerne. J. Reine Angew. Math. **317**, 157–199 (1980)
11. Jantzen, J.C.: Zur Reduction modulo p unipotenter Charaktere endlicher Chevalley-Gruppen. Math. Z. **181**, 97–128 (1982)
12. Lusztig, G.: Hecke algebras and Jantzen's generic decomposition patterns. Adv. Math. **37**, 121–164 (1980)
13. Verma, D.N.: The role of the affine Weyl groups in the representation theory of algebraic Chevalley groups and their Lie algebras, in I. M. Gelfand (ed.), Lie group and their representations, pp. 653–705. New York: Halsted 1975
14. Wang, J.-p.: Sheaf cohomology on G/B and tensor products of Weyl modules. J. Algebra **77**, 162–185 (1982)

Received November 15, 1984

Math. Z. 189, 529–554 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Die Existenz einer Lévy-Abbildung auf einem homogenen Raum

Hansmartin Zeuner

Mathematisches Institut der Universität, Auf der Morgenstelle 10, D-7400 Tübingen, Bundesrepublik Deutschland

1. Einleitung

Ein wichtiges Hilfsmittel beim Studium von Faltungshalbgruppen ist die Lévy-Hinčin-Formel. Diese Formel setzt in einem gewissen Sinn jede Faltungshalbgruppe aus drei Grundbausteinen zusammen, nämlich dem Translationsanteil, dem Gauß-Anteil und dem Lévy-Anteil. Da bekanntlich eine eineindeutige Beziehung zwischen einer Faltungshalbgruppe und ihrem infinitesimalen Erzeuger, einer dissipativen Distribution, besteht (vgl. [3], Prop. 3 und 4), beschränken wir uns in dieser Arbeit im wesentlichen auf die Betrachtung dissipativer Distributionen. Unter einer Lévy-Hinčin-Formel verstehen wir damit eine Zerlegung einer dissipativen Distribution in die Summe aus einem primitiven, einem quadratischen und einem durch das Lévy-Maß bestimmten Anteil.

Diese Zerlegung ist im Fall einer Faltungshalbgruppe $(\mu_t: t \geqq 0)$ von Wahrscheinlichkeitsmaßen mit Idempotent $\mu_0 = \varepsilon_1$ vollständig dargestellt in [13]. Es lag nahe, zu fragen, ob eine Lévy-Hinčin-Formel auch in den Fällen existiert, wo $(\mu_t: t \geqq 0)$ aus komplexen Maßen der Norm $\leqq 1$ besteht, beziehungsweise, wo für das Idempotent die Einschränkung $\mu_0 = \varepsilon_1$ fallengelassen wird. Eine Antwort im ersten Fall mit der Einschränkung $\mu_0 = \varepsilon_1$ wird in [15] gegeben.

Im zweiten Fall ist das Idempotent μ_0 von der Form $\mu_0 = \chi \omega_K$ (wobei ω_K das Haar-Maß auf einer kompakten Untergruppe K von G und χ ein stetiger Homomorphismus von K in $\mathbb{C}\setminus\{0\}$ ist). Im Spezialfall des Idempotents $\mu_0 = \omega_K$ kann man die Faltungshalbgruppe interpretieren als Faltungshalbgruppe auf dem homogenen Raum G/K, invariant bez. der Multiplikation von links mit Elementen aus K, vgl. dazu [8], 1.4.2. In einigen Spezialfällen, insbesondere für viele Gelfand-Paare (G, K), ist eine Lévy-Hinčin-Formel schon bekannt, vgl. dazu die Paragraphen 7.1, 7.2 und 7.3 in [8] sowie [10].

Das Hauptproblem im allgemeinen Fall eines beliebigen Idempotents besteht in der Existenz einer Lévy-Abbildung, die das entscheidende Hilfsmittel ist, um aus einem Lévy-Maß eine dissipative Distribution zu gewinnen. Zusätzlich zu den in [13] geforderten Eigenschaften muß nun noch die Invarianz der Lévy-Abbildung unter der Faltung mit dem Idempotent gefordert werden, was

eine geringfügige Modifikation der Definition aus [13] und eine Verfeinerung der Beweismethoden notwendig macht.

In dieser Arbeit wird nun die Existenz einer Lévy-Abbildung im Fall eines beliebigen Idempotents bewiesen. Dieses Ergebnis wird dazu verwendet, um im Spezialfall einer Faltungshalbgruppe von Wahrscheinlichkeitsmaßen mit Idempotent ω_K eine Lévy-Hinčin-Formel zu beweisen. Der allgemeine Fall von komplexen Maßen der Norm $\leqq 1$ wird unter Benutzung des hier Bewiesenen in einer späteren Arbeit abgehandelt.

Der Verfasser möchte Herrn Professor Dr. H. Heyer für die Beratung bei der Anfertigung dieser Arbeit herzlich danken.

2. Bezeichnungen

Es sei G eine lokalkompakte Gruppe mit neutralem Element 1. Mit G_1 bezeichnen wir die Zusammenhangskomponente der 1 und ω_G sei ein linkes Haar-Maß. Wenn G kompakt ist, sei ω_G normiert. Für jedes $g \in G$ seien σ_g bzw. τ_g die Abbildungen $x \to xg$ bzw. $x \to gx$ von G in sich. Ist f eine Funktion auf G und $g \in G$, so definiert man die Funktionen $\tilde{f}$, $\sigma_g(f)$ bzw. $\tau_g(f)$ auf G durch $\tilde{f}(x) := f(x^{-1})$ für alle $x \in G$ und $\sigma_g(f) := f \circ \sigma_{g^{-1}}$ bzw. $\tau_g(f) := f \circ \tau_{g^{-1}}$. $\mathscr{C}(G)$ bezeichne die Menge aller stetigen komplexwertigen Funktionen auf G. Den Träger von $f \in \mathscr{C}(G)$ definiert man als $\operatorname{supp}(f) := \overline{\{g \in G : f(g) \neq 0\}} \subset G$. $\mathscr{K}(G)$ sei die Menge aller $f \in \mathscr{C}(G)$ mit kompaktem Träger. Für jede beschränkte Funktion f auf G setzen wir $\|f\|_\infty := \sup_{x \in G} |f(x)|$. Mit $\mathscr{E}(G)$ bezeichnen wir den topologischen Vektorraum der komplexwertigen regulären Funktionen auf G und setzen $\mathscr{D}(G) := \mathscr{E}(G) \cap \mathscr{K}(G)$; beide Funktionenräume werden in [2] mit Hilfe der Lösung von Hilberts 5. Problem definiert.

Seien nun und im folgenden K eine kompakte Untergruppe von G und χ ein Charakter auf K, d.h. ein stetiger Homomorphismus von K in $\mathbb{T}$, die Gruppe der komplexen Zahlen vom Betrag 1. Ist $\mathscr{A}(G)$ ein Raum von Funktionen auf G, so seien $\mathscr{A}_{\mathbb{R}}(G)$ bzw. $\mathscr{A}^b(G)$ bzw. $\mathscr{A}_+(G)$ die Teilmengen der reellwertigen bzw. beschränkten bzw. nicht negativen Funktionen aus $\mathscr{A}(G)$ und

$$\mathscr{A}(G; \chi) := \{f \in \mathscr{A}(G) : \tau_k(f) = \sigma_k(f) = \chi(k) f \text{ für alle } k \in K\}.$$

In dem Fall, daß χ der triviale Charakter 1_K (d.h. die konstante Funktion mit Wert 1 auf K) ist, gelte die Abkürzung $\mathscr{A}(G; K) := \mathscr{A}(G; 1_K)$.

Der (topologische) Dual von $\mathscr{D}(G)$, der Raum $\mathscr{D}'(G)$ aller Distributionen auf G, umfaßt in kanonischer Weise die Menge $\mathscr{M}(G)$ der komplexen Radon-Maße auf G. Mit $\mathscr{M}^b(G)$ bzw. $\mathscr{M}_+(G)$ bzw. $\mathscr{M}^1(G)$ bezeichnen wir die Mengen der beschränkten bzw. der positiven bzw. der Wahrscheinlichkeitsmaße. Ähnlich wie für Maße kann man auch für beliebige Distributionen $T \in \mathscr{D}'(G)$ den Träger $\operatorname{supp}(T)$ definieren und für zwei Distributionen, von denen mindestens eine kompakten Träger besitzt, das Faltungsprodukt $T_0 * T_1$ (vgl. [8], Seite 6).

Für $T \in \mathscr{D}'(G)$ und $g \in G$ seien $\sigma_g(T)$, $\tau_g(T)$, $\check{T}$, $\bar{T}$ bzw. T^* die Distributionen $T \circ \sigma_{g^{-1}}$, $T \circ \tau_{g^{-1}}$, $f \to T(\tilde{f})$, $f \to \overline{T(\bar{f})}$ bzw. $\bar{\check{T}}$. Ähnlich wie oben setzen wir $\mathscr{D}'(G; \chi) := \{T \in \mathscr{D}'(G) : \chi \omega_K * T * \chi \omega_K = T\}$ mit der Abkürzung $\mathscr{D}'(G; K) := \mathscr{D}'(G; 1_K)$.

3. Dissipative Distributionen und Faltungshalbgruppen

3.1 Definition. Eine Distribution $T\in\mathscr{D}'(G;\chi)$ heißt *χ-dissipativ*, falls

$$\operatorname{Re} T(f)\leqq 0 \quad \text{für alle } f\in\mathscr{D}(G;\chi) \quad \text{mit } f(1)=\|f\|_\infty$$

gilt.

3.2. *Bemerkung.* Nach [3], Lemme 1 läßt sich jede χ-dissipative Distribution T zerlegen als Summe einer Distribution T_0 mit kompaktem Träger und eines Maßes $\mu\in\mathscr{M}^b(G)$. Damit läßt sich T auf $\mathscr{E}^b(G)$ fortsetzen durch

$$T(f):=T_0(f)+\int f\,d\mu \quad \text{für alle } f\in\mathscr{E}^b(G).$$

3.3. Definition. Eine χ-dissipative Distribution T heißt *normiert*, falls $T+\varepsilon\cdot\chi\,\omega_K$ für kein $\varepsilon>0$ χ-dissipativ ist. Eine 1_K-dissipative Distribution heißt *positiv normiert*, falls $T(1_G)=0$.

3.4. Eine Distribution T ist genau dann positiv normiert und 1_K-dissipativ, wenn sie normiert und ein „K-laplacien généralisé“ im Sinne von [3] ist, das heißt reell ist und

$$T(f)\leqq 0 \quad \text{für alle } f\in\mathscr{D}_{\mathbb{R}}(G;K) \quad \text{mit } f(1)\geqq 0 \quad \text{und} \quad f(1)\geqq f \quad \text{erfüllt.}$$

Beweis. Sei T positiv normiert und 1_K-dissipativ. Zuerst wird gezeigt, daß T reell ist, das heißt, daß $T(f)\in\mathbb{R}$ für alle $f\in\mathscr{D}_{\mathbb{R}}(G)$ gilt. Sei zunächst $f\in\mathscr{E}^b_{\mathbb{R}}(G;K)$ und $f(1)=0$. Für die Funktionen

$$\phi:=(\|f\|_\infty+1)\,1_G-\sqrt{(\|f\|_\infty+1)^2\,1_G-f^2}\in\mathscr{E}^b_{\mathbb{R}}(G;K)$$

und

$$f_n^{\pm}:=n(\|f\|_\infty+1)\,1_G-\frac{1}{n}\phi\pm if\in\mathscr{E}^b(G;K) \qquad (n\geqq 1)$$

gilt dann

$$\phi(1)=0\leqq\phi\leqq(\|f\|_\infty+1)\,1_G \quad \text{und} \quad f_n^{\pm}(1)=\|f_n^{\pm}\|_\infty.$$

Daraus folgt

$$0\geqq\operatorname{Re} T(f_n^{\pm})=-\frac{1}{n}\operatorname{Re} T(\phi)\mp\operatorname{Im} T(f),$$

das heißt

$$|\operatorname{Im} T(f)|\leqq\frac{1}{n}\operatorname{Re} T(\phi) \qquad \text{für alle } n\geqq 1.$$

Dies ist nur möglich, wenn

$$\operatorname{Im} T(f)=0 \qquad \text{für alle } f\in\mathscr{E}^b_{\mathbb{R}}(G;K) \text{ mit } f(1)=0.$$

Sei nun $f\in\mathscr{D}_{\mathbb{R}}(G;K)$ beliebig. Für $f':=f-f(1)\,1_G\in\mathscr{E}^b_{\mathbb{R}}(G;K)$ gilt dann $f'(1)=0$, nach obigem also $T(f)=T(f')\in\mathbb{R}$. Damit ist bewiesen, daß $T\in\mathscr{D}'_{\mathbb{R}}(G;K)$.

Sei nun $f\in\mathscr{D}_{\mathbb{R}}(G;K)$ und $f(1)\geqq f$, $f(1)\geqq 0$. Für $f'':=f+\|f\|_\infty\,1_G\in\mathscr{E}^b_{\mathbb{R}}(G;K)$ gilt dann $f''(1)=\|f\|_\infty$, woraus $T(f)=\operatorname{Re} T(f'')\leqq 0$ folgt. Dies beweist, daß T ein „K-laplacien généralisé“ ist.

Wenn umgekehrt T reell und $T(f)\leqq 0$ für alle $f\in\mathscr{D}_{\mathbb{R}}(G;K)$ mit $f(1)\geqq f$ und $f(1)\geqq 0$, so gilt für jede Funktion $f'\in\mathscr{D}(G;K)$ mit $f'(1)=\|f'\|_\infty$ $\operatorname{Re} f'(1)\geqq\operatorname{Re} f'$ und $\operatorname{Re} f'(1)\geqq 0$ und damit $\operatorname{Re} T(f')=T(\operatorname{Re} f')\leqq 0$. Also ist T 1_K-dissipativ.

Sei nun T außerdem normiert. Wir setzen $\varepsilon := -T(1_G) \geqq 0$. Dann gilt für jede Funktion $f \in \mathscr{D}_{\mathbb{R}}(G; K)$ mit $f(1) \geqq f$ und $f(1) \geqq 0$

$$(T + \varepsilon \cdot \omega_K)(f) = T(f) - f(1)\, T(1_G) = T(f - f(1)\, 1_G) \leqq 0,$$

denn

$$0 = (f - f(1)\, 1_G)(1) \geqq f - f(1)\, 1_G.$$

Nach dem obigen Schluß ist also $T + \varepsilon \cdot \omega_K 1_K$-dissipativ. Da T normiert ist, folgt $T(1_G) = \varepsilon = 0$. T ist also positiv normiert.

3.5. Definition. Unter einer *Faltungshalbgruppe* verstehen wir einen stetigen Homomorphismus $t \to \mu_t$ von $(\mathbb{R}_+, +)$ in die Faltungsalgebra $\mathscr{M}^b(G)$ (die mit der schwachen Topologie versehen ist), derart daß $\|\mu_t\| \leqq 1$ für alle $t \geqq 0$.

3.6. Bekanntlich gibt es eine kompakte Untergruppe K und einen Charakter χ auf K mit $\mu_0 = \chi\, \omega_K$ (vgl. [6], Kapitel IV Satz 5.2). Die Menge aller Faltungshalbgruppen $(\mu_t : t \geqq 0)$ mit Idempotent $\mu_0 = \chi\, \omega_K$ steht mit der Menge aller χ-dissipativen Distributionen $T \in \mathscr{D}'(G; \chi)$ vermöge

$$T(f) = \lim_{t \to 0} \frac{1}{t} (\mu_t(f) - \mu_0(f)) \qquad \text{für alle } f \in \mathscr{D}(G)$$

in eineindeutiger Beziehung. Genau dann ist $(\mu_t : t \geqq 0)$ eine Faltungsgruppe von Wahrscheinlichkeitsmaßen auf G, falls $\chi = 1_K$ und T positiv normiert ist (vgl. [3], Prop. 3 und 4).

4. Primitive Distributionen

Die in diesem Paragraphen behandelten primitiven Distributionen sind aus zwei Gründen interessant: einerseits sind sie das einfachste Beispiel für dissipative Distributionen und bilden einen wichtigen Bestandteil der Lévy-Hinčin-Formel, andererseits sind sie ein entscheidendes Hilfsmittel bei der Konstruktion einer Lévy-Abbildung.

4.1. Definition. Eine Distribution $T \in \mathscr{D}'(G; \chi)$ heißt *χ-primitiv*, falls T und $-T$ χ-dissipativ sind. Die Menge der χ-primitiven Distributionen wird mit $\mathscr{P}(G; \chi)$ bezeichnet; es gelte die Abkürzung $\mathscr{P}(G; K) := \mathscr{P}(G; 1_K)$. Die Elemente der Menge $\mathscr{P}_{\mathbb{R}}(G; K)$ der reellen 1_K-primitiven Distributionen heißen *echt K-primitiv*.

4.2. $\mathscr{P}(G; \chi)$ ist ein reeller Vektorraum.

4.3. Definition. Eine Distribution $T \in \mathscr{D}'(G; \chi)$ heißt *lokal*, falls $\operatorname{supp}(T) \subset K$ gilt.

4.4. *Bemerkung.* Für jede Faltungshalbgruppe $(\mu_t : t \geqq 0)$, deren erzeugende Distribution T lokal ist, gilt

$$\operatorname{supp}(\mu_t) \subset K \cdot G_1 \qquad \text{für alle } t \geqq 0.$$

Beweis. Man sieht leicht, daß $K \cdot G_1 = \bigcap \{H \leqq G : H \text{ offen}, K \subset H\}$. Es genügt also zu zeigen, daß

$\operatorname{supp}(\mu_t) \subset H$ für alle $t \geqq 0$ und alle offenen Untergruppen H von G, die K enthalten.
Um dies zu beweisen, betrachten wir $f \in \mathscr{D}(G; \chi)$ mit $f(1) = \sup_{h \in H} |f(h)|$. Da T lokal ist, folgt daraus aber $\operatorname{Re} T(f) \leqq 0$ und mit [3], Prop. 6 schließt man $\operatorname{supp}(\mu_t) \subset H$ für alle $t \geqq 0$. □

4.5. Jede χ-primitive Distribution is lokal.
[[Sei nämlich $f \in \mathscr{D}(G; \chi)$ mit $\operatorname{supp}(f) \cap K = \varnothing$ gegeben. Dann gibt es ein $f_0 \in \mathscr{D}(G; \chi)$ mit $f_0(1) = \|f_0\|_\infty > \|f\|_\infty$ und $\operatorname{supp}(f_0) \subset \complement \operatorname{supp}(f)$. Da T χ-primitiv ist, gilt $\operatorname{Re} T(f_0) = 0$, und für alle $z \in \mathbb{T}$ hat man $\operatorname{Re}(zT(f)) = \operatorname{Re}(T(f_0 + zf)) = 0$. Daraus folgt $T(f) = 0$.]]

4.6. Definition. Eine *Faltungsgruppe* auf G ist ein schwach stetiger Homomorphismus $t \to \mu_t$ von $(\mathbb{R}, +)$ in $\mathscr{M}^b(G)$ mit $\|\mu_t\| \leqq 1$ für alle $t \in \mathbb{R}$.

4.7. Eine χ-dissipative Distribution $T \in \mathscr{D}'(G)$ ist genau dann χ-primitiv, wenn eine Faltungsgruppe $(\mu_t : t \in \mathbb{R})$ auf G mit Idempotent $\chi \omega_K$ existiert, sodaß $(\mu_t : t \geqq 0)$ die von T erzeugte Faltungshalbgruppe ist.
[[Wie in [3] konstruiert man eine Operatorhalbgruppe auf einem geeigneten Banach-Raum und wendet aus Kapitel IX von [14] die Theoreme in § 7 und § 9 an.]]

Um eine vollständige Darstellung aller χ-primitiver Distributionen zu erhalten, kann man sich nun mit Hilfe von [12], II.2.4 auf den Fall von positiv normierten Distributionen zurückziehen. Dazu müssen einige Bezeichnungen eingeführt werden.

4.8. *Bezeichnungen.* Seien K eine Untergruppe von G und χ ein Charakter auf K. Von nun an bezeichne π_K die kanonische Projektion von G auf G/K und K' die kompakte Untergruppe $\{(k, \chi(k)) : k \in K\}$ von $G \times \mathbb{T}$. Für jede Funktion $f \in \mathscr{C}(G)$ sei Zf die Funktion

$(g, z) \to zf(g)$ auf $G \times \mathbb{T}$.

Es gilt $Zf \in \mathscr{K}(G \times \mathbb{T})$, wenn $f \in \mathscr{K}(G)$,

$Zf \in \mathscr{D}(G \times \mathbb{T})$, wenn $f \in \mathscr{D}(G)$, sowie

$Zf \in \mathscr{C}(G \in \mathbb{T}; K')$, falls $f \in \mathscr{C}(G; \chi)$.

Man kann daher für jedes $\mu \in \mathscr{M}(G \times \mathbb{T})$ ein Maß $Z\mu \in \mathscr{M}(G)$ definieren durch $Z\mu(f) := \mu(Zf)$ für alle $f \in K(G)$. Ebenso sei für $T \in \mathscr{D}'(G \times \mathbb{T})$ die Distribution $ZT \in \mathscr{D}'(G)$ erklärt durch $ZT(f) := T(Zf)$ für alle $f \in \mathscr{D}(G)$. Aus $T \in \mathscr{D}'(G \times \mathbb{T}; K')$ folgt $ZT \in \mathscr{D}'(G; \chi)$.

4.9. Hilfssatz. *Sei $(\mu_t : t \in \mathbb{R})$ eine Faltungsgruppe auf G mit Idempotent $\chi \omega_K$. Dann gibt es genau eine Faltungsgruppe $(\nu_t : t \in \mathbb{R})$ von Wahrscheinlichkeitsmaßen auf $G \times \mathbb{T}$ mit $Z\nu_t = \mu_t$ für alle $t \in \mathbb{R}$.*

Es gilt $\nu_0 = \omega_{K'}$.

Beweis. Wegen $\mu_t * \mu_{-t} = \mu_0 = \chi \omega_K$ gilt $\|\mu_t\| = 1$ für alle $t \in \mathbb{R}$. Nach [12], Proposition II.1.1 gibt es eindeutig bestimmte Wahrscheinlichkeitsmaße ν_t ($t \in \mathbb{R}$) auf $G \times \mathbb{T}$ mit

$$Z\nu_t = \mu_t \quad \text{für alle } t \in \mathbb{R}.$$

Die nach [12], Corollaire II.2.5 konstruierten Faltungshalbgruppen $(\nu_t' : t \geqq 0)$ und $(\nu_t'' : t \geqq 0)$ von Wahrscheinlichkeitsmaßen auf $G \times \mathbb{T}$ mit

$$Z\nu_t' = \mu_t \quad \text{und} \quad Z\nu_t'' = \mu_{-t} \quad \text{für alle } t \geqq 0$$

stimmen wegen der Eindeutigkeitsaussage aus [12], II.1.1 und $\|\mu_t\| = 1$ für alle $t \in \mathbb{R}$ mit $(\nu_t : t \geqq 0)$ und $(\nu_{-t} : t \geqq 0)$ überein. Aus [12], II.2.2 Lemme 1 folgt $\nu_t * \nu_s = \nu_{t+s}$ für alle $s, t \in \mathbb{R}$. Deshalb ist $(\nu_t : t \geqq 0)$ eine Faltungsgruppe mit Idempotent $\omega_{K'}$. □

4.10. Satz. *$(\mu_t : t \in \mathbb{R})$ ist genau dann eine Faltungsgruppe mit $\mu_0 = \chi\omega_K$, falls eine Einparameteruntergruppe $(g_t : t \in \mathbb{R})$ von G mit*

$$g_t k g_{-t} k^{-1} \in \operatorname{Ker}\chi \quad \textit{für alle } t \in \mathbb{R} \quad \textit{und} \quad k \in K$$

und ein $u \in \mathbb{R}$ existieren, derart, daß

$$\mu_t = e^{iut} \cdot \tau_{g_t}(\chi\omega_K) \quad \textit{für alle } t \in \mathbb{R}.$$

Beweis. Daß jede derartige Familie $(\mu_t : t \in \mathbb{R})$ eine Faltungsgruppe ist, rechnet man leicht nach. Den Beweis dafür, daß jede Faltungsgruppe $(\mu_t : t \in \mathbb{R})$ von der angegebenen Form ist, führen wir in zwei Schritten.

I. Seien zunächst $\chi = 1_K$ und $\mu_t \in \mathcal{M}^1(G)$ für alle $t \in \mathbb{R}$. Wir setzen $M_t := \operatorname{supp}(\mu_t)$ und wählen $a_t \in M_t$. Dann gilt $\overline{M_s M_t} = M_{s+t}$ für alle $s, t \in \mathbb{R}$, $a_t K \subset \overline{M_t M_0} = M_t$ und $a_{-t} M_t \subset \overline{M_{-t} M_t} = M_0 = K$ für alle $t \in \mathbb{R}$, folglich $a_t K \subset M_t \subset a_{-t}^{-1} K$ und schließlich

$$a_t K = M_t = a_{-t}^{-1} K \quad \text{für alle } t \in \mathbb{R}.$$

Ebenso sieht man

$$K a_t = M_t = K a_{-t}^{-1} \quad \text{für alle } t \in \mathbb{R}.$$

Aus $a_t K = K a_t$ folgt, daß a_t im Normalisator $N(K)$ von K ist, also

$$\operatorname{supp}(\mu_t) \subset \mathrm{N(K)} \quad \text{für alle } t \in \mathbb{R}.$$

Wir können uns also auf $G = N(K)$ beschränken, das heißt K als Normalteiler von G annehmen.

Für die Faltungsgruppe $(\pi_K(\mu_t) : t \in \mathbb{R})$ auf der Faktorgruppe G/K gilt

$$\operatorname{supp} \pi_K(\mu_t) = \pi_K(M_t) = \{\pi_K(a_t)\},$$

das heißt

$$\pi_K(\mu_t) = \varepsilon_{x_t} \quad \text{für alle } t \in \mathbb{R}$$

(mit $x_t := \pi_K(a_t)$). Da $(\pi_K(\mu_t) : t \in \mathbb{R})$ eine Faltungsgruppe ist, gilt

$$x_s x_t = x_{s+t} \quad \text{für alle } s, t \in \mathbb{R},$$

und x_0 ist das Einselement von G/K.

Aus der schwachen Stetigkeit von $t \to \pi_K(\mu_t) = \varepsilon_{x_t}$ folgt $\lim_{t \to 0} x_t = x_0 = 1$, das heißt $(x_t : t \in \mathbb{R})$ ist eine Einparameteruntergruppe von G/K.

Nach [11], 4.15, Theorem 1 existiert eine Einparameteruntergruppe $(g_t: t\in\mathbb{R})$ von G mit $\pi_K(g_t)=x_t$ für alle $t\in\mathbb{R}$; daraus folgt $g_t K=a_t K$ und somit $\mu_t=\tau_{g_t}(\omega_K)$ für alle $t\in\mathbb{R}$.

II. Seien nun μ_t und χ beliebig. Nach 4.9 existiert eine Faltungsgruppe $(\nu_t: t\in\mathbb{R})$ auf $G\times\mathbb{T}$ mit

$$Z\nu_t=\mu_t \quad \text{für alle } t\in\mathbb{R} \quad \text{und} \quad \nu_0=\omega_{K'}.$$

Nach I existieren Einparameteruntergruppen $(g_t: t\in\mathbb{R})$ bzw. $(z_t: t\in\mathbb{R})$ auf G bzw. $\mathbb{T}$ mit

$$\nu_t=\tau_{(g_t, z_t)}(\omega_{K'}) \quad \text{für alle } t\in\mathbb{R}.$$

Also existiert $u\in\mathbb{R}$ mit $z_t=e^{iut}$ für alle $t\in\mathbb{R}$, und aus $(g_t, z_t)\in N(K')$ folgt

$$\chi(g_t k g_{-t})=\chi(k)$$

und weiter

$$g_t k g_{-t} k^{-1}\in\operatorname{Ker}\chi \quad \text{für alle } t\in\mathbb{R} \quad \text{und} \quad k\in K.$$

Schließlich hat man

$$\mu_t=Z(\tau_{(g_t, e^{iut})}(\omega_{K'}))=e^{iut}\tau_{g_t}(\chi\,\omega_K)$$

für alle $t\in\mathbb{R}$. □

4.11. Korollar. *Die Abbildung* $Z: \mathscr{P}_{\mathbb{R}}(G\times\mathbb{T}; K')\to\mathscr{P}(G;\chi)$ *ist ein Isomorphismus.*

Es folgen nun noch einige algebraische Eigenschaften der Menge der χ-primitiven Distributionen, die bei der Konstruktion von Lévy-Abbildungen im 6. Paragraph wichtig werden.

4.12. *Notation.* $\mathscr{P}_{\mathbb{C}}(G;\chi):=\mathscr{P}(G;\chi)+i\mathscr{P}(G;\chi)$.

Die Zerlegung der Elemente von $\mathscr{P}_{\mathbb{C}}(G;\chi)$ ist eindeutig, wie aus dem folgenden Hilfssatz ersichtlich ist.

4.13. Hilfssatz *a)* $\mathscr{P}(G;\chi)\cap i\mathscr{P}(G;\chi)=\{0\}$.

b) Sei N ein kompakter Normalteiler von G. Falls $K\cap N$ nicht in $\operatorname{Ker}\chi$ *enthalten ist, gilt* $\pi_N(\mathscr{D}'(G;\chi))=\{0\}$.

c) Sei wieder N ein kompakter Normalteiler von G, und gelte $K\cap N\subset\operatorname{Ker}\chi$. *Dann ist*

$$\chi_N: KN/N\to\mathbb{T}, \quad kN\to\chi(k)$$

ein (wohldefinierter) Charakter auf KN/N; es gilt

$$\pi_N(\mathscr{P}(G;\chi))=\mathscr{P}(G/N;\chi_N)$$

sowie

$$\{T\in\mathscr{P}(G;\chi): \pi_N(T)=iu\cdot\chi_N\,\omega_{KN/N} \text{ für ein } u\in\mathbb{R}\}=\mathscr{P}(KN;\chi),$$
$$\{T\in\mathscr{P}_{\mathbb{C}}(G;\chi): \pi_N(T)=c\cdot\chi_N\,\omega_{KN/N} \text{ für ein } c\in\mathbb{C}\}=\mathscr{P}_{\mathbb{C}}(KN;\chi).$$

d) Ist L ein abgeschlossener Normalteiler von G und zugleich Lie-Gruppe, dann ist

$$\dim_{\mathbb{R}}(\mathscr{P}(KL;\chi))<\infty.$$

Beweis. a) Aus 4.10 folgt $T^* = -T$ für alle $T \in \mathscr{P}(G;\chi)$. Für jedes $T \in \mathscr{P}(G;\chi) \cap i\mathscr{P}(G;\chi)$ gilt dann $T^* = -T$ und $(iT)^* = -iT$. Daraus folgt $-iT = (iT)^* = -iT^* = +iT$, das heißt $T = 0$.

b) Seien $k \in K \cap N$ mit $\chi(k) \neq 1$ und $T \in \mathscr{D}'(G;\chi)$. Dann folgt aus

$$\overline{\chi(k)} \cdot \pi_N(T) = \pi_N(\tau_k(T)) = \tau_{kN}(\pi_N(T)) = \pi_N(T)$$

sofort $\pi_N(T) = 0$.

c) Ist dagegen $K \cap N \subset \operatorname{Ker}\chi$, so impliziert $kN = k'N$ $k^{-1}k' \in K \cap N \subset \operatorname{Ker}\chi$, das heißt $\chi(k) = \chi(k')$.

χ_N ist also wohldefiniert.

Zum Beweis von $\pi_N(\mathscr{P}(G;\chi)) = \mathscr{P}(G/N;\chi_N)$ sei $T \in \mathscr{P}(G/N;\chi_N)$; die von T erzeugte Faltungsgruppe sei $(e^{iut}\tau_{x_t}(\chi_N \omega_{KN/N}) : t \in \mathbb{R})$ mit $u \in \mathbb{R}$ und einer Einparameteruntergruppe $(x_t : t \in \mathbb{R})$ von G/N. Nach [11], 4.15, Theorem 1 existiert eine Einparameteruntergruppe $(g_t : t \in \mathbb{R})$ von G mit

$$g_t N = x_t \quad \text{für alle } t \in \mathbb{R}.$$

$T'' \in \mathscr{P}(G;\{1\})$ sei definiert durch

$$T''(f) := iuf(1) + \frac{d}{dt} f(g_t)|_{t=0}$$

und

$$T' := \chi\omega_K * T * \chi\omega_K \in \mathscr{P}(G;\chi).$$

Damit gilt für alle $f \in \mathscr{D}(G/N;\chi_N)$

$$\begin{aligned}
\pi_N(T')(f) &= T'(f \circ \pi_N) \\
&= T''(f \circ \pi_N) \\
&= iu \cdot f(1N) + \frac{d}{dt} f(g_t N)|_{t=0} \\
&= iu \cdot f(1N) + \frac{d}{dt} f(x_t)|_{t=0} = T(f),
\end{aligned}$$

das heißt $\pi_N(T') = T$.

Da die letzte der Identitäten in c) aus der vorletzten folgt, ist noch zu zeigen, daß

$$\{T \in \mathscr{P}(G;\chi) : \pi_N(T) = iu \cdot \chi_N \omega_{KN/N} \text{ für ein } u \in \mathbb{R}\} = \mathscr{P}(KN;\chi).$$

Sei also $T \in \mathscr{P}(G;\chi)$ und $(e^{iat}\tau_{g_t}(\chi\omega_K) : t \in \mathbb{R})$ die von T erzeugte Faltungsgruppe. Dann gilt folgende Kette von Äquivalenzen:

$$\begin{aligned}
&\bigvee_{u \in \mathbb{R}} \pi_N(T) = iu \cdot \chi\omega_{KN/N} \\
\Leftrightarrow\ &\bigvee_{u \in \mathbb{R}} \bigwedge_{t \in \mathbb{R}} e^{iat}\tau_{g_t N}(\chi_N \omega_{KN/N}) = \pi_N(e^{iat}\tau_{g_t}(\chi\omega_K)) = e^{iut}\chi_N\omega_{KN/N} \\
\Leftrightarrow\ &\bigwedge_{t \in \mathbb{R}} g_t N \in KN/N \\
\Leftrightarrow\ &\bigwedge_{t \in \mathbb{R}} g_t \in KN \\
\Leftrightarrow\ &T \in \mathscr{P}(KN;\chi).
\end{aligned}$$

d) $\mathscr{P}(KL;\chi)\cong\mathscr{P}_{\mathbb{R}}(KL\times\mathbb{T};K')$ wegen 4.11

$=\mathscr{P}_{\mathbb{R}}(K'(L\times\mathbb{T})\cap N(K');K')$ wegen 4.10

$=\mathscr{P}_{\mathbb{R}}(K'((L\times\mathbb{T})\cap N(K'));K')$

$\cong\mathscr{P}_{\mathbb{R}}(K'((L\times\mathbb{T})\cap N(K'))/K';\{1\})$ wegen c).

