Pacific Journal of Mathematics

Volume 188 No. 2 April 1999

PACIFIC JOURNAL OF MATHEMATICS

http://www.pjmath.org

Founded in 1951 by

E. F. Beckenbach (1906–1982) F. Wolf (1904–1989)

See inside back cover or www.pjmath.org for submission instructions.

Regular subscription rate for 2006: $425.00 a year (10 issues). Special rate: $212.50 a year to individual members of supporting institutions. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840 is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS
at the University of California, Berkeley 94720-3840
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

WILLMORE–CHEN TUBES ON HOMOGENEOUS SPACES IN WARPED PRODUCT SPACES

J. Arroyo, M. Barros, and O.J. Garay

We present a new method to obtain Willmore–Chen submanifolds in spaces endowed with warped product metrics and fibers being a given homogeneous space. The main points are: First the invariance of the variational problem of Willmore–Chen with respect to the conformal changes in the ambient space metric. Second, the principle of symmetric criticality which allows us to relate the problem with that for generalized elastic curves in the conformal structure on the base.

We obtain some applications of our method, including one, to get a rational one parameter family of Willmore tori in the standard 3-sphere shaped on an associated family of closed free elastic curves in the standard hyperbolic 2-plane.

We also get a 3-dimensional Riemannian manifold which is foliated with leaves being nontrivial Willmore tori.

We start by recalling that Willmore–Chen submanifolds in a Riemannian manifold (M, g) are the critical points of the Willmore–Chen functional $\mathcal{W}$, which is defined by

$$\mathcal{W}(N) = \int_N (\alpha^2 + \tau_e)^{\frac{n}{2}} dv, \tag{1}$$

where N is an n-dimensional compact submanifold immersed in M. The mean curvature function and the extrinsic scalar curvature function of N in (M, g) are α and τ_e respectively and dv denotes the volume element relative to the induced metric by g in N. This functional is known to be invariant under conformal changes of the metric g of M, [**9**]. When $n = 2$, it coincides with the Willmore functional and then its critical points are the Willmore surfaces. Minimal surfaces in the standard sphere (of any dimension) are obvious examples of Willmore surfaces. Articles showing different methods to construct examples of non-minimal Willmore surfaces in spheres are known in the literature (see for example, [**3**, **7**, **10**, **12**, **15**]) even in spaces with no constant curvature (see, for example, [**1**, **2**]).

However the first non-trivial known examples of Willmore–Chen submanifolds (of course with dimension greater than two, exactly they have dimension four) were obtained in [**6**].

The Euler-Lagrange equations associated with the Willmore functional, in spaces of constant curvature, were computed in [**16**]. Now the nice family of Willmore tori, in the standard 3-sphere, obtained in [**15**], can also be directly obtained as a family of solutions of these Euler-Lagrange equations. In fact, it is not difficult to see that these equations for Hopf tori reduce to the Euler-Lagrange equations for certain elastic curves on the standard 2-sphere, [**11**].

Let (M_1, g_1) and (M_2, g_2) be two Riemannian manifolds with dimensions n_1 and n_2 respectively. Given a positive function f on M_1 (assume $\inf f > 0$ if M_1 is not compact) we define on $M_1 \times M_2$ the following metric

$$g = g_1 + f^2 g_2, \tag{2}$$

where we omit the pulling back via the canonical projections of M onto its factors. The Riemannian manifold (M, g) is called the *warped product* of (M_1, g_1) (which is called the base) and (M_2, g_2) (which is called the fiber) with *warping* function f, (see [**8**, **13**] for details on the subject). The above warped product is simply denoted by $M_1 \times_f M_2$.

From now on (M, g) will denote a warped product with fiber (M_2, g_2) being a homogeneous space. Let G be the group of isometries of (M_2, g_2). The natural action of G on (M_2, g_2) is obviously transitive and it can be extended to an action on (M, g) through isometries as follows

$$M \times G \longrightarrow M; \qquad p.g = (p_1, p_2).g = (p_1, p_2.g), \tag{3}$$

for any $p = (p_1, p_2) \in M = M_1 \times M_2$ and $g \in G$. The orbit of this action through $p = (p_1, p_2) \in M$ is given by

$$[p] = [(p_1, p_2)] = \{p_1\} \times M_2.$$

The $(n_2 + 1)$-dimensional submanifolds in (M, g) which are G-invariant are characterized in the following proposition whose proof is evident.

Proposition. *Let N be an $(n_2 + 1)$-dimensional submanifold in (M, g). Then N is G-invariant if and only if there exists a curve γ in (M_1, g_1) such that $N = \gamma \times_f M_2$.*

To understand the following theorem better, we will say that a *r-generalized elastica* in a Riemannian manifold (P, h) is a curve, which is a critical point of the functional

$$\mathcal{F}^r(\gamma) = \int_\gamma (\kappa^2)^{\frac{r+1}{2}} ds, \tag{4}$$

(κ denotes the curvature function of γ in (P,h)), defined on the manifold consisting only of regular closed curves or curves which satisfy a given first order boundary data. The Euler-Lagrange equations relative to this variational problem were computed in [**6**] when $r = 3$. However the same computation can be adapted for an arbitrary value of r. In particular, when $r = 1$, we have the usual concept of free elastica in the sense of [**11**].

Theorem. *Let $(M,g) = M_1 \times_f M_2$ and (M_2, g_2) a compact homogeneous space of dimension n_2. Let γ be a closed curve immersed in (M_1, g_1). The submanifold $N = \gamma \times_f M_2$ is a Willmore–Chen submanifold in (M,g) if and only if γ is a n_2-generalized elastica in $\left(M_1, \frac{1}{f^2} g_1\right)$.*

Proof. Since the Willmore–Chen variational problem is invariant under conformal changes in the ambient space metric, we make the following change of the metric $g = g_1 + f^2 g_2$ on M, just define

$$\tilde{g} = \frac{1}{f^2} g = \frac{1}{f^2} g_1 + g_2.$$

Now the Willmore–Chen submanifolds in (M,g) are those in $(M,\tilde{g})$. Moreover we can take advantage from the Riemannian product structure of $(M,\tilde{g})$.

Let $\mathcal{N}$ be the smooth manifold consisting of the (n_2+1)-dimensional compact submanifolds in $(M,\tilde{g})$. The Willmore–Chen functional $\mathcal{W} : \mathcal{N} \longrightarrow R$ is given by (1) and it certainly is invariant under the G-action (3) on (M,g), that is $\mathcal{W}(N) = \mathcal{W}(N,g)$ for any $g \in G$. We put $\mathcal{N}_G$ to denote the submanifold of $\mathcal{N}$ made up of those submanifolds which are G-invariant. According to the Proposition, we have

$$\mathcal{N}_G = \{\gamma \times M_2 / \gamma \ \text{ is a closed curve in } \ M_1\}.$$

We also write $\mathcal{C}$ and $\mathcal{C}_G$ to denote the sets of critical points of $\mathcal{W}$ and $\mathcal{W}|_{\mathcal{N}_G}$ on $\mathcal{N}$ and $\mathcal{N}_G$ respectively. The first one is nothing but the set of Willmore–Chen submanifolds.

The principle of symmetric criticality , [**14**], can be applied here, consequently we have

$$\mathcal{C} \cap \mathcal{N}_G = \mathcal{C}_G.$$

Next we obtain G-symmetric Willmore–Chen submanifolds by first computing $\mathcal{W}$ on $\mathcal{N}_G$,

$$\mathcal{W}(\gamma \times M_2) = \int_{\gamma \times M_2} (\alpha^2 + \tau_e)^{\frac{n_2+1}{2}} ds\ dv_2, \tag{5}$$

where ds denotes the arc-length of γ in $\left(M_1, \frac{1}{f^2} g_1\right)$ and dv_2 is the volume element of (M_2, g_2). Having in mind that $(M, \tilde{g})$ is a Riemannian product, it is not difficult to see that τ_e vanishes identically and $\alpha^2 = \frac{1}{(1+n_2)^2}\kappa^2$, where κ is the curvature function of γ in $\left(M_1, \frac{1}{f^2} g_1\right)$. Therefore (5) can be written as

$$\mathcal{W}(\gamma \times M_2) = \frac{\mathrm{vol}(M_2, g_2)}{(1+n_2)^{1+n_2}} \int_\gamma (\kappa^2)^{\frac{n_2+1}{2}} ds,$$

which proves the statement.

From now on, we are going to get some applications of the above established result. To get examples of Willmore tori in the standard 3-sphere we proceed as follows. Let Ω be the open hemisphere, in the unit 2-sphere, defined in R^4 by $x_1 > 0$ and $x_2 = 0$. We denote its standard metric of constant curvature 1, by g_o. Let $f : \Omega \longrightarrow R$ be the positive function defined as the x_1 projection. Then $\Sigma = \Omega \times S^1$ is the unit 3-sphere where one geodesic was removed. The standard metric $\bar{g}_o$ on Σ is $g_o + f^2 dt^2$, with the obvious meaning. In others words, $(\Sigma, \bar{g}_o)$ is the *warped product* of (Ω, g_o) and (S^1, dt^2) with warping function f.

To better understand the next result, notice that Ω endowed with the metric $g = (1/f^2)g_o$ is nothing but the standard hyperbolic 2-plane with constant Gaussian curvature -1. Consequently we can apply the above stated theorem to have:

Corollary 1. *Let γ be an immersed closed curve in Ω. The torus $T_\gamma = \gamma \times S^1$ is Willmore in Σ if and only if γ is a free elastica in the hyperbolic plane (Ω, g).*

The complete classification of free elastica in the standard hyperbolic plane was achieved in [**11**]. Besides the m-fold cover η_o^m of the so called *hyperbolic equator* η_o (that is the geodesic circle of radius $\sinh^{-1}(1)$ in (Ω, g)), there exists an integer two parameter family of free elasticae in (Ω, g), $\{\eta^{m,n} / m > 1 \ \ \textit{and} \ \ \frac{1}{2} < \frac{m}{n} < \frac{\sqrt{2}}{2}\}$, (see [**11**] for a geometrical description of this family). Therefore we obtain:

Corollary 2. *There exist infinitely many Willmore tori in the standard* 3-*sphere Σ. This family includes the following two subfamilies:*

1) $\{T_{\eta_o^m} = \eta_o^m \times S^1 / m$ *is a non zero integer,*$\}$ *and*
2) $\{T_{\eta^{m,n}} = \eta^{m,n} \times S^1 / m > 1$ *and* $\frac{1}{2} < \frac{m}{n} < \frac{\sqrt{2}}{2}\}$.

Remark. (1) The reader should compare the Willmore tori just obtained in the standard 3-sphere with those obtained in [**15**]. Those came from nonfree elasticae in the standard 2-sphere. To be precise, they are Hopf tori shaped on elastic curves in the 2-sphere but the elasticity is constrainted with a Lagrange multiplier $\lambda = 4$ (when one regards the Hopf map as a Riemannian submersion). Also, these tori are conformal to those obtained in [**12**].

(2) We also notice, (see once more [**11**]) that the total squared curvature of any immersed closed curve in (Ω, g) satisfies

$$\int_\gamma \kappa^2(s)ds \geq 4\pi, \tag{6}$$

and the equality in (6) holds if and only if γ is a circular free elastica in (Ω, g), that is the hyperbolic equator η_o which is a geodesic in (Ω, g_o). Consequently we obtain the following connection with the popular and well known Willmore conjecture

$$\mathcal{W}(T_\gamma) \geq 2\pi^2, \tag{7}$$

and the equality in (7) holds if and only if T_γ is the Clifford torus in $(\Sigma, \bar{g}_o)$.

The following construction combined with the above studied argument allow us to give examples of Chen-Willmore submanifolds associated with certain metrics defined on $M = S^1 \times S^1 \times P$, where P denotes any compact homogeneous space (with a given metric ds^2) of dimension, say r. We start with any metric g on a torus $T = S^1 \times S^1$, as it is known that g is conformal to some flat metric, say g_o, on T. In others words, there exists a positive function f on T such that $g = f^2 g_o$. We consider $M = T \times P$ endowed with the metric $\tilde{g} = g + f^2 ds^2$ and so $\left(M, \frac{1}{f^2}\tilde{g}\right)$ is the Riemannian product of the flat torus (T, g_o) and (P, ds^2). Now for any closed free elastica γ in the flat torus (T, g_o), the tube $\Upsilon_\gamma = \gamma \times P$ is a Willmore–Chen submanifold in $(M, \tilde{g})$.

Let R^3 be the Euclidean 3-space endowed with its canonical metric $<,>$. We define the following positive function on R^3, $f(p) = 1 + \frac{1}{2}|p|$, where the point $p \in R^3$ is identified with its position vector (relative to some origin in R^3) and $|p| = < p, p >^{\frac{1}{2}}$. We choose any compact, r-dimensional, homogeneous space, say (P, ds^2) as above. Then we have:

Corollary 3. *There exists a rational one-parameter family of $(r+1)$-dimensional Willmore–Chen submanifolds in $M = R^3 \times_f P$ for any r-dimensional compact homogeneous space (P, ds^2).*

Proof. In M with the metric $<,> + f^2 ds^2$ we make a conformal change to get $\frac{1}{f^2} <,> + ds^2$. Then $(R^3, \frac{1}{f^2} <,>)$ is nothing but the once punctured

3-sphere, Σ. Now we can use an argument similar to that used in [**6**], to show the following statement: For any non-zero rational number q, there exists a closed helix β_q in Σ which is a r-generalized elastica in Σ. Then we apply the main theorem to conclude that $\beta_q \times_f P$ is a Willmore–Chen submanifold in M.

The last application is dedicated to obtain a 3-dimensional Riemannian manifold which admit a foliation with leaves being non-trivial Willmore tori, we will call it a *Willmore foliation* and we will say that the manifold is *Willmore foliated.*

We start from a plane immersed curve $\gamma : I \subset R \longrightarrow R^2$ which we assumed to be parametrized by arc-length s. Let $M = \gamma \times S^1 \times S^1$ endowed with the metric $g = ds^2 + f_1^2 dt^2 + f_2^2 du^2$, where f_1 and f_2 are two positive functions on γ. Certainly g is conformal to the Riemannian product metric $\tilde{g}$ on $M = N \times S^1$ given by $\tilde{g} = g_o + du^2$, where $g_o = dv^2 + \frac{f_1^2}{f_2^2} dt^2$ and $N = \gamma \times S^1$. Notice that we made a change of parameters in γ, to be precise we considered $\frac{ds}{dv} = f_1(s)$ on γ. Now a suitable choice of both f_1 and f_2 on γ, allows us to regard (N, g_o) as a surface of revolution in R^3.

On the other hand the elasticity of the parallels in a surface of revolution was discussed in [**5**]. The following statement was proved there: Besides the right cylinders (all whose parallels are geodesics and so trivially free elastic curves), the only surfaces of revolution with all the parallels being free elasticae are what were called the *trumpet* surfaces, such surfaces are free of geodesic parallels, therefore we have, (see [**5**] for details):

Corollary 4. *Let b and c be a couple of real numbers with $c > 0$ and $I = \left(-\frac{2}{c}, 0\right) \bigcup \left(0, \frac{2}{c}\right)$. We define the plane curve (a trumpet) $\gamma : I \subset R \longrightarrow R^2$ by*

$$\gamma(s) = \left(\frac{c}{4} s^2, \frac{s}{2} \sqrt{1 - \frac{c^2}{4} s^2} - \frac{1}{c} \arccos \frac{cs}{2} + b \right).$$

Let f_1 and f_2 be two positive functions on I with $\frac{f_1(s)}{f_2(s)} = \frac{c}{4} s^2$. Then $M = \gamma \times S^1 \times S^1$ endowed with the metric $g = ds^2 + f_1^2 dt^2 + f_2^2 du^2$ admits a nontrivial Willmore foliation with leaves being nontrivial Willmore tori.

Acknowledgement. The authors would like to express their thanks to the referee for his valuable suggestions.

References

[1] M. Barros, *Free elasticae and Willmore tori in warped product spaces*, Glasgow Math. J., **40** (1988), 263-270.

[2] ———, *Willmore tori in non standard 3-spheres*, Math. Proc. Cambridge Phil. Soc., **121** (1997), 321-324.

[3] M. Barros and B.Y. Chen, *Stationary 2-type surfaces in a hypersphere*, J. Math. Soc. Japan, **39** (1987), 627-648.

[4] M. Barros, A. Ferrández, P. Lucas and M.A. Meroño, *Helicoidal filaments in the 3-sphere*, Preprint.

[5] M. Barros and O.J. Garay, *Free elastic parallels in a surface of revolution*, Amer. Math. Monthly, **103** (1996), 149-156.

[6] ———, *Hopf submanifolds in S^7 which are Willmore–Chen submanifolds*, Math. Z., **228** (1998), 121-129.

[7] M. Barros, O.J. Garay and D.A. Singer, *Elasticae with constant slant in the complex projective plane and new examples of Wilmore tori in five spheres*, to appear in Tôhoku Math. J.

[8] A.L. Besse, *Einstein manifolds*, Springer-Verlag, 1987.

[9] B.Y. Chen, *Some conformal invariants of submanifolds and their applications*, Boll. Un. Mat. Ital., **10** (1974), 380-385.

[10] N. Ejiri, *A counter-example for Weiner open question*, Indiana Univ. Math. J., **31** (1982), 209-211.

[11] J. Langer and D. Singer, *The total squared curvature of closed curves*, J. Diff. Geom., **20** (1984), 1-22.

[12] ———, *Curves in the hyperbolic plane and mean curvature of tori in 3-space*, Bull. London Math. Soc., **16** (1984), 531-534.

[13] B. O'Neill, *Semi-Riemannian Geometry*, Academic Press, 1983.

[14] R.S. Palais, *Critical points theory and minimax principle*, Proc. Sympos. Pure Math., **15** (1970), 185-212.

[15] U. Pinkall, *Hopf tori in S^3*, Invent. Math., **81** (1985), 379-386.

[16] J.L. Weiner, *On a problem of Chen, Willmore et. al.*, Indiana Univ. Math. J., **27** (1978), 19-35.

Received April 30, 1997 and revised April 27, 1998. This work was partially supported by a DGICYT Grant No. PB97-0784.

Univ. de Granada
18071 Granada
Spain

Univ. de Granada
18071 Granada
Spain
E-mail address: mbarros@goliat.ugr.es

Universidad del Pais Vasco
Bilbao
Spain
E-mail address: mtpgabeo@lg.ehu.es

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

FREDHOLM PROPERTIES OF TOEPLITZ OPERATORS ON DIRICHLET SPACES

Guangfu Cao

In this paper, the Fredholm properties of some Toeplitz operators on Dirichlet spaces be discussed, and the essential spectra of Toeplitz operators with symbols in C^1 or $H_1^\infty + C^1$ be computed.

1. Introduction.

Let $\mathbb{D}$ be the unit disk in the complex plane $\mathbb{C}$, $dA = \frac{1}{\pi}dxdy$ be the normalized area measure on $\mathbb{D}$. $L^{2,1}$ is the space of functions $u : \mathbb{D} \to \mathbb{C}$, for which the norm

$$\|u\|_{\frac{1}{2}} = \left[\int_{\mathbb{D}} \left(\left|\frac{\partial u}{\partial z}(z)\right|^2 + \left|\frac{\partial u}{\partial \bar{z}}(z)\right|^2\right) dA\right]^{\frac{1}{2}} < \infty.$$

$L^{2,1}$ is a Hilbert space with the inner product

$$\langle u, v\rangle_{\frac{1}{2}} = \left\langle \frac{\partial u}{\partial z}, \frac{\partial v}{\partial z}\right\rangle_{L^2(dA)} + \left\langle \frac{\partial u}{\partial \bar{z}}, \frac{\partial v}{\partial \bar{z}}\right\rangle_{L^2(dA)}.$$

The Dirichlet space, $\mathcal{D}$, is the subspace of all analytic functions g in $L^{2,1}$ with $g(0) = 0$. Let P be the orthogonal projection from $L^{2,1}$ into $\mathcal{D}$. P is an integral operator represented by

$$P(u)(w) = \int_{\mathbb{D}} \frac{\partial u}{\partial z}\overline{\frac{\partial K(z,w)}{\partial z}}dA,$$

where $K(z,w) = \sum_{k=1}^{\infty} \frac{z^k \bar{w}^k}{k}$ is the reproducing kernel of $\mathcal{D}$ (see R. Rochberg and Z.J. Wu [**5**] and Wu [**7**]). Let G be a domain in $\mathbb{C}$, define

$$C^1(G) = \left\{u | u, \frac{\partial u}{\partial z}, \frac{\partial u}{\partial \bar{z}} \in C(G)\right\},$$
$$C^1(\bar{G}) = \left\{u | u, \frac{\partial u}{\partial z}, \frac{\partial u}{\partial \bar{z}} \in C(\bar{G})\right\},$$

and

$$H_1^\infty(G) = \{f \in H(G) | f' \in H^\infty(G)\},$$

where $C(G)$(or $C(\bar{G})$) and $H(G)$ are respectively the set of continuous functions on G(or $\bar{G}$) and the one of analytic functions on G.

In [**5**] and [**7**], R. Rochberg and Z.J. Wu define the Toeplitz operator with symbol μ as

$$T_\mu f = \int_{\mathbb{D}} f'(z)\overline{\frac{\partial K(z,w)}{\partial z}}d\mu \qquad (\forall f \in \mathcal{D}),$$

where μ is a measure on $\mathbb{D}$. If $d\mu = \varphi dA, \varphi \in L^\infty(\mathbb{D}, dA)$, one can define

$$T_\varphi f = \int_{\mathbb{D}} f'(z)\overline{\frac{\partial K(z,w)}{\partial z}}\varphi dA.$$

We see easily that $T_\varphi f \neq P(\varphi f)$ in general, in fact, $P(\varphi f)$ may be undefined for $\varphi \in L^\infty(\mathbb{D}, dA)$. However, it seems to be more natural that the Toeplitz operators be defined as the form $P(\varphi f)$. In this paper, we try to define the Toeplitz operators with some special symbols as following

Definition 1. Suppose $\varphi \in C^1(\mathbb{D})$, the operator

$$T_\varphi f(w) = P(\varphi f)(w) = \langle \varphi f, K\rangle_{\frac{1}{2}} = \int_{\mathbb{D}} \frac{\partial(\varphi f)}{\partial z}\overline{\frac{\partial K(z,w)}{\partial z}}dA(z)$$

is said to be the Toeplitz operator with symbol φ.

We will compute the spectra and essential spectra of these operators. For convenience, we use frequently the notation $\widetilde{T_\varphi}$ to denote the Toeplitz operator with symbol φ on Bergman space.

By the way, our results are also true for Toeplitz operators on weighted Dirichlet spaces $\mathcal{D}_\alpha$ ($\alpha < 1$). If $\alpha = \frac{1}{2}$, then $\mathcal{D}_{\frac{1}{2}} = \mathcal{D}$, the usual Dirichlet space (c.f. [**5**], [**7**]).

2. Toeplitz operators with symbols in $C^1(\bar{\mathbb{D}})$.

Throughout this paper, we use the symbol "$\langle .,. \rangle$" to represent the inner product in $L^2(\mathbb{D}, dA)$, and "$\langle .,. \rangle_{\frac{1}{2}}$" to that in $\mathcal{D}$.

Define the norm in $C^1(\bar{\mathbb{D}})$ as

$$\|\varphi\|_* = \max_{z\in\bar{\mathbb{D}}} \max\left\{|\varphi|, \left|\frac{\partial\varphi}{\partial z}\right|, \left|\frac{\partial\varphi}{\partial\bar{z}}\right|\right\}, \qquad (\forall\varphi \in C^1(\bar{\mathbb{D}})).$$

It is well-known that $C^1(\bar{\mathbb{D}})$ is a norm-closed algebra relative to $\|.\|_*$.

Lemma 1. *For any $\varphi \in C^1(\bar{\mathbb{D}})$, $H_\varphi f =: (I-P)(\varphi f)$ $(\forall f \in \mathcal{D})$ is a compact operator from $\mathcal{D}$ to $\mathcal{D}^\perp$.*

Proof. For any $f \in \mathcal{D}, g \in \mathcal{D}^{\perp}$

$$\begin{aligned}\langle H_\varphi f, g\rangle_{\frac{1}{2}} = \langle \varphi f, g\rangle_{\frac{1}{2}} &= \left\langle \frac{\partial \varphi}{\partial z} f + \varphi \frac{\partial f}{\partial z}, \frac{\partial g}{\partial z}\right\rangle + \left\langle \frac{\partial \varphi}{\partial \bar{z}} f, \frac{\partial g}{\partial \bar{z}}\right\rangle \\ &= \left\langle \frac{\partial \varphi}{\partial z} f, \frac{\partial g}{\partial z}\right\rangle + \left\langle \varphi \frac{\partial f}{\partial z}, \frac{\partial g}{\partial z}\right\rangle + \left\langle \frac{\partial \varphi}{\partial \bar{z}} f, \frac{\partial g}{\partial \bar{z}}\right\rangle.\end{aligned}$$

Note $g \perp \mathcal{D}$, so $\frac{\partial g}{\partial z} \perp L_a^2$ (L_a^2 is the classical Bergman space on $\mathbb{D}$). Let P_a is the orthogonal projection from $L^2(\mathbb{D}, dA)$ to L_a^2, then for any $f \in \mathcal{D}$ and $g \in \mathcal{D}^{\perp}$, we have

$$\left\langle \varphi f', \frac{\partial g}{\partial z}\right\rangle = \left\langle (I - P_a)(\varphi f'), \frac{\partial g}{\partial z}\right\rangle = \left\langle \widetilde{H}_\varphi f', \frac{\partial g}{\partial z}\right\rangle,$$

where $\widetilde{H}_\varphi$ is the Hankel operator from L_a^2 to $L_a^{2\perp}$ with symbol φ. Since $\varphi \in C^1(\bar{\mathbb{D}})$, $\widetilde{H}_\varphi$ is a compact operator. Now assume that $\{f_k\} \subset \mathcal{D}$ is a sequence which converges weakly 0 and satisfies $\|f_k\|_{\mathcal{D}} = 1$, we prove first that $\|f_k\|_{L^2} \to 0$ ($k \to \infty$). In fact, without loss of generality, assume $f_k(z) = \sum_{n=1}^{\infty} a_n^{(k)} z^k$, then $a_n^{(k)} \to 0$ ($k \to \infty, \forall n$) by $f_k \xrightarrow{w} 0$. Note $\|f_k\|_{\mathcal{D}}^2 = \|f_k'\|_{L^2}^2 = 1$, so $\sum_{n=1}^{\infty} |a_n^{(k)}|^2 n = 1$ ($\forall k$), thus $\sum_{n=1}^{\infty} |a_n^{(k)}|^2 \le 1$, furthermore, for any $\epsilon > 0$, there is a N_0 which is independent on k such that $\sum_{n \ge N_0+1}^{\infty} |a_n^{(k)}|^2 \frac{1}{n+1} < \frac{\epsilon}{2}$. Fix such a N_0, then

$$\|f_k\|_{L^2}^2 = \sum_{n=1}^{\infty} \left|a_n^{(k)}\right|^2 \frac{1}{n+1} \le \sum_{n=1}^{N_0} \left|a_n^{(k)}\right|^2 \frac{1}{n+1} + \frac{\epsilon}{2}.$$

Since $a_n^{(k)} \to 0$ ($\forall n, k \to \infty$), there is a K_0 such that for any $k > K_0$, $\sum_{n=1}^{N_0} |a_n^{(k)}|^2 \frac{1}{n+1} < \frac{\epsilon}{2}$, thus $\|f_k\|_{L^2}^2 < \frac{\epsilon}{2}$ ($k > K_0$), hence $\|f_k\|_{L^2} \to 0$. Note

$$\begin{aligned}\|H_\varphi f_k\|_{L^{2,1}}^2 &= \langle H_\varphi f_k, H_\varphi f_k\rangle_{\frac{1}{2}} \\ &\le \left\|\frac{\partial \varphi}{\partial z} f_k\right\|_{L^2} \|H_\varphi f_k\|_{L^{2,1}} + \left\|\widetilde{H}_\varphi f_k'\right\|_{L^2} \|H_\varphi f_k\|_{L^{2,1}} \\ &\quad + \left\|\frac{\partial \varphi}{\partial \bar{z}} f_k\right\|_{L^2} \|H_\varphi f_k\|_{L^{2,1}} \\ &\le \|\varphi\|_* \|f_k\|_{L^2} \|H_\varphi f_k\|_{L^{2,1}} + \left\|\widetilde{H}_\varphi f_k'\right\|_{L^2} \|H_\varphi f_k\|_{L^{2,1}} \\ &\quad + \|\varphi\|_* \|f_k\|_{L^2} \|H_\varphi f_k\|_{L^{2,1}},\end{aligned}$$

we know easily that $\|H_\varphi f_k\|_{L^{2,1}} \to 0$. This shows that H_φ is compact. □

Proposition 2. *For any $\varphi \in C^1(\bar{\mathbb{D}})$, T_φ is a bounded operator on $\mathcal{D}$.*

Proof. For any $f, g \in \mathcal{D}$,

$$\langle T_\varphi f, g\rangle_{\frac{1}{2}} = \langle \varphi f, g\rangle_{\frac{1}{2}} = \left\langle \frac{\partial(\varphi f)}{\partial z}, g'\right\rangle = \left\langle \frac{\partial \varphi}{\partial z} f, g'\right\rangle + \langle \varphi f', g'\rangle,$$

hence

$$\begin{aligned}\left|\langle T_\varphi f, g\rangle_{\frac{1}{2}}\right| &\le \left\|\frac{\partial\varphi}{\partial z} f\right\|_{L^2} \|g'\|_{L^2} + \|\varphi f'\|_{L^2}\|g'\|_{L^2}\\ &\le \|\varphi\|_*(\|f\|_{L^2}\|g\|_{\mathcal{D}} + \|f\|_{\mathcal{D}}\|g\|_{\mathcal{D}}).\end{aligned}$$

It is not difficult to check $\|f\|_{L^2} \le \|f\|_{\mathcal{D}}$ for any $f \in \mathcal{D}$ (since $f(0) = 0$ for any $f \in \mathcal{D}$), hence $|\langle T_\varphi f, g\rangle_{\frac{1}{2}}| \le 2\|\varphi\|_*\|f\|_{\mathcal{D}}\|g\|_{\mathcal{D}}$. This shows that $\|T_\varphi\| \le 2\|\varphi\|_*$. $\square$

Remark. In general, $\|T_\varphi\| \ne \|\varphi\|_*$ even $\varphi \in C^1(\mathbb{D}) \cap A(\mathbb{D})$ ($A(\mathbb{D})$ be the disk algebra), it may be that $\|\varphi\|_{C^1} < \|T_\varphi\| < 2\|\varphi\|_{C^1}$. For instance, if $\varphi(z) = z$, then $\|\varphi\|_* = \|\varphi\|_\infty = 1$, but $\|T_\varphi\| \ge \sqrt{2}$.

Proposition 3. *Let* $L^{\infty,1} = \{f \in L^{2,1} | f, \frac{\partial f}{\partial z}, \frac{\partial f}{\partial \bar z} \in L^\infty(\mathbb{D}, dA)\}$, *then for any* $\varphi \in L^{\infty,1}$ *and* $\psi \in C^1(\bar{\mathbb{D}})$ $T_\varphi T_\psi - T_{\varphi\psi} \in \mathcal{K}(\mathcal{D})$ (*the compact operator algebra on* $\mathcal{D}$.)

Proof. For any $f, g \in \mathcal{D}$, as a direct computing, we see that

$$\begin{aligned}\langle (T_\varphi T_\psi - T_{\varphi\psi})f, g\rangle_{\frac{1}{2}} &= \left\langle \frac{\partial\varphi}{\partial z} T_\psi f, \frac{\partial g}{\partial z}\right\rangle - \left\langle \frac{\partial\varphi}{\partial z}(\psi f), \frac{\partial g}{\partial z}\right\rangle\\ &\quad + \left\langle \varphi\frac{\partial(-H_\psi f)}{\partial z}, \frac{\partial g}{\partial z}\right\rangle,\end{aligned}$$

set $g = (T_\varphi T_\psi - T_{\varphi\psi})f$, we have

$$\begin{aligned}\|(T_\varphi T_\psi - T_{\varphi\psi})f\|_{\mathcal{D}}^2 \le \|\varphi\|_*[&\|T_\psi f\|_{L^2} + \|\psi\|_*\|f\|_{L^2}\\ &+ \|H_\psi f\|_{L^{2,1}}]\|(T_\varphi T_\psi - T_{\varphi\psi})f\|_{\mathcal{D}}.\end{aligned}$$

Note for any sequence $\{f_k\} \subset \mathcal{D}$ which converges weakly 0 and satisfies $\|f_k\|_{\mathcal{D}} = 1$ (certainly, $T_\psi f_k \xrightarrow{w} 0$ in $\mathcal{D}$), $\|T_\psi f_k\|_{L^2} \to 0, \|f_k\|_{L^2} \to 0$, hence

$$\|(T_\varphi T_\psi - T_{\varphi\psi})f_k\|_{\mathcal{D}} \to 0$$

by H_ψ is compact. Futhermore, $T_\varphi T_\psi - T_{\varphi\psi}$ is compact. $\square$

Lemma 4. *For any* $\varphi \in L^{\infty,1}$, $T_\varphi^* - T_{\bar\varphi} \in \mathcal{K}(\mathcal{D})$.

Proof. Similar to the proof of Proposition 2, it is easy to see that T_φ is bounded for any $\varphi \in L^{\infty,1}$. Now assume $\varphi \in L^{\infty,1}$, then for any $f \cdot g \in \mathcal{D}$,

$\langle (T_\varphi^* - T_{\bar\varphi})f.g\rangle_{\frac{1}{2}} = \langle \frac{\partial f}{\partial z}, \frac{\partial \varphi}{\partial z} g\rangle - \langle f, \frac{\partial \varphi}{\partial \bar z}\frac{\partial g}{\partial z}\rangle$. Thus

$$\begin{aligned}|\langle (T_\varphi^* - T_{\bar\varphi})f, g\rangle_{\frac{1}{2}}| &\leq \left\|\frac{\partial f}{\partial z}\right\|_{L^2}\left\|\frac{\partial \varphi}{\partial z} g\right\|_{L^2} + \|f\|_{L^2}\left\|\frac{\partial \varphi}{\partial \bar z}\frac{\partial g}{\partial z}\right\|_{L^2}\\ &\leq \|\varphi\|_*\|g\|_{L^2}\|f\|_{\mathcal{D}} + \|\varphi\|_*\|f\|_{L^2}\|g\|_{\mathcal{D}},\end{aligned}$$

where $\|\varphi\|_* = \operatorname{esssup}_{z\in\mathbb{D}} \max\{|\varphi(z)|, |\frac{\partial\varphi}{\partial z}(z)|, |\frac{\partial\varphi}{\partial \bar z}(z)|\}$. Hence for any $\{f_k\} \subset \mathcal{D}, \|f_k\| = 1, f_k \xrightarrow{w} 0$, we have

$$\|(T_\varphi^* - T_{\bar\varphi})f_k\|_{\mathcal{D}}^2 \leq \|\varphi\|_*\|(T_\varphi^* - T_{\bar\varphi})f_k\|_{L^2}\|f_k\|_{\mathcal{D}} + \|\varphi\|_*\|f_k\|_{L^2}\|(T_\varphi^* - T_{\bar\varphi})f_k\|_{\mathcal{D}}.$$

Note $(T_\varphi^* - T_{\bar\varphi})f_k \xrightarrow{w} 0$, so $\|(T_\varphi^* - T_{\bar\varphi})f_k\|_{L^2} \to 0$. Consequently, $\|(T_\varphi^* - T_{\bar\varphi})f_k\|_{\mathcal{D}} \to 0$. This shows that $T_\varphi^* - T_{\bar\varphi}$ is compact. □

Remark. It is well-known that $T_\varphi^* = T_{\bar\varphi}$ for any $\varphi \in L^\infty(\mathbb{D}, dA)$ on L_a^2, but it is not difficult to check $T_\varphi^* \neq T_{\bar\varphi}$ in general. For example, if $\varphi(z) = \sum_{n=1}^\infty \frac{z^n}{n^2}$, then $\varphi \in H^\infty$, but $\|\varphi'\|_\infty = \infty$, one can check easily that T_φ is unbounded, however, $T_{\bar\varphi}$ is always bounded for every $\varphi \in H^\infty$, so is $T_{\bar\varphi}^*$, hence $T_{\bar\varphi}^* - T_\varphi$ is unbounded.

Proposition 5. *If $\varphi \in C^1(\bar{\mathbb{D}})$, then T_φ is a compact operator on $\mathcal{D}$ if and only if $\varphi|_{\partial\mathbb{D}} \equiv 0$.*

Proof. Assume $\varphi|_{\partial\mathbb{D}} \equiv 0$, then $\widetilde{T}_{|\varphi|^2}$ is a compact operator on $L_a^2(\mathbb{D}, dA)$, so for any $\{f_k\} \subset L_a^2$, $f_k \xrightarrow{w} 0$, $\|f_k\|_{L^2} = 1$, we have $\|\widetilde{T}_{|\varphi|} f_k\|_{L^2} \to 0$ $(k \to \infty)$. If T_φ is not compact on $\mathcal{D}$, then there is a sequence $\{F_k\} \subset \mathcal{D}$,$\|F_k\|_{\mathcal{D}} = 1$, $F_k \xrightarrow{w} 0$ such that $\|T_\varphi F_k\|_{\mathcal{D}} \nrightarrow 0$, thus $\|\varphi F_k\|_{L^{2,1}} \nrightarrow 0$, furthermore

$$\left\langle \frac{\partial(\varphi F_k)}{\partial z}, \frac{\partial(\varphi F_k)}{\partial z}\right\rangle + \left\langle \frac{\partial(\varphi F_k)}{\partial \bar z}, \frac{\partial(\varphi F_k)}{\partial \bar z}\right\rangle \nrightarrow 0.$$

Note

$$\frac{\partial(\varphi F_k)}{\partial z} = \frac{\partial\varphi}{\partial z}F_k + \varphi\frac{\partial F_k}{\partial z},\ \frac{\partial(\varphi F_k)}{\partial \bar z} = \frac{\partial\varphi}{\partial \bar z}F_k,$$

thus

$$\left\langle \frac{\partial(\varphi F_k)}{\partial \bar z}, \frac{\partial(\varphi F_k)}{\partial \bar z}\right\rangle \leq \|\varphi\|_*^2\|F_k\|_{L^2}^2 \to 0,$$

and

$$\begin{aligned}\left\|\frac{\partial(\varphi F_k)}{\partial z}\right\|_{L^2}^2 &= \left\langle \frac{\partial\varphi}{\partial z}F_k, \frac{\partial\varphi}{\partial z}F_k\right\rangle + \left\langle \frac{\partial\varphi}{\partial z}F_k, \varphi F_k'\right\rangle\\ &\quad + \left\langle \varphi F_k', \frac{\partial\varphi}{\partial z}F_k\right\rangle + \langle \varphi F_k', \varphi F_k'\rangle.\end{aligned}$$

Since $\|F_k'\|_{L^2} = \|F_k\|_{\mathcal{D}} = 1$, and $F_k' \xrightarrow{w} 0$ in L_a^2,

$$\langle \varphi F_k', \varphi F_k'\rangle = \langle |\varphi|^2 F_k', F_k'\rangle = \langle T_{|\varphi|^2} F_k', F_k'\rangle \to 0,$$

consequently, $\|\frac{\partial(\varphi F_k)}{\partial z}\|^2_{L^2} \to 0$ by $|\langle\frac{\partial\varphi}{\partial z}F_k, \frac{\partial\varphi}{\partial z}F_k\rangle| \le \|\varphi\|^2_*\|F_k\|^2_{L^2}$ and $|\langle\frac{\partial\varphi}{\partial z}F_k, \varphi F'_k\rangle| \le \|\varphi\|^2_*\|F_k\|_{L^2}\|F'_k\|_{L^2}$, and $\|F_k\|_{L^2} \to 0$. This contradicts that $\|\varphi F_k\|_{L^{2,1}} \nrightarrow 0$. Hence T_φ must be a compact operator on $\mathcal{D}$.

Conversely, assume that T_φ is compact on $\mathcal{D}$. We need to prove that $\varphi|_{\partial\mathbb{D}} = 0$. Otherwise, $\widetilde{T}_{|\varphi|}$ is not compact on $L^2_a(\mathbb{D})$, thus there is a sequence $\{f_k\} \subset L^2_a$, $\|f_k\|_{L^2} = 1$, $f_k \xrightarrow{w} 0$ such that $\|\widetilde{T}_{|\varphi|}f_k\|_{L^2} \nrightarrow 0$, thus $\||\varphi|f_k\|_{L^2} \nrightarrow 0$, that is $\int_{\mathbb{D}} |\varphi|^2|f_k|^2 dA \nrightarrow 0$. But T_φ is compact on $\mathcal{D}$, so for $F_k = \int_0^z f_k dw$, $\|T_\varphi F_k\|_{\mathcal{D}} \to 0$, further $\|T^*_\varphi T_\varphi F_k\|_{\mathcal{D}} \to 0$. Note $T^*_\varphi - T_{\overline{\varphi}} \in \mathcal{K}$, so $\|T_{|\varphi|^2}F_k\|_{\mathcal{D}} \to 0$ (since $F_k \xrightarrow{w} 0$ in $\mathcal{D}$ and $\|F_k\|_{\mathcal{D}} = 1$), thus $\langle T_{|\varphi|^2}F_k, F_k\rangle_{\frac{1}{2}} \to 0$, hence

$$\left\langle \frac{\partial(|\varphi|^2 F_k)}{\partial z}, F'_k \right\rangle \to 0,$$

that is

$$\left\langle \frac{\partial(|\varphi|^2)}{\partial z} F_k, F'_k \right\rangle + \langle |\varphi|^2 F'_k, F'_k\rangle \to 0.$$

Note

$$\langle |\varphi|^2 F'_k, F'_k\rangle = \langle |\varphi|^2 f_k, f_k\rangle = \int_{\mathbb{D}} |\varphi|^2 |f_k|^2 dA \nrightarrow 0,$$

and

$$\left|\left\langle \frac{\partial(|\varphi|^2)}{\partial z} F_k, F'_k \right\rangle\right| \le \||\varphi|^2\|_* \|F_k\|_{L^2}\|f_k\|_{L^2} \to 0,$$

hence $\langle T_{|\varphi|^2}F_k, F_k\rangle_{\frac{1}{2}} \nrightarrow 0$, this contradiction shows that $\varphi|_{\partial\mathbb{D}} \equiv 0$. □

Remark. If $\varphi \notin C^1(\bar{\mathbb{D}})$, then T_φ may be non-compact even $\varphi \in C(\bar{\mathbb{D}})$ and $\varphi|_{\partial\mathbb{D}} = 0$. For instance, let $\varphi = \sqrt{1-|z|^2}$, then it is not difficult to check T_φ is not a compact operator. In fact, T_φ is also a unbounded operator on $\mathcal{D}$.