Da abgeschlossene Untergruppen und Faktoren von Lie-Gruppen wieder Lie-Gruppen sind, ist

$$(K'((L\times\mathbb{T})\cap N(K')))/K'\cong((L\times\mathbb{T})\cap N(K'))/(K'\cap(L\times\mathbb{T})\cap N(K'))$$

eine Lie-Gruppe. Deren (endlich-dimensionale) Lie-Algebra stimmt aber nach [13], Lemma 2 gerade mit der Menge der echt $\{1\}$-primitiven Distributionen überein.

5. Quadratische Distributionen

Ein komplizierteres Beispiel für χ-dissipative Distributionen bilden die χ-quadratischen Distributionen. Es stellt sich heraus, daß die Existenz von χ-quadratischen Distributionen nur von der Topologie von G/K abhängt. Am Ende des Paragraphen wird die Äquivalenz unserer Definition mit der in [13] gezeigt.

5.1. Definition. Eine χ-dissipative Distribution T auf G heißt *χ-quadratisch*, falls T normiert und lokal ist und $T^*=T$ gilt.

T heißt *echt K-quadratisch*, wenn T 1_K-quadratisch und positiv normiert ist.

5.2. Satz. *Von 0 verschiedene χ-quadratische Distributionen T auf G existieren genau dann, wenn G/K nicht total unzusammenhängend ist. In diesem Fall gibt es sogar echt K-quadratische Distributionen.*

Beweis. Wenn G/K total unzusammenhängend ist, gilt $K\cdot G_1=K$. Für die zu T gehörige Faltungshalbgruppe $(\mu_t: t\geqq 0)$ gilt nach 4.4 $\operatorname{supp}(\mu_t)\subset KG_1=K$, das heißt $\mu_t=e^{at}\chi\omega_K$ für alle $t\geqq 0$ und ein $a\in\mathbb{C}$. Wegen $T^*=T$ und der Normiertheit von T ist nur $T=0$ möglich.

Sei nun G/K nicht total unzusammenhängend. Dann existieren $X\in\mathscr{P}_{\mathbb{R}}(G;\{1\})$ und $f\in\mathscr{D}_{\mathbb{R}}(G)$ mit $f(k)=0$ für alle $k\in K$ und $X(f)\neq 0$. Wir betrachten zuerst den Fall $\chi=1_K$. Für die Funktion $f_0:=f^2$ gilt deshalb $X*X(f_0)>0$, $f_0\geqq 0$ und $f_0(k)=0$ für alle $k\in K$. Die gleichen Eigenschaften gelten auch für $f_0*\omega_K$, denn

$$(X*X)(f_0*\omega_K)=\int(X*X)(\sigma_k(f_0))\,d\omega_K(k)>0.$$

[[Wegen $f_0\geqq 0$ und $f_0|K=0$ ist für alle $k\in K$ $(X*X)(\sigma_k(f_0))\geqq 0$, es gilt $(X*X)(\sigma_1(f_0))=(X*X)(f_0)>0$ und die Abbildung $k\to\sigma_k(f_0)$ von K in $\mathscr{D}(G)$ ist stetig nach [2], Prop. 3.]] Auf analoge Weise kommt man schließlich zu

$$(\omega_K*X*X*\omega_K)(f_0)=X*X(\omega_K*f_0*\omega_K)>0.$$

$T':=\omega_K*X*X*\omega_K\neq 0$ ist also eine lokale, 1_K-dissipative Distribution mit $T'^*=T'$. Wegen $T'(1_G)=0$ ist T' echt K-quadratisch.

Wenn schließlich χ ein beliebiger Charakter auf K ist, wählt man eine Funktion $\phi \in \mathcal{D}(G;\chi)$ mit $1=\phi(1)=\|\phi\|_\infty$ und $\tilde{\phi}=\bar{\phi}$. $T:=\bar{\phi}\cdot T'$ (definiert durch $T(f)=T'(\bar{\phi}f)$) ist dann χ-quadratisch.

5.3. Als Beispiel folgt nun die Untersuchung von quadratischen Distributionen auf irreduziblen symmetrischen Paaren von kompaktem oder nicht kompaktem Typ, im Sinn von [7], S. 174 und 306. Nach [7], S. 307 und 389 enthält die Lie-Algebra Lie(G) von G Lie(K) als maximale echte Unteralgebra und es existiert ein Untervektorraum M von Lie(G) mit Lie$(G)=$Lie$(K)\oplus M$ (direkte Summe), $\mathrm{Ad}(k)M\subset M$ für $k\in K$, $[M,M]\subset$Lie(K) und $[\mathrm{Lie}(K),M]\subset M$. Die Einschränkung der Killing-Form B auf $M\times M$ ist entweder positiv oder negativ definit; es sei $(X_j: j\in p)$ eine Orthonormalbasis von M (dabei sei $p:=\{0,1,\ldots,p-1\}$).

Im kompakten Fall ist der folgende Satz eine Konsequenz aus [4], Corollaire 1.1 auf S. 36.

5.4. Satz. *Unter den Voraussetzungen von* 5.4 *sind sämtliche echt K-quadratischen Distributionen auf G gegeben durch*

$$\lambda\cdot\sum_{j\in p} X_j^{*2} * \omega_K \quad \text{mit } \lambda\geqq 0.$$

Beweis. Der triviale Fall $\dim G=\dim K$ sei oBdA ausgeschlossen. Wegen 5.2 genügt es zu beweisen, daß jede echt K-quadratische Distribution $T\neq 0$ von der angegebenen Form ist. Nach [9], Theorem 9.1 gilt

$$T=\sum_{j,l\in p} c_{jl}\, X_j * X_l * \omega_K$$

mit einer positiv semidefiniten reellen Matrix $c:=(c_{jl})_{j,l\in p}\neq 0$. Für jedes $k\in K$ sei $b(k)$ die (wegen $B(\mathrm{Ad}(k)X,\mathrm{Ad}(k)Y)=B(X,Y)$ für alle $X,Y\in M$) orthogonale $p\times p$-Matrix mit

$$\mathrm{Ad}(k)X_j=\sum_{l\in p} b(k)_{jl}X_l \quad \text{für alle } j\in p.$$

Es folgt

$$b(k^{-1})={}^t b(k) \quad \text{und} \quad c=b(k^{-1})\cdot c\cdot b(k) \quad \text{für alle } k\in K.$$

Sei $V\neq\{0\}$ ein Eigenraum von c zu einem Eigenwert $\lambda>0$. Für alle $k\in K$ gilt dann $b(k)V=V$. Man rechnet leicht nach, daß

$$L:=\{Y+\sum_{j\in p} x_j X_j:\ Y\in\mathrm{Lie}(K),(x_j: j\in p)\in V\}$$

eine Unteralgebra von Lie(G) ist, die Lie(K) enthält. Nach Voraussetzung 5.3 ist $L=$Lie(G), also $V=\mathbb{R}^p$. Somit ist c das λ-fache der Einheitsmatrix, das heißt

$$T=\lambda\cdot\sum_{j\in p} X_j^{*2} * \omega_K.$$

5.5. Als Beispiel betrachten wir das Gelfand-Paar

$$G:=SO(2p),\qquad K:=\left\{\begin{pmatrix} A & B\\ -B & A\end{pmatrix}: A+iB\in U(p)\right\}\cong U(p) \quad \text{für } p\geqq 3.$$

In den Bezeichnungen von [7], S. 339 ff. und S. 350 rechnet man leicht nach, daß der von den Matrizen

$$X_{j,l} := E_{j,l} - E_{l,j} - E_{p+j,p+l} + E_{p+l,p+j}$$

und

$$Y_{j,l} := E_{p+j,l} - E_{p+l,j} + E_{j,p+l} - E_{l,p+j} \qquad (j \in l \in p)$$

erzeugte Untervektorraum M von Lie(G) die Voraussetzungen von 5.3 erfüllt, und daß

$$B(X_{j,l}, X_{m,n}) = B(Y_{j,l}, Y_{m,n}) = -8(p-1)\,\delta_{jm}\,\delta_{ln}$$

und $B(X_{j,l}, Y_{m,n}) = 0$ für alle $j \in l \in p$ und $m \in n \in p$ gilt.

5.6. Satz. *Für $G := SO(2p)$ und $K := U(p)$ $(p \geqq 3)$ gilt:*

a) $\mathscr{P}_{\mathbb{R}}(G; K) = \{0\}$.

b) Sämtliche echt K-quadratischen Distributionen sind gegeben durch

$$\sum_{j \in l \in p} (X_{j,l}^{*2} + Y_{j,l}^{*2}) * \omega_K \qquad \text{mit } \lambda \geqq 0.$$

Beweis. a) Da K nicht Normalteiler von G ist und Lie(K) eine maximale echte Unteralgebra von Lie(G) ist, ist $N(K)/K$ diskret. Folglich existieren außer der trivialen keine weiteren Einparameteruntergruppen von G, die die Bedingung von 4.10 erfüllen. Daraus folgt $\mathscr{P}_{\mathbb{R}}(G; K) = \{0\}$.

b) ist direkte Konsequenz von 5.4.

5.7. *Notation.* Sei χ ein Charakter auf einer kompakten Untergruppe K von G. Mit $\mathscr{N}_\chi$ bezeichnen wir die Menge aller Funktionen $f \in \mathscr{D}(G; \chi)$ mit $T(f) = 0$ für alle $T \in \mathscr{P}(G; \chi)$; es gelte die Abkürzung $\mathscr{N}_K := \mathscr{N}_{1_K}$.

$\mathscr{M}_K$ sei die Menge aller $f \in \mathscr{D}(G; K)$ mit $f \geqq 0 = f(1)$.

5.8. Hilfssatz. *a) Für jedes $f \in \mathscr{N}_\chi$ existiert ein $f_0 \in \mathscr{M}_K$ mit $|f| \leqq f_0$.*

b) Für jedes $f \in \mathscr{D}(G; K)$ existiert ein $f_1 \in \mathscr{M}_K$, derart daß für jedes $\varepsilon > 0$ $f - \check{f} \leqq \varepsilon f_1$ in einer Umgebung der 1 gilt.

c) $\mathscr{N}_K$ ist der von $\mathscr{M}_K$ erzeugte Untervektorraum von $\mathscr{D}(G; K)$.

Beweis. a) Im Falle, daß G eine Lie-Gruppe ist, erhält man die Behauptung durch Anwendung der Taylor'schen Formel auf die Funktion $gK \to |f(g)|^2$ auf G/K.

Als nächstes betrachten wir den Spezialfall, daß G Lie-projektiv ist. Dann gibt es einen kompakten Normalteiler N von G, so daß G/N eine Lie-Gruppe ist und ein $f_N \in \mathscr{D}(G/N)$ mit $f = f_N \circ \pi_N$ existiert. Es gilt oBdA $K \cap N \subset \operatorname{Ker} \chi$ (andernfalls folgt $f = 0$ und es ist nichts zu beweisen). Man kann $\chi_N(kN) := \chi(k)$ für alle $k \in K$ definieren und erhält $f_N \in \mathscr{D}(G/N; \chi_N)$ und $T(f_N) = 0$ für alle $T \in \mathscr{P}(G/N; \chi_N)$ wegen 4.13c). Deshalb existiert eine Funktion $f_N' \in \mathscr{D}(G/N; KN/N)$ mit $|f_N| \leqq f_N'$ und $f_N'(1N) = 0$. $f_0 := f_N' \circ \pi_N$ leistet dann das Gewünschte.

Falls G eine beliebige lokalkompakte Gruppe ist, gibt es nach [6], S. 137 eine offene Lie-projektive Untergruppe G_0, die K enthält. Wir bekommen ein $f_0' \in \mathscr{D}(G_0; K)$ mit $f_0'(1) = 0$ und $f_0' \geqq |f|_{G_0}|$. Es gibt $g_j \in G$ $(j \in n)$ mit $g_0 = 1$ und $\operatorname{supp}(f) \subset \bigcup_{j \in n} g_j G_0$ sowie eine Funktion $f_1 \in \mathscr{D}_+(G_0)$, die auf $G_0 \cap \bigcup_{j \in n} g_j^{-1} \operatorname{supp}(f)$ größer als $\|f\|_\infty$ ist.

Definiert man

$$f'(g) := \begin{cases} f_0'(g) & \text{falls } g \in G_0, \\ f_1(g_j^{-1} g) & \text{falls } g \in g_j G_0 \text{ und } j \geqq 1, \\ 0 & \text{falls } g \notin \bigcup_{j \in n} g_j G_0, \end{cases}$$

so gilt $f' \in \mathscr{D}_+(G)$ und $f'(k) = 0$ für alle $k \in K$. Man hat auch $|f| \leqq f'$, denn auf G_0 und $\complement \bigcup_{j \in n} g_j G_0$ ist dies klar, und für $g \in g_j G_0$ ist entweder $g \in \operatorname{supp}(f)$ und $f'(g) = f_1(g_j^{-1} g) \geqq \|f\|_\infty \geqq |f(g)|$, oder $f(g) = 0$. $f_0 := \omega_K * f' * \omega_K$ leistet also das Gewünschte.

b) wird analog wie a) bewiesen.

c) Die Inklusion $\mathscr{M}_K \subset \mathscr{N}_K$ ist klar. Sei umgekehrt $f \in \mathscr{N}_K$. Dann gilt $\operatorname{Re} f$, $\operatorname{Im} f \in \mathscr{N}_K$; es existieren daher nach a) Funktionen $f_0, f_1 \in \mathscr{M}_K$ mit $|\operatorname{Im} f| \leqq f_0$ und $|\operatorname{Re} f| \leqq f_1$. Daraus folgt

$$(f_0 - \operatorname{Im} f) \in \mathscr{M}_K, \quad (f_1 - \operatorname{Re} f) \in \mathscr{M}_K$$

und

$$f = f_1 - (f_1 - \operatorname{Re} f) + i f_0 - i(f_0 - \operatorname{Im} f). \quad \square$$

5.9. Satz. *Sei T eine 1_K-dissipative Distribution auf G.*

a) Äquivalent sind:

(i) *T ist echt K-primitiv,*

(ii) *$T(f_0 \tilde{f}_1) = T(f_0) f_1(1) - T(f_1) f_0(1)$ für alle $f_0, f_1 \in \mathscr{D}(G; K)$.*

b) Äquivalent sind:

(iii) *T ist echt K-quadratisch,*

(iv) *$T(f_0 f_1) + T(f_0 \tilde{f}_1) = 2T(f_0) f_1(1) + 2T(f_1) f_0(1)$ für alle $f_0, f_1 \in \mathscr{D}(G; K)$.*

Beweis. a) (i)$\Rightarrow$(ii): Nach 4.10 existiert eine Einparameteruntergruppe $(g_t : t \in \mathbb{R})$ von G mit

$$T(f) = \frac{d}{dt} f(g_t)|_{t=0} \quad \text{für alle } f \in \mathscr{D}(G; K).$$

Für alle $f_0, f_1 \in \mathscr{D}(G; K)$ gilt somit

$$\begin{aligned} T(f_0 \tilde{f}_1) &= \frac{d}{dt} (f_0(g_t) f_1(g_{-t}))|_{t=0} \\ &= \frac{d}{dt} f_0(g_t)|_{t=0} \cdot f_1(1) - f_0(1) \frac{d}{dt} f_1(g_t)|_{t=0} \\ &= T(f_0) f_1(1) - T(f_1) f_0(1). \end{aligned}$$

(ii)$\Rightarrow$(i): Sei $\phi \in \mathscr{D}(G; K)$ mit $\phi = \tilde{\phi}$.

Dann gilt $T(\phi^2) = T(\phi)\phi(1) - T(\phi)\phi(1) = 0$.

Nach [3], Lemme 1 kann man für jedes $\varepsilon > 0$ die Funktion ϕ so wählen, daß $|T(1_G) - T(\phi^2)| < \varepsilon$. Daraus folgt $T(1_G) = 0$, T ist also positiv normiert und nach 3.4 reell.

Sei nun $f\in\mathscr{D}(G;K)$ beliebig und $\phi\in\mathscr{D}(G;K)$ so gewählt, daß $\phi=\tilde{\phi}$ und $\phi=1$ auf $K\cup\mathrm{supp}(f)$. Dann gilt

$$T^*(f)=\check{T}(f)=T(\tilde{f})=T(\phi^2\tilde{f})=T(\phi^2)f(1)-T(f)\phi^2(1)=-T(f),$$

das heißt $-T=T^*$ ist χ-dissipativ und T χ-primitiv.

b) (iii)$\Rightarrow$(iv): Zuerst wird (iv) in einigen Spezialfällen bewiesen:

(1) (iv) gilt, wenn $f_0\in\mathscr{M}_K$ und $f_1(1)=0$.

[[Für jedes $\varepsilon>0$ gibt es eine Umgebung von K, auf der $f_1\leqq\varepsilon$, das heißt

$$f_0(f_1-\varepsilon)\leqq 0=(f_0(f_1-\varepsilon))(1).$$

Es folgt $T(f_0(f_1-\varepsilon))\leqq 0$ für alle $\varepsilon>0$, also $T(f_0f_1)\leqq 0$. Da diese Ungleichung auch für $-f_1,\tilde{f}_1$ und $-\tilde{f}_1$ statt für f_1 richtig ist, gilt

$$T(f_0f_1)=T(f_0\tilde{f}_1)=0$$

und schließlich

$$T(f_0f_1)+T(f_0\tilde{f}_1)=0=2T(f_0)f_1(1)+2T(f_1)f_0(1).]]$$

(2) (iv) gilt, wenn $\tilde{f}_0=-f_0$.

[[Aus $T(f_0)=T^*(f_0)=T(\tilde{f}_0)=-T(f_0)$ folgt $T(f_0)=0$. Deshalb gilt

$$\begin{aligned}T(f_0f_1)+T(f_0\tilde{f}_1)&=T(f_0f_1)+T(\tilde{f}_0f_1)=T(f_0f_1)-T(f_0f_1)=0\\&=2T(f_0)f_1(1)+2T(f_1)f_0(1).]]\end{aligned}$$

(3) (iv) gilt, wenn f_0 auf einer Umgebung von $K\cup\mathrm{supp}(f_1)$ konstant $=1$ ist.

[[Da T lokal und positiv normiert ist, gilt $T(f_0)=0$. Deshalb ist

$$T(f_0f_1)+T(f_0\tilde{f}_1)=2T(f_1)=2T(f_0)f_1(1)+2T(f_1)f_0(1).]]$$

Entscheidend für den Beweis ist nun, daß sowohl die rechte, als auch die linke Seite von (iv) eine symmetrische Bilinearform von (f_0,f_1) sind. Deshalb folgt aus (1) und 5.8 c), daß (iv) für alle $f_0,f_1\in\mathscr{D}(G;K)$ mit $\tilde{f}_0=f_0$ und $f_0(1)=f_1(1)=0$ gilt; wegen (2) gilt (iv) für alle $f_0,f_1\in\mathscr{D}(G;K)$ mit $f_0(1)=f_1(1)=0$ und für beliebige $f_0,f_1\in\mathscr{D}(G;K)$ folgt (iv) durch zweimalige Anwendung von (3).

(iv)$\Rightarrow$(iii): Zuerst wird gezeigt, daß T lokal ist. Sei $g\notin K$, dann kann man Umgebungen $U_0=KU_0K$ von g und $U_1=KU_1K$ von 1 mit $U_0\cap U_1=\varnothing$ und eine Funktion $\phi\in\mathscr{D}(G;K)$ mit $\phi=\tilde{\phi}$, $\phi(1)=1$ und $\mathrm{supp}(\phi)\subset U_1$ wählen. Für alle $f\in\mathscr{D}(G;K)$ mit Träger in U_0 gilt dann $f\cdot\phi=f\cdot\tilde{\phi}=0$ und deshalb wegen (iv)

$$T(f)=T(f)\phi(1)+T(\phi)f(1)=\tfrac{1}{2}(T(f\phi)+T(f\tilde{\phi}))=0.$$

Um zu zeigen, daß T positiv normiert ist, wählt man eine Funktion $\phi\in\mathscr{D}(G;K)$, die auf einer Umgebung von K konstant $=1$ ist. Da T lokal ist, folgt

$$2T(\phi)=T(\phi\cdot\phi)+T(\phi\cdot\tilde{\phi})=2T(\phi)\phi(1)+2T(\phi)\phi(1)=4T(\phi).$$

Deshalb ist $T(\phi)=0$, woraus $T(1_G)=0$ folgt.

Zuletzt wird $T^*=T$ gezeigt. Sei $f\in\mathscr{D}(G;K)$ beliebig und $\phi\in\mathscr{D}(G;K)$ so gewählt, daß auf einer Umgebung von $K\cup\mathrm{supp}(f)\cup\mathrm{supp}(\check{f})$ $\phi=1$ gilt. Es folgt $T(\phi)=0$ und

$$T(f)+T(\check{f})=T(\phi f)+T(\phi\check{f})=2T(f)\,\phi(1)+T(\phi)f(1)=2T(f).$$

Deswegen ist $T(f)=T(\check{f})$ für alle $f\in\mathscr{D}(G;K)$ und folglich $T=\check{T}=T^*$. □

6. Existenz einer Lévy-Abbildung

Die in diesem Paragraphen gegebene Definition einer Lévy-Abbildung weicht aus mehreren Gründen von der in [13] ab. Zum einen ist es im Falle eines nicht-reellen Idempotents $\mu_0=\chi\omega_K$ nicht möglich, einen zum Begriff der echten Primitivität (die dem Begriff der primitiven Form in [13] entspricht) analogen Begriff zu finden, weshalb der konstante Term in der Funktion g in [13], Lemma 6 in die Lévy-Abbildung eingegliedert werden muß. Zum anderen soll die Definition auf eine Anwendung des Satzes II.2.4 von Roth [12] abgestimmt werden. Dazu wird der Begriff der Verträglichkeit eingeführt.

6.1. Definition. *Lévy-Abbildung für* $(G;\chi)$ heißt jede lineare Abbildung Γ von $\mathscr{D}(G;\chi)$ in sich mit

(LA1) $\quad T(\Gamma(f))=T(f)$ für alle $T\in\mathscr{P}(G;\chi)$ und $f\in\mathscr{D}(G;\chi)$,

(LA2) $\quad (f\to\Gamma(\tilde{\chi}\,\omega_K * f * \tilde{\chi}\,\omega_K)(g))\in\mathscr{P}_{\mathbb{C}}(G;\chi)$ für alle $g\in G$.

Im Falle $\chi=1_K$ spricht man von einer *Lévy-Abbildung für* $(G;K)$.

Bezeichnet K' wie in 4.8 die kompakte Untergruppe $\{(k,\chi(k)):k\in K\}$ von $G\times\mathbb{T}$, so heißt eine Lévy-Abbildung für $(G\times\mathbb{T};K')$ *mit der Abbildung Z verträglich* (oder kurz *verträglich*), wenn sie die Menge

$$\mathscr{D}_Z(G\times\mathbb{T};K'):=\{f\in\mathscr{D}(G\times\mathbb{T};K'):f(g,z)=zf(g,1)\quad\text{für alle } g\in G,\ z\in\mathbb{T}\}$$

in sich abbildet.

Eigenschaften:

6.2. Für jede Lévy-Abbildung Γ gilt $\Gamma\circ\Gamma=\Gamma$.

6.3. Jede Konvexkombination von Lévy-Abbildungen ist wieder eine Lévy-Abbildung.

Wenn Γ ein Lévy-Abbildung für $(G;K)$ ist, so sind auch $\bar{\Gamma}:f\to\overline{\Gamma(\bar{f})}$ und $\mathrm{Re}\,\Gamma:=\frac{1}{2}(\Gamma+\bar{\Gamma})$ Lévy-Abbildungen für $(G;K)$.

6.4. Jede Lévy-Abbildung Γ ist stetig.
Dies ist eine Konsequenz von (LA2).

6.5. Sei Γ' eine mit der Abbildung Z verträgliche Lévy-Abbildung für $(G\times\mathbb{T};K')$. Dann wird durch

$$\Gamma(f)(g) := \Gamma'(Zf)(g,1) \quad \text{für alle } f \in \mathscr{D}(G;\chi)$$

eine Lévy-Abbildung Γ für $(G;\chi)$ definiert mit

$$Z(\Gamma(f)) = \Gamma'(Zf) \quad \text{für alle } f \in \mathscr{D}(G;\chi).$$

Beweis. Die letzte Behauptung folgt aus der Verträglichkeit. Zum Nachweis von (LA1) betrachten wir $T \in \mathscr{P}(G;\chi)$. Nach 4.11 existiert eine Distribution $T' \in \mathscr{P}_{\mathbb{R}}(G \times \mathbb{T}; K')$ mit $T = ZT'$, und es gilt für alle $f \in \mathscr{D}(G;\chi)$

$$T(\Gamma(f)) = T'(Z(\Gamma f)) = T'(\Gamma'(Zf)) = T'(Zf) = T(f).$$

Die Bedingung (LA2) folgt aus

$$\Gamma(\tilde{\chi}\,\omega_K * f * \tilde{\chi}\,\omega_K)(g) = \Gamma'(Z(\tilde{\chi}\,\omega_K * f * \tilde{\chi}\,\omega_K))(g,1) = \Gamma'(\omega_{K'} * Zf * \omega_{K'})(g,1)$$

für alle $f \in \mathscr{D}(G)$ und $g \in G$. □

6.6. Um die Beziehung zwischen der hier gewählten Definition einer Lévy-Abbildung und der in [13] herzustellen, nehmen wir an, daß $\chi = 1_K$ und Γ eine reelle Lévy-Abbildung für $(G;K)$ ist. Wir wählen $\phi \in \mathscr{D}_{\mathbb{R}}(G;K)$ mit $1 = \phi(1) = \|\phi\|_\infty$ und setzen $\psi := \Gamma(\phi)$. Setzt man

$$\Gamma_0(f) := \tfrac{1}{2}\Gamma(f - f(1)\psi) - \tfrac{1}{2}\Gamma(f - f(1)\psi)^\sim \quad \text{für } f \in \mathscr{D}(G;K),$$

so ist Γ_0 eine Lévy-Abbildung im Sinn von [13]. Durch

$$\Gamma_1(f) := \tfrac{1}{2} f(1)(\psi + \tilde{\psi}) + \Gamma_0(f - f(1)\psi) \quad \text{für } f \in \mathscr{D}(G;K)$$

wird eine reelle Lévy-Abbildung für $(G;K)$ definiert mit

$$\Gamma_1(f)^\sim = \Gamma_1(f) \quad \text{falls } f \in \mathscr{D}(G;K) \text{ und } f = \tilde{f}$$

und

$$\Gamma_1(f)^\sim = -\Gamma_1(f) \quad \text{falls } f \in \mathscr{D}(G;K) \text{ und } f(1) = 0. \quad \square$$

Nun sind die wichtigsten Eigenschaften von Lévy-Abbildungen formuliert; es fehlt aber noch der Beweis für deren Existenz. Dieser benutzt die Lösung von Hilberts fünftem Problem. Der entscheidende Schritt ist dabei die Konstruktion einer Lévy-Abbildung für eine Lie-projektive Gruppe.

6.7. *Bezeichnungen.* Von jetzt bis zum Hilfssatz 6.12 sei G Lie-projektiv. Wir setzen

$$\mathscr{F} := \{L \subset G : L \text{ kompakter Normalteiler, } G/L \text{ Lie-Gruppe}\}$$

$$\mathscr{F}_0 := \{N \subset G : N \text{ kompakter Normalteiler, } N \cap K \subset \operatorname{Ker}\chi\}$$

und

$$\mathscr{F}_1 := \mathscr{F} \cap \mathscr{F}_0.$$

6.8. Hilfssatz. *Es gilt* $\mathscr{F}_1 \neq \varnothing$, $\bigcap \mathscr{F}_1 = \{1\}$ *und aus* $M, N \in \mathscr{F}_1$ *folgt* $M \cap N \in \mathscr{F}_1$.

Beweis. Zuerst zeigen wir, daß für jedes $L_0 \in \mathscr{F}$ und $\beta \in \chi(L_0 \cap K) \setminus \{1\}$ ein $L_1 \in \mathscr{F}$ mit $L_1 \subset L_0$ und $\beta \notin \chi(L_1 \cap K)$ existiert:

Sei $k \in L_0 \cap K$ mit $\chi(k) = \beta$. Da $k \cdot (L_0 \cap \operatorname{Ker} \chi) \neq 1$ kompakt ist, folgt aus der endlichen Durchschnittseigenschaft und

$$\bigcap_{L \in \mathscr{F}} k \cdot (L_0 \cap \operatorname{Ker} \chi) \cap L = \varnothing,$$

daß ein $L \in \mathscr{F}$ existiert mit $k \cdot (L_0 \cap \operatorname{Ker} \chi) \cap L = \varnothing$. Für $L_1 := L_0 \cap L \in \mathscr{F}$ gilt dann $\beta \notin \chi(L_1 \cap K)$.

Zum Beweis der ersten Behauptung beginnen wir mit einem beliebigen $L_0 \in \mathscr{F}$. Der obige Schluß ergibt ein $L_1 \in \mathscr{F}$ mit $L_1 \subset L_0$ und $-1 \notin \chi(L_1 \cap K)$. Die kompakte Untergruppe $\chi(L_1 \cap K)$ von $\mathbb{T}$ ist somit endlich und wir können in endlich vielen analogen Schritten ein $L_2 \in \mathscr{F}$ finden mit $L_2 \subset L_1$ und $\chi(L_2 \cap K) = \{1\}$, d.h. $L_2 \in \mathscr{F}_1$. Da der Durchschnitt einer Menge aus $\mathscr{F}_1$ mit einer aus $\mathscr{F}$ stets wieder in $\mathscr{F}_1$ liegt und $\bigcap \mathscr{F} = \{1\}$, gelten die beiden anderen Behauptungen. □

6.9. *Bezeichnungen.* Für jedes $N \in \mathscr{F}_0$ setzen wir

$$\mathscr{D}^N(G \times \mathbb{T}; K') := \{f \in \mathscr{D}(G \times \mathbb{T}; K') : Tf = 0 \text{ für alle } T \in \mathscr{P}((N \times \{1\}) K'; K')\}.$$

Weiter setzen wir $\mathscr{P}_{\mathbb{C}}(G \times \mathbb{T}; K')/\mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K')$ und $\mathscr{D}^N(G \times \mathbb{T}; K')$ in Dualität durch

$$\langle T + \mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K') | f \rangle := T(f)$$

für alle

$$(T + \mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K')) \in \mathscr{P}_{\mathbb{C}}(G \times \mathbb{T}; K')/\mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K')$$

und $f \in \mathscr{D}^N(G \times \mathbb{T}; K')$.

6.10. *Bezeichnungen.* Sei $N \in \mathscr{F}_0$. Wie in 4.13 setzen wir

$$\chi_N : \pi_N(K) \to \mathbb{T}, \qquad \chi_N(kN) := \chi(k)$$

sowie

$$K'_N := \{(x, \chi_N(x)) : x \in \pi_N(K)\} = \pi_{N \times \{1\}}(K').$$

Wegen 4.13 c) gilt für die Abbildung

$$\pi_{N \times \{1\}} : \mathscr{P}_{\mathbb{C}}(G \times \mathbb{T}; K') \to \mathscr{P}_{\mathbb{C}}\left(\frac{G}{N} \times \mathbb{T}; K'_N\right)$$

$\operatorname{Ker} \pi_{N \times \{1\}} \subset \mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K')$, und somit existiert eine eindeutig bestimmte lineare Abbildung

$$\Phi_G : \mathscr{P}_{\mathbb{C}}\left(\frac{G}{N} \times \mathbb{T}; K'_N\right) \to \mathscr{P}_{\mathbb{C}}(G \times \mathbb{T}; K')/\mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K')$$

mit

$$\Phi_G \circ \pi_{N \times \{1\}} = \pi_{\mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K')}.$$

In den folgenden Hilfssätzen sei eine verträgliche Lévy-Abbildung Γ für $\left(\frac{G}{N} \times \mathbb{T}; K'_N\right)$ gegeben.

Für jedes $x \in \frac{G}{N} \times \mathbb{T}$ sei $T_x \in \mathscr{P}_{\mathbb{C}}\left(\frac{G}{N} \times \mathbb{T}; K'_N\right)$ die durch

$$T_x(f) := \Gamma(\omega_{K'_N} * f * \omega_{K'_N})(x)$$

definierte Distribution auf $\frac{G}{N} \times \mathbb{T}$.

6.11. Hilfssatz. *Sei $f \in \mathscr{D}^N(G \times \mathbb{T}; K')$. Dann ist die Funktion*

$$F: G \times \mathbb{T} \to \mathbb{C} \qquad (g, z) \to \langle \Phi_G(T_{(gN, z)}) | f \rangle$$

ein Element von $\mathscr{D}(G \times \mathbb{T}; K')$ und erfüllt

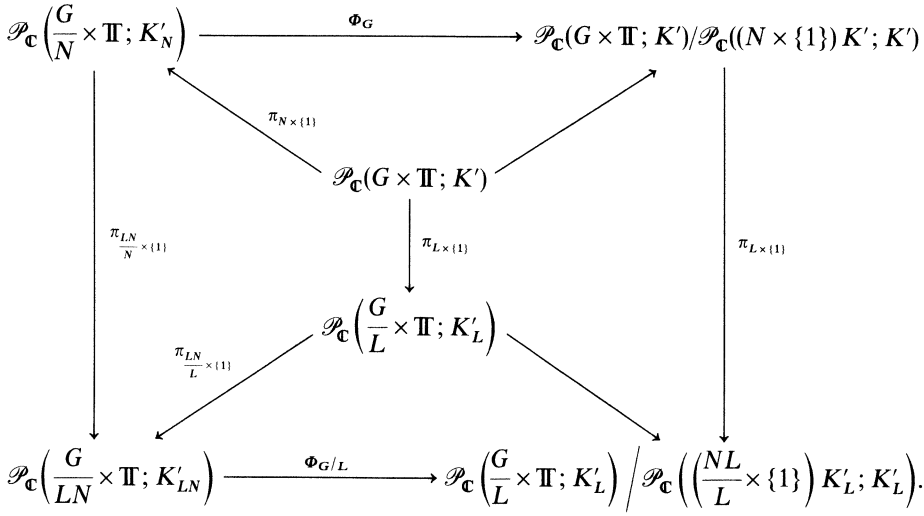
$$T(F) = T(f) \quad \text{für alle } T \in \mathscr{P}(G \times \mathbb{T}; K').$$

Gilt zusätzlich $f \in \mathscr{D}_Z(G \times \mathbb{T}; K')$, so ist auch $F \in \mathscr{D}_Z(G \times \mathbb{T}; K')$.

Beweis. Nach 6.8 existiert ein $L \in \mathscr{F}_1$ mit $LN \in \mathscr{F}_1$ (das heißt $\pi_N(K) \cap \pi_N(L) \subset \operatorname{Ker} \chi_N$) derart, daß es ein $\bar{f} \in \mathscr{D}^{\pi_L(N)}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$ gibt mit

$$f = \bar{f} \circ \pi_{L \times \{1\}}.$$

Wir betrachten nun das kommutative Diagramm



Hieraus sieht man, daß für alle $(g, z) \in G \times \mathbb{T}$

$$F(g, z) = \langle \Phi_{G/L}(\pi_{\frac{LN}{N} \times \{1\}}(T_{(gN, z)})) | \bar{f} \rangle$$

gilt.