Theorem 6. *Suppose $\varphi \in C^1(\bar{\mathbb{D}})$, then $\sigma_e(T_\varphi) = \varphi(\partial\mathbb{D})$.*

Proof. Without loss of generality, assume $0 \in \varphi(\partial\mathbb{D})$, thus there is a $\zeta \in \partial\mathbb{D}$ such that $\varphi(\zeta) = 0$. Write $f_k(z) = (\frac{1+z\bar{\zeta}}{2})^k$, then $\frac{f_k}{\|f_k\|_{\mathcal{D}}} \xrightarrow{w} 0$, and $\frac{\|T_\varphi f_k\|_{\mathcal{D}}}{\|f_k\|_{\mathcal{D}}} \le \frac{\|\varphi f_k\|_{L^{2,1}}}{\|f_k\|_{\mathcal{D}}}$. Clearly

$$\begin{aligned}\|\varphi f_k\|^2_{L^{2,1}} &= \left\langle \frac{\partial\varphi}{\partial z}f_k + \varphi f'_k, \frac{\partial\varphi}{\partial z}f_k + \varphi f'_k \right\rangle + \left\langle \frac{\partial\varphi}{\partial\bar{z}}f_k, \frac{\partial\varphi}{\partial\bar{z}}f_k \right\rangle \\ &= \left\langle \frac{\partial\varphi}{\partial z}f_k, \frac{\partial\varphi}{\partial z}f_k \right\rangle + \left\langle \varphi f'_k, \frac{\partial\varphi}{\partial z}f_k \right\rangle + \left\langle \frac{\partial\varphi}{\partial z}f_k, \varphi f'_k \right\rangle \\ &\quad + \langle \varphi f'_k, \varphi f'_k\rangle + \left\langle \frac{\partial\varphi}{\partial\bar{z}}f_k, \frac{\partial\varphi}{\partial\bar{z}}f_k \right\rangle,\end{aligned}$$

note

$$\frac{|\langle \frac{\partial\varphi}{\partial z} f_k, \frac{\partial\varphi}{\partial z} f_k\rangle + \langle \varphi f_k', \frac{\partial\varphi}{\partial z} f_k\rangle + \langle \frac{\partial\varphi}{\partial z} f_k, \varphi f_k'\rangle + \langle \frac{\partial\varphi}{\partial \bar z} f_k, \frac{\partial\varphi}{\partial \bar z} f_k\rangle|}{\|f_k\|_{\mathcal{D}}^2} \longrightarrow 0,$$

and

$$\begin{aligned}\frac{\langle \varphi f_k', \varphi f_k'\rangle}{\|f_k\|_{\mathcal{D}}^2} &= \frac{\int_{\mathbb{D}} |\varphi|^2 |f_k'|^2 dA}{\int_{\mathbb{D}} |f_k'|^2 dA}\\ &= \frac{\int_{\mathbb{D}} |\varphi|^2 |(\frac{1+z\bar\zeta}{2})^{2(k-1)}| dA}{\int_{\mathbb{D}} |(\frac{1+z\bar\zeta}{2})^{2(k-1)}| dA}\\ &= \frac{\int_{\mathbb{D}} |\varphi|^2 |f_{k-1}|^2 dA}{\int_{\mathbb{D}} |f_{k-1}|^2 dA} \to |\varphi(\zeta)|^2 = 0\end{aligned}$$

(since f_k is a peak function at ζ). Thus $\frac{\|T_\varphi f_k\|_{\mathcal{D}}}{\|f_k\|_{\mathcal{D}}} \to 0$, this shows that T_φ can not be Fredholm, hence $0 \in \sigma_e(T_\varphi)$. That is $\varphi(\partial\mathbb{D}) \subset \sigma_e(T_\varphi)$.

Conversely, if $0 \notin \varphi(\partial\mathbb{D})$, then $|\varphi(\zeta)| > \epsilon_0 > 0$ for any $\zeta \in \partial\mathbb{D}$, thus $\widetilde{T}_\varphi$ is Fredholm on L_a^2, if T_φ is not Fredholm on $\mathcal{D}$, then there is a sequence $\{F_k\} \subset \mathcal{D}$ with $\|F_k\|_{\mathcal{D}} = 1, F_k \xrightarrow{w} 0$ such that $\|T_\varphi F_k\|_{\mathcal{D}} \to 0$, or $\|T_\varphi^* F_k\|_{\mathcal{D}} \to 0$. Since $\widetilde{T}_\varphi$ is Fredholm on L_a^2, there is an $S \in L(L_a^2)$ such that $[ST_\varphi] = [T_\varphi S] = [I]$ in the Calkin algebra $\frac{L(L_a^2)}{\mathcal{K}}$, thus $S^* T_\varphi^* = I + K,\ K \in \mathcal{K}(L_a^2)$, assume $\|T_\varphi^* F_k\|_{\mathcal{D}} \to 0$. (For the case of $\|T_\varphi F_k\|_{\mathcal{D}} \to 0$, we can complete the proof similarly.) Write $f_k = F_k'$, then $\|f_k\|_{L^2} = 1$, and $f_k \xrightarrow{w} 0$ in L_a^2, thus $\langle S^* T_\varphi^* f_k, f_k\rangle \to 1$, that is $\langle T_\varphi^* f_k, S f_k\rangle \to 1$.

On the other hand,

$$\begin{aligned}\left\langle T_\varphi^* F_k, \int_0^z (Sf_k)(w)dw \right\rangle_{\frac12} &= \left\langle F_k, \varphi \int_0^z (Sf_k)(w)dw \right\rangle_{\frac12}\\ &= \left\langle f_k, \frac{\partial\varphi}{\partial z} \int_0^z (Sf_k)(w)dw + \varphi S f_k \right\rangle\\ &= \left\langle f_k, \frac{\partial\varphi}{\partial z} \int_0^z S f_k dw \right\rangle + \langle f_k, \varphi S f_k\rangle.\end{aligned}$$

Since $Sf_k \xrightarrow{w} 0$ in L_a^2, $\|\int_0^z Sf_k dw\|_{L_a^2} \to 0$, thus $\langle f_k, \frac{\partial\varphi}{\partial z} \int_0^z Sf_k(w)dw\rangle \to 0$. Note $\langle f_k, \varphi S f_k\rangle = \langle T_\varphi^* f_k, S f_k\rangle \to 1 \neq 0$, this shows that $\langle T_\varphi^* F_k, \int_0^z (Sf_k)(w)dw\rangle_{\frac12} \to 1$, thus

$$\|T_\varphi^* F_k\|_{\mathcal{D}} \|S\|_{L(L_a^2)} \geq \left\langle T_\varphi^* F_k, \int_0^z S f_k dw \right\rangle_{\frac12} \to 1,$$

this contradicts that $\|T_\varphi^* F_k\|_{\mathcal{D}} \to 0$. Hence T_φ must be Fredholm, that is $0 \notin \sigma_e(T_\varphi)$. This follows the theorem. $\square$

Lemma 7. *For any $\varphi \in C^1(\bar{\mathbb{D}})$, $\|T_\varphi\|_e = \|[T_\varphi]\| = \|\varphi|_{\partial\mathbb{D}}\|_\infty$, where $\|T_\varphi\|_e$ denotes the essential norm of T_φ.*

Proof. Assume $\zeta \in \partial\mathbb{D}$ such that $|\varphi(\zeta)| = \|\varphi|_{\partial\mathbb{D}}\|_\infty$, set $f_\zeta(z) = \frac{1+z\bar{\zeta}}{2}$, then $f_k(z) = \frac{f_\zeta^k}{\|f_\zeta^k\|_{\mathcal{D}}} \xrightarrow{w} 0$, thus for any $K \in \mathcal{K}$, $\|Kf_k\| \to 0$. Note

$$\left|\langle T_\varphi f_k, f_k\rangle_{\frac{1}{2}}\right| = \left|\left\langle \frac{\partial\varphi}{\partial z} f_k, f_k'\right\rangle + \langle \varphi f_k', f_k'\rangle\right| \to |\varphi(\zeta)| = \|\varphi|_{\partial\mathbb{D}}\|_\infty$$

and $|\langle (T_\varphi + K)f_k, f_k\rangle_{\frac{1}{2}}| \le \|T_\varphi + K\|$. So $\|\varphi|_{\partial\mathbb{D}}\|_\infty \le \|T_\varphi + K\|$ for any $K \in \mathcal{K}$. Hence $\|\varphi|_{\partial\mathbb{D}}\|_\infty \le \|T_\varphi\|_e$.

Conversely, by Lemma 1.2 in C.K. Fong [**3**], there is an orthogonal sequence $\{f_k\} \subset \mathcal{D}$ such that $\|T_\varphi f_k\| \to \|T_\varphi\|_e$, so for any $\epsilon > 0$, there is a k_0 such that for any $k > k_0$, $\|T_\varphi f_k\| > \|T_\varphi\|_e - \epsilon$. Since $\varphi \in C^1(\bar{\mathbb{D}})$, there is a $0 < r < 1$ such that $|\varphi(z)| < \|\varphi|_{\partial\mathbb{D}}\|_\infty + \epsilon$ for any $|z| > r$. Note

$$\begin{aligned}
\|T_\varphi f_k\|_{\mathcal{D}}^2 \le \|\varphi f_k\|_{L^{2,1}}^2 = \langle \varphi f_k, \varphi f_k\rangle_{\frac{1}{2}} \\
&= \langle \varphi f_k', \varphi f_k'\rangle + \left\langle \frac{\partial\varphi}{\partial z} f_k, \varphi f_k'\right\rangle \\
&\quad + \left\langle \varphi f_k', \frac{\partial\varphi}{\partial z} f_k\right\rangle + \left\langle \frac{\partial\varphi}{\partial z} f_k, \frac{\partial\varphi}{\partial z} f_k\right\rangle \\
&\quad + \left\langle \frac{\partial\varphi}{\partial \bar{z}} f_k, \frac{\partial\varphi}{\partial \bar{z}} f_k\right\rangle,
\end{aligned}$$

clearly,

$$\left\langle \frac{\partial\varphi}{\partial z} f_k, \varphi f_k'\right\rangle + \left\langle \varphi f_k', \frac{\partial\varphi}{\partial z} f_k\right\rangle + \left\langle \frac{\partial\varphi}{\partial z} f_k, \frac{\partial\varphi}{\partial z} f_k\right\rangle + \left\langle \frac{\partial\varphi}{\partial \bar{z}} f_k, \frac{\partial\varphi}{\partial \bar{z}} f_k\right\rangle \to 0,$$

so there is a k_1 such that for any $k > k_1$, $\|T_\varphi f_k\|^2 \le |\langle \varphi f_k', \varphi f_k'\rangle| + \epsilon$. Since $f_k \xrightarrow{w} 0$, we know that $f_k \overset{\text{uniformly}}{\longrightarrow} 0$ on any compact subset of $\mathbb{D}$, so for any $0 < t < 1$, $\int_{\{|z| \le t\}} |f_k|^2 dA \to 0$ $(k \to \infty)$, assume $f_k(z) = \sum_{i=1}^\infty a_n^{(k)} z^n$, thus

$$\sum_{i=1}^\infty a_n^{(k)} \frac{t^{2(n+1)}}{n+1} = \int_{\{|z| \le t\}} |f_k|^2 dA \to 0.$$

Without loss of generality, assume $t > r$, thus

$$\int_{\{|z| \le r\}} |f_k'|^2 dA = \sum_{i=1}^\infty \left|a_n^{(k)}\right|^2 n r^{2n} = \sum_{i=1}^\infty \left|a_n^{(k)}\right|^2 \frac{t^{2(n+1)}}{n+1} n(n+1) t^{-2} \left(\frac{r}{t}\right)^{2n},$$

since $(\frac{r}{t})^{2n}n(n+1)t^{-2} \to 0$ $(n \to \infty)$, it is clearly that $\int_{\{|z|\leq r\}} |f_k'|^2 dA \to 0$, so there is a k_2 such that for any $k > k_2$, $\int_{\{|z|\leq r\}} |f_k'|^2 dA < \epsilon$, consequently,

$$\begin{aligned}
|\langle \varphi f_k', \varphi f_k' \rangle| &= \int_{\mathbb{D}} |\varphi|^2 |f_k'|^2 dA \\
&= \int_{\{|z|\leq r\}} |\varphi|^2 |f_k'|^2 dA + \int_{\{|z|>r\}} |\varphi|^2 |f_k'|^2 dA \\
&\leq \|\varphi\|_\infty^2 \int_{\{|z|\leq r\}} |f_k'|^2 dA + (\|\varphi|_{\partial\mathbb{D}}\|_\infty + \epsilon)^2 \int_{\{|z|>r\}} |f_k'|^2 dA \\
&\leq \|\varphi\|_\infty^2 \int_{\{|z|\leq r\}} |f_k'|^2 dA + (\|\varphi|_{\partial\mathbb{D}}\|_\infty + \epsilon)^2 \\
&< \|\varphi\|_\infty^2 \epsilon + (\|\varphi|_{\partial\mathbb{D}}\|_\infty + \epsilon)^2.
\end{aligned}$$

Furthermore, for any $k > \max\{k_0, k_1, k_2\}$

$$\begin{aligned}
\|T_\varphi f_k\|^2 &\leq \|\varphi\|_\infty^2 \int_{\{|z|\leq r\}} |f_k'|^2 dA + (\|\varphi|_{\partial\mathbb{D}}\|_\infty + \epsilon)^2 + \epsilon \\
&\leq (\|\varphi\|_\infty^2 + 1)\epsilon + (\|\varphi|_{\partial\mathbb{D}}\|_\infty + \epsilon)^2.
\end{aligned}$$

Hence $\lim_{k\to\infty} \|T_\varphi f_k\| \leq \|\varphi|_{\partial\mathbb{D}}\|_\infty$ by the arbitrarity of ϵ. This shows that $\|T_\varphi\|_e \leq \|\varphi|_{\partial\mathbb{D}}\|_\infty$. We are done. $\square$

Theorem 8. *Let $\mathcal{I}=\mathcal{I}(C^1)$ be the C^*-algebra generated by $\{T_\varphi | \varphi \in C^1(\bar{\mathbb{D}})\}$, then the commutator ideal $\mathcal{C}(C^1)$ of $\mathcal{I}$ equals $\mathcal{K}(\mathcal{D})$, and $\frac{\mathcal{I}}{\mathcal{K}} \cong C(\partial\mathbb{D})$. Consequently, following short sequence*

$$(*) \qquad 0 \to \mathcal{K} \to \mathcal{I} \to C(\partial\mathbb{D}) \to 0$$

is exact.

Proof. By Proposition 3, we know that $\mathcal{C}(C^1) \subset \mathcal{K}$, since $\mathcal{K}$ is minimal, $\mathcal{C}(C^1) = \mathcal{K}$. By Lemma 4, $[T_\varphi^*] = [T_{\bar{\varphi}}]$ in $\frac{\mathcal{I}}{\mathcal{K}}$. Define $\xi : \{[T_\varphi] | \varphi \in C^1(\bar{\mathbb{D}})\} \to C(\partial\mathbb{D})$ as $\xi([T_\varphi]) = \varphi|_{\partial\mathbb{D}}$, it is easy to see that ξ is well-defined, and one-to-one by Proposition 5. By Proposition 3, Lemma 4 and Lemma 7, we see that ξ is an isometric *-homomorphism. Hence ξ can be extended to $\frac{\mathcal{I}}{\mathcal{K}}$, in fact, for any $[T] \in \frac{\mathcal{I}}{\mathcal{K}}$, there is a sequence $[T_{\varphi_k}]$ such that $\|[T_{\varphi_k}] - [T]\| \to 0$, thus $\|(\varphi_k - \varphi_j)|_{\partial\mathbb{D}}\|_\infty = \|[T_{\varphi_k - \varphi_j}]\| = \|[T_{\varphi_k}] - [T_{\varphi_j}]\| \to 0$. Hence there is a $\varphi \in C(\partial\mathbb{D})$ such that $\|\varphi_k|_{\partial\mathbb{D}} - \varphi\|_\infty \to 0$. Let $\xi([T]) = \varphi$, then ξ is well-defined on $\frac{\mathcal{I}}{\mathcal{K}}$, and $\|\xi([T])\|_\infty = \|\varphi\|_\infty = \lim_{k\to\infty} \|\varphi_k|_{\partial\mathbb{D}}\|_\infty = \lim_{k\to\infty} \|[T_{\varphi_k}]\| = \|[T]\|$, so ξ is an isometry from $\frac{\mathcal{I}}{\mathcal{K}}$ into $C(\partial\mathbb{D})$. For any $\varphi \in C(\partial\mathbb{D})$, there is a polynomial sequence $\{p_k\} \subset C^1(\bar{\mathbb{D}})$ such that $\|p_k|_{\partial\mathbb{D}} - \varphi\|_\infty \to 0$, thus $\|(p_k - p_j)|_{\partial\mathbb{D}}\|_\infty \to 0$. Furthermore, $\|[T_{p_k}] - [T_{p_j}]\| \to 0$, so there is a $[T] \in \frac{\mathcal{I}}{\mathcal{K}}$ such that $\|[T_{p_k}] - [T]\| \to 0$, hence $\xi([T]) = \varphi$, i.e. ξ is a surjection onto $C(\partial\mathbb{D})$. This shows that ξ is a *-isomorphism between $\frac{\mathcal{I}}{\mathcal{K}}$ and $C(\partial\mathbb{D})$. The proof is thus completed. $\square$

3. Toeplitz operators with symbols in $H_1^\infty + C^1(\bar{\mathbb{D}})$.

It is well-known that $H^\infty + C(\partial\mathbb{D})$ is a norm-closed algebra (c.f. R.G. Douglas [**2**]), Rudin [**6**] proved that $H^\infty(\partial B_n) + C(\partial B_n)$ is also a norm-closed algebra, where B_n is the unit ball of $\mathbb{C}^n$. The Toeplitz operators with symbols in $H^\infty + C$ on Hardy or Bergman spaces have many important properties, their essential spectra and Fredholm index can be completely determined by their symbols. In this section, we prove that $H_1^\infty + C^1(\bar{\mathbb{D}})$ is a norm-closed space relative to a suitable norm, and obtain the representation of essential spectra of the Toeplitz operators with symbols in $H_1^\infty + C^1(\bar{\mathbb{D}})$.

For any $\varphi \in H_1^\infty + C^1(\bar{\mathbb{D}})$, define

$$\|\varphi\|_* = \sup_{z\in\mathbb{D}} \max\left\{|\varphi|, \left|\frac{\partial\varphi}{\partial z}\right|, \left|\frac{\partial\varphi}{\partial\bar{z}}\right|\right\},$$

it is clearly that $\|\,.\,\|_*$ is a norm on $H_1^\infty + C^1(\bar{\mathbb{D}})$.

Theorem 9. *$H_1^\infty + C^1(\bar{\mathbb{D}})$ is a closed space relative to the norm $\|\,.\,\|_*$.*

Proof. Our proof is similar to that of Rudin [**6**]. Assume $\varphi \in cl(H_1^\infty + C^1(\bar{\mathbb{D}}))$, to prove $\varphi \in H_1^\infty + C^1(\bar{\mathbb{D}})$, we first prove that for any $\psi \in H_1^\infty + C^1(\bar{\mathbb{D}})$, there are $\psi_1 \in H_1^\infty$, and $\psi_2 \in C^1(\bar{\mathbb{D}})$ such that $\psi = \psi_1 + \psi_2$ and $\|\psi_1\|_* \le 3\|\psi\|_*, \|\psi_2\|_* \le 2\|\psi\|_*$. In fact, if $\psi = \tilde{\psi}_1 + \tilde{\psi}_2 \in H_1^\infty + C^1(\bar{\mathbb{D}})$, $\tilde{\psi}_1 \in H_1^\infty$, $\tilde{\psi}_2 \in C^1(\bar{\mathbb{D}})$, then write $\tilde{\psi}_2^r(z) = \tilde{\psi}_2(rz)$, clearly, $\|\tilde{\psi}_2^r - \tilde{\psi}_2\|_\infty \to 0$ $(r \to 1^-)$, and $\frac{\partial\tilde{\psi}_2^r}{\partial z}(z) = r\frac{\partial\tilde{\psi}_2}{\partial z}(rz)$, $\frac{\partial\tilde{\psi}_2^r}{\partial\bar{z}}(z) = r\frac{\partial\tilde{\psi}_2}{\partial\bar{z}}(rz)$, so $\|\tilde{\psi}_2^r - \tilde{\psi}_2\|_* \to 0$. Fix a r_0 such that $\|\tilde{\psi}_2^{r_0} - \tilde{\psi}_2\|_* \le \|\psi\|_*$. Set

$$\psi_2 = \tilde{\psi}_2 - \tilde{\psi}_2^{r_0} + \psi^{r_0},\ \psi_1 = \tilde{\psi}_1 - \tilde{\psi}_1^{r_0},$$

then $\psi = \psi_1 + \psi_2$, and $\|\psi_1\|_* = \|\psi - \psi_2\|_* \le 3\|\psi\|_*$, $\|\psi_2\|_* \le 2\|\psi\|_*$. Since $\varphi \in cl(H_1^\infty + C^1(\bar{\mathbb{D}}))$, there are $\varphi_i \in H_1^\infty + C^1(\bar{\mathbb{D}})$, such that $\|\varphi_i\|_* \le 2^{-i}$ $(i \ge 2)$, and $\varphi = \sum_{i=1}^\infty \varphi_i$ (in fact, there is a sequence $\{p_k\} \subset H_1^\infty + C^1(\bar{\mathbb{D}})$ such that $\|p_k - \varphi\|_* \to 0$, thus there is a subsequence $\{p_k\}$ such that $\|p_{k_{i+1}} - p_{k_i}\|_* < \frac{1}{2^i}$, write $\varphi_1 = p_{k_1}, \varphi_i = p_{k_{i+1}} - p_{k_i}$ $(i \ge 2)$, then $\sum_{i=1}^\infty \varphi_i = \varphi$ and $\|\varphi_i\|_* < \frac{1}{2^i}$ $(i \ge 2)$). For each φ_i, there is a $\varphi_i^{(1)} \in H_1^\infty$, $\varphi_i^{(2)} \in C^1(\bar{\mathbb{D}})$ such that $\|\varphi_i^{(1)}\|_* \le 3\|\varphi_i\|_*$, $\|\varphi_i^{(2)}\|_* \le 2\|\varphi_i\|_*$, $\varphi_i = \varphi_i^{(1)} + \varphi_i^{(2)}$, thus $\varphi^{(1)} = \sum_{i=1}^\infty \varphi_i^{(1)} \in H_1^\infty$ (since H_1^∞ is closed relative to $\|.\|_*$), $\varphi^{(2)} = \sum_{i=1}^\infty \varphi_i^{(2)} \in C^1(\bar{\mathbb{D}})$, and $\varphi = \varphi^{(1)} + \varphi^{(2)}$. That is $\varphi \in H_1^\infty + C^1(\bar{\mathbb{D}})$. ◻

Proposition 10. *If $\varphi \in H_1^\infty + C^1(\bar{\mathbb{D}})$ satisfies $\varphi|_{\partial\mathbb{D}} = 0$, then T_φ is a compact operator on $\mathcal{D}$, where $\varphi|_{\partial\mathbb{D}}$ denotes the radial boundary values of φ.*

Proof. Suppose $\varphi = \varphi_1 + \varphi_2, \varphi_1 \in H_1^\infty, \varphi_2 \in C^1(\bar{\mathbb{D}})$, since $\varphi|_{\partial\mathbb{D}} = 0$, $\varphi_1|_{\partial\mathbb{D}} = -\varphi_2|_{\partial\mathbb{D}}$, thus $\varphi_1|_{\partial\mathbb{D}} \in C(\partial\mathbb{D})$, further $\varphi_1(z) = P[\varphi_1|_{\partial\mathbb{D}}] \in H_1^\infty \cap A(\mathbb{D})$ (where $P[\varphi|_{\partial\mathbb{D}}]$ denotes the Poisson integral of $\varphi|_{\partial\mathbb{D}}$), consequently, $\varphi \in C(\bar{\mathbb{D}}) \cap (H_1^\infty + C^1(\bar{\mathbb{D}}))$.

Suppose $\{f_k\} \subset \mathcal{D}$, $\|f_k\|_{\mathcal{D}} = 1$, $f_k \xrightarrow{w} 0$, then

$$\begin{aligned}
\|T_\varphi f_k\|_{\mathcal{D}}^2 &\le \|\varphi f_k\|_{L^{2,1}}^2 = \langle \varphi f_k, \varphi f_k \rangle_{\frac{1}{2}} \\
&= \left\langle \frac{\partial \varphi}{\partial z} f_k + f_k' \varphi, \frac{\partial \varphi}{\partial z} f_k + f_k' \varphi \right\rangle + \left\langle \frac{\partial \varphi}{\partial \bar{z}} f_k, \frac{\partial \varphi}{\partial \bar{z}} f_k \right\rangle \\
&= \left\| \frac{\partial \varphi}{\partial z} f_k \right\|_{L^2}^2 + 2\mathrm{Re} \int_{\mathbb{D}} f_k' \varphi \overline{\frac{\partial \varphi}{\partial z} f_k} dA + \|\varphi f_k'\|_{L^2}^2 + \left\| \frac{\partial \varphi}{\partial \bar{z}} f_k \right\|_{L^2}^2 \\
&\le \|\varphi\|_*^2 \|f_k\|_{L^2}^2 + 2\|\varphi\|_*^2 \|f_k\|_{\mathcal{D}} \|f_k\|_{L^2} \\
&\quad + \|\varphi f_k'\|_{L^2}^2 + \|\varphi\|_*^2 \|f_k\|_{L^2}^2.
\end{aligned}$$

Since $\|f_k\|_{\mathcal{D}} = 1$ and $f_k \xrightarrow{w} 0$ in $\mathcal{D}$, $\|f_k\|_{L^2} \to 0$, so

$$\|\varphi\|_*^2 \|f_k\|_{L^2}^2 + 2\|\varphi\|_*^2 \|f_k\|_{\mathcal{D}} \|f_k\|_{L^2} + \|\varphi\|_*^2 \|f_k\|_{L^2}^2 \to 0.$$

Note $\varphi \in C(\bar{\mathbb{D}})$, and $\varphi|_{\partial \mathbb{D}} = 0$, hence for any $\epsilon > 0$, there is a $0 < \delta < 1$ such that $|\varphi(z)| < \epsilon$ for any $|z| > \delta$, thus

$$\begin{aligned}
\int_{\mathbb{D}} |\varphi|^2 |f_k'|^2 dA &\le \epsilon^2 \int_{\{|z|>\delta\}} |f_k'|^2 dA + \int_{\{|z|\le\delta\}} |\varphi|^2 |f_k'|^2 dA \\
&\le \epsilon^2 + \|\varphi\|_*^2 \int_{\{|z|\le\delta\}} |f_k'|^2 dA.
\end{aligned}$$

By $f_k \xrightarrow{w} 0$, we know that $\int_{\{|z|\le\delta\}} |f_k'|^2 dA \to 0$, hence $\int_{\mathbb{D}} |\varphi|^2 |f_k'|^2 dA \to 0$ $(k \to \infty)$, furthermore $\|T_\varphi f_k\|_{\mathcal{D}}^2 \to 0$, this shows that T_φ is compact on $\mathcal{D}$. ☐

Proposition 11. *If $\varphi \in H_1^\infty(\mathbb{D})$, then on Dirichlet space $\mathcal{D}$, $\sigma(T_\varphi) = \overline{\varphi(\mathbb{D})}$.*

Proof. Note for any $f \in \mathcal{D}$,

$$\langle f, T_\varphi^* K \rangle_{\frac{1}{2}} = \langle T_\varphi f, K \rangle_{\frac{1}{2}} = \langle \varphi f, K \rangle_{\frac{1}{2}} = \varphi(z) f(z),$$

so $T_\varphi^* K(z,w) = \overline{\varphi(z)} K(z,w)$, this shows that $\overline{\varphi(\mathbb{D})} \subset \sigma(T_\varphi)$.

Conversely, if $0 \notin \overline{\varphi(\mathbb{D})}$, then there is a $\delta > 0$ such that $|\varphi(z)| > \delta > 0$ $(\forall z \in \mathbb{D})$, thus $\varphi^{-1} \in H^\infty$. Note

$$\left| \frac{\partial(\varphi^{-1})}{\partial z} \right| = \left| -\frac{\varphi'}{\varphi^2} \right| \le \frac{1}{\delta^2} |\varphi'| \le \frac{1}{\delta^2} \|\varphi\|_* < \infty,$$

we have $\varphi^{-1} \in H_1^\infty$, so $T_{\varphi^{-1}} T_\varphi = T_\varphi T_{\varphi^{-1}} = I$, that is T_φ is invertible, hence $\sigma(T_\varphi) \subset \overline{\varphi(\mathbb{D})}$. The proof is thus complete. ☐

Theorem 12. *If $\varphi = \varphi_1 + \varphi_2 \in H_1^\infty + C^1(\bar{\mathbb{D}})$, where $\varphi_1 \in H_1^\infty$, $\varphi_2 \in C^1(\bar{\mathbb{D}})$, then on Dirichlet space $\mathcal{D}$,*

$$\sigma_e(T_\varphi) = \overline{\cap_{0<\delta<1} \{\varphi(z) \| z| > \delta\}} = \varphi(\partial \mathbb{D}).$$

Proof. Without loss of generality, assume $0 \in \overline{\cap_{0<\delta<1}\{\varphi(z)||z|>\delta\}}$, then there is a sequence $\{z_k\} \subset \mathbb{D}$ with $z_n \to \zeta \in \partial\mathbb{D}$, such that $\varphi(z_n) \to 0$, thus $\tilde{T}_\varphi$ is not Fredholm on L^2_a (c.f. G.F. Cao [**1**] or MacDonald [**4**]), hence there is a sequence $\{f_k\} \subset L^2_a$ with $\|f_k\|_{L^2} = 1, f_k \xrightarrow{w} 0$ such that $\|\tilde{T}_\varphi f_k\|_{L^2} \to 0$ or $\|\tilde{T}^*_\varphi f_k\|_{L^2} = \|\tilde{T}_{\bar{\varphi}} f_k\|_{L^2} \to 0$. If T_φ is Fredholm on $\mathcal{D}$, then there is a bounded operator S on $\mathcal{D}$ such that $ST_\varphi - I, T_\varphi S - I \in \mathcal{K}(\mathcal{D})$, further $S^*T^*_\varphi - I, T^*_\varphi S^* - I \in \mathcal{K}(\mathcal{D})$. Write $F_k(z) = \int_0^z f_k(w)dw$, then $F_k \in \mathcal{D}$ and $\|F_k\|_{\mathcal{D}} = \|f_k\|_{L^2} = 1$, so $\|F_k\|_{L^2} \to 0$. In addition, it is clearly that $F_k \xrightarrow{w} 0$ in $\mathcal{D}$. Thus

$$\begin{aligned}
\lim_{k\to\infty} \langle ST_\varphi F_k, F_k\rangle_{\frac{1}{2}} &= \lim_{k\to\infty} \langle T_\varphi S F_k, F_k\rangle_{\frac{1}{2}} \\
&= \lim_{k\to\infty} \left\langle T^*_\varphi S^* F_k, F_k\right\rangle_{\frac{1}{2}} \\
&= \lim_{k\to\infty} \left\langle S^* T^*_\varphi F_k, F_k\right\rangle_{\frac{1}{2}} = 1.
\end{aligned}$$

Hence, without loss of generality, we can assume that $\|\tilde{T}^*_\varphi f_k\|_{L^2} \to 0$ (similarly for the case of $\|\tilde{T}_\varphi f_k\|_{L^2} \to 0$). Set $G_k = SF_k$, then $\|G_k\|_{\mathcal{D}} \le \|S\|\|F_k\|_{\mathcal{D}} = \|S\|$, and $G_k \xrightarrow{w} 0$, so $\|G_k\|_{L^2} \to 0$, and

$$\begin{aligned}
\left|\langle T_\varphi S F_k, F_k\rangle_{\frac{1}{2}}\right| &= \left|\langle \varphi(SF_k)', f_k\rangle + \left\langle \frac{\partial\varphi}{\partial z} SF_k, f_k \right\rangle\right| \\
&\le \left|\left\langle (SF_k)', \tilde{T}^*_\varphi f_k \right\rangle\right| + \|\varphi\|_* \|SF_k\|_{L^2}\|f_k\|_{L^2} \\
&\le \|(SF_k)'\|_{L^2} \left\|\tilde{T}^*_\varphi f_k\right\|_{L^2} + \|\varphi\|_* \|G_k\|_{L^2}\|f_k\|_{L^2} \\
&= \|G_k\|_{\mathcal{D}} \left\|\tilde{T}^*_\varphi f_k\right\|_{L^2} + \|\varphi\|_* \|G_k\|_{L^2}\|f_k\|_{L^2} \to 0.
\end{aligned}$$

This contradicts that $\langle T_\varphi S F_k, F_k\rangle_{\frac{1}{2}} \to 1$. hence T_φ must be non-Fredholm. That is $0 \in \sigma_e(T_\varphi)$, consequently $\overline{\cap_{0<\delta<1}\{\varphi(z)||z|>\delta\}} \subset \sigma_e(T_\varphi)$.

Conversely, assume $0 \notin \overline{\cap_{0<\delta<1}\{\varphi(z)||z|>\delta\}}$, thus there are $\epsilon, \delta > 0$ such that $|\varphi(z)| > \epsilon$ for any $|z| > \delta$, we prove that T_φ is Fredholm on $\mathcal{D}$. Otherwise, there is a sequence $\{F_k\} \subset \mathcal{D}$ with $\|F_k\|_{\mathcal{D}} = 1, F_k \xrightarrow{w} 0$, such that $\|T_\varphi F_k\|_{\mathcal{D}} \to 0$ or $\|T^*_\varphi F_k\|_{\mathcal{D}} \to 0$. Similar to above proof, we can assume $\|T^*_\varphi F_k\|_{\mathcal{D}} \to 0$ (if $\|T_\varphi F_k\|_{\mathcal{D}} \to 0$, the proof will be simpler). It is well-known that $\xrightarrow{\sim} T_\varphi$ is Fredholm on L^2_a if $|\varphi(z)| > \epsilon$ for any $|z| > \delta$, hence there is a $S \in L(L^2_a)$ such that $S\tilde{T}_\varphi - I, \tilde{T}_\varphi S - I \in \mathcal{K}(L^2_a)$, further, $S^*\tilde{T}^*_\varphi - I, \tilde{T}^*_\varphi S^* - I \in \mathcal{K}(L^2_a)$. Set $f_k = F'_k, g_k = Sf_k$, then $f_k \xrightarrow{w} 0$, $g_k \xrightarrow{w} 0$ in L^2_a and $\|f_k\|_{L^2_a} = \|F_k\|_{\mathcal{D}} = 1$, $\|g_k\|_{L^2_a} \le \|S\|\|f_k\|_{L^2_a} = \|S\|$, so $\langle S^*\tilde{T}^*_\varphi f_k, f_k\rangle \to$

1, that is $\langle \tilde{T_\varphi^*} f_k, Sf_k\rangle \to 1$. Note

$$\begin{aligned}\left\langle T_\varphi^* F_k, \int_0^z Sf_k(w)dw \right\rangle_{\frac{1}{2}} &= \left\langle F_k, \varphi \int_0^z Sf_k dw \right\rangle_{\frac{1}{2}} \\ &= \left\langle f_k, \left(\frac{\partial \varphi}{\partial z} \int_0^z Sf_k dw + \varphi S f_k \right) \right\rangle_{L^2} \\ &= \left\langle f_k, \frac{\partial \varphi}{\partial z} \int_0^z g_k dw \right\rangle + \langle f_k, \varphi g_k \rangle\end{aligned}$$

and $g_k \xrightarrow{w} 0$, so $\| \int_0^z g_k dw\|_{L^2} \to 0$, and

$$\left| \left\langle f_k, \frac{\partial \varphi}{\partial z} \int_0^z g_k dw \right\rangle \right| \le \|f_k\|_{L^2_a} \|\varphi\|_* \left\| \int_0^z g_k dw \right\|_{L^2} \to 0,$$

but $\langle f_k, \varphi g_k\rangle = \langle \tilde{T_\varphi^*} f_k, Sf_k\rangle \to 1 \ne 0$, hence $\langle T_\varphi^* F_k, \int_0^z g_k dw\rangle_{\frac{1}{2}} \to 1 \ne 0$, this contradicts that $\|T_\varphi^* F_k\|_{\mathcal{D}} \to 0$. It shows that T_φ must be Fredholm on $\mathcal{D}$. That is $0 \notin \sigma_e(T_\varphi)$. Note the functions with derivatives in H^∞ are continuous in the closed unit disk, hence $H_1^\infty + C^1(\bar{\mathbb{D}}) \subset C(\bar{\mathbb{D}})$, consequently, $\overline{\cap_{0<\delta<1}\{\varphi(z) | |z| > \delta\}} = \varphi(\partial\mathbb{D})$. We are done. □

Remark. If $\varphi \in H^\infty$, then $T_{\bar\varphi}$ is always bounded on $\mathcal{D}$, in fact, for any $f, g \in \mathcal{D}$,

$$\begin{aligned}\left|\langle T_{\bar\varphi} f, g\rangle_{\frac{1}{2}}\right| &= \left|\langle \bar\varphi f, g\rangle_{\frac{1}{2}}\right| = |\langle \bar\varphi f', g'\rangle| \\ &= \left|\left\langle \tilde{T_{\bar\varphi}} f', g'\right\rangle\right| \le \left\|\tilde{T_{\bar\varphi}}\right\| \|f'\|_{L^2} \|g'\|_{L^2} \le \|\varphi\|_\infty \|f\|_{\mathcal{D}} \|g\|_{\mathcal{D}},\end{aligned}$$

hence $\|T_{\bar\varphi} f\| \le \|\varphi\|_\infty \|f\|_{\mathcal{D}}$. However, for $\varphi \in H^\infty$, T_φ may be unbounded on $\mathcal{D}$. Let

$$L_1^{\infty,1} = \left\{ f \in L^{2,1} \middle| f, \frac{\partial f}{\partial z} \in L^\infty(\mathbb{D}, dA) \right\},$$

then $L_1^{\infty,1} \cup \overline{H^\infty}$ is perhaps the most suitable symbol space of Toeplitz operators which are bounded. Also, we can prove that if $\varphi \in L_1^{\infty,1}$ has a compact support set, then T_φ is compact on $\mathcal{D}$.

4. An index formula of Toeplitz operators.

The classical index formula shows that for $\varphi \in C(\partial\mathbb{D})$, if $|\varphi| > \epsilon > 0$, then T_φ is a Fredholm operator on Hardy space $H^2(\partial\mathbb{D})$, and $\mathrm{Ind}T_\varphi = -\mathrm{wind}\varphi$. If $\varphi \in C(\bar{\mathbb{D}})$ with $|(\varphi|_{\partial\mathbb{D}})(\zeta)| > \epsilon > 0$, then $\mathrm{Ind}T_\varphi = -\mathrm{wind}\varphi|_{\partial\mathbb{D}}$ on Bergman space $L_a^2(\mathbb{D})$. The proof of these index formulas is relative to the topology homotopy of symbol functions, it can not be directly extended to the case of Dirichlet space since we do not know whether there is a C^1-function H_t which is continuous with respect to $t \in [0, 1]$ such that $H_0 = \varphi, H_1 = \psi$ if $\varphi, \psi \in C^1(\bar{\mathbb{D}})$ and $\mathrm{wind}\,\varphi|_{\partial\mathbb{D}} = \mathrm{wind}\psi|_{\partial\mathbb{D}}$. In this section, we use the short exact sequence $(*)$ to prove an analogy of the above index formulas for C^1-symbols.

Theorem 13. *Suppose $\varphi \in C^1(\bar{\mathbb{D}})$ such that $T_{\bar{\varphi}}$ is Fredholm on $\mathcal{D}$, then*

$$\mathrm{Ind}T_{\varphi} = -\mathrm{wind}\varphi|_{\partial\mathbb{D}}.$$

Proof. By Theorem 8, if $\varphi, \psi \in C^1(\bar{\mathbb{D}})$ such that T_φ and T_ψ are Fredholm on $\mathcal{D}$ and $\mathrm{wind}\varphi|_{\partial\mathbb{D}} = \mathrm{wind}\psi|_{\partial\mathbb{D}}$, then there is a $H_t \in C([0,1] \times \partial\mathbb{D})$ such that $H_t \in GC(\partial\mathbb{D})$ (the set of invertible elements in $C(\partial\mathbb{D})$) for each $t \in [0,1]$ and $H_0 = \varphi|_{\partial\mathbb{D}}, H_1 = \psi|_{\partial\mathbb{D}}$. Note ξ is an isometry isomorphism, so $\xi^{-1}(H_t)$ is continuous with respect to t. On the other hand, H_t^{-1} is also continuous on $[0,1] \times \partial\mathbb{D}$, and $\xi^{-1}(H_t)\xi^{-1}(H_t^{-1}) = \xi^{-1}(H_t^{-1})\xi^{-1}(H_t) = \xi^{-1}(H_t H_t^{-1}) = [I]$. Hence, $\xi^{-1}(H_t)$ is invertible in $\frac{\mathcal{I}}{\mathcal{K}}$. Furthermore, it is easy to see that $\mathrm{Ind}T_\varphi = \mathrm{Ind}T_\psi$.

Now suppose $\varphi \in C^1(\bar{\mathbb{D}})$ such that T_φ is Fredholm on $\mathcal{D}$ and $\mathrm{wind}\varphi|_{\partial\mathbb{D}} = k$, note T_z is Fredholm on $\mathcal{D}$ with $\mathrm{Ind}T_z = -1 = -\mathrm{wind}z|_{\partial\mathbb{D}}$, let

$$\tilde{z}^k = \begin{cases} z^k, & \text{if } k > 0, \\ \bar{z}^{-k}, & \text{otherwise}; \end{cases}$$

we see that $\mathrm{wind}\varphi|_{\partial\mathbb{D}} = \mathrm{wind}\tilde{z}^k|_{\partial\mathbb{D}}$, thus $\mathrm{Ind}T_\varphi = \mathrm{Ind}T_{\tilde{z}^k}$ by above proof, consequently, $\mathrm{Ind}T_\varphi = -k = -\mathrm{wind}\varphi|_{\partial\mathbb{D}}$. The proof is thus complete. □

Proposition 14. *Suppose $\varphi \in H^\infty$ such that $T_{\bar{\varphi}}$ is Fredholm on $\mathcal{D}$, then*

$$\mathrm{Ind}T_{\bar{\varphi}} = -\lim_{r \to 1^-} \mathrm{wind}\bar{\varphi}_r|_{\partial\mathbb{D}},$$

where $\varphi_r(z) = \varphi(rz)$.

Proof. If $f \in \mathrm{Ker}T_{\bar{\varphi}}$, then for any $g \in \mathbb{D}$,

$$0 = \langle T_{\bar{\varphi}} f, g\rangle_{\frac{1}{2}} = \langle \bar{\varphi} f, g\rangle_{\frac{1}{2}} = \left\langle \bar{\varphi} f', g' \right\rangle = \left\langle \tilde{T}_{\bar{\varphi}} f', g' \right\rangle,$$

since $\{g' | g \in \mathcal{D}\} = L_a^2$, $\tilde{T}_{\bar{\varphi}} f' = 0$, hence $\{f' | f \in \mathrm{Ker}T_{\bar{\varphi}}\} \subset \mathrm{Ker}\widetilde{T}_{\bar{\varphi}}$. Consequently, assume $\tilde{T}_{\bar{\varphi}} f = 0, f \in L_a^2$, set $F(z) = \int_0^z f dw$, then $F \in \mathcal{D}$, and for any $G \in \mathcal{D}$,

$$\langle T_{\bar{\varphi}} F, G\rangle_{\frac{1}{2}} = \left\langle \bar{\varphi} F', G' \right\rangle = \left\langle \widetilde{T}_{\bar{\varphi}} f, G' \right\rangle = 0,$$

so $T_{\bar{\varphi}} F = 0$, further $\{\int_0^z f(w)dw | f \in \mathrm{Ker}\widetilde{T}_{\bar{\varphi}}\} \subset \mathrm{Ker}T_{\bar{\varphi}}$. This shows that

$$\dim\, \mathrm{Ker}T_{\bar{\varphi}} = \dim\, \mathrm{Ker}\tilde{T}_{\bar{\varphi}}.$$

Now assume $f \in \mathcal{D}$ such that $T_{\bar{\varphi}}^* f = 0$, then for any $g \in \mathcal{D}$,

$$\begin{aligned} 0 &= \left\langle T_{\bar{\varphi}}^* f, g \right\rangle_{\frac{1}{2}} = \langle f, \bar{\varphi} g\rangle_{\frac{1}{2}} \\ &= \left\langle f', \bar{\varphi} g' \right\rangle = \left\langle \varphi f' g' \right\rangle = \left\langle \tilde{T}_\varphi f', g' \right\rangle, \end{aligned}$$

hence $\tilde{T}_\varphi f' = 0$, that is $\{f' | f \in \mathrm{Ker} T^*_{\bar{\varphi}}\} \subset \mathrm{Ker}\tilde{T}_\varphi$.

Similarly to above proof, we have also $\{\int_0^z f(w)dw | f \in \mathrm{Ker}\tilde{T}_\varphi\} \subset \mathrm{Ker}T^*_{\bar{\varphi}}$. Hence dim $\mathrm{Ker}T^*_{\bar{\varphi}} =$ dim $\mathrm{Ker}\tilde{T}_\varphi$, consequently,

$$\mathrm{Ind}T_{\bar{\varphi}} = -\mathrm{Ind}\tilde{T}_\varphi = \lim_{r\to 1^-} \mathrm{wind}\varphi_r|_{\partial\mathbb{D}} = -\lim_{r\to 1^-} \mathrm{wind}\bar{\varphi}_r|_{\partial\mathbb{D}}.$$

$\square$

Acknowledgement. The referee gives many invaluable suggestions and indicates that the functions with derivatives in H^∞ are continuous in the closed unit disk. The author is indebted to him for his generous help.

References

[1] G.F. Cao and Sh.L. Sun, *The connectivity of essential spectra of Toeplitz operators with symbols in $H^\infty + C$*, Chinese Ann. of Math., **1** (1995), 93-102.

[2] R.G. Douglas, *Banach algebra techniques in operator theory*, Academic Press, 1972.

[3] C.K. Fong, *On the essential maximal numerical range*, Acta Sci. Math., **41** (1979), 307-315.

[4] McDonald, *Fredholm properties of a class of Toeplitz operators on the ball*, Indiana Math. J., **26** (1977), 567-576.

[5] R. Rochberg and Z.J. Wu, *Toeplitz operators on Dirichlet spaces*, Integr. Equat. Oper. Th., **15** (1992), 325-342.

[6] W. Rudin, *Function theory in the unit ball of $\mathbb{C}^n$*, Springer-Verlag, 1980.

[7] Z.J. Wu, *Hankel and Toeplitz operators on Dirichlet spaces*, Integr. Equat. Oper. Th., **15** (1992), 503-525

Received July 30, 1997 and revised January 15, 1998. This work was supported by National Natural Science Foundation of China.

Sichuan University
Chengdu 610064
China
E-mail address: nic2601@scuu.edu.cu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

CONGRUENCE OF TWO-DIMENSIONAL SUBSPACES IN $M_2(K)$ (CHARACTERISTIC $\neq 2$)

B. Corbas and G.D. Williams

The structure and classification up to isomorphism of a naturally arising class of local rings is determined. Although we are primarily interested in the case of a finite residue field K, our results apply in fact over any field K of characteristic $\neq 2$. The problem is shown to be equivalent to that of classifying two-dimensional subspaces of $M_2(K)$ up to congruence, and it is in these terms that the question is addressed.

1. Introduction.

In investigating the structure of finite local rings one is led to consider such a ring of the form $R = K \oplus J$ in which $K = \mathbf{F}_q$ and the Jacobson radical J is such that $J^3 = 0$ and both J/J^2 and J^2 are two-dimensional over $R/J = K$. Rings with $J^3 = 0$ form a natural object of study, the case $J^2 = 0$ having long been settled [**2**, **3**]. If $J = Kx_1 \oplus Kx_2 \oplus J^2$ and $J^2 = Ky_1 \oplus Ky_2$, then we may write $x_ix_j = \alpha_{ij}y_1 + \beta_{ij}y_2$ ($\alpha_{ij}, \beta_{ij} \in K$) and these four products span J^2. The ring structure is determined by the pair of (2×2) matrices $A = (\alpha_{ij})$, $B = (\beta_{ij})$, which are linearly independent over K, and any pair of independent matrices defines such a ring.We wish to determine the number of isomorphism classes of such rings and to find normal forms for the pair of matrices A, B defining them. Chikunji [**1**] has shown that there are 10 classes for $q = 2$ and, on the basis of computer calculations for $q = 3, 5, 7$, has conjectured that when q is odd the number of classes is $3q + 5$. It is also conjectured that exactly three of these rings are commutative. Our purpose here is *inter alia*, to prove these conjectures.

If (x_1', x_2', y_1', y_2') is a new basis of J with corresponding matrices A', B', then x_1', x_2' are linear combinations of x_1, x_2, y_1, y_2. Since $J^3 = 0$, we may assume that the coefficients of y_1, y_2 are zero and write $x_i' = p_{1i}x_1 + p_{2i}x_2$, so that $P = (p_{ij})$ is the transition matrix from the basis $(\overline{x}_1, \overline{x}_2)$ of J/J^2 to the basis $(\overline{x}_1', \overline{x}_2')$. Equally, let $Q = (q_{ij})$ be the transition matrix from the basis (y_1, y_2) of J^2 to (y_1', y_2'). If we now calculate $x_i'x_j'$ and compare coefficients of y_i we obtain equations which, in matrix form, are

$$\begin{cases} P^tAP = q_{11}A' + q_{12}B' \\ P^tBP = q_{21}A' + q_{22}B' \end{cases}.$$

Evidently, the problem of classifying our rings up to isomorphism amounts to that of classifying pairs of linearly independent matrices (A, B) under the above relation of *equivalence*, P and Q being arbitrary invertible matrices, and it is to this problem of linear algebra that the paper is devoted. We shall, in fact, solve it over an arbitrary field of characteristic $\neq 2$ and will consider all pairs, independent or otherwise. The approach we take is to first of all deal with pairs of *symmetric* matrices (corresponding to commutative rings) and then to use the fact that a general equivalence class may be represented by the sum of one of the standard symmetric pairs already found with an antisymmetric pair. This is similar to an idea used in [**4**] for congruence of single matrices.

2. The symmetric case.

We first establish some notation. Let X be the set of all pairs (A, B) of (2×2) matrices over a field K. The group GL_2 acts on the right on X by congruence: $(A, B) \cdot P = (P^t AP, P^t BP)$ and on the left via $Q \cdot (A, B) = (q_{11}A + q_{12}B, q_{21}A + q_{22}B)$, where $Q = (q_{ij})$. These two actions are permutable and define a (left) action of $G = GL_2 \times GL_2$ on X:

$$(P, Q) \cdot (A, B) = Q \cdot (A, B) \cdot P^{-1}.$$

By restriction, G acts on the subset Y consisting of pairs with A, B linearly independent. This amounts to studying the congruence action (via P) of GL_2 on the set $\mathcal{Y}$ of 2-dimensional subspaces of $M_2(K)$, Q just representing a change of basis in a given subspace. In the same way, the whole action of G on X may be reinterpreted as an action of GL_2 on the set $\mathcal{X}$ of subspaces of dimension ≤ 2. Two pairs in the same G-orbit will be called *equivalent*.

G also acts by restriction on the set S of pairs with A, B symmetric. *Assuming henceforth that* char $K \neq 2$, we determine these orbits first. To avoid a plague of parentheses we omit these around ordered pairs of displayed matrices.

Theorem 1. *The following table gives a complete set of representatives for the orbits of G on S, together with their stabilizers:*

Representative	Stabilizing elements (P, Q)
1. $\begin{pmatrix}1 & \\ & 0\end{pmatrix}, \begin{pmatrix} & 1\\ 1 & \end{pmatrix}$	$\begin{pmatrix}a & 0\\ c & d\end{pmatrix}, \begin{pmatrix}a^2 & 0\\ 2ac & ad\end{pmatrix}$
2. $\begin{pmatrix}1 & \\ & \delta\end{pmatrix}, \begin{pmatrix} & 1\\ 1 & \end{pmatrix}$	$\begin{pmatrix}a & \pm\delta c\\ c & \pm a\end{pmatrix}, \begin{pmatrix}a^2+\delta c^2 & \pm 2\delta ac\\ 2ac & \pm(a^2+\delta c^2)\end{pmatrix}$
3. $\begin{pmatrix}0 & \\ & 0\end{pmatrix}, \begin{pmatrix}0 & \\ & 0\end{pmatrix}$	All
4. $\begin{pmatrix}1 & \\ & 0\end{pmatrix}, \begin{pmatrix}0 & \\ & 0\end{pmatrix}$	$\begin{pmatrix}a & 0\\ c & d\end{pmatrix}, \begin{pmatrix}a^2 & l\\ 0 & n\end{pmatrix}$
5. $\begin{pmatrix}1 & \\ & \delta\end{pmatrix}, \begin{pmatrix}0 & \\ & 0\end{pmatrix}$	$\begin{pmatrix}a & \mp\delta c\\ c & \pm a\end{pmatrix}, \begin{pmatrix}a^2+\delta c^2 & l\\ 0 & n\end{pmatrix}$

In 2) *and* 5) δ *runs through a set of coset representatives of* K^{*2} *in* K^**.*

Write $P = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$ and $Q = \begin{pmatrix}k & l\\ m & n\end{pmatrix}$. Before giving the proof it is useful to record that if $A = \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$, then:

$$P^t AP = \begin{pmatrix}a^2\alpha + ac(\beta+\gamma) + c^2\delta & ab\alpha + ad\beta + bc\gamma + cd\delta\\ ab\alpha + ad\gamma + bc\beta + cd\delta & b^2\alpha + bd(\beta+\gamma) + d^2\delta\end{pmatrix}.$$

In particular we have:

A	$\begin{pmatrix}1 & \\ & 0\end{pmatrix}$	$\begin{pmatrix}1 & \\ & \delta\end{pmatrix}$	$\begin{pmatrix} & 1\\ 1 & \end{pmatrix}$
P^tAP	$\begin{pmatrix}a^2 & ab\\ ab & b^2\end{pmatrix}$	$\begin{pmatrix}a^2+c^2\delta & ab+cd\delta\\ ab+cd\delta & b^2+d^2\delta\end{pmatrix}$	$\begin{pmatrix}2ac & ad+bc\\ ad+bc & 2bd\end{pmatrix}$

Note also that (P, Q) fixes a pair $\Pi = (A, B) \Leftrightarrow \Pi \cdot P = Q \cdot \Pi$.

Proof of Theorem 1. Consider first *independent* pairs (A, B) in S. We claim that any such pair in equivalent to one with $B = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$. To prove this it is enough to show that every 2-dimensional subspace W of the space V of symmetric matrices contains an *isotropic* matrix, in the sense that it is nonsingular and the associated quadratic form represents zero. For all isotropic matrices are congruent to this one. If W equals the space of diagonal matrices, then it contains the isotropic matrix $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$. If not, then, since $\dim V = 3$, W is spanned by a diagonal matrix and a non-diagonal matrix $\begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix}$. We may clearly modify the latter so that α or δ equals 0, and then it is isotropic.

So now let (A, B) be independent, with $B = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$. We may take A to be diagonal, $A = \begin{pmatrix} \alpha & \\ & \delta \end{pmatrix}$. Under congruence by $P = B$, if necessary, we may assume that $\alpha \neq 0$, and then, via a suitable Q, that $\alpha = 1$.

We now determine when two pairs $\Pi = \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ and $\Pi' = \begin{pmatrix} 1 & \\ & \delta' \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ are equivalent. This happens when there exist P, Q as above such that $\Pi \cdot P = Q \cdot \Pi'$, or in other words:

(1)
$$\begin{pmatrix} a^2 + c^2\delta & ab + cd\delta \\ ab + cd\delta & b^2 + d^2\delta \end{pmatrix}, \begin{pmatrix} 2ac & ad + bc \\ ad + bc & 2bd \end{pmatrix} = \begin{pmatrix} k & l \\ l & k\delta' \end{pmatrix}, \begin{pmatrix} m & n \\ n & m\delta' \end{pmatrix}.$$

Comparing diagonal terms gives

$$\begin{cases} b^2 + d^2\delta = \delta'(a^2 + c^2\delta) \\ bd = \delta' ac \end{cases} . \tag{2}$$

Squaring these and subtracting 4δ times the second from the first leads to $b^2 - d^2\delta = \pm\delta'(a^2 - c^2\delta)$. According to the sign, there are two cases:

$$\text{(i)} \quad \begin{cases} b^2 = \delta' a^2 \\ \delta d^2 = \delta\delta' c^2 \end{cases} \qquad \text{or} \quad \text{(ii)} \quad \begin{cases} b^2 = \delta\delta' c^2 \\ \delta d^2 = \delta' a^2 \end{cases} .$$

In either case it follows from nonsingularity of P that if $\delta' = 0$, then $b = 0$, $d \neq 0$ and $\delta = 0$. By symmetry we deduce that $\delta = 0 \Leftrightarrow \delta' = 0$. The stabilizer in this case is given by the single condition $b = 0$, and the form of Q follows from (1).

Assume now that $\delta, \delta' \neq 0$. Case (i) cannot now arise, as is shown by the second equation of (2), the first of (i) and nonsingularity of P. It follows from (ii) that Π and Π' are equivalent $\Leftrightarrow \delta, \delta'$ are in the same square-class. The form of the stabilizer results at once.

We are left with the *dependent* pairs (A, B) in S. Via Q we may assume that $B = 0$, and then (via P) that $A = \begin{pmatrix} \alpha & \\ & \delta \end{pmatrix}$. If $A \neq 0$, then again (via P) we may assume $\alpha \neq 0$, and finally (via Q) that $\alpha = 1$. This gives the remaining types in the table. As for equivalence, these cannot be equivalent to independent pairs, so we only have to examine equivalence between $\Pi = \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$ and $\Pi' = \begin{pmatrix} 1 & \\ & \delta' \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$. The condition $\Pi \cdot P = Q \cdot \Pi'$ this time gives

$$\begin{cases} b^2 + d^2\delta = \delta'(a^2 + c^2\delta) \\ ab = -cd\delta \end{cases} . \tag{3}$$

(3) is of exactly the same form as (2): we merely have to interchange a, d and replace δ by $-\delta'$, δ' by $-\delta$. It follows that Π and Π' are equivalent $\Leftrightarrow \delta, \delta'$ are in the same (possibly zero) square-class. Once more, the form of the stabilizers results immediately. $\square$

3. The general case.

Consider now an *arbitrary* pair $\Pi = (A, B)$. This decomposes uniquely as the sum $\Pi = \Pi_s + \Pi_a$ of a *symmetric* pair $\Pi_s = (A_s, B_s)$ and an *antisymmetric* pair $\Pi_a = (A_a, B_a)$. One checks at once that this decomposition commutes with the action: $((P, Q) \cdot \Pi)_s = (P, Q) \cdot \Pi_s$ and $((P, Q) \cdot \Pi)_a = (P, Q) \cdot \Pi_a$. In particular:

$$(P, Q) \text{ fixes } \Pi \Leftrightarrow \text{ it fixes each of } \Pi_s \text{ and } \Pi_a.$$

Let $\mathcal{S}$ be the set of symmetric representatives in Theorem 1. We now have:

Proposition 1. (i) *Each equivalence class contains a pair* $\Sigma + \mathrm{T}$, *where* $\Sigma \in \mathcal{S}$ *and* T *is antisymmetric. Moreover, the class determines* Σ *uniquely.* (ii) *If* $\Pi = \Sigma + \mathrm{T}$ *and* $\Pi' = \Sigma + \mathrm{T}'$ *(similarly), then* $(P, Q) \cdot \Pi = \Pi' \Leftrightarrow (P, Q)$ *stabilizes* Σ *and* $(P, Q) \cdot \mathrm{T} = \mathrm{T}'$.

We also record the following evident lemma. Henceforth let $J = \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$.

Lemma 1. *If* $\mathrm{T} = (\alpha \mathrm{J}, \beta \mathrm{J})$ *and* $\mathrm{T}' = (\alpha' \mathrm{J}, \beta' \mathrm{J})$ *are antisymmetric pairs and* $\Delta = \det P$, *then*

$$(P, Q) \cdot \mathrm{T} = \mathrm{T}' \Leftrightarrow \begin{cases} k\alpha + l\beta = \Delta\alpha' \\ m\alpha + n\beta = \Delta\beta' \end{cases} . \tag{4}$$

Prop. 1 shows that each equivalence class has an underlying type in $\mathcal{S}$, and each type is a union of equivalence classes. We now analyze these types in turn, keeping the notation established above:

1) $\Sigma = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$: $(P,Q)\cdot\Pi = \Pi'$ if and only if (P,Q) is as in line 1 of the table in Theorem 1 and (4) holds, which amounts to $\begin{cases} a\alpha = d\alpha' \\ 2c\alpha + d\beta = d\beta'. \end{cases}$

If $\alpha = 0$, then $\alpha' = 0$ and $\beta' = \beta$. Thus there is *one orbit for each* $\beta \in K$, corresponding to $\mathrm{T} = (0, \beta\mathrm{J})$. The stabilizer for each of these is all of $\mathrm{Stab}(\Sigma)$. If $\alpha \neq 0$, we may take $a = 1$, $d = \alpha$, $c = -\beta/2$ to get $\alpha' = 1$, $\beta' = 0$, resulting in *one more orbit* given by $\mathrm{T} = (\mathrm{J}, 0)$. The stabilizer is given by the equations $a = d$, $c = 0$, hence consists of the pairs $(P,Q) = (aI, a^2I)$.

2) $\Sigma = \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$: Let $O_{2,\lambda} = \left\{ \begin{pmatrix} x & \mp\lambda y \\ y & \pm x \end{pmatrix} : x^2 + \lambda y^2 = 1 \right\}$ be the orthogonal group of the quadratic form $(1, \lambda)$. The form of (P,Q) shows that $Q/\Delta \in O_{2,-\delta}$ and Equations (4) say that Q/Δ sends (α,β) to (α',β'). Hence these vectors have the same length with respect to the form $(1,-\delta)$, in other words $\alpha^2 - \delta\beta^2 = \alpha'^2 - \delta\beta'^2$.

Conversely, let (α,β) and (α',β') be non-zero vectors satisfying this condition. Then by Witt's Extension Theorem (cf. [**4**, Prop. 3]) there exists $R = \begin{pmatrix} x & \pm\delta y \\ y & \pm x \end{pmatrix}$ in $O_{2,-\delta}$ (so that $x^2 - \delta y^2 = 1$) sending (α,β) to (α',β'). We can now choose a, c such that $R = Q/\Delta$. Namely, if $x \neq \mp 1$, let $a = \delta^{-1}(1 \pm x)$, $c = \pm\delta^{-1}y$ and if $x = \mp 1$, let $a = 0$, $c = 1$. Now (4) holds, so Π and Π' are equivalent.

Thus, apart from the symmetric class (given by $\mathrm{T} = (0,0)$), there is *one orbit for each element of* K *represented* (*non-trivially*) *by the form* $\alpha^2 - \delta\beta^2$, corresponding to $\mathrm{T} = (\alpha\mathrm{J}, \beta\mathrm{J})$.

The stabilizers are easily found from (4), with $\alpha' = \alpha$ $\beta' = \beta$.

If $P = \begin{pmatrix} a & \delta c \\ c & a \end{pmatrix}$, this condition becomes $\begin{cases} c(c\alpha + a\beta) = 0 \\ c(a\alpha + \delta c\beta) = 0 \end{cases}$, which reduces to $c = 0$, P being nonsingular. Thus $(P,Q) = (aI, a^2I)$.

If $P = \begin{pmatrix} a & -\delta c \\ c & -a \end{pmatrix}$, it amounts to $a\alpha = \delta c\beta$, so that $(a,c) = \mu(\delta\beta,\alpha)$ $(\mu \neq 0)$. The only other condition which must be met is that $\Delta \neq 0$, or equivalently $\alpha^2 - \delta\beta^2 \neq 0$. Provided this is so, the stabilizer contains elements of this second type, namely $(P,Q) = \mu\begin{pmatrix} \delta\beta & -\delta\alpha \\ \alpha & -\delta\beta \end{pmatrix}, \delta\mu^2\begin{pmatrix} \alpha^2 + \delta\beta^2 & -2\delta\alpha\beta \\ 2\alpha\beta & -(\alpha^2 + \delta\beta^2) \end{pmatrix}$. Otherwise such elements do not arise.

3) $\Sigma = \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$: Taking $P = I$ and Q arbitrary shows that in addition to the symmetric class there is just *one orbit* with $\mathrm{T} \neq 0$. We may, for example, take $\mathrm{T} = (\mathrm{J}, 0)$. The stabilizer then consists of all pairs $(P, Q) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \Delta & l \\ 0 & n \end{pmatrix}$.

4) $\Sigma = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$: Here (4) becomes $\begin{cases} a^2\alpha + l\beta = ad\alpha' \\ n\beta = ad\beta' \end{cases}$, which implies that $\beta = 0 \Leftrightarrow \beta' = 0$. As well as the symmetric class we have the cases:

(i) $\beta' \neq 0$: This is equivalent to the case $(\alpha, \beta) = (0, 1)$ as follows by taking $a = d = 1$, $l = \alpha' - \alpha$, $n = \beta'$. So we get *one orbit* corresponding to $\mathrm{T} = (0, \mathrm{J})$. The stabilizer consists of the pairs $(P, Q) = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ 0 & ad \end{pmatrix}$.

(ii) $\alpha' \neq 0$, $\beta' = 0$: This is equivalent to $(\alpha, \beta) = (1, 0)$ (take $a = \alpha'$, $d = 1$), and there is again *one orbit*, given by $\mathrm{T} = (\mathrm{J}, 0)$. The stabilizer consists of the pairs $(P, Q) = \begin{pmatrix} a & 0 \\ c & a \end{pmatrix}, \begin{pmatrix} a^2 & l \\ 0 & n \end{pmatrix}$.

5) $\Sigma = \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$: Now (4) is $\begin{cases} (a^2 + \delta c^2)\alpha + l\beta = \pm(a^2 + \delta c^2)\alpha' \\ n\beta = \pm(a^2 + \delta c^2)\beta' \end{cases}$, leading again to $\beta = 0 \Leftrightarrow \beta' = 0$. Apart from the symmetric class we must consider:

(i) $\beta' \neq 0$: As before, this reduces to *one orbit*, given by $\mathrm{T} = (0, \mathrm{J})$. The stabilizer is the set of all $(P, Q) = \begin{pmatrix} a & \mp\delta c \\ c & \pm a \end{pmatrix}, \begin{pmatrix} a^2 + \delta c^2 & 0 \\ 0 & \pm(a^2 + \delta c^2) \end{pmatrix}$.

(ii) $\alpha' \neq 0$, $\beta' = 0$: It follows that $\alpha = \pm\alpha'$, and thus that the distinct orbits are given by $\mathrm{T} = (\alpha\mathrm{J}, 0)$, α running over $K^*/\{\pm 1\}$. To calculate the stabilizers we put $\alpha = \alpha'$, $\beta = \beta' = 0$ in the equations above. This forces the sign to be $+$, and hence the stabilizer in the set of $(P, Q) = \begin{pmatrix} a & -\delta c \\ c & a \end{pmatrix}$, $\begin{pmatrix} a^2 + \delta c^2 & l \\ 0 & n \end{pmatrix}$.

We collect our results in the next theorem. Since we have dealt already with the symmetric classes in Theorem 1, we confine ourselves to the rest:

Theorem 2. *The following table gives a complete set of representatives for the orbits of G on $X - S$ (the* non-symmetric *classes), together with their stabilizers:*

	Representative	Stabilizing elements (P, Q)
1a.	$\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1+\beta \\ 1-\beta & \end{pmatrix}$ $(\beta \in K^*)$	$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ 2ac & ad \end{pmatrix}$
1b.	$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$	$\begin{pmatrix} a & \\ & a \end{pmatrix}, \begin{pmatrix} a^2 & \\ & a^2 \end{pmatrix}$
2a.	$\begin{pmatrix} 1 & \alpha \\ -\alpha & \delta \end{pmatrix}, \begin{pmatrix} & 1+\beta \\ 1-\beta & \end{pmatrix}$ in 1-1 correspondence with the values in K represented by $\alpha^2 - \delta\beta^2$, for each $\delta \in K^*/K^{*2}$	$\begin{pmatrix} a & \\ & a \end{pmatrix}, \begin{pmatrix} a^2 & \\ & a^2 \end{pmatrix}$ if $\alpha^2 - \delta\beta^2 = 0$ Otherwise, the above pairs plus: $\mu \begin{pmatrix} \delta\beta & -\delta\alpha \\ \alpha & -\delta\beta \end{pmatrix}$, $\delta\mu^2 \begin{pmatrix} \alpha^2 + \delta\beta^2 & -2\delta\alpha\beta \\ 2\alpha\beta & -(\alpha^2 + \delta\beta^2) \end{pmatrix}$
3a.	$\begin{pmatrix} & 1 \\ -1 & \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} \Delta & l \\ 0 & n \end{pmatrix}$
4a.	$\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ 0 & ad \end{pmatrix}$
4b.	$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ c & a \end{pmatrix}, \begin{pmatrix} a^2 & l \\ 0 & n \end{pmatrix}$
5a.	$\begin{pmatrix} 1 & \\ & \delta \end{pmatrix}, \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}$	$\begin{pmatrix} a & \mp\delta c \\ c & \pm a \end{pmatrix}$, $\begin{pmatrix} a^2 + \delta c^2 & 0 \\ 0 & \pm(a^2 + \delta c^2) \end{pmatrix}$
5b.	$\begin{pmatrix} 1 & \alpha \\ -\alpha & \delta \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$ $(\alpha \in K^*/\{\pm 1\})$	$\begin{pmatrix} a & -\delta c \\ c & a \end{pmatrix}, \begin{pmatrix} a^2 + \delta c^2 & l \\ 0 & n \end{pmatrix}$

By inspection from Theorems 1 and 2 we also have:

Corollary 1. *The orbits of G on Y (the* linearly independent *classes) are given by lines* 1, 2, 1a, 1b, 2a, 4a *and* 5a.

4. Finite Fields.

We now specialize the foregoing to the finite field $K = \mathbf{F}_q$ (q odd). In this case $|G| = q^2(q-1)^2(q^2-1)^2$, $|X| = q^8$, $|Y| = q(q^3-1)(q^4-1)$ and $|S| = q^6$. There are two square-classes in K^*, represented by 1 and a fixed non-square ε. Over $\mathbf{F}_q$ quadratic forms of rank ≥ 2 are universal (cf. [**5**] for example), so that for each of $\delta = 1$, ε the form $\alpha^2 - \delta\beta^2$ takes all values in K^*. In addition, when $\delta = 1$ it represents 0, but not when $\delta = \varepsilon$. Let χ denote the quadratic character of K.

From the previous results we can now easily determine the number of equivalence classes and their sizes:

Theorem 3. *The following table gives, for each type of representative, the sizes of the stabilizer and equivalence class and the number of classes:*

Rep.	|Stabilizer|	|Class|	Number of classes
1	$q(q-1)^2$	$q(q^2-1)^2$	1
2	$2(q-1)(q-\chi(\delta))$	$\frac{1}{2}q^2(q-1)(q^2-1)(q+\chi(\delta))$	2
3	$\|G\|$	1	1
4	$q^2(q-1)^3$	$(q+1)(q^2-1)$	1
5	$2q(q-1)^2(q-\chi(-\delta))$	$\frac{1}{2}q(q^2-1)(q+\chi(-\delta))$	2
1a	$q(q-1)^2$	$q(q^2-1)^2$	$q-1$
1b	$q-1$	$q^2(q-1)(q^2-1)^2$	1

2a	$\begin{cases} q-1 & \text{if } \alpha^2-\delta\beta^2=0 \\ 2(q-1) & \text{if not} \end{cases}$	$\begin{cases} q^2(q-1)(q^2-1)^2 \\ \frac{1}{2}q^2(q-1)(q^2-1)^2 \end{cases}$	$\begin{cases} 1 \\ 2(q-1) \end{cases}$
3a	$q^2(q-1)^2(q^2-1)$	q^2-1	1
4a	$q(q-1)^2$	$q(q^2-1)^2$	1
4b	$q^2(q-1)^2$	$(q^2-1)^2$	1
5a	$2(q-1)(q-\chi(-\delta))$	$\frac{1}{2}q^2(q-1)(q^2-1)(q+\chi(-\delta))$	2
5b	$q(q-1)^2(q-\chi(-\delta))$	$q(q^2-1)(q+\chi(-\delta))$	$q-1$

In all there are $4q+10$ *classes, of which* 7 *are symmetric. For the linearly independent pairs, the number of classes is* $3q+5$, *and* 3 *of these are symmetric.*

Proof. It is only necessary to observe, for lines 2, 5, 5a and 5b, that if $\xi \in K^*$ then the number of solutions of $\alpha^2-\xi\beta^2 \neq 0$ is $(q-1)(q-\chi(\xi))$. Note also in line 5b that there are $\frac{1}{2}(q-1)$ classes for each of $\delta = 1, \varepsilon$. □

As a check on the arithmetic, one readily verifies that the sum of all the class sizes is $q^8 = |X|$. For the symmetric classes the sum is $q^6 = |S|$.

Corollary 2. *For the finite local rings of the Introduction, there are* $3q+5$ *isomorphism classes* (q *odd*). *Of these,* 3 *are commutative.*

References

[1] C.J. Chikunji, *On the Classification of Finite Rings*, Ph. D. Thesis, University of Reading, 1996.

[2] B. Corbas, *Rings with few zero divisors*, Math. Ann., **181** (1969), 1-7.

[3] ______, *Finite rings in which the product of any two zero divisors is zero*, Archiv der Math., **XXI** (1970), 466-469.

[4] B. Corbas and G.D. Williams, *Congruence classes in* $M_3(\mathbf{F}_q)$ (q *odd*), Preprint, University of Reading, 1997.

[5] N. Jacobson, *Basic Algebra* I, Freeman, San Francisco, 1974.

Received June 6, 1997.

University of Reading
Reading RG6 6AX
England

University of Reading
Reading RG6 6AX
England
E-mail address: G.D.Williams@reading.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

CONGRUENCE OF TWO-DIMENSIONAL SUBSPACES IN $M_2(K)$ (CHARACTERISTIC 2)

B. Corbas and G.D. Williams

Two-dimensional subspaces of $M_2(K)$ are here classified up to congruence, K being any field of characteristic 2. This complements the authors' earlier solution of the problem over fields of characteristic $\neq 2$. There is again a corresponding conclusion for the structure of a certain class of local rings.

1. Introduction.

This is a sequel to [**2**] and shares motivation, notation and preliminaries with that paper. To recall, the problem is to classify ordered pairs (A, B) of (2×2) matrices over a field K up to *equivalence*, where (A, B) is equivalent to (A', B') if there exist invertible matrices P and $Q = (q_{ij})$ such that

$$\begin{cases} P^t AP = q_{11}A' + q_{12}B' \\ P^t BP = q_{21}A' + q_{22}B' \end{cases}.$$

If $\langle A, B\rangle$ is the subspace of $M_2(K)$ spanned by A and B, we may equally speak of $\langle A, B\rangle$ and $\langle A', B'\rangle$ being *congruent* via P. Recall also that if X is the set of all the pairs (A, B), then GL_2 acts on the right on X by $(A, B).P = (P^t AP, P^t BP)$ and on the left by $Q\cdot(A, B) = (q_{11}A+q_{12}B, q_{21}A+q_{22}B)$, and that thereby $G = GL_2 \times GL_2$ acts on the left via $(P, Q) \cdot (A, B) = Q \cdot (A, B) \cdot P^{-1}$.

One motivation for this problem is that its solution enables us to classify up to isomorphism a certain naturally arising class of local rings, namely those of the form $R = K \oplus J$, where the Jacobson radical J is such that $J^3 = 0$ and both J/J^2 and J^2 have dimension two over K. Such rings have been considered by, among others, Chikunji [**1**], at least when K is finite.

In the companion paper to the present we have solved the classification problem over any field of characteristic $\neq 2$, and we turn our attention here to the case where char $K = 2$, which we assume henceforth. Our earlier strategy (that of splitting a pair into the sum of a symmetric and an antisymmetric pair) is thus not viable anymore and we have to follow a different approach. We remark that we deal here with an arbitrary field. If we confine ourselves to finite or, more generally, perfect fields then a number of subtleties in the ensuing discussion disappear, and the treatment is correspondingly shorter. In the final section we apply our results to the case $K = \mathbf{F}_q$ (q even).

We dispense first with the simple case in which A and B are *linearly dependent*, or in other words $\dim\langle A, B\rangle \leq 1$. Here we may take $A = 0$ and it is clear that $(0, B)$ is equivalent to $(0, B')$ if and only if there exist $P \in GL_2(K)$ and $\lambda \in K^*$ such that $P^tBP = \lambda B'$. We shall say that B and B' are *projectively congruent* in this case.

Here and throughout we will write $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\Delta = \det P$. If $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, then $P^tAP = \begin{pmatrix} a^2\alpha + ac(\beta+\gamma) + c^2\delta & ab\alpha + ad\beta + bc\gamma + cd\delta \\ ab\alpha + ad\gamma + bc\beta + cd\delta & b^2\alpha + bd(\beta+\gamma) + d^2\delta \end{pmatrix}$. Let $D_{\text{off}}(A) = \beta - \gamma$ be the difference (or sum) of the off-diagonal entries in A. The following evident lemma will often be used in the sequel:

Lemma 1. $\det(P^tAP) = \Delta^2 \det(A)$ *and* $D_{\text{off}}(P^tAP) = \Delta D_{\text{off}}(A)$.

We now classify matrices up to projective congruence:

Proposition 1. *The distinct projective congruence classes are represented by:* $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$, $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & \\ & \xi \end{pmatrix}$ *and* $\begin{pmatrix} 1 & 1 \\ & \delta \end{pmatrix}$ *with* $\xi \in K^*/K^{*2}$ *and* $\delta \in K$.

[We use $\xi \in K^*/K^{*2}$ to mean that ξ runs over a complete set of representatives for the cosets of K^{*2} in K^*, and will use similar abbreviations throughout.]

Proof. If the bilinear form represented by A is alternating, then A is congruent to $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$ or $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$. Otherwise, by congruence, we may assume that A is of the form $\begin{pmatrix} \alpha & \beta \\ & \delta \end{pmatrix}$ with $\alpha \neq 0$. Since we are only interested in projective congruence we may take $\alpha = 1$. But if $\beta \neq 0$, then A is now congruent to $\begin{pmatrix} 1 & 1 \\ & \delta/\beta^2 \end{pmatrix}$. So we may assume that $A = \begin{pmatrix} 1 & \beta \\ & \delta \end{pmatrix}$, with $\beta = 0$ or 1. These two cases are not equivalent, since one is symmetric and the other not.

Consider now $P^t \begin{pmatrix} 1 & \beta \\ & \delta \end{pmatrix} P = \lambda \begin{pmatrix} 1 & \beta \\ & \delta' \end{pmatrix}$. From Lemma 1, $\Delta^2\delta = \lambda^2\delta'$, so that δ and δ' are in the same multiplicative square-class, and moreover $\Delta\beta = \lambda\beta$. If $\beta = 1$, then $\Delta = \lambda$ and so $\delta = \delta'$. If $\beta = 0$ and conversely δ, δ' are in the same square-class, say $\delta' = d^2\delta$ ($d \neq 0$), then $\begin{pmatrix} 1 & \\ & \delta \end{pmatrix}$ is congruent to $\begin{pmatrix} 1 & \\ & \delta' \end{pmatrix}$ via $P = \begin{pmatrix} 1 & \\ & d \end{pmatrix}$. $\square$

We turn now to the main case of pairs (A, B) with A and B linearly independent. Although there is not the same necessity as in [**2**] to deal with the case of A and B symmetric first, we still find it convenient to do so here, since there are some differences in the detail of how we treat the symmetric and asymmetric cases.

2. The Symmetric Case.

In this section we classify linearly independent pairs (A, B) in which both matrices are symmetric. Observe first that such a pair is equivalent to one in which $A = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ or $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}$. For if $W = \langle A, B \rangle$ equals the space of diagonal matrices, then it contains the identity, and this is congruent over $\mathbf{F}_2$ to $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}$, via $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. If not, then since $\dim W = 2$ and the space of all symmetric matrices has dimension 3, W is spanned by a diagonal matrix D and a non-diagonal matrix $\begin{pmatrix} \alpha & \beta \\ \beta & \delta \end{pmatrix}$. Since $D \neq 0$ we may assume that one of α, δ is zero and indeed, by congruence, that $\alpha = 0$. Now we may take $\beta = 1$. But $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}$ is projectively congruent to $\begin{pmatrix} & 1 \\ 1 & \delta \end{pmatrix}$ (for any $\delta \neq 0$), via $P = \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}$. So we may assume that $\delta = 0$ or 1, establishing our claim.

In order to classify symmetric pairs we need two groups. The first is the usual additive subgroup of K in characteristic 2, namely $\Gamma = \{x^2 + x : x \in K\}$. The other is the group $\mathcal{K}$ of all bijective functions $K \to K$, $x \mapsto \lambda x + \mu$ where μ belongs to the subfield F of squares in K and $\lambda \in F^*$. Thus $\mathcal{K}$ acts naturally on K. It is, of course, the semidirect product of F^* by F.

We now classify the symmetric pairs, or in other words the orbits of G on the subset S of X consisting of symmetric pairs.

Theorem 1. *The following table gives a complete set of representatives for the orbits of G on S:*

1. $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}$ $(\delta \in \mathcal{K}\backslash K)$	2. $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}$ $(\delta \in K/\Gamma)$
3. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 0 & \\ & 0 \end{pmatrix}$	4. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$
5. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$	6. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 1 & \\ & \xi \end{pmatrix}$ $(\xi \in K^*/K^{*2})$

Proof. The dependent pairs 3) to 6) have been dealt with in Section 1. For the independent pairs there are two cases to consider, as explained above.

1) $A = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$: We may take $B = \begin{pmatrix} \alpha & \\ & \delta \end{pmatrix}$. At least one of α, δ is non-zero. Via P we may assume that $\alpha \neq 0$, and then that $\alpha = 1$. Consider now equivalence between $\Pi = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}$ and $\Pi' = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & \\ & \delta' \end{pmatrix}$. Suppose δ and δ' are in the same $\mathcal{K}$-orbit, say $\delta' = x^2 + y^2\delta$ $(y \neq 0)$. With $P = \begin{pmatrix} 1 & x \\ & y \end{pmatrix}$ and $Q = \begin{pmatrix} y & \\ x & 1 \end{pmatrix}$ one immediately checks that $\Pi \cdot P = Q \cdot \Pi'$, whence the pairs are equivalent.

Conversely, if the pairs are equivalent via $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $Q = \begin{pmatrix} k & l \\ m & n \end{pmatrix}$ (which notation we will take as standard henceforth), then in particular $P^t \begin{pmatrix} 1 & \\ & \delta \end{pmatrix} P = \begin{pmatrix} n & m \\ m & n\delta' \end{pmatrix}$ and taking determinants shows that $\Delta^2\delta = m^2 + n^2\delta'$. Note that $n \neq 0$, for example by Prop. 1, and thus δ, δ' are in the same $\mathcal{K}$-orbit.

2) $A = \begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}$: As before, we may take $B = \begin{pmatrix} \alpha & \\ & \delta \end{pmatrix}$ and further assume that $\alpha \neq 0$ and then that $\alpha = 1$. For $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}$ is equivalent to $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 0 & \\ & 1 \end{pmatrix}$ and hence to $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$, which appears above. Consider equivalence between $\Pi = \begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \\ & \delta \end{pmatrix}$ and $\Pi' = \begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix},$

$\begin{pmatrix} 1 & \\ & \delta' \end{pmatrix}$. Suppose that $\delta' - \delta = x^2 + x \in \Gamma$. Taking $P = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}$ and $Q = \begin{pmatrix} 1 & \\ x & 1 \end{pmatrix}$, we find that $\Pi \cdot P = Q \cdot \Pi'$.

If, conversely, Π and Π' are equivalent via a general P and Q, then

$$\begin{aligned} (1) \qquad & \begin{pmatrix} c^2 & \Delta + cd \\ \Delta + cd & d^2 \end{pmatrix}, \begin{pmatrix} a^2 + c^2\delta & ab + cd\delta \\ ab + cd\delta & b^2 + d^2\delta \end{pmatrix} \\ & = \begin{pmatrix} l & k \\ k & k + l\delta' \end{pmatrix}, \begin{pmatrix} n & m \\ m & m + n\delta' \end{pmatrix}, \end{aligned}$$

from which we deduce the equations $\begin{cases} \Delta + cd + d^2 = \delta' c^2 \\ ab + cd\delta + b^2 + d^2\delta = \delta'(a^2 + c^2\delta) \end{cases}$.

If $c = 0$, the first equation gives $a = d$, and then the second implies that

$$(2) \qquad \delta' - \delta = x^2 + x \in \Gamma, \qquad \text{with } x = b/a.$$

If $c \neq 0$, then

$$(3) \qquad \delta' = \frac{\Delta + cd + d^2}{c^2}.$$

By symmetry, replacing P by $P^{-1} = \frac{1}{\Delta}\begin{pmatrix} d & b \\ c & a \end{pmatrix}$, we also have

$$\delta = \frac{\Delta^{-1} + ac\Delta^{-2} + a^2\Delta^{-2}}{c^2\Delta^{-2}} = \frac{\Delta + ac + a^2}{c^2}.$$

Thus

$$(4) \qquad \delta' - \delta = y^2 + y, \quad \text{with } y = \frac{a + d}{c}.$$

Hence again $\delta' - \delta \in \Gamma$.

Finally, types 1) and 2) above do not overlap, since clearly the space $\left\langle \begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \\ & \delta \end{pmatrix} \right\rangle$ contains no non-zero alternating matrix, whereas type 1) spaces do. □

3. The Asymmetric Case.

We now consider the orbits of G on the set $X - S$ of asymmetric pairs. If such a pair (A, B) is linearly independent, then the two-dimensional space $\langle A, B \rangle$ has non-trivial intersection with the three-dimensional space of symmetric matrices in $M_2(K)$. Hence we may assume that A is *symmetric* and B is *not*. Equivalence of such pairs clearly determines A up to projective congruence and so we may take A to be $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$, $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ or $\begin{pmatrix} 1 & \\ & \xi \end{pmatrix}$ $(\xi \in K^*/K^{*2})$. These cases do not overlap and we will treat them in turn. When it comes to the last type, however, it turns out that the analysis proceeds in a more streamlined way if we replace $\begin{pmatrix} 1 & \\ & \xi \end{pmatrix}$ by the matrix $\begin{pmatrix} \eta & 1 \\ 1 & 1 \end{pmatrix}$, with $\eta = 1+\xi$, to which it is congruent via $P = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. If $\begin{pmatrix} 1 & \\ & \xi' \end{pmatrix}$ similarly corresponds to $\begin{pmatrix} \eta' & 1 \\ 1 & 1 \end{pmatrix}$, then by Prop. 1 $\begin{pmatrix} 1 & \\ & \xi \end{pmatrix}$ is projectively congruent to $\begin{pmatrix} 1 & \\ & \xi' \end{pmatrix}$ $\Leftrightarrow$ there exists $\nu \neq 0$ such that $\xi' = \nu^2\xi$, or equivalently $\eta' = (1+\nu^2) + \nu^2\eta$. If $\mathcal{H}$ denotes the subgroup of the group $\mathcal{K}$ of Section 1 consisting of just those functions for which $\mu = 1+\lambda$, we have proved:

Proposition 2. *$\begin{pmatrix} 1 & \\ & \xi \end{pmatrix}$ is congruent to $\begin{pmatrix} \eta & 1 \\ 1 & 1 \end{pmatrix}$, with $\eta = 1+\xi$. Moreover $\begin{pmatrix} \eta & 1 \\ 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} \eta' & 1 \\ 1 & 1 \end{pmatrix}$ are projectively congruent if and only if η and η' are in the same $\mathcal{H}$-orbit.*

In addition to the various groups already introduced we need one more class of additive subgroups of K. Namely, for any t, $\eta \in K$ let $x_\lambda = \frac{\lambda(t+1)+t+\eta}{\lambda^2+\eta}$ and let $\mathcal{R}(t,\eta) = \{x_\lambda : \lambda \in K,\ \lambda^2 \neq \eta\} \cup \{0\}$. If $\lambda \neq \mu$, one easily verifies that $x_\lambda + x_\mu = x_\nu$ with $\nu = \frac{\lambda\mu+\eta}{\lambda+\mu}$, and it follows that $\mathcal{R}(t,\eta)$ is an additive subgroup of K. We will assume that $\eta \neq 1$.

Remark. For $t = 1$ we have

$$\mathcal{R}(1,\eta) = \left\{ \frac{1+\eta}{\lambda^2+\eta} : \lambda^2 \neq \eta \right\} \cup \{0\} = \left\{ \frac{\xi}{\mu^2+\xi} : \mu^2 \neq \xi \right\} \cup \{0\},$$

where $\xi = 1+\eta$. For $t \neq 1$ we may describe $\mathcal{R}(t,\eta)$ in an alternative, more natural way as follows. Namely, it is the inverse image under the additive homomorphism $K \to K$, $a \mapsto a^2 + a$ of the subgroup $\theta\Gamma$ of K, where $\theta = \frac{1+t^2}{1+\eta}$. To see this, the reader is invited to verify firstly that if $y_\lambda = \frac{(\lambda+t)(1+\eta)}{(1+t)(\lambda^2+\eta)}$ then $x_\lambda^2 + x_\lambda = \theta(y_\lambda^2 + y_\lambda)$, and secondly that if $x^2 + x = \theta(y^2+y)$ (some x, y), then $x = x_\lambda$ where $\lambda = \frac{xt(1+\eta)+y(t+\eta)(1+t)}{x(1+\eta)+y(1+t^2)}$. We shall not need this in the sequel.