Sei $(T_j: j\in n)$ eine $\mathbb{C}$-Basis von $\mathscr{P}_{\mathbb{C}}\left(\frac{G}{LN}\times\mathbb{T}; K'_{LN}\right)$. Dann gibt es eindeutig bestimmte $\alpha_j(gN,z)\in\mathbb{C}$, so daß

$$\pi_{\frac{LN}{N}\times\{1\}}(T_{(gN,z)})=\sum_{j\in n}\alpha_j(gN,z)\,T_j \quad \text{für alle } (g,z)\in G\times\mathbb{T}.$$

Da Funktionen $f_j\in\mathscr{D}\left(\frac{G}{LN}\times\mathbb{T}; K'_{LN}\right)$ $(j\in n)$ existieren mit $T_j(f_l)=\delta_{jl}$ für alle $j,l\in n$, gilt

$$\alpha_j=\Gamma(f_j\circ\pi_{\frac{LN}{N}\times\{1\}})\in\mathscr{D}\left(\frac{G}{N}\times\mathbb{T}; K'_N\right) \quad \text{für alle } j\in n.$$

Daraus folgt aber

$$F=\sum_{j\in n}\langle\Phi_{G/L}(T_j)|\bar{f}\rangle\cdot\alpha_j\circ\pi_{N\times\{1\}}\in\mathscr{D}(G\times\mathbb{T}; K').$$

Nun sei $T\in\mathscr{P}(G\times\mathbb{T}; K')$ und

$$\pi_{LN\times\{1\}}(T)=\sum_{j\in n}\beta_j T_j \quad (\text{mit } \beta_j\in\mathbb{C} \text{ für alle } j\in n).$$

Wegen

$$\begin{aligned}T(\alpha_j\circ\pi_{N\times\{1\}})&=\pi_{N\times\{1\}}(T)(\Gamma(f_j\circ\pi_{\frac{LN}{N}\times\{1\}}))\\&=\pi_{N\times\{1\}}(T)(f_j\circ\pi_{\frac{LN}{N}\times\{1\}})\\&=\pi_{LN\times\{1\}}(T)(f_j)\\&=\beta_j\end{aligned}$$

haben wir

$$\begin{aligned}T(F)&=\sum_{j\in n}\langle\Phi_{G/L}(T_j)|\bar{f}\rangle\cdot T(\alpha_j\circ\pi_{N\times\{1\}})\\&=\langle\Phi_{G/L}(\sum_{j\in n}\beta_j T_j)|\bar{f}\rangle\\&=\langle\Phi_{G/L}(\pi_{LN\times\{1\}}(T))|\bar{f}\rangle\\&=\langle\pi_{L\times\{1\}}(\Phi_G(\pi_{N\times\{1\}}(T)))|\bar{f}\rangle\\&=\langle\Phi_G(\pi_{N\times\{1\}}(T))|f\rangle\\&T(f).\end{aligned}$$

Nun wird noch die letzte Behauptung bewiesen. Sei dazu $f\in\mathscr{D}_Z(G\times\mathbb{T}; K')$. Die Dimension von

$$\left\{(T_j f': j\in n): f'\in\mathscr{D}_Z\left(\frac{G}{LN}\times\mathbb{T}; K'_{LN}\right)\right\}$$

ist $n-1$; dies folgt aus

$$\left\{(T_j f': j\in n): f'\in\mathscr{D}_Z\left(\frac{G}{LN}\times\mathbb{T}; K'_{LN}\right)\right\}=\left\{(ZT_j f'': j\in n): f''\in\mathscr{D}\left(\frac{G}{LN}; \chi_{LN}\right)\right\},$$

aus 4.11 und der Tatsache, daß $\left(\text{wegen } \dim_{\mathbb{R}}\left(\mathscr{P}\left(\frac{G}{LN}\times\mathbb{T};K'_{LN}\right)\right)=n\right)$ die Dimension von $\mathscr{P}_{\mathbb{R}}\left(\frac{G}{LN}\times\mathbb{T};K'_{LN}\right)$ $n-1$ ist.

Man kann also die T_j und f_j so modifizieren, daß

$$f_j\in\mathscr{D}_Z\left(\frac{G}{LN}\times\mathbb{T};K'_{LN}\right) \quad \text{für alle } j\in n-1.$$

Es folgt

$$T_{n-1}(f')=0 \quad \text{für alle } f'\in\mathscr{D}_Z\left(\frac{G}{LN}\times\mathbb{T};K'_{LN}\right).$$

T_{n-1} ist somit ein Vielfaches der Distribution

$$f'\to\frac{d}{dt}(f'*\omega_{K'_{LN}})(1,e^{it})|_{t=0}-i\,\omega_{K'_{LN}}(f') \quad \left(f'\in\mathscr{D}\left(\frac{G}{LN}\times\mathbb{T}\right)\right).$$

Also enthält auch die Klasse $\Phi_{G/L}(T_{n-1})$ ein Element dieser Gestalt, und folglich ist

$$\langle\Phi_{G/L}(T_{n-1})|\bar{f}\rangle=0.$$

Aus

$$f_j\circ\pi_{\frac{LN}{N}\times\{1\}}\in\mathscr{D}_Z\left(\frac{G}{N}\times\mathbb{T};K'_N\right) \quad \text{für } j\in n-1$$

folgt

$$\alpha_j=\Gamma(f_j\circ\pi_{\frac{LN}{N}\times\{1\}})\in\mathscr{D}_Z\left(\frac{G}{N}\times\mathbb{T};K'_N\right) \quad \text{für } j\in n-1$$

wegen der Verträglichkeit und damit

$$F=\sum_{j\in n-1}\langle\Phi_{G/L}(T_j)|\bar{f}\rangle\cdot\alpha_j\circ\pi_{N\times\{1\}}\in\mathscr{D}_Z(G\times\mathbb{T};K'). \quad \square$$

6.12. Hilfssatz. *Seien G Lie-projektiv, K eine kompakte Untergruppe von G, χ ein Charakter auf K und N eine Lie-Gruppe, die kompakter Normalteiler von G ist mit $N\cap K\subset\operatorname{Ker}\chi$. Wenn für $\left(\frac{G}{N}\times\mathbb{T};K'_N\right)$ eine verträgliche Lévy-Abbildung Γ existiert, dann existiert für $(G\times\mathbb{T};K')$ eine verträgliche Lévy-Abbildung Γ' mit*

$$\Gamma'\circ\pi_{N\times\{1\}}=\pi_{N\times\{1\}}\circ\Gamma.$$

Beweis. $(T_l: l\in m)$ sei eine Basis von $\mathscr{P}_{\mathbb{R}}((N\times\{1\})K';K')$ (vgl. 4.13 d)) und

$$T_m:=i\,\omega_{K'}\in\mathscr{P}((N\times\{1\})K';K').$$

Wegen $N\cap K\subset\operatorname{Ker}\chi$ existiert eine Funktion $\phi\in\mathscr{D}\left(\frac{G}{N};\chi_N\right)$ mit $\phi(1N)=-i$.

Wir setzen

$$\bar{f}_m := \Gamma(Z\phi) \in \mathscr{D}_Z\left(\frac{G}{N} \times \mathbb{T}; K'_N\right)$$

und $f_m := \bar{f}_m \circ \pi_{N \times \{1\}} \in \mathscr{D}_Z(G \times \mathbb{T}; K')$. Wegen

(†) $T(f) = 0$ für alle $T \in \mathscr{P}_{\mathbb{R}}((N \times \{1\}) K'; K')$ und $f \in \mathscr{D}(G \times \mathbb{T}; (N \times \{1\}) K')$

gilt $T_l(f_m) = 0$ für alle $l \in m$ und aus $f_m(1,1) = -i$ folgt $T_m(f_m) = 1$. Wegen 4.11 und

$$\mathscr{D}_Z(G \times \mathbb{T}; K') = \{Zf : f \in \mathscr{D}(G; \chi)\}$$

gibt es für jedes $T \in \mathscr{P}_{\mathbb{R}}((N \times \{1\}) K'; K') \setminus \{0\}$ ein $f \in \mathscr{D}_Z(G \times \mathbb{T}; K')$ mit $T(f) \neq 0$.

Deshalb existieren Funktionen $f_l \in \mathscr{D}_Z(G \times \mathbb{T}; K')$ $(l \in m)$ mit

$$T_j(f_l) = \delta_{jl} \quad \text{für } j, l \in m.$$

Indem man anstelle von f_l die Funktion $f_l - T_m(f_l) f_m$ benutzt, erreicht man

$$T_j(f_l) = \delta_{jl} \quad \text{für } j, l \in m+1.$$

Wir setzen nun

$$\Gamma''(f) := \sum_{j \in m+1} T_j(f) \cdot f_j \quad \text{für alle } f \in \mathscr{D}(G \times \mathbb{T}; K').$$

Γ'' ist eine lineare Abbildung von $\mathscr{D}(G \times \mathbb{T}; K')$ in $\mathscr{D}_Z(G \times \mathbb{T}; K')$, die die folgenden Bedingungen erfüllt:

(1) $T(\Gamma''(f)) = T(f)$ für alle $T \in \mathscr{P}((N \times \{1\}) K'; K')$, $f \in \mathscr{D}(G \times \mathbb{T}; K')$.

[[Wenn $T = \sum_{l \in m+1} \beta_l T_l$ mit $\beta_l \in \mathbb{R}$, so gilt

$$T(\Gamma''(f)) = \sum_{l, j \in m+1} \beta_1 T_j(f) T_l(f_j) = \sum_{l \in m+1} \beta_l T_l(f) = T(f).]]$$

(2) Für alle $(g, z) \in G \times \mathbb{T}$ gehört die Abbildung

$$(f \to \Gamma''(\omega_{K'} * f * \omega_{K'})(g, z)) \quad \text{zu } \mathscr{P}_{\mathbb{C}}((N \times \{1\}) K'; K').$$

[[Diese Abbildung ist nämlich gerade $\sum_{j \in m+1} f_j(g, z) \cdot T_j$.]]

Nach (1) gilt also $f - \Gamma'' f \in \mathscr{D}^N(G \times \mathbb{T}; K')$ für alle $f \in \mathscr{D}(G \times \mathbb{T}; K')$. Mit den Bezeichnungen von 6.9 und 6.10 setzen wir nun

$$\Gamma'(f)(g, z) := \langle \Phi_G(T_{(gN, z)}) | f - \Gamma'' f \rangle + \Gamma'' f(g, z)$$

für alle $g \in G$, $z \in \mathbb{T}$ und $f \in \mathscr{D}(G \times \mathbb{T}; K')$.

Nach 6.11 ist Γ' eine lineare Abbildung von $\mathscr{D}(G \times \mathbb{T}; K')$ in sich. Weiter folgt aus $f \in \mathscr{D}_Z(G \times \mathbb{T}; K')$ $f - \Gamma'' f \in \mathscr{D}_Z(G \times \mathbb{T}; K')$ und damit auch $\Gamma'(f) \in \mathscr{D}_Z(G \times \mathbb{T}; K')$. Γ' ist also verträglich.

Nun werden die beiden Bedingungen aus der Definition der Lévy-Abbildung nachgeprüft:

(LA1) Seien $T\in\mathscr{P}(G\times\mathbb{T};K')$, $f\in\mathscr{D}(G\times\mathbb{T};K')$ und F wie in 6.11 definiert, wobei man von $f-\Gamma''f$ anstelle von f ausgeht. Dann gilt

$$T(\Gamma' f)=T(F)+T(\Gamma'' f)=T(f-\Gamma'' f)+T(\Gamma'' f)=T(f).$$

(LA2) Seien $(g,z)\in G\times\mathbb{T}$ und $T\in P_{\mathbb{C}}(G\times\mathbb{T};K')$ ein Element der Klasse $\Phi_G(T_{(gN,z)})$. Wegen

$$\Gamma'(f)(g,z)=T(f-\Gamma'' f)+\Gamma'' f(g,z)$$

und (2) ist nur noch

$$(f\to T(\Gamma''(\omega_{K'}*f*\omega_{K'})))\in\mathscr{P}_{\mathbb{C}}(G\times\mathbb{T};K')$$

zu zeigen.

Dies folgt aber aus der Zerlegung von T und Γ'' in die Summe eines primitiven und das i-fache eines primitiven Teils.

Zum Beweis für die letzte Behauptung betrachten wir eine Funktion $\bar f\in\mathscr{D}\left(\frac{G}{N}\times\mathbb{T};K'_N\right)$. Für

$$f:=\bar f\circ\pi_{N\times\{1\}}\in\mathscr{D}(G\times\mathbb{T};(N\times\{1\})K')$$

gilt dann wegen (†):

$$\Gamma''(f)=\sum_{j\in m+1}T_j(f)f_j=T_m(f)f_m.$$

Damit haben wir für alle $(g,z)\in G\times\mathbb{T}$

$$\begin{aligned}\Gamma'(\bar f\circ\pi_{N\times\{1\}})(g,z)&=\langle\Phi_G(T_{(gN,z)})\,|\,(\bar f-T_m(f)\bar f_m)\circ\pi_{N\times\{1\}}\rangle+T_m(f)f_m(g,z)\\&=\langle\pi_{N\times\{1\}}(\Phi_G(T_{(gN,z)}))\,|\,\bar f-T_m(f)\bar f_m\rangle+T_m(f)f_m(g,z)\\&=\Gamma(\bar f)(gN,z)-T_m(f)\Gamma(\bar f_m)(gN,z)+T_m(f)\bar f_m(gN,z)\\&=\Gamma(\bar f)\circ\pi_{N\times\{1\}}(g,z)\\&=(\pi_{N\times\{1\}}\circ\Gamma)(\bar f)(g,z).\end{aligned}$$

Bei der vorletzten Umformung wird benutzt, daß nach Wahl von $\bar f_m$ und wegen 6.2 $\Gamma(\bar f_m)=\bar f_m$ gilt. □

6.13. Satz. *Seien G eine lokalkompakte Gruppe, K eine kompakte Untergruppe und χ ein Charakter auf K.*

Dann existiert eine mit der Abbildung Z verträgliche Lévy-Abbildung für $(G\times\mathbb{T};K')$.

Beweis. I. Zunächst nehmen wir an, daß G Lie-projektiv ist. Es gelten die Bezeichnungen 6.7, 6.9 und 6.10. Die Menge

$$\mathscr{M}:=\left\{(N,\Gamma_N):N\in\mathscr{F}_0,\quad\Gamma_N\text{ verträgliche Lévy-Abb. für }\left(\frac{G}{N}\times\mathbb{T};K'_N\right)\right\}$$

ist geordnet durch

$$(M, \Gamma_M) < (N, \Gamma_N) :\Leftrightarrow N \subset M \quad \text{und} \quad \pi_{\frac{M}{N}\times\{1\}} \circ \Gamma_M = \Gamma_N \circ \pi_{\frac{M}{N}\times\{1\}}.$$

Auf die gleiche Weise wie (1) und (2) in 6.12 beweist man für jedes $L \in \mathscr{F}_1$ die Existenz einer verträglichen Lévy-Abbildung Γ_L für $\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$. $\mathscr{M}$ ist also nicht leer.

Außerdem ist $\mathscr{M}$ induktiv geordnet:

Sei etwa J eine totalgeordnete Menge, $(L_j, \Gamma_j) \in \mathscr{M}$ für alle $j \in J$ und $(L_j, \Gamma_j) < (L_l, \Gamma_l)$ wenn $j < l$; dann setzen wir $L := \bigcap_{j\in J} L_j$. Für jedes

$$f \in \mathscr{D}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right) = \bigcup_{j\in J} \left\{ \bar{f} \circ \pi_{\frac{L_j}{L}\times\{1\}} : \bar{f} \in \mathscr{D}\left(\frac{G}{L_j} \times \mathbb{T}; K'_{L_j}\right)\right\}$$

– etwa von der Form $f = \bar{f} \circ \pi_{(L_j/L)\times\{1\}}$ – setzen wir

$$\Gamma_L(f) := \Gamma_j(\bar{f}) \circ \pi_{(L_j/L)\times\{1\}}.$$

Γ_L ist eine (wohldefinierte) Abbildung von $\mathscr{D}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$ in sich, die $\mathscr{D}_Z\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$ in sich abbildet, also verträglich ist. Γ_L ist sogar eine Lévy-Abbildung, denn:

(LA1) Für alle $T \in \mathscr{P}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$ und $f = \bar{f} \circ \pi_{(L_j/L)\times\{1\}} \in \mathscr{D}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$ gilt

$$\begin{aligned} T(\Gamma_L(f)) &= T(\Gamma_j(\bar{f}) \circ \pi_{(L_j/L)\times\{1\}}) = \pi_{(L_j/L)\times\{1\}}(T)(\Gamma_j(\bar{f})) \\ &= \pi_{(L_j/L)\times\{1\}}(T)(\bar{f}) = T(\bar{f} \circ \pi_{(L_j/L)\times\{1\}}) = T(f). \end{aligned}$$

(LA2) Sei $x \in \frac{G}{L} \times \mathbb{T}$ und für jedes $j \in J$ $T_j + iT'_j$ die nach 4.13 eindeutige Zerlegung von

$$(f \to \Gamma_j(\omega_{K'_{L_j}} * f * \omega_{K'_{L_j}})(\pi_{(L_j/L)\times\{1\}}(x))) \in \mathscr{P}_{\mathbb{C}}\left(\frac{G}{L_j} \times \mathbb{T}; K'_{L_j}\right).$$

Für die Distributionen $T: f \to T_j(\bar{f})$ und $T': f \to T'_j(\bar{f})$ (wobei $f = \bar{f} \circ \pi_{(L_j/L)\times\{1\}}$) gilt damit

$$(T + iT')(f) = \Gamma_L(\omega_{K'_L} * f * \omega_{K'_L})(x) \quad \text{für alle } f \in \mathscr{D}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$$

und

$$\pi_{(L_j/L)\times\{1\}}(T) = T_j, \quad \pi_{(L_j/L)\times\{1\}}(T') = T'_j \quad \text{für alle } j \in J,$$

also $T, T' \in \mathscr{P}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right)$ und schließlich

$$(f \to \Gamma_L(\omega_{K'_L} * f * \omega_{K'_L})(x)) = T + iT' \in \mathscr{P}_{\mathbb{C}}\left(\frac{G}{L} \times \mathbb{T}; K'_L\right).$$

Es ist sehr leicht zu sehen, daß für alle $j \in J$ $(L, \Gamma_L) > (L_j, \Gamma_j)$ gilt. Nach dem Lemma von Zorn besitzt $\mathscr{M}$ also ein maximales Element (N, Γ_N).

Nun sei $L \in \mathscr{F}$ beliebig. Wendet man Hilfssatz 6.12 auf $\left(\frac{G}{L \cap N}, \frac{N}{L \cap N}, \pi_{L \cap N}(K)\right)$ anstatt (G, N, K) an, so erhält man eine verträgliche Lévy-Abbildung $\Gamma_{L \cap N}$ für

$$\left(\frac{G}{L \cap N} \times \mathbb{T}; \pi_{(L \cap N) \times \{1\}}(K)\right)$$

mit

$$\Gamma_{L \cap N} \circ \pi_{(N/L \cap N) \times \{1\}} = \pi_{(N/L \cap N) \times \{1\}} \circ \Gamma_N,$$

das heißt $(L \cap N, \Gamma_{L \cap N}) > (N, \Gamma_N)$.

Da (N, Γ_N) aber maximal war, folgt $L \supset N$ für alle $L \in \mathscr{F}$, also

$$N \subset \bigcap_{L \in \mathscr{F}} L = \{1\}.$$

$\Gamma_{\{1\}}$ ist also die gesuchte verträgliche Lévy-Abbildung für $(G \times \mathbb{T}; K')$.

II. Im Falle einer beliebigen lokalkompakten Gruppe G existiert nach [6], S. 137 eine offene Lie-projektive Untergruppe G_0, die K enthält. Für $(G_0 \times \mathbb{T}; K')$ existiert eine verträgliche Lévy-Abbildung Γ_0, und man muß nur

$$\Gamma(f)(g, z) := \begin{cases} \Gamma_0(f \mid (G_0 \times \mathbb{T}))(g, z) & g \in G_0 \\ 0 & g \notin G_0 \end{cases}$$

setzen, um eine verträgliche Lévy-Abbildung Γ für $(G \times \mathbb{T}; K')$ zu erhalten. □

6.14 Korollar. a) *Für jede lokalkompakte Gruppe G und jeden Charakter χ auf einer kompakten Untergruppe von G existiert eine Lévy-Abbildung.*

b) *Für jede lokalkompakte Gruppe G und jede kompakte Untergruppe K von G existiert eine Lévy-Abbildung, die reell ist und folgende Eigenschaften hat:*

$$\Gamma(f)^{\sim} = \Gamma(f), \quad \text{falls } f \in \mathscr{D}(G; K) \text{ und } \tilde{f} = f,$$
$$\Gamma(f)^{\sim} = -\Gamma(f), \quad \text{falls } f \in \mathscr{D}(G; K) \text{ und } f(1) = 0.$$

Beweis. a) folgt aus 6.13 und 6.5.

b) folgt aus a), 6.3 und 6.6. □

7. Die Lévy-Hinčin-Formel für Faltungshalbgruppen von Wahrscheinlichkeitsmaßen auf homogenen Räumen

Das in den vorherigen Paragraphen Bewiesene wird nun verwendet, um die Lévy-Hinčin-Formel für Faltungshalbgruppen von Wahrscheinlichkeitsmaßen auf homogenen Räumen zu beweisen, das heißt nach 3.6 eine Zerlegung jeder positiv normierten, 1_K-dissipativen Distribution in einen echt K-primitiven, einen echt K-quadratischen Anteil und den Lévy-Anteil. Der Letztere kann mit Hilfe eines Maßes auf $G \backslash K$, des Lévy-Maßes, dargestellt werden, das im folgenden eingeführt wird.

7.1. Definition. *Ein Radon-Maß η auf $G\setminus K$ heißt positives K-Lévy-Maß, falls*

(PL0) $\eta\in\mathscr{M}_+(G\setminus K)$,
(PL1) $\tau_k(\eta)=\sigma_k(\eta)=\eta$ für alle $k\in K$,
(PL2) $f\,|\,(G\setminus K)\in L^1(\eta)$ für alle $f\in\mathscr{N}_K$,
(PL3) $\eta(\complement U)<\infty$ für jede offene Umgebung U von K.

7.2. Beispiel. *Sei T eine positiv normierte 1_K-dissipative Distribution auf G. Dann läßt sich die Abbildung, die jeder in einer Umgebung von K verschwindenden Funktion $f\in\mathscr{D}(G)$ die Zahl $T(f)$ zuordnet, zu einem positiven Radon-Maß η auf $G\setminus K$ fortsetzen. η ist ein positives K-Lévy-Maß.*

Beweis. Die erste Behauptung folgt aus [3], Lemme 1. Da T positiv normiert ist, folgt $\eta\geqq 0$, denn für alle $f\in\mathscr{D}_+(G)$, die in einer Umgebung von K verschwinden, gilt $-f\leqq -f(1)=0$ und folglich $-\eta(f)=T(-f)\leqq 0$. (PL1) folgt aus $T\in\mathscr{D}'(G;K)$, (PL3) läßt sich von [3], Lemme 1 ableiten. Zum Beweis von (PL2) können wir uns wegen 5.8 c) auf $f\in\mathscr{M}_K$ beschränken, d.h. $f\in\mathscr{D}(G;K)$ und $f(1)=0\leqq f$. Für alle Funktionen $\phi\in\mathscr{D}(G;K)$ mit $0\leqq\phi\leqq 1$, die in einer Umgebung von K konstant 1 sind, gilt $(f\phi)(1)=0\leqq f\phi$, also $T(f\phi)\geqq 0$ und $T((1-\phi)f)\leqq T(f)$. Daraus folgt aber

$$\eta(f\,|\,(G\setminus K))=\sup_{\phi} T((1-\phi)f)\leqq T(f),\qquad \text{das heißt } f\,|\,(G\setminus K)\in L^1(\eta).$$

7.3. Definition. *Ist T eine positiv normierte 1_K-dissipative Distribution auf G, so heißt das in 7.2 definierte Maß η das Lévy-Maß der Distribution T.*

7.4. Ist η das Lévy-Maß von T, so gilt $\int\limits_{G\setminus K} f\,d\eta\leqq T(f)$ für alle $f\in\mathscr{M}_K$. Dies folgt aus dem Beweis von 7.2.

7.5. Hilfssatz. a) *Sei Γ eine reelle Lévy-Abbildung für $(G;K)$ und η ein positives K-Lévy-Maß. Dann ist für jedes $f\in\mathscr{D}(G)$ die Einschränkung von $f_0:=\omega_K*f*\omega_K-\Gamma(\omega_K*f*\omega_K)$ auf $G\setminus K$ η-integrierbar und die Abbildung*

$$A_\eta^\Gamma\colon\mathscr{D}(G)\to\mathbb{C},\ f\to\int\limits_{G\setminus K} f_0\,d\eta$$

ein Element von $\mathscr{D}'(G;K)$.

b) *Sind Γ und Γ' zwei reelle Lévy-Abbildungen und η ein positives K-Lévy-Maß, so gilt*

$$A_\eta^\Gamma-A_\eta^{\Gamma'}=T+a\omega_K\quad \textit{mit}\quad T\in\mathscr{P}_{\mathbb{R}}(G;K)\ \textit{und}\ a\in\mathbb{R}.$$

Beweis. a) Die erste Behauptung folgt aus (PL2), da wegen (LA1) $f_0\in\mathscr{N}_K$ gilt. Der Beweis der Stetigkeit von A_η^Γ erfolgt mit Hilfe von 6.4 analog wie der Beweis von Hilfssatz 4.2 aus [5].

b) Wir setzen $T_0:=A_\eta^\Gamma-A_\eta^{\Gamma'}$, $a:=T_0(1_G)$ und $T:=T_0-a\omega_K$. Für jedes $f\in\mathscr{D}(G;K)$ mit $f\leqq f(1)$ und $f(1)\geqq 0$ wählen wir eine Funktion $\phi\in\mathscr{D}(G;K)$, die in einer Umgebung von $\operatorname{supp}(f)\cup K$ den Wert $f(1)$ annimmt. Somit gilt $T_0(\phi)=af(1)$. Da Γ und Γ' reell sind, gilt nach (LA2) $\Gamma(f-\phi)=\Gamma'(f-\phi)=0$. Daraus folgt

$$T(f)=T_0(f)-af(1)=T_0(f-\phi)=\int(\Gamma'(f-\phi)-\Gamma(f-\phi))\,d\eta=0.$$

Da T außerdem reell ist, gilt $T\in\mathscr{P}_{\mathbb{R}}(G;K)$.

7.6. Satz (Lévy-Hinčin-Formel). *Sei Γ eine reelle Lévy-Abbildung für $(G;K)$.*

a) *Für jedes positive K-Lévy-Maß η auf $G\setminus K$ existiert eine eindeutig bestimmte reelle Zahl a_η^Γ, derart daß $L_\eta^\Gamma:=A_\eta^\Gamma+a_\eta^\Gamma\omega_K$ positiv normiert und 1_K-dissipativ ist. Das Lévy-Maß von L_η^Γ ist η.*

b) *Ist Γ' eine weitere reelle Lévy-Abbildung für $(G;K)$, so gilt $L_\eta^\Gamma-L_\eta^{\Gamma'}\in\mathscr{P}_{\mathbb{R}}(G;K)$.*

c) *Sei T eine positiv normierte 1_K-dissipative Distribution auf G mit Lévy-Maß η. Dann existieren eindeutig bestimmte echt K-primitive bzw. echt K-quadratische Distributionen P bzw. Q auf G, so daß*

$$T=P+Q+L_\eta^\Gamma.$$

Beweis. a) Wählt man $a_\eta^\Gamma:=\int(\Gamma(1_G)-1_G)\,d\eta\in\mathbb{R}$, so gilt $L_\eta^\Gamma(1_G)=0$ und für alle $f\in\mathscr{D}_{\mathbb{R}}(G;K)$ mit $f(1)\geqq f$ und $f(1)\geqq 0$ gilt wegen $\Gamma(f)=f(1)\,\Gamma(1_G)$:

$$\begin{aligned}L_\eta^\Gamma(f)&=\int(f-f(1)\,\Gamma(1_G))\,d\eta+f(1)\cdot\int(\Gamma(1_G)-1_G)\,d\eta\\&=\int(f-f(1)\,1_G)\,d\eta\leqq 0.\end{aligned}$$

Damit ist L_η^Γ positiv normiert und 1_K-dissipativ.

Die Eindeutigkeit folgt aus $0=L_\eta^\Gamma(1_G)=a_\eta^\Gamma+A_\eta^\Gamma(1_G)$. Für jede Funktion $f\in\mathscr{D}(G;K)$, die in einer Umgebung von K verschwindet, gilt $\Gamma(f)=0$ und folglich $L_\eta^\Gamma(f)=A_\eta^\Gamma(f)=\int f\,d\eta$. Also ist η das Lévy-Maß von L_η^Γ.

b) folgt aus 7.5 b) und der positiven Normiertheit von L_η^Γ und $L_\eta^{\Gamma'}$.

c) Die Distribution $T':=T-L_\eta^\Gamma$ ist reell, lokal und erfüllt $T'(1_G)=0$. Für alle $f\in\mathscr{D}_{\mathbb{R}}(G;K)$ mit $f(1)\geqq f$ und $f(1)\geqq 0$ gilt

$$\begin{aligned}T'(f)&=T(f)-\int(f-\Gamma(f))\,d\eta-f(1)\int(\Gamma(1_G)-1_G)\,d\eta\\&=T(f)+\int(f(1)\,1_G-f)\,d\eta\leqq T(f)+T(f(1)\,1_G-f)\\&=0\end{aligned}$$

wegen 7.4. T' ist also positiv normiert, 1_K-dissipativ und lokal. Nun setzt man $Q:=\frac{1}{2}(T'+T'^*)$ und $P:=\frac{1}{2}(T'-T'^*)$. Q ist nach Definition 5.1 echt K-quadratisch. Um zu beweisen, daß P positiv normiert und 1_K-primitiv ist, genügt es, für jedes $f\in\mathscr{D}_{\mathbb{R}}(G;K)$ mit $f\leqq f(1)$ und $f(1)\geqq 0$

$$P(f)=\tfrac{1}{2}T'(f-\tilde{f})=0$$

zu zeigen. Nach 5.8 c) existiert $\psi\in\mathscr{M}_K$, derart, daß für alle $\varepsilon>0$ $\varepsilon\psi\geqq f-\tilde{f}$ in einer Umgebung von K gilt. Da T' lokal, positiv normiert und 1_K-dissipativ ist, folgt $\varepsilon T'(\psi)\geqq T'(f-\tilde{f})=2P(f)$ für alle $\varepsilon>0$, das heißt $P(f)\leqq 0$. Wegen $P^*=-P$ ist P echt K-primitiv. Nach Konstruktion gilt

$$T=T'+L_\eta^\Gamma=P+Q+L_\eta^\Gamma.$$

Es bleibt die Eindeutigkeit von P und Q zu zeigen. Diese folgt aber daraus, daß man $T'=T-L_\eta^\Gamma$ auf genau eine Weise in die Summe einer symmetrischen und einer antisymmetrischen Distribution aufspalten kann.

7.7. Bemerkung. Die Zerlegung der Distribution T in 7.6 c) ist zwar bei vorgegebener Lévy-Abbildung Γ eindeutig, bei Änderung von Γ ändert sich L_η^Γ wegen b) und folglich auch der echt K-primitive Anteil P. Dagegen ist der echt K-quadratische Anteil Q unabhängig von Γ; auch dies folgt aus 7.6 b).

Bibliographie

1. Berg, Ch., Forst, G.: Potential theory on locally compact Abelian groups. Berlin-Heidelberg-New York: Springer 1975
2. Bruhat, F.: Distributions sur un groupe localement compact et applications à l'étude des groupes p-adiques. Bull. Soc. Math. Fr. **89**, 43–75 (1961)
3. Duflo, M.: Représentations de semi-groupes de mesures sur un groupe localement compact. Ann. Inst. Fourier (Grenoble) **28/3**, 225–249 (1978)
4. Gebuhrer, M.-O.: Processus de Markov relativistes. Noyaux de Hunt invariants et Laplaciens généralisés sur les espaces riemanniens symétriques de type compact. Thèse, Université Louis Pasteur, Strasbourg I (1973)
5. Hazod, W.: Über die Lévy-Hinčin-Formel auf lokalkompakten topologischen Gruppen. Z. Wahrscheinlichkeitstheor. Verw. Geb. **25**, 301–322 (1973)
6. Hazod, W.: Stetige Faltungshalbgruppen von Wahrscheinlichkeitsmaßen und erzeugende Distributionen. Lecture Notes in Mathematics Vol. **595**. Berlin-Heidelberg-New York: Springer 1977
7. Helgason, S.: Differential geometry and symmetric spaces. New York-London: Academic Press 1962
8. Heyer, H.: Convolution semigroups of probability measures on Gelfand pairs. Expo. Math. **1**, 3–45 (1983)
9. Hunt, G.A.: Semigroups of measures on a Lie group. Trans. Am. Math. Soc. **81**, 264–294 (1956)
10. Lasser, R.: On the Lévy-Hinčin formula for commutative hypergroups. In: Proceedings of the 7th Conference on Probability Measures on Groups, pp. 298–308, Lecture Notes in Mathematics Vol. **1064**. Berlin-Heidelberg-New York-Tokyo: Springer 1984
11. Montgomery, D., Zippin, L.: Topological Transformation Groups, 2nd edition. New York: Interscience Publishers 1964
12. Roth, J.P.: Opérateurs dissipatifs et semi-groupes dans les espaces de fonctions continues. Ann. Inst. Fourier (Grenoble) **26/4**, 1–97 (1976)
13. Siebert, E.: Über die Erzeugung von Faltungshalbgruppen auf beliebigen lokalkompakten Gruppen. Math. Z. **131**, 313–333 (1973)
14. Yosida, K.: Functional Analysis, 6th edition. Berlin-Heidelberg-New York: Springer 1980
15. Zeuner, H.: Complex Lévy Measures. In: Proceedings of the 7th Conference on Probability Measures on Groups, pp. 471–480, Lecture Notes in Mathematics Vol. **1064**. Berlin-Heidelberg-New York-Tokyo: Springer 1984

Eingegangen am 27. August 1984

Math. Z. 189, 555–560 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

In a Nonreflexive Space the Subdifferential is Not Onto

Simon Fitzpatrick and Bruce Calvert

Department of Mathematics, University of Auckland, Auckland, New Zealand

1. Introduction

Letting X be a real Banach space and $\varphi: X \to \mathbb{R} \cup \{\infty\}$ a proper lower semicontinuous convex function, $\partial\varphi: X \to 2^{X^*}$ is defined by $f \in \partial\varphi(x)$ iff for all y in X, $\varphi(y) \geqq \varphi(x) + (f, y-x)$. Let B be the closed unit ball in X. In case $\varphi(x) = \psi_B(x)$ defined as 0 on B and ∞ off B, $\partial\psi_B(x) = \{f \in X^*: f \text{ attains its sup on } B \text{ at } x\}$. R.C. James showed $\partial\psi_B$ is onto iff X is reflexive. We refer only to his more recent proof [5], and show that with little change it extends to deal with a general φ. We do not try to make this note self contained, but refer to James [5], following his setting out for ease of reference.