We are now in a position to classify the asymmetric pairs.

Theorem 2. *The following table gives a complete set of representatives for the orbits of G on $X - S$:*

7. $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 0 \\ 1 & \end{pmatrix}$	8. $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ \gamma & \end{pmatrix}$ $(\gamma \neq 1)$
9. $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ & 1 \end{pmatrix}$	10. $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & \delta \end{pmatrix}$ $(\delta \in K/\Gamma)$
11. $\begin{pmatrix} \eta & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} t+\eta\delta & 1 \\ & \delta \end{pmatrix}$ $(\eta \in \mathcal{H}\backslash(K-\{1\}),\ t \in K,\ \delta \in K/\mathcal{R}(t,\eta))$	
12. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & \delta \end{pmatrix}$ $(\delta \in K)$	

Proof. The dependent pairs 12) have been discussed in Section 1, so we now consider independent pairs $\Pi = (A, B)$ in which A is symmetric and B is not. If $\Pi' = (A, B')$ is a second such pair (with the same A) and $\Pi \cdot P = Q \cdot \Pi'$, then, with the usual notation, $P^tAP = kA + lB'$ and $P^tBP = mA + nB'$. Since A is symmetric and B' is not, we have $l = 0$. Moreover, from Lemma 1 we have the following facts:

$$k = \Delta \quad \textit{if } A \textit{ is nonsingular} \quad \text{and} \tag{5}$$

$$\Delta D_{\text{off}}(B) = nD_{\text{off}}(B'). \tag{6}$$

We now analyze the three types for A alluded to at the start of the section.

7)-9) $A = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$: We may take $B = \begin{pmatrix} 0 & \beta \\ \gamma & \delta \end{pmatrix}$. There are two cases: (a) $\delta = 0$, (b) $\delta = 1$.

(a) If $B = \begin{pmatrix} & \beta \\ \gamma & \end{pmatrix}$, $B' = \begin{pmatrix} & \beta' \\ \gamma' & \end{pmatrix}$ and $\Pi \cdot P = Q \cdot \Pi'$, then $P^tAP = kA$ implies that $b = 0$. Then $P^tBP = mA + nB'$ gives $\begin{pmatrix} ac(\beta+\gamma) & ad\beta \\ ad\gamma & 0 \end{pmatrix} = \begin{pmatrix} m & n\beta' \\ n\gamma' & 0 \end{pmatrix}$ and in particular:

$$\begin{cases} ad\beta = n\beta' \\ ad\gamma = n\gamma' \end{cases}. \tag{7}$$

Since ad and n are non-zero, $\beta = 0 \Leftrightarrow \beta' = 0$. If $\beta = 0$, we may assume that $\gamma = 1$. If not, we may assume that $\beta = 1$. If now $\beta = \beta' = 1$, then (7) gives $\gamma = \gamma'$.

(b) Any pair $\begin{pmatrix}1 & \\ & 0\end{pmatrix}, \begin{pmatrix} & \beta \\ \gamma & 1\end{pmatrix}$ is equivalent to $\begin{pmatrix}1 & \\ & 0\end{pmatrix}, \begin{pmatrix}0 & 1 \\ & 1\end{pmatrix}$ via $P = \begin{pmatrix}1 & 0 \\ \gamma & \beta+\gamma\end{pmatrix}$, $Q = \begin{pmatrix}1 & 0 \\ \beta\gamma & \beta^2+\gamma^2\end{pmatrix}$.

Finally, cases (a) and (b) do not overlap. For if the P-transform of this last pair equals the Q-transform of a type (a) pair, for a general P, Q, it follows easily that $b = d = 0$, contradicting nonsingularity of P. This deals with types 7)-9) in the table.

10) $A = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$: We may clearly take $B = \begin{pmatrix}\alpha & 1 \\ & \delta\end{pmatrix}$. Moreover, we may assume $\alpha = 1$. For if $\alpha \neq 0$, (A, B) is equivalent to $\left(A, \begin{pmatrix}1 & 1 \\ & \alpha\delta\end{pmatrix}\right)$ via $P = \begin{pmatrix}1 & \\ & \alpha\end{pmatrix}$, $Q = \begin{pmatrix}\alpha & \\ & \alpha\end{pmatrix}$; and if $\alpha = 0$, it is equivalent to $\left(A, \begin{pmatrix}1 & 1 \\ & 0\end{pmatrix}\right)$ via $P = \begin{pmatrix}1+\delta & 1 \\ 1 & 0\end{pmatrix}$, $Q = \begin{pmatrix}1 & 0 \\ 1 & 1\end{pmatrix}$.

Consider now equivalence between $\Pi = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix}1 & 1 \\ & \delta\end{pmatrix}$ and $\Pi' = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix}1 & 1 \\ & \delta'\end{pmatrix}$. If $\delta' - \delta = x^2 + x \in \Gamma$, then $\Pi \cdot P = Q \cdot \Pi'$, where $P = \begin{pmatrix}1 & x \\ 0 & 1\end{pmatrix}$ and $Q = \begin{pmatrix}1 & 0 \\ x & 1\end{pmatrix}$. If, conversely, Π is equivalent to Π' by means of a general P, Q, then from (6) $\Delta = n$. But also $\det(P^tBP) = \det(mA + nB')$, giving $\Delta^2\delta = n^2\delta' + m^2 + mn$, and then

$$\delta' - \delta = x^2 + x, \quad \text{with } x = m/n. \tag{8}$$

We have accounted for type 10).

11) $A = \begin{pmatrix}\eta & 1 \\ 1 & 1\end{pmatrix}$: Here $\eta \neq 1$ is fixed and we may again take $B = \begin{pmatrix}\alpha & 1 \\ & \delta\end{pmatrix}$. Suppose that for some P, Q we have $\Pi \cdot P = Q \cdot \Pi'$, where $\Pi = (A, B)$, $\Pi' = (A, B')$ and $B' = \begin{pmatrix}\alpha' & 1 \\ & \delta'\end{pmatrix}$. From (5) and (6) we have $k = n = \Delta$. Moreover $P^tAP = kA$, so that $P^tA = kAP^{-1}$, i.e. $\begin{pmatrix}\eta a + c & a + c \\ \eta b + d & b + d\end{pmatrix} = \begin{pmatrix}\eta d + c & \eta b + a \\ c + d & a + b\end{pmatrix}$. Thus $d = a$, $c = \eta b$ and $P = \begin{pmatrix}a & b \\ \eta b & a\end{pmatrix}$. Now $P^tBP = mA + nB'$ becomes:

$$\begin{pmatrix}a^2\alpha + \eta ab + \eta^2 b^2\delta & ab\alpha + a^2 + \eta ab\delta \\ ab\alpha + \eta b^2 + \eta ab\delta & b^2\alpha + ab + a^2\delta\end{pmatrix} = \begin{pmatrix}m\eta + n\alpha' & m + n \\ m & m + n\delta'\end{pmatrix}. \tag{9}$$

Adding the $(1,1)$-entry to η times the $(2,2)$-entry gives $\Delta(\alpha+\eta\delta) = n(\alpha' + \eta\delta')$ and, since $\Delta = n$, $\alpha+\eta\delta = \alpha'+\eta\delta'$. Thus $t = \alpha+\eta\delta$ is an *invariant.*

Adding the bottom entries in (9) and using $n = a^2+\eta b^2$ leads to $n(\delta'-\delta) = ab\alpha+\eta b^2+\eta ab\delta+b^2\alpha+ab+\eta b^2\delta = ab(t+1)+b^2(t+\eta)$. Either $b=0$, and then $\delta'=\delta$, or $b\neq 0$, in which case $\delta'-\delta = \frac{\lambda(t+1)+t+\eta}{\lambda^2+\eta}$, where $\lambda = a/b$. This is the typical element x_λ of the group $\mathcal{R}(t,\eta)$ defined earlier. Conversely, if $\alpha+\eta\delta = \alpha'+\eta\delta' = t$ and $\delta'-\delta = x_\lambda$ (some λ), then Π and Π' are equivalent via $P = \begin{pmatrix}\lambda & 1\\ \eta & \lambda\end{pmatrix}$, $Q = \begin{pmatrix}\Delta & 0\\ \lambda t+\eta & \Delta\end{pmatrix}$. This deals with type 11) and the theorem is proved. $\square$

4. Finite Fields.

In this section we apply our results to the finite field $K = \mathbf{F}_q$ (q even) and determine the number of equivalence classes and the class sizes. We begin by determining the stabilizers of the various representatives 1)-12) of Theorems 1, 2. Since all elements of K are squares, it follows that $\mathcal{K}$ acts transitively on K and $\mathcal{H}$ acts transitively on $K-\{1\}$. Thus in 1) we need only consider $\delta = 0$, and in 11) we may take $\eta = 0$. In 6) we take $\xi = 1$. The homomorphism $K \to K$, $x \mapsto x^2+x$ has kernel $\{0,1\}$, so that Γ is of index two in K and in 2) and 10) we may take δ to be 0 or a fixed element $\varepsilon \notin \Gamma$. As for 11), we have $\mathcal{R}(0,0) = \mathcal{R}(1,0) = K$. If $t \neq 0,1$ then $\mathcal{R}(t,0) = \{tx^2+(t+1)x : x \in K\}$ has index two in K and we take δ to be 0 or a fixed element $\varepsilon_t \notin \mathcal{R}(t,0)$. Equivalently $tX^2+(t+1)X+\varepsilon_t$ is a fixed irreducible polynomial.

We now have:

Theorem 3. *The stabilizers of the various representatives are given as follows:*

Representative	Stabilizing elements (P,Q)
1. $\begin{pmatrix} & 1\\ 1 & \end{pmatrix}, \begin{pmatrix}1 & \\ & 0\end{pmatrix}$	$\begin{pmatrix}a & 0\\ c & d\end{pmatrix}, \begin{pmatrix}ad & \\ & a^2\end{pmatrix}$
2. $\begin{pmatrix} & 1\\ 1 & 1\end{pmatrix}, \begin{pmatrix}1 & \\ & \delta\end{pmatrix}$ $(\delta = 0, \varepsilon)$	$\begin{pmatrix}a & a+c\delta\\ c & a\end{pmatrix}, \begin{pmatrix}a^2+c^2\delta & c^2\\ a^2 & a^2+c^2\delta\end{pmatrix}$ and $\begin{pmatrix}a & c\delta\\ c & a+c\end{pmatrix}, \begin{pmatrix}a^2+c^2+c^2\delta & c^2\\ c^2\delta & a^2+c^2\delta\end{pmatrix}$
3. $\begin{pmatrix}0 & \\ & 0\end{pmatrix}, \begin{pmatrix}0 & \\ & 0\end{pmatrix}$	All

4. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} k & 0 \\ m & \Delta \end{pmatrix}$
5. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} k & 0 \\ m & a^2 \end{pmatrix}$
6. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$	$\begin{pmatrix} a & c \\ c & a \end{pmatrix}, \begin{pmatrix} k & 0 \\ m & \Delta \end{pmatrix}$
7. $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 0 \\ 1 & \end{pmatrix}$	$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ ac & ad \end{pmatrix}$
8. $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} & 1 \\ \gamma & \end{pmatrix}$ $(\gamma \neq 1)$	$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ ac(1+\gamma) & ad \end{pmatrix}$
9. $\begin{pmatrix} 1 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ & 1 \end{pmatrix}$	$\begin{pmatrix} a & \\ & a \end{pmatrix}, \begin{pmatrix} a^2 & \\ & a^2 \end{pmatrix}$
10. $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & \delta \end{pmatrix}$ $(\delta = 0, \varepsilon)$	$\begin{pmatrix} a & a + c\delta \\ c & a \end{pmatrix}, \begin{pmatrix} \Delta & 0 \\ \Delta & \Delta \end{pmatrix}$ and $\begin{pmatrix} a & c\delta \\ c & a + c \end{pmatrix}, \begin{pmatrix} \Delta & 0 \\ 0 & \Delta \end{pmatrix}$
11a. $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} t & 1 \\ & 0 \end{pmatrix}$ $(t = 0, 1)$	$\begin{pmatrix} a & \\ & a \end{pmatrix}, \begin{pmatrix} a^2 & \\ & a^2 \end{pmatrix}$
11b. $\begin{pmatrix} & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} t & 1 \\ & \delta \end{pmatrix}$ $(t \neq 0, 1;\ \delta = 0, \varepsilon_t)$	$\begin{pmatrix} a & a(1+t^{-1}) \\ 0 & a \end{pmatrix}, \begin{pmatrix} a^2 & 0 \\ a^2(t+1) & a^2 \end{pmatrix}$ and the above
12. $\begin{pmatrix} 0 & \\ & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ & \delta \end{pmatrix}$ $(\delta \in K)$	$\begin{pmatrix} a & c\delta \\ c & a + c \end{pmatrix}, \begin{pmatrix} k & 0 \\ m & \Delta \end{pmatrix}$

Proof. Most of this is very easily extracted from the proofs of Theorems 1, 2. Only lines 2, 6, 10 and 12 require explanation, and we deal with these in turn.

2) If $\Pi \cdot P = Q \cdot \Pi$, then Equation (1) in the proof of Theorem 1 holds, with $\delta = \delta'$. There are two cases:

i) $c = 0$: Then $a = d$ and from (2) $x = 0$ or 1, i.e. $b = 0$ or a.

ii) $c \neq 0$: From (4) $y = 0$ or 1, i.e. $d = a$ or $a + c$. If $d = a$, then (3) gives $\delta = \frac{a+b}{c}$, so that $b = a + c\delta$. If $d = a + c$, (3) gives $\delta = \frac{b}{c}$ and so $b = c\delta$.

Thus in either case $P = \begin{pmatrix} a & a + c\delta \\ c & a \end{pmatrix}$ or $\begin{pmatrix} a & c\delta \\ c & a + c \end{pmatrix}$, and the form of Q results at once from (1).

6) The equation $P^tBP = nB$ here gives $P^t = nP^{-1}$. By determinants we have $n = \Delta$, and so $\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} d & b \\ c & a \end{pmatrix}$, whence $d = a$, $b = c$.

10) From (8) in the proof of Theorem 2 we have $x = 0$ or 1, and thus $P^tB = (mA + nB)P^{-1}$, with $m = 0$ or n. Recall that $n = \Delta$. We deal with the cases:

i) $m = 0$: The matrix equation becomes $\begin{pmatrix} a & a + c\delta \\ b & b + d\delta \end{pmatrix} = \begin{pmatrix} d + c & b + a \\ c\delta & a\delta \end{pmatrix}$, and so $b = c\delta$, $d = a + c$.

ii) $m = n$: This time $\begin{pmatrix} a & a + c\delta \\ b & b + d\delta \end{pmatrix} = \begin{pmatrix} d & b \\ d + c\delta & b + a\delta \end{pmatrix}$, and thus $b = a + c\delta$, $d = a$.

12) The reasoning is similar to the previous two cases and we omit it. □

We may now determine the number of equivalence classes over $\mathbf{F}_q$ and their sizes. Note that $|G| = q^2(q-1)^2(q^2-1)^2$, $|X| = q^8$ and $|S| = q^6$. It is convenient to introduce the additive character $\psi : \mathbf{F}_q \to \{\pm 1\}$ given by $\psi(x) = 1$ $(x \in \Gamma)$, -1 $(x \notin \Gamma)$.

Theorem 4. *The following table gives, for each type of representative, the sizes of the stabilizer and equivalence class and the number of classes:*

Rep.	|Stabilizer|	|Class|	Number of classes
1	$q(q-1)^2$	$q(q^2-1)^2$	1
2	$2(q-1)(q-\psi(\delta))$	$\frac{1}{2}q^2(q-1)(q^2-1)(q+\psi(\delta))$	2
3	$\|G\|$	1	1
4	$q^2(q-1)^2(q^2-1)$	q^2-1	1
5	$q^2(q-1)^3$	$(q+1)(q^2-1)$	1
6	$q^2(q-1)^2$	$(q^2-1)^2$	1
7	$q(q-1)^2$	$q(q^2-1)^2$	1
8	$q(q-1)^2$	$q(q^2-1)^2$	$q-1$
9	$q-1$	$q^2(q-1)(q^2-1)^2$	1
10	$2(q-1)(q-\psi(\delta))$	$\frac{1}{2}q^2(q-1)(q^2-1)(q+\psi(\delta))$	2
11a	$q-1$	$q^2(q-1)(q^2-1)^2$	2
11b	$2(q-1)$	$\frac{1}{2}q^2(q-1)(q^2-1)^2$	$2q-4$
12	$q(q-1)^2(q-\psi(\delta))$	$q(q^2-1)(q+\psi(\delta))$	q

In all there are $4q+8$ *classes, of which* 7 *are symmetric. For the linearly independent pairs, the number of classes is* $3q+4$*, and* 3 *of these are symmetric.*

Proof. We count the number of solutions (a, c) of the equation $a^2+ac+c^2\delta = 0$, for a given δ. If $c = 0$, then $a = 0$. If $c \neq 0$, then $\delta = \left(\frac{a}{c}\right)^2 + \frac{a}{c}$ and there are two solutions for a if $\delta \in \Gamma$, and none otherwise. So the total number of solutions is $1 + 2(q-1) = 2q-1$ $(\delta \in \Gamma)$, 1 $(\delta \notin \Gamma)$.

It follows that the number of solutions of $a^2+ac+c^2\delta \neq 0$ is $q^2-2q+1 = (q-1)^2$ $(\delta \in \Gamma)$, q^2-1 $(\delta \notin \Gamma)$. In either case this equals $(q-1)(q-\psi(\delta))$. This explains lines 2, 10 and 12, and the rest is clear. □

Note that a simple check shows that the sum of all the class sizes is $q^8 = |X|$, as should be, and that the sum of the symmetric class sizes is $q^6 = |S|$.

Corollary. *For the finite local rings of the Introduction, there are* $3q+4$ *isomorphism classes* (q *even*)*. Of these,* 3 *are commutative.*

This agrees with the result for $q = 2$ in [**1**], found by computer search.

References

[1] C.J. Chikunji, *On the Classification of Finite Rings*, Ph.D. Thesis, University of Reading, 1996.

[2] B. Corbas and G.D. Williams, *Congruence of Two-Dimensional Subspaces in* $M_2(K)$ (*characteristic* $\neq 2$), Pacific J. of Math., **188**(**2**) (1999), 225-235.

Received August 11, 1997.

University of Reading
Reading RG6 6AX
England

University of Reading
Reading RG6 6AX
England
E-mail address: G.D.Williams@reading.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

ON THE TQFT REPRESENTATIONS OF THE MAPPING CLASS GROUPS

Louis Funar

We prove that the image of the mapping class group by the representations arising in the $SU(2)$-TQFT is infinite, provided that the genus $g \geq 2$ and the level of the theory $r \neq 2, 3, 4, 6$ (and $r \neq 10$ for $g = 2$). In particular it follows that the quotient groups $\mathcal{M}_g/N(t^r)$ by the normalizer of the r-th power of a Dehn twist t are infinite if $g \geq 3$ and $r \neq 2, 3, 4, 6, 8, 12$.

1. Introduction.

Witten [**50**] constructed a TQFT in dimension 3 using path integrals and afterwards several rigorous constructions arose, like those using the quantum group approach ([**39**, **25**]), the Temperley-Lieb algebra ([**30**, **31**]), the theory based on the Kauffman bracket ([**4**, **5**]) or that obtained from the mapping class group representations and the conformal field theory ([**27**]).

Any TQFT gives rise to a tower of representations of the mapping class groups $\mathcal{M}_g$ in all genera g and this tower determines in fact the theory, up to the choice of the vacuum vector (see [**10**, **5**, **48**]). The aim of this paper is to answer whether the image of the mapping class groups is finite or not under such representations.

There is some evidence supporting the finiteness of this image group. First, in the Abelian $U(1)$-theory the representations can be identified with the monodromy of a system of theta functions. The latter is explicitly computed (see e.g. [**11**, **15**]) and it is easy to see that it factors through a finite extension (due to the projective ambiguity) of $Sp(2g, \mathbf{Z}/r\mathbf{Z})$, where r is the level of the theory. For a general Lie group G, the monodromy associated to the genus 1 surfaces may also be determined (see [**18**, **11**]) using some formulas of Kac (see [**21**, **22**]) and again it factors through a finite extension of $Sp(2g, \mathbf{Z}/r\mathbf{Z})$, where r is now the shifted level. This has already been suggested by the fact that the Reshetikhin-Turaev invariant for lens spaces $L_{a,b}$ ($a \leq b$ running over the positive integers) takes only a finite number of distinct values, namely for the cosets mod r of a and b (see for instance [**29**]). For low levels $r = 4, 6$ the whole tower of representations was described by Blanchet and Masbaum in [**6**] and by Wright ([**51**, **52**]) and in particular the images are finite groups. On the other hand all TQFTs are associated to conformal field theories (abb. CFT) in dimension 2 (see e.g. [**10**, **50**]) and the finiteness question appeared also in the context of classification of rational conformal field theories. For instance in [**36**] one asks whether the algebraic CFT have finite monodromy (which is equivalent to our problem for some classes of TQFTs, as for example the $SU(2)$-TQFT). Some of the irreducible representations of $SL(2, \mathbf{Z}/r\mathbf{Z})$ which admit extensions as monodromies of some CFT in all genera were discussed in [**8**, **9**]. Also in [**24**] the action of $SL(2, \mathbf{Z})$ on the conformal blocks was computed for all quantum doubles, and it could be proved that the image is always finite. Gilmer obtained in another way (see [**14**]) the finiteness of the image for $g = 1$, in the $SU(2)$ theory, result which seems to be known in the conformal field theory community, and noticed also by M. Kontsevich. Meantime Stanev and Todorov [**41**] have a partial answer to this question in the case of the 4-punctured sphere, as we will explain below.

This is the motivation for our main result:

Theorem 1.1. *The image $\rho(\mathcal{M}_g)$ of the mapping class group $\mathcal{M}_g$ under the representation ρ arising in the $SU(2)$-TQFT (in both the BHMV and RT versions) and respectively $SO(3)$-TQFT is infinite provided that $g \geq 2$, $r \neq 2, 3, 4, 6$, and if $g = 2$ also $r \neq 10$.*

Let us mention that ρ is only a projective representation and thus its image is well defined up to scalar multiplication by roots of unity of order $4r$. To explain briefly what means the two versions BHMV and RT we recall that the $SU(2)$-TQFT was constructed either using the Kauffman bracket - and this is the BHMV version from [**5**] - or else using the coloured Jones polynomial - and this is the RT version from [**39**, **25**]. The invariants obtained for closed 3-manifolds are "almost" the same, but their TQFT extensions are different. That is the reason for considering here both of them, though as it should be very unlikely that the mapping class group representations do not share the same properties, in the two related cases.

Before we proceed let us outline the relationship with the results from [**41**], where the Schwarz problem is considered for the $\widehat{su(2)}$ Knizhnik-Zamolodchikov equation. The authors determined whether the image of the mapping class group of the 4-punctured sphere is finite, thereby solving a particular case of our problem, however in a slightly different context. It would remain to identify the following two representations of the mapping class groups (in arbitrary genus):

- one is that arising from the conformal field theory based on the $\widehat{su(2)}$ Knizhnik-Zamolodchikov equation. Tsuchiya, Ueno and Yamada [**43**] constructed the CFT using tools from algebraic geometry, for all Riemann surfaces.
- the other one is that arising in the RT-version of the $SU(2)$-TQFT.

There are some naturally induced representations of braid groups in both approaches, which can be proved to be the same by the explicit computations of Tsuchyia and Kanie [**42**].

Presumably the two representations of the mapping class groups are also equivalent, but a complete proof of this fact does not exist, on author's knowledge. First it should be established that the CFT extends to a TQFT in 3 dimensions, which is equivalent to control the behaviour of conformal blocks sheaves over the compactification divisor on the moduli space of curves. Observe that a different and direct construction of the associated TQFT can be given ([**27**, **28**, **12**]) if we assume the CFT has all the properties claimed by the physicists. Notice that a complete solution of that problem would furnish an entirely geometric description of the TQFT following Witten's prescriptions, in which the mapping class group representation is the monodromy of a projectively flat connection on some vector bundle of non-abelian theta functions over the Teichmuller space.

Thus we cannot deduce directly from [**41**] the finiteness of the mapping class group representation without assuming the previous unproved claim. Our purpose is to use instead the BHMV approach which has a simple and firmly established construction. We consider braid group representations using basically the monodromy of the holed spheres. The data we obtain is similar to that obtained by all the other means, hence also to that from [**41**, **42**]. Specifically, the idea of the proof of the main theorem is to identify a certain subspace of the space on which $\mathcal{M}_g$ acts, which is invariant to the action of a subgroup of $\mathcal{M}_g$, the last being a quotient of a pure braid group P_n, $n \geq 3$. Next we observe that the action of P_n extends naturally to an action of the whole braid group B_n, and this it turns to factor through the Hecke algebra $H_n(q)$ of type A_{n-1} at a root of unity q. This was inspired by the computations done by Tsuchyia and Kanie ([**42**], see also [**41**]) of the monodromy in the conformal field theory on $\mathbf{P}^1$. Now an easy modification of the Jones theorem ([**20**]), about the generic infiniteness of the image of B_n in Hecke algebra representations, will settle our question.

For fixing the notations, we denote by r the level, which is supposed to be in this sequel the order of the roots of unity which appear in the definition of the invariants for the RT-version and respectively a quarter (or half) of it for the BHMV-version for $SU(2)$ (respectively $SO(3)$).

The groups $\mathcal{M}_g/N(t^r)$, quotients of $\mathcal{M}_g$ by the normalizer of a power of a Dehn twist, were previously considered for $r = 2, 3$ by Humphries in [**17**], and it was shown that these are finite groups for $r = 2$ and arbitrary g, and infinite for $g = 2$ and $r \geq 3$. This solved the problem 28 asked by Birman in [**2**], p. 219. We derive a generalization of that, to all other genera g, namely:

Corollary 1.2. *The quotient groups $\mathcal{M}_g/N(t^r)$ are infinite for $g \geq 3$, $r \neq 2, 3, 4, 6, 8, 12$.*

Proof. It is well-known (see [**35**], p. 379) that the image of a Dehn twist $\rho(t)$, in some nice basis, is a diagonal matrix whose entries are $(-1)^j A^{j^2+2j}$, where A is a $2r$-th root of unity. It follows that, for odd r, $\rho(t)^r$, and for even r, $\rho(t)^{2r}$ respectively, is a scalar matrix in this particular basis, and furthermore it is a scalar matrix in any other basis. Therefore, the image group $\rho(\mathcal{M}_g)$(modulo multiplication by roots of unity of order $4r$) is a quotient of $\mathcal{M}_g/N(t^r)$ (and respectively $\mathcal{M}_g/N(t^{2r})$), and now the claim follows. □

Notice that the proof given by Humphries used the Jones representation [**19**] of $\mathcal{M}_2$ which arises as follows: The group $\mathcal{M}_2$ is viewed as a quotient of the braid group B_6, and some Hecke algebra representation factors throughout $\mathcal{M}_2$. For $g > 2$ it is only a proper subgroup of $\mathcal{M}_g$ which is a quotient of B_{2g+2}, so that an extension of the Jones representation to higher genus is not obvious.

It seems that not only the representations have infinite image, but the set of values the $SU(2)$-invariant (at a given level r) takes on the closed 3-manifolds of fixed Heegaard genus g, is also infinite. Our result does not imply this stronger statement, because the infinite image we found comes from a subgroup of $K \subset \mathcal{M}(F)$ of homeomorphisms of the surface extending to the handlebody. In fact when we twist the gluing map of a Heegaard splitting by an element of K we obtain a manifold homeomorphic to the former one. However it is very likely that the same method could be refined to yield this stronger statement.

We think that our theorem holds also for the case $g = 2, r = 10$ and the corollary is true more generally for $g \geq 3, r \geq 3$. The same ideas can be used for the $SU(N)$-TQFT to show that, in general, the corresponding representations of $\mathcal{M}_g$ have infinite images.

Acknowledgements. This work was done during author's visit at Columbia University, whose hospitality is gratefully acknowledged. We are thankful to Roland Bacher, Joan Birman, Christian Blanchet, Razvan Gelca, Patrick Gilmer, Gregor Masbaum, Jerome Los, Vlad Sergiescu, Gretchen Wright for their suggestions and comments and to the referee for pointing out some errors in the previous version and simplifying the proof.

2. Preliminaries.

2.1. Hecke algebras and Temperley-Lieb algebras.

We will outline briefly, for the sake of completeness, some basic notions concerning the Hecke algebras (see [**49**] for more details). Recall that the Hecke algebra of type A_{n-1} is the algebra over $\mathbf{C}$ generated by $1, g_1, ..., g_{n-1}$ and the following relations:

$$g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1},\ i = 1, 2, ..., n-2,$$

$$g_ig_j = g_jg_i, \mid i-j \mid > 1,$$

$$g_i^2 = (q-1)g_i + q, \ i = 1, 2, ..., n-1,$$

where $q \in \mathbf{C} - \{0\}$ is a complex parameter. Denote this algebra by $H_n(q)$. It is known (see e.g. [**7**], p. 54-55) that $H_n(q)$ is isomorphic to the group algebra $\mathbf{C}S_n$ of the symmetric group S_n, provided that q is not a root of unity.

Notice that $H_n(q)$ is the quotient of the group algebra CB_n of the braid group B_n. The braid group is usually presented as generated by $g_1, ..., g_{n-1}$, together with the first two relations from above. In particular there is a natural representation of B_n in $H_n(q)$.

From the quadratic relation satisfied by g_i it follows that g_i has at most two spectral values. For $q \neq -1$ set f_i for the spectral projection corresponding to the eigenvalue -1; then $g_i = q - (1+q)f_i$, and another presentation of $H_n(q)$ can be obtained in terms of the generators $1, f_1, ..., f_{n-1}$, as follows:

$$f_if_{i+1}f_i - q(1+q)^{-2}f_i = f_{i+1}f_if_{i+1} - q(1+q)^{-2}f_{i+1}, \ i = 1, 2, ..., n-2,$$

$$f_if_j = f_jf_i, \mid i-j \mid > 1,$$

$$f_i^2 = f_i, \ i = 1, 2, ..., n-1.$$

The irreducible representations of Hecke algebras are well-known in the case when $H_n(q)$ are semi-simple, which means that q is not a root of unity. The structure of $H_n(q)$ at roots of unity is more complicated (see for instance [**49**]) and as we will be concerned with this situation precisely we introduce also some smaller quotients (after [**16**, **20**]), namely the Temperley-Lieb algebras.

The algebras $A_{\beta,n}$ (following the convention from [**16**], Section 2.8) are generated over $\mathbf{C}$ by $1, e_1, ..., e_{n-1}$ and the relations:

$$e_ie_{i+1}e_i = e_ie_{i-1}e_i = \beta^{-1}e_i, \ i = 1, 2, ..., n-2,$$

$$e_ie_j = e_je_i, \mid i-j \mid > 1,$$

$$e_i^2 = e_i, \ i = 1, 2, ..., n-1.$$

Remark that $A_{\beta,n}$ is a quotient of the Hecke algebra $H_n(q)$, for $\beta = 2 + q + q^{-1}$. In fact, the image of the projector f_i is $1 - e_i$. If we replace $e_i = 1 - f_i$ we find a presentation of $A_{\beta,n}$ as the quotient of $H_n(q)$ by adding the set of relations (see also Prop. 2.11.1 from [**16**], p. 123):

$$g_ig_{i+1}g_i + g_ig_{i+1} + g_{i+1}g_i + g_i + g_{i+1} + 1 = 0, \ \text{for } i = 1, n-1.$$

It is known that the algebra $A_{\beta,3}$ is semi-simple for $\beta \neq 1$ (or equivalently, $r \neq 3$) and $A_{\beta,4}$ is semi-simple for $\beta \neq 1, 2$ (equivalently $r \neq 3, 4$). Moreover, when semi-simple these are multi-matrix algebras: $A_{\beta,3} = M_2(\mathbf{C}) \oplus \mathbf{C}$, and $A_{\beta,4} = M_3(\mathbf{C}) \oplus M_2(\mathbf{C}) \oplus \mathbf{C}$, (see Theorem 2.8.5, p. 98 from [**16**]).

2.2. Mapping class group representations.

Most of the material presented here comes from [**30**, **40**, **34**]. Let A be a fixed complex number and M be a compact oriented 3-manifold. The skein module $S(M)$ is the vector space generated by the isotopy classes (rel ∂M) of framed links, quotiented by the (skein) relations from Figure 1.

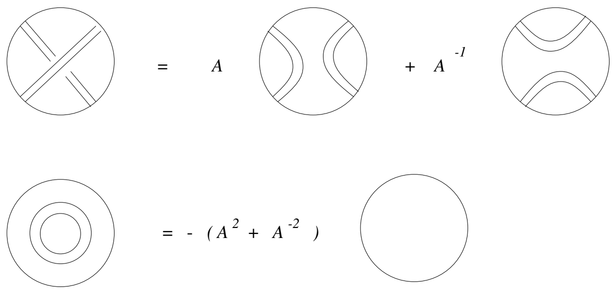


Figure 1. Skein relations.

For example $S(S^3)$ is one dimensional (as a module over $\mathbf{Z}[A, A^{-1}]$), with basis the empty link; the image of the framed link $L \subset S^3$ in $S(S^3)$ is the value of the Kauffman bracket evaluated at A. Notice that, in order to construct the TQFT we must specialize A to be a primitive $2r$-th root of unity. For even r we obtain the $SU(2)$-TQFT (level $\frac{r}{2}$ with our convention) and for odd r we obtain the $SO(3)$-TQFT (of level r this time).

The skein space for the 3-ball with $2n$ boundary (framed) points has an algebra structure, by representing the framed link in a planar projection sitting into a rectangle, and separating the points into two groups of n on opposite sides. The multiplication is given by the juxtaposition of diagrams, and the algebra thus obtained can be identified with the the Temperley-Lieb algebra $A_{\beta,n}$, for a suitable β. The generators for $A_{\beta,n}$ are the elements $1_n, e_1, ..., e_{n-1}$ pictured in Figure 2.

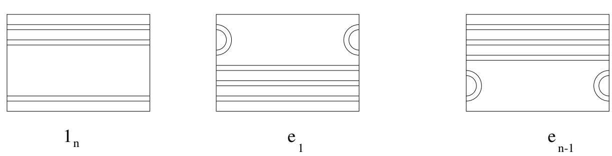


Figure 2. Generators of the Temperley-Lieb algebra.

Now the Jones-Wenzl idempotents $f^{(n)} \in A_{\beta,n}$ are uniquely determined by the conditions $f^{(n)2} = f^{(n)}$, $f^{(n)}e_i = e_i f^{(n)} = 0$, for $i = 1, 2, ..., n-1$, whenever A is such that all $\Delta_i = (-1)^i \frac{A^{2i+1} - A^{-2i-1}}{A^2 - A^{-2}}$ for $i = 0, 1, ..., n-1$ are non-zero. This implies that $f^{(n)}x = xf^{(n)} = \lambda_x f^{(n)}$, for all x, with a suitable chosen complex number λ_x.

Denote in a planar diagram by a line labeled with n (in a small rectangle) the element $1_n \in A_{\beta,n}$, and by a line with a dash labeled n the insertion of the element $f^{(n)} \in A_{\beta,n}$. This will give a convenient description for the elements of skein modules.

One construction for the $SU(2)$-invariants via skein modules, was given in [**30**, **31**] and latter extended to a TQFT in [**5**], and to higher $SU(N)$-invariants recently in [**32**].

Let us outline first the construction of the conformal blocks, which are the vector spaces associated to surfaces via the TQFT. Decompose the sphere S^3 as the union of two handlebodies H of genus g, and H' with a small cylinder $F \times I$ over the surface $F = \partial H = \partial H'$ inserted between them. There is a map

$$\langle , \rangle : S(H) \times S(H') \longrightarrow S(S^3) = \mathbf{C},$$

induced by the Kauffman bracket and the union of links. In [**5**] it was shown that, if A is a primitive $4r$-th root of unity, then

$$W(F) = S(H)/\ker\langle , \rangle$$

is the space associated to the surface F by the $SU(2)$-TQFT at level r. Here ker denotes the left kernel of the bilinear form $\langle , \rangle$. This space has however a more concrete description. If i, j, k satisfy the following conditions:

$$0 \le i, j, k \le r-1, \ | i - j | \le k \le i + j, \ i + j + k \text{ is even},$$

then we can define an element of the skein space of the 3-ball with $i + j + k$ boundary points, given by inserting $f^{(i)}, f^{(j)}, f^{(k)}$ in the diagram, and therefore connecting up with no crossings (see Figure 3).

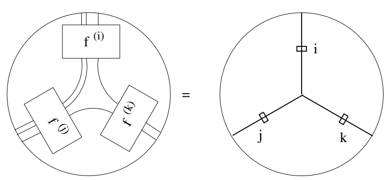


Figure 3. Vertex elements in the skein modules.

Now the triple (i, j, k) is called admissible if, additionally to the previous conditions, it satisfies $0 \le i, j, k \le r-2$ and $i+j+k \le 2(r-2)$. Furthermore let consider the standard 3-valent graph in H which is the standard spine of the handlebody H, and label its edges with integers $i_1, i_2, ..., i_{3g-3}$, such that all labels incident to a vertex form an admissible triple. We form an element of $S(H)$ by inserting idempotents $f^{(i)}$ along the edges of the graph and triodes, like we did above at vertices. It is shown in [**5**, **40**] that the vectors we obtain this way form a basis of the quotient space $W(F)$.

For a 3-valent graph Γ, possibly with leaves and some of the edges already carrying a label, we denote by $W(\Gamma)$ the space generated by the set of labelings of (non-labeled) edges which have the property that all triples from incident edges are admissible. An easy extension of the arguments in [**5**, **40**] shows that $W(\Gamma)$ is isomorphic to $W(F)$, provided that Γ is some closed 3-valent graph of genus g.

If K and K' are the subgroups of the mapping class group $\mathcal{M}(F)$ of F consisting of the classes of those homeomorphisms which extend to the handlebodies H and H' respectively, then we have natural actions of K on $S(H)$, and K' on $S(H')$. Moreover these actions descend to actions on the quotient $W(F)$. One of these two actions, say that of K on H, has a simple meaning: Consider an element $x \in S(H)$, which is a representative of the class $[x] \in W(F)$, $u \in K$, then $u(x) = [\varphi(x)] \in W(F)$, where φ is a homeomorphism of H whose restriction at F, modulo isotopy, is u. The other action, that of K' on $W(F)$ can be described in a similar manner, using the non-degenerate bilinear form $\langle , \rangle$ on $W(F)$. Namely, $u(x)$, for $u \in K'$, $x \in W(F)$ is defined by the equality:

$$\langle ux, y\rangle = \langle x, [u(y')]\rangle,$$

holding for any $y \in W(F)$; on the right hand side $y' \in S(H')$ is a lift of y, and the action of K' on $S(H')$ is the geometric one.

Moreover we have an induced action of the free group generated by K and K' on $W(F)$. It is shown in [**40**, **34**] that this action descends to a central extension of the mapping class group $\mathcal{M}(F)$. This is the representation coming from TQFT. Actually we can build up the TQFT starting from that representation. The main idea is that, if we cut a closed 3-manifold M along a (closed embedded) surface F into two pieces M_1 and M_2, then the invariant $Z(M)$ can be recovered from the invariants $Z(M_i)$ associated to M_i (which are vectors in the space $W(F)$) as follows:

$$Z(M) = \langle Z(M_1), Z(M_2)\rangle.$$

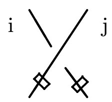
If we want to glue back now M_1 to M_2 using an additional twist $\varphi \in \mathcal{M}(F)$ then we can compute also the invariant of the resulting manifold $M_1 \cup_\varphi M_2$ using the (projective) representation $\rho : \mathcal{M}(F) \longrightarrow GL(W(F))$, defined above:

$$Z(M_1 \cup_\varphi M_2) = \langle \rho(\varphi)Z(M_1), Z(M_2)\rangle.$$

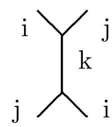
We skipped the complications arising from the projective ambiguity, which is a root of unity, which amount to consider a supplementary structure (framing) on the manifold. This gives a simple formula for the invariant in terms of Heegaard splittings. In fact the vector $Z(H) = Z(H') \in W(F)$, associated to the handlebody is corresponding to the graph of genus g whose labels are all 0 (up to a normalization factor, which we skip for simplicity). Then $Z(H \cup_\varphi H')$, the invariant of the closed manifold obtained by gluing two handlebodies along their common surface F using the homeomorphism φ, is now $\langle \rho(\varphi)Z(H), Z(H')\rangle$.

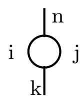
2.3. Transformation rules for planar diagrams in the skein modules.

In order to make explicit computations we will freely use the recipes from [**35**] which allows us to transform planar diagrams representing elements in the skein module of the 3-ball (with some boundary points) into simpler planar diagrams, eventually arriving to linear combinations of the elements of a fixed basis. For completeness we include these rules below.

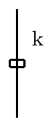


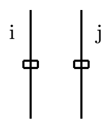
$$= \sum_k \delta(k;i,j)\frac{\langle k\rangle}{\langle i,j,k\rangle}$$



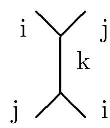


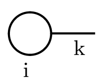
$$= \delta_n^k \frac{\langle i,j,k\rangle}{\langle k\rangle}$$



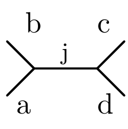


$$= \sum_k \frac{\langle k\rangle}{\langle i,j,k\rangle}$$

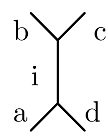


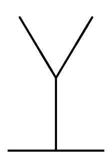


$$= 0 \qquad \text{for } k \geq 1$$

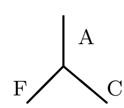


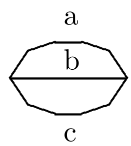
$$= \sum_i \left\{ \begin{array}{ccc} a & b & i \\ c & d & j \end{array} \right\}$$



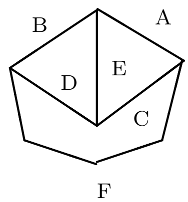


$$= \frac{\left\langle \begin{array}{ccc} A & B & E \\ D & C & F \end{array} \right\rangle}{\langle A,F,C\rangle}$$





$$= \langle a, b, c\rangle$$



$$= \left\langle \begin{array}{ccc} A & B & E \\ D & C & F \end{array} \right\rangle$$

where

$$\langle k\rangle = (-1)^k[k+1] = (-1)^k \frac{A^{2k+2} - A^{-2k-2}}{A^2 - A^{-2}},$$

$$\delta(c; a, b) = (-1)^k A^{ij-k(i+j+k+2)},$$

$$\langle a, b, c\rangle = (-1)^{i+j+k} \frac{[i+j+k+1]![i]![j]![k]!}{[i+j]![i+k]![j+k]!}.$$

Here i, j, k are the internal colors given by

$$i = \frac{b+c-a}{2},\ j = \frac{a+c-b}{2},\ k = \frac{b+a-c}{2},$$

and $[n]! = [1][2]...[n]$.

Consider now A, B, C, D, E, F such that (A, B, E), (B, D, F), (E, D, C), (A, C, F) are admissible triples and make some notations: $\Sigma = A + B + C + D + E + F$, $a_1 = \frac{A+B+E}{2}$, $a_2 = \frac{B+D+F}{2}$, $a_3 = \frac{E+D+C}{2}$, $a_4 = \frac{A+C+F}{2}$, $b_1 = \frac{\Sigma-A-F}{2}$, $b_2 = \frac{\Sigma-B-C}{2}$, $a_1 = \frac{\Sigma-A-D}{2}$.

The tetrahedron coefficient is defined as:

$$\left\langle \begin{array}{ccc} A & B & E \\ D & C & F \end{array} \right\rangle = \frac{\prod_i \prod_j [b_i - a_j]!}{[A]![B]![C]![D]![E]![F]!} \left(\begin{array}{ccccccc} a_1 & & a_2 & & a_3 & & a_4 \\ & b_1 & & b_2 & & b_3 & \end{array} \right),$$

where

$$\left(\begin{array}{ccccccc} a_1 & & a_2 & & a_3 & & a_4 \\ & b_1 & & b_2 & & b_3 & \end{array} \right) = \sum_{\max a_j \leq \zeta \leq \min a_j} \frac{(-1)^\zeta [\zeta+1]!}{\prod_i [b_i - \zeta]! \prod_j [\zeta - a_i]!}.$$

The quantum 6j-symbol of [**35**] is given by the formula:

$$\left\{ \begin{array}{ccc} a & b & i \\ c & d & j \end{array} \right\} = \frac{\langle i\rangle \left\langle \begin{array}{ccc} i & b & c \\ j & d & a \end{array} \right\rangle}{\langle i, a, d\rangle \langle i, b, c\rangle}.$$

3. Proof of the theorem.

3.1. Outline.

Consider a surface F of genus g and let $\Gamma \subset H$ be a 3-valent graph embedded in the handlebody H. Suppose that the graph Γ' shown in Figure 4 is a subgraph of Γ. Then Γ' can be viewed as the spine of the $(n+2)$-holed sphere $F' \subset F$, which is the intersection of a regular neighborhood of Γ' (in $\mathbf{R}^3$) with F. Consider a partial labeling $\Gamma'(n,m)$ of Γ' as shown on the right hand side in the figure. Notice that the leaf with label 0 can be removed without affecting the space $W(\Gamma'(n,m))$.

Lemma 3.1. *For a suitably chosen* Γ, *of genus* $g \geq 4$, *there exist* $m \geq 0, n \geq 5$, *such that* $W(\Gamma'(n,m)) \subset W(\Gamma)$ *and* $\dim W(\Gamma'(n,m)) \geq 2$. *For* $g = 3$ *we have* $W(\Gamma'(4,2)) \subset W(\Gamma)$, *and for* $g = 2$ $W(\Gamma'(3,1)) \subset W(\Gamma)$.

Proof. We connect among them the leaves of Γ' using some planar arcs, in order to obtain a closed graph of minimal genus, and such that the labels agree when making connections. $\square$

Figure 4. The graphs $\Gamma'(n,m)$.

Fix now once for all the embedding of graphs $\Gamma' \subset \Gamma$ as in the lemma, and denote by $V = V(n,m) \subset W(\Gamma)$ the image of $W(\Gamma'(n,m))$. Consider the curves $\gamma_{ij} \subset F'$, $1 \leq i,j \leq n$, drawn on the $(n+2)$-holed sphere F' which encircle the holes i and j like in Figure 5 and the set of curves γ_i which are loops around the holes. The Dehn twists $T_{\gamma_{ij}}$ and T_{γ_i} generate a subgroup S of the mapping class group $\mathcal{M}(F)$.

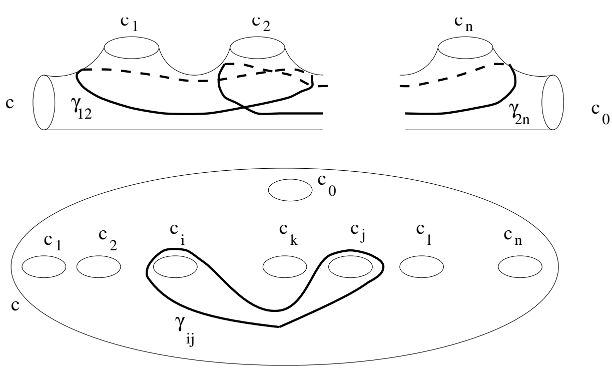


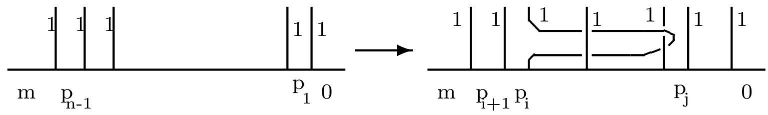
Figure 5. The curves γ_{ij}.

Proposition 3.1. *The subspace $V(n,m) \subset W(F)$ is $\rho(S)$-invariant. Moreover the image $\rho\,|_{V(3,1)}\,(S) \subset GL(V(3,1))$ is an infinite group, provided that $g \geq 2$ and the level $r \neq 2,3,4,6,10$. For $r = 10$ and $g \geq 3$ the image of $\rho\,|_{V(4,2)}\,(S) \subset GL(V(4,2))$ is infinite.*

The first part of the proposition follows from a more general fact concerning sub-surfaces $F' \subset F$ and subgroups S of $\mathcal{M}(F)$ of classes of homeomorphisms which keep F' invariant up to an isotopy, and send each boundary component into itself. Assume that a labeling of the boundary components of F' was fixed: This amounts to fix a labeling for the leaves of the subgraph Γ', the spine of F'. Then the subspace $W(\Gamma') \subset W(\Gamma)$ is invariant by the action of S on $W(F)$. Moreover, consider now that F' may be sent by a larger group S' into a sub-surface F'' which is isotopic to F, but the boundary components may be permuted among themselves. We claim now, that a subspace $W(\Gamma')$, associated to a labeling of the boundary components, is sent by such a homeomorphism into the subspace $W(\Gamma')$ associated to the permuted labeling on the boundary components. In particular the space V from the proposition is not only invariant under $\rho(S)$, but also under larger groups which could permute the n boundary components c_i, $i = 1,2,...,n$, since all their labels are identical.

Another observation is that the action of $\mathcal{M}(F')$ on the space $W(\Gamma')$, where Γ' has one external edge e (corresponding to the boundary component $c_e \subset \partial F'$) is the same as the action of $\mathcal{M}(F' \cup_{c_e} D^2)$ on the space $W(\Gamma'')$; here $F' \cup_{c_e} D^2$ is the result of gluing a disk on the circle c_e, and Γ'' is Γ with the edge e removed from it. This way we see that $\rho(S)$ acts like $\mathcal{M}(F' \cup_{c_0} D^2)$ on the given space. This will help to find out the corresponding extension to the braid group.

Before we explain this action, remark that all Dehn twists along γ_{ij}, γ_i are elements of the subgroup $K \subset \mathcal{M}(F)$ of classes of homeomorphisms extending to the handlebody H. Therefore, according to the discussion in the previous section, the action of $T_{\gamma_{ij}}$ (or T_{γ_i}) on $V(n,m)$ has a simple expression in the skein module of the 3-ball with $(n+2)$-boundary points: Just perform the Dehn twist on the 3-ball which is a regular neighborhood of the graph Γ', viewed as part of the handlebody H, whose spine is Γ. This is equivalent with twisting the i-th and j-th legs of the graph Γ'. We have to apply further the skein relations, in order to compute the latter element in terms of the given basis of V, where all the legs are straight. Notice that the representative graphs considered in Sections 2.2 and 2.3 are framed graphs, and the framings considered in the planar pictures are the blackboard ones. When the Dehn twist $T_{\gamma_{ij}}$ acts on the spine graph Γ' then the framings of the strands i and j are altered. Then the action of T_{γ_i} on the i-th strand is the change of its framing by one unit. Then the element $A_{ij} = T_{\gamma_{ij}} T_{\gamma_i}^{-1} T_{\gamma_j}^{-1}$ acts on $V(n,m)$ as follows:



Remark that the T_{γ_i}'s commute with all the other $T_{\gamma_{jk}}$ because their support curves can be made disjoint. These formulas make up a representation of the pure braid group P_n, which extends to the whole braid group in the obvious manner: Consider that the i-th and $i+1$-th legs are only half-twisted. This defines the action of the i-th generator g_i of the braid group B_n. In fact, looking at the generators A_{ij} of P_n as elements of B_n, their action on V consists in twisting the corresponding strands of Γ', modulo Reidemester moves in plane. On the other hand the fact that we obtained a representation of B_n is checked the same manner: The relation $g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1}$ translates into the third Reidemester move, which is obviously satisfied in the skein module. We continue to denote by ρ the representation of B_n on V. In fact this should enter in the computation of the action of elements in the larger group $\mathcal{M}(F)$, so actually is "part" of ρ but for bigger genus.

The main ingredient in Proof of Proposition 3.1 is:

Proposition 3.2. *The representations*

$$(-A^{-1})\rho : B_3 \longrightarrow \mathrm{End}(V(3,1))$$

and respectively

$$(-A^{-1})\rho : B_4 \longrightarrow \mathrm{End}(V(4,2))$$

factor throughout the Temperley-Lieb algebras $A_{\beta,3}$ (and respectively $A_{\beta,4}$) where $\beta = 2 + q + q^{-1}$ and the parameter $q = A^{-4}$.

Let us explain what is meant by $c\rho$ where $c \in \mathbf{C}$ is a constant. This is another representation of the braid group, which is defined on the generators by $(c\rho)(g_i) = c\rho(g_i)$. Therefore if w is a word in the generators g_i we have $(c\rho)(w) = c^{|w|}\rho(w)$, where $|w|$ is the sum of exponents appearing in the word w. Since the braid relations are homogeneous in the generators g_i this is well-defined.

We will prove that the image of the braid group is infinite in $A_{\beta,n}$. This was done by Jones in [**20**] for one value of A (and a slightly different context), but the proof extends to an arbitrary primitive root of unity. We can be more precise. Let us consider $M_2(\mathbf{C})$ as the image of $A_{\beta,3}$ by the natural projection, and the factor $M_3(\mathbf{C})$ as the quotient of $A_{\beta,4}$, when both are semi-simple.

Proposition 3.3. *The image of B_3 in $M_2(\mathbf{C})$ (via $A_{\beta,3}$) is infinite provided that $r \neq 2, 3, 4, 6, 10$. For $r = 10$ the image of B_4 in $M_3(\mathbf{C})$ (via $A_{\beta,4}$) is infinite.*

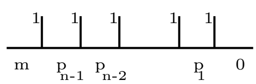
Now the algebra $A_{\beta,n}$ $(n = 3, 4)$ is semi-simple for r outside the excepted range $r \neq 3, 4$ and $A_{\beta,3} = M_2(\mathbf{C}) \oplus \mathbf{C}$. The representation $\tilde{\rho}$ is not abelian and then the induced map $A_{\beta,3} \longrightarrow M_2(\mathbf{C}) = \mathrm{End}(V(3,1))$ must be the canonical projection (up to an automorphism). Thus the image of B_3 by ρ should be infinite. For $r = 10$ we work within B_4 and $A_{\beta,4}$ and we are forced then to restrict to $g \geq 3$. Again $\tilde{\rho}$ is not abelian and we will see in the next section that it is irreducible. Therefore the induced map $A_{\beta,4} \longrightarrow M_3(\mathbf{C}) = \mathrm{End}(V(4,2))$ is again the projection on the corresponding factor. This establishes Proposition 3.1, because P_n is of finite index in B_n.

Eventually recall that ρ, at the mapping class group level, is only a projective representation, and it can be also considered as a representation of a finite extension (depending on the level) of $\mathcal{M}_g$. The image group stays then in the unitary group modulo roots of unity of order $4r$ and thus the claim of Theorem 1.1 (concerning the BHMV-version) is proved. The present proof (see Section 3.3) shows actually that the image in the projective unitary group (i.e. modulo $U(1)$) is also infinite. □

3.2. Proof of Proposition 3.2.

The TQFT considered here is the one constructed in [**5**], for A a primitive $2r$-th root of unity. We suppose for simplicity that r is even, hence we are working with the $SU(2)$-TQFT. The same arguments hold verbatim for the representation associated to the $SO(3)$-TQFT, with only minor modifications in the range of colors.

Lemma 3.2.1. *A basis for $V(n,m)$ is provided by the labeled graphs $L(\mathbf{p})$ below,*

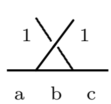


whose labels are in one-to-one correspondence with

$$B(V) = \{\mathbf{p} = (p_0, p_1, ..., p_n); p_i \in \mathbf{Z}_+,$$
$$p_0 = 0, p_n = m, p_i \leq 2r-2, |p_i - p_{i+1}| = 1, i = 0, ..., n\}.$$

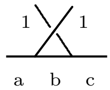
Proof. It follows immediately from the admissibility conditions on the triples $(p_i, p_{i+1}, 1)$. □

Then the computation of $\rho(g_i)$ is reduced to that of $g_i L(\mathbf{p})$, in the skein module. Observe now that the only values of the labels p_j which may change when g_i is applied are p_{i-1}, p_i, p_{i+1}. This will be also seen during the explicit computation.

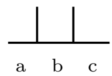
Actually we have to compute 

and according to the lemma it suffices to consider that $|a-b| = |b-c| = 1$.

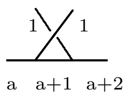
Lemma 3.2.2. *Suppose that* $|a-b| = |b-c| = 1$ *and* $|a-c| = 2$. *Then*



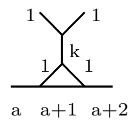
$$= A$$



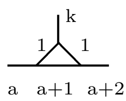
Proof. Suppose for simplicity that $c = a+2, b = a+1$. Then according to [**35**] and Section 2.3 we have



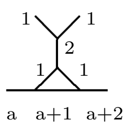
$$= \sum_k \delta(k;1,1)\frac{\langle k\rangle}{\langle 1,1,k\rangle}$$



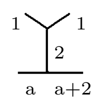
Therefore the triple $(1,1,k)$ has to be admissible, so that $k \in \{0,2\}$. Also the open tetrahedron



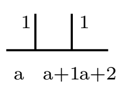
vanishes if $(k,a,a+2)$ is not admissible, so that there is only one possibility left, namely $k=2$. We get rid of the triangular face by using the formula:



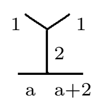
$$= \frac{\left\langle \begin{array}{ccc} 2 & 1 & 1 \\ a+1 & a+2 & a \end{array} \right\rangle}{\langle 2,a,a+2\rangle}$$



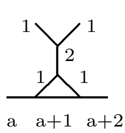
We eventually perform a fusion in order to express the right hand side in terms of the the usual basis $L(\mathbf{p})$. The formula of the fusing is:



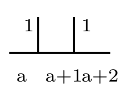
$$= \left\{ \begin{array}{ccc} a & 1 & 2 \\ 1 & a+2 & a+1 \end{array} \right\}$$



where the quantum 6-j symbol involved, namely $\left\{ \begin{array}{ccc} a & 1 & 2 \\ 1 & a+2 & a+1 \end{array} \right\}$, it turns out to be equal to 1, for all a. This implies that

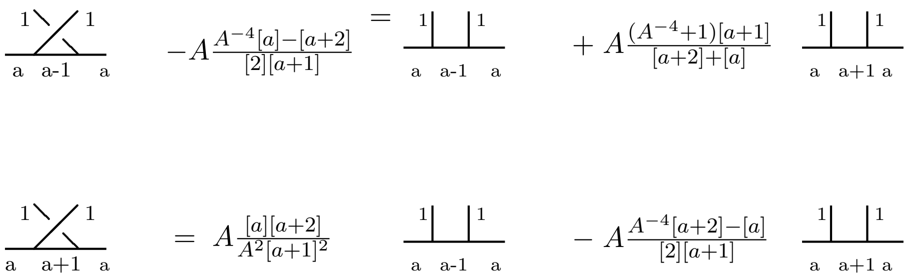


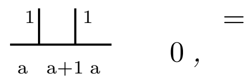
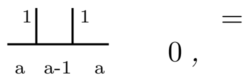
$$=$$



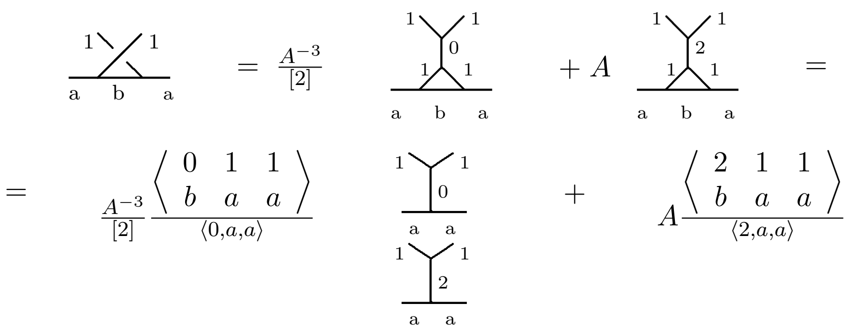
and the lemma follows. $\square$

Lemma 3.2.3. *The following identities hold:*

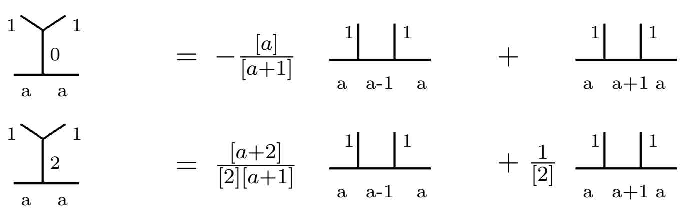


Here is to be understood that for $a = 2r-2$ we have  *and for $a = 0$ the equality*  *holds.*

Proof. We have, like in the previous lemma, the following formula:



We have to perform a fusing, in order to arrive to the standard basis of V. Using the computations of 6j-symbols appearing in this particular fusing we obtain that:



The convention is that diagrams whose labels form non-admissible triples are vanishing. Explicit computations now yield our claim. $\square$

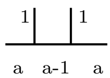
Lemma 3.2.4. *The following numerical identities are satisfied:*

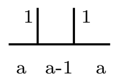
$$[a] + [a+2] = [2][a+1], \; \frac{[a+1](1+A^{-4})}{[a]+[a+2]} = \frac{1}{A^2},$$

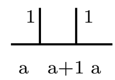
$$\frac{[a+2]A^{-4}-[a]}{[2][a+1]}=\frac{A^{-4}-1}{1-A^{4+4a}},\quad \frac{[a]A^{-4}-[a+2]}{[2][a+1]}=\frac{A^{-4}-1}{1-A^{-4-4a}}.$$

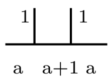
The proof is a mere computation.

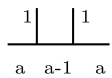
Set now $q = A^{-4}$. Then Lemma 3.2.3 can be reformulated as

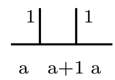
$\rho(g_i)$ 

$= -A\frac{q-1}{1-q^{a+1}}$ 

$+ \frac{1}{A}$ 

$\rho(g_i)$ 

$= \nu^2 A^3$ 

$- A\frac{q-1}{1-q^{-a-1}}$ 

where $\nu^2 = \frac{(1-q^{a+2})(1-q^{a+1})(1-q^{-a})(1-q^{-a-1})}{(1+q)^2}$.

We specify now to the cases $n = 3$, $m = 1$, and $n = 4, m = 2$. Then the vector space $V(n,m)$ has dimension 2 and respectively 3. The Hecke algebra relations follow immediately from the formulas above. The additional relation defining $A_{\beta,n}$ is verified by a direct computation.

Let us check out the case $n = 3$. The two vectors which span $V(3,1)$ have the labels $(0, p_1, p_2, m) \in \{(0,1,0,1), (0,1,2,1)\}$. Then the previous formulas read:

$$\tilde{\rho}(g_1) = \begin{pmatrix} -1 & 0 \\ 0 & A^{-4} \end{pmatrix},$$

$$\tilde{\rho}(g_2) = \begin{pmatrix} \frac{1}{A^4(1+A^4)} & -\frac{A^4+A^{-4}+1}{A^2(A^4+A^{-4}+2)} \\ -A^{-2} & -\frac{1}{1+A^{-4}} \end{pmatrix},$$

where $\tilde{\rho} = (-A^{-1})\rho$. This implies that

$$\tilde{\rho}(g_1g_2) = \begin{pmatrix} -\frac{1}{A^4(1+A^4)} & \frac{A^4+A^{-4}+1}{A^2(A^4+A^{-4}+2)} \\ -A^{-6} & -\frac{1}{A^4(1+A^{-4})} \end{pmatrix},$$
$$\tilde{\rho}(g_2g_1) = \begin{pmatrix} -\frac{1}{A^4(1+A^4)} & -\frac{A^4+A^{-4}+1}{A^6(A^4+A^{-4}+2)} \\ A^{-2} & -\frac{1}{A^4(1+A^{-4})} \end{pmatrix},$$

and

$$\tilde{\rho}(g_1g_2g_1) = \begin{pmatrix} \frac{1}{A^4(1+A^4)} & \frac{A^4+A^{-4}+1}{A^6(A^4+A^{-4}+2)} \\ A^{-6} & -\frac{1}{A^8(1+A^{-4})} \end{pmatrix}.$$

Therefore the relation $1 + \tilde{\rho}(g_1) + \tilde{\rho}(g_2) + \tilde{\rho}(g_1g_2) + \tilde{\rho}(g_2g_1) + \tilde{\rho}(g_1g_2g_1) = 0$ holds, which proves the first part of Proposition 3.1.

Let us give the explicit matrices for $n = 4$, $m = 2$ and an ad-hoc proof of the irreducibility. The space $V(4,2)$ is spanned by the vectors of labels

$(0, p_1, p_2, p_3, m) \in \{(0,1,0,1,2), (0,1,2,1,2), (0,1,2,3,2)\}$. Then within this basis we have from above:

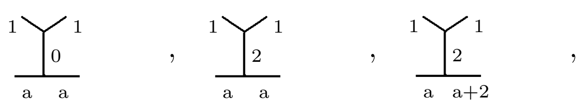
$$\tilde{\rho}(g_1) = \begin{pmatrix} A^{-4} & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$\tilde{\rho}(g_2) = \begin{pmatrix} -\frac{A}{1+A^{-4}} & -A^{-2} & 0 \\ -\frac{(1-A^{12})(1-A^4)^3}{A^{34}} & \frac{1}{1+A^8} & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$\tilde{\rho}(g_3) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{1+A^{-4}+A^{-8}} & -A^{-2} \\ 0 & -\frac{(1+A^8)^2(A^4-1)^2(A^{12}-1)^2}{A^{66}} & \frac{A^{-12}}{1+A^{-4}+A^{-8}} \end{pmatrix}.$$

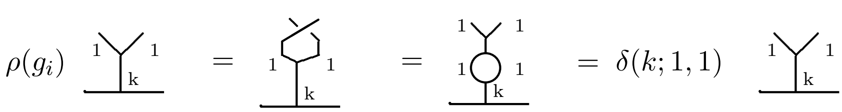
Assume that the B_4 representation $\tilde{\rho}$ is not irreducible. Then the $A_{\beta,4}$-module $V(4,2)$ is completely reducible and there is at least one simple factor of dimension 1. Equivalently, $V(4,2)$ contains a 1-dimensional B_4-invariant subspace say $\mathbf{C}w$, for some non-zero vector w. There exist then the scalars λ_i such that $\tilde{\rho}(g_i)w = \lambda_i w$. The group relations imply $\lambda_1 = \lambda_2 = \lambda_3 = \lambda$ (since the matrices above are non-singular) and from the relations in $A_{\beta,4}$ we derive $\lambda = -1$. However the condition $\tilde{\rho}(g_1)w + w = 0$ yields $w_1 = 0$, further $\tilde{\rho}(g_2)w + w = 0$ adds the constraint $w_2 = 0$ and the last identity shows that w vanishes. This proves that $V(4,2)$ is irreducible. In particular $\tilde{\rho}$ factors throughout the projection $A_{\beta,4} \longrightarrow M_3(\mathbf{C})$. □

Remarks 3.2.5. We could show from the very beginning of the proof that the representation of B_n factors through the Hecke algebra $H_n(q)$ with $q = A^{-4}$. Let assume we are interested in the action of g_j. Observe that the vectors



and the corresponding ones with $a+2$ replaced by $a-2$, (having the vertical strand on the j-th position) span all of $V(n,m)$. Indeed using the fusing matrices (which are invertible) we can relate this system to the standard basis $L(\mathbf{p})$.

But now these are precisely the eigenvectors for g_i, because we have the following relations:



This implies that the eigenvalues of $\rho(g_i)$ are $-A^{-3}$ and A, so that shifting ρ by a factor of $-A^{-1}$ will change them into -1 and A^{-4}, as in the usual presentation of $H_n(q)$ with $q = A^{-4}$.

It can be proved (but this is beyond the scope of this note) that the representation $\tilde{\rho}$ is precisely the Hecke algebra representation $\pi_\lambda^{(r)}$ associated to the Young diagram $\lambda = [\frac{n+m}{2}, \frac{n-m}{2}]$ as considered by Wenzl in [**49**]. This is clear for $n = 3, m = 1$. The direct computational approach is somewhat cumbersome for the general case. However from [**3**] we can derive a still more general equivalence between the Hecke algebras representation associated to a Young diagram λ with N rows and that arising in the previous construction for the $SU(N)$-TQFT, where the label m is replaced by the "color" λ. This follows from the irreducibility of the latter, using the technique from [**3**].

3.3. Proof of Proposition 3.3.

In [**20**] a proof for Proposition 3.3 is given for the case when $q = \exp\left(\frac{2\pi\sqrt{-1}}{r}\right)$, but the argument generalizes easily to all primitive roots of unity. We outline it below, for the sake of completeness.

It is known that $A_{\beta,3}$ is semi-simple and splits as $A_{\beta,3} = M_2(\mathbf{C}) \oplus \mathbf{C}$, for all $\beta \neq 1$ (see the Theorem 2.8.5, p. 98 from [**16**]). It suffices then to show that the images $\pi(g_1)$ and $\pi(g_2)$ in the factor $M_2(\mathbf{C})$ generate an infinite group. Observe first that $\pi(g_1)$ and $\pi(g_2)$ (and respectively $\tilde{\rho}(g_1)$ and $\tilde{\rho}(g_2)$) do not commute with each other. Thus the $A_{\beta,3}$-module $V(3,1)$ is isomorphic to the simple non-trivial factor $M_2(\mathbf{C})$. As a consequence it suffices to see what happens with the images of these two generators, when restricted to this summand. The B_3 representation on $M_2(\mathbf{C})$ is also unitarizable when q is a root of unity according to Proposition 3.2, p. 257 from [**20**]. Thus it makes sense to consider the images $\iota(g_1)$ and $\iota(g_2)$ in $SO(3) = U(2)/\mathbf{C}^*$. We have then the following decomposition in orthogonal projectors:

$$\iota(g_i) = qe_i - (1 - e_i),$$

so that the order of $\iota(g_i)$ in $SO(3)$ is $2r$ if r is odd, $r/2$ if $r = 2(4)$ and r if $r = 0(4)$, because q is a primitive root of unity of order r. As $r \neq 1$ these two elements cannot belong to a cyclic or dihedral subgroup of $SO(3)$. But no other subgroups have elements of order bigger than 5. Thus for $r = 5, 7, 8, 9$ or $r \geq 11$, the image in $M_2(\mathbf{C})$ of the subgroup generated by g_1 and g_2 is infinite.

When $r = 10$ we have to work within B_4, likewise to [**20**], p. 269. We already saw that the $A_{\beta,4}$-module $V(4,2)$ is irreducible and identified therefore with the simple factor $M_3(\mathbf{C})$ from $A_{\beta,4}$. Moreover the representation of B_4 on this factor was explicitly found out in [**19**], and it is the tensor product of the Burau and parity representations. It is also shown that the Burau representation contains elements of infinite order, for instance $g_1 g_2 g_3^{-1}$.

We may wonder whether an element of infinite order in the image can be explicitly given for $r \neq 10$. Since we have to consider only the matrices $\iota(B_n)$ in $SO(3)$, it is very likely that the element $g_1^{-1}g_2$ has infinite order.

Remark 3.1. Once we obtained the fact that the image of $\mathcal{M}_g$ is infinite at a particular primitive root of unity, we may argue also as follows: The Galois group $\mathrm{Gal}(\overline{\mathbf{Q}};\mathbf{Q})$ acts on the set of roots of unity, as well as on the entries of the matrices $\rho(x)$, with $x \in \mathcal{M}_g$. It suffices to prove that the two actions of $\mathrm{Gal}(\overline{\mathbf{Q}};\mathbf{Q})$ are compatible to each other, in order to conclude that the image group is infinite at all roots of unity. This argument was pointed to me by Gregor Masbaum.

3.4. The RT version.

Lickorish [**31**] established the relationship among the invariants obtained via the Temperley-Lieb algebra (basically those from [**4**]) $I(M, A)$ and the Reshetikhin-Turaev invariant $\tau_r(M)$ (see [**25**]), for closed oriented 3-manifolds M, as follows:

$$I\left(M, -\exp\frac{\pi\sqrt{-1}}{2r}\right) = \exp\left(\frac{(6-3r)b_1(M)\pi\sqrt{-1}}{4r}\right)\tau_r(M),$$

where $b_1(M)$ is the first Betti number of the manifold M. Roughly speaking the two invariants are the same up to a normalization factor. There are however two associated TQFTs, still very close to each other:

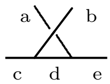
1) The TQFT based on the Kauffman bracket, as described in [**5**], which arises in a somewhat canonical way; in fact any invariant of closed 3-manifolds extends to a TQFT via this procedure (see [**5**, **10**] for details). The associated mapping class group representation we denote it by ρ^K.
2) The TQFT based on the Jones polynomial, as described in [**25**] (see also [**13**]). The associated mapping class group representation we denote it by ρ^J, and may be computed using the definitions from conformal field theory like in [**37**]. A derivation of this representation, and the reconstruction of the invariant from it was first given by Kohno [**27**] (see also [**44**, **45**, **12**]).

The two representations are similar: The associated spaces on which they act are naturally isomorphic. This means that in both theories $W(F)$ has a distinguished basis given by labelings of 3-valent graphs, with the same set of labels. Basically both theories are built up using some variants of the quantum 6j-symbols:

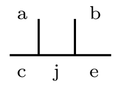
1) In [**35**] these are identified with the tetrahedron coefficients, (see also [**23**]); the relationship with the usual 6j-symbol (coming from representation theory) was outlined in [**38**].
2) In the case of ρ^J the 6j-symbols are coming from the representation theory of $U_q(sl_2)$ and where described in [**26**].

Consider now the analogous subspace $V(n,m) = W(\Gamma'(n,m))$ of $W(F)$, as in 3.1. We have again an action of the braid group B_n on V, but this time the interpretation is no longer related to skein modules of the ball. Here the graph Γ is considered to be embedded in the surface F, giving a rigid structure on F [**10**, **48**]. This means that there is a pants decomposition c of F with the property that all circles in c are transversal to Γ, the intersection of Γ with every trinion is the suspension of 3 points (topologically, the space underlying the figure Y). Remark that c and Γ determine uniquely an identification of F with a fixed and decomposed surface, up to an isotopy.

This time twisting the strands of the labeled graphs in $L(\mathbf{p})$ can be expressed in terms of the data of conformal field theory (see [**27**]). Specifically, we have:



$$= \sum_j B_{dj} \left[\begin{array}{cc} a & b \\ c & e \end{array} \right]$$



where the matrix B is the so-called braiding matrix. The braiding matrix can be expressed in terms of the fusing matrix F (see [**27**, **37**]) by the following formula:

$$B_{ij} \left[\begin{array}{cc} j_2 & j_3 \\ j_1 & j_4 \end{array} \right]$$
$$= (-1)^{j_1+j_4-i-j}/2 \exp\left(\pi\sqrt{-1}(\Delta_{j_1} + \Delta_{j_4} - \Delta_i - \Delta_j)\right) F_{ij} \left[\begin{array}{cc} j_1 & j_3 \\ j_2 & j_4 \end{array} \right],$$

where

$$\Delta_j = \frac{j(j+1)}{4r}.$$

We use the same set of labels for the graphs, namely integers running from 1 to $2r-2$ as before, instead of the traditional half-integer labels from [**26**, **27**, **23**]. Set also $q = \exp \frac{2\pi\sqrt{-1}}{r}$, and $[n] = \frac{q^n - q^{-n}}{q - q^{-1}}$.

The natural choice for the fusing matrix F is (see [**27**], p. 213-214, [**46**]):

$$F_{ij} \left[\begin{array}{cc} j_1 & j_3 \\ j_2 & j_4 \end{array} \right] = \left\{ \begin{array}{ccc} j_1 & j_3 & i \\ j_2 & j_4 & j \end{array} \right\}_{KR},$$

where $\{,\}_{KR}$ denotes the quantum 6j-symbols of Kirillov and Reshetikhin.

Using the computations from [**26**], and those from 2.3 we find that the only non-trivial braiding matrix for $j_2 = j_3 = 1$ is that with $j_1 = j_4$, and its value is therefore:

$$B \left[\begin{array}{cc} 1 & 1 \\ a & a \end{array} \right] = \left(\begin{array}{cc} -q^{a+\frac{1}{4}} \left(\frac{[a]}{[2][a+1]} \right)^{1/2} & -q^{-\frac{1}{4}} \left(\frac{[a+2]}{[2][a+1]} \right)^{1/2} \\ -q^{-\frac{1}{4}} \left(\frac{[a+2]}{[2][a+1]} \right)^{1/2} & -q^{-a-\frac{3}{4}} \left(\frac{[a]}{[2][a+1]} \right)^{1/2} \end{array} \right).$$

Notice that the braiding matrices arising in conformal field theory were previously computed by Tsuchyia and Kanie in [**42**]. Their result, used however a different normalization and the matrices are not identical, but equivalent up to a power of q. In fact, in our case, the representation $q^{1/4}\rho^J$ is also equivalent to $\tilde{\rho}$, for $n = 3, 4$. As an immediate consequence the representation ρ^J has an infinite image, too, under the same condition as ρ^K. This ends the proof of the main theorem.

References

[1] L.C. Biederharn and J.D. Louck, *Angular momentum in Quantum Physics*, Encycloped. Math. and Its Appl., no. 8, Addison-Wesley, 1981.

[2] J. Birman, *Braids, links and mapping class groups*, Ann. Math. Studies, Princeton Univ. Press, no. 82, 1974.

[3] C. Blanchet, *Hecke algebras and* 3*-manifolds quantum invariants*, Preprint, 1997.

[4] C. Blanchet, N. Habegger, G. Masbaum and P. Vogel, *Three-manifold invariants derived from the Kauffman bracket*, Topology, **31** (1992), 685-699.

[5] ______, *Topological quantum field theories derived from the Kauffman bracket*, Topology, **34** (1995), 883-927.

[6] C. Blanchet and G. Masbaum, *Topological quantum field theories for surfaces with spin structures*, Duke Math. J., **82** (1996), 229-267.

[7] N. Bourbaki, *Algèbres de Lie*, IV, V, VI, Masson, 1981.

[8] W. Eholzer, *Fusion algebras induced by representations of the modular group*, Internat. J. Modern Phys. A, **8**(**20**) (1993), 3495-3507.

[9] W. Eholzer and N.P.Skoruppa, *Conformal characters and theta series*, Lett. Math. Phys., **35** (1995), 197-211.

[10] L. Funar, 2+1-D *Topological quantum field theory and* 2-D *conformal field fheory*, Commun. Math. Phys., **171** (1995), 405-458.

[11] ______, *Thêta functions, root systems and* 3*-manifold invariants*, J. Geom. Phys., **17** (1995), 61-82.

[12] ______, TQFT *for general Lie algebras and open* 3*-manifolds*, J. Math. Sci. Univ. Tokyo, **4** (1997), 121-181.

[13] R. Gelca, *The $SL(2, C)$ Topological Quantum Field Theory with corners*, Preprint q-alg/9604006, 1996.

[14] P. Gilmer, *On the WRT representations of the mapping class groups*, Proc. A.M.S., to appear.

[15] T. Gocho, *The topological invariant of three manifolds based on the $U(1)$ gauge theory*, J. Fac. Sci. Univ. Tokyo, Sect. 1A, Math., **39** (1992), 169-184.

[16] F.M. Goodman, P. de la Harpe and V.F.R. Jones, *Coxeter graphs and tower of algebras*, MSRI Series, Berkeley, Springer, New-York, 1989.

[17] S.P. Humphries, *Normal closures of powers of Dehn twists in mapping class groups*, Glasgow J. Math., **34** (1992), 313-317.

[18] L. Jeffrey, *Chern-Simons-Witten invariants of lens spaces and torus bundles*, Commun. Math. Phys., **147** (1992), 563-604.

[19] V.F.R. Jones, *Hecke algebras and a polynomial invariant for knots*, Ann. of Math., **126** (1987), 335-388.

[20] ———, *Braid groups, Hecke algebras and type* II_1*-factors*, in 'Geometric methods in operator algebras', Proc. US-Japan Seminar, H. Araki, E.G. Effros Eds., 1986, 242-271.

[21] V. Kac, *Infinite dimensional Lie algebras*, Progress in Math., 1983.

[22] V.G. Kac and D.H. Petersen, *Infinite dimensional Lie algebras, theta functions and modular forms*, Adv. Math., **53** (1984), 125-264.

[23] L. Kauffman and S.L. Lins, *Temperley-Lieb Recoupling Theory and Invariants of* 3*-manifolds*, Ann. Math. Studies 134, Princeton Univ. Press, 1994.

[24] T. Kerler, *Mapping class group actions on quantum doubles*, Commun. Math. Phys., **168** (1995), 353-388.

[25] R. Kirby and P. Melvin, *The* 3*-manifold invariants of Witten and Reshetikhin-Turaev for* $sl(2, C)$, Invent. Math., **105** (1991), 473-545.

[26] A.N. Kirillov and N.Yu. Reshetikhin, *Representation of the algebra of* $U_q(sl_2)$*, q-orthogonal polynomials and invariants for links*, in 'Infinite dimensional Lie algebras and groups', Adv. Ser. Math. Phys., V. Kac, Editor, 1988, 285-339.

[27] T. Kohno, *Topological invariants for* 3*-manifolds using representations of mapping class groups* I, Topology, **31** (1992), 203-230.

[28] ———, *Topological invariants for* 3*-manifolds using representations of mapping class groups* II, Contemporary Math., **175** (1994), 193-217.

[29] B.H. Li and T.J. Li, *Generalized Gaussian sums and Chern-Simons-Witten-Jones invariants of lens spaces*, J. Knot Theory and Its Ramif., **5** (1996), 183-224.

[30] R. Lickorish, *Three-manifold invariants and the Temperley-Lieb algebra*, Math. Annalen., **290** (1991), 657-670.

[31] ———, *Calculations with the Temperley-Lieb algebra*, Comment. Math. Helv., **67** (1992), 571-591.

[32] ———, *Sampling the* $SU(n)$*-invariants*, J. Knot Theory and Its Ramif., **6** (1997), 45-60.

[33] D. Littlewood, *The theory of group characters*, Oxford Univ. Press, 1940.

[34] G. Masbaum and J. Roberts, *On central extensions of mapping class groups*, Math. Annalen., **302** (1995), 131-150.

[35] G. Masbaum and P. Vogel, 3*-valent graphs and the Kauffman bracket*, Pacific J. Math., **164** (1994), 361-381.

[36] S.D. Mathur and A. Sen, *Group theoretic classification of rational conformal field theories with algebraic characters*, Nucl. Phys. B, **327** (1989), 725-743.

[37] G. Moore and N. Seiberg, *Classical and quantum field theory*, Commun. Math. Phys., **123** (1989), 177-254.

[38] S. Piunikhin, *Reshetikhin-Turaev and Kontsevich-Kohno-Crane invariants for* 3*-manifolds coincide*, J. Knot Theory and its Ramif., **2** (1993), 65-95.

[39] N. Reshetikhin and V. Turaev, *Invariants of three-manifolds via link polynomials and quantum groups*, Invent. Math., **103** (1991), 547-597.

[40] J. Roberts, *Skeins and mapping class groups*, Math. Proc. Cambridge Phil. Soc., **115** (1994), 53-77.

[41] Y.S. Stanev and I.T. Todorov, *On the Schwarz problem for the* $\widehat{su(2)}$ *Knizhnik-Zamolodchikov equation*, Lett. Math. Phys., **35** (1995), 123-134.

[42] A. Tsuchyia and Y. Kanie, *Vertex operators in two dimensional conformal field theory on P^1 and monodromy representations of braid groups*, Adv. Studies Pure Math., Tokyo, **16** (1988), 297-372.

[43] A. Tsuchyia, K. Ueno and Y. Yamada, *Conformal field theory on universal family of stable curves with gauge symmetries*, Adv. Studies Pure Math., Tokyo, **19** (1989), 459-566.

[44] V. Turaev, *Quantum invariants of links and* 3*-valent graphs in* 3*-manifolds*, Publ. Math. I.H.E.S., **77** (1993), 121-171.

[45] ———, *Quantum invariants of knots and* 3*-manifolds*, Vol. 18. de Gruyter Studies in Math., 1994.

[46] V. Turaev and O.Viro, *State-sum invariants of* 3*-manifolds and quantum* $6j$*-symbols*, Topology, **31** (1992), 865-902.

[47] V. Turaev and H. Wenzl, *Quantum invariants of* 3*-manifolds associated with classical simple Lie algebras*, International Journal of Math., **4** (1993), 323-358.

[48] K. Walker, *Chern-Simons-Witten invariants of* 3*-manifolds*, Preprint, 1993.

[49] H. Wenzl, *Hecke algebras of type A_n and subfactors*, Invent. Math., **92** (1988), 349-388.

[50] E. Witten, *Quantum field theory and the Jones polynomial*, Commun. Math. Phys., **121** (1989), 351-399.

[51] G. Wright, *The Reshetikhin-Turaev representation of the mapping class groups*, J. Knot Theory and Its Ramif., **3** (1994), 547-574.

[52] ———, *The Reshetikhin-Turaev representation of the mapping class groups at sixth roots of unity*, J. Knot Theory and Its Ramif., **5** (1996), 721-741.

Received June 10, 1997 and revised December 15, 1997.

Institute Fourier, 8P74
Univ. Grenoble I
38402 Saint-Martin-d'Hères cedex
France
E-mail address: funar@fourier.ujf-grenoble.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

UBIQUITY OF GEOMETRIC FINITENESS IN MAPPING CLASS GROUPS OF HAKEN 3-MANIFOLDS

Sungbok Hong and Darryl McCullough

For a Haken 3-manifold M with incompressible boundary, we prove that the mapping class group $\mathcal{M}$ acts properly discontinuously on a contractible simplicial complex, with compact quotient. This implies that every torsionfree subgroup of finite index in $\mathcal{M}$ is geometrically finite. Also, a simplified proof of the fact that torsionfree subgroups of finite index in $\mathcal{M}$ exist is given. All results are given for mapping class groups that preserve a boundary pattern in the sense of K. Johannson. As an application, we show that if F is a nonempty compact 2-manifold in ∂M such that $\partial M - F$ is incompressible, then the classifying space B$Diff(M$ rel $F)$ of the diffeomorphism group of M relative to F has the homotopy type of a finite aspherical complex.

1. Introduction.

The mapping class group $\mathcal{H}(F)$ of a 2-manifold is the group of isotopy classes of diffeomorphisms, $\pi_0(\mathrm{Diff}(F))$. As is well-known, it acts properly discontinuously on a Teichmüller space, which is topologically a Euclidean space (traditionally only the action of the orientation-preserving classes was considered). This classical setup was refined by Harer [**8**, **9**], who found an ideal triangulation of Teichmüller space for which $\mathcal{H}(F)$ acts simplicially, then constructed a contractible simplicial complex in the first barycentric subdivision of this triangulation which is invariant and has finite quotient under the action. Consequently, if Γ is any torsionfree subgroup of finite index in $\mathcal{H}(F)$ (and such subgroups always exist), the quotient of the action of Γ on this complex is a finite $K(\Gamma, 1)$-complex. That is, Γ is a *geometrically finite* group.

For a compact 3-manifold M, a great deal is known about $\mathcal{H}(M)$ (see for example [**20**] or the surveys [**25**, **27**]). In the present paper, we extend the aforementioned property of torsionfree subgroups of $\mathcal{H}(F)$ to the case of Haken 3-manifolds. To postpone the introduction of boundary patterns, we state here only a corollary of our main theorem. In the corollary, $\mathcal{H}(M \text{ rel } W)$ means $\pi_0(\mathrm{Diff}(M \text{ rel } W))$, where $\mathrm{Diff}(M \text{ rel } W)$ is the group of diffeomorphisms of M which restrict to the identity on W.