We note the result of Brezis [3] on surjectivity of $\partial\varphi$ in a Hilbert space extends to a general Banach space, using the main result.

In addition, we pose the question as to whether the result extends to maximal monotone operators, and give some small results in this setting; it extends to bilipschitzian angled bounded maps. This is complementary to existence results as in Zini [7], Gossez [4], and Brezis and Browder [2].

2. Modifications of Theorems 1 and 2 of James [5]

We recall that the dual convex function $\varphi^*: X^* \to \mathbb{R} \cup \{\infty\}$ is given by $\varphi^*(g) = \sup\{g(x) - \varphi(x): x \in X\}$.

Proposition 1. *Let X be a Banach space, and $\varphi: X \to \mathbb{R} \cup \{\infty\}$ a lower semicontinuous proper convex function. The following are equivalent:*

(i) *For all $g \in X^*$, $g - \varphi$ is bounded above.*
(ii) *The function φ^* has only finite real values.*
(iii) *For all $g \in X^*$ and $k \in \mathbb{R}$, the set $\{x \in X: (g-\varphi)(x) \geqq k\}$ is bounded.*

Proof. (i)$\Rightarrow$(iii). Let $g \in X^*$, $k \in \mathbb{R}$, and let $K = \{x: (g, x) - \varphi(x) \geqq k\}$. Given $f \in X^*$, take $h \in \mathbb{R}$ such that for all $x \in X$, $(f+g, x) - \varphi(x) \leqq h$. For $x \in K$, $(f, x) \leqq h - k$. Thus K is weakly bounded, hence bounded.

(iii)$\Rightarrow$(i). Given $g\in X^*$, taking $k=0$ there is $M\geqq 0$ such that if $\|x\|\geqq M$ then $(\varphi-g)(x)\geqq 0$. By Hahn Banach, using the separation theorem to separate the epigraph of $\varphi-g$ from some point outside it, there is $f\in X^*$ and $h\geqq 0$ such that $f(x)-h\leqq(\varphi-g)(x)$ for all $x\in X$. Hence, if $\|x\|\leqq M$ then $(\varphi-g)(x)\geqq -M\|f\|-h$. Hence $g-\varphi\leqq M\|f\|+h$.

(i)$\Rightarrow$(ii). By definition of φ^*. q.e.d.

Theorem 1. *Let X be a real separable Banach space. Let $\varphi: X\to\mathbb{R}\cup\{\infty\}$ be l.s.c. and convex, with $D(\varphi)$ having nonempty interior. Suppose that for all $g\in X^*$ the function $g-\varphi$ is bounded above. The following are equivalent:*

(i) *X is not reflexive.*

(ii) *There is $x_0\in X$, $\theta>0$, and a bounded sequence (f_n) in X^* converging weak star to 0, with $\varphi^*(f)-(f, x_0)\geqq\theta$ for all $f\in\operatorname{co}\{f_n\}$.*

(iii) *There exists $x_0\in X$, $\theta>0$, such that if $\lambda_n>0$, $\sum\lambda_n=1$, then there exists $\alpha>0$, and a bounded sequence (g_n) in X^*, g_n converging weak star to 0, with $\varphi^*(g)-(g, x_0)=\alpha$, where $g=\sum\lambda_n g_n$, and for each n,*

$$\varphi^*\left(\sum_{i=1}^{n}\lambda_i g_i\right)-\left(\sum_{i=1}^{n}\lambda_i g_i, x_0\right)<\alpha\left(1-\theta\sum_{i=n+1}^{\infty}\lambda_i\right).$$

(iv) $R(\partial\varphi)\neq X^*$.

Proof. Replacing φ by $x\to\varphi(x_0+x)$ where $x_0\in X$, we may assume $0\in\operatorname{int}D(\varphi)$ and $x_0=0$, for (i)$\Rightarrow$(ii), and simply $x_0=0$ for (ii)$\Rightarrow$(iii) and (iii)$\Rightarrow$(iv).

(i)$\Rightarrow$(ii). There is $E>0$ and $\delta>0$ such that if $\|x\|\leqq\delta$ then $\varphi(x)\leqq E$. Let $f\in X^*$. For $x\in X$ with $\|x\|\leqq\delta$, $f(x)\leqq\varphi(x)+\varphi^*(f)\leqq E+\varphi^*(f)$. Hence, $\|f\|\leqq\delta^{-1}(E+\varphi^*(f))$.

As in (i)$\Rightarrow$(ii) of Theorem 1 of [5], take $F\in X^{**}$, $\|F\|<1$, $d(F, X)>\theta_0>0$. Let $\{x_n\}$ be dense in X. Take $M>E/(\delta\theta_0)$. For each n, by Helly's Theorem, take $f_{n-1}\in X^*$ with $\|f_{n-1}\|<M$, $f_{n-1}(x_i)=0$, $1\leqq i\leqq n-1$, $F(f_{n-1})=M\theta_0$. Then $(f_n, x)\to 0$ for all $x\in X$.

For $f\in\operatorname{co}\{f_n\}$, $M\theta_0=F(f)\leqq\|F\|\|f\|\leqq\|f\|\leqq\delta^{-1}(E+\varphi^*(f))$. Letting $\theta=\delta M\theta_0-E$, we have $\theta>0$ and $\varphi^*(f)\geqq\delta M\theta_0-E=\theta$.

(ii)$\Rightarrow$(iii). This follows from the following lemma, which is proved in the same way as Lemma 1 of [5].

Lemma 1. *Let $\theta\in(0, 1)$, $\{f_n\}$ bounded in X^*, $\varphi^*(f)\geqq\theta$ for all $f\in\operatorname{co}\{f_n\}$. If $\lambda_n>0$, $\sum_{n=1}^{\infty}\lambda_n=1$, then there exists $\alpha\geqq\theta$ and a sequence (g_n) with $g_n\in\operatorname{co}\{f_n, f_{n+1}, \ldots\}$, $\varphi^*\left(\sum_{n=1}^{\infty}\lambda_n g_n\right)=\alpha$, and for all n, $\varphi^*\left(\sum_{1}^{n}\lambda_i g_i\right)<\alpha\left(1-\theta\sum_{i=n+1}^{\infty}\lambda_i\right)$.*

(iii)$\Rightarrow$(iv). The proof of (iii)$\Rightarrow$(iv) of [5] shows that if $g=\sum_{i=1}^{\infty}\lambda_i g_i$ then for $x\in X$, $(g, x)-\varphi(x)<\varphi^*(g)$. This implies $g\notin R(\partial\varphi)$, since if $g\in\partial\varphi(z)$, then for all x, $(g, x)-\varphi(x)\leqq(g, z)-\varphi(z)$, giving $\varphi^*(g)=(g, z)-\varphi(z)$.

(iv)$\Rightarrow$(i). If X is reflexive, and $g\in X^*$, then for large k the set $K=\{x\in X: \varphi(x)\leqq(g, x)+k\}$ is nonempty and weakly compact by Prop. 1. Thus $\varphi(x)-(g, x)$ attains its minimum over K, hence over X, at a point z, giving $g\in\partial\varphi(z)$. Hence $R(\partial\varphi)=X^*$. q.e.d.

We recall from James [5] that for (f_n) a sequence in X^*, $L\{f_n\}$ denotes the set of $w \in X^*$ with

$$\liminf (f_n, x) \leqq (w, x) \leqq \limsup (f_n, x)$$

for all $x \in X$.

Theorem 2. *Let X be a real Banach space. Let $\varphi: X \to \mathbb{R} \cup \{\infty\}$ be l.s.c., convex, with $\operatorname{int} D(\varphi) \neq \phi$. Suppose that for all $g \in X^*$ the function $g - \varphi$ is bounded above. The following are equivalent:*

(i) *X is not reflexive.*

(ii) *There exists $\theta \in (0, 1)$, a bounded sequence (f_n) in X^*, and a subspace X_0 of X such that for all $w \in X_0^{\perp}$ and $f \in \operatorname{co}\{f_n\}$, $\varphi^*(f - w) \geqq \theta$, and $(f_n, x) \to 0$ for all $x \in X_0$.*

(iii) *There is $\theta \in (0, 1)$ such that if $\lambda_n > 0$ and $\sum \lambda_n = 1$, then there exists $\alpha \geqq \theta$ and a bounded sequence (g_n) in X^* such that if $w \in L(g_n)$ then $\varphi^*\left(\sum_{i=1}^{\infty} \lambda_i (g_i - w)\right) = \alpha$ and $\varphi^*\left(\sum_{i=1}^{n} \lambda_i (g_i - w)\right) < \alpha\left(1 - \theta \sum_{i=n+1}^{\infty} \lambda_i\right)$ for each n.*

(v) *$R(\partial \varphi) \neq X^*$.*

Proof. (i) $\Rightarrow$ (ii). Let X_0 be a nonreflexive separable subspace of X. Take $\theta > 0$, f_n bounded in X^*, $(f_n, x) \to 0$ for $x \in X_0$, and such that for $f \in \operatorname{co}\{f_n\}$, $(\varphi|_{X_0})^*(f) > \theta$. For $w \in X_0^{\perp}$,

$$\varphi^*(f - w) = \sup_{x \in X} \{(f - w, x) - \varphi(x)\} \geqq \sup_{x \in X_0} \{(f - w, x - \varphi(x)\} = \sup_{x \in X_0} (f, x) - \varphi(x_0)\} \geqq \theta.$$

(ii) $\Rightarrow$ (iii). This is proved by the following Lemma, which is proved in the same way as Lemma 2 of [5].

Lemma 2. *Let X be a real Banach space, $0 < \theta < 1$, $\{f_n\}$ bounded in X^*, $\varphi^*(f - w) \geqq \theta$ for all $f \in \operatorname{co}\{f_n\}$, and for all $w \in L\{f_n\}$.*

If $\lambda_n > 0$, $\sum \lambda_n = 1$ then there exists α, $\theta \leqq \alpha$, and a bounded sequence (g_n) in X^*, such that for all $w \in L\{g_n\}$, $\varphi^*\left(\sum_{i=1}^{\infty} \lambda_i (g_i - w)\right) = \alpha$ and for each n,

$$\varphi^*\left(\sum_{i=1}^{n} \lambda_i (g_i - w)\right) < \alpha\left(1 - \sum_{i=n+1}^{\infty} \lambda_i\right).$$

(iii) $\Rightarrow$ (iv). As in (iii) $\Rightarrow$ (iv) of Theorem 2 of [5], take Δ, $0 < \Delta < \frac{1}{2}\theta^2$, and take $\lambda_n > 0$ so that for all n, $\lambda_{n+1} < \varepsilon_n \lambda_n$, where (ε_n) is a sequence of real numbers converging downwards to 0.

Take $k_0 \in \mathbb{R}$, $k_0 < \varphi^*\left(\sum_{i=1}^{\infty} \lambda_i (g_i - w)\right)$. The set

$$K_0 = \{x \in X : (\textstyle\sum \lambda_i (g_i - w), x) - \varphi(x) > k_0\}$$

is nonempty, and bounded. Suppose $\|x\| \leqq K$ for $x \in K_0$. Take M with $\|g_i\| \leqq M$ for all i. Take N such that for $n \geqq N$, $\varepsilon_n < \Delta/(MK)$. Hence $\lambda_{n+1} < \Delta (MK)^{-1} \lambda_n$. Let $x \in K_0$. Since $\liminf g_i(x) \leqq w(x)$, and $\theta \leqq \alpha$, there is $n \geqq N$ such that $((g_{n+1} - w), x) < \theta^2 - 2\Delta \leqq \alpha\theta - 2\Delta$. Then

$$\left(\sum_{i=1}^{\infty}\lambda_i(g_i-w),\,x\right)-\varphi(x)$$

$$<\left(\sum_{i=1}^{n}\lambda_i(g_i-w),\,x\right)+(\alpha\theta-2\Delta)\,\lambda_{n+1}+\left(\sum_{i=n+2}^{\infty}\lambda_i(g_i-w),\,x\right)-\varphi(x)$$

$$\leqq\varphi^*\left(\sum_{i=1}^{n}\lambda_i(g_i-w)\right)+(\alpha\theta-2\Delta)\,\lambda_{n+1}+2MK\sum_{n+2}^{\infty}\lambda_i$$

$$<\alpha\left(1-\theta\sum_{n+1}^{\infty}\lambda_i\right)+(\alpha\theta-2\Delta)\,\lambda_{n+1}+2\Delta\sum_{n+1}^{\infty}\lambda_i$$

$$=\alpha-(\alpha\theta-2\Delta)\sum_{n+2}^{\infty}\lambda_i<\alpha=\varphi^*\left(\sum_{i=1}^{\infty}\lambda_i(g_i-w)\right).$$

Hence $\sum_{i=1}^{\infty}\lambda_i(g_i-w)\notin R(\partial\varphi)$.

(iv)$\Rightarrow$(i). As in Theorem 1. q.e.d.

The following result changes round the suppositions of Theorem 2. It gives an extension of the Hilbert space case of Brezis [3, Theorem 2.13].

Theorem 3. *Let X be a real Banach space. Let $\varphi: X\to\mathbb{R}\cup\{\infty\}$ be l.s.c., convex, and have* int $D(\varphi)$ *nonempty. The following are equivalent:*

(a) $R(\partial\varphi)=X^*$.

(b) *For all $g\in X^*$ the function $g-\varphi$ is bounded above and X is relexive.*

Proof. (b)$\Rightarrow$(a). By (iv)$\Rightarrow$(i) of Theorem 2.

(a)$\Rightarrow$(b). Let $g\in X^*$. If $g\in\partial\varphi(z)$ then for all $x\in X$, $\varphi(x)\geqq\varphi(z)+(g,x-z)$, giving the first condition. By (i)$\Rightarrow$(iv) of Theorem 2, X is reflexive.

3. Monotone Operators

We ask: If X is a Banach space and $A: X\to 2^{X^*}$ is a monotone operator with $R(A)=X^*$ and int $D(A)\neq\phi$, must X be reflexive?

Theorem 4. *Let X be a Banach space and let $G\subseteq X$ be open and let $c>0$, and let $A: G\to X^*$ be Lipschitz and satisfy for x, y in G, (*)$(Ax-Ay,\,x-y)\geqq c\,\|x-y\|^2$. Then X is reflexive.*

Proof. Without loss of generality, $0\in G$. Let $i: X_0\subseteq X$ be a separable subspace. Let $A_0=i^*Ai$. A_0 satisfies (*). By Aronszajn, [1, Theorem 1], since A_0 is Lipschitz, there exists a point $x\in G$ with $\frac{1}{t}(A_0(x+tu)-A_0(x))$ converging to $DA_0(x)u$, $DA_0(x)$ a bounded linear operator, for each $u\in X_0$. Note Aronszajn requires in the second to last paragraph of [1, p. 163] that the range of the mapping be the conjugate of a separable space, with the domain being separable. The properties (A) and (B) of a Gateaux τ-differential at the bottom of [1, p. 161] hold with τ the weak star topology in our case. Taking limits gives

$$(DA_0(x)u, u) \geqq c\|u\|^2,$$

so X_0 is Hilbertizable, and reflexive.

By Eberlein-Shmulyan, X is relexive.

Definition 1. Let X be a Banach space. $A: X \to 2^{X^*}$ is angle bounded if there is $\sigma > 0$ such that for all u, v, w in X,

$$(Au - Aw, w - v) \leqq \sigma(Au - Av, u - v).$$

Corollary. *Let X be a Banach space and let $A: X \to 2^{X^*}$ be Lipschitz on an open set G with Lipschitz inverse on the open set $A(G)$. If A is angle bounded on G then X is reflexive.*

Proof. By Brezis and Browder [2, p. 158], for u in G, and v near u,

$$(Au - Av, u - v) \geqq \frac{1}{4L}\|Au - Av\|^2$$

where L is the Lipschitz constant of A. Since A has Lipschitz inverse, for some $c > 0$,

$$(Au - Av, u - v) \geqq c\|u - v\|^2,$$

and the conditions of Theorem 3 are satisfied.

We give another linearization proof, changing (*).

Theorem 5. *Let X be a real Banach space, $c > 0$, $0 \leqq a < \frac{1}{2}$, $K > 0$. Let $A: X \to X^*$ satisfy $c\|x\|^2 \leqq (Ax, x) \leqq K\|x\|^2$ for $\|x\|$ large. Suppose that for $x \in X$, $t \to A(tx)$ is continuous. Suppose for all x, y, $\|A(x+y) - (A(x) + A(y))\| \leqq K\|x\|^a \|y\|^a$.*

Then X is reflexive.

Proof. By Rassias [6], A has a derivative at infinity L, $Lx = \lim_{n \to \infty} 2^{-n} A(2^n x)$. Then $K\|x\|^2 \geqq (Lx, x) \geqq c\|x\|^2$ and X is Hilbertizable.

Note for the next result that if A is linear, monotone and onto then A^{-1} is locally bounded, hence a bounded linear operator.

Theorem 6. *Let X be a Banach space, $A: X \to X^*$ a linear monotone and surjective map. Let $\pi: X \to X^{**}$ be the usual isometry. If $\|(A + A^*\pi)/2\| < \|A^{-1}\|^{-1}$ then X is relexive.*

Proof. The inequality above gives $B = (A - A^*\pi)/2$ being invertible. Hence $B^*: X^{**} \to X$ is an isomorphism, and since $-B = B^*\pi$, we have $X^{**} = \pi(X)$.

References

1. Aronszajn, N.: Differentiability of Lipschitzian mappings between Banach spaces. Studia Math. **57**, 147–160 (1976)
2. Brezis, H., Browder, F.E.: Nonlinear Integral equations and systems of Hammerstein type. Adv. Math. **18**, 115–147 (1975)

3. Brezis, H.: Operateurs Maximaux Monotones. Amsterdam: North Holland 1973
4. Gossez, J.P.: Opérateurs monotones non lineaires dans les espaces de Banach non reflexifs. J. Math. Anal. Appl. **34**, 371–395 (1971)
5. James, R.C.: Reflexivity and the sup of linear functionals. Isr. J. Math. **13**, 289–300 (1972)
6. Rassias, J.H.: On approximation of approximately linear mappings by linear mappings. J. Funct. Anal. **46**, 126–130 (1982)
7. Zini, E.: A propos de quelques opérateurs non lineaires. Math. Z. **141**, 111–138 (1975)

Received September 13, 1984

Math. Z. 189, 561–570 (1985)

Mathematische Zeitschrift
© Springer-Verlag 1985

Obere Abschätzung für die Anzahl der B-Zwillinge auf kurzen Intervallen

Gerhard Bantle

Universität Ulm, Math. III, Oberer Eselsberg, D-7900 Ulm, Bundesrepublik Deutschland

1. Einleitung

Es sei B die Menge aller natürlichen Zahlen, die sich als Summe zweier Quadrate ganzer Zahlen darstellen lassen. Das Paar $(n, n+1)$ nennen wir B-Zwilling, wenn sowohl $n \in B$ als auch $n+1 \in B$ ist. Ferner sei

$$B_2(x,k,l) = |\{n \leqq x : n \in B, n+1 \in B, n \equiv l(k)\}|.$$

1964 zeigte G. Rieger [12] mit dem Sieb die obere Abschätzung

$$B_2(x,1,1) \ll \frac{x}{\log x}. \tag{1.1}$$

Hooley [3] und Indlekofer [4] zeigten 1974 unabhängig voneinander, die analoge untere Abschätzung

$$B_2(x,1,1) \gg \frac{x}{\log x}.$$

Das Problem einer Asymptotik für B-Zwillinge scheint äußerst schwierig zu sein und ist noch ungelöst.

1978 übertrug Kelly [9] in seiner Dissertation die obigen Resultate auf kurze Intervalle und zeigte insbesondere mit der Methode von Rieger:

$$B_2(x,1,1) - B_2(x-x^\theta,1,1) \ll \frac{x^\theta}{\log x} \quad \text{für } \theta > \frac{1}{2}.$$

Das Ergebnis dieser Arbeit ist der folgende

Satz. *Sei $A>0$, $1 \leqq k \leqq (\log x)^A$, $(k, l(l+1)) = 1$, dann existiert eine nur von A und θ abhängende Konstante $c>0$, so daß*

$$B_2(x,k,l) - B_2(x-x^\theta,k,l) \leqq \frac{c}{k} \prod_{\substack{p \equiv 3 (4) \\ p \mid k}} \frac{p}{p-2} \frac{x^\theta}{\log x} \quad \textit{für } \theta > \frac{75}{307}. \tag{1.2}$$

Bemerkungen. 1971 hat Wolke in [14] obere Abschätzungen für multiplikative Funktionen auf schnell wachsenden Folgen angegeben, aus denen sich sofort (1.1) ableiten läßt; für Abschätzungen auf kurzen Intervallen sind diese Methoden jedoch nicht geeignet.

Die zu (1.2) analoge untere Abschätzung wurde in [1] für $\theta \geqq 0.8521$ gezeigt.

Das Kelly'sche Resultat $\theta > \frac{1}{2}$ läßt sich direkt auf $\theta > \frac{1}{3}$ ausdehnen, die zusätzliche Verbesserung auf $\theta > \frac{75}{307}$ beruht im wesentlichen auf der nichttrivialen Behandlung von Summen, deren Summationsintervalle eine Länge kleiner 1 haben, mit klassischen Van der Corput-Methoden (vgl. Bemerkung nach Lemma 2).

2. Bezeichnungen, ein Siebresultat

Mit p bezeichnen wir wie üblich eine Primzahl, mit $\mathbb{P}$ die Menge aller Primzahlen und mit $\mathscr{P}$ eine Teilmenge von $\mathbb{P}$; ferner sei $\mathscr{P}(z) = \{p \in \mathscr{P}: p < z\}$ und speziell

$$\mathscr{P}_i^{(k)}(z) = \{p < z: p \equiv i(4), p \nmid k\}, \quad \mathscr{P}_i^{(k)} = \mathscr{P}_i^{(k)}(\infty) \quad \text{und} \quad \mathscr{P}_i = \mathscr{P}_i^{(1)}(\infty).$$

Beim Sieb verwenden wir die von Halberstam und Richert in [2] eingeführten Bezeichnungen:

Sei $\mathscr{A}$ eine endliche Folge ganzer Zahlen, dann bezeichnen wir

$$S(\mathscr{A}, \mathscr{P}, z) := |\{a \in \mathscr{A}: (a, \mathscr{P}(z)) = 1\}| = \sum_{\substack{a \in \mathscr{A} \\ (a, \mathscr{P}(z)) = 1}} 1 \tag{2.1}$$

als Siebfunktion.

Ist $Q(\mathscr{P})$ die Menge aller positiven quadratfreien ganzen Zahlen, deren Primfaktoren alle in $\mathscr{P}$ liegen, dann sei für $d \in Q(\mathscr{P})$

$$|\mathscr{A}_d| = |\{a \in \mathscr{A}: a \equiv 0(d)\}|. \tag{2.2}$$

$X > 1$ sei eine Approximation für $|\mathscr{A}|$; $r(\mathscr{A}, 1) := |\mathscr{A}| - X$.

$\omega(d)$ sei eine auf $Q(\mathscr{P})$ multiplikative Funktion und so gewählt, daß

$$r(\mathscr{A}, d) := |\mathscr{A}_d| - \frac{\omega(d)}{d} X \tag{2.3}$$

im Mittel bei Summation über d klein bleibt.

Vereinbaren wir noch, wie üblich, $\omega(d) = 0$, falls $d \notin Q(\mathscr{P})$ und setzen für $z \geqq 2$

$$V(z) = \prod_{p < z} \left(1 - \frac{\omega(p)}{p}\right), \tag{2.4}$$

dann gilt folgendes

Lemma 1 (vgl. [5], Theorem 1). *Für alle $p \in \mathscr{P}$ gelte $0 < \frac{\omega(p)}{p} < 1$; ferner existiere eine Konstante $C \geqq 2$ so, daß für alle $2 \leqq w_1 < w_2$:*

$$\frac{V(w_1)}{V(w_2)} < \frac{\log w_2}{\log w_1}\left(1+\frac{C}{\log w_1}\right),$$

dann gilt für $2 \leqq z \leqq D \leqq z^3$ *mit einer absoluten O-Konstanten*

$$S(\mathscr{A},\mathscr{P},z) \leqq XV(z)\left\{2e^{\gamma}\frac{\log z}{\log D}+O\left(\frac{e^{\sqrt{C}}}{(\log D)^{\frac{1}{3}}}\right)\right\}+\sum_{\substack{d<D\\ d\mid\mathscr{P}(z)}}|r(\mathscr{A},d)|.$$

Dieses Lemma ist ein Spezialfall des Rosser-Iwaniec Siebes.

3. Grundlegende Aufspaltung

Bekanntlich ist $n \in B$, d.h. als Summe von 2 Quadratzahlen darstellbar, genau dann, wenn n keinen Primteiler aus der Restklasse 3 mod 4 in ungerader Vielfachheit enthält, d.h., wenn sich n darstellen läßt als

$$n = a^2 v \quad \text{mit } (a, 2\mathscr{P}_1)=1 \quad \text{und} \quad (v,\mathscr{P}_3)=1.$$

Wegen $(k, l(l+1))=1$ erhalten wir sofort

$$B_2(x,k,l)-B_2(x-x^{\theta},k,l)=\sum_{\substack{a\leqq\sqrt{x}\\(a,b)=1\\(ab,2k\mathscr{P}_1^{(k)})=1}}\ \sum_{b\leqq\sqrt{x+1}}\ \sum_{\substack{v,w\\ \frac{x-x^{\theta}}{a^2}<v\leqq\frac{x}{a^2}\\ a^2v+1=b^2w\\ (vw,\mathscr{P}_3^{(k)})=1\\ a^2v\equiv l(k)}} 1. \tag{3.1}$$

Ist (v_0, w_0) mit $1 \leqq v_0 \leqq b^2$, $1 \leqq w_0 \leqq a^2$ eine Lösung von $b^2 w_0 - a^2 v_0 = 1$, so sind alle Lösungen von $b^2 w - a^2 v = 1$ gegeben durch

$$\begin{aligned} v &= v_0 + b^2 n \\ w &= w_0 + a^2 n. \end{aligned} \tag{3.2}$$

Wir erhalten also aus (3.1)

$$\begin{aligned} B_2(x,k,l)-B_2(x-x^{\theta},k,l) \leqq & \sum_{\substack{a,b\\ 1\leqq a^2b^2k\leqq x^{2\theta}\\ (a,b)=1\\ (ab,2k\mathscr{P}_1^{(k)})=1}}\ \sum_{\substack{\frac{x-x^{\theta}}{a^2b^2}-\frac{v_0}{b^2}<n\leqq\frac{x}{a^2b^2}-\frac{v_0}{b^2}\\ ((b^2n+v_0)\cdot(a^2n+w_0),\mathscr{P}_3^{(k)})=1\\ a^2b^2n\equiv l-a^2v_0(k)}} 1 \\ & + \left\{ \sum_{\substack{a,b\\ x^{2\theta}\leqq a^2b^2k\leqq x^{\theta}\\ (ab,k)=1}} + \sum_{a\leqq\sqrt{\frac{x^{\theta}}{k}\frac{1}{\log x}}}\ \sum_{b\leqq\sqrt{\frac{x^{\theta}}{k}\frac{1}{\log x}}} \right\}_{\substack{(ab,k)=1\\ a^2b^2k\geqq x^{\theta}}}\ \sum_{\substack{\frac{x-x^{\theta}}{a^2b^2}-\frac{v_0}{b^2}<n\leqq\frac{x}{a^2b^2}-\frac{v_0}{b^2}\\ a^2b^2n\equiv l-a^2v_0(k)}} 1 \\ & + \left\{ \sum_{\sqrt{\frac{x^{\theta}}{k}\frac{1}{\log x}}\leqq a\leqq\sqrt{x}}\ \sum_{(ab,2\mathscr{P}_1)=1} + \sum_{\sqrt{\frac{x^{\theta}}{k}\frac{1}{\log x}}\leqq b\leqq\sqrt{x+1}}\ \sum_{(ab,2\mathscr{P}_1)=1} \right\}_{\substack{\frac{x-x^{\theta}}{a^2}<v\leqq\frac{x}{a^2}\\ a^2v+1=b^2w\\ (vw,\mathscr{P}_3)=1}}\ \sum_{v,w} 1 \\ := & \sum{}^{(1)}+\sum{}^{(2)}+\sum{}^{(3)}+\sum{}^{(4)}+\sum{}^{(5)}. \end{aligned} \tag{3.3}$$

Zu $\sum^{(1)}$. Seien $\overline{a^2}$ bzw. $\overline{b^2}$ Lösungen der Kongruenzen

$$a^2\overline{a^2}\equiv 1(k) \quad \text{bzw.} \quad b^2\overline{b^2}\equiv 1(k), \tag{3.4}$$

dann hat ein in $\sum^{(1)}$ gezähltes n die Form $n=\overline{b^2}(\overline{a^2}\,l-v_0)+km$. Mit

$$M(a,b)=\{(c_1m+d_1)(c_2m+d_2)\colon x_0-y_0<m\leqq x_0\}, \tag{3.5}$$

wobei

$$x_0=\frac{x}{a^2b^2k}-\frac{v_0+b^2\overline{b^2}(\overline{a^2}\,l-v_0)}{b^2k}, \qquad y_0=\frac{x^\theta}{a^2b^2k}, \qquad c_1=b^2k, \qquad c_2=a^2k,$$

$$d_1=b^2\overline{b^2}(\overline{a^2}\,l-v_0)+v_0 \quad \text{und} \quad d_2=a^2\overline{b^2}(\overline{a^2}\,l-v_0)+w_0,$$

folgt für $z\geqq 1$ nach Definition der Siebfunktion in (2.1)

$$\sum^{(1)}\leqq \sum_{\substack{a,b\\ 1\leqq a^2b^2k\leqq x^{\frac{5}{6}\theta}\\ (a,b)=1\\ (ab,2k\mathscr{P}_1^{(k)})=1}} S(M(a,b),\mathscr{P}_3^{(k)},z). \tag{3.6}$$

Zu $\sum^{(2)}$ *und* $\sum^{(3)}$. $\sum^{(2)}$ läßt sich abschätzen mit

$$\sum^{(2)}\leqq \sum_{\substack{a,b\\ x^{\frac{5}{6}\theta}\leqq a^2b^2k\leqq x^\theta\\ (ab,k)=1}} \sum_{x_0-\frac{x^\theta}{a^2b^2k}<m\leqq x_0} 1\ll\frac{x^\theta}{k}\sum_{d\geqq x^{\frac{5}{6}\theta}k^{-\frac{1}{2}}}\frac{\tau(d)}{d^2}\ll\frac{x^\theta}{k\log^2 x}. \tag{3.7}$$

Für $\sum^{(3)}$ ergibt sich wegen $\dfrac{x^\theta}{a^2b^2k}\leqq 1$ sofort

$$\sum^{(3)}\ll \sum_{a\leqq\sqrt{\frac{x^\theta}{k}}\frac{1}{\log x}} \sum_{b\leqq\sqrt{\frac{x^\theta}{k}}\frac{1}{\log x}} 1\ll\frac{x^\theta}{k\log^2 x}. \tag{3.8}$$

Zu $\sum^{(4)}$. Wegen der Eindeutigkeit der Primfaktorzerlegung gibt es zu gegebenem a und v höchstens ein Paar b, w mit $a^2v+1=b^2w$, $(b,2\mathscr{P}_1)=1$, $(w,\mathscr{P}_3)=1$, so daß

$$\begin{aligned}\sum^{(4)}&\leqq\left\{\sum_{\sqrt{\frac{x^\theta}{k}}\frac{1}{\log x}\leqq a\leqq x^{\frac{1}{2}-\frac{1}{2}\theta}k\log^2 x}+\sum_{x^{\frac{1}{2}-\frac{1}{2}\theta}k\log^2 x\leqq a\leqq\sqrt{x}}\right\}\sum_{\frac{x-x^\theta}{a^2}<v\leqq\frac{x}{a^2}}1\\ &:=\sum^{(41)}+\sum^{(42)}.\end{aligned} \tag{3.9}$$

Für $\sum^{(42)}$ erhalten wir durch Vertauschung der Summationen

$$\sum^{(42)}\leqq \sum_{v\leqq\frac{x^\theta}{k^2\log^4 x}} \sum_{\sqrt{\frac{x-x^\theta}{v}}<a\leqq\sqrt{\frac{x}{v}}} 1\leqq \sum_{v\leqq\frac{x^\theta}{k^2\log^4 x}}(1+x^{\theta-\frac{1}{2}}v^{-\frac{1}{2}})\ll\frac{x^\theta}{k\log^2 x}. \tag{3.10}$$

Diese Abschätzung gilt auch für $\sum^{(41)}$, falls $\theta>\frac{1}{3}$ und $x\geqq x_0(\theta, A)$ ist, denn

$$\sum^{(41)}\leqq \sum_{\sqrt{\frac{x^\theta}{k}}\frac{1}{\log x}\leqq a\leqq x^{\frac{1}{2}-\frac{1}{2}\theta}k\log^2 x}\left(1+\frac{x^\theta}{a^2}\right)\ll\frac{x^\theta}{k\log^2 x}. \tag{3.11}$$

Falls $\theta\leqq\frac{1}{3}$ ist, führen wie einen Parameter u mit $\theta\leqq u\leqq 1$ ein und schätzen $\sum^{(41)}$ für $x\geqq x_0(\varepsilon, A)$ ab durch

$$\sum^{(41)}\leqq \sum_{\sqrt{\frac{x^\theta}{k}}\frac{1}{\log x}\leqq a\leqq x^{\frac{1}{2}(1-u)}}\ \sum_{\frac{x-x^\theta}{a^2}<v\leqq\frac{x}{a^2}} 1+\sum_{x^{\theta-\varepsilon}\leqq v\leqq x^u}\ \sum_{\sqrt{\frac{x-x^\theta}{v}}<a\leqq\sqrt{\frac{x}{v}}} 1. \tag{3.12}$$

Da sich $\sum^{(5)}$ analog abschätzen läßt, erhalten wir aus (3.3) und (3.6)–(3.12)

Lemma 2. *Sei* $\varepsilon>0$, $z\geqq 1$, $\theta\leqq u\leqq 1$, $1\leqq k\leqq(\log x)^A$, $(k, l(l+1))=1$,

$$\chi_{\frac{1}{3}}(\theta)=\begin{cases}1 & \text{falls } \theta\leqq\frac{1}{3}\\ 0 & \text{falls } \theta>\frac{1}{3},\end{cases}$$

dann existiert ein $x_0(\theta, A, \varepsilon)$, *so daß für* $x\geqq x_0(\theta, A, \varepsilon)$

$$\begin{aligned}B_2(x,k,l)-B_2(x-x^\theta,k,l)\leqq &\sum_{\substack{a,b\\ 1\leqq a^2b^2k\leqq x^{\frac{2}{3}\theta}\\ (a,b)=1\\ (ab,2k\mathscr{P}_1^{(k)})=1}} S(M(a,b),\mathscr{P}_3^{(k)},z)\\ +\chi_{\frac{1}{3}}(\theta)\Bigg\{&\sum_{x^{\theta-\varepsilon}\leqq v\leqq x^u}\ \sum_{\sqrt{\frac{x-x^\theta}{v}}<a\leqq\sqrt{\frac{x}{v}}} 1+\sum_{x^{\theta/2-\varepsilon}\leqq a\leqq x^{\frac{1}{2}-\frac{1}{2}u}}\ \sum_{\substack{x-x^\theta<n\leqq x\\ n\equiv O(a^2)}} 1\\ +&\sum_{x^{\theta-\varepsilon}\leqq w\leqq x^u}\ \sum_{\sqrt{\frac{x+1-x^\theta}{w}}<b\leqq\sqrt{\frac{x+1}{w}}} 1+\sum_{x^{\theta/2-\varepsilon}\leqq b\leqq x^{\frac{1}{2}-\frac{1}{2}u}}\ \sum_{\substack{x+1-x^\theta<n\leqq x+1\\ n\equiv O(b^2)}} 1\Bigg\}\\ +O\left(\frac{x^\theta}{k\log^2 x}\right)&:=HT+R_1+R_2+S_1+S_2+O\left(\frac{x^\theta}{k\log x}\right).\end{aligned} \tag{3.13}$$

Bemerkung. Die Abschätzung des Haupttermes HT im nächsten Abschnitt mit Hilfe von Lemma 1, verläuft, abgesehen von der Übertragung auf Restklassen, vollkommen analog zur Arbeit von Kelly. Die Fehlerterme R_i bzw. S_i tauchen nur für $\theta\leqq\frac{1}{3}$ auf, d.h. kommen bei Kelly nicht vor. Offensichtlich können wir uns bei diesen Fehlertermen auf die Abschätzung von R_1 und R_2 beschränken.