Corollary 7.10. *Let M be a Haken 3-manifold and let W be a compact 2-dimensional submanifold of ∂M such that $\partial M - W$ is incompressible. Then any torsionfree subgroup of finite index in $\mathcal{H}(M \text{ rel } W)$ is geometrically finite.*

In this result, W, or $\partial M - W$, or both may be empty. For Haken 3-manifolds, it was already known that $\mathcal{H}(M \text{ rel } W)$ *contains* a finite-index subgroup which is geometrically finite [**26**], so the new information from the Main Theorem is that *every* torsionfree subgroup of finite index in $\mathcal{H}(M \text{ rel } W)$ is geometrically finite. It is a longstanding and apparently difficult open question whether a torsionfree finite extension of a geometrically finite group must be geometrically finite. Of course, if this were known, the result from [**26**] would imply our strengthened version.

Our method of proof is to construct a topological action of $\mathcal{H}(M \text{ rel } W)$ on a contractible simplicial complex (whose quotient is compact), and as we show in Section 3 this is sufficient to deduce the geometric finiteness of torsionfree finite-index subgroups. It would be interesting to give a more direct construction of a contractible complex, along the lines of those developed by Harer, admitting a *simplicial* action of $\mathcal{H}(M \text{ rel } W)$ with finite quotient.

The Kontsevich Conjecture (Problem 3.48 in the new version of R. Kirby's problem list [**20**]) asserts that the classifying space $\mathrm{BDiff}(M \text{ rel } \partial M)$ has the homotopy type of a finite complex when M is a compact 3-manifold with nonempty boundary. This conjecture was recently proven for the case of Haken 3-manifolds in [**14**]. As observed there, the Kontsevich Conjecture for a Haken 3-manifold M is equivalent to the assertion that $\mathcal{H}(M \text{ rel } \partial M)$ is geometrically finite. In Section 8 we use our results on geometric finiteness to deduce a generalization of the result from [**14**]:

Theorem 8.7. *Let M be a Haken 3-manifold with incompressible boundary, and let F be a nonempty compact 2-manifold in ∂M such that $\partial M - F$ is incompressible. Then $\mathrm{BDiff}(M \text{ rel } F)$ has the homotopy type of a finite complex.*

The proof makes use of the extension of Nielsen's theorem to 3-manifolds made by Heil and Tollefson [**15**].

We will work in the context of 3-manifolds with boundary patterns. This lends greater generality to the results, and allows us to make direct use of Johannson's powerful characteristic submanifold theory for Haken manifolds. As is well-known, the characteristic submanifold was discovered and exploited independently by Johannson [**17**] and Jaco and Shalen [**16**]. We use Johannson's formulation because it is ideally suited to working with homotopy equivalences and homeomorphisms of 3-manifolds. In Section 2 we provide a brief exposition of the portion of Johannson's theory that we will use. In Section 3 we introduce a generalization of geometric finiteness, called (for lack of imagination) *almost geometric finiteness*. Any torsionfree finite-index subgroup of an almost geometrically finite group is geometrically finite. Using a theorem of Kamishima, Lee, and Raymond [**18**], we prove in Proposition 3.3 a key fact: An extension of a virtually finitely generated abelian group by an almost geometrically finite group is almost geometrically finite. The rest of the proof then follows the general approach of [**26**] to show that $\mathcal{H}(M)$ is an extension of this form, and hence is almost geometrically finite. It is necessary to work with a bit more precision than was needed in [**26**], since one can no longer evade difficulties by passing to subgroups of finite index in $\mathcal{H}(M)$. We also give a new proof that torsionfree subgroups of finite index in $\mathcal{H}(M)$ exist. This was proven in [**26**] by a very complicated argument; the new proof uses an algebraic fact from [**28**] to give a much shorter and more transparent proof. The final section contains the application to the Kontsevich Conjecture.

The authors thank BSRI-1422 and Nondirect Research Fund (K.R.F.) for support of their collaborative work, and acknowledge helpful discussions with Kyung-bai Lee and Leonard Rubin.

2. Johannson's characteristic submanifold theory.

We give here a brief review of the basic definitions of Johannson's formulation of the characteristic submanifold. We refer the reader to [**17**] for the original presentation, and also to Chapter 2 of [**2**] for a more extensive expository treatment with a number of examples.

A *boundary pattern* $\underline{\underline{m}}$ for an n-manifold M is a finite set of compact, connected $(n-1)$-manifolds in ∂M, such that the intersection of any i of them is empty or consists of $(n-i)$-manifolds. Thus when $n=3$, two elements of the boundary pattern intersect in a collection of arcs and circles, while if three elements meet, their intersection consists of a finite collection of points at which three intersection arcs meet. It is important in arguments to distinguish between elements of $\underline{\underline{m}}$ and the points of M which lie in them, and we will always be precise in this distinction. The symbol $|\underline{\underline{m}}|$ will mean the set of points of ∂M that lie in some element of $\underline{\underline{m}}$. When $|\underline{\underline{m}}|=\partial M$, $\underline{\underline{m}}$ is said to be *complete*. Provided that ∂M is compact, we define the *completion* of $\underline{\underline{m}}$ to be the complete boundary pattern $\overline{\underline{\underline{m}}}$ which is the union of $\underline{\underline{m}}$ and the set of components of the closure of $\partial M - |\underline{\underline{m}}|$. In particular, the set of boundary components of M is the completion of the empty boundary pattern $\overline{\underline{\underline{\emptyset}}}$.

Suppose $(X, \underline{\underline{x}})$ is an admissibly imbedded codimension-zero submanifold of $(M, \underline{\underline{m}})$, which is admissibly imbedded in $(M, \overline{\underline{\underline{m}}})$. The latter assumption guarantees that $X \cap \partial M = |\underline{\underline{x}}|$, that $|\underline{\underline{x}}|$ is contained in the topological interior of $|\underline{\underline{m}}|$ in ∂M, and that an element of $\underline{\underline{x}}$ which does not meet any other element of $\underline{\underline{x}}$ must be imbedded in the manifold interior of an element of $\underline{\underline{m}}$. Let $\underline{\underline{x}}''$ denote the collection of components of the frontier of X in M. To *split M along X* means to construct the manifold with boundary pattern $(\overline{M-X}, \underline{\underline{\widetilde{m}}} \cup \underline{\underline{x}}'')$, where the elements of $\underline{\underline{\widetilde{m}}}$ are the closures of the components of $F-(X\cap F)$ for $F\in \underline{\underline{m}}$. The boundary pattern $\underline{\underline{\widetilde{m}}}\cup\underline{\underline{x}}''$ is called the *proper boundary pattern* on $\overline{M-X}$.

Maps which respect boundary pattern structures are called admissible. Precisely, a map f from $(M,\underline{\underline{m}})$ to $(N,\underline{\underline{n}})$ is called *admissible* when $\underline{\underline{m}}$ is the disjoint union

$$\underline{\underline{m}} = \coprod_{G\in\underline{\underline{n}}} \{\text{components of } f^{-1}(G)\}.$$

An admissible map $h\colon (K,\underline{\underline{k}}) \to (X,\underline{\underline{x}})$, where K is an arc or a circle and $(X,\underline{\underline{x}})$ is a 2- or 3-manifold, is called *inessential* if it is admissibly homotopic to a constant map (the constant map might not be admissible, but all the other maps in the homotopy must be admissible), otherwise it is called *essential.* A map $f\colon (X,\underline{\underline{x}}) \to (Y,\underline{\underline{y}})$ between 2- or 3-manifolds (not necessarily of the same dimension) is called *essential* if for any essential path or loop $h\colon (K,\underline{\underline{k}}) \to (X,\underline{\underline{x}})$, the composition $fh\colon (K,\underline{\underline{k}}) \to (Y,\underline{\underline{y}})$ is essential. Notice that when $\underline{\underline{x}}$ is empty, this simply says that f is injective on fundamental groups.

The group of admissible isotopy classes of admissible diffeomorphisms from $(M,\underline{\underline{m}})$ to $(M,\underline{\underline{m}})$ is denoted by $\mathcal{H}(M,\underline{\underline{m}})$. The classes that preserve the orientation of each component are indicated by a plus subscript, as in $\mathcal{H}_+(M,\underline{\underline{m}})$. The classes relative to the subset $|\underline{\underline{m_1}}|$, where $\underline{\underline{m_1}} \subseteq \underline{\underline{m}}$, are denoted by $\mathcal{H}(M,\underline{\underline{m}} \text{ rel } |\underline{\underline{m_1}}|)$. Suppose $\langle h\rangle \in \mathcal{H}(M,\underline{\underline{m}})$. Since $h^{-1}(|\underline{\underline{m}}|) = |\underline{\underline{m}}|$, h must carry each free side of $(M,\underline{\underline{m}})$ diffeomorphically to a free side of $(M,\underline{\underline{m}})$. Therefore h is admissible for $(M,\overline{\underline{\underline{m}}})$. Thus when working with mapping class groups of manifolds with boundary patterns, the requirement that the boundary pattern be complete is not at all restrictive.

An *i-faced disc* is a 2-disc whose boundary pattern is complete and has i components. Observe that each element of $\underline{\underline{m}}$ is incompressible if and only if whenever D is an admissibly imbedded 1-faced disc in $(M,\underline{\underline{m}})$, the boundary of D bounds a disc in $|\underline{\underline{m}}|$ which is contained in a single element of $\underline{\underline{m}}$. For most of Johannson's theory, a somewhat stronger condition is needed. The boundary pattern $\underline{\underline{m}}$ of a 3-manifold M is called *useful* when the boundary of every admissibly imbedded i-faced disc in $(M,\underline{\underline{m}})$ with $i\leq 3$ bounds a disc D in ∂M such that $D\cap(\cup_{F\in\underline{\underline{m}}}\partial F)$ is the cone on $\partial D\cap(\cup_{F\in\underline{\underline{m}}}\partial F)$. Notice that $\overline{\underline{\underline{\emptyset}}}$ is a useful boundary pattern on M if and only ∂M is incompressible.

A Seifert fibering on a 3-manifold $(V,\underline{\underline{v}})$ with boundary pattern is called an *admissible Seifert fibering* when the elements of $\underline{\underline{v}}$ are the preimages of the components of a boundary pattern of the orbit surface. Consequently the elements of $\underline{\underline{v}}$ must be tori or fibered annuli.

Assume that V carries a fixed structure as an I-bundle over B. Each component of the associated ∂I-bundle is a 2-manifold in ∂V, called a *lid* of the I-bundle. There are two lids when the bundle is a product, and one when it is twisted. Let $\underline{\underline{b}}$ be a boundary pattern on B. The preimages of the elements of $\underline{\underline{b}}$ form a collection of squares and annuli in ∂V, called the *sides* of the I-bundle. The lid or lids, together with the sides, if any, form a boundary pattern $\underline{\underline{v}}$ on V. When V carries this boundary pattern, the fibering is called an *admissible I-fibering* of V as an I-bundle over $(B, \underline{\underline{b}})$. We emphasize that for an admissible I-fibering, *the lids are always elements of the boundary pattern.*

We are now ready to introduce the characteristic submanifold. An *admissible I-bundle* or *Seifert fibered space* $(X, \underline{\underline{x}})$ in a 3-manifold $(M, \underline{\underline{m}})$ is an I-bundle or Seifert fibered space imbedded in M such that the inclusion defines admissible maps $(X, \underline{\underline{x}}) \to (M, \underline{\underline{m}})$ and $(X, \underline{\underline{x}}) \to (M, \overline{\underline{\underline{m}}})$. An admissible I-bundle or Seifert fibered space in M is called *essential* when its frontier is an essential surface in $(M, \underline{\underline{m}})$. This implies that the inclusion of $(X, \underline{\underline{x}})$ into $(M, \underline{\underline{m}})$ is an essential map.

We suppose that $(M, \underline{\underline{m}})$ is a Haken 3-manifold with useful boundary pattern. A disjoint collection $(\Sigma, \underline{\underline{\sigma}})$ of essential admissible I-bundles and Seifert fibered spaces is a *characteristic submanifold* for $(M, \underline{\underline{m}})$ if

1) $(\Sigma, \underline{\underline{\sigma}})$ is full, i.e. the union of Σ with any of the components of $\overline{M - \Sigma}$ is not a disjoint union of essential admissible I-bundles and Seifert fibered spaces,
2) (Engulfing property) any essential admissible I-bundle or Seifert fibered space $(X, \underline{\underline{x}})$ in $(M, \underline{\underline{m}})$ is admissibly isotopic into $(\Sigma, \underline{\underline{\sigma}})$, and
3) (Enclosing property) any essential map $f\colon (T, \underline{\underline{t}}) \to (M, \underline{\underline{m}})$ of a square, annulus, or torus into $(M, \underline{\underline{m}})$ is admissibly homotopic to a map with image in Σ.

One may combine Proposition 9.4, Corollaries 10.9 and 10.10 and Theorem 12.5 of [**17**] to see that every Haken 3-manifold with useful boundary pattern has a characteristic submanifold, and that the characteristic submanifold is unique up to admissible isotopy.

A Haken 3-manifold $(M, \underline{\underline{m}})$ whose completed boundary pattern $\overline{\underline{\underline{m}}}$ is useful is called *simple* if every component of the characteristic submanifold of $(M, \overline{\underline{\underline{m}}})$ is a regular neighborhood of an element of $\overline{\underline{\underline{m}}}$. When the boundary consists of tori, the interior of a simple 3-manifold admits a hyperbolic structure of finite volume. Then, by Mostow Rigidity and Waldhausen's theorem, $\mathcal{H}(M, \overline{\underline{\underline{\emptyset}}})$ is isomorphic to the group of isometries, hence is finite. One of the deep applications of Johannson's theory is the following generalization of this fact (although it was proven before the dissemination of Thurston's work on hyperbolic structures of 3-manifolds). It appears as Proposition 27.1 of [**17**].

Theorem 2.1 (Finite Mapping Class Group Theorem)**.** *Let* $(M, \underline{\underline{m}})$ *be a simple* 3*-manifold with complete and useful boundary pattern. Then* $\mathcal{H}(M, \underline{\underline{m}})$ *is finite.*

3. Almost geometric finiteness.

We say that a group G is *almost geometrically finite* if it acts smoothly and properly discontinuously on a contractible simplicial complex L, such that L/G is compact. A subgroup of finite index in an almost geometrically finite group is almost geometrically finite. Almost geometrically finite groups are finitely generated (any group acting properly discontinuously on a connected locally compact space with compact quotient must be finitely generated). Every torsionfree subgroup of finite index in an almost geometrically finite group must be geometrically finite:

Lemma 3.1. *Let G be an almost geometrically finite group, and let Γ be a torsionfree subgroup of finite index in G. Then Γ is geometrically finite.*

Proof. Let G act properly discontinuously on the contractible complex L with compact quotient. Since Γ has finite index, L/Γ is compact, and since Γ is torsionfree, it acts freely on L. Therefore L/Γ is locally an ANR and hence an ANR (see for example Theorem 5.4.5 of [**32**]). By [**35**], a finite-dimensional compact ANR is homotopy equivalent to a finite simplicial complex, and therefore Γ is geometrically finite. □

We caution that in general, an almost geometrically finite group need not contain any geometrically finite subgroups of finite index: Schneebeli [**30**] constructed extensions $1 \to \mathbf{Z}/k \to E \to Q \to 1$ where Q is geometrically finite but E contains no torsionfree subgroups of finite index.

By Lemma 3.1, a group is geometrically finite if and only if it is almost geometrically finite and torsionfree. Simple examples of almost geometrically finite groups are finite groups (acting on a point) and finitely generated abelian groups (project to the quotient by the torsion subgroup and act on $\mathbf{R}^n$ by translations). By work of Borel and Serre [**1**], many arithmetic groups including $\mathrm{GL}(n, \mathbf{Z})$ are almost geometrically finite. Finitely generated virtually free groups are another interesting example; by [**19**] these are exactly the groups that act simplicially on trees with finite quotient and finite vertex stabilizers. In our context, mapping class groups of 2-manifolds are one of the most important examples:

Lemma 3.2. *Let B be a 2-manifold, not necessarily connected, of finite type with compact boundary, and let $\underline{\underline{b}}$ be a boundary pattern for B. Then $\mathcal{H}(B, \underline{\underline{b}})$ is almost geometrically finite.*

Proof. Write $\underline{\underline{b}} = \underline{\underline{a}} \cup \underline{\underline{c}}$ where the elements of $\underline{\underline{a}}$ are arcs and those of $\underline{\underline{c}}$ are circles. Let B_0 be obtained by removing all boundary circles that do not contain an element of $\underline{\underline{a}}$. Let P' consist of one point from the interior of each element of $\underline{\underline{a}}$ and let P'' consist of the points that are the intersection of two different elements of $\underline{\underline{a}}$. It is not difficult to see that $\mathcal{H}(B, \underline{\underline{b}})$ is isomorphic to a subgroup of finite index in $\mathcal{H}(B_0, P' \cup P'')$. Thus we are reduced to proving the lemma for $\mathcal{H}(B, P)$, where P is a finite subset of ∂B that meets every boundary component of B.

We will first prove the lemma under the assumption that B is connected. Suppose first that $\chi(B) \geq 0$. For S^2 with 0, 1, or 2 punctures, $\mathbf{RP}^2$ with 0 or 1 puncture, D^2 with 0 or 1 puncture, the Möbius band, or the Klein bottle, $\mathcal{H}(B, P)$ is finite, for the annulus it is virtually infinite cyclic, and for the torus it is $\mathrm{GL}(2, \mathbf{Z})$. From now on we assume that $\chi(B) < 0$.

Suppose first that P is empty, so B is either closed or is a punctured closed surface. If B is closed and orientable, then $\mathcal{H}(B)$ acts properly discontinuously on Teichmüller space. (Classically, from the viewpoint of conformal surfaces, only the action of the orientation-preserving classes was considered, but using the hyperbolic viewpoint as in [**5**], $\mathcal{H}(B)$ acts as well.) This action extends to the Harvey bordification [**10**, **11**] of Teichmüller space, on which it acts smoothly with compact quotient. Lifting a smooth triangulation gives the necessary simplicial structure to conclude that $\mathcal{H}(B)$ is almost geometrically finite. If B is orientable and punctured, then $\mathcal{H}(B)$ acts on the contractible complex Y^0 constructed by Harer [**8**, **9**] with compact quotient. The analogues of these for nonorientable surfaces are given in Section 2 of [**26**]. Suppose now that P is nonempty. Lemma 1.2 of [**14**] uses [**13**] to construct a contractible complex on which $\mathcal{H}(B, P)$ acts properly discontinuously.

We now assume that B is not connected. Let $B_1, \ldots, B_n$ be the components of B, and let $P_i = P \cap B_i$. Assume first that each (B_i, P_i) is diffeomorphic to (B_1, P_1). From the connected case, there is a contractible simplicial complex L such that $\mathcal{H}(B_1, P_1)$ acts properly discontinuously on L. Let L_i be a copy of L on which $\mathcal{H}(B_i, P_i)$ acts. Fix diffeomorphisms $f_i \colon (B_1, P_1) \to (B_i, P_i)$ and equivariant simplicial isomorphisms $\phi_i \colon L_1 \to L_i$. By means of these identifications, if $\langle h \rangle \in \mathcal{H}(B, P)$, and h carries B_i to B_j, we may regard $\langle h \rangle$ as carrying L_i to L_j. For let $h_i \colon B_i \to B_j$ be the restriction of h to B_i. Then let h_i send L_i to L_j by $\phi_j \phi_i^{-1} \circ \langle f_i f_j^{-1} h_i \rangle$. Now we define an action of $\mathcal{H}(B, P)$ on $\prod_{i=1}^{n} L_i$. Given h, define a permutation σ by $h(B_i) = B_{\sigma(i)}$. Then, put $\langle h \rangle(x_1, \ldots, x_n) = (\langle h_{\sigma^{-1}(1)} \rangle(x_{\sigma^{-1}(1)}), \ldots, \langle h_{\sigma^{-1}(n)} \rangle(x_{\sigma^{-1}(n)}))$. This action is properly discontinuous with compact quotient, since the finite index subgroup of $\mathcal{H}(B, P)$ that preserves each component acts as a product of properly discontinuous actions each with compact quotient.

For the general case, partition the components of (B, P) into maximal subsets such that each subset consists of pairwise diffeomorphic components.

Let $F_1, \ldots, F_m$ be the unions of the components in the subsets, and let $Q_i = P \cap F_i$. By the previous paragraph, each $\mathcal{H}(F_i, Q_i)$ acts properly discontinuously on a contractible complex K_i. Since $\mathcal{H}(B, P)$ is the direct product of the $\mathcal{H}(F_i, Q_i)$, it acts properly discontinuously on the product of the K_i. $\square$

A useful observation is that if there is a surjective homomorphism $Q_1 \to Q_2$ with finite kernel, and Q_2 is almost geometrically finite, then Q_1 is also almost geometrically finite. In our proof that mapping class groups of Haken 3-manifolds are almost geometrically finite, we will use the following generalization of this observation, which is a direct consequence of a theorem of Kamishima, Lee, and Raymond [**18**]:

Proposition 3.3. *Let $1 \to V \to G \to Q \to 1$ be an exact sequence of groups, such that V contains a finitely generated abelian group of finite index and Q is almost geometrically finite. Then G is almost geometrically finite.*

Proof. Consider a subgroup $\mathbf{Z}^n$ of finite index in V. Under conjugation by elements of G, its orbit consists of finitely many subgroups, whose intersection is a normal subgroup isomorphic to $\mathbf{Z}^n$. Let Q' be the quotient of G by this subgroup. Since Q' maps onto Q with finite kernel, it is almost geometrically finite. Therefore we may assume that V itself is free abelian of rank n. From Proposition 2.2 of [**18**], there exists a Seifert construction for this data, that is, a properly discontinuous action of G on $\mathbf{R}^n \times L$, where L is any contractible simplicial complex on which Q acts properly discontinuously. The action takes $\mathbf{R}^n$-fibers (over points of L) to $\mathbf{R}^n$-fibers, and the quotient of each $\mathbf{R}^n$-fiber is a quotient of a compact flat manifold by a finite group action. Therefore $(\mathbf{R}^n \times L)/G$ is compact, so G is almost geometrically finite. $\square$

4. Laudenbach's Theorem.

Throughout this section let M be a Haken 3-manifold, $M \neq D^3$, and let $G \neq S^2$ be a compact connected properly imbedded incompressible surface in M. Let m_0 be a basepoint in G. If G is not closed, choose m_0 in ∂G. Denote by $\mathrm{Imb}^0(G, M)$ the space of smooth imbeddings of G in M that take m_0 to m_0 and ∂G to ∂G. The inclusion map of G into M is understood to be the basepoint of $\mathrm{Imb}^0(G, M)$. The following result is proven on pp. 49-62 of [**22**].

Theorem 4.1. $\pi_1(\mathrm{Imb}^0(G, M)) = \{1\}$.

From this we deduce the result we will need.

Theorem 4.2. *Assume that G is not an annulus in a solid torus or a boundary-parallel disc. Let f and g be diffeomorphisms of M which preserve G and fix m_0. Let H be an isotopy from f to g, whose trace at m_0 lies in*

$\pi_1(G, m_0)$. Then H is deformable relative to $M \times \partial I$ to an isotopy through diffeomorphisms that preserve G. Moreover,

(i) *If H is relative to ∂G, then the new isotopy may be chosen to be relative to ∂G.*
(ii) *If H is relative to ∂M, then the new isotopy may be chosen to be relative to ∂M.*
(iii) *If H is relative to ∂M, and f and g agree on G, then the new isotopy may be chosen to be relative to $G \cup \partial M$.*
(iv) *If F is an incompressible surface (not necessarily connected) in M, disjoint from G, such that H preserves F, then the new isotopy may be chosen to agree with H on F.*

Note that by virtue of part (iv), the theorem applies to G that are not connected, provided that the trace condition is satisfied for a basepoint in each component of G.

Proof. Replacing f by $g^{-1}f$, we may assume that g is the identity. If $\partial G \neq \emptyset$ and the hypothesis of one of cases (i), (ii), or (iii) holds, then H is already an isotopy relative to m_0. Otherwise, we may change f by isotopy in a neighborhood of G so that the trace of H at m_0 is trivial, and then (using [**24**]) so that H is an isotopy relative to m_0. Consequently H induces the identity automorphism on $\pi_1(M, m_0)$.

Let f_1 be the restriction of f to G. We will first reduce to the situation when f_1 is the identity. In case (iii), this already holds. In cases (i) and (ii), H is an isotopy relative to ∂G or ∂M. As in Lemma 7.3 of [**34**], f_1 is homotopic to the identity preserving ∂G, so f_1 is isotopic to the identity. Choose a basepoint in each component of ∂G. Suppose that γ is an arc in G connecting two of them. The restriction of H to γ gives an isotopy in M from $f(\gamma)$ to γ, and since $\pi_2(M, G)$ is trivial (because $\pi_2(M) = 0$ and $\pi_1(G) \to \pi_1(M)$ is injective), $f(\gamma)$ is homotopic and hence isotopic to γ in G. Thus by changing f by isotopy preserving G and relative to ∂M, we may assume that f_1 preserves a set of arcs (disjoint except for their endpoints) connecting the basepoints in the boundary circles. This, together with the fact that f_1 is isotopic to the identity relative to m_0, implies that f_1 is isotopic to the identity relative to ∂G. After changing f by isotopy preserving G and relative to ∂M, we may assume that f_1 is the identity. Finally, suppose that none of (i), (ii), or (iii) holds. The trace condition implies that f induces the identity automorphism on $\pi_1(M, m_0)$. Since $f(G) = G$, the restriction f_1 of f to G induces the identity automorphism on $\pi_1(G, m_0)$. Therefore f_1 is homotopic to the identity. Any orientation-preserving diffeomorphism of a compact 2-manifold that is homotopic to the identity is isotopic to the identity, unless G is a disc or annulus and the diffeomorphism is orientation-reversing. If f_1 is orientation-reversing, then f must reverse the sides of G. If G were a disc, then since f induces the identity automorphism on $\pi_1(M, m_0)$, M would have to be a 3-ball, so G would be a boundary-parallel disc. If G were an annulus, then for the same reason M would have to be a solid torus, also excluded. So f_1 is isotopic to the identity, and since f_1 induces the identity automorphism on $\pi_1(G, m_0)$, we may assume that the isotopy preserves m_0. Extending such an isotopy of f_1 to an isotopy of f, we may assume that f_1 is the identity on G.

Next we need to achieve that H preserves ∂G. In (i), (ii), or (iii), no alteration is needed. Otherwise, let C be the component of ∂G that contains m_0, let B be the component of ∂M that contains C, and consider the trace τ of $H|_B$ at m_0. Note that C is not contractible in B, since G is not a boundary-parallel disc. Suppose that τ is not trivial in $\pi_1(B)$. Suppose for contradiction that τ does not lie in $\pi_1(C)$. Since f is the identity on C, τ lies in the normalizer of $\pi_1(C)$ in $\pi_1(B)$. Therefore B is a torus, and since τ is essential in $\pi_1(B)$ and trivial in $\pi_1(M)$, M must be a solid torus. Since τ is not in $\pi_1(C)$, C is not contractible in $\pi_1(M)$, so G is an annulus, the case excluded by hypothesis. So we may assume that τ lies in $\pi_1(C)$. Since τ is trivial in $\pi_1(M)$, G must be a disc. There is an isotopy of M that preserves G and moves m_0 around C. Juxtaposing H with the correct multiple of this isotopy, we may assume that τ is trivial in $\pi_1(B)$. By [**6**], $\pi_1(\mathrm{Imb}((C, m_0), (B, m_0)))$ is trivial unless B is a 2-sphere, a case excluded since G is not a boundary-parallel disc, so H may be deformed in a neighborhood of B so that it preserves C at each level. Repeating for the other components of ∂G, we may assume that ∂G is preserved at each level of H.

The restriction j_t of H to $G \times I$ defines a loop in $\pi_1(\mathrm{Imb}^0(G, M))$. Applying Theorem 4.1, this loop is contractible, so there exists a 2-parameter family $j_{t,s}$, $0 \leq s, t \leq 1$, such that $j_{t,0} = j_t$, and $j_{t,1}$, $j_{0,s}$, and $j_{1,s}$ are the inclusions for each t and s. Define $J_{t,0} = H_t$, $J_{0,s} = f$, and $J_{1,s} = 1_M$. By the isotopy extension theorem (i.e. the fact that $\mathrm{Diff}^0(M) \to \mathrm{Imb}^0(G, M)$ is a Serre fibration) this extends to a 2-parameter family $J_{t,s}$ of diffeomorphisms of M. Letting $K_t = J_{t,1}$, we have an isotopy from f to 1_M relative to G, and the existence of an isotopy preserving G together with statements (i), (ii), and (iii) are established.

For (iv), notice that all of our alterations to H may be performed so as not to change H outside a neighborhood of G. Therefore if F is another incompressible surface (not necessarily connected) which is preserved by H, we cut along F and apply the previous argument to obtain a new isotopy agreeing with H on F. □

There is a 2-dimensional analogue of Theorem 4.2.

Theorem 4.3. *Let $G \neq S^2$ be a surface and k an arc or circle essentially imbedded in G. If k is a circle, let m_0 be a basepoint in k. Let f and g be diffeomorphisms of G which preserve G. Let H be an isotopy from f to g*

through diffeomorphisms fixing m_0*, if* k *is a circle, or preserving* ∂k*, if* k *is an arc. Then* H *is deformable relative to* $G \times \partial I$ *to an isotopy through diffeomorphisms that preserve* k*. Moreover,*

(i) *If* H *is relative to* ∂M*, then the new isotopy may be chosen to be relative to* ∂M*.*
(ii) *If* H *is relative to* ∂M*, and* f *and* g *agree on* G*, then the new isotopy may be chosen to be relative to* $G \cup \partial M$*.*
(iii) *If* ℓ *is a* 1*-manifold in* G*, disjoint from* k*, such that* H *preserves* ℓ*, then the new isotopy may be chosen to agree with* H *on* ℓ*.*

The proof is analogous to the proof of Theorem 4.2, but much simpler.

5. Exceptional Seifert-fibered 3-manifolds.

In the next section we will give a general treatment of the mapping class groups of Seifert-fibered Haken 3-manifolds, but there are a few exceptional cases to which it will not apply. We address those cases in the present section. The manifolds (all assumed to be admissibly fibered with complete boundary pattern) are:

(E1) The S^1-bundle over the annulus, with boundary pattern $\underline{\underline{\overline{\emptyset}}}$.
(E2) The S^1-bundle over the Möbius band, with boundary pattern $\underline{\underline{\overline{\emptyset}}}$.
(E3) An S^1-bundle over the torus.
(E4) An S^1-bundle over the Klein bottle.
(E5) The Hantzsche-Wendt manifold, which is the closed flat 3-manifold with Seifert invariants $\{-1; (n_2, 1); (2, 1), (2, 1)\}$ (see [**29**] pp. 133, 138, [**3**] pp. 478-481, [**7**, **33**, **36**]).
(E6) A Haken manifold which fibers over S^2 with three exceptional orbits.

Proposition 5.1. *Let* $(\Sigma, \underline{\underline{\sigma}})$ *be admissibly fibered as a Seifert* 3*-manifold of one of the exceptional types* (E1)-(E4)*. Then* $\mathcal{H}(\Sigma, \underline{\underline{\sigma}})$ *is almost geometrically finite.*

Proof. For (E1), from Proposition 3.4.1 of [**26**] we have $\mathcal{H}(\Sigma, \underline{\underline{\sigma}})$ isomorphic to $\mathbf{Z}/2 \times \mathrm{GL}(2, \mathbf{Z})$ (the $\mathbf{Z}/2$ factor is generated by reflection in the I-fibers of the I-bundle structure). So $\mathcal{H}(\Sigma, \underline{\underline{\sigma}})$ is virtually free. For (E2), $\mathcal{H}(\Sigma, \underline{\underline{\sigma}})$ is finite, by Proposition 3.4.2 of [**26**]. For (E3), if the Euler class is zero then Σ is the 3-torus, with $\mathcal{H}(\Sigma) \cong \mathrm{Out}(\pi_1(\Sigma)) \cong \mathrm{GL}(3, \mathbf{Z})$. If the Euler class is n, then $\pi_1(\Sigma) \cong \langle x, y, t \mid [x, t] = [y, t] = 1, [x, y] = t^n \rangle$. The center is $\mathbf{Z}$ generated by t, and the quotient of $\pi_1(\Sigma)$ by its center is $\mathbf{Z} \times \mathbf{Z}$ generated by the images of x and y. Sending $\mathrm{Out}(\pi_1(\Sigma))$ to $\mathrm{Aut}(\mathbf{Z} \times \mathbf{Z}) = \mathrm{GL}(2, \mathbf{Z})$ is surjective and splits (the three automorphisms determined by sending (1) x to xy, y to y, and t to t, (2) x to x, y to xy, and t to t, and (3) x to y, y to x, and t to t^{-1} define the splitting). Elements of the kernel are the automorphisms $\phi(i, j)$ that send x to xt^i, y to yt^j, and t to t. Conjugation by x equals $\phi(0,n)$ and conjugation by y equals $\phi(-n,0)$, so the kernel of $\mathrm{Out}(\pi_1(\Sigma)) \to \mathrm{GL}(2,\mathbf{Z})$ is $\mathbf{Z}/n \times \mathbf{Z}/n$, showing that $\mathcal{H}(\Sigma) \cong \mathrm{Out}(\pi_1(\Sigma))$ is almost geometrically finite. The manifolds of type (E4) are analyzed in Proposition 3.4.4 of [**26**]. If the Euler class is zero then there is a homomorphism from $\mathrm{Out}(\pi_1(M))$ to $\mathrm{PGL}(2,\mathbf{Z})$ with finite kernel. Since $\mathrm{PSL}(2,\mathbf{Z})$ is virtually free, it is almost geometrically finite. If the Euler class is nonzero, $\mathrm{Out}(\pi_1(\Sigma))$ is finite. □

Proposition 5.2. *Let M be one of the exceptional types* (E5)-(E6). *Then $\mathcal{H}(M)$ is finite.*

Proof. For (E5), it is from [**3**] (although the correct structure of the group was later given in [**36**]). The lemma in Section 3.4 of [**26**] gives case (E6). □

6. Fibered 3-manifolds.

We first treat the case of I-bundles, over surfaces which are not necessarily connected.

Lemma 6.1. *Suppose that $(\Sigma,\underline{\underline{\sigma}})$ is admissibly I-fibered over the compact 2-manifold $(B,\underline{\underline{b}})$. Let $p\colon \Sigma \to B$ be the projection. Then $\mathcal{H}(\Sigma,\underline{\underline{\sigma}})$ is isomorphic to a semidirect product $F \circ \mathcal{H}(B,\underline{\underline{b}})$ where F is a direct sum of copies of $\mathbf{Z}/2$, one for each component of Σ, and the action of an element of $\mathcal{H}(B,\underline{\underline{b}})$ on F is to permute the copies of $\mathbf{Z}/2$ exactly as it permutes the corresponding components of Σ.*

Proof. Let $i\colon (B,\underline{\underline{b}}) \to (\Sigma,\underline{\underline{\sigma}})$ be the 0-section of the I-bundle (where I is regarded as $[-1,1]$ and the structure group is reduced to $\mathbf{Z}/2$ generated by reflection in I, and the twisting is given by the orientation homomorphism). For each component of Σ there is the involution given by reflection in the I-fibers, and this is not admissibly isotopic to the identity since it is orientation-reversing. These give the generators of F.

Define $j\colon \mathcal{H}(B,\underline{\underline{b}}) \to \mathcal{H}(\Sigma,\underline{\underline{\sigma}})$ by extending diffeomorphisms on $i(B)$ to Σ linearly in each fiber, choosing the unique way to do this that is orientation-preserving on Σ. To see that this is injective, let $\langle h\rangle$ be an element of $\mathcal{H}(B,\underline{\underline{b}})$ such that $j(\langle h\rangle)$ is trivial in $\mathcal{H}(\Sigma,\underline{\underline{\sigma}})$. Then h is orientation-preserving, since otherwise the restriction of $j(h)$ of the lid or lids of Σ cannot be isotopic to the identity. Projecting an admissible isotopy from $j(h)$ to 1_Σ down to $i(B)$ gives an admissible homotopy from h to 1_B. This implies that h is admissibly isotopic to 1_B (see for example Lemma 2.19 of [**2**]). Corollary 5.9 of [**17**] shows that the image of j is the entire group of orientation-preserving mapping elements of $\mathcal{H}(\Sigma,\underline{\underline{\sigma}})$. Then, it is clear that the subgroups F and $j(\mathcal{H}(\Sigma,\underline{\underline{\sigma}}))$ generate $\mathcal{H}(\Sigma,\underline{\underline{\sigma}})$, and F is normal, and the lemma follows. □

Throughout the remainder of this section, $(\Sigma,\underline{\underline{\sigma}})$ will denote a Haken 3-manifold with complete and useful boundary pattern, which admits an admissible Seifert fibering over $(B, \underline{\underline{b}})$. A diffeomorphism of Σ is called *fiber-preserving* if it carries each fiber Σ to a fiber of Σ, and is called *vertical* if it takes each fiber to itself. By $\mathcal{H}_f(\Sigma, \underline{\underline{\sigma}})$ we indicate the mapping classes of fiber-preserving diffeomorphisms (that is, fiber-preserving diffeomorphisms modulo isotopy through fiber-preserving diffeomorphisms). There is a natural homomorphism $\mathcal{H}_f(\Sigma, \underline{\underline{\sigma}}) \to \mathcal{H}(\Sigma, \underline{\underline{\sigma}})$.

Theorem 6.2. *If* $(\Sigma, \underline{\underline{\sigma}})$ *is not one of* (E1)-(E6), *then* $\mathcal{H}_f(\Sigma, \underline{\underline{\sigma}}) \to \mathcal{H}(\Sigma, \underline{\underline{\sigma}})$ *is an isomorphism.*

Proof. Provided that $(\Sigma, \underline{\underline{\sigma}})$ is not an exception (E1)-(E5), [**33**] or Theorem 8.1.7 of [**29**] shows that the homomorphism is surjective. If it is not an exception (E6), then it has a vertical incompressible surface, and the argument of p. 85-86 of [**34**] shows that it is injective. $\square$

Define $\mathcal{H}^0(\Sigma, \underline{\underline{\sigma}})$ to be the elements of $\mathcal{H}(\Sigma, \underline{\underline{\sigma}})$ that contain a vertical diffeomorphism. As in Lemma 25.2 of [**17**] (see also Lemma 3.5.3 of [**26**]), we have the following calculation.

Proposition 6.3. *Let* Σ *be Seifert-fibered over* $(B, \underline{\underline{b}})$*, with no component of* Σ *an exceptional case* (E1)-(E6). *Then* $\mathcal{H}^0_+(\Sigma, \underline{\underline{\sigma}}) \cong H_1(B, |\underline{\underline{b}}|)$.

The rough idea behind this result is that the generators of $\mathcal{H}^0_+(\Sigma, \underline{\underline{\sigma}})$ are Dehn twists (see [**17**] or Section 3.3 of [**26**] for a definition of Dehn twist) about vertical tori and annuli, which obey the same homological relations as their image circles and arcs in B. When B is nonorientable, there is another type of generator supported in a neighborhood of a vertical Klein bottle; its square is a Dehn twist about the boundary of a regular neighborhood of the Klein bottle.

Let E be the exceptional points of B, that is, the images of the exceptional orbits of Σ. This is a finite subset of the interior of B. Denote by $\rho\colon \mathcal{H}(\Sigma, \underline{\underline{\sigma}}) \to \mathcal{H}(B - E, \underline{\underline{b}})$ the homomorphism induced by projection of fiber-preserving homomorphisms to the base surface. Define $\mathcal{H}_0(B - E, \underline{\underline{b}})$ to be the subgroup of $\mathcal{H}(B - E, \underline{\underline{b}})$ consisting of the classes $\langle f \rangle$ such that f is admissibly isotopic to a map which is the identity on $|\underline{\underline{b}}| = \partial B$, and f permutes the punctures of B trivially. Since $\underline{\underline{b}}$ consists of arcs and circles (and since B is of finite type), $\mathcal{H}_0(B - E, \underline{\underline{b}})$ has finite index in $\mathcal{H}(B - E, \underline{\underline{b}})$.

Proposition 6.4. *The image of* $\rho\colon \mathcal{H}_f(\Sigma, \underline{\underline{\sigma}}) \to \mathcal{H}(B - E, \underline{\underline{b}})$ *contains* $\mathcal{H}_0(B - E, \underline{\underline{b}})$*, hence has finite index in* $\mathcal{H}(B - E, \underline{\underline{b}})$*. If* $\partial B \neq \emptyset$*, then there exists a homomorphism* $s\colon \mathcal{H}_0(B - E, \underline{\underline{b}}) \to \mathcal{H}_f(\Sigma, \underline{\underline{\sigma}})$ *such that* ρs *is the identity, and such that each* $s(\langle f \rangle)$ *has a representative which is the identity on* $\partial\Sigma$.

Proof. The proposition follows from the special case when B is connected. The connected case appears a Proposition 25.3 of [**17**] and Theorem 3.5.2 of [**26**]. The splitting is constructed using a cross-section $i\colon B - E \to \Sigma - \widetilde{E}$, where $\widetilde{E}$ is the union of the exceptional fibers. For $\langle f\rangle \in \mathcal{H}_0(B-E,\underline{\underline{b}})$, one may assume that f is the identity on ∂B, so (since $s(f)$ is always selected to be orientation-preserving) that $s(f)$ is the identity on $i(\partial B)$ and hence on $\partial\Sigma$. $\square$

Propositions 6.3 and 6.4, combined with Lemma 3.2, yield immediately the following:

Theorem 6.5. *Let* $(\Sigma,\underline{\underline{\sigma}})$ *be an admissibly Seifert-fibered Haken* 3*-manifold with complete and useful boundary pattern* σ*. Assume that* $(\Sigma,\underline{\underline{\sigma}})$ *is not one of the exceptional manifolds* (E1)-(E6)*. Then there is an exact sequence*

$$1 \longrightarrow V \longrightarrow \mathcal{H}(\Sigma,\underline{\underline{\sigma}}) \longrightarrow Q \longrightarrow 1$$

where V *has a finitely generated abelian subgroup of finite index, and* Q *is almost geometrically finite.*

Applying Proposition 3.3, we have:

Corollary 6.6. *Let* $(\Sigma,\underline{\underline{\sigma}})$ *be an admissibly Seifert-fibered Haken* 3*-manifold with complete and useful boundary pattern* σ*. Assume that* $(\Sigma,\underline{\underline{\sigma}})$ *is not one of the exceptional manifolds* (E1)-(E6)*. Then* $\mathcal{H}(\Sigma,\underline{\underline{\sigma}})$ *is almost geometrically finite.*

7. Haken manifolds.

Throughout this section we assume that $(M,\underline{\underline{m}})$ is a Haken manifold with a complete and useful boundary pattern. We allow the possibility that ∂M is empty. Denote by Σ the characteristic submanifold of $(M,\underline{\underline{m}})$. Let $\underline{\underline{\sigma}}$ be the proper boundary pattern on Σ. The following result was proved in [**26**], but we present a less abbreviated proof here. A reference for the *Sol* geometry is Theorem 5.5 of [**31**].

Proposition 7.1. *Suppose that* M *is not a torus bundle over the circle that admits a Sol structure. Then* $\mathcal{H}(M,\Sigma,\underline{\underline{m}}) \to \mathcal{H}(M,\underline{\underline{m}})$ *is an isomorphism.*

Proof. Since the characteristic submanifold is unique up to admissible isotopy, the homomorphism is surjective. For injectivity, let $\langle f\rangle \in \mathcal{H}(M,\Sigma,\underline{\underline{m}})$ and suppose that $H\colon M\times I \to M$ is an admissible isotopy from f to 1_M. We must find an isotopy that preserves the frontier of Σ.

Let F be a component of the frontier of Σ. We claim that $f(F) = F$. Suppose not. The restriction of H to $F\times I$ is an admissible map from an I-bundle or Seifert fiber space into $(M,\underline{\underline{m}})$, so by Proposition 13.1 of [**17**], it is admissibly homotopic into Σ. That is, two components of the frontier of Σ are admissibly homotopic in Σ. By Proposition 19.1 of [**17**], these components must be admissibly parallel in Σ, that is, the component of Σ containing F is of the form $F\times I$ and f interchanges of the components of its boundary. Therefore f is isotopic to a diffeomorphism that preserves F

and interchanges its sides. Since f is isotopic to the identity, Lemma 7.4 of [**34**] shows that this is impossible.