Mit Hilfe einer neuartigen Technik von Iwaniec (vgl. [7] und [8]) bzw. mit einem Analogon zum Erdös-Turan Theorem (vgl. [6], Lemma 6) transformieren wir R_1 und R_2 auf Exponentialsummen, die wir dann mit den klassischen Van der Corput-Methoden behandeln.

4. Abschätzungen des Haupttermes HT

Wir zeigen zunächst

Lemma 3. *Sei* $a^2b^2k \leqq x^{\frac{9}{10}\theta}$, $(ab, 2k\mathscr{P}_1^{(k)})=1$, *dann gilt für*

$$z=x^{\frac{1}{20}\theta} \quad \textit{und} \quad \omega^*(p)=\begin{cases}2 & \textit{falls } p\in\mathscr{P}_3\\ 0 & \textit{sonst}\end{cases}$$

$$S(M(a,b),\mathscr{P}_3^{(k)},z)$$
$$\leqq \frac{2}{ab\phi(a)\phi(b)} \prod_{\substack{p\equiv 3(4)\\ p\mid k}} \frac{p}{p-2} \prod_p \left(1-\frac{\omega^*(p)}{p}\right)\frac{p}{p-1}\prod_p\left(1+\frac{1}{p(p-2)}\right)$$
$$\cdot \frac{x^\theta}{k\log z}\left(1+O\left(\frac{1}{(\log x)^{\frac{1}{3}}}\right)\right),$$

wobei die O-Konstante von k unabhängig ist.

Beweis. Die zu siebende Menge $M(a,b)$, vgl. (3.5), ist ein Spezialfall von [2], Kap. 1, Beisp. 4; wir wählen also $X=y_0=\dfrac{x^\theta}{a^2b^2k}$.

Wegen $c_1c_2(c_1d_2-c_2d_1)=a^2b^2k^3$ ist die Wahl

$$\omega(p)=\begin{cases}2 & \text{falls } p\in\mathscr{P}_3^{(k)}, p\nmid ab\\ 1 & \text{falls } p\in\mathscr{P}_3^{(k)}, p\mid ab\\ 0 & \text{sonst}\end{cases} \tag{4.1}$$

sinnvoll. Für $d\in Q(\mathscr{P}_3^{(k)})$ ist

$$|M(a,b)_d|=\omega(d)\left(\frac{y_0}{d}+\delta\right), \quad |\delta|\leqq 1$$

und damit $|r(M(a,b),d)|\leqq 2$ - vgl. (2.2) und (2.3). - Wir erhalten also aus Lemma 1 für $D=z=x^{\frac{1}{20}\theta}$ - *vgl.* [2], *Lemma* 2.3 -

$$S(M(a,b),\mathscr{P}_3^{(k)},z)\leqq 2e^\gamma XV(z)\left(1+O\left(\frac{1}{(\log x)^{\frac{1}{3}}}\right)\right). \tag{4.2}$$

Für $V(z)$ - vgl. (2.4) - folgt aus (4.1) wegen $(ab,k)=1$

$$V(z)=\prod_{\substack{p<z\\ p\equiv 3(4)\\ p\nmid kab}}\left(1-\frac{2}{p}\right)\prod_{\substack{p<z\\ p\equiv 3(4)\\ p\mid ab}}\left(1-\frac{1}{p}\right)$$
$$\leqq \prod_{\substack{p<z\\ p\equiv 3(4)}}\left(1-\frac{2}{p}\right)\prod_{p\mid ab}\frac{p-1}{p-2}\prod_{\substack{p\equiv 3(4)\\ p\mid k}}\frac{p}{p-2}.$$

Nach [2], Lemma 5.3 ist

$$\prod_{\substack{p<z\\ p\equiv 3(4)}}\left(1-\frac{2}{p}\right)=\prod_p\left(1-\frac{\omega^*(p)}{p}\right)\frac{p}{p-1}\frac{e^{-\gamma}}{\log z}\left(1+O\left(\frac{1}{\log z}\right)\right).$$

Wegen

$$\prod_{p|ab}\frac{p-1}{p-2}=\frac{ab}{\phi(ab)}\prod_{p|ab}\frac{(p-1)^2}{p(p-2)}$$

gilt also

$$V(z)\leqq\frac{ab}{\phi(ab)}\prod_{\substack{p\equiv 3(4)\\ p|k}}\frac{p}{p-2}\prod_{p}\left(1-\frac{\omega^*(p)}{p}\right)\frac{p}{p-1}$$

$$\cdot\prod_{p}\left(1+\frac{1}{p(p-2)}\right)\frac{e^{-\gamma}}{\log z}\left(1+O\left(\frac{1}{\log z}\right)\right).$$

Durch Einsetzen dieser Abschätzung in (4.2) folgt sofort die Behauptung von Lemma 3.

Mit Lemma 3 erhalten wir für den Hauptterm HT - vgl. (3.13) -

$$\mathrm{HT}=\sum_{\substack{a,b\\ 1\leqq a^2b^2k\leqq x^{\frac{2}{5}\theta}\\ (ab,2k\mathscr{P}_1^{(k)})=1\\ (a,b)=1}} S(M(a,b),\mathscr{P}_3^{(k)},z)\ll\frac{x^\theta}{k\log x}\prod_{\substack{p\equiv 3(4)\\ p|k}}\frac{p}{p-2}\sum_{d}\frac{\tau(d)}{d\phi(d)}$$

und folglich

$$\mathrm{HT}\ll\prod_{\substack{p\equiv 3(4)\\ p|k}}\frac{p}{p-2}\frac{x^\theta}{k\log x}. \tag{4.3}$$

5. Abschätzungen von R_1

In diesem Abschnitt zeigen wir

Lemma 4. *Sei* $\frac{1}{5}+\frac{6}{5}\varepsilon\leqq\theta\leqq 1$, $\theta\leqq u\leqq\min(\frac{2}{3}\theta+\frac{1}{3},6\theta-1)-6\varepsilon$, *dann gilt*

$$R_1=\sum_{x^{\theta-\varepsilon}\leqq v\leqq x^u}\sum_{\sqrt{\frac{x-x^\theta}{v}}<a\leqq\sqrt{\frac{x}{v}}}1\ll x^{\theta-\varepsilon}.$$

Offensichtlich gilt mit $\psi(x)=x-[x]-\frac{1}{2}$

$$R_1=\sum_{x^{\theta-\varepsilon}\leqq v\leqq x^u}\left(\left[\sqrt{\frac{x}{v}}\right]-\left[\sqrt{\frac{x-x^\theta}{v}}\right]\right)$$

$$=\sum_{x^{\theta-\varepsilon}\leqq v\leqq x^u}\left(\sqrt{\frac{x}{v}}-\sqrt{\frac{x-x^\theta}{v}}\right)+O\left(\log x\max_{x^{\theta-\varepsilon}\leqq N\leqq x^u}\left\{\left|\sum_{N\leqq v<2N}\psi\left(\sqrt{\frac{x}{v}}\right)\right.\right.\right.$$

$$\left.\left.\left.+\left|\sum_{N\leqq v<2N}\psi\left(\sqrt{\frac{x-x^\theta}{v}}\right)\right|\right\}\right):=R_{11}+R_{12}. \tag{5.1}$$

Für R_{11} erhalten wir sofort

$$R_{11}\ll x^{\theta-\frac{1}{2}}\sum_{v\leqq x^u}\frac{1}{\sqrt{v}}\ll x^{\theta-\varepsilon}. \tag{5.2}$$

Zur Abschätzung von R_{12} verwenden wir ein Analogon zum Erdös-Turan Theorem:

Lemma 5 (vgl. [6] Lemma 6). *Sei (x_n) eine Folge reeller Zahlen, dann gilt mit einer absoluten O-Konstanten*

$$\left|\sum_{z_1 \leqq n \leqq z_2} \psi(x_n)\right| \ll \frac{z_2 - z_1}{Q} + \sum_{1 \leqq q \leqq Q} \frac{1}{q} \left|\sum_{z_1 \leqq n \leqq z_2} e(q x_n)\right|,$$

wobei, wie üblich, $e(x) := e^{2\pi i x}$ sei.

Für die Abschätzung der bei der Anwendung von Lemma 5 auf R_{12} auftretenden Exponentialsummen benötigen wir

Lemma 6

$$\sum_{N \leqq v < 2N} e\left(\frac{q\sqrt{\zeta}}{\sqrt{v}}\right) \ll \begin{cases} \dfrac{N^{\frac{3}{2}}}{q\sqrt{\zeta}} & \text{falls } q\sqrt{\zeta} \leqq N^{\frac{3}{2}} \\ (q\sqrt{\zeta})^{\frac{1}{2}} N^{-\frac{1}{4}} & \text{falls } q\sqrt{\zeta} \gg N^{\frac{3}{2}}. \end{cases}$$

Die Abschätzung für $q\sqrt{\zeta} \leqq N^{\frac{3}{2}}$ erhalten wir aus [13] Lemma 4.2 und Lemma 4.8; und die restliche Abschätzung aus [10] bzw. [11], mit dem Exponentenpaar $(\frac{1}{2}, \frac{1}{2})$.

Aus Lemma 5 und Lemma 6 erhalten wir mit $Q = \dfrac{N}{x^{\theta-\varepsilon}}$ und $\zeta = x$ bzw. $\zeta = x - x^{\theta}$ sofort

$$\left|\sum_{N \leqq v < 2N} \psi\left(\sqrt{\frac{\zeta}{v}}\right)\right| \ll x^{\theta-\varepsilon} + N^{\frac{3}{2}} \zeta^{-\frac{1}{2}} + Q^{\frac{1}{2}} \zeta^{\frac{1}{4}} N^{-\frac{1}{4}} \ll x^{\frac{3}{2}u - \frac{1}{2}} + x^{\frac{1}{4}u + \frac{1}{4} - \frac{1}{2}\theta + \frac{1}{2}\varepsilon} + x^{\theta-\varepsilon}.$$

Für

$$u \leqq \min(\tfrac{2}{3}\theta + \tfrac{1}{3}, 6\theta - 1) - 6\varepsilon \tag{5.3}$$

folgt hieraus

$$R_{12} \ll x^{\theta-\varepsilon},$$

und damit zusammen mit (5.2) die Behauptung von Lemma 4.

Abschätzung von R_2. In diesem Abschnitt zeigen wir

Lemma 7. *Sei $u = \min(\frac{2}{3}\theta + \frac{1}{3}, 6\theta - 1) - 6\varepsilon$, $\theta \geqq \frac{75}{307} + 2\varepsilon$, dann gilt*

$$R_2 = \sum_{x^{\theta/2-\varepsilon} \leqq a \leqq x^{\frac{1}{2}-\frac{1}{2}u}} \sum_{\substack{x - x^{\theta} < n \leqq x \\ n \equiv 0 (a^2)}} 1 \ll x^{\theta-\varepsilon}.$$

Zur Abschätzung von R_2 verwenden wir die Poisson'sche Summenformel in der Form von

Lemma 8 (vgl. [7], Lemma 1). *Sei $0 < y_1 < y_2$ und $f(x)$ eine auf $\mathbb{R}$ einmal stetig differenzierbare Funktion mit $f(y_1) = f(y_2)$, dann gilt*

$$\sum_{\substack{y_1 < n \leqq y_2 \\ n \equiv l (q)}} f(n) = \frac{1}{q} \sum_{h=-\infty}^{\infty} e\left(-h\frac{l}{q}\right) \int_{y_1}^{y_2} f(t) e\left(\frac{ht}{q}\right) dt.$$

Zur Anwendung dieses Lemmas ersetzen wir in R_2 die charakteristische Funktion des Intervalls $[x-x^\theta, x]$ durch eine auf $\mathbb{R}$ beliebig oft differenzierbare Funktion $f(x)$ mit folgenden Eigenschaften (vgl. [7] bzw. [8])

$$\begin{aligned} &\text{(i) } f(u)=\begin{cases}1 & \text{für } x-x^\theta<u\leqq x\\ 0 & \text{für } u\geqq x+x^\theta \text{ bzw. } u\leqq x-2x^\theta\end{cases}\\ &\text{(ii) } 0\leqq f(u)\leqq 1\\ &\text{(iii) } f^{(q)}(u)\ll(x^\theta)^{-q} \quad \text{für alle } q\geqq 0, \text{ wobei die } O\text{-Konstante}\\ &\qquad\qquad\qquad\qquad\quad \text{höchstens von } q \text{ abhängt.}\end{aligned} \tag{6.1}$$

Damit erhalten wir

$$R_2\ll\log x \max_{x^{\theta/2-\varepsilon}\leqq N\leqq x^{\frac{1}{2}-\frac{1}{2}u}} \sum_{N\leqq a<2N} \sum_{\substack{m\\ m\equiv 0(a^2)}} f(m). \tag{6.2}$$

Lemma 8 auf die innerste Summe angewendet, ergibt

$$\sum_{m\equiv 0(a^2)} f(m)=\frac{1}{a^2}\sum_{h=-\infty}^{\infty}\int f(t)\,e\left(\frac{ht}{a^2}\right)dt. \tag{6.3}$$

Der Term $h=0$ liefert zu R_2 den Beitrag

$$\ll x^\theta\log x\sum_{a\geqq x^{\theta/2-\varepsilon}}\frac{1}{a^2}\ll x^{\theta-\varepsilon},$$

d.h. er geht in der O-Abschätzung von Lemma 7 auf.

Wegen (6.1) erhalten wir durch partielle Integration

$$\int f(t)\,e\left(\frac{ht}{a^2}\right)dt=\left(\frac{2\pi i h}{a^2}\right)^{-q}\int f^{(q)}(t)\,e\left(\frac{ht}{a^2}\right)dt\ll x^\theta\left(\frac{a^2}{|h|x^\theta}\right)^q$$

und damit mit $H=N^2x^{-\theta+\varepsilon}$

$$\frac{1}{a^2}\sum_{|h|\geqq H}\left|\int f(t)\,e\left(\frac{ht}{a^2}\right)dt\right|\ll\left(\frac{N^2}{Hx^\theta}\right)^{q-1}\ll x^{-\varepsilon(q-1)}.$$

Für genügend großes q liefert also auch dieser Teil der Summe auf der rechten Seite von (6.3) zu R_2 einen Beitrag, der in der O-Abschätzung von Lemma 7 aufgeht. Zum Beweis von Lemma 7 genügt es somit wegen (6.1) zu zeigen:

$$\begin{aligned} R_2^*:=\frac{1}{N^2}\sum_{1\leqq h\leqq H}\sum_{N\leqq a<N'} e\left(\frac{ht}{a^2}\right)\ll x^{-\varepsilon} \quad &\text{für alle } N\in[x^{\theta/2-\varepsilon}, x^{\frac{1}{2}-\frac{1}{2}u}]\\ &\text{für alle } N'\in[N,2N]\\ &\text{für alle } t\in[x-2x^\theta, x+x^\theta].\end{aligned} \tag{6.4}$$

Mit der Theorie der Exponentenpaare erhalten wir analog zu Lemma 6 aus [10], [11] bzw. ([13], Lemma 4.2 und 4.8):

Lemma 9. *Sei* (K,L) *ein Exponentenpaar,* $0 \leqq K \leqq \frac{1}{2} \leqq L \leqq 1$, *dann gilt*

$$\left| \sum_{N \leqq a \leqq N'} e\left(\frac{ht}{a^2}\right) \right| \ll \begin{cases} \dfrac{N^3}{ht} & \text{falls } ht \leqq \dfrac{1}{2+\varepsilon} N^3 \\ (ht)^K N^{L-3K} & \text{falls } ht \gg N^3. \end{cases}$$

Wegen $N \leqq x^{\frac{1}{2}-\frac{1}{2}u}$ *folgt hieraus sofort mit* $H = N^2 x^{-\theta+\varepsilon}$

$$\begin{aligned} R_2^* &\ll x^{\varepsilon}(N x^{-1} + H^{K+1} x^K N^{L-3K-2}) \\ &\ll x^{-\frac{1}{2}-\frac{1}{2}u+\varepsilon} + x^{\frac{1}{2}(L+K)-\frac{1}{2}u(L-K)-\theta(K+1)+\varepsilon(K+2)}. \end{aligned}$$

Mit dem Exponentenpaar $(\frac{4}{49}, \frac{75}{98})$ und der in Lemma 7 angegebenen Wahl von u folgt die Behauptung (6.4), falls $\theta \geqq \frac{75}{307} + 2\varepsilon$ ist. Das Exponentenpaar $(\frac{4}{49}, \frac{75}{98})$ erhalten wir mit der Operatorfolge $ABABA^3(\frac{1}{2}, \frac{1}{2})$; dabei sei wie üblich BA^q der durch [10] Theorem 4 und A der durch [10] Theorem 1 beschriebene Operator (vgl. auch [11]).

Der θ-Bereich läßt sich durch andere Operatorenfolgen noch geringfügig vergrößern.

Der Beweis von Satz 1 folgt nun sofort aus Lemma 2, (4.3), Lemma 4 und Lemma 7.

Literatur

1. Bantle, G.: Untere Abschätzung für die Anzahl der B-Zwillinge auf kurzen Intervallen, to appear in Acta Arith. **45** (1985)
2. Halberstam, H., Richert, H.-E.: Sieve Methods. London: Academic Press 1974
3. Hooley, C.: On the intervals between numbers, that are sums of two squares III. J. für Reine Angew. Math. **267**, 207–218 (1974)
4. Indlekofer, K.H.: Scharfe untere Abschätzung für die Anzahlfunktion der B-Zwillinge. Acta Arith. **26**, 207–212 (1974)
5. Iwaniec, H.: Rosser's sieve. Acta Arith. **36**, 171–202 (1980)
6. Iwaniec, H.: On the Brun Titchmarsh theorem. J. Math. Soc. Japan **34**, 95–122 (1982)
7. Iwaniec, H., Deshouillers, J.M.: An additive divisor problem. J. Lond. Math. Soc. II. Sèr. **26**, 1–14 (1982)
8. Iwaniec, H., Laborde, M.: P_2 in short intervals. Ann. Inst. Fourier Grenoble **31** (4), 37–56 (1981)
9. Kelly, P.J.: The number of B-twins in an interval. Dissertation, Nottingham, 1978
10. Phillips, E.: The zeta-function of Riemann; further developments of van der Corput's method. Quart. J. Math. Oxford **4**, 209–225 (1933)
11. Rankin, R.A.: Van der Corput's Method and the theory of exponent pairs. Quart. J. Math. Oxford (2), **6**, 147–153 (1955)
12. Rieger, G.J.: Aufeinanderfolgende Zahlen als Summe von 2·Quadraten. Indag. Math. **27**, 208–220 (1965)
13. Titchmarsh, E.C.: The theory of the Riemann Zeta Function. Oxford: Clarendon Press
14. Wolke, D.: Multiplikative Funktionen auf schnell wachsenden Folgen. J. Reine Angew. Math., **251**, 54–67 (1971)

Eingegangen am 20. Oktober 1983;
eingegangen in endgültiger Form am 21. Dezember 1984

Math. Z. 189, 571–592 (1985)

Mathematische Zeitschrift
© Springer-Verlag 198

QF-1 Algebras of Local-Colocal Type

Ryohei Makino

Department of Mathematics, Fukushima University, Fukushima 960-12, Japan

Introduction

As a generalization of QF algebras, Thrall [14] defined QF-1 algebras as follows: For a finite dimensional algebra A over a field, A is said to be QF-1 if every faithful representation of A coincides with its own second commutator. He proposed to give an internal characterization of QF-1 algebras, but this problem remains unsolved at the present time. Of course there are results which are partial solutions of Thrall's problem in special cases (Camillo and Fuller [2], Dlab and Ringel [3], Fuller [4], the auther [7] and Ringel [9]), and many other results concerning QF-1 algebras. However, what is worth noting, it seems that QF-1 algebras which are not QF obtained from these results are all algebras of local-colocal type (i.e. algebras each of whose indecomposable modules has either a unique maximal submodule or a unique minimal submodule). For example, Ringel [9] characterized QF-1 algebras with radical square zero as those of local-colocal type satisfying the socle conditions. Tachikawa [13] proved that left serial QF-1 algebras are of left colocal type (i.e. algebras each of whose indecomposable left modules has a unique minimal submodule), and by adding some conditions the auther [7] characterized left serial QF-1 algebras. In this paper we shall give a solution of Thrall's problem to the case of algebras of local-colocal type. Our result contains those mentioned above as special cases.

Let M be a left module over a ring A, and K, L submodules of M. Suppose that the following conditions are satisfied:

(1) $M=K+L$;

(2) Both K and L are serial modules with $|K|, |L| \geqq 2$, where K denotes the composition length of K (a module is said to be serial if its composition series is unique);

(3) $K \cap L$ coincides with the socle $S(M)$ of M.

Then we shall denote the sum $M=K+L$ by $K \nabla L$. We shall say a module M to be of type I in case $M=K \nabla L$ for suitable submodules K and L of M. By Tachikawa's theorems in [10] and [11] it is known that if A is a finite dimen-

sional algebra of local-colocal type over a field P with the radical N, then indecomposable left A-modules M are classified into the next four types;

(1) type I,

(2) the P-dual module $\mathrm{Hom}_P(M, P)$ is of type I,

(3) the top M/NM of M is simple and the radical NM of M is of type I,

(4) serial.

We shall call indecomposable A-modules M satisfying (2) and (3) to be of type II and type III, respectively. Now we can state the following main theorem: Let A be a finite dimensional algebra of local-colocal type over a field P, N the radical of A, e a primitive idempotent and K, L serial submodules of Ae. Then in order that A may be QF-1 it is necessary and sufficient that the following conditions are satisfied.

(1) If Ae is serial and Ne is projective, then every composition factor of Ae is isomorphic to a minimal left ideal. And the same holds for a primitive right ideal eA that is serial.

(2) If Ae is of type II, then every composition factor of Ae is isomorphic to a minimal left ideal. And the same holds for a primitive right ideal eA that is of type II.

(3) Let Ae be of type III and $Ne = K \nabla L$. If K is projective, then every composition factor of $K \nabla S^2(L)$ is isomorphic to a minimal left ideal, where $S^2(L)$ is the second socle of L. And the same holds for a primitive right ideal eA that is of type III.

(4) Let Ae be of type II and $Ne = K \oplus L$. If the P-dual module $(Ae/K)^* = \mathrm{Hom}_P(Ae/K, P)$ is projective, then K^* is embedded into a primitive right ideal. And the same holds for a primitive right ideal eA that is of type II.

(5) Let Ae be of type II and $Ne = K \oplus L$. Let f be a primitive idempotent such that $Af/Nf \simeq K/NK$ and $\phi: eA \to fN$ the projective cover of fN. If $(Ae/K)^*$ is projective, $(Ae/L)^*$ is nonprojective and eA is of type II, then the kernel $\mathrm{Ker}(\phi)$ of ϕ is nonserial. And the same holds for a primitive right ideal eA that is of type II.

It should be noted that each condition in this characterization can be checked if the ideal-structure of A is known. In [7] the argument is based on the balancedness of a minimal faithful projective module. On the other hand, in this paper a normal minimal faithful module takes a place of a minimal faithful projective module. Here, a normal minimal faithful module is a direct sum of all modules of type III and all serial modules which are maximal relative to a certain order (see Sect. 2 for the precise definition of a normal minimal faithful module). The condition (5) in the characterization is concerned with the balancedness of a normal minimal faithful module.

Tachikawa [12] proved that if A is a maximal left quotient ring (equivalently the injective hull $E({}_AA)$ is balanced), then the $\mathscr{S}$-injective hull $E_{\mathscr{S}}({}_AA)$ is balanced for any strongly complete Serre class $\mathscr{S}$. Ringel [9] proved that a maximal left and right quotient algebra A of local-colocal type having square zero radical is QF-1. In its proof he showed that every minimal faithful left A-module is equivalent to $E_{\mathscr{S}}({}_AA)$ for some $\mathscr{S}$, and applied Tachikawa's result. In general, however, there exists a QF-1 algebra having a minimal faithful left module which is never equivalent to $E_{\mathscr{S}}({}_AA)$ for any $\mathscr{S}$ (cf. [7, Example 3.4]). Even if

M is a normal minimal faithful left module over a QF-1 algebra A of local-colocal type, it does not hold that M is equivalent to $E_{\mathscr{S}}({}_AA)$ for a suitable class $\mathscr{S}$. In Sect. 4 we shall prove that a normal minimal faithful left A-module is balanced if A is an algebra of local-colocal type satisfying several conditions. In the course of its proof we define a certain strongly complete Serre class ω. And applying Tachikawa's result we get the balanced module $E_\omega({}_AA)$, which is not equivalent to a normal minimal faithful module. So, to accomplish the proof we need more tools, one of which is Fuller's result in [5] (Lemma 4.8).

1. Preliminaries and Fundamental Properties of Modules over an Algebra of Local-Colocal Type

Throughout this paper A denotes a finite dimensional algebra with a unit element over a field P, and N the radical of A. All A-modules are unitary and finitely generated. Let M be a left A-module. By $S^i(M)$ we denote the annihilator of N^i in M, and call it the S^i-part of M. Especially $S(M)=S^1(M)$ is the socle of M. We write $T(M)=M/NM$, $E(M)$ for the injective hull of M, $|M|$ for the composition length of M and M^* for the P-dual module of M. In this paper homomorphisms will be always written on the side opposite the scalars.

A left A-module ${}_AM$ can be considered as a right B-module, where $B=\mathrm{End}({}_AM)$ is the endomorphism ring of ${}_AM$. We call $C=\mathrm{End}(M_B)$ the biendomorphism ring of ${}_AM$ and denote it by $\mathrm{BiEnd}({}_AM)$. If the canonical ring homomorphism from A into C is surjective, then ${}_AM$ is said to be balanced. A QF-1 algebra is precisely an algebra having the property that each of its finitely generated faithful modules is balanced.

Tachikawa [11] established the following structure theorem on an algebra of local-colocal type.

Theorem 1.1 [Tachikawa, 11]. *In order that an algebra A may be of local-colocal type it is necessary and sufficient that the following conditions are satisfied for all primitive idempotents e and f in A.*

(1) *Each of Ne and eN is either serial or a sum of two serial ideals;*

(2) *If Ne is a sum of two serial ideals and $T(Af)$ is embedded into a proper factor module of Ne, then fA is serial. And the same holds for a right ideal eN which is a sum of two serial ideals.*

From now on we shall assume throughout that A is an algebra of local-colocal type except in Lemmas 3.1, 3.7, 4.5, 4.7, 4.8 and 5.1. As mentioned in Introduction, it follows from Tachikawa's theorem in [11] that each of non-serial indecomposable A-modules has one of type I, type II and type III. If M is a left A-module of type II, then the radical NM of M is a direct sum of two serial modules. So we write $M|K\oplus L$ for the meaning of an indecomposable A-module M of type II with the radical $K\oplus L$, where K and L are serial submodules of M.

Lemma 1.2. *Let M and L be serial left A-modules with $|M|=|L|\geqq 2$. If $S(M)\simeq S(L)$ and $S^2(M)/S(M)\simeq S^2(L)/S(L)$ ($T(M)\simeq T(L)$ and $T(NM)\simeq T(NL)$), then $M\simeq L$.*

Proof. Let e be a primitive idempotent such that $T(Ae) \simeq S^2(M)/S(M) \simeq S^2(L)/S(L)$. If Ae is serial, obviously we have $S^2(M) \simeq S^2(L)$. If Ae is non-serial, then the injective hull of $S(M) \simeq S(L)$ is serial by the condition (2) of Theorem 1.1. Thus also in this case we have $S^2(M) \simeq S^2(L)$. The conclusion follows by [11, Proposition 4.1]. The parenthetical version is obtained by the duality.

Lemma 1.3 [Tachikawa, 11]. *For left A-modules* $M_1 = K_1 \nabla L_1$ *and* $M_2 = K_2 \nabla L_2$ $(M_1 | K_1 \oplus L_1$ *and* $M_2 | K_2 \oplus L_2)$, $M_1 \simeq M_2$ *if and only if* $\{K_1, L_1\} \simeq \{K_2, L_2\}$, *i.e.* $K_1 \simeq K_2$, $L_1 \simeq L_2$ *or* $K_1 \simeq L_2$, $L_1 \simeq K_2$.

Proof. "Only if" is obvious. For the proof of "if" see the proof of [11, Proposition 3.4] or [6, Theorem 2.3].

The next lemma follows at once from Theorem 1.1 and Lemma 1.2.

Lemma 1.4. (1) *For left A-modules* $K_1 \nabla L_1$ *and* $K_2 \nabla L_2$ *if* $K_1 \simeq K_2$, *then one of* L_1 *and* L_2 *is embedded into the other.*

(2) *For left A-modules* $M_1 | K_1 \oplus L_1$ *and* $M_2 | K_2 \oplus L_2$ *if* $M_1/K_1 \simeq M_2/K_2$, *then one of* M_1/L_1 *and* M_2/L_2 *is a homomorphic image of the other.*

Let M_i $(i = 1, 2)$ be left A-modules, and S_i $(i = 1, 2)$ submodules of M_i respectively. Assume that there exists an isomorphism $\theta: S_1 \to S_2$. Then we can construct a factor module $L = M_1 \oplus M_2/\{(m, -(m)\theta) | m \in S_1\}$. Following [10] we shall denote L by $\mathrm{Int}_\theta(M_1, M_2)$ and call it the interlacing module. It should be noted that an A-module $K \nabla L$ is isomorphic to $\mathrm{Int}_\theta(K, L)$, where θ is the identity on $K \cap L$.

Lemma 1.5. *For a left A-module* $K \nabla L$ *if* $S(K \nabla L)$ *is a homomorphic image of a serial primitive ideal Ae, then either K or L is projective.*

Proof. There exist primitive idempotents f and g such that K and L are homomorphic images of Af and Ag respectively. Then by Theorem 1.1 both Af and Ag are serial. Thus we have $K \simeq Af/N^k f$ and $L \simeq Ag/N^l g$ for some integers $k, l \geqq 2$. Since $K \nabla L$ has the simple socle $K \cap L$, by [10, Lemma 1.3] there exists an isomorphism $\theta: N^{k-1} f/N^k f \to N^{l-1} g/N^l g$ which can not be extended to an isomorphism $N^{k-2} f/N^k f \to N^{l-2} g/N^l g$. Suppose that both $N^k f$ and $N^l g$ are nonzero. Then we have $N^{k-1} f/N^{k+1} f \simeq N^{l-1} g/N^{l+1} g \simeq Ae/N^2 e$, since Ae is serial. Therefore θ can be lifted to an isomorphism $\sigma: N^{k-1} f/N^{k+1} f \to N^{l-1} g/N^{l+1} g$. We can easily see that the interlacing module $\mathrm{Int}_\sigma(Af/N^{k+1} f, Ag/N^{l+1} g)$ is a nonserial module with the serial S^2-part. But the existence of such a module contradicts [11, Proposition 2.5].

Lemma 1.6. *Let X be a serial left A-module. Then*

(1) *If X is embedded into a left A-module* $K \nabla L$, *then X is embedded into either K or L;*

(2) *If X is a homomorphic image of a left A-module* $M | K \oplus L$, *then X is a homomorphic image of either* M/K *or* M/L.

Proof. It suffices only to show (1), because (1) and (2) are dual statements of each other. Obviously $T(X) = S^n(X)/S^{n-1}(X)$ is embedded into $S^n(K \nabla L)/$

$S^{n-1}(K \nabla L) \simeq S^n(K)/S^{n-1}(K) \oplus S^n(L)/S^{n-1}(L)$, where $n=|X|$. If $T(X) \simeq S^n(K)/S^{n-1}(K)$, then we have $X \simeq S^n(K)$. For, the projective cover of $S^n(K)/S^{n-1}(K)$ is serial by Theorem 1.1.

2. Minimal Faithful Modules

In this section, similarly as [7, Lemma 1.3] we shall determine the form of minimal faithful A-modules. Here a module M is said to be minimal faithful if M is faithful and any proper direct summand of M is unfaithful.

By $\mathfrak{S}_l$ let us denote the class of all serial left A-modules and all left A-modules of type III, and let $\bar{\mathfrak{S}}_l$ be the totality of the isomorphism classes of modules belonging to $\mathfrak{S}_l$. We define an order $\leqq$ on $\bar{\mathfrak{S}}_l$ as follows: For $X, Y \in \mathfrak{S}_l$

$$\bar{X} \leqq \bar{Y} \quad \text{if and only if } X \text{ is isomorphic to a subquotient of } Y,$$

where $\bar{X}$ and $\bar{Y}$ are isomorphism classes of X and Y, respectively. By $\mathfrak{N}_l$ we shall denote the totality of left A-modules in $\mathfrak{S}_l$ whose isomorphism classes are maximal with respect to this order $\leqq$. We observe that every left A-module M of type III belongs to $\mathfrak{N}_l$, since M is both projective and injective by [11, Proposition 3.1]. Similarly we define $\mathfrak{N}_r$ for right A-modules. It is obvious that $M \in \mathfrak{N}_l$ if and only if $M^* \in \mathfrak{N}_r$ for a left A-module M.

Lemma 2.1. *Every left A-module M in $\mathfrak{N}_l$ is either projective or injective.*

Proof. Suppose that M is not projective. Then M is serial. If the projective cover Ae of M is serial, then we have $Ae \simeq M$ by the maximality of M. This is a contradiction. It follows from Theorem 1.1 that the injective hull $E(M)$ is serial. Then again by the maximality of M we obtain $M \simeq E(M)$.

Lemma 2.2. *For each left A-module M in $\mathfrak{N}_l$ there exists an element a in A such that $aM \neq 0$ and $aX=0$ if $X \in \mathfrak{S}_l$ and $X \not\simeq M$.*

Proof. At first assume $M \simeq Ae$, where e is a primitive idempotent. Take a non-zero element a of $S(Ae)$. Suppose $aY \neq 0$ for some $Y \in \mathfrak{S}_l$, and let $y \in Y$ be an element such that $ay \neq 0$. Then the homomorphism θ: $Ae \to Y$ defined by $(be)\theta = bey$ $(be \in Ae)$ is a monomorphism, since the socle $S(Ae)$ is simple. It follows $Ae \simeq Y$ by the maximality. As a consequence we conclude that a is a required element. Next if M is not projective, then M^* is projective by Lemma 2.1. As M^* belongs to $\mathfrak{N}_r$, the same holds for M^*. Considering the fact that the annihilator of an A-module equals to that of its P-dual module, we obtain the desired conclusion.

Proposition 2.3. *Let*

$$M = \left(\bigoplus_{i=1}^{r} X_i | K_i \oplus L_i\right) \oplus \left(\bigoplus_{j=r+1}^{s} M_{2j-1} \nabla M_{2j}\right) \oplus \left(\bigoplus_{k=2s+1}^{t} M_k\right)$$

be a left A-module, where $M_k (2s+1 \leqq k \leqq t)$ are either serial or of type III. Put $M_{2i-1} = X_i/K_i$ and $M_{2i} = X_i/L_i$ for all $1 \leqq i \leqq r$. Then M is minimal faithful if and only if the following conditions are satisfied:

(1) *At least one of* M_{2i-1} *and* M_{2i} *belongs to* $\mathfrak{N}_l$ *for each* $1 \leqq i \leqq s$.
(2) M_k *belongs to* $\mathfrak{N}_l$ *for each* $2s+1 \leqq k \leqq t$.
(3) *For any* $Y \in \mathfrak{N}_l$ *there exists* $M_i (1 \leqq i \leqq t)$ *such that* $M_i \simeq Y$.
(4) *If* M_i *and* $M_j (1 \leqq i, j \leqq t, i \neq j)$ *are isomorphic to each other and belong to* $\mathfrak{N}_l$, *then* $\{i,j\} = \{2l-1, 2l\}$ *for some* $1 \leqq l \leqq s$.

Proof. This can be proved in almost the same way as the proof of [7, Lemma 1.3]. Assume that M is minimal faithful. Then, since each of indecomposable direct summands of M is either generated or cogenerated by $M_i (1 \leqq i \leqq t)$, we obtain (3) by Lemma 2.2. Suppose that neither M_{2i-1} nor M_{2i} belong to $\mathfrak{N}_l$ for some $1 \leqq i \leqq r$. Then, each of M_{2i-1} and M_{2i} is a proper subquotient of some A-module in $\mathfrak{N}_l$, and hence of some $M_k (1 \leqq k \leqq t,\ k \neq 2i-1,\ k \neq 2i)$. Since $X_i | K_i \oplus L_i$ is cogenerated by M_{2i-1} and M_{2i}, the deletion of $X_i | K_i \oplus L_i$ from M still leaves a faithful module. This is a contradiction. Since quite the same argument is possible for $r+1 \leqq i \leqq s$, (1) holds. Similarly we see that (2) holds. Applying Lemma 1.4, it is easy to show (4).

Next assume that all the conditions are satisfied. Then since every indecomposable left A-module is either generated or cogenerated by A-modules in $\mathfrak{S}_l$, we see that M is faithful by (3). The same argument as in the proof of [7, Lemma 1.3] leads to the minimality of M.

Let $M_1, \cdots, M_n$ be mutually nonisomorphic left A-modules in $\mathfrak{N}_l$ and satisfy that any module in $\mathfrak{N}_l$ is isomorphic to some $M_i (1 \leqq i \leqq n)$. Then the direct sum $M = M_1 \oplus \cdots \oplus M_n$ is minimal faithful by Proposition 2.3. We shall call this module M a normal minimal faithful left A-module. It should be noted that in case A is of left colocal type, a normal minimal faithful left A-module is just a minimal faithful projective left A-module (cf. [5]), and that normal minimal faithful left and right modules are P-dual modules of each other.

3. The Proof of Necessity

The purpose of this section is to prove that if A is QF-1, then the following hold:

(a) If a primitive left ideal Ae is serial and Ne is projective, then every composition factor of Ae is isomorphic to a minimal left ideal.

(b) If a primitive left ideal Ae is of type II, then every composition factor of Ae is isomorphic to a minimal left ideal.

(c) Let Ae be a primitive left ideal of type III and $Ne = K \nabla L$. If K is projective, then every composition factor of $K \nabla S^2(L)$ is isomorphic to a minimal left ideal.

(d) Let $Ae | K \oplus L$ be a primitive left ideal of type II. If $(Ae/K)^*$ is projective, then K^* is embedded into a primitive right ideal.

(e) Let $Ae | K \oplus L$ be a primitive left ideal of type II. Let f be a primitive idempotent such that $T(Af) \simeq T(K)$ and $\phi: eA \to fN$ the projective cover of fN. If $(Ae/K)^*$ is projective, $(Ae/L)^*$ is nonprojective and eA is of type II, then the kernel $\mathrm{Ker}(\phi)$ is nonserial.

Fuller [5] showed that a minimal faithful projective left A-module is uniquely determined up to isomorphism.

Lemma 3.1. *Assume that a primitive left ideal Ae does not appear as a direct summand of a minimal faithful projective left A-module M. If A is* QF-1, *then every composition factor of Ae is isomorphic to a minimal left ideal.*

Proof. It is clear that any factor module Ae/X is not generated by M. Therefore, by Morita's criterion [8] Ae/X is cogenerated by M. Thus any simple submodule of Ae/X is embedded into M.

For the balancedness of the injective hull $E({}_AA)$ (or maximal left quotient rings) there are many investigations. For example, it is well known that $E({}_AA)$ is balanced if and only if $E({}_AA)/{}_AA$ is cogenerated by $E({}_AA)$. So in our case it holds that $E({}_AA)$ is balanced if and only if every simple submodule of $E({}_AA)/{}_AA$ is isomorphic to a minimal left ideal.

Proposition 3.2. *If A is* QF-1, *then* (a) *is satisfied.*

Proof. Let $Af \simeq Ne$, where f is a primitive idempotent. Then, obviously Af is a primitive left ideal which does not appear as a direct summand of a minimal faithful projective left A-module. It follows directly from Lemma 3.1 that every composition factor of Ne appears in the left socle $S({}_AA)$. Moreover, as $E(Ne) \supset Ae \supset Ne$, $T(Ae)$ is embedded into $E(Af)/Af (\simeq E(Ne)/Ne)$, and hence into $S({}_AA)$ since $E({}_AA)$ is balanced. This shows the proposition.

Let $K \bigtriangledown L$ be a left A-module of type I. If K is embedded into a serial left A-module K', then by making the interlacing module we obtain a left A-module $X \bigtriangledown Y$ with $X \simeq K'$ and $Y \simeq L$. Hence if $K \bigtriangledown L$ is injective, then there does not exist a serial left A-module M into which K is properly embedded. Taking P-dual, for a primitive left ideal $Ae | K \oplus L$ of type II there does not exist a serial left A-module M such that Ae/K is a proper homomorphic image of M.

Lemma 3.3. *If a primitive left ideal Ae is of type II, then $S(E(Ae)/Ae) \simeq T(Ae)$.*

Proof. Let $Ae = Ae | K \oplus L$. Then $E(Ae) = E(K) \oplus E(L)$, and both $U = E(K)$ and $V = E(L)$ are serial by Theorem 1.1. Let U' and V' be submodules of U and V such that $U'/K = S(U/K)$ and $V'/L = S(V/L)$, respectively. Then since $Ae/K \oplus L$ is simple, we have $U' \oplus V' \supset Ae \supset K \oplus L$. It is impossible that Ae equals to either $U' \oplus L$ or $K \oplus V'$, since Ae is indecomposable. It follows $U'/K \simeq V'/L \simeq T(Ae)$ and $E(Ae)/Ae \simeq \mathrm{Int}_\theta(U/K, V/L)$ where θ: $U'/K \to V'/L$ is an isomorphism. Therefore it suffices to show that the socle of $E(Ae)/Ae$ is simple. If either U/K or V/L is simple, this is obvious. So we may assume $|U/K|$, $|V/L| \geqq 2$. If $K \neq L$, then we have $S(U/U') \neq S(V/V')$ by Lemma 1.3. Thus θ can not be extended to $S^2(U/K) \to S^2(V/L)$. From [10, Lemma 1.3] it follows that the socle of $\mathrm{Int}_\theta(U/K, V/L)$ is simple, as desired.

By the duality, the proof will be completed if we show the following: For an exact sequence

$$0 \to Z \to Af \oplus Af \to X \bigtriangledown Y \to 0,$$

$T(Z)$ is simple, where $X \bigtriangledown Y$ is an injective left A-module with $X \simeq Y$ and Af is a primitive left ideal. We may assume that X and Y are nonprojective, because this is obvious in not the case. Then the projective cover Ag of $S(X \bigtriangledown Y)$ is not serial by Lemma 1.5, and $S(Af)$ is embedded into a proper factor module of

Ng. Thus the injective hull of Af is serial by Theorem 1.1. Since $X \nabla Y$ is injective, there does not exist a serial left A-module into which X is properly embedded. Consequently we conclude that Af is injective. Now, suppose that $Z(\subset Af \oplus Af)$ is decomposable, and let $Z = Z' \oplus Z''$. Then we have $Af \oplus Af = E(Z') \oplus E(Z'')$ and $E(Z') \simeq E(Z'') \simeq Af$. It follows $X \nabla Y \simeq Af \oplus Af/Z \simeq E(Z')/Z' \oplus E(Z'')/Z''$. However, this is a contradiction. It is easily seen that the top of any indecomposable submodule of $Af \oplus Af$ is simple. So we conclude that $T(Z)$ is simple. The proof is completed.

Proposition 3.4. *If A is* QF-1, *then* (b) *is satisfied.*

Proof. From Lemma 3.3 we immediately see that $T(Ae)$ appears in $S({}_AA)$. Let f be a primitive idempotent such that $fNe \neq 0$. Then, taking P-dual, $(Ae)^*$ is an injective right A-module of type I and $((Ae)^*/S((Ae)^*))f \neq 0$. Thus there exists a submodule $K \nabla L$ of $(Ae)^*$ with $T(K) \simeq T(fA)$. Now according to Morita's criterion, $K \nabla L$ is either generated or cogenerated by a minimal faithful projective right A-module M. If $K \nabla L$ is cogenerated by M, then $K \nabla L$ is embedded into some indecomposable direct summand of M since the socle $S(K \nabla L)$ is simple. But this is impossible, because the injective hull of $K \nabla L$ is of type I. Hence $K \nabla L$ is generated by M. It follows that the simple module $T(fA)$ is isomorphic to a submodule of $T(M)$, and hence $T(Af)$ is embedded into $S({}_AA)$ (cf. [5]). This shows the proof.

Proposition 3.5. *If A is* QF-1, *then* (c) *is satisfied.*

Proof. Every composition factor of K is embedded into $S({}_AA)$ by Lemma 3.1. It is easily seen that $E(K)/K = Ae/K$ is a serial module with the radical isomorphic to $L/S(L)$ (and the top isomorphic to $T(Ae)$). Since K is projective, $S^2(L)/S(L)$ $(\simeq S(E(K)/K))$ is embedded into $S({}_AA)$.

The same proof as in [7, Proposition 2.1] is valid for showing (d). So in the next proposition we give the gist of proof.

Proposition 3.6. *If A is* QF-1, *then* (d) *is satisfied.*

Proof. It is obvious that $(Ae)^* = U \nabla V$, $(Ae/K)^* \simeq U$, $(Ae/L)^* \simeq V$ and $K^* \simeq V/S(V)$ for suitable submodules U and V of $(Ae)^*$. Since $U \nabla V$ is injective and U is projective, U belongs to $\mathfrak{N}_r$. Hence by Proposition 2.3 we can make a minimal faithful right A-module $M = (U \nabla V) \oplus M_1 \oplus \cdots \oplus M_n$, where each M_i is in $\mathfrak{N}_r$. Suppose that $V/S(V)$ is not embedded into any primitive right ideal. Then $V/S(V)$ is not embedded into any module in $\mathfrak{N}_r$, since $V/S(V)$ does not belong to $\mathfrak{N}_r$. By the same argument as in the proof of [7, Proposition 2.1] it follows $U \supsetneq \cap \mathrm{Ker}(\psi)$ and $U \not\subset \sum \mathrm{Im}(\psi)$. Here, $\cap \mathrm{Ker}(\psi)$ (resp. $\sum \mathrm{Im}(\psi)$) denotes the intersection (resp. sum) of kernels (resp. images) of all homomorphisms from $U \nabla V$ $\left(\text{resp. } \bigoplus_{i=1}^{n} M_i\right)$ to $\bigoplus_{i=1}^{n} M_i$ (resp. $U \nabla V$) and all endomorphisms of $U \nabla V$ which are not automorphisms. But this contradicts [7, Lemma 1.4]. So we have shown (d).

The following lemma due to Camillo is fundamental in proving (e).

Lemma 3.7 [Camillo, 1]. *If* $M=\bigoplus_{i=1}^{n} M_i$ *is a direct sum of left* A*-modules and* $a_i(1\leqq i\leqq n)$ *are elements in* A*, then the map defined by* $\sum_{i=1}^{n} m_i\mapsto \sum_{i=1}^{n} a_i m_i\in \bigoplus_{i=1}^{n} M_i$ *is an element of* $\mathrm{BiEnd}({}_AM)$ *if and only if* $(a_i-a_j)\mathrm{Tr}(M_i, M_j)=0$ *for all* $1\leqq i,j\leqq n$*. Here* $\mathrm{Tr}(M_i, M_j)$ *is the sum of images of all homomorphisms from* M_i *to* M_j*.*

Proposition 3.8. *If* A *is* QF-1*, then* (e) *is satisfied.*

Proof. Let U and V be the annihilators of K and L in $(Ae)^*$, respectively. Then we have $(Ae)^*=U\bigtriangledown V$, $(Ae/K)^*\simeq U$ and $(Ae/L)^*\simeq V$. Moreover, $S^2(V)$ is a homomorphic image of a serial primitive right ideal fA and $S(U\bigtriangledown V)\simeq T(eA)$. Suppose that $\mathrm{Ker}(\phi)$ is serial, and let $eA|X\oplus Y$. Then by Lemma 1.6 we may assume $fN\simeq eA/X$. Put $W=eA/X$ and $Z=eA/Y$.

Now as (b) holds, eA appears as a direct summand of a minimal faithful projective right A-module. Hence by Proposition 2.3 either W or Z belongs to $\mathfrak{N}_r$. But since W is properly embedded into fA, W does not belong to $\mathfrak{N}_r$. Thus we conclude that Z belongs to $\mathfrak{N}_r$, and at the same time that $W\not\simeq Z$. Again by (b), fA appears as a direct summand of a minimal faithful projective right A-module. As fA is serial, fA belongs to $\mathfrak{N}_r$. Further, U also belongs to $\mathfrak{N}_r$, since $U\bigtriangledown V$ is injective and U is projective by assumption. It is easily seen that Z, fA and U are mutually nonisomorphic. Therefore we obtain a normal minimal faithful right A-module $M=Z\oplus fA\oplus U\oplus M_1\oplus\cdots\oplus M_n$, where each M_i belongs to $\mathfrak{N}_r$. Now we can see

(i) $T(eA)$ does not appear in composition factors of any right A-module of type III;

(ii) For any serial right A-module H, $T(eA)$ appears at most once in composition factors of H;

(iii) In Z, fA, U and $M_1,\cdots,M_n$, only U has the socle isomorphic to $T(eA)$;

(iv) Z, fA and $M_1,\cdots,M_n$ are classified into three; modules having a subquotient isomorphic to Z/ZN^2, modules having a subquotient isomorphic to W/WN^2 and modules do not have a composition factor isomorphic to $T(eA)$.

(i) and (ii) follow from the assumption that both eA and Ae are of type II. To see (iii) suppose that there exists a serial module H in $\mathfrak{N}_r$ such that $H\not\simeq U$ and $S(H)\simeq T(eA)$. Then, by making the interlacing module we obtain a module $I\bigtriangledown J$ with $I\simeq U$ and $J\simeq H$. Since both V and H are not properly embedded into any serial module, it follows from Lemma 1.4 that $V\simeq H$. Thus V belongs to $\mathfrak{N}_r$, and hence V is projective. But this contradicts the assumption that $(eA/L)^*$ is nonprojective. This shows (iii). From Lemma 1.6 we see that any serial right A-module H with $|H|=2$ and $T(H)\simeq T(eA)$ is isomorphic to either W/WN^2 or Z/ZN^2. Considering this, (iv) follows at once from (ii) and (iii). Now we may assume

(i) $M_i(1\leqq i\leqq s)$ are serial modules having a subquotient isomorphic to Z/ZN^2;

(ii) $M_i(s+1\leqq i\leqq t)$ are serial modules having a subquotient isomorphic to W/WN^2;

(iii) $M_i(t+1\leqq i\leqq n)$ do not have a composition factor isomorphic to $T(eA)$.

We define a map d on the normal minimal faithful module $M = Z \oplus fA \oplus U \oplus M_1 \oplus \cdots \oplus M_n$ as follows:

$$(m)d = \begin{cases} 0 & \text{if } m \in Z^I \\ me & \text{if } m \in (fA)^I \\ 0 & \text{if } m \in U^I \\ 0 & \text{if } m \in (M_i)^I (1 \leqq i \leqq s) \\ me & \text{if } m \in (M_i)^I (s+1 \leqq i \leqq t) \\ 0 & \text{if } m \in (M_i)^I (t+1 \leqq i \leqq n), \end{cases}$$

where $m \in M$ and $(X)^I$ denotes the image of the inclusion $X \to M$. Then, applying Lemma 3.7 we can conclude $d \in \mathrm{BiEnd}(M_A)$. For example, we shall show $\mathrm{Tr}(M_k, U)e = 0$ for $s+1 \leqq k \leqq t$. Since W does not belong to $\mathfrak{N}_r$, the top $T(M_k)$ is not isomorphic to $S(U) \simeq T(eA)$. Thus if $l \geqq 1$ is an integer such that $S^{l+1}(M_k)/S^l(M_k) \simeq T(eA)$ (consequently $S^{l+1}(M_k)/S^{l-1}(M_k) \simeq W/WN^2$), then $l+2 \leqq |M_k|$. As $W/WN^2 \simeq fN/fN^3$, we have $S^{l+2}(M_k)/S^{l+1}(M_k) \simeq T(fA) \simeq S^2(V)/S(V)$. Since $U \not\simeq V$, we have $S^{l+2}(M_k)/S^{l+1}(M_k) \not\simeq S^2(U)/S(U)$. Therefore we get $\mathrm{Tr}(M_k, U) = 0$. For other cases we can easily see that the condition of Lemma 3.7 is satisfied. Since M is balanced by the assumption that A is QF-1, d is obtained by the right multiplication with an element of A. But d acts as non-zero on the subquotient fN/fN^2 of fA, and acts as zero on the socle $S(U)$ of U. As $fN/fN^2 \simeq S(U)$, this is a contradiction. We complete the proof.

4. The Proof of the Balancedness of a Normal Minimal Faithful Module

The purpose of this section is to show that a normal minimal faithful left A-module is balanced on the assumption of the conditions obtained in Sect. 3. Throughout this section we shall assume that A satisfies the conditions (a), (b), (c), (d), (e) in the preceding section and the conditions (a)*, (b)*, (c)*, (d)*, (e)*, where by $(\cdot)^*$ we mean the same condition as $(\cdot)$ for a primitive right ideal eA. At first we prove the balancedness of $E({}_AA)$.

Lemma 4.1. *$E({}_AA)$ is balanced.*

Proof. It is sufficient to show that every simple submodule of $E(Ae)/Ae$ is embedded into $S({}_AA)$ for any primitive left ideal Ae. This is trivial if Ae is of type III. In case Ae is of type II, this is immediate by the condition (b) and Lemma 3.3. So let Ae be serial.

Case (i) where $E(Ae)$ is serial. Let K be a submodule of $E(Ae)$ such that $E(Ae) \supset K \supset Ae$ and $K/Ae = S(E(Ae)/Ae)$. Let Af be the projective cover of K. Then there exists an exact sequence $Nf \to Ae \to 0$. Hence either $Nf \simeq Ae$ or Nf is decomposable (i.e. Af is of type II). If $Nf \simeq Ae$, then $S(E(Ae)/Ae)(\simeq T(Af))$ is embedded into $S({}_AA)$ by (a). If Af is of type II, then we have the same from (b).

Case (ii) where $E(Ae)$ is of type I. Let $E(Ae) = U \nabla V$. We may assume $U \supset Ae$ by Lemma 1.6. Let K be a submodule of U such that $U \supset K \supset Ae$ and $S(U/Ae) = K/Ae$. Then it is easy to see $S(E(Ae)/Ae) \simeq K/Ae \oplus S^2(V)/S(V)$. Notice that every submodule X of U with $X \supset Ae$ is projective. Since K is projective, K/Ae is embedded into $S({}_AA)$ by (a). The same holds for $S^2(V)/S(V)$ by (d)*.

Case (iii) where $E(Ae)$ is of type III. Let $NE(Ae)=U\nabla V$. We may assume $U\supset Ae$ by Lemma 1.6. If K is a submodule of U such that $U\supset K\supset Ae$ and $K/Ae=S(U/Ae)$, then we have $S(E(Ae)/Ae)\simeq K/Ae\oplus S^2(V)/S(V)$. Since U is projective, we conclude that both K/Ae and $S^2(V)/S(V)$ are embedded into $S({}_AA)$ by (c).

Let Π (resp. Π') denote the totality of simple left A-modules S satisfying the following (*) (resp. (**)):

(*) There exists an injective left A-module $K\nabla L$ such that K is projective, L is nonprojective and $S^2(L)/S(L)\simeq S$.

(**) There exists an injective left A-module $K\nabla L$ such that $K\simeq L$ and $S^2(L)/S(L)\simeq S$.

Put $\Pi\cup\Pi'=\Omega$. Then we have

Lemma 4.2. *Let a primitive left ideal Ae be the projective cover of a simple left A-module S belonging to Ω, and a primitive left ideal Af the projective cover of Ne. Then*

(1) *If Af is not of type III, then there exists a serial left A-module X such that Ne is a proper homomorphic image of X;*

(2) *If Af is of type III, then there exists a serial left A-module X such that N^2e is a proper homomorphic image of X.*

Proof. Since $T(Af)$ is isomorphic to the socle of an injective left A-module of type I, Af appears as a direct summand of a minimal faithful projective left A-module by (b)$^\#$.

(1) In case Af is serial, obviously $Af=X$ is a required module. Let $Af=Af|K\oplus L$ be of type II. At first assume $S\in\Pi$. By Lemma 1.6 we may assume $Ne\simeq Af/N^tK\oplus L$ for some integer t. According to (e)$^\#$ the kernel of the projective cover ϕ: $Af\to Ne$ is nonserial. Thus we have $N^tK\neq 0$. Hence $Af/L=X$ is a required module. Next assume $S\in\Pi'$. Let $U\nabla V$ be an injective left A-module with $U\simeq V$ and $S^2(V)/S(V)\simeq S$. If $S^2(V)$ is projective, then Ne is simple. Thus (1) is trivially satisfied. So let $S^2(V)$ be nonprojective. Then, since $S(Ae)$ is embedded into a proper factor module of Nf, we note that Ae is quasi-injective by Theorem 1.1. Now there exists an exact sequence

$$0\to Y\to Ae\oplus Ae\to S^2(U)\nabla S^2(V)\to 0.$$

Considering that $Ae\oplus Ae$ is quasi-injective, from the same argument as in the proof of Lemma 3.3 it follows that Y is a module of type II having the radical $NY=Z\oplus W$ with $Z\simeq W$ and $T(Y)\simeq T(Af)$. From this we have $K\simeq L$, equivalently $Af/K\simeq Af/L$. Since Af is a direct summand of a minimal faithful module, Af/K belongs to $\mathfrak{N}_l$ by Proposition 2.3. Therefore Ne is a proper homomorphic image of the serial module Af/K.

(2) Let $Nf=K\nabla L$. It is obvious that Ne is a homomorphic image of $M=Af/S(Af)$, and there exist submodules U and V of M such that $M=M|U\oplus V$, $U\simeq K/S(K)$ and $V\simeq L/S(L)$. By Lemma 1.6 we may assume that Ne is a homomorphic image of M/U. Then N^2e is a homomorphic image of $N(M/U)\simeq V$. Hence N^2e is a proper homomorphic image of L. We complete the proof.

Lemma 4.3. *Let M be a serial left A-module, and K a submodule of M. Assume that there does not exist a serial module X such that K is a proper homomorphic image of X. Then*

(1) *If the projective cover of K is not of type III, then no composition factor of M/K belongs to Ω;*

(2) *Assume that M is not properly embedded into any serial left A-module. If neither the injective hull nor the projective cover of K are of type III, then M belongs to $\mathfrak{N}_l$.*

Proof. (1) Suppose that a simple module $S=L/NL$ belongs to Ω, where $M\supset L\supset NL\supset K$. Let Ae be the projective cover of S, and Af the projective cover of Ne. If Af is not of type III, then according to Lemma 4.2 (1) there exists a serial left A-module X such that Ne is a proper homomorphic image of X. Since NL is a homomorphic image of Ne, NL is a proper homomorphic image of X. As a consequence we conclude that K is a proper homomorphic image of a suitable submodule of X. But this is a contradiction.

If Af is of type III, then we have $N^2L\supset K$ by the assumption on K. According to Lemma 4.2 (2), there exists a serial left A-module X such that N^2e is a proper homomorphic image of X. From these, similarly we have a contradiction.

(2) It is obvious that there does not exist a serial left A-module having a proper subquotient isomorphic to M. Suppose that M is isomorphic to a subquotient of a primitive left ideal Ae of type III with the radical $Ne=U\nabla V$. Then, by the assumption on K it follows that K is isomorphic to a subquotient of $Ne/S(Ae)\simeq U/S(U)\oplus V/S(V)$, and hence of either $U/S(U)$ or $V/S(V)$. Thus K is a proper homomorphic image of a suitable submodule of U or V. This is a contradiction.

A Serre class is a nonempty class $\mathscr{S}$ of left A-modules such that if

$$0\to X'\to X\to X''\to 0$$

is an exact sequence of left A-modules X, X' and X'', then $X\in\mathscr{S}$ if and only if X' and X'' are in $\mathscr{S}$. $\mathscr{S}$ is called to be strongly complete if $\mathscr{S}$ is closed under arbitrary infinite direct sums. Every strongly complete Serre class is obtained by a suitable injective left A-module W as follows: $\mathscr{S}=\mathscr{S}(W)=\{X|X$: left A-module, $\mathrm{Hom}_A(X,W)=0\}$. Following C.L. Walker and E.A. Walker [15] a left A-module V is said to be $\mathscr{S}$-injective if $\mathrm{Hom}_A(Y,V)\to\mathrm{Hom}_A(X,V)\to 0$ is exact, whenever $0\to X\to Y\to Y/X\to 0$ is exact and $Y/X\in\mathscr{S}$. The $\mathscr{S}$-injective hull $E_\varphi(M)$ of any left A-module M is obtained as the largest submodule $\hat{M}$ of the injective hull $E(M)$ of M such that $\hat{M}/M\in\mathscr{S}$.

Let $S_1,\cdots,S_n$ be mutually nonisomorphic simple left A-modules in Ω such that any simple left A-module in Ω is isomorphic to some $S_i(1\leqq i\leqq n)$. By ω we denote a strongly complete Serre class obtained by an injective left A-module $\bigoplus_{i=1}^{n}E(S_i)$, i.e. $\omega=\left\{X|X\text{: left } A\text{-module, } \mathrm{Hom}_A\left(X,\bigoplus_{i=1}^{n}E(S_i)\right)=0\right\}$. We note that $X\in\omega$ if and only if any simple subquotient of X does not belong to Ω.

Lemma 4.4. *For each primitive left ideal Ae of A, the ω-injective hull $E_\omega(Ae)$ of Ae is one of the following*

(1) *of type III;*

(2) *a serial left A-module in $\mathfrak{N}_l$;*

(3) *a direct sum of two serial left A-modules in $\mathfrak{N}_l$;*

(4) *isomorphic to a left A-module $K \nabla L$ such that K is in $\mathfrak{N}_l$, L is a submodule of a module in $\mathfrak{N}_l$ and $S^2(K) \neq S^2(L)$.*

Proof. Case (i) where Ae is of type III. Obvious.

Case (ii) where Ae is serial. At first, if $E(Ae)$ is serial, then applying Lemma 4.3 (1) we see that every composition factor of $E(Ae)/Ae$ does not belong to Ω. Thus we have $E(Ae)/Ae \in \omega$. This implies $E_\omega(Ae) = E(Ae)$. Moreover, $E(Ae)$ belongs to $\mathfrak{N}_l$ by Lemma 4.3 (2). Therefore in this case $E_\omega(Ae)$ is of the form (2). Next, let $E(Ae) = K \nabla L$. We may assume $K \supset Ae$ by Lemma 1.6. Then since K is not properly embedded into any serial module, K belongs to $\mathfrak{N}_l$. Further, similarly as above we have $K/Ae \in \omega$, and hence $K \subset E_\omega(Ae)$ $(\subset E(Ae))$. Note that $K \nabla S^i(L)$ $(i = 0, 1, 2, \cdots)$ are the totality of submodules of $K \nabla L$ containing K. Now if L belongs to $\mathfrak{N}_l$ and $S^2(K) \neq S^2(L)$, then obviously $E_\omega(Ae)$ is of the form (4). If not the case, by the definition of Ω we have $S^2(L)/S(L) \in \Omega$. From this, it follows $E_\omega(Ae) = K$. Thus $E_\omega(Ae)$ is of the form (2). Finally let $E(Ae)$ be of type III. Suppose that a simple left A-module $T(Af)$ in Ω is isomorphic to a composition factor of $E(Ae)/Ae$, where f is a primitive idempotent. Then clearly $T(Af)$ is embedded into a proper factor module of $NE(Ae)$. Thus $T(Nf)$ is isomorphic to a composition factor of $NE(Ae)$. Since the injective hull of any composition factor of $NE(Ae)$ is not of type I, this is a contradiction. As a consequence we conclude $E_\omega(Ae) = E(Ae)$.

Case (iii) where Ae is of type II. Let $Ae = Ae|K \oplus L$. Then we can easily see that every composition factor of $E(Ae)/Ae$ is isomorphic to either $T(Ae)$ or a composition factor of $(E(Ae/K)/Ae/K) \oplus (E(Ae/L)/Ae/L)$. Here we note that both $E(Ae/K)$ and $E(Ae/L)$ are serial. Now since Ae is of type II, $T(Ae)$ is not in Ω. Since there does not exist a serial module X such that Ae/K is a proper homomorphic image of X, every composition factor of $E(Ae/K)/Ae/K$ is not in Ω by Lemma 4.3 (1). Of course the same holds for $E(Ae/L)/Ae/L$. From these it follows $E_\omega(Ae) = E(Ae)$ $(\simeq E(Ae/K) \oplus E(Ae/L))$. Since both $E(Ae/K)$ and $E(Ae/L)$ belong to $\mathfrak{N}_l$ by Lemma 4.3 (2), we conclude that $E_\omega(Ae)$ is of the form (3). We complete the proof.

Lemma 4.5. *Let $M_i (1 \leqq i \leqq n)$ be left A-modules and $M_i = K_{2i-1} + K_{2i}$ $(1 \leqq i \leqq n)$, where K_{2i-1} and K_{2i} are submodules of M_i. Assume that $K_{2i-1} \cap K_{2i}$ is a homomorphic image of some K_l for each $1 \leqq i \leqq n$. Further assume that for any homomorphism $\phi: M_p \to M_q$, the image $(K_{2p-1})\phi$ of K_{2p-1} is contained in either K_{2q-1} or K_{2q} and also the image $(K_{2p})\phi$ of K_{2p} is contained either K_{2q-1} or K_{2q}. If $M = \bigoplus_{i=1}^{n} M_i$ is balanced, then so is $K = \bigoplus_{j=1}^{2n} K_j$.*

Proof. By $(m)_i$ (resp. $[k]_j$) we denote the image of an element $m \in M_i$ (resp. $k \in K_j$) by the inclusion $M_i \to M = \bigoplus_{i=1}^{n} M_i$ $\left(\text{resp. } K_j \to K = \bigoplus_{j=1}^{2n} K_j\right)$. Let d be any element of $\mathrm{BiEnd}({}_AK)$, and for each $1 \leqq j \leqq 2n$ let d_j be the map on K_j induced by d, that is

$$d\left(\sum_{j=1}^{2n} [k_j]_j\right) = \sum_{j=1}^{2n} [d_j(k_j)]_j.$$

We can show $d_{2i-1}(k)=d_{2i}(k)$ for all $k\in K_{2i-1}\cap K_{2i}$. To see this, let $\alpha\colon K_l\to K_{2i-1}\cap K_{2i}$ be an epimorphism, and let $\alpha'\colon K\to K$ be the homomorphism defined by $\left(\sum_{j=1}^{2n}[k_j]_j\right)\alpha'=[(k_l)\alpha]_{2i-1}$. Take an element $x\in K_l$ such that $(x)\alpha=k$. Then we have

$$\begin{aligned}[d_{2i-1}(k)]_{2i-1}&=[d_{2i-1}((x)\alpha)]_{2i-1}=d([(x)\alpha]_{2i-1})\\&=d(([x]_l)\alpha')=(d([x]_l))\alpha'=([d_l(x)]_l)\alpha'=[(d_l(x))\alpha]_{2i-1}.\end{aligned}$$

Thus we have $d_{2i-1}(k)=(d_l(x))\alpha$. Similarly we get $d_{2i}(k)=(d_l(x))\alpha$. Therefore $d_{2i-1}(k)=d_{2i}(k)$, as required. Now we define a map $\bar{d}_i$ on M_i by

$$\bar{d}_i(k_1+k_2)=d_{2i-1}(k_1)+d_{2i}(k_2),$$

where $k_1\in K_{2i-1}$ and $k_2\in K_{2i}$. By what we have shown above, we can easily see that $\bar{d}_i$ is well-defined. Further we define a map $\bar{d}$ on M by

$$\bar{d}\left(\sum_{i=1}^{n}(m_i)_i\right)=\sum_{i=1}^{n}(\bar{d}_i(m_i))_i.$$

We want to show $\bar{d}\in\mathrm{BiEnd}({}_AM)$. Let $e_i\in\mathrm{End}({}_AM)$ be the projection on M_i, and let Φ be any element of $\mathrm{End}({}_AM)$. Then since $\Phi=\sum_{i,j}e_i\Phi e_j$, it is sufficient to show $\bar{d}\left(\left(\sum_{i=1}^{n}(m_i)_i\right)e_p\Phi e_q\right)=\left(\bar{d}\left(\sum_{i=1}^{n}(m_i)_i\right)\right)e_p\Phi e_q$ for all $1\leqq p,q\leqq n$. Let $\phi\colon M_p\to M_q$ be the homomorphism induced by $e_p\Phi e_q$. Then we have

$$\begin{aligned}x&=\left(\bar{d}\left(\sum_{i=1}^{n}(m_i)_i\right)\right)e_p\Phi e_q=\left(\sum_{i=1}^{n}(\bar{d}_i(m_i))_i\right)e_p\Phi e_q\\&=((\bar{d}_p(m_p))\phi)_q=((\bar{d}_p(k_1+k_2))\phi)_q=((d_{2p-1}(k_1))\phi)_q+((d_{2p}(k_2))\phi)_q,\end{aligned}$$

and

$$\begin{aligned}y&=\bar{d}\left(\left(\sum_{i=1}^{n}(m_i)_i\right)e_p\Phi e_q\right)=\bar{d}(((m_p)\phi)_q)=(\bar{d}_q((m_p)\phi))_q\\&=(\bar{d}_q((k_1+k_2)\phi))_q=(\bar{d}_q((k_1)\phi))_q+(\bar{d}_q((k_2)\phi))_q,\end{aligned}$$

where $m_p=k_1+k_2(k_1\in K_{2p-1}, k_2\in K_{2p})$. We can see $(d_{2p-1}(k_1))\phi=\bar{d}_q((k_1)\phi)$ and $(d_{2p}(k_2))\phi=d_q((k_2)\phi)$. For example, consider the case where $(K_{2p-1})\phi\subset K_{2q-1}$. Let $\psi\colon K_{2p-1}\to K_{2q-1}$ be the restriction of ϕ, and $\Psi\colon K\to K$ the homomorphism induced by ψ. Then we have

$$\begin{aligned}[(d_{2p-1}(k_1))\phi]_{2q-1}&=[(d_{2p-1}(k_1))\psi]_{2q-1}\\&=([d_{2p-1}(k_1)]_{2p-1})\,\Psi=(d([k_1]_{2p-1}))\,\Psi=d(([k_1]_{2p-1})\,\Psi)\\&=d([(k_1)\psi]_{2q-1})=[d_{2q-1}((k_1)\psi)]_{2q-1}=[\bar{d}_q((k_1)\phi)]_{2q-1}.\end{aligned}$$

Thus we obtain $(d_{2p-1}(k_1))\phi=\bar{d}_q((k_1)\phi)$. For the case where $(K_{2p-1})\phi\subset K_{2q}$, the same is possible. Similarly we obtain $(d_{2p}(k_2))\phi=\bar{d}_q((k_2)\phi)$. This shows $x=y$. Hence we get $\bar{d}\in\mathrm{BiEnd}({}_AM)$, as desired. Since M is balanced, $\bar{d}$ is obtained

by the left multiplication with an element of A. And so is d by the definition of $\bar{d}$. This means that K is balanced.