By isotopy preserving Σ, we may assume that f fixes a basepoint m_0 in the interior of F. We claim that the trace of H at m_0 is in the subgroup $\pi_1(F, m_0)$. When ∂M is nonempty, Corollary 18.2 of [**17**] applies to prove the claim. When ∂M is empty, the argument in Lemma 18.1 of [**17**] shows that if the claim is false then the components of Σ and $\overline{M-\Sigma}$ adjacent to F are each diffeomorphic to the product of the torus and an interval. By maximality of Σ, this is only possible when M is a torus bundle over the circle which admits a *Sol* structure, which is excluded by hypothesis. This establishes the claim.

Since F is a square, annulus, or torus, there is an isotopy on F from the identity to the identity whose trace is equal to the trace of H. So it is possible to change f by an admissible isotopy with support in a neighborhood of F, so that the trace of the isotopy from f to the identity of M is trivial. Then f must induces the identity automorphism on $\pi_1(M, m_0)$. Since f preserves F, it also induces the identity isomorphism on $\pi_1(F, m_0)$.

Let f_1 be the restriction of f to F. We will show that f_1 preserves each element of the boundary pattern of F. If F is a torus, there is nothing to prove. Suppose F is an annulus, so that the boundary pattern consists of the two boundary circles. If f_1 interchanges them, then since f_1 induces the identity automorphism it must be orientation-reversing. But f preserves the sides of F (because f preserves Σ, or alternatively using Lemma 7.4 of [**34**] again), so f would be orientation-reversing and could not be isotopic to the identity. Suppose that F is a square. Its boundary pattern consists of the four edges. Suppose for contradiction that f_1 moves some edge to a different edge. Since f is admissibly homotopic to the identity, it must preserve each element of $\underline{\underline{m}}$. Since adjacent edges cannot lie in the same element of $\underline{\underline{m}}$, f_1 must interchange a pair of opposite edges. If it interchanges one pair of opposite edges, but preserves each of the other two edges, then f_1 is orientation-reversing, a contradiction as in the case when F is an annulus. Therefore f must interchange both pairs of opposite edges. Since F is a square, the component of Σ containing F must be an I-bundle, and since opposite edges of F lie in the same component of the boundary pattern, one pair lies in the lid and the other pair are joined by a square S which is contained in an element F' of $\underline{\underline{m}}$. Now the I-bundle cannot be fibered over the disc, because since F and S are sides which meet in two fibers, the I-bundle would be fibered over a 2-faced disc and would not be essentially imbedded. This is impossible since it is a component of Σ. So the center circle of the annulus $F \cup S$ is essential in M. Since f interchanges opposite edges of F, it must send this element of $\pi_1(M)$ to its negative, which is impossible since f induces the identity automorphism on $\pi_1(M, m_0)$.

Fix a component F of the frontier of Σ. Suppose F is a torus. Since the trace of the isotopy at m_0 is trivial, Theorem 4.2 applies to show that the isotopy from f to the identity may be deformed to preserve F at each stage.

Suppose F is an annulus. Consider a basepoint b_0 in a boundary circle C of F. Since f_1 is the identity on F, and the trace of H at m_0 is trivial, the trace of H at b_0 is also trivial in $\pi_1(M, b_0)$. Since the elements of $\underline{\underline{m}}$ are incompressible, the trace at b_0 is trivial in $\pi_1(G)$ where G is the element of $\underline{\underline{m}}$ that contains b_0. Applying Theorem 4.3, we may deform the isotopy admissibly in a neighborhood of ∂M so that C preserved at each level of the isotopy. Repeating, we assume that all of ∂F is preserved, and then apply Theorem 4.2.

Now suppose F is a square. Let b_0 be a corner where two edges meet, so $b_0 \in G_1 \cap G_2$ for two elements of the boundary pattern. Again, the trace of H at b_0 is trivial in $\pi_1(G_1)$ and $\pi_1(G_2)$. If it is not trivial in the fundamental group of the component k of $G_1 \cap G_2$ that contains it, then that component is a circle and G_1 and G_2 must be discs. Then M would be a 3-ball with boundary pattern $\{G_1, G_2\}$ and M could not contain the essentially imbedded square F. We conclude that the trace of H at b_0 is trivial in $\pi_1(k)$, so we may assume that H preserves b_0. Repeating, we assume that H preserves each of the four corners of F. Now using Theorem 4.3 we may assume that H preserves each of the four edges of F, and apply Theorem 4.2 to assume that H preserves F.

Repeating this for each component of the frontier of Σ, not disturbing those already adjusted, shows that f was the trivial element of $\mathcal{H}(M, \Sigma, \underline{\underline{m}})$. $\square$

From now until when we reach Theorem 7.7, we assume that $M \neq \Sigma$. Let $H = \overline{M - \Sigma}$, and let $\underline{\underline{h}}$ be the proper boundary pattern on H. Define R_H to be the image of the restriction $\mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to \mathcal{H}(H, \underline{\underline{h}})$. From Lemma 4.2.1 of [**26**] we have the following fact.

Lemma 7.2. *R_H is finite.*

Actually, according to Theorem 2.1, $\mathcal{H}(H_i, \underline{\underline{h_i}})$ is itself finite for all components of $(H, \underline{\underline{h}})$ except the case when $(H_i, \underline{\underline{h_i}}) = (T^2 \times I, \underline{\underline{\emptyset}})$. For these a special argument is given in [**26**]. The idea is that on the adjacent Seifert-fibered component(s) of Σ, any diffeomorphism must be fiber-preserving up to isotopy. The fibers are linearly independent in $\pi_1(T^2 \times I) \cong \mathbf{Z} \times \mathbf{Z}$, and only $\pm I$ can preserve two linearly independent elements of $\mathbf{Z} \times \mathbf{Z}$, allowing at most two possibilities for the restriction to $\mathcal{H}(H_i, \underline{\underline{h_i}})$.

Since $M \neq \Sigma$, the components of Σ, with their proper boundary patterns, are of four kinds.

(i) The components which can be admissibly fibered as I-bundles with their lids in ∂M. Their union will be denoted by I.

(ii) Those that are diffeomorphic to $S^1 \times S^1 \times I$ with boundary pattern $\overline{\underline{\underline{\phi}}}$, but are not as in (i) (that is, exceptional case (E1)). These must be Seifert-fibered (since as I-bundles they could not have both lids in ∂M), and must have frontier equal to one or both of their boundary components. Their union will be denoted by T.

(iii) Those that are diffeomorphic to the twisted I-bundle over the Klein bottle with boundary pattern $\overline{\underline{\underline{\phi}}}$ (that is, exceptional case (E2)). Their union will be denoted by K.

(iv) The components that are not as in (i), (ii), or (iii). These are Seifert-fibered, since they are not as in (i), and their fiberings are unique up to admissible isotopy, since they are not as in (ii) or (iii). Their union will be denoted by S.

The proper boundary pattern of I is denoted $\underline{\underline{i}}$, and similarly for K, T, and S. Clearly each of I, K, T, and S is preserved by each element of $\mathcal{H}(M, \Sigma, \underline{\underline{m}})$. Let R_T be the image of the restriction $\mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to \mathcal{H}(T, \underline{\underline{t}})$.

Lemma 7.3. *R_T is finite.*

Proof. Consider the composition

$$\mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to R_T \to \mathcal{H}(\mathrm{Fr}(T)),$$

where $\mathrm{Fr}(T)$ is the frontier. Since this equals the composition

$$\mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to \mathcal{H}(H, \underline{\underline{s}}) \to \mathcal{H}(\mathrm{Fr}(T)),$$

and $\mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to \mathcal{H}(H, \underline{\underline{s}})$ has finite image by Lemma 7.2, it follows that $R_T \to \mathcal{H}(\mathrm{Fr}(T))$ has finite image. But it is also injective, since any diffeomorphism of T isotopic to the identity on a boundary component is isotopic to the identity. Therefore R_T is finite. □

Lemma 7.4. *$\mathcal{H}(K, \underline{\underline{k}})$ is finite.*

Proof. Since $(K, \underline{\underline{k}})$ fibers admissibly as an I-bundle, $\underline{\underline{k}} = \overline{\underline{\underline{\emptyset}}}$. By Lemma 6.1, if X is the twisted I-bundle over the Klein bottle Y, we have $\mathcal{H}(X, \overline{\underline{\underline{\emptyset}}}) \cong \mathcal{H}(Y)$, and by [**23**] the latter is $\mathbf{Z}/2 \times \mathbf{Z}/2$. Since K has finitely many components, $\mathcal{H}(K, \underline{\underline{k}})$ is also finite. □

Consider the restriction homomorphism

$$\mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to R_H \times R_T \times \mathcal{H}(K, \underline{\underline{k}}) \times \mathcal{H}(I, \underline{\underline{i}}) \times \mathcal{H}(S, \underline{\underline{s}}).$$

Let $(B, \underline{\underline{b}})$ be the quotient surface of $(S, \underline{\underline{s}})$, and let E be the image of the exceptional fibers. By Theorem 6.2 and Proposition 6.4, there is a homomorphism $\mathcal{H}(S, \underline{\underline{s}}) \to \mathcal{H}(B - E, \underline{\underline{b}})$, and composing with this in the $\mathcal{H}(S, \underline{\underline{s}})$ factor, we obtain a homomorphism

$$\Psi \colon \mathcal{H}(M, \Sigma, \underline{\underline{m}}) \to R_H \times R_T \times \mathcal{H}(K, \underline{\underline{k}}) \times \mathcal{H}(I, \underline{\underline{i}}) \times \mathcal{H}(B - E, \underline{\underline{b}}).$$

Proposition 7.5. *The image of Ψ has finite index.*

Proof. By Lemmas 7.2, 7.3, and 7.4, we need only examine $\mathcal{H}(I, \underline{\underline{i}})$ and $\mathcal{H}(B-E, \underline{\underline{b}})$. Let $\mathcal{H}_0(I, \underline{\underline{i}})$ be the subgroup of $\mathcal{H}(I, \underline{\underline{i}})$ consisting of the elements containing representatives which are the identity on the frontier of I, and recall the subgroup $\mathcal{H}_0(B-E, \underline{\underline{b}})$ of $\mathcal{H}(B-E, \underline{\underline{b}})$ defined shortly before Proposition 6.4. Since the mapping class groups of the square, annulus, arc, and circle are finite, these are finite index subgroups. It suffices to show that $\mathcal{H}_0(I, \underline{\underline{i}}) \times \mathcal{H}_0(B-E, \underline{\underline{b}})$ is contained in the image of Ψ.

Proposition 6.4 shows that the image of the homomorphism $\rho\colon \mathcal{H}_f(S, \underline{\underline{s}}) \to \mathcal{H}(B-E, \underline{\underline{b}})$ contains $\mathcal{H}_0(B-E, \underline{\underline{b}})$, and that there is a homomorphism $s\colon \mathcal{H}_0(B-E, \underline{\underline{b}}) \to \mathcal{H}_f(S, \underline{\underline{s}})$ with ρs equal to the identity. Moreover, each $s(\langle h \rangle)$ can be represented by a diffeomorphism which is the identity on ∂S. So given an element in $\mathcal{H}_0(I, \underline{\underline{i}}) \times \mathcal{H}_0(B-E, \underline{\underline{b}})$, it can be represented on $I \cup S$ by a diffeomorphism which is the identity on the frontier of $I \cup S$. Extending this to M using the identity map on $M - (I \cup S)$ produces an element of $\mathcal{H}(M, \Sigma, \underline{\underline{m}})$ that Ψ carries to the given element. $\square$

Define $\mathcal{K}(M, \Sigma, \underline{\underline{m}})$ to be the kernel of Ψ.

Proposition 7.6. *$\mathcal{K}(M, \Sigma, \underline{\underline{m}})$ is finitely generated and abelian.*

Proof. Let g be a vertical diffeomorphism of S which is isotopic to the identity on $\mathrm{Fr}(S)$. Each such g extends to a diffeomorphism of M which is the identity outside a regular neighborhood of S. The image of $\mathcal{K}(M, \Sigma, \underline{\underline{m}})$ in $\mathcal{H}(S, \underline{\underline{s}})$ is contained in the subgroup of $\mathcal{H}_+^0(S, \underline{\underline{s}})$ consisting of elements representable by maps which are isotopic to the identity on the frontier of S. By Proposition 6.3, $\mathcal{H}_+^0(S, \underline{\underline{s}})$ is finitely generated and abelian, hence so is the image of $\mathcal{K}(M, \Sigma, \underline{\underline{m}})$. Choose a set of generators for the image, and for each an extension to M. Let X_1 be the resulting subset of $\mathcal{H}(M, \Sigma, \underline{\underline{m}})$. These commute with each other and with any diffeomorphism supported in a regular neighborhood of $\mathrm{Fr}(H)$. Choose a set of generators X_2 (Dehn twists about components of $\mathrm{Fr}(H)$) for the elements of $\mathcal{H}(M, \Sigma, \underline{\underline{m}})$ representable by diffeomorphisms supported in a regular neighborhood of $\mathrm{Fr}(H)$. The set $X = X_1 \cup X_2$ is a finite set of commuting elements, and we claim that it generates $\mathcal{K}(M, \Sigma, \underline{\underline{m}})$. Given any element of $\mathcal{K}(M, \Sigma, \underline{\underline{m}})$, we may compose this element with elements in X_1 to assume that its restriction to S is isotopic to the identity. But all elements of $\mathcal{K}(M, \Sigma, \underline{\underline{m}})$ are isotopic to the identity when restricted to each of I, K, T, and H, so when the restriction to S is also isotopic to the identity, the element is a product of elements in X_2. $\square$

Theorem 7.7. *Let M be a Haken manifold and $\underline{\underline{m}}$ a boundary pattern on M whose completion is useful. Then $\mathcal{H}(M, \underline{\underline{m}})$ is virtually torsionfree and almost geometrically finite.*

Proof. If M is a torus bundle over S^1 that admits a *Sol* structure, then by Proposition 4.1.2 of [**26**], $\mathcal{H}(M)$ is finite. Otherwise, by Proposition 7.1, we may work instead with $\mathcal{H}(M,\Sigma,\underline{\underline{m}})$.

First we combine previous results to prove that $\mathcal{H}(M,\underline{\underline{m}})$ is almost geometrically finite. If $M=\Sigma$, then Lemma 6.1 and Corollary 6.6 apply. Suppose that $M\neq\Sigma$. By Lemma 6.1, $\mathcal{H}(I,\underline{\underline{i}})$ is almost geometrically finite. By Proposition 7.5, the image of Ψ is almost geometrically finite. By Proposition 7.6, the kernel of Ψ is finitely generated abelian. By Proposition 3.3, $\mathcal{H}(M,\underline{\underline{m}})$ is almost geometrically finite.

It was proven in [**26**] that $\mathcal{H}(M,\underline{\underline{m}})$ is virtually torsionfree, but we give here a proof that is much simpler.

If $M=\Sigma$, then Lemma 3.5.9 of [**26**] completes the proof, so we assume that $M\neq\Sigma$. Let $(Q,\underline{\underline{q}})$ be obtained as follows. Let $(B_1,\underline{\underline{b_1}})$ be the base surface of $(I,\underline{\underline{i}})$, and let $(B_2,\underline{\underline{b_2}})$ be the base surface of $(S,\underline{\underline{s}})$, and E the image of the exceptional fibers. Remove each component $(X,\underline{\underline{x}})$ from $(B_1,\underline{\underline{b_1}})\cup(B_2,\underline{\underline{b_2}})$ for which $\mathcal{H}(X-E,\underline{\underline{x}})$ is finite, and call the remainder $(Q,\underline{\underline{q}})$. Let

$$\pi\colon R_H\times R_T\times\mathcal{H}(K,\underline{\underline{k}})\times\mathcal{H}(I,\underline{\underline{i}})\times\mathcal{H}(B_2-E,\underline{\underline{b_2}})\longrightarrow\mathcal{H}(Q-E,\underline{\underline{q}})$$

be the projection, and note that π has finite kernel. Put $\Psi'=\pi\Psi$ and $V=(\Psi')^{-1}(F)$ where F is the subgroup of $\mathcal{H}(I,\underline{\underline{i}})$ from Lemma 6.1. Then we have an exact sequence

$$1\to V\longrightarrow\mathcal{H}(M,\underline{\underline{m}})\xrightarrow{\Psi'}\mathcal{H}(Q-E,\underline{\underline{q}})$$

where V is virtually abelian. Let $\mathcal{G}(Q-E,\underline{\underline{q}})$ denote the subgroup of $\mathcal{H}(Q-E,\underline{\underline{q}})$ consisting of the elements that preserve each element of $\underline{\underline{q}}$ and each point of intersection of two elements of $\underline{\underline{q}}$, and let $\mathcal{G}_0(Q-E,\underline{\underline{q}})$ be the corresponding subgroup of $\mathcal{H}_0(Q-E,\underline{\underline{q}})$. Let $\mathcal{H}'(M,\underline{\underline{m}})$ denote the subgroup $(\Psi')^{-1}(\mathcal{G}_0(Q-E,\underline{\underline{q}}))$; as in Proposition 6.4, there is a splitting $s\colon\mathcal{G}_0(Q-E,\underline{\underline{q}})\to\mathcal{H}'(M,\underline{\underline{m}})$ so we have a semidirect product

$$\mathcal{H}'(M,\underline{\underline{m}})=V\circ\mathcal{G}_0(Q-E,\underline{\underline{q}})\ .$$

Let $W=H_1(B_1)\times V$, and consider the *holomorph* of W, $W\circ\mathrm{Aut}(W)$, which is defined to be the semidirect product in which $\mathrm{Aut}(W)$ acts naturally on W. We claim that $\mathrm{Aut}(W)$ is virtually torsionfree. Write W as $W_1\circ T$ where W_1 is free abelian and T is finitely generated. Since T is characteristic, there is a homomorphism $\mathrm{Aut}(W)\to\mathrm{Aut}(W_1)$. It is easily checked that the kernel is a finite group of the form $\mathrm{Hom}(W_1,T)\circ\mathrm{Aut}(T)$, and moreover it splits. Since $\mathrm{Aut}(W_1)\cong\mathrm{GL}(n,\mathbf{Z})$ is virtually torsionfree, it follows that $\mathrm{Aut}(W)$ is virtually torsionfree. Since W is also virtually torsionfree, and W is finitely generated, Lemma 6.8 of [**28**] implies that $W\circ\mathrm{Aut}(W)$ is virtually torsionfree.

The action in the semidirect product $V \circ \mathcal{G}_0(Q-E, \underline{\underline{q}})$ together with the natural action of $\mathcal{H}_0(B_1)$ on $H_1(B_1)$ defines a homomorphism from $\mathcal{G}_0(Q-E, \underline{\underline{q}})$ to $W \circ \mathrm{Aut}(W)$. Since the latter is virtually torsionfree, it suffices to show that the kernel is torsionfree. The kernel is contained in the subgroup of $\mathcal{G}_0(Q-E, \underline{\underline{q}})$ consisting of elements that act trivially on $H_1(Q-E)$. Since we have removed all disc components from $Q-E$, elements of the kernel preserve each component of $Q-E$, so we may assume that $Q-E$ is connected. Let P consist of one point from each arc of $\underline{\underline{q}}$ together with all points that are intersections of two distinct arcs of $\underline{\underline{q}}$. It is easy to see that $\mathcal{G}_0(Q-E, \underline{\underline{q}})$ is isomorphic to a subgroup of finite index in $\mathcal{H}(Q-E \text{ rel } P)$. If P is empty, then it is well-known that the subgroup of $\mathcal{H}(Q-E)$ inducing the identity on homology is torsionfree (see for example [**4**]). Suppose P is not empty. Let $\mathrm{Diff}_1(Q-E)$ denote the subgroup of $\mathrm{Diff}(Q-E)$ consisting of the elements which induce the identity on $H_1(Q-E)$ and preserve each component of $\partial(Q-E)$. There is a fibration

$$\mathrm{Diff}_1(Q-E \text{ rel } P) \to \mathrm{Diff}_1(Q-E) \to \mathrm{imb}(P, \partial(Q-E))$$

where $\mathrm{imb}(P, \partial(Q-E))$ is the connected component of the inclusion in the space of imbeddings $\mathrm{Imb}(P, \partial(Q-E))$. Clearly $\mathrm{imb}(P, \partial(Q-E)) \cong (S^1)^k$, where k is the number of boundary circles of $Q-E$ that contain points of P. From the homotopy exact sequence of this fibration, we have an exact sequence

$$0 \to \pi_1(\mathrm{Diff}_1(Q-E)) \to \mathbf{Z}^k \to \mathcal{H}_1(Q-E \text{ rel } P) \to \mathcal{H}_1(Q-E) \to 1$$

where $\mathcal{H}_1(Q-E \text{ rel } P) = \pi_0(\mathrm{Diff}_1(Q-E \text{ rel } P))$ and $\mathcal{H}_1(Q-E) = \pi_0(\mathrm{Diff}_1(Q-E))$. Now $\pi_1(\mathrm{Diff}_1(Q-E))$ is nontrivial only when $Q-E$ is an annulus and $k=2$ (a disc, a Möbius band, or an annulus with $k=1$ have already been excluded by the definition of $Q-E$), and in this case the generator of $\pi_1(\mathrm{Diff}_1(Q-E))$ is carried to an element of the form $(\pm 1, \pm 1) \in \mathbf{Z} \times \mathbf{Z}$. Therefore we have an exact sequence

$$0 \to \mathbf{Z}^\ell \to \mathcal{H}_1(Q-E \text{ rel } P) \to \mathcal{H}_1(Q-E) \to 1.$$

We have already seen that $\mathcal{H}_1(Q-E)$ is torsionfree, and therefore $\mathcal{H}_1(Q-E \text{ rel } P)$ is torsionfree. □

To deduce some corollaries, it is convenient to introduce a general technique for extending results about $\mathcal{H}(M, \underline{\underline{m}})$ to relative mapping class groups. Assume that $\underline{\underline{m}}$ is complete. Suppose $\underline{\underline{m_1}} \subset \underline{\underline{m}}$. Take a fine triangulation of $|\underline{\underline{m_1}}|$, which includes as vertices the points of $\partial|\underline{\underline{m_1}}|$ that are intersections of three elements of $\underline{\underline{m}}$, and let $\underline{\underline{m_2}}$ be the complete boundary pattern on M consisting of the elements of $\underline{\underline{m}} - \underline{\underline{m_1}}$ and the 2-cells dual to the triangulation of $|\underline{\underline{m_1}}|$. Choose the triangulation so that at every vertex in $|\underline{\underline{m_1}}|$, at least four triangles meet (for example, take the barycentric subdivision of the triangulation of $|\underline{\underline{m_1}}|$ without introducing the barycenters of 1-simplices that lie in the boundary). Then, each dual 2-cell has at least four sides. This ensures the following.

(i) If $\underline{\underline{m}}$ is useful, then $\underline{\underline{m_2}}$ is useful.

(ii) If $\underline{\underline{m}} - \underline{\underline{m_1}}$ consists of the components of $\overline{\partial M - |\underline{\underline{m_1}}|}$, and these components are incompressible, then $\underline{\underline{m_2}}$ is useful.

The important property of $\underline{\underline{m_2}}$ is the following.

Proposition 7.8. *$\mathcal{H}(M, \underline{\underline{m}}$ rel $|\underline{\underline{m_1}}|)$ is isomorphic to a subgroup of finite index in $\mathcal{H}(M, \underline{\underline{m_2}})$.*

Proof. Consider the natural homomorphism from $\mathcal{H}(M, \underline{\underline{m}}$ rel $|\underline{\underline{m_1}}|)$ to $\mathcal{H}(M, \underline{\underline{m_2}})$. We claim this is injective and has image of finite index. Suppose f and g are the identity on $|\underline{\underline{m_1}}|$ and are equivalent in $\mathcal{H}(M, \underline{\underline{m_2}})$. Any admissible isotopy must preserve the intersection of any collection of dual 2-cells, so must preserve the points where three dual 2-cells in W meet. Therefore we may assume the isotopy is relative to the dual 0-cells. Also, it must preserve the intersection of any two dual 2-cells, i.e. the dual 1-cells. By the Alexander trick applied to each dual 1-cell $\times I$, we may make the admissible isotopy to be relative to the intersections of the dual cells. Then, by the Alexander trick applied to each dual 2-cell $\times$ I, we may make the isotopy relative to all of $|\underline{\underline{m_1}}|$. To show the image has finite index, let $\mathcal{H}(M, \underline{\underline{m_2}})$ act on the union of the set of all the dual cells. This defines a homomorphism to a finite permutation group, and by an argument similar to the proof of injectivity one shows that any element in the kernel is admissibly (for $\underline{\underline{m_2}}$) isotopic to the identity on $|\underline{\underline{m_1}}|$, so is in the image. □

Using $\underline{\underline{m_2}}$ we can now deduce the corollaries of the main theorem.

Corollary 7.9. *Let $(M, \underline{\underline{m}})$ be Haken with complete and useful boundary pattern, and let W be a union of elements of $\underline{\underline{m}}$. Then $\mathcal{H}(M, \underline{\underline{m}}$ rel $W)$ is virtually torsionfree and almost geometrically finite.*

Proof. By remark (i) above and Proposition 7.8, there is a complete and useful boundary pattern $\underline{\underline{m_2}}$ on M so that $\mathcal{H}(M, \underline{\underline{m}}$ rel $W)$ is isomorphic to a subgroup of finite index in $\mathcal{H}(M, \underline{\underline{m_2}})$. By Theorem 7.7, $\mathcal{H}(M, \underline{\underline{m_2}})$ is almost geometrically finite, hence so is $\mathcal{H}(M, \underline{\underline{m}}$ rel $W)$. □

Corollary 7.10. *Let M be a Haken 3-manifold and let W be a compact 2-dimensional submanifold of ∂M such that $\partial M - W$ is incompressible. Then $\mathcal{H}(M$ rel $W)$ is virtually torsionfree and almost geometrically finite.*

Proof. Let $\underline{\underline{m_1}}$ be the set of components of W, and let $\underline{\underline{m}}$ be the union of $\underline{\underline{m_1}}$ with the set of components of $\overline{\partial M - W}$. The boundary pattern constucted above $\underline{\underline{m_2}}$ is complete, and by (ii) above it is useful. Since $\mathcal{H}(M$ rel $W) = \mathcal{H}(M, \underline{\underline{m_2}}$ rel $|\underline{\underline{m_1}}|)$, Corollary 7.9 applies. □

8. The Kontsevich Conjecture.

Throughout this section, the symbol $f \cong g$ means that f and g are isotopic relative to the boundary of the manifold on which they are defined. We first isolate a couple of technical steps that will be needed in later arguments.

Lemma 8.1. *Let M be a Haken manifold containing an incompressible surface G, not necessarily connected. Let f and g be two diffeomorphisms of M which are homotopic relative to ∂M. Then $f \cong g$. If f and g agree on G, and the homotopy has trivial trace at a basepoint in each component of G, then they are isotopic relative to $G \cup \partial M$.*

Proof. Replacing f by $g^{-1}f$, we may assume that f is orientation-preserving, g is the identity, and f restricts to the identity on G. As in Theorem II.6.1 of [**22**], the homotopy can be deformed relative to $M \times \partial I \cup \partial M \times I$ to an isotopy. In particular, the traces at basepoints on components of G remain trivial. By Theorem 4.2, we may assume that the isotopy is relative to $G \cup \partial M$. □

Lemma 8.2. *Let M be a compact 3-manifold, each of whose components is Haken with incompressible boundary. Assume that each component M_0 of M has the property that if g is a diffeomorphism from M_0 to itself such that $g^n \cong 1_{M_0}$, then there exists a diffeomorphism h of M_0 such that $g \cong h$ with $h^n = 1_{M_0}$. Then M itself has this property.*

Proof. It suffices to consider the case when g acts transitively on the components of M. Let M_1 be a component and let i be the smallest positive integer such that g^i preserves M_1. Since M_1 has the property, $g^i \cong h$ such that $h^{n/i} = 1_{M_1}$. Let $M_2 = g^{-1}(M_1)$. By isotopy relative to the boundary of M_2, we may change $g|_{M_2}$ to hg^{1-i}, then g has order n on each component of M. □

We will need an extension of Theorem 3 of [**15**].

Proposition 8.3. *Let M be Haken with nonempty incompressible boundary. Assume that each component of ∂M is a torus. If g is a diffeomorphism from M to itself such that $g^n \cong 1_M$, then $g \cong h$ such that $h^n = 1_M$.*

Proof. Consider the characteristic submanifold Σ of $(M, \overline{\underline{\underline{\emptyset}}})$. Since ∂M consists of tori, each component of the frontier of Σ is a torus, and is boundary parallel exactly when it lies in a component of Σ which is a regular neighborhood of a component of ∂M. We induct on the number of components of the frontier of Σ that are not parallel onto ∂M. If there are none, M is either simple or a Seifert fiber space. By Theorem 3 of [**15**] and Lemma 8.1, the theorem is true for M. (In the remainder of the proof, Lemma 8.1 must be used in similar fashion to strengthen conclusions from [**15**], but we will no longer mention these individually.)

We induct on the number of components of the frontier of Σ that are not parallel into ∂M. Since Σ is unique up to isotopy, we may assume that $g(\Sigma)=\Sigma$. Let F be a component of the frontier of Σ such that F is not parallel into ∂M. Let $\widehat{F}=\cup g^i(F)$, a collection of components of the frontier of Σ. By induction and Lemma 8.2, it suffices to show that g^n is isotopic to 1_M relative to $F\cup\partial M$.

Fix an isotopy $H\colon g^n \cong 1_M$. The proof of Lemma 9(ii) of [**15**] contains most of the arguments needed to obtain our conclusion, so we only explain the changes needed. We refer to the notation used there. The first paragraph of the proof (of Case (ii)) in [**15**] is not needed; since $\pi_1(M)$ is centerless the condition $h(\gamma)\simeq\gamma$ holds automatically, as shown by the argument for Lemma 6 of [**15**]. For the next paragraph, we know that M cannot fiber over S^1 with fiber F, since ∂M is nonempty, therefore the first half of the paragraph shows that the restriction of the homotopy to $\widehat{F}\times I$ is homotopic relative to $\widehat{F}\times\partial I$ to a map into $\widehat{F}\times I$. Therefore the homotopy at the start of the third paragraph may also be assume to be relative to ∂M. The remainder of the proof makes some rather delicate adjustments to achieve that the restriction of g to $\widehat{F}$ is periodic and that the homotopy has trivial trace at a basepoint in each component of $\widehat{F}$. Since these changes take place only in a regular neighborhood of $\widehat{F}$, the resulting homotopy is still relative to ∂M. Then Lemma 8.1 yields from this an isotopy relative to $F\cup\partial M$, to complete the inductive step and the proof. □

Theorem 8.4. *Let M be a Haken 3-manifold such that ∂M is nonempty and incompressible. Then $\mathcal{H}(M \operatorname{rel} \partial M)$ is torsionfree.*

Proof. Let $\langle g\rangle\in\mathcal{H}(M \operatorname{rel} \partial M)$ and suppose that for $n>1$, $g^n\cong 1_M$. We must show that $g\cong 1_M$.

Let T be the union of the torus boundary components of M. Suppose first that $T=\partial M$. By Proposition 8.3, $g\cong h$ with $h^n=1_M$. Since h is the identity on ∂M, a theorem of M. H. A. Newman (see Proposition 3.1 of [**21**]) shows that $h=1_M$. Now suppose that T is not empty but $T\neq\partial M$. Form N by gluing two copies of M together along $\partial M-T$, and let $D(g)$ be the diffeomorphism of N defined by taking g on each copy of M. Since $D(g)^n\cong 1_N$, the previous case shows that $D(g)\cong 1_N$. Let $H\colon N\times I\to N$ be an isotopy from $D(g)$ to 1_N. By Lemmas 7.2 and 7.3 of [**34**], H may be deformed to a homotopy that preserves G. Therefore the trace of H at a point in G lies in G. Since $D(g)$ is the identity on G, the trace is a central element of $\pi_1(G)$. Since G is not a torus, the center of $\pi_1(G)$ is trivial. By Theorem 4.2, H may be deformed to an isotopy relative to G. Repeating, we have an isotopy relative to $\partial M-W$, so $g\cong 1_M$. This completes the case when T is not empty.

Now suppose that no component of ∂M is a torus. Let G be a boundary component, and choose an essential simple closed curve γ in G. Let G_1 be a regular neighborhood of γ in G. Let W be $S^1 \times S^1 \times I$, and let G_2 be a regular neighborhood of $S^1 \times \{s_0\} \times \{0\}$ in $S^1 \times S^1 \times \{0\}$ for some $s_0 \in S^1$. Form N by identifying G_1 with G_2 and let G_0 be the incompressible surface in N obtained from G_1 and G_2. Since G_0 is incompressible in M and W, N is Haken. Extend g to a diffeomorphism f of N using the identity on W. The isotopy $g^n \cong 1_M$ extends using the identity on W to an isotopy $f^n \cong 1_N$. Since N has a torus boundary component, the previous case implies that $f \cong 1_N$. By Theorem 4.2, $f \cong 1_N$ relative to G_1, and therefore $g \cong 1_M$. □

Now we will weaken the hypothesis.

Theorem 8.5. *Let M be a Haken 3-manifold and let F be a nonempty compact 2-manifold in ∂M. Then $\mathcal{H}(M \text{ rel } F)$ is torsionfree.*

Proof. Fix $\langle g \rangle \in \mathcal{H}(M \text{ rel } F)$ with g^n isotopic to 1_M relative to F for some $n > 1$. We must prove that g is isotopic to 1_M relative to F.

Let W be the union of the boundary components of M that meet F. On each component X of $\overline{W - F}$, we have $g^n|_X \cong 1_X$. Since ∂X is nonempty, Lemma 1.2 of [**14**] shows that $\mathcal{H}(X \text{ rel } \partial X)$ is torsionfree. Therefore $g|_X \cong 1_X$, so we may change g by isotopy relative to F so that $g|_W = 1_W$. Also, $\pi_1(\text{Diff}(X \text{ rel } \partial X))$ is trivial, so we may assume that g^n is isotopic to 1_M relative to W.

Form N by attaching two copies of M along $\partial M - W$. Then $D(g)^n \cong 1_N$, so by Theorem 8.4, $D(g) \cong 1_N$. As in the proof of Theorem 8.4, the trace of an isotopy from $D(g)$ to 1_N relative to ∂N at each component G of $\partial M - W$ lies in G, so by Theorem 4.2 we may deform the isotopy so that it preserves $\partial M - W$. Therefore g is isotopic to 1_M relative to W and hence relative to F. □

Using Corollary 7.10, we have the following immediate consequence.

Theorem 8.6. *Let M be a Haken 3-manifold with incompressible boundary, and let F be a nonempty compact 2-manifold in ∂M, such that $\partial M - F$ is incompressible. Then $\mathcal{H}(M \text{ rel } F)$ is geometrically finite.*

From this we will obtain the following generalized version of the Kontsevich Conjecture for Haken manifolds.

Theorem 8.7. *Let M be a Haken 3-manifold with incompressible boundary, and let F be a nonempty compact 2-manifold in ∂M such that $\partial M - F$ is incompressible. Then $\text{BDiff}(M \text{ rel } F)$ has the homotopy type of a finite complex.*

Proof. We will show that $\pi_i(\text{Diff}(M \text{ rel } F)) = 0$ for $i \geq 1$. Since $\text{BDiff}(M \text{ rel } F)$ is connected and $\pi_{i+1}(\text{BDiff}(M \text{ rel } F)) \cong \pi_i(\text{Diff}(M \text{ rel } F))$ for $i \geq 1$, this implies that $\text{BDiff}(M \text{ rel } F)$ is a $K(\mathcal{H}(M \text{ rel } F), 1)$-complex, so Theorem 8.6 shows that $\mathrm{BDiff}(M \text{ rel } F)$ has the homotopy type of a finite complex.

By the main theorem of [**12**], the homotopy groups $\pi_i(\mathrm{Diff}(M \text{ rel } \partial M))$ vanish for $i \geq 1$, which gives the assertion when $F = \partial M$. Otherwise, let $W = \overline{\partial M - F}$. Restricting diffeomorphisms to W is the projection map of a fibration

$$\mathrm{Diff}(M \text{ rel } \partial M) \to \mathrm{Diff}(M \text{ rel } F) \to \mathrm{Diff}(W \text{ rel } \partial W).$$

From the homotopy exact sequence of this fibration, we have using [**12**] again, that $\pi_i(\mathrm{Diff}(M \text{ rel } F)) \cong \pi_i(\mathrm{Diff}(W \text{ rel } \partial W)) = 0$ for $i \geq 2$. We also obtain an exact sequence

$$0 \to \pi_1(\mathrm{Diff}(M \text{ rel } F)) \to \pi_1(\mathrm{Diff}(W \text{ rel } \partial W)) \to \mathcal{H}(M \text{ rel } \partial M).$$

No component of W is a 2-sphere, so elements in $\pi_1(\mathrm{Diff}(W \text{ rel } \partial W))$ are classified by their traces (nontrivial elements occur only for tori). Since the traces of an isotopy from 1_M to 1_M at different basepoints are freely homotopic, and F is nonempty, all traces of an element of $\pi_1(\mathrm{Diff}(M \text{ rel } F))$ must be trivial in $\pi_1(M)$. Since W is incompressible, the restriction of an element of $\pi_1(\mathrm{Diff}(M \text{ rel } F))$ to W has trivial trace in each component of W, so is trivial in $\pi_1(\mathrm{Diff}(W \text{ rel } \partial W))$. Therefore $\pi_1(\mathrm{Diff}(M \text{ rel } F))$ is trivial. □

References

[1] A. Borel and J.-P. Serre, *Corners and arithmetic groups*, Comment. Math. Helv., **48** (1973), 436-491.

[2] R. Canary and D. McCullough, *Homotopy equivalences of 3-manifolds and deformation theory of Kleinian groups*, in preparation.

[3] L.S. Charlap and A.T. Vasquez, *Compact flat Riemannian manifolds* III; *the group of affinities*, Amer. J. Math., **95** (1973), 471-494.

[4] H. Farkas and I. Kra, *Riemann Surfaces*, Springer-Verlag Graduate Texts in Mathematics 71, 1980.

[5] A. Fathi, F. Laudenbach and V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque, **66-67** (1979).

[6] A. Gramain, *Le type d'homotopie du groupe des difféomorphismes d'une surface compacte*, Ann. scient. Éc. Norm. Sup., **6**(**4**) (1973), 53-66.

[7] W. Hantzsche and W. Wendt, *Drei dimensionale Euklidische Raumformen*, Math. Ann., **110** (1934), 593-611.

[8] J. Harer, *The virtual cohomological dimension of the mapping class group of an orientable surface*, Invent. Math., **84** (1986), 157-176.

[9] ———, *The cohomology of the moduli space of curves*, in 'Theory of Moduli', ed. E. Sernesi, Springer-Verlag Lecture Notes in Mathematics, **1337** (1988), 138-221.

[10] W. Harvey, *Geometric structure of surface mapping class groups*, in 'Homological Group Theory', ed. C.T.C. Wall, London Math. Soc. Lecture Note Series, **36** (1979), 255-269.

[11] ———, *Boundary structure of the modular group*, in 'Riemann Surfaces and Related Topics: Proceedings of the 1978 Stony Brook Conference', ed. I. Kra and B. Maskit, Annals of Math. Study, **97** (1981), 245-251.

[12] A. Hatcher, *Homeomorphisms of sufficiently large P^2-irreducible 3-manifolds*, Topology, **15** (1976), 343-347.

[13] ———, *On triangulations of surfaces*, Topology Appl., **40** (1991), 189-194.

[14] A. Hatcher and D. McCullough, *Finiteness of classifying spaces of relative diffeomorphism groups of Haken 3-manifolds*, Geometry and Topology, **1** (1997), 91-109.

[15] W. Heil and J. Tollefson, *On Nielsen's theorem for 3-manifolds*, Yokohama Math. J., **35** (1987), 1-20.

[16] W. Jaco and P. Shalen, *Seifert fibered spaces in 3-manifolds*, Mem. Amer. Math. Soc., **220** (1979), 1-192.

[17] K. Johannson, *Homotopy Equivalences of 3-manifolds with Boundary*, Springer-Verlag Lecture Notes in Mathematics, **761** (1979).

[18] Y. Kamishima, K.B. Lee and F. Raymond, *The Seifert construction and its application to infranilmanifolds*, Quart. J. Math., **34** (1983) 433-452.

[19] A. Karrass, A. Pietrowski and D. Solitar, *Finite and infinite cyclic extensions of free groups*, J. Australian Math. Soc., **16** (1972), 458-466.

[20] R. Kirby, *Problems in low-dimensional topology*, in 'Geometric Topology Part II', ed. W. Kazez, AMS/IP Studies in Advanced Mathematics, **2.2** (1997), 35-473.

[21] J.S. Lee, *Almost periodic homeomorphisms and p-adic transformation groups on compact 3-manifolds*, Proc. Amer. Math. Soc., **121** (1994), 267-273.

[22] F. Laudenbach, *Topologie de dimension trois. Homotopie et isotopie*, Astŕisque, **12** (1974), 11-152.

[23] R. Lickorish, *Homeomorphisms of non-orientable two-manifolds*, Proc. Cambridge Phil. Soc., **59** (1963), 307-317.

[24] G.S. McCarty, *Homeotopy groups*, Trans. Amer. Math. Soc., **106** (1963), 293-304.

[25] D. McCullough, *Mappings of reducible 3-manifolds*, in 'Proceedings of the Semester on Topology of the Stefan Banach International Mathematical Center', ed. H. Toruńczyk, Banach Center Publications, Warsaw, (1986), 61-76.

[26] ———, *Virtually geometrically finite mapping class groups of 3-manifolds*, J. Diff. Geom., **33** (1991), 1-65.

[27] ———, *3-Manifolds And Their Mappings*, Global Analysis Center Lecture Note Series 26, Seoul National University, 1995.

[28] D. McCullough and A. Miller, *Symmetric automorphisms of free products*, Mem. Amer. Math. Soc., **582** (1996), 1-97.

[29] P. Orlik, *Seifert Manifolds*, Springer-Verlag Lecture Notes in Mathematics, **291** (1972).

[30] H. Schneebeli, *On virtual properties and group extensions*, Math. Z., **159** (1978), 159-167.

[31] P. Scott, *The geometries of 3-manifolds*, Bull. London Math. Soc., **15** (1983), 401-487.

[32] J. van Mill, *Infinite-Dimensional Topology, Prerequisites and Introduction*, North-Holland Mathematical Library, **43** (1989).

[33] F. Waldhausen, *Eine Klasse von* 3*-dimensionalen Mannigfaltigkeiten* I, II, Invent. Math., **3** (1967), 308-333; **4**(1967), 87-117.

[34] ______, *On irreducible* 3*-manifolds which are sufficiently large*, Ann. of Math., **87** (1968), 56-88.

[35] J. West, *Mapping Hilbert manifolds to ANR's: A solution of a conjecture of Borsuk*, Ann. of Math., **106**(**2**) (1977), 1-18.

[36] B. Zimmermann, *On the Hantzsche-Wendt manifold*, Monatsh. Math., **110** (1990), 321-327.

Received July 14, 1997.

Korea University
Seoul 136-701
Korea
E-mail address: shong@semi.korea.ac.kr

University of Oklahoma
Norman, OK 73019
E-mail address: dmccullough@math.ou.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

CONTINUOUS FAMILIES OF NONNEGATIVE DIVISORS

Guido Lupacciolu and Edgar Lee Stout

In this work it is shown that, under appropriate hypotheses, the multiplicative Cousin problem on complex manifolds admits solutions that depend continuously on parameters.

1. Introduction.

The bulk of this paper is devoted to the study of continuous families of nonnegative divisors on a complex manifold. In particular our work leads to the solution, for a broad class of complex manifolds, of two problems proposed by Stoll [**St**, Problem A and Problem B, p. 155 and pp. 201-202].