Lemma 4.6. *If the injective hull of a left A-module $K \nabla L$ is of type I and $S^2(K) \neq S^2(L)$, then there exists no simple A-module which appears in composition factors of both $K/S(K)$ and $L/S(L)$.*

Proof. Suppose $S^{p+1}(K)/S^p(K) \simeq S^{q+1}(L)/S^q(L)$ for some $p, q \geqq 1$. Since $S^2(K) \neq S^2(L)$, we have $p \neq q$. So we may assume $p > q$. Then we have $S^{p-q+1}(K)/S^{p-q}(K) \simeq S(K \nabla L)$. Thus $S(K \nabla L)$ is a homomorphic image of a serial primitive left ideal. Since $S^{q+1}(L)$ is nonprojective, applying Lemma 1.5 to $S^2(K) \nabla S^{q+1}(L)$ we see that $S^2(K)$ is projective. From (b) it follows that $S^2(K)$ appears as a direct summand of a minimal faithful projective left A-module. Hence we conclude $K = S^2(K)$. But this is a contradiction.

The following Lemma 4.7 due to Tachikawa and Lemma 4.8 due to Fuller are applied to the proof of the balancedness of a normal minimal faithful module.

Lemma 4.7 [Tachikawa, 12]. *Let $\mathscr{S}$ and $\mathscr{S}_0$ be strongly complete Serre classes. If $\mathscr{S} \subset \mathscr{S}_0$ and $E_{\mathscr{S}_0}(A)$ is balanced, then so is $E_{\mathscr{S}}(A)$.*

Lemma 4.8 [Fuller, 5]. *Let M and K be left A-modules. If M both generates and cogenerates K, then the restriction* Res: $\mathrm{BiEnd}({}_AM \oplus K) \to \mathrm{BiEnd}({}_AM)$ *is an isomorphism.*

Proposition 4.9. *A normal minimal faithful left A-module is balanced.*

Proof. Since $E({}_AA) = E_{\mathfrak{M}}({}_AA)$ where $\mathfrak{M}$ is the class of all left A-modules, $E_\omega({}_AA)$ is balanced by Lemmas 4.1 and 4.7. If $1 = \sum_{i=1}^{n} e_i$ is a sum of mutually orthogonal primitive idempotents, then $E_\omega({}_AA) = \bigoplus_{i=1}^{n} E_\omega(Ae_i)$. Hence we know from Lemma 4.4 that $E_\omega({}_AA)$ is a direct sum $E_\omega({}_AA) = \left(\bigoplus_{i=1}^{r} M_i \nabla K_i\right) \oplus \left(\bigoplus_{j=r+1}^{s} M_j\right)$, where $M_i (1 \leqq i \leqq s)$ are in $\mathfrak{N}_l$, K_i are submodules of modules in $\mathfrak{N}_l$ and $S^2(M_i) \neq S^2(K_i)$ $(1 \leqq i \leqq r)$. It is easily seen that $\bigoplus_{i=1}^{s} M_i$ is equivalent to a normal minimal faithful left A-module, i.e. for any $X \in \mathfrak{N}_l$, X is isomorphic to some M_i.

By the condition (b)$^\#$ the projective cover of $M_i \cap K_i = S(M_i \nabla K_i)$ appears as a direct summand of a minimal faithful projective left A-module for each $1 \leqq i \leqq r$. Thus $M_i \cap K_i$ is a homomorphic image of some $M_l (1 \leqq l \leqq s)$ for each $1 \leqq i \leqq r$. Moreover, every serial submodule of $M_i \nabla K_i$ is contained in either M_i or K_i by Lemma 4.6. Therefore, applying Lemma 4.5 to modules

$$\left(\bigoplus_{i=1}^{r} M_i \nabla K_i\right) \oplus \left(\bigoplus_{j=r+1}^{s} M_j\right) \quad \text{and} \quad \left(\bigoplus_{i=1}^{s} M_i\right) \oplus \left(\bigoplus_{i=1}^{r} K_i\right)$$

we see that $\left(\bigoplus_{i=1}^{s} M_i\right) \oplus \left(\bigoplus_{i=1}^{r} K_i\right)$ is balanced. Since each K_i is embedded into some module in $\mathfrak{N}_l$, obviously $\bigoplus_{i=1}^{r} K_i$ is cogenerated by $\bigoplus_{i=1}^{s} M_i$. Moreover, $\bigoplus_{i=1}^{r} K_i$

is generated by $\overset{s}{\underset{i=1}{\oplus}} M_i$, since the projective cover of each K_i belongs to $\mathfrak{N}_l$ by (b)*. Thus by Lemma 4.8 we conclude that $\overset{s}{\underset{i=1}{\oplus}} M_i$ is balanced, and hence so is a normal minimal faithful left A-module. We complete the proof.

5. The Proof of Sufficiency

In this section we shall prove that the conditions assumed in the preceding section are sufficient for A to be QF-1. Our proof can be regarded as a modification of that of [7, Theorem 2.5].

Let $M_i (1 \leqq i \leqq n)$ be left A-modules, and $K_j (1 \leqq j \leqq s)$ submodules of $M_{\alpha_j} (1 \leqq j \leqq s)$ respectively. Put $X_j = M_{\alpha_j}/K_j$ if $1 \leqq j \leqq r$ and $X_j = K_j$ if $r+1 \leqq j \leqq s$, where $i \leqq r \leqq s$, and put $M = \overset{n}{\underset{i=1}{\oplus}} M_i$ and $X = \overset{s}{\underset{j=1}{\oplus}} X_j$. By $(m)_i$ (resp. $[x]_j$) we denote the image of $m \in M_i$ (resp. $x \in X_j$) by the inclusion $M_i \to M$ (resp. $X_j \to X$). Let d be an element of $\mathrm{BiEnd}({}_AM)$. By d_i we denote the map on M_i induced by d, that is

$$d\left(\sum_{i=1}^{n} (m_i)_i\right) = \sum_{i=1}^{n} (d_i(m_i))_i.$$

Then $D_i = \{d_i \mid d \in \mathrm{BiEnd}({}_AM)\}$ is a subring of $\mathrm{BiEnd}({}_AM_i)$. Assume that K_j is a D_{α_j}-submodule of M_{α_j} for each $1 \leqq j \leqq s$. Then we can define a map $d' : X \to X$ by

$$d'\left(\sum_{j=1}^{r} [\bar{m}_j]_j + \sum_{j=r+1}^{s} [k_j]_j\right) = \sum_{j=1}^{r} [\overline{d_{\alpha_j}(m_j)}]_j + \sum_{j=r+1}^{s} [d_{\alpha_j}(k_j)]_j,$$

where $\bar{m}_j = m_j + K_j \in M_{\alpha_j}/K_j = X_j$ $(1 \leqq j \leqq r)$ and $k_j \in K_j = X_j$ $(r+1 \leqq j \leqq s)$. With these setting we have the following lemma.

Lemma 5.1. *Assume that for any homomorphism $\phi : X_p \to X_q$ there exists a homomorphism $\Phi : M_{\alpha_p} \to M_{\alpha_q}$ such that the diagram*

$$\begin{array}{ccc} M_{\alpha_p} & \xrightarrow{\exists_\Phi} & M_{\alpha_q} \\ {\scriptstyle \sigma_p}\uparrow\downarrow & & \uparrow\downarrow{\scriptstyle \sigma_q} \\ X_p & \xrightarrow{\phi} & X_q \end{array}$$

commute, where $\sigma_p : M_{\alpha_p} \to M_{\alpha_p}/K_p = X_p$ is the canonical epimorphism if $1 \leqq p \leqq r$ and $\sigma_p : X_p = K_p \to M_{\alpha_p}$ is the inclusion if $r+1 \leqq p \leqq s$. Then d' belongs to $\mathrm{BiEnd}({}_AX)$ *for any element d of* $\mathrm{BiEnd}({}_AM)$.

Proof. Let $\Phi : M_{\alpha_p} \to M_{\alpha_q}$ be a homomorphism, and $\Psi : M \to M$ the homomorphism induced by Φ. If $m \in M_{\alpha_p}$, then as an element of M we have

$$\begin{aligned} ((d_{\alpha_p}(m))\Phi)_{\alpha_q} &= ((d_{\alpha_p}(m))_{\alpha_p})\ \Psi = (d((m)_{\alpha_p}))\ \Psi \\ &= d(((m)_{\alpha_p})\Psi) = d(((m)\Phi)_{\alpha_q}) = (d_{\alpha_q}((m)\Phi))_{\alpha_q}. \end{aligned}$$

Hence we get

$$(d_{\alpha_p}(m))\Phi = d_{\alpha_q}((m)\Phi). \tag{*}$$

Now let ψ be any element of $\mathrm{End}({}_AX)$, and e_j the projection on X_j. It is sufficient to show that

$$x = \left(d'\left(\sum_{j=1}^{r}[\bar{m}_j]_j + \sum_{j=r+1}^{s}[k_j]_j\right)\right)e_p\psi e_q$$

is equal to

$$y = d'\left(\left(\sum_{j=1}^{r}[\bar{m}_j]_j + \sum_{j=r+1}^{s}[k_j]_j\right)e_p\psi e_q\right)$$

for all $1 \leqq p, q \leqq s$. Let $\phi: X_p \to X_q$ be the homomorphism induced by $e_p\psi e_q$, and $\Phi: M_{\alpha_p} \to M_{\alpha_q}$ a homomorphism such that the above diagram commute for ϕ. Now, consider the case where $r+1 \leqq p \leqq s$ and $1 \leqq q \leqq r$. In this case we have

$$\begin{aligned} x &= \left(\sum_{j=1}^{r}[\overline{d_{\alpha_j}(m_j)}]_j + \sum_{j=r+1}^{s}[d_{\alpha_j}(k_j)]_j\right)e_p\psi e_q \\ &= [(d_{\alpha_p}(k_p))\phi]_q = [\overline{(d_{\alpha_p}(k_p))\Phi}]_q \end{aligned}$$

and

$$y = d'([(k_p)\phi]_q) = d'([\overline{(k_p)\Phi}]_q) = [\overline{d_{\alpha_q}((k_p)\Phi)}]_q.$$

Considering similarly for the other cases, we have

$$x = \begin{cases} [\overline{(d_{\alpha_p}(m_p))\Phi}]_q \\ [(d_{\alpha_p}(m_p))\Phi]_q \\ [\overline{(d_{\alpha_p}(k_p))\Phi}]_q \\ [(d_{\alpha_p}(k_p))\Phi]_q \end{cases} \quad \text{and} \quad y = \begin{cases} [\overline{d_{\alpha_q}((m_p)\Phi)}]_q & \text{if } 1 \leqq p \leqq r \text{ and } 1 \leqq q \leqq r \\ [d_{\alpha_q}((m_p)\Phi)]_q & \text{if } 1 \leqq p \leqq r \text{ and } r+1 \leqq q \leqq s \\ [\overline{d_{\alpha_q}((k_p)\Phi)}]_q & \text{if } r+1 \leqq p \leqq s \text{ and } 1 \leqq q \leqq r \\ [d_{\alpha_q}((k_p)\Phi)]_q & \text{if } r+1 \leqq p \leqq s \text{ and } r+1 \leqq q \leqq s. \end{cases}$$

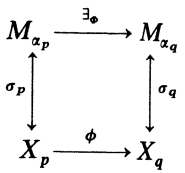
Thus, from (*) we have always $x = y$.

Let

$$M = \left(\bigoplus_{i=1}^{r} M_i | K_{2i-1} \oplus K_{2i}\right) \oplus \left(\bigoplus_{j=r+1}^{s} K_{2j-1} \nabla K_{2j}\right) \oplus \left(\bigoplus_{k=s+1}^{t} M_k\right)$$

be a given minimal faithful left A-module, where $M_k (s+1 \leqq k \leqq t)$ are either serial or of type III. Put $X_{2i-1} = M_i/K_{2i-1} (1 \leqq i \leqq r)$, $X_{2i} = M_i/K_{2i} (1 \leqq i \leqq r)$ and $M_j = K_{2j-1} \nabla K_{2j} (r+1 \leqq j \leqq s)$, $X_{2j-1} = K_{2j-1} (r+1 \leqq j \leqq s)$, $X_{2j} = K_{2j} (r+1 \leqq j \leqq s)$ and $K_{k+s} = X_{k+s} = M_k (s+1 \leqq k \leqq t)$. Let $\alpha_j = [j+1/2]$ if $1 \leqq j \leqq 2s$ and $\alpha_j = j-s$ if $2s+1 \leqq j \leqq s+t$. Then

Lemma 5.2. *If both X_p and $X_q (1 \leqq p, q \leqq s+t)$ are in $\mathfrak{N}_1$, then for any homomorphism $\phi: X_p \to X_q$ there exists a homomorphism $\Phi: M_{\alpha_p} \to M_{\alpha_q}$ such that the diagram*



commute, where $\sigma_p: M_{\alpha_p} \to X_p$ *is the canonical epimorphism if* $1 \leqq p \leqq 2r$ *and* $\sigma_p: X_p \to M_{\alpha_p}$ *is the inclusion if* $2r+1 \leqq p \leqq s+t$.

Proof. We can see that if $\psi: K \to Z$ is a nonmonomorphism, then ψ can be extended to $\Psi: K \nabla L \to Z$. In fact, the homomorphism $\Psi: K \nabla L \to Z$ defined by $(k+l)\Psi = (k)\psi$ is an extension of ψ. By taking P-dual, we know that if $\psi: Z \to U/V$ is a nonepimorphism, then ψ can be lifted to $\Psi: Z \to U | V \oplus W$. In case $\alpha_p \neq \alpha_q$, $\phi: X_p \to X_q$ is neither a monomorphism nor an epimorphism by Proposition 2.3. From these the lemma can be easily verified except for the case where $1 \leqq \alpha_p = \alpha_q \leqq s$. According to [11, Proposition 4.1] serial modules over an algebra of local-colocal type have the same properties as those over an algebra of left colocal type. Hence for the case where $r+1 \leqq \alpha_p = \alpha_q \leqq s$, the proof in [7, Theorem 2.5] is valid. For the case where $1 \leqq \alpha_p = \alpha_q \leqq r$, we can also obtain the conclusion by the duality.

Lemma 5.3. *If A satisfies the conditions* (a), (b), (c), (d), (e) *and* (a)*, (b)*, (c)*, (d)*, (e)* *in the previous sections, then any minimal faithful left A-module*

$$M = \left(\bigoplus_{i=1}^{r} M_i | K_{2i-1} \oplus K_{2i}\right) \oplus \left(\bigoplus_{j=r+1}^{s} K_{2j-1} \nabla K_{2j}\right) \oplus \left(\bigoplus_{k=s+1}^{t} M_k\right)$$

is balanced, where $M_k (s+1 \leqq k \leqq t)$ *are serial or of type III.*

Proof. By $M_i (1 \leqq i \leqq t)$, $K_j (1 \leqq j \leqq s+t)$ and $X_j (1 \leqq j \leqq s+t)$ we denote the same modules as before. Let $D_i = \{d_i | d \in \mathrm{BiEnd}({}_AM)\}$, where d_i is the map induced by d, and let $\Gamma = \{\text{integer } j | 1 \leqq j \leqq s+t, X_j \text{ belongs to } \mathfrak{N}_l\}$. We can show that if $j \in \Gamma$, then K_j is a D_{α_j}-submodule of M_{α_j}. This is trivial in case $2s+1 \leqq j \leqq s+t$. To see this for the case $2r+1 \leqq j \leqq 2s$, for example, assume that $K_{2p-1} = X_{2p-1} (r+1 \leqq p \leqq s)$ belongs to $\mathfrak{N}_l$. Then $K_{2p}/S(K_{2p})$ is embedded into a primitive left ideal by (d)*, and hence into some $M_i (1 \leqq i \leqq t)$. It follows at once that K_{2p-1} is a D_p-submodule of M_p (cf. the proof of [7, Theorem 2.5]). Considering dually, we get the desired conclusion for the case $1 \leqq j \leqq 2r$.

Now let d be any element of $\mathrm{BiEnd}({}_AM)$. By what we have shown above, we can define a map d' on $X = \bigoplus_{j \in \Gamma} X_j$ as follows:

$$d'\Big(\sum_{\substack{j \in \Gamma \\ 1 \leqq j \leqq 2r}} [\bar{m}_j]_j + \sum_{\substack{j \in \Gamma \\ 2r+1 \leqq j \leqq s+t}} [k_j]_j\Big) = \sum_{\substack{j \in \Gamma \\ 1 \leqq j \leqq 2r}} [\overline{d_{\alpha_j}(m_j)}]_j + \sum_{\substack{j \in \Gamma \\ 2r+1 \leqq j \leqq s+t}} [d_{\alpha_j}(k_j)]_j,$$

where $[x]_j$ denote the image of $x \in X_j$ by the inclusion $X_j \to X$. Then, applying Lemmas 5.1 and 5.2 we get $d' \in \mathrm{BiEnd}({}_AX)$. Since .by Proposition 2.3 X is equivalent to a normal minimal faithful left A-module, X is balanced by Proposition 4.9. Hence d' is obtained by the left multiplication with an element a of A. From the definition of d' it follows that if $j \in \Gamma$ and $1 \leqq j \leqq 2r$, then $d_{\alpha_j}(m) - am \in K_j (\forall m \in M_{\alpha_j})$, and that if $j \in \Gamma$ and $2r+1 \leqq j \leqq s+t$, then $d_{\alpha_j}(k) = ak (\forall k \in K_j)$.

Let us show that if $M_p/K_{2p-1} (1 \leqq p \leqq r)$ is in $\mathfrak{N}_l$, then $d_p(k) = ak (\forall k \in K_{2p})$. Since $d_p(k) - ak \in K_{2p-1}$ and $K_{2p-1} \cap K_{2p} = 0$, it is sufficient to show that K_{2p} is a D_p-submodule of M_p. This has been shown, in case M_p/K_{2p} is in $\mathfrak{N}_l$. So assume that M_p/K_{2p} is not in $\mathfrak{N}_l$. Since $S(M_p/K_{2p})$ is embedded into $S({}_AA)$ by

(b), the injective hull $E(M_p/K_{2p})$ is a serial module in $\mathfrak{N}_l$. Thus M_p/K_{2p} is properly embedded into a serial module in $\mathfrak{N}_l$, and hence into some $M_q(1 \leqq q \leqq t)$. Using a homomorphism $\phi: M_p \to M_q$ with $\mathrm{Ker}(\phi) = K_{2p}$, it is easy to see that K_{2p} is a D_p-submodule of M_p.

Now, the proof will be completed if we show $d_i(m) = am (\forall m \in M_i)$ for all $1 \leqq i \leqq t$. This is obvious for $s \leqq i \leqq t$. Let us show it for $1 \leqq i \leqq r$. By Proposition 2.3 we may assume that M_i/K_{2i-1} is in $\mathfrak{N}_l$. If also M_i/K_{2i} is in $\mathfrak{N}_l$, then from $d_i(m) - am \in K_{2i-1}$, $d_i(m) - am \in K_{2i}$ and $K_{2i-1} \cap K_{2i} = 0$ we have $d_i(m) = am$. So let M_i/K_{2i} be not in $\mathfrak{N}_l$. Then, as above, M_i/K_{2i} is properly embedded into a module Y in $\mathfrak{N}_l$. From this we can obtain a monomorphism $\psi: M_i/K_{2i} \to M_j (1 \leqq j \leqq t)$ such that d_j acts as a-multiplication on the image $(M_i/K_{2i})\psi$. For, if $Y \simeq M_p/K_{2p-1} (1 \leqq p \leqq r)$, then M_i/K_{2i} is embedded into K_{2p}, and $d_p(k) = ak (\forall k \in K_{2p})$ as shown above. By using ψ it is easily seen that $d_i(m) - am \in K_{2i} (\forall m \in M_i)$, and it follows $d_i(m) = am (\forall m \in M_i)$. The dual argument is possible for $r \leqq i \leqq s$ (cf. the proof of [7, Theorem 2.5]). We complete the proof.

By Morita's criterion the proof of sufficiency will be completed if we show the next lemma.

Lemma 5.4. *If A satisfies the conditions* (a), (b), (c), (d), (e) *and* (a)*, (b)*, (c)*, (d)*, (e)* *in the previous sections, then M either generates or cogenerates U for any minimal faithful left A-module M and any indecomposable left A-module U.*

Proof. Case (i) where U is of type I. Let $U = K \triangledown L$. If the injective hull $E(U)$ of U is of type III, obviously U is cogenerated by M. If not the case, the projective cover K' of K is a serial module in $\mathfrak{N}_l$ by (b)*. Hence M has an indecomposable direct summand isomorphic to either K' or $X \triangledown Y$ with $X \simeq K'$. Of course, the same holds for L. Taking Lemmas 1.2 and 1.4 into consideration it follows that M either generates or cogenerates U (cf. the proof of [7, Theorem 3.1]).

Case (ii) where U is of type II. By duality we get the proof from (i).

Case (iii) where U is of type III. Trivial.

Case (iv) where U is serial. Note that if U is either embedded into a module in $\mathfrak{N}_l$ or a homomorphic image of a module in $\mathfrak{N}_l$, then U is either generated or cogenerated by M. If $E(U)$ is of type I, then U is either generated or cogenerated by M. For, in this case U is a homomorphic image of a module in $\mathfrak{N}_l$ by (b)*. Further we can show that if $S(U)$ is embedded into $S({}_AA)$, then U is either generated or cogenerated by M. This is obvious if $S(U)$ is embedded into a primitive left ideal of type III. If $S(U)$ is embedded into a primitive left ideal of type II, or if $S(U)$ is embedded into a serial primitive ideal having the serial injective hull, then U is embedded into $E(S(U))$ and $E(S(U))$ is in $\mathfrak{N}_l$.

Now, if U is isomorphic to a subquotient of a primitive left ideal of type II, then $S(U)$ is embedded into $S({}_AA)$ by (b). So we may assume that U is non-isomorphic to any subquotient of a primitive left ideal of type II.

Let Ae be the projective cover of U. Then Ae is not of type II. In case Ae is of type III, obviously M generates U. So we may assume that Ae is serial, and moreover that Ae is not in $\mathfrak{N}_l$. At first consider the case where $E(Ae)$ is serial. Since $E(Ae)$ is in $\mathfrak{N}_l$ by Lemma 4.3, there exists a submodule X of $E(Ae)$ such that $E(Ae) \supset X \supset Ae$ and $NX = Ae$. It is easily seen that the projective

cover of X is neither of type II nor of type III. Therefore X is projective. From (a) it follows that $S(U)$ is embedded into $S({}_AA)$. Next assume that $E(Ae)=K\nabla L$ is of type I, and $K\supset Ae$. Since K is in $\mathfrak{N}_l$ by Lemma 4.3, there exists a submodule X of K such that $K\supset X\supset Ae$ and $NX=Ae$. Since X is projective, again by (a) $S(U)$ is embedded into $S({}_AA)$. Finally assume that $E(Ae)$ is of type III. Let $NE(Ae)=K\nabla L$ and $K\supset Ae$. Then K is projective. Thus, from (c) we conclude that $S(U)$ is embedded into $S({}_AA)$. This proves the lemma.

Thus we obtain

Theorem 5.5. *Let A be an algebra of local-colocal type. Then in order that A may be* QF-1 *it is necessary and sufficient that the following conditions are satisfied.*

(1) *If a primitive left ideal Ae is serial and Ne is projective, then every composition factor of Ae is isomorphic to a minimal left ideal. And the same holds for a primitive right ideal eA that is serial.*

(2) *If a primitive left ideal Ae is of type II, then every composition factor of Ae is isomorphic to a minimal left ideal. And the same holds for a primitive right ideal eA that is of type II.*

(3) *Let Ae be a primitive left ideal of type III and $Ne=K\nabla L$. If K is projective, then every composition factor of $K\nabla S^2(L)$ is isomorphic to a minimal left ideal. And the same holds for a primitive right ideal eA that is of type III.*

(4) *Let $Ae|K\oplus L$ be a primitive left ideal of type II. If $(Ae/K)^*$ is projective, then K^* is embedded into a primitive right ideal. And the same holds for a primitive right ideal eA that is of type II.*

(5) *Let $Ae|K\oplus L$ be a primitive left ideal of type II. Let f be a primitive idempotent such that $T(Af)\simeq T(K)$, and $\phi\colon eA\to fN$ the projective cover of fN. If $(Ae/K)^*$ is projective, $(Ae/L)^*$ is nonprojective and eA is of type II, then the kernel* $\mathrm{Ker}(\phi)$ *is nonserial. And the same holds for a primitive right ideal eA that is of type II.*

Remark. In fact, the conditions (a)$^\#$, (c)$^\#$ and (e) are not used in the proofs. So, in Theorem 5.5 we may replace the conditions with (a), (b), (b)$^\#$, (c), (d), (d)$^\#$ and (e)$^\#$ (it is known by examples that each of these conditions are independent of all other conditions (see Example 6.2)).

6. Examples

In the following c_{ij} mean matrix units in the full matrix ring $M_n(P)$ over a field P.

Example 6.1. Let A be the subalgebra of full matrix ring $M_{22}(P)$ such that the elements

$$e_1=c_{11}+c_{18,18},\quad e_2=c_{22}+c_{66}+c_{19,19},\quad e_3=c_{33}+c_{21,21},$$
$$e_4=c_{44}+c_{99}+c_{22,22},\quad e_5=c_{55}+c_{77}+c_{10,10}+c_{12,12},\quad e_6=c_{88}+c_{15,15},$$
$$e_7=c_{11,11}+c_{13,13},\quad e_8=c_{14,14},\quad e_9=c_{16,16},\quad e_{10}=c_{17,17},$$

$e_{11}=c_{20,20}, c_{21}+c_{19,18}, c_{43}+c_{22,21}, c_{51}, c_{52}+c_{76}, c_{53}, c_{54}+c_{10,9}, c_{86}, c_{87},$
$c_{11,9}, c_{11,10}+c_{13,12}, c_{14,12}, c_{14,13}, c_{16,15}, c_{18,17}, c_{19,17}, c_{21,20}, c_{22,20}$

form a P-basis of A. We know that A is of local-colocal type by checking the conditions of Theorem 1.1. It is not hard to verify that A satisfies the conditions of Theorem 5.5. Hence A is a QF-1 algebra. According to [7, Corollary 3.2], if $K \nabla L$ is a nonserial indecomposable left module over a QF-1 algebra of left colocal type, then $|K|=2$ or $|L|=2$. On the other hand, we can see that $(e_5A)^*$ is a left A-module of the form $K \nabla L$ with $|K|=|L|=3$.

Example 6.2. Let A be the subalgebra of full matrix ring $M_{12}(P)$ such that the elements

$$e_1=c_{11}+c_{77}+c_{10,10}, \quad e_2=c_{22}+c_{44}+c_{88}, \quad e_3=c_{33},$$
$$e_4=c_{55}, \quad e_5=c_{66}+c_{12,12}, \quad e_6=c_{99},$$
$$e_7=c_{11,11}, c_{21}+c_{87}, c_{31}, c_{54}, c_{76}, c_{86}, c_{10,9}, c_{12,11}$$

form a P-basis of A. Then this algebra A is of local-colocal type. As easily checked, A satisfies the conditions (1), (2), (3) and (4) of Theorem 5.5, but does not satisfy (5). Hence the condition (5) of Theorem 5.5 is independent of the others. Observe that the condition (5) is not used in the proof of Lemma 4.1. Thus both $E({}_AA)$ and $E(A_A)$ are balanced. But a normal minimal faithful left A-module is not balanced. For, as can be seen by the proof of Proposition 3.8, the condition (5) is obtained from the other conditions (1), (2), (3), (4) and the assumption of the balancedness of a normal minimal faithful module.

References

1. Camillo, V.P.: Balanced rings and a problem of Thrall. Trans. Am. Math. Soc. **149**, 143–153 (1970)
2. Camillo, V.P., Fuller, K.R.: Balanced and QF-1 algebras. Proc. Am. Math. Soc. **34**, 373–378 (1972)
3. Dlab, V., Ringel, C.M.: Rings with the double centralizer property. J. Algebra **22**, 480–501 (1972)
4. Fuller, K.R.: Generalized uniserial rings and their Kupisch series. Math. Z. **106**, 248–260 (1968)
5. Fuller, K.R.: Double centralizers of injectives and projectives over artinian rings. Ill. J. Math. **14**, 658–664 (1970)
6. Fuller, K.R.: Biserial rings. In: Proceedings of a Conference on Ring Theory (Waterloo 1978). Lecture Notes in Mathematics **734**, pp. 64–90. Berlin-Heidelberg-New York: Springer 1979
7. Makino, R.: Double centralizers of minimal faithful modules over left serial algebras. J. Algebra (To appear)
8. Morita, K.: On algebras for which every faithful representation is its own second commutator. Math. Z. **69**, 429–434 (1958)
9. Ringel, C.M.: Socle conditions for QF-1 rings. Pac. J. Math. **41**, 309–336 (1973)
10. Tachikawa, H.: On rings for which every indecomposable right module has a unique maximal submodule. Math. Z. **71**, 200–222 (1959)
11. Tachikawa, H.: On algebras of which every indecomposable representation has an irreducible one as the top or the bottom Loewy constituent. Math. Z. **75**, 215–227 (1961)
12. Tachikawa, H.: Double centralizers and dominant dimensions. Math. Z. **116**, 79–88 (1970)

13. Tachikawa, H.: Balancedness and left serial algebras of finite type. In: Proceedings of the International Conference (Ottawa 1974). Lecture Notes in Mathematics **488**, pp. 351–378. Berlin-Heidelberg-New York: Springer 1975
14. Thrall, R.M.: Some generalizations of quasi-Frobenius algebras. Trans. Am. Math. Soc. **64**, 173–183 (1948)
15. Walker, E.A., Walker, C.L.: Quotient categories and rings of quotients. Rocky Mt. J. Math. **2**, 513–555 (1972)

Received February 7, 1984; received in final form October 30, 1984

Math. Z. 189, 593–602 (1985)

Mathematische Zeitschrift

© Springer-Verlag 1985

Algebraic Cycles on Compact Shimura Surface

K.F. Lai

Department of Mathematics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong

1. Introduction

1.1. Let F be a real quadratic number field and D be a quaternion algebra over F. We associate to D a Shimura surface S^K and study Tate's conjecture [16] on algebraic cycles in S^K by the method of Harder-Langlands-Rapaport [5].

1.2. Given a smooth projective variety X of complex dimension n we have the Hodge decomposition

$$H^k(X,\mathbb{C}) = \bigoplus_{p+q=k} H^{pq}$$

where $0 \leqq k \leqq 2n$. Let $\mathfrak{A}^p(X)$ denote the $\mathbb{Q}$-linear space of the algebraic subvarieties of X of codimension p. Poincare duality gives the cycle map

$$d_p\colon \mathfrak{A}^p(X) \to H^{2p}(X,\mathbb{Q}) \cap H^{pp}.$$

The Hodge conjecture states that this map is surjective.

Tate [16] investigated the ℓ-adic analogue of the Hodge conjecture. Suppose that X above is defined over a number field F. We have the ℓ-adic cycle map

$$d_p\colon \mathfrak{A}^p(X) \otimes \mathbb{Q}_\ell \to H_\ell^{2p}(\bar{X})(p)$$

where $H_\ell^{2p}(\bar{X})(p)$ is the ℓ-adic cohomology group twisted by roots of unity [16]. The action of $\mathrm{Gal}(\bar{F}/F)$ on $\bar{X}$ gives a continuous ℓ-adic representation

$$\rho_{p,\ell}\colon \mathrm{Gal}(\bar{F}/F) \to \mathrm{Aut}_{\mathbb{Q}_\ell}(H_\ell^p(\bar{X})).$$

For each finite extension E of F, let $\mathfrak{A}^p(X,E)$ denote the subspace of $d_{p,\ell}({}^p(X) \otimes \mathbb{Q}_\ell)$ left fixed by $\mathrm{Gal}(\bar{F}/F)$. Tate [16] conjectured (i) that

$$\mathfrak{A}^p(X,E) = [H_\ell^{2p}(\bar{X})(p)]^{\mathrm{Gal}(\bar{F}/E)}.$$

One can attach to the representation $\rho_{p,\ell}$ a L-function $L^p(X/F,s)$ which is defined by an Euler product (denoted by Φ_p in Serre [14]); this product converges for $\mathrm{Re}(s) > 1 + p/2$.