If $\mathcal{M}$ is a complex-analytic manifold[1] of complex dimension $N \geq 1$, we denote by $\mathfrak{D}^+(\mathcal{M})$ the set of nonnegative divisors on $\mathcal{M}$. Moreover we denote by $\mathfrak{D}^+_P(\mathcal{M})$ the subset of $\mathfrak{D}^+(\mathcal{M})$ consisting of the principal divisors, i.e., the divisors of holomorphic functions. (It will also be convenient to use the notation that $\mathfrak{D}^+(\mathcal{M};p)$ denotes the set of all nonnegative divisors with support not containing the point $p \in \mathcal{M}$, and that $\mathfrak{D}^+_P(\mathcal{M};p)$ is the set of principal divisors in $\mathfrak{D}^+(\mathcal{M};p)$.) An element $D \in \mathfrak{D}^+(\mathcal{M})$ can be understood as a formal sum $D = \sum m_j V_j$ with each m_j a nonnegative integer and with $\{V_j\}_{j=1,\dots}$ a locally finite family of irreducible complex hypersurfaces in $\mathcal{M}$. It is natural to identify this divisor with the *current of integration over* D, that is, the current defined by

$$D(\alpha) = \langle \alpha, D \rangle = \sum m_j \int_{V_j} \alpha,$$

for each $\mathcal{C}^\infty$ compactly supported $(N-1, N-1)$-form, α, on $\mathcal{M}$. Thus, $\mathfrak{D}^+(\mathcal{M})$ may be considered as a subset of $\mathcal{D}_{1,1}(\mathcal{M}) = \mathcal{D}'^{(1,1)}(\mathcal{M})$, the space of bihomogeneous currents on $\mathcal{M}$ of type $(1,1)$, which is dual to the space $\mathcal{D}^{(N-1,N-1)}(\mathcal{M})$ of compactly supported smooth forms of bidegree $(N-1, N-1)$ on $\mathcal{M}$.

From the point of view of functional analysis there are two natural topologies on $\mathcal{D}'^{(1,1)}(\mathcal{M})$: The *weak* topology* and the *strong topology*. It follows that $\mathfrak{D}^+(\mathcal{M})$ inherits from $\mathcal{D}'^{(1,1)}(\mathcal{M})$ two topologies: The relative topology induced by the weak* topology and the relative topology induced by

[1] In what follows, we assume all our manifolds to be countable at infinity. In the absence of explicit mention to the contrary, they are also assumed to be connected.

the strong topology. Stoll [**St**] introduced a third topology on the space $\mathfrak{D}^+(\mathcal{M})$, a topology particularly suited to the study of the "normal families of nonnegative divisors". In Section 3 of the paper we establish that on $\mathfrak{D}^+(\mathcal{M})$ these three topologies coincide; the unique topology they determine will be seen eventually to be separable and metrizable.

Stoll [**St**, Problem B, p. 155 and p. 202] posed the following problem: *Let $\mathcal{M}$ be a complex-analytic manifold. Suppose that every nonnegative divisor on $\mathcal{M}$ is a principal divisor. Let $\mathfrak{R}$ be a set of nonnegative divisors on $\mathcal{M}$. Does there exist a continuous map $h : \mathfrak{R} \to \mathcal{O}(\mathcal{M})$ such that* $\operatorname{Div} h(D) = D$ *for every $D \in \mathfrak{R}$?*

Stoll [**St**] solved this problem in certain cases involving domains in $\mathbb{C}^N$. In particular, he proved the following result [**St**, Theorems 1.9, 2.25 and 3.6].

Theorem 1.0. *Let Ω be a domain in $\mathbb{C}^N$ that contains the closed unit ball $\bar{\mathbb{B}}_N$. There is a continuous map $h : \mathfrak{D}^+(\Omega; 0) \to \mathcal{O}(\mathbb{B}_N)$ such that for each $D \in \mathfrak{D}^+(\Omega; 0)$,* Div *$h(D) = D|\mathbb{B}_N$.*

We shall exploit this result systematically in the sequel. The appendix to the paper gives a development of the theorem from a point of view somewhat different from that used by Stoll.

The main thrust of the present work is to obtain generalizations of this result of Stoll.

We shall establish in Theorem 6.4 that *if $\mathcal{M}$ is a complex manifold with $H^1(\mathcal{M}, \mathcal{O}) = 0$ and $H^1(\mathcal{M}, \mathbb{Z}) = 0$, then there exists a continuous map $\varsigma : \mathfrak{D}^+_P(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ with* Div *$\varsigma(D) = D$ for all $D \in \mathfrak{D}^+_P(\mathcal{M})$*. There is a strong result in the converse direction, Theorem 6.6.

The proof will depend in an essential way on Theorem 1.0 quoted above and on certain topological methods.

We shall see in Section 7 that the conclusion of Theorem 6.4 can also be drawn if $\mathcal{M}$ is a domain in a Stein manifold that satisfies the geometric condition $H^1(\mathcal{M}, \mathbb{Z}) = 0$.

In Section 8 we show that there is an analogue of Theorem 6.4 in which the role of the space $\mathcal{O}(\mathcal{M})$ is played by the space $\mathcal{L}(\mathcal{M})$ of global sections of the holomorphic line bundle $\mathcal{L}$ over $\mathcal{M}$.

The above results can be reformulated in terms of the solvability of the Second Cousin Problem for nonnegative divisors with continuous dependence on parameters as follows. Given a continuous family $\{D_x : x \in X\}$ in $\mathfrak{D}^+_P(\mathcal{M})$, parametrized by a topological space X, we ask: Does there exist a continuous function $F : X \times \mathcal{M} \to \mathbb{C}$ such that $F(x, \cdot) \in \mathcal{O}(\mathcal{M})$ and $\operatorname{Div} F(x, \cdot) = D_x$ for all $x \in X$?

Consider the continuous map $\mu : X \to \mathfrak{D}^+_P(\mathcal{M})$ defined by $\mu(x) = D_x$, for all $x \in X$. The existence of F amounts to the existence of a continuous map

$\tilde{\mu} : X \to \mathcal{O}(\mathcal{M})$ with $\operatorname{Div} \tilde{\mu}(x) = \mu(x)$ for all $x \in X$, for, if either of them exists, then we may define the other by $F(x, \cdot) = \tilde{\mu}(x)$, for all $x \in X$.

It is well to indicate with a simple example that for the kind of lifting problem we are concerned with, some geometric conditions must be imposed on the manifold in question. This shows up already in the plane.

Set $\Omega = \mathbb{C} \setminus \{0\}$, the punctured plane. Denote by γ the unit circle in $\mathbb{C}$. Define $\psi : [0,1] \to \mathcal{O}(\Omega) \setminus \{0\}$ by

$$\psi(t)(z) = tz + (1-t)\frac{1}{z}.$$

Define $\phi = \operatorname{Div} \circ \psi : [0,1] \to \mathfrak{D}^+(\Omega) = \mathfrak{D}^+_P(\Omega)$. This is continuous and satisfies $\phi(0) = \phi(1)$.

We ask: Does this map lift to a continuous map $\tilde{\phi} : [0,1] \to \mathcal{O}(\Omega)$ that satisfies $\tilde{\phi}(0) = \tilde{\phi}(1)$ and $\operatorname{Div} \circ \tilde{\phi} = \phi$?

Suppose such a $\tilde{\phi}$ to exist. Without loss of generality, $\tilde{\phi}(0) = \psi(0)$. $\big($If not, replace $\tilde{\phi}$ by $\frac{\psi(0)}{\tilde{\phi}(0)}\tilde{\phi}$.$\big)$

Define $\chi(t) = \frac{1}{2\pi i}\int_\gamma d\log\left(\frac{\tilde{\phi}(t)}{\psi(t)}\right)$.

The function χ is an integer that depends continuously on t and so is constant. But compute:

$$\chi(0) = \frac{1}{2\pi i}\int_\gamma d\log\left(\frac{\tilde{\phi}(0)}{\psi(0)}\right) = 0.$$

Also,

$$\begin{aligned}\chi(1) &= \frac{1}{2\pi i}\int_\gamma d\left(\frac{\tilde{\phi}(1)}{\psi(1)}\right)\\ &= \frac{1}{2\pi i}\int_\gamma d\log\left(\frac{\psi(0)}{\psi(1)}\right)\\ &= \frac{1}{2\pi i}\int_\gamma d\log z^{-2} = -2.\end{aligned}$$

Contradiction.

This simple example shows that to obtain lifting theorems of the kind we are concerned with, it is necessary to impose *some* geometric condition on the manifolds under consideration.

Stoll [**St**, Problem A, p. 155 and pp. 201–202] posed also the following problem:

Let $\mathcal{M}$ be a complex-analytic manifold. Suppose that every nonnegative divisor on $\mathcal{M}$ is a principal divisor. Let $\{D_\lambda\}_{\lambda\in\Lambda}$ be a normal family of nonnegative divisors on $\mathcal{M}$. Does there exist a normal family $\{h_\lambda\}_{\lambda\in\Lambda}$ of *holomorphic functions on $\mathcal{M}$ and a compact subset K of $\mathcal{M}$ such that, for each $\lambda \in \Lambda$, $\operatorname{Div} h_\lambda(D_\lambda) = D_\lambda$ and $h_\lambda(a_\lambda) = 1$ for some $a_\lambda \in K$?*

Stoll [**St**] solved this problem in the case that $\mathcal{M} = \mathbb{C}^N$.

In Section 9 we show that if there is a continuous map $\varsigma : \mathfrak{D}^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ with $\operatorname{Div}\circ\varsigma = id$, then the problem has a positive solution.

Remark. Our motivation for undertaking this work was the desire to apply some of the results obtained here to the study of certain hulls that generalize polynomially convex hulls. In addition to the well known polynomially convex hulls, another collection of hulls has been introduced by Basner [**Bas**]. Given a compact subset X of $\mathbb{C}^N$ there is a hull $h_q(X)$ defined for each integer q in the range $1 \leq q \leq N-1$. The hull $h_q(X)$ is defined in terms of polynomial mappings from $\mathbb{C}^N$ to $\mathbb{C}^q$. By using the results of the present paper, we show that the hull $h_q(X)$ can be described in terms of continuous families of polynomial maps from $\mathbb{C}^N$ to $\mathbb{C}^{q+1}$. This work will be published in a subsequent paper.

We shall need to use a selection theorem proved by Michael [**Mi**]: *If E and F are Fréchet spaces and if $T : E \to F$ is a surjective continuous linear transformation, then there is a continuous map $\varsigma : F \to E$ such that $T \circ \varsigma$ is the identity map on F.* In general the selection map ς cannot be chosen to be linear; it can always be chosen to be homogeneous. A perspicuous proof of the result is given in [**Ru1**].

Given an indexed family of sets $\{S_\alpha\}_{\alpha\in A}$, we understand $\{S_{\alpha\beta}\}_{\alpha,\beta\in A}$ to be the family of intersections $S_{\alpha\beta} = S_\alpha \cap S_\beta$. Similar notation will be used for triple, quadruple,..., intersections.

We shall also use the notation that $\mathbb{C}^*$ denotes the collection of nonzero complex numbers and that $\mathcal{C}(X)$ denotes the space of continuous $\mathbb{C}$-valued functions on the space X.

A referee has drawn our attention to the papers of McGrath [**MG**] and Siu [**Si**]. McGrath's thesis extends Stoll's results to the setting of polycylinders in $\mathbb{C}^N$. In Siu's paper further results are obtained that extend those of Stoll and that are closely related to the present work.

It should be remarked that the possibility of solving the First Cousin Problem with continuous dependence on real parameters is essentially due to Oka [**Ok**]; it was discussed in [**Na**].

2. The Topologies on $\mathfrak{D}^+(\mathcal{M})$.

The space $\mathfrak{D}^+(\mathcal{M})$ has three natural topologies, two with functional-analytic roots, the other based in function theory.

The first of these is the *weak* topology* in which a net $\{D_\alpha\}_{\alpha\in A}$ in $\mathfrak{D}^+(\mathcal{M})$ converges to $D_o \in \mathfrak{D}^+(\mathcal{M})$ if and only if for every compactly supported smooth form ϑ of bidegree $(N-1, N-1)$ on $\mathcal{M}$

$$\lim_{\alpha \in A} \int_{D_\alpha} \vartheta = \int_{D_o} \vartheta.$$

The second functional-analytic topology is the *strong topology* in which a net $\{D_\alpha\}_{\alpha \in A}$ in $\mathfrak{D}^+(\mathcal{M})$ converges to $D_o \in \mathfrak{D}^+(\mathcal{M})$ if and only if it converges in the weak* sense and if, moreover, the convergence is uniform on *bounded* sets in the space $\mathcal{D}^{N-1,N-1}(\mathcal{M})$.

Recall that if $\mathcal{D}(\mathcal{M})$ denotes the space of compactly supported functions on $\mathcal{M}$, then a subset $\mathcal{B}$ of $\mathcal{D}(\mathcal{M})$ is bounded if there is a fixed compact set K in $\mathcal{M}$ with $\mathrm{supp} f \subset K$ for all $f \in \mathcal{B}$. It is required, moreover that the set K be decomposed into a union of closed subsets $K_1, ..., K_r$, with each K_j contained in an open set U_j on which there are global smooth coordinates. For each j, there is a sequence of positive constants $\{k_\nu\}_{\nu=1,2,...}$ with the property that for each $f \in \mathcal{B}$ and each $j = 1, ..., r$, the derivatives of order less than ν of f with respect to the coordinates in U_j are bounded uniformly on K_j by k_ν. This gives rise to the notion of bounded set in the space of forms $\mathcal{D}^{N-1,N-1}(\mathcal{M})$.

The third topology we shall consider is that introduced in function-theoretic terms by Stoll [**St**]. This topology is defined by the condition that a net $\{D_\alpha\}_{\alpha \in A}$ in $\mathfrak{D}^+(\mathcal{M})$ converges to $D_o \in \mathfrak{D}^+(\mathcal{M})$ if and only if there is an open cover $\mathcal{V} = \{V_j\}_{j=1,...}$ of $\mathcal{M}$ such that for each $\alpha \in A$, there is an $f_{j\alpha} \in \mathcal{O}(V_j)$ such that $\mathrm{Div} f_{j\alpha} = D_\alpha | V_j$ and such that $f_{j\alpha}$ converges uniformly on compacta in V_j to $f_{j_o} \in \mathcal{O}(V_j)$, f_{j_o} a function with divisor $D_o | V_j$. Stoll verifies that this prescription does specify a topology on the space $\mathfrak{D}^+(\mathcal{M})$.

The principal goal of the present section is to show the equivalence of these three topologies. In this connection, certain results are immediate. It is plain that strong convergence implies convergence in the weak* sense. It is less evident, but essentially known, that convergence in the sense of Stoll's topology implies strong convergence. This is contained in a theorem of Andreotti and Norguet [**AN**] to the effect that *if $\mathcal{N}$ is a connected complex manifold, then the map that associates to a nonzero-function $f \in \mathcal{O}(\mathcal{N})$ its divisor, viewed as a current, is continuous when the space $\mathcal{O}(\mathcal{N})$ is endowed with the usual topology of uniform convergence on compacta and the space of currents is endowed with its strong topology, viewed as the dual of the topological vector space $\mathcal{D}^{N-1,N-1}(\mathcal{N})$ of compactly supported $(N-1, N-1)$-forms on $\mathcal{N}$.*

To complete the proof of the equivalence of the three topologies on $\mathfrak{D}^+(\mathcal{M})$, it suffices to show that on this space, convergence in the weak* sense implies convergence in the sense of the topology of Stoll.

This depends on Stoll's characterization of normal families in the space $\mathfrak{D}^+(\mathcal{M})$. Recall that, by definition, a subset of $\mathfrak{D}^+(\mathcal{M})$ is a normal family if it is a relatively compact subset, when the space of divisors is endowed with the topology introduced by Stoll.

For normal families, Stoll [**St**] gives the following characterization: *A subset $\mathfrak{N}$ of $\mathfrak{D}^+(\mathcal{M})$ is relatively compact with respect to to Stoll's topology if and only if $\mathfrak{N}$ is bounded on every compact subset $K \subset \mathcal{M}$ in the sense that, for a fixed Hermitian metric on $\mathcal{M}$ with associated fundamental form ω, there is a positive constant L_K such that for each $D \in \mathfrak{N}$, the area of D in K given by $D(\chi_K \omega^{N-1}) = \frac{1}{(N-1)!} \int_D \chi_K \omega^{N-1}$ is not more than L_K.* (Here χ_K denotes the characteristic function of the set K.)

Fix now a Hermitian metric on $\mathcal{M}$, and let ω denote its associated fundamental form. Let $\{D_\lambda\}_{\lambda \in \Lambda}$ be a net in $\mathfrak{D}^+(\mathcal{M})$ that converges in the relative weak* topology to $D_o \in \mathfrak{D}^+(\mathcal{M})$. Fix a relatively compact open set U in $\mathcal{M}$, and let χ be a compactly supported $\mathcal{C}^\infty$ function on $\mathcal{M}$ that is indentically one on a neighborhood of $\bar{U}$ and is everywhere nonnegative.

As the net converges in the weak* sense, there is a constant L large enough that for some $\lambda_o \in \Lambda$ if $\lambda > \lambda_o$, then

$$\int_{D_\lambda} \chi \omega^{N-1} < L.$$

By the characterization of normal families quoted above, it follows that the family $\{D_\lambda | U : \lambda > \lambda_o\}$ is a normal family. Accordingly, the net $\{D_\lambda | U\}_{\lambda > \lambda_o}$ has a cluster point, say $\tilde{D}$, in $\mathfrak{D}^+(U)$ with respect to Stoll's topology. As convergence in the sense of Stoll's topology implies convergence in the weak* sense, $\tilde{D}$ can only be the limit $D_o | U$. (Stoll proves that the topology he introduces satisfies the Hausdorff separation axiom.) This implies that the net $\{D_\lambda | U\}_{\lambda \in \Lambda}$ converges to $D_o | U$ in the sense of Stoll's topology.

As the open, relatively compact subsets of $\mathcal{M}$ constitute an open cover for $\mathcal{M}$, it follows, as we wished, that the initial net $\{D_\lambda\}_{\lambda \in \Lambda}$ converges to D_o in the sense of Stoll's topology.

We have now reached the desired conclusion that *the three naturally defined topologies on the space $\mathfrak{D}^+(\mathcal{M})$ of nonnegative divisors on $\mathcal{M}$ coincide.* In the sequel, we shall speak simply of *the* topology on $\mathfrak{D}^+(\mathcal{M})$.

The proof just given of the equivalence of the three topologies on the space $\mathfrak{D}^+(\mathcal{M})$ depends explicitly on the mechanism developed in [**St**]. It is of interest, and of some importance at a later point in this work, that there is a rather simple, direct proof of the equivalence of the weak* and strong topologies on this space. It runs as follows.

What is to be shown is that a net $\{D_\iota\}_{\iota \in I}$ converges to $D_o \in \mathfrak{D}^+(\mathcal{M})$ in the relative weak* topology if and only if it converges to D_o in the relative strong topology.

That strong convergence implies weak* convergence is evident.

Conversely, we consider a net $\{D_\iota\}_{\iota\in I}$ in $\mathfrak{D}^+(\mathcal{M})$ that converges in the weak* sense to D_o, and we show it to converge strongly.

To this end, consider a bounded set $\mathcal{B} \subset \mathfrak{D}^{N-1,N-1}(\mathcal{M})$. As $\mathcal{B}$ is bounded, there is a compact set $X \subset \mathcal{M}$ such that each $\alpha \in \mathcal{B}$ has support in X. If χ is a smooth function on $\mathcal{M}$, then the set $\chi\mathcal{B}$ defined by

$$\chi\mathcal{B} = \{\chi\alpha : \alpha \in \mathcal{B}\}$$

is a bounded set in $\mathfrak{D}^{N-1,N-1}(\mathcal{M})$. If $\chi_1 + \ldots + \chi_q = 1$, then

$$\mathcal{B} = \chi_1\mathcal{B} + \ldots + \chi_q\mathcal{B}.$$

If $\{D_\iota\}_{\iota\in I}$ converges uniformly on each $\chi_k\mathcal{B}$ to χD_o, then $\{D_\iota\}_{\iota\in I}$ converges uniformly on $\mathcal{B}$ to D_o. This remark, coupled with the existence of smooth partitions of unity on $\mathcal{M}$, permits us to localize our problem: We assume from here on that $\mathcal{M} = \mathbb{B}_N(2)$, the open ball of radius 2 centered in the origin in $\mathbb{C}^N$, and that $X = \overline{\mathbb{B}}_N$, the closed unit ball.

Since $\mathcal{B}$ is bounded, there exist positive constants k_r, $r = 0, 1, \ldots$, such that for each $\alpha \in \mathcal{B}$, if

$$\alpha = \sum_{|I|,|J|=N-1} \alpha_{IJ} dz^I \wedge d\bar{z}^J,$$

then

$$\left|\frac{\partial^{|P|+|Q|}\alpha_{IJ}}{\partial z^P \partial \bar{z}^Q}(z)\right| \le k_{|P|+|Q|}$$

for all $z \in \mathbb{B}_N(2)$ and all multi-indices P, Q.

Denote by χ a real-valued nonnegative smooth function on $\mathbb{C}^N$ with $\chi = 1$ on a neighborhood of $\overline{\mathbb{B}}_N(1)$ and with $\chi(z) = 0$ if $|z| > \frac{3}{2}$. Let ω be the Kähler form $\frac{i}{2}\sum_{j=1}^N dz_j \wedge d\bar{z}_j$.

By hypothesis

$$\lim_\iota D_\iota(\chi\omega^p) = D_o(\chi\omega^p). \tag{1}$$

For each $\iota \in I$ write $D_\iota = \sum_j m_{\iota j} V_{\iota j}$ for suitable positive integers $m_{\iota j}$ and suitable irreducible complex hypersurfaces $V_{\iota j}$ in $\mathbb{B}_N(2)$. Then (1) implies the existence of an index $\iota_o \in I$ such that for a positive constant C_o and for all multi-indices I, J with $|I| = |J| = N - 1$

$$\left|\sum_j m_{\iota j} \int_{V_{\iota j}\cap\overline{\mathbb{B}}_N} dz^I \wedge \bar{z}^J\right| \le C_o \tag{2}$$

provided $\iota_o \le \iota$.

Let $\alpha = \sum_{|I|,|J|=N-1} \alpha_{IJ} dz^I \wedge d\bar{z}^J \in \mathcal{B}$ and compute: For $\iota_o \leq \iota$ we have

(3)

$$\begin{aligned}
&|D_\iota(\alpha) - D_o(\alpha)| \\
&= \Bigg| \sum_j m_{\iota j} \int_{V_{\iota j}} \sum_{|I|=|J|=N-1} \alpha_{IJ} dz^I \wedge d\bar{z}^J \\
&\quad - \sum_k m_{ok} \int_{V_{ok}} \sum_{|I|=|J|=N-1} \alpha_{IJ} dz^I \wedge d\bar{z}^J \Bigg| \\
&\leq \sum_{|I|=|J|=N-1} \left| \sum_j m_{\iota j} \int_{V_{\iota j}} \alpha_{IJ} dz^I \wedge d\bar{z}^J - \sum_k m_{ok} \int_{V_{ok}} \alpha_{IJ} dz^I \wedge d\bar{z}^K \right|.
\end{aligned}$$

For each $\iota \in \mathcal{I}$ and for all multi-indices I and J, define a measure $\mu_{\iota IJ}$ on $\mathbb{B}_N(2)$ by

$$\int g d\mu_{\iota IJ} = \sum_j m_{\iota j} \int_{V_{\iota j} \cap \overline{\mathbb{B}}_N} g dz^I \wedge d\bar{z}^J,$$

for $g \in \mathcal{C}(\overline{\mathbb{B}}_N)$ and define μ_{oIJ}, associated with D_o in a similar way. By (2), $\|\mu_{\iota IJ}\|$ and $\|\mu_{oIJ}\|$ are bounded uniformly in $\iota, \iota_o < \iota$, say by the constant C_1. Let $\lambda = 8N$ and define $\psi_{\iota IJ}(\zeta)$ for $\zeta \in \mathbb{C}^N$ by

$$\psi_{\iota IJ}(\zeta) = \gamma \int |z - \zeta|^{2\lambda - 4N} \log |z - \zeta| d\mu_{\iota IJ}(z).$$

Define ψ_{oIJ} similarly. With a suitable choice of constant γ, these functions satisfy the equations

$$\Delta^\lambda \psi_{\iota IJ} = \mu_{\iota IJ}, \quad \text{and} \quad \Delta^\lambda \psi_{oIJ} = \mu_{oIJ}$$

in the sense of distributions. See [**GS**, p. 282]. (Here Δ denotes the Laplacian on $\mathbb{R}^{2N} = \mathbb{C}^N$.)

We may write

$$\begin{aligned}
&\left| \sum_j m_{\iota j} \int_{V_{\iota j}} \alpha_{IJ} dz^I \wedge d\bar{z}^I - \sum_k m_{ok} \int_{V_{ok}} \alpha_{IJ} dz^I \wedge d\bar{z}^J \right| \\
&= \left| \int_{\overline{\mathbb{B}}_N} \alpha_{IJ} d(\mu_{\iota IJ} - \mu_{oIJ}) \right| \\
&= \left| \int_{\overline{\mathbb{B}}_N} \Delta^\lambda \alpha_{IJ}(\zeta)(\psi_{\iota IJ}(\zeta) - \psi_{oIJ}(\zeta)) \omega^N(\zeta) \right| \\
&\leq M \int_{\overline{\mathbb{B}}_N} |\psi_{\iota IJ}(\zeta) - \psi_{oIJ}(\zeta)| \omega^N(\zeta)
\end{aligned}$$

for a suitable positive constant M.

Call the last integral $L_{\iota IJ}$.

The definition of the $\psi_{\iota IJ}$'s shows that they are uniformly bounded and that they satisfy a uniform Lipschitz condition on $\mathbb{B}_N(2)$. Moreover, the measures $\mu_{\iota IJ}$ converge in the weak* sense (in the space of measures) to μ_{oIJ}. Thus, $\psi_{\iota IJ}(\zeta) \to \psi_{oIJ}(\zeta)$ for each $\zeta \in \mathbb{B}_N(2)$.

From this it follows that $\lim_\iota L_{\iota IJ} = 0$: The numbers $L_{\iota IJ}$ are uniformly bounded, so the net $\{L_{\iota IJ}\}_\iota$ has cluster points. Let t_o be such a cluster point, and let $\{L_{\iota_\beta IJ}\}_{\beta \in B}$ be a subnet of such that $\lim_\beta L_{\iota_\beta IJ} = t_o$.

The net $\{\psi_{\iota_\beta IJ} - \psi_{oIJ}\}_{\beta \in B}$ is a net of continuous functions that are uniformly bounded and uniformly equicontinuous. Consequently, there is a subnet, which we may suppose to be $\{\psi_{\iota_\beta IJ} - \psi_{oIJ}\}_{\beta \in B}$ itself, that converges *uniformly* on $\overline{\mathbb{B}}_N$. As $\psi_{\iota IJ} \to \psi_{oIJ}$ pointwise, the limit can only be zero. As the convergence is uniform, we have

$$\lim_\beta L_{\iota_\beta IJ} = 0.$$

Consequently, $t_o = 0$. This means that $\lim_\iota L_{\iota IJ} = 0$.

The proposition is proved.

It is worth noting that this argument applies, *mutatis mutandis,* to yield the corresponding result for the space of holomorphic p-chains on M.

3. The Topology of $\mathfrak{D}^+(\mathcal{M})$.

We have now to treat certain aspects of $\mathfrak{D}^+(\mathcal{M})$ and $\mathfrak{D}^+_P(\mathcal{M})$ as topological spaces, the topology being that discussed in the preceding section.

Lemma 3.1. *If $p \in \mathcal{M}$, then the set of $D \in \mathfrak{D}^+(\mathcal{M})$ with $p \notin \mathrm{supp} D$ is open in $\mathfrak{D}^+(\mathcal{M})$.*

Proof. Fix a divisor $D_o \in \mathfrak{D}^+(\mathcal{M})$ with $p \notin \mathrm{supp} D_o$. Let ω be the fundamental form for some Hermitian metric on $\mathcal{M}$. Fix neighborhoods $U'' \Subset U' \Subset U$ of p with $U \cap \mathrm{supp} D_o = \emptyset$. Let χ be a nonnegative $\mathcal{C}^\infty$ function on $\mathcal{M}$ with χ identically one on U' and χ identically zero on $\mathcal{M} \setminus U$. There is a constant $C > 0$ small enough that if $D \in \mathfrak{D}^+(\mathcal{M})$ and $\mathrm{supp} D \cap U'' \neq \emptyset$, then $\int_{D_o} \chi \omega^{N-1} > C$. It follows that the set

$$\left\{ D \in \mathfrak{D}^+(\mathcal{M}) : \int_D \chi \omega^{N-1} < C/2 \right\}$$

is a (weak*) neighborhood V of D_o with the property that if $D \in V$, then $\mathrm{supp} D$ does not contain p. The lemma is proved.

Lemma 3.2. *Let p be a fixed point in the complex manifold $\mathcal{M}$. There exists a countable family $\{\varphi_j\}_{j \in \mathbb{N}}$ of maps from the unit ball $\mathbb{B}_N$ into $\mathcal{M}$ with the following properties:*

i. *Each* φ_j *carries* $\mathbb{B}_N$ *biholomorphically onto a domain* Ω_j *in* $\mathcal{M}$.
ii. *Each* φ_j *carries* $0 \in \mathbb{B}_N$ *to the point* p.
iii. *For suitable* $\rho_j \in (0,1)$, $\mathcal{M} = \cup\{\varphi(\rho_j \mathbb{B}_N) : j = 1, \dots\}$.

Proof. This depends on knowing that every pair of points in a complex manifold is contained in a biholomorphic copy of the ball.

Granted this, for each point $q \in \mathcal{M}$, fix a biholomorphic map ψ_q from $\mathbb{B}_N$ onto a domain in $\mathcal{M}$ that carries the origin to the point p and the range of which contains the point q. If $\rho_q \in (0,1)$ is large enough, then the point q will lie in the set $\psi_q(\rho_q \mathbb{B}_N)$.

Let $\{K_j\}_{j\in\mathbb{N}}$ be a sequence of compact sets in $\mathcal{M}$ with union $\mathcal{M}$. A finite number of the sets $\psi_j(\rho_j \mathbb{B}_N)$ will cover a given K in the sequence, so the result follows.

To realize that it is possible to find a biholomorphic copy of the ball that contains a given pair of points in an arbitrary complex manifold, consider points p and q in $\mathcal{M}$. There is a real-analytic arc in $\mathcal{M}$ that contains both p and q. Then as a compact (real) line segment in $\mathbb{C}^N$ has a neighborhood basis that consists of biholomorphic copies of the ball, we see that the desired ball exists. The lemma is proved.

We next prove that the space $\mathfrak{D}^+(\mathcal{M};p)$ is metrizable.

Fix a point $p \in \mathcal{M}$. Let $\Omega_j, j \in \mathbb{N}$, be biholomorphic copies of $\mathbb{B}_N$ in $\mathcal{M}$ as introduced in the preceding lemma. Let Ω_j' be the corresponding concentric balls that cover $\mathcal{M}$.

There are restriction maps

$$r : \mathfrak{D}^+(\mathcal{M};p) \to \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j;p)$$

and

$$r' : \mathfrak{D}^+(\mathcal{M};p) \to \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j';p)$$

given by

$$r(D) = \{D|\Omega_j : j \in \mathbb{N}\} \in \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j;p)$$

and

$$r'(D) = \{D|\Omega_j' : j \in \mathbb{N}\} \in \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j';p).$$

In addition, there is a restriction map

$$\tilde{r} : \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j;p) \to \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j';p)$$

defined in the evident way. These maps are all continuous when the product spaces are endowed with the product topologies. Plainly $\tilde{r} \circ r = r'$. Notice that the map r is a homeomorphism from $\mathfrak{D}^+(\mathcal{M};p)$ onto the closed subset

$$\{\{D_j : j \in \mathbb{N}\} \in \Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j;p) : D_j|\Omega_j \cap \Omega_k = D_k|\Omega_j \cap \Omega_k\}$$

of $\Pi_{j\in\mathbb{N}} \mathfrak{D}^+(\Omega_j;p)$. Similarly, r' is a homeomorphism onto its range.

Consider now the map

$$\sigma_j : \mathfrak{D}^+(\Omega_j;p) \to \mathcal{O}(\Omega_j')$$

given by Stoll that has the property that $\mathrm{Div} \circ \sigma_j(D) = D|\Omega'_j$. These maps taken together give a map

$$\sigma : \Pi_{j\in\mathbb{N}}\mathfrak{D}^+(\Omega_j;p) \to \oplus_{j\in\mathbb{N}}\mathcal{O}(\Omega'_j)$$

given by

$$\sigma(\{D_j : j \in \mathbb{N}\}) = \{\sigma_j(D_j) : j \in \mathbb{N}\}.$$

Also let

$$\delta : \oplus_{j\in\mathbb{N}}\mathcal{O}(\Omega_j) \to \Pi_{j\in\mathbb{N}}\mathfrak{D}^+(\Omega_j)$$

and

$$\delta' : \oplus_{j\in\mathbb{N}}\mathcal{O}(\Omega'_j) \to \Pi_{j\in\mathbb{N}}\mathfrak{D}^+(\Omega'_j)$$

be the divisor maps

$$\delta(\{f_j : j \in \mathbb{N}\}) = \{\mathrm{Div} f_j : j \in \mathbb{N}\}$$

and with δ' defined in a similar way. These maps are continuous.

The map $\delta' \circ \sigma$ carries $r\mathfrak{D}^+(\mathcal{M};p)$ onto $r'\mathfrak{D}^+(\mathcal{M};p)$.

Finally, notice that if $D \in \mathfrak{D}^+(\mathcal{M};p)$, then ${r'}^{-1} \circ \delta' \circ \sigma \circ r(D)$ is simply D. This implies that $\sigma \circ r$ is a homeomorphism of $\mathfrak{D}^+(\mathcal{M};p)$ onto a subset of the metrizable space $\Pi_{j\in\mathbb{N}}\mathcal{O}(\Omega'_j)$.

We have reached the following conclusion:

Proposition 3.3. *The space $\mathfrak{D}^+(\mathcal{M};p)$ is metrizable.*

Consequently, $\mathfrak{D}^+(\mathcal{M};p)$ and all its topological subspaces are paracompact.

4. Lifts of $\varphi : X \to \mathfrak{D}^+_P(\mathcal{M};p)$.

The main result of this section is the following intermediate proposition, which will serve as a stepping stone to more general results.

Proposition 4.1. *Let $\mathcal{M}$ be a complex manifold with $H^1(\mathcal{M},\mathcal{O}) = 0$. Let $p \in \mathcal{M}$. If X is a connected paracompact space with $H^1(X,\mathbb{Z}) = 0$, then given a continuous map $\psi : X \to \mathfrak{D}^+_P(\mathcal{M};p)$, there is a continuous map $\tilde{\psi} : X \to \mathcal{O}(\mathcal{M})$ with* $\mathrm{Div} \circ \tilde{\psi} = \psi$.

Proof. Again introduce a sequence $\{\Omega_j\}_{j\in\mathbb{N}}$ of (biholomorphic copies of) balls in $\mathcal{M}$ and concentric balls Ω'_j , all centered at the point $p \in \mathcal{M}$ and with $\cup\Omega'_j = \mathcal{M}$. Let $\mathcal{U} = \{U_j\}_{j\in\mathbb{N}}$ be a locally finite refinement of the cover $\{\Omega'_j\}_{j\in\mathbb{N}}$ of $\mathcal{M}$ by connected domains U_j with the property that the intersections U_{jk} are all contractible. That such covers exist is shown in [**BT**, p. 42]. Denote by $\tau : \mathbb{N} \to \mathbb{N}$ a refining map so that for each j, $U_j \subset \Omega'_{\tau(j)}$.

Also, let $\sigma_j : \mathfrak{D}^+(\Omega_j;p) \to \mathcal{O}(\Omega'_j)$, be the map used above that satisfies $\mathrm{Div} \circ \sigma_j(D) = D|\Omega'_j$ for all $D \in \mathfrak{D}^+(\Omega_j;p)$. Define $r_j : \mathfrak{D}^+(\mathcal{M};p) \to$

$\mathfrak{D}^+(\Omega_j; p)$ to be the restriction map so that $r_j(D) = D|\Omega'_j$. The map r_j is continuous and satisfies

$$\mathrm{Div} \circ \sigma_j \circ r_j(D) = D|\Omega'_j$$

for all j.

For each $j = 1, \dots,$ let $F_j : \mathfrak{D}^+_P(\mathcal{M}) \times \Omega'_j \to \mathbb{C}$ be defined by

$$F_j(D, z) = (\sigma_j \circ r_j(D))(z).$$

These functions have the following properties: i. $F_j \in \mathcal{C}(\mathfrak{D}^+_P(\mathcal{M}) \times \Omega'_j)$, ii. $F_j(D, \cdot) \in \mathcal{O}(\Omega'_j)$ for each j, and iii. $\mathrm{Div}F_j(D, \cdot)|\Omega'_{jk} = \mathrm{Div}F_k(D, \cdot)|\Omega'_{jk}$ for all j, k.

Define $f_{jk} : \mathfrak{D}^+_P(\mathcal{M}; p) \times U_{jk} \to \mathbb{C}$ by

$$f_{jk} = (F_{\tau(j)}|U_{jk})(F_{\tau(k)}|U_{jk})^{-1}.$$

This is a continuous zero-free function with the property that for all D, $f_{jk}(D, \cdot)$ is holomorphic on Ω'_{jk} for each choice of j, k.

With $\psi : X \to \mathfrak{D}^+_P(\mathcal{M}, p)$ as in the theorem, let the map $\psi_{jk} : X \times U_{jk} \to \mathbb{C}$ be given for each choice of j and k by

$$\psi_{jk}(x, z) = f_{jk}(\psi(x), z);$$

it is continuous and zero-free. Moreover, for each $x \in X$, the function $\psi_{jk}(x, \cdot)$ is holomorphic on U_{jk}.

By hypothesis, $H^1(X, \mathbb{Z}) = 0$, so the group $H^1(X \times U_{jk}, \mathbb{Z})$ also vanishes. Accordingly, the function ψ_{jk} has a continuous logarithm, to be denoted by λ_{jk}, on $X \times U_{jk}$. For each $x \in X$, $\lambda_{jk}(x, \cdot)$ is holomorphic on U_{jk}.

Define $c_{\alpha\beta\gamma}$ on $X \times U_{\alpha\beta\gamma}$ by

$$c_{\alpha\beta\gamma} = \lambda_{\beta\gamma} - \lambda_{\alpha\gamma} + \lambda_{\alpha\beta}.$$

This is some value of log 1. The data $\{c_{\alpha\beta\gamma}\}_{\alpha,\beta,\gamma\in\mathbb{N}}$ constitute a 2-cocycle for the covering $\mathcal{U}_X = \{X \times U_j\}_{j\in\mathbb{N}}$ of $X \times \mathcal{M}$ with values in the group $2\pi i\mathbb{Z}$. For fixed $x \in X$, this cocycle defines the Chern class of the set $\{\psi_{jk}(x, \cdot)\}_{j,k\in\mathbb{N}}$ of Cousin II data on $\mathcal{M}$. As the associated divisor is principal, this cocycle is a coboundary: There are integers $\{n_{\alpha\beta}\}_{\alpha,\beta\in\mathbb{N}}$ such that

$$c_{\alpha\beta\gamma} = 2\pi i(n_{\beta\gamma} - n_{\alpha\gamma} + n_{\alpha\beta}).$$

As X is connected, continuity implies that for each fixed $z \in \mathcal{M}$, $c_{\alpha\beta\gamma}$ is independent of x.

Given $j, k \in \mathbb{N}$, define a continuous map $\lambda'_{jk} : X \to \mathcal{O}(U_{jk})$ by

$$\lambda'_{jk} = \lambda_{jk} - n_{jk}.$$

For each $x \in X$, the family $\lambda'(x) = \{\lambda'_{jk}(x)\}_{jk\in\mathbb{N}}$ defines a 1-cocycle for the covering $\mathcal{U}$ with values in the sheaf $\mathcal{O} = \mathcal{O}_{\mathcal{M}}$, that is, an element of the space $Z^1(\mathcal{U}, \mathcal{O})$. Thus, we have a continuous map $\lambda' : X \to Z^1(\mathcal{U}, \mathcal{O})$, given by $x \mapsto \lambda'(x)$, for all $x \in X$.

For every nonnegative integer p the space $C^p(\mathcal{U},\mathcal{O})$ of p-chains for the covering $\mathcal{U}$ with values in the sheaf $\mathcal{O}$ is the direct sum $\oplus_{\beta_o,\dots,\beta_p\in\mathbb{N}}\mathcal{O}(U_{\beta_o\dots\beta_p})$ of the Fréchet spaces $\mathcal{O}(U_{\beta_o\dots\beta_p})$. As this is a *countable* direct sum, $C^p(\mathcal{U},\mathcal{O})$ is itself a Fréchet space.

The coboundary map $\delta^p : C^p(\mathcal{U},\mathcal{O}) \to C^{p+1}(\mathcal{U},\mathcal{O})$ is a continuous linear map, so its kernel $Z^p(\mathcal{U},\mathcal{O})$ is a closed subspace of $C^p(\mathcal{U},\mathcal{O})$, hence a Fréchet space.

By hypothesis, $H^1(\mathcal{M},\mathcal{O})$ vanishes, so $H^1(\mathcal{U},\mathcal{O}) = 0$ as well, whence the range of the coboundary map $\delta^0 : C^0(\mathcal{U},\mathcal{O}) \to C^1(\mathcal{U},\mathcal{O})$ is closed and so is a Fréchet space.

Accordingly, by the selection theorem of Michael quoted in the Introduction, there is a continuous map

$$\varsigma : Z^1(\mathcal{U},\mathcal{O}) \to C^0(\mathcal{U},\mathcal{O})$$

such that $\delta^0\circ\varsigma$ is the identity. Consider then the continuous map $f = \varsigma\circ\lambda' : X \to C^0(\mathcal{U},\mathcal{O})$. We have $f = \{f_j\}_{j\in\mathbb{N}}$, where each f_j is a continuous map from X into $\mathcal{O}(U_j)$, and the f_j's satisfy $f_j - f_k = \lambda'_{jk}$ on U_{jk}, for all $j,k\in\mathbb{N}$. It follows that $F_{\tau(j)}e^{-2\pi i f_j} = F_{\tau(k)}e^{-2\pi i f_k}$ on U_{jk}, for all $j,k\in\mathbb{N}$ and all $x\in X$.

Hence the map we seek is the map $\tilde{\psi} : X \to \mathcal{O}(\mathcal{M})$ given by

$$\tilde{\psi}(x)(z) = F_{\tau(k)}(x,z)e^{-2\pi i f_k(x)(z)}$$

for all $k\in\mathbb{N}$, $x\in X$ and $z\in U_k$.

We point out that the proof just given contains implicitly also the proof of the following fact:

Proposition 4.2. *Let $\mathcal{M}$ be a complex-analytic manifold of dimension $N \geq 1$. Let X be a topological space that is paracompact and connected and has the further property that the Čech cohomology group $H^1(X,\mathbb{Z}) = 0$. Let $\mu : X \to \mathfrak{D}^+(\mathcal{M})$ be a continuous map. Then all the divisors $\mu(x)$, $x \in X$ have the same Chern class.*

5. The Topology of $\mathfrak{D}^+_P$, Continued.

The following lemma will be used at several points below.

Lemma 5.1. *If E is an infinite dimensional locally convex Fréchet space, then for all $n = 0,1,\dots$, the homotopy group $\pi_n(E\setminus\{0\})$ vanishes.*

Remark. This is known: $E\setminus\{0\}$ is known to be contractible. The reason we include Lemma 5.1, with a proof, is that it can be proved in a very elementary way by an argument for which we have no reference. It seems worth our while to include the argument for the convenience of the reader.

That $E\setminus\{0\}$ is contractible appears in the literature as follows. (See [**BP**].) By a theorem of Anderson and Kadec [**BP**, Theorem 5.2, p. 189], every infinite-dimensional separable Fréchet space is homeomorphic to $\mathbb{R}^{\mathbb{N}}$. On the other hand, by a theorem of Anderson, if A is a compact subset of $\mathbb{R}^{\mathbb{N}}$ (or more generally a countable union of compacta), then $\mathbb{R}^{\mathbb{N}} \setminus A$ is homeomorphic to the whole space $\mathbb{R}^{\mathbb{N}}$. This follows from [**BP**, Theorem