Tate [16] conjectured (ii) that $L^{2p}(X/E,s)$ has a pole at $s=p+1$ of order equal to

$$\dim_{\mathbb{Q}_\ell}(H_\ell^{2p}(\bar{X})(p))^{\mathrm{Gal}(\bar{F}/E)}.$$

In the rest of this paper we shall mean by Tate's conjecture over E both (i) and (ii) above.

1.3. Tate's conjecture has been verified for $p=1$ by Pohlmann [13] in case X is an abelian variety of C.M. type and by V. Kumar Murty [11] in case X is a quotient of $J_1(N)$ associated to noncomplex multiplication newforms of weight 2. J. Tunnel (unpublished) proved Tate's conjecture for the product of two modular curves. Inspired by Tunnel's methods, Harder-Langlands-Rapoport [5] proved Tate's conjecture for compactified Hilbert modular surfaces with respect to cyclotomic extensions E. In this note we make the observation that with the help of a suitable trace formula [7], from the results of Harder-Langland-Rapoport we can deduce Tate's conjecture over cyclotomic extension for a compact Shimura surface S^K associated to a quaternion algebra satisfying the assumptions given in §1.4. We shall assume that the readers are familar with [5] and avoid repeating their proofs.

Inside the subspace $\mathfrak{A}$ of algebraic cycles in $H^2(S^K)$ is the subspace $\mathfrak{Z}$ of Hirzebruch-Zagier cycles (cf. [6]). The cohomology group $H^2(S^K)$ (resp. $\mathfrak{A}$, $\mathfrak{Z}$) is decomposed into a direct sum parametrized by characters ω of $\mathrm{Gal}(E/\mathbb{Q})$ and automorphic representations π' (see §2.6). In the case when π' is distinguished (for definition see §3), with the help of the relative trace formula [7] we can deduce from [5] that the algebraic cycles are Hirzebruch-Zagier cycles. In all the cases we establish that the dimension of the subspace of $\mathrm{Gal}(\bar{\mathbb{Q}}/E)$ invariant algebraic cycles parametrized by (ω,π') is equal to the order of the pole of the corresponding factor of the L function of $H^2(S^K)$; for this purpose we use implicitly the results of Langlands [10] which expresses the zeta function of S^K in terms of automorphic L-functions and we take into the account of the action of the Hecke operators.

1.4. We introduce some notations which we shall use unless otherwise specified.

$\mathbb{Q}$ is the field of rational numbers. p, ℓ are prime numbers. $\mathbb{A}$ (resp. $\mathbb{A}_f$) is the ring of (resp. finite) adeles of $\mathbb{Q}$. Fix algebraic closures $\bar{\mathbb{Q}}$, $\bar{\mathbb{Q}}_\ell$ of $\mathbb{Q}$, $\mathbb{Q}_\ell$ respectively. Choose once for all embeddings $\bar{\mathbb{Q}}\hookrightarrow\bar{\mathbb{Q}}_\ell\hookrightarrow\mathbb{C}$.

Assumption. F is a real quadratic extension of $\mathbb{Q}$. B is a quaternion algebra over $\mathbb{Q}$. Put $D=B\underset{\mathbb{Q}}{\otimes}F$. Assume that D is a quaternion (division) algebra over F which splits at the real places and F splits over the places where B ramifies.

H' is the algebraic group defined over $\mathbb{Q}$ whose $\mathbb{Q}$-rational points $H'(\mathbb{Q})=B^\times$. G' is the algebraic group over $\mathbb{Q}$ such that $G'(\mathbb{Q})=D^\times$.

Define h: $C^\times\to G(\mathbb{R})$ to be the $\mathbb{R}$-homomorphism such that $h(\sqrt{-1})$ is the right multiplication on $D\underset{\mathbb{Q}}{\otimes}\mathbb{R}\cong M(2,\mathbb{R})\times M(2,\mathbb{R})$ by $\left(\begin{pmatrix}0&1\\-1&0\end{pmatrix},\begin{pmatrix}0&1\\-1&0\end{pmatrix}\right)$. Let K_∞ be the centralizer of h in $G'(\mathbb{R})$. For a sufficiently small compact open

subgroup K of $G'(\mathbb{A}_f)$, the above data define a proper smooth Shimura surface S^K over $\mathbb{Q}$ and

$$S^K(\mathbb{C}) = G'(\mathbb{Q}) \backslash G'(\mathbb{A}) / K_\infty K.$$

2. Cohomology

2.1. Let ξ be an absolutely irreducible finite dimensional $\mathbb{Q}$-rational representation of G'. For any field E containing $\mathbb{Q}$ we associate to ξ a local system $\mathscr{F}_\xi(E)$ on $S^K(\mathbb{C})$ of vector spaces over E.

By universal coefficient theorem we have

$$H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(E)) \approx H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{Q})) \underset{\mathbb{Q}}{\otimes} E.$$

2.2. Suppose $g \in G'(\mathbb{A}_f)$, K' is another open compact subgroup of $G'(\mathbb{A}_f)$ and $K' \subseteq gKg^{-1}$. The homomorphism $G'(\mathbb{A}) \to G'(\mathbb{A})$: $h \mapsto hg$ defines a morphism

$$r(g): S^{K'}(\mathbb{C}) \to S^K(\mathbb{C}).$$

Now take $K' = gKg^{-1} \cap K$ and write $\mathscr{F}'_\xi(\mathbb{C})$ for the local system on $S^{K'}(\mathbb{C})$ defined by ξ. We define the Hecke operator

$$T(g): H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{C})) \to H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{C}))$$

to be the composition of

$$\begin{aligned} H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{C})) &\to H^i(S^{K'}(\mathbb{C}), r(g)^* \mathscr{F}_\xi(\mathbb{C})) \\ &\to H^i(S^{K'}(\mathbb{C}), \mathscr{F}'_\xi(\mathbb{C})) \to H^i(S^K(\mathbb{C}), r(1)_* \mathscr{F}'_\xi(\mathbb{C})) \to H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{C})) \end{aligned}$$

where the first arrow is given by the definition of $r(g)^*$, the third arrow is an isomorphism because the Leray spectral sequence of a finite morphism degenerates, the second and the fourth arrows are given respectively by the correspondences $r(g)^* \mathscr{F}_\xi(\mathbb{C}) \to \mathscr{F}'_\xi(\mathbb{C})$, $r(1)_* \mathscr{F}'_\xi(\mathbb{C}) \to \mathscr{F}_\xi(\mathbb{C})$ (see [9], p. 390, 391 resp.).

2.3. Let $\mathrm{Coh}_1(\xi)$ be the set of automorphic representation $\pi' = \pi'_\infty \otimes \pi'_f$ of $G'(\mathbb{A})$ such that $\pi'_\infty \cong \xi^{-1}$ if ξ is one dimensional and be empty otherwise. Let $\mathrm{Coh}_0(\xi)$ be the set of automorphic representations $\pi' = \pi'_\infty \otimes \pi'_f$ of $G'(\mathbb{A})$ such that π'_∞ lies in the discrete series and the relative Lie algebra cohomology $H^2(\mathfrak{g}, K_\infty, \pi'_\infty \otimes \xi) \neq 0$ ($\mathfrak{g}$ is the Lie algebra of $G'(\mathbb{R})$). Let

$$\mathrm{Coh}(\xi) = \mathrm{Coh}_0(\xi) \cup \mathrm{Coh}_1(\xi).$$

Since $S^K(\mathbb{C})$ is compact, we can compute $H^2(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{C}))$ by means of L^2 cohomology and obtain a decomposition ([2, 3, 9])

$$H^2(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{C})) = \bigoplus_{\pi' \in \mathrm{Coh}(\xi)} H_{\pi', \infty}(\mathbb{C}) \otimes H_{\pi', f}(\mathbb{C})$$

where $H_{\pi',f}(\mathbb{C})$ is the finite dimensional space of K invariant vectors in the space of π'_f and

$$\dim_{\mathbb{C}} H_{\pi',\infty}(\mathbb{C}) = \begin{cases} 2 \text{ if } \pi'_\infty \text{ is 1 dimensional,} \\ 4 \text{ if } \pi'_\infty \text{ lies in the discrete series.} \end{cases}$$

2.4. We can associate to ξ an étale sheaf of $\mathbb{Q}_\ell$ vector spaces $\mathscr{F}_\xi(\mathbb{Q}_\ell)_{\text{ét}}$ on $(S^K_{\mathbb{C}})_{\text{ét}}$. By comparison theorem,

$$H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\mathbb{Q}_\ell)) \approx H^i((S^K_{\mathbb{C}})_{\text{ét}}, \mathscr{F}_\xi(\mathbb{Q}_\ell)_{\text{ét}}).$$

Composing with the specialization map, we have $H^i(S^K(\mathbb{C}), \mathscr{F}_\xi(\bar{\mathbb{Q}}_\ell)) \approx H^i((S^K_{\bar{\mathbb{Q}}})_{\text{ét}}, \mathscr{F}_\xi(\bar{\mathbb{Q}}_\ell)_{\text{ét}})$. We shall drop the suffix ét and the index $\bar{\mathbb{Q}}$ in the rest of this paper.

On the étale cohomology groups $H^i(S^K, \mathscr{F}_\xi(\bar{\mathbb{Q}}_\ell))$ we also have Hecke operators $T(g)$ operating. It follows from Langlands yoga ([9], p. 37) that we can derive from 2.3 a decomposition of $H^i(S^K, \mathscr{F}_\xi(\bar{\mathbb{Q}}_\ell))$ over $\pi' \in \text{Coh}(\xi)$ in such a way that the action of the Hecke operators on $H_{\pi',v}(\mathbb{C})$ (resp. $H_{\pi',v}(\bar{\mathbb{Q}}_\ell)$) can be obtained by tensoring with $\mathbb{C}$ (resp. $\bar{\mathbb{Q}}_\ell$) the action of the Hecke operators on $H_{\pi',f}(\bar{\mathbb{Q}})$, for $v = \infty, f$. We have

$$H^2(S^K, \mathscr{F}_\xi(\bar{\mathbb{Q}}_\ell)) = \bigoplus_{\pi' \in \text{Coh}(\xi)} H_{\pi',\infty}(\bar{\mathbb{Q}}_\ell) \otimes H_{\pi',f}(\bar{\mathbb{Q}}_\ell).$$

2.5. Write $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ for the Galois group of $\bar{\mathbb{Q}}/\mathbb{Q}$. For $\sigma \in \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, let σ act on $\text{Spec}\,\bar{\mathbb{Q}} \to \text{Spec}\,\mathbb{Q}$ to the left so that the corresponding action on $\bar{\mathbb{Q}}$ is $x \to \sigma^{-1}(x)$, and let σ act on

$$S^K_{\bar{\mathbb{Q}}} = S^K \times_{\text{Spec}\mathbb{Q}} \text{Spec}\,\bar{\mathbb{Q}}$$

through the action of σ on the second factor. This yields a representation ρ of $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on $H^i(S^K_{\bar{\mathbb{Q}}}, \mathscr{F}_\xi(\bar{\mathbb{Q}}_\ell))$. $\rho(\sigma)$ acts to the right while $T(g)$ acts to the left and the two actions commute by construction. It follows that corresponding to the decomposition in § 2.4 we have ([9], p. 401)

$$\rho = \bigoplus_{\pi' \in \text{Coh}(\xi)} \rho(\pi') \otimes 1.$$

From now on we shall omit $\mathscr{F}_\xi$ in the notation when ξ is the trivial representation.

2.6. Let α: $\text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{Q}_\ell^\times$ denote the Tate character - it is defined by the conditions -

$$\sigma(\zeta) = \zeta^{\alpha(\sigma)}$$

for $\sigma \in \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ and $\zeta^{\ell^n} = 1$.

Fix a cyclotomic extension E of $\mathbb{Q}$.

For a character ω of $\text{Gal}(E/\mathbb{Q})$, let

$$\mathfrak{B}(\omega) = \{x \in H^2(S^K_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell) \mid \sigma \cdot x = \alpha(\sigma)\omega^{-1}(\sigma)x,\ \forall \sigma \in \text{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})\}$$

(note that $\omega(\sigma) \in \bar{\mathbb{Q}} \subset \bar{\mathbb{Q}}_\ell$). Elements of $\mathfrak{B}(\omega)$ are called Tate cycles.

Let $\mathfrak{A}$ be the subspace of $H^2(S^K_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell)$ generated by algebraic cycles defined over cyclotomic extensions of $\mathbb{Q}$. Put $\mathfrak{A}(\omega)=\mathfrak{A}\cap\mathfrak{B}(\omega)$.

Let L_∞ be the centralizer of the homomorphism

$$C^\times \to H'(\mathbb{R}): a+b\sqrt{-1}\mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

and L be a sufficiently small compact open subgroup of $H'(\mathbb{A}_f)$ and J^L be the Shimura curve over $\mathbb{Q}$ such that $J^L(\mathbb{C})=H'(\mathbb{Q})\backslash H'(\mathbb{A})/L_\infty L$. There is a homomorphism $H'\to G'$ defined over $\mathbb{Q}$ which embeds H' diagonally in G' and defines a morphism $J^L\to S^K$ over $\mathbb{Q}$ if the image of L lies inside K.

Poincaré duality yields a homomorphism

$$H^0(J^L_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell)\to H^2(S^K_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell).$$

Let $\mathfrak{Z}$ be the subspace of $H^2(S^K_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell)$ spanned by the images of $H^0(J^L_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell)$ (for all suitable L) under the action of the Hecke operators $T(g)$, $g\in G(\mathbb{A}_f)$. Put $\mathfrak{Z}(\omega)=\mathfrak{Z}\cap\mathfrak{B}(\omega)$. Elements of $\mathfrak{Z}$ are called Hirzebruch-Zagier cycles.

Corresponding to the decomposition in §2.4, we have the following decompositions:

$$\mathfrak{A}(\omega)=\bigoplus_{\pi'\in \mathrm{Coh}(1)} \mathfrak{A}(\omega,\pi')\otimes H_{\pi',f}(\bar{\mathbb{Q}}_\ell),$$

$$\mathfrak{B}(\omega)=\bigoplus_{\pi'\in \mathrm{Coh}(1)} \mathfrak{B}(\omega,\pi')\otimes H_{\pi',f}(\bar{\mathbb{Q}}_\ell),$$

$$\mathfrak{Z}(\omega)=\bigoplus_{\pi'\in \mathrm{Coh}(1)} \mathfrak{Z}(\omega,\pi')\otimes \mathrm{H}_{\pi',f}(\bar{\mathbb{Q}}_\ell).$$

Let $\mathscr{S}$ be a finite set of places of $\mathbb{Q}$ (including ∞ and ℓ) outside which S^K has good reduction. Let $C(\omega,\pi')$ be the order of pole at $s=2$ of the L-function $L_{\mathscr{S}}(s,\omega\rho(\pi'))$.

Let $A(\omega,\pi')$, $B(\omega,\pi')$, $Z(\omega,\pi')$ be the dimension of $\mathfrak{A}(\omega,\pi')$, $\mathfrak{B}(\omega,\pi')$, $\mathfrak{Z}(\omega,\pi')$ respectively. Then Tate's conjecture over the cyclotomic extension E for algebraic cycles in $H^2(S^K_{\bar{\mathbb{Q}}}, \bar{\mathbb{Q}}_\ell)$ follows immediately from the following

2.7. **Theorem.** $Z(\omega,\pi')\leqq A(\omega,\pi')=B(\omega,\pi')=C(\omega,\pi')$.

First of all, it follows from definitions that $Z(\omega,\pi')\leqq A(\omega,\pi')\leqq B(\omega,\pi')$. We shall deal with the case when π'_∞ lies in the discrete series and then the case when π'_∞ is the trivial representation.

3. Distinguished Representation

Let H be $GL(2)_{\mathbb{Q}}$ and G be $R_{F/\mathbb{Q}}GL(2)_F$. Let ε be the character of the extension $F/\mathbb{Q}$, ω be a character of $\mathbb{Q}^\times\backslash\mathbb{Q}^\times(\mathbb{A})$ of finite order.

Recall ([15]) that an automorphic representation [1] π of G is *distinguished* (with respect to ω^{-1}) if (i) π_∞ lies in the discrete series; (ii) $\pi=\pi^0\otimes\lambda^{-1}$ where (iii) the automorphic representation π^0 of G is the image by base change [11] of a cuspidal automorphic representation $\pi^0_{\mathbb{Q}}$ of H with central character $\omega^0_{\mathbb{Q}}$ and (iv) λ is a character of $F^\times\backslash F^\times(\mathbb{A})$ whose restriction to $\mathbb{Q}^\times\backslash\mathbb{Q}^\times(\mathbb{A})$ is $\omega^0_{\mathbb{Q}}\omega\varepsilon^{-1}$. (Note that the central character of π^0 is $\omega^0=\omega^0_{\mathbb{Q}}\circ N_{F/\mathbb{Q}}$.)

We say that an automorphic representation π' of G' is *distinguished* (with respect to ω^{-1}) if the automorphic representation π of G associated to π' by the Jacquet-Langlands correspondence ([8], p. 494) is distinguished.

4. Hirzebruch-Zagier Cycles

We study here the space $\mathfrak{Z}(\omega, \pi')$.

4.1. Let $\mathfrak{g}, \mathfrak{k}$ be the Lie algebra of $G'(\mathbb{R})$, K_∞ respectively. K_∞ acts on $\mathfrak{g}/\mathfrak{k}$ via the adjoint representation. Fix a basis $\{\omega_i | 1 \leqq i \leqq 6\}$ of $\Lambda^2 \mathfrak{g}/\mathfrak{k}$. Then an element of $\mathrm{Hom}_{K_\infty}(\Lambda^2 \mathfrak{g}/\mathfrak{k}, \pi'_\infty) \otimes H_{\pi', f}(\mathbb{C})$ can be written as $\sum_{i=1}^{6} f_i \omega_i$ where f_i is a square integrable automorphic form on $G'(\mathbb{Q}) \backslash G'(\mathbb{A}) / K_\infty K$. The ω_i can be taken as left invariant 2 forms on $G'(\mathbb{R})$ right invariant under K_∞ and so on $G'(\mathbb{A})/K_\infty K$.

Now consider the Hirzebruch-Zagier cycle J^L of S^K (image of L lies in K). The space $H^0(J^L(\mathbb{C}), \mathbb{C})$ can be identified with the space of locally constant functions on $H'(\mathbb{Q}) \backslash H'(\mathbb{A}) / L_\infty L$. Poincaré duality yields a homomorphism $H^0(J^L(\mathbb{C}), \mathbb{C}) \to H^2(S^K(\mathbb{C}), \mathbb{C})$ and a pairing: for $\mu \in H^0(J^L(\mathbb{C}), \mathbb{C})$, $\sum f_i \omega_i \in H_{\pi', \infty}(\mathbb{C}) \otimes H_{\pi', f}(\mathbb{C})$,

4.2. $\langle \mu, \sum f_i \omega_i \rangle = \int_{H'(\mathbb{Q}) \backslash H'(\mathbb{A}) / L_\infty L} \mu \sum f_i \nu_i$

where ν_i is the 2 form on $H'(\mathbb{R})$ lifted from ω_i via the embedding $H'(\mathbb{R}) \hookrightarrow G'(\mathbb{R})$.

Extend ω to a character of $H'(\mathbb{A})$ by composing with the norm map. Let I be the representation obtained by inducing ω from $H'(\mathbb{A})$ to $G'(\mathbb{A})$. For f in the space of π', let $\mathscr{T}_{\pi'} f$ be the function on $G'(\mathbb{A})$ given by

$$g \mapsto \int_{H'(\mathbb{Q}) Z_{H'}(\mathbb{R}) \backslash H'(\mathbb{A})} \omega^{-1}(h) f(hg)\, dh.$$

Here $Z_{H'}$ is the centre of H'. Then $\mathscr{T}_{\pi'} \in \mathrm{Hom}_{G'(\mathbb{A})}(\pi', I)$.

4.3. **Lemma.** *When* $H_{\pi', \infty}(\bar{\mathbb{Q}}_\ell) \otimes H_{\pi', f}(\bar{\mathbb{Q}}_\ell) \neq 0$, *it is not orthogonal to* $\mathfrak{Z}(\omega, \pi')$ *if and only if* $\mathscr{T}_{\pi'} \neq 0$.

Let k be a finite extension of $\mathbb{Q}$ containing the values of ω. Let $\mathfrak{Z}(\omega, \pi')_k$ be defined in the same way as $\mathfrak{Z}(\omega, \pi')$ but with k replacing $\bar{\mathbb{Q}}_\ell$. Then $\mathfrak{Z}(\omega, \pi')_k \otimes \bar{\mathbb{Q}}_\ell = \mathfrak{Z}(\omega, \pi')$. Now 4.3 follows from 4.2, the definition of $\mathscr{T}_{\pi'}$ and the fact that the Hecke operators are self-adjoint with respect to the pairing.

4.4. **Corollary.** $Z(\omega, \pi') > 0$ *if and only if* $\mathscr{T}_{\pi'} \neq 0$.

By the lemma, $\mathscr{T}_{\pi'} \neq 0$ is equivalent to $\mathfrak{Z}(\omega, \pi') \neq 0$.

4.5. **Lemma.** *Let* π *correspond to* π' *under the Jacquet-Langlands correspondence and* $\mathscr{T}_\pi$ *be defined in a similar way. Then* $\mathscr{T}_{\pi'} \neq 0$ *if and only if* $\mathscr{T}_\pi \neq 0$.

This is proved in [7].

4.6. It is proved in [5] that (i) $\mathscr{T}_\pi \neq 0$ if and only if π is distinguished with respect to ω^{-1}; (ii) $C(\omega,\pi)\geqq 0$ and $C(\omega,\pi)>0$ if and only if π is distinguished, in which case $C(\omega,\pi)=1$.

The next lemma follows from the definitions of distinguished representation and the Jacquet-Langlands correspondence.

4.7. **Lemma.** *Let* $\pi'\in \mathrm{Coh}_0(1)$.

(i) $Z(\omega,\pi')>0$ *if and only if* π' *is distinguished.*

(ii) $C(\omega,\pi')\geqq 0$, *and* $C(\omega,\pi')>0$ *if and only if* π' *is distinguished, in which case* $C(\omega,\pi')=1$.

5. Tate classes

We study in this section $B(\omega,\pi')$ for $\pi'\in \mathrm{Coh}_0(1)$.

5.1. **Lemma.** π' *is distinguished implies* $B(\omega,\pi')=1$.

If π' is distinguished, then so is the corresponding representation π. By [5] (Satz 4.5.7), there exists an irreducible 2 dimensional ℓ-adic representation $\tau=\tau(\pi')$ of $\mathrm{Gal}(\bar{F}/F)$ such that for almost all places v of F,

$$L_v(s,\tau(\pi'))=L_v(s-\tfrac{1}{2},\pi')$$

where the right hand side is the local automorphic L function of π'. Fix an element σ from $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})-\mathrm{Gal}(\bar{F}/F)$. Then

$$\eta\mapsto \tau(\eta)\times\tau(\sigma\eta\sigma^{-1})\times\eta,$$
$$\sigma\mapsto 1\times\tau(\sigma^2)\times\sigma$$

($\eta\in\mathrm{Gal}(\bar{F}/F)$) defines a homomorphism ψ: $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})\to {}^L G'_{\mathbb{Q}_\ell}$, where ${}^L G'$ is the L-group of G'. We extend the natural action r of $GL(2,\mathbb{C})\times GL(2,\mathbb{C})$ on $\mathbb{C}^2\otimes\mathbb{C}^2$ to a representation of ${}^L G'$ by setting

$$r(\mathrm{Id}\times\mathrm{Id}\times\sigma)(x\otimes y)=y\otimes x.$$

Furthermore we identify $\mathbb{C}^2\otimes\mathbb{C}^2$ with the space of 2×2 matrics (cf. [5] §4.3). Eichler-Shimura relation [8] implies that $\rho(\pi')$ is equivalent to the irreducible representation $r\circ\psi$ on the space of 2×2 matrices. The condition

$$\rho(\pi')\cdot x=a\omega^{-1}\cdot x$$

on $\mathfrak{B}(\omega,\pi')$ then correspond to the condition

$$\tau(\sigma\eta\sigma^{-1})=\chi(\eta)\,C\tau(\eta)\,C^{-1}$$

for all $\eta\in\mathrm{Gal}(\bar{F}/F)$, all 2×2 matrices C and $\chi=\alpha\cdot\omega^{-1}\cdot\det^{-1}\tau$. It follows that $B(\omega,\pi')\leqq 1$. But π' is distinguished implies $0<Z(\omega,\pi')\leqq B(\omega,\pi')$. Therefore $B(\omega,\pi')=1$.

5.2. **Lemma.** *$B(\omega,\pi')>0\Rightarrow\pi'$ distinguished.*

If one replaces the intersection cohomology in [5], §4 by cohomology then what remains, word for word, yields a proof of the fact that $\bar{\pi}'=\tilde{\pi}'\otimes\chi$, where $\chi=\omega^{-1}$ restricted to $\mathrm{Gal}(\bar{F}/F)$. If π is the representation of G corresponding to π' under the Jacquet-Langlands correspondence, then $\bar{\pi}=\tilde{\pi}\otimes\chi$ and from this it follows that π is distinguished ([5], §4) and so is π'.

5.3. **Corollary.** *$B(\omega,\pi')=0$ or 1. $B(\omega,\pi')=1$ if and only if π' is distinguished.*

6. Trivial at the Infinite Place

We shall assume in this section $\pi'\in\mathrm{Coh}_e(1)$. That is $\pi'_\infty\cong 1$ and so π' is one dimensional, $\pi'=\nu\circ\mathrm{Norm}$ for some character ν of $F^\times\backslash F^\times(\mathbb{A})$ of finite order. Let $\chi_{\pi'}$ be the restriction of ν to $\mathbb{Q}^\times\backslash\mathbb{A}^\times$ and identify it with the corresponding character on the Galois group.

6.1. **Lemma.** *If $H_{\pi',f}(\bar{\mathbb{Q}}_\ell)\neq 0$ then for almost all p we have*

$$L_p(s,\rho(\pi'))=\frac{1}{1-\dfrac{p\chi_{\pi'}(p)}{p^s}}\,\frac{1}{1-\dfrac{p\varepsilon(p)\chi_{\pi'}(p)}{p^s}}.$$

For $i=1,3$, $H^i(S^K,\mathbb{Q}_\ell)=0$. For $i=0,4$ we have

$$\mathrm{H}^i(S^K,\bar{\mathbb{Q}}_\ell)=\bigoplus_{\pi'\in\mathrm{Coh}_e(1)} H^i_{\pi',\infty}(\bar{\mathbb{Q}}_\ell)\otimes H^i_{\pi',f}(\bar{\mathbb{Q}}_\ell)$$

and $\dim H^{i,\pi'}_\infty(\bar{\mathbb{Q}}_\ell)=1$. Now 6.1 follows from the known value of the zeta function of the Shimura surface [10] and the fact that the Galois group acts by $\chi_{\pi'}$, $\alpha^2\chi_{\pi'}$ respectively on $H^0_{\pi',\infty}(\bar{\mathbb{Q}}_\ell)$, $H^4_{\pi',\infty}(\bar{\mathbb{Q}}_\ell)$ respectively.

6.2. **Corollary.** *$B(\omega,\pi')=C(\omega,\pi')=1$ if $\omega^{-1}=\chi_{\pi'}$ or $\varepsilon\chi_{\pi'}$ and is zero otherwise.*

This follows from 6.1 which gives the eigen characters of the action of $\rho(\pi')$ on $H_{\pi',\infty}(\mathbb{Q}_\ell)$ as $\chi_{\pi'}$ and $\varepsilon\cdot\chi_{\pi'}$.

6.3. **Lemma.** *$Z(\omega,\pi')=1$ if $\omega^{-1}=\chi_{\pi'}$ and is zero otherwise.*

Note that 4.3 is actually valid for $\pi\in\mathrm{Coh}_e(1)$. That is $\mathfrak{Z}(\omega,\pi')\neq 0$ if and only if $\mathscr{T}_{\pi'}\neq 0$ and so the lemma follows.

6.4. **Lemma.** *$A(\omega,\pi')=1$ if $\omega^{-1}=\chi_{\pi'}$ or $\varepsilon\chi_{\pi'}$ and is zero otherwise.*

As $A(\omega,\pi')\leqq B(\omega,\pi')$, Corollary 6.2 implies $A(\omega,\pi')\leqq 1$ and if $A(\omega,\pi')=1$ then $\omega^{-1}=\chi_{\pi'}$ or $\varepsilon\chi_{\pi'}$. From $Z(\omega,\pi')\leqq A(\omega,\pi')\leqq 1$ and Lemma 6.3, we see that if $\omega^{-1}=\chi_{\pi'}$, then $A(\omega,\pi')=1$. It remains to show that if $\omega^{-1}=\varepsilon\chi_{\pi'}$ then $A(\omega,\pi')\neq 0$. This follows from 6.1 and the fact that Lefschetz theorem on $(1-1)$ class implies that

$$\bigoplus_{\pi'\in\mathrm{Coh}_e(1)} H_{\pi',\infty}(\mathbb{Q})\otimes H_{\pi',f}(\mathbb{Q})$$

consists only of algebraic cycles if it consists of $(1-1)$ forms only. We let $\mathfrak{g}^i$ $(i=1,2)$ be the Lie algebra of $GL(2,\mathbb{R})$ and K_∞^i $(i=1,2)$ be $O(2,\mathbb{R})$. Then the relative Lie algebra cohomology of $G'(\mathbb{R})=GL(2,\mathbb{R})\times GL(2,\mathbb{R})$ is $\sum H^i(\mathfrak{g}^1,K_\infty^1,\pi'_\infty)\otimes H^j(\mathfrak{g}^2,K_\infty^2,\pi'_\infty)$. The contribution to $H^2(\mathfrak{g},K_\infty,\pi'_\infty)$ is the sum of $H^0(\mathfrak{g}^1,K_\infty^1,\pi'_\infty)\otimes H^2(\mathfrak{g}^2,K_\infty^2,\pi'_\infty)$ and $H^2(\mathfrak{g}^1,K_\infty^1,\pi'_\infty)\otimes H^0(\mathfrak{g}^2,K_\infty^2,\pi'_\infty)$ and we are done.

7. End of Theorem 2.7

Theorem 2.7 follows from 4.7 and 5.3 if $\pi'\in\mathrm{Coh}_0(1)$ and from 6.2, 6.3 and 6.4 if $\pi'\in\mathrm{Coh}_e(1)$.

8. Remarks

8.1 From now on let us drop the convention adopted in § 1.4.

Let us suppose that we are given a quaternion algebra D over a real quadratic field F such that (i) D is unramified at infinity, (ii) if w, $\bar{w}$ are two different primes of F dividing a common prime v of $\mathbb{Q}$ then D has the same invariant at both w and $\bar{w}$, (iii) if w is the unique prime dividing a prime v of $\mathbb{Q}$, then the invariant of D is zero. In this situation, $D=B\otimes_{\mathbb{Q}}F$ for some quaternion algebra B over $\mathbb{Q}$. Let $\mathscr{S}_\eta$ be the set of primes v of $\mathbb{Q}$ which splits into two different primes w, $\bar{w}$ of F at which D has invariant $1/2$ (i.e. D_w and $D_{\bar{w}}$ are ramified). We define the parity η of D to be the number of elements in $\mathscr{S}_\eta$.

8.2. If η is even, then we can choose a B such that the assumption in § 1.4 is satisfied, because the only constraint on the invariants of B is that their sum be zero mod 1. Therefore, by Theorem 2.7, Tate's conjecture over cyclotomic field is true for the Shimura surface S^K (cf. § 1.4) constructed from an indefinite quaternion algebra of even parity over a real quadratic field.

8.3. If η is odd, then for each prime p of $\mathbb{Q}$ which does not split in F, choose a quaternion algebra B^p over $\mathbb{Q}$ which is ramified at all the primes of $\mathscr{S}_\eta\cup\{p\}$, unramified otherwise and $D=B^p\otimes_{\mathbb{Q}}F$. Let $\mathscr{B}=\{B^p\}$. For each B such that $D=B\otimes_{\mathbb{Q}}F$, we can construct the space $\mathfrak{Z}_B$ of Hirzebruch-Zagier cycles arising from the Shimura curve constructed from B as in § 2.6. Similarly we have the space $\mathfrak{Z}_B(\omega,\pi')$ and the operator $\mathscr{T}_{B,\pi'}$. Fix an ω as in § 2.6, the question is: given a distinguished automorphic representation π' of G' (as § 1.4), is there an $B\in\mathscr{B}$ such that $\mathscr{T}_{B,\pi'}\neq 0$? Tate's conjecture for an D with odd parity would follow from an affirmative answer to this question.

I would like to thank Professor R.P. Langlands for the suggestion of and many conversations on this problem and the I.A.S. for a year's hospitality.

References

1. Borel, A., Jacquet, H.: Automorphic forms and automorphic representations. Proc. Symp. Pure Math. **33**, 189–202 (1979)
2. Borel, A., Jacquet, H., Wallach, N.: Continuous cohomology, discrete subgroups, and representations of reductive groups. Annals of Math. Studies. Princeton: Princeton University Press 1980
3. Casselman, W.: Hasse-Weii ζ-function. Proc. Symp. Pure Math. **33-2**, 141–164 (1979)
4. Deligne, P.: Variétés de Shimura. Proc. Symp. Pure Math. **33**, 247–290 (1979)
5. Harder, G., Langlands, R.P., Rapaport, M.: Algebraische Zyklen auf Hilbert-Blumenthal Flächen. Preprint
6. Hirzebruch, F., Zagier, D.: Intersection numer of curves on Hilbert modular surfaces and modular forms of Nebentypus. Invent. Math. **36**, 57–113 (1967)
7. Jacquet, H., Lai, K.F.: Sur une formule des traces relative. C.R. Acad. Sci. Paris t 296 (juin 1983) 959–963
8. Jacquet, H., Langlands, R.P.: Automorphic forms on $GL(2)$. Springer Lecture Notes in Math. **114**. Berlin-Heidelberg New York: Springer 1970
9. Langlands, R.P.: Modular forms and $l$adic representations. Springer Lecture Notes in Math. **349**, pp. 362–500. Berlin Heidelberg New York: Springer 1973
10. Langlands, R.P.: On the zeta-functions of some simple Shimura varieties. Can. J. Math. **31**, 1121–1216 (1979)
11. Langlands, R.P.: Base change for $GL(2)$, Annals of Math. Studies. Princeton: Princeton University Press 1980
12. Kumar Murty, V.: Algebraic cycles on abelian varieties. Duke Math. J. **50**, 487–504 (1983)
13. Pohlmann, H.: Algebraic cycles on abelian varieties of complex multiplication type. Ann. Math. **88**, 161–180 (1968)
14. Serre, J.P.: Abelian ℓ-adic representations and elliptic curves. New York: Benjamin 1968
15. SGA5, Cohomologie ℓ-adique et Fonctions L. Springer Lecture Notes in Math. **589**. Berlin Heidelberg New York: Springer 1977
16. Tate, J.: Algebraic cycles and poles of zeta functions, 93–110, in Arithmetical Algebraic Geometry, ed. by O.F.G. Schilling. New York: Harpers & Row 1965

Received July 12, 1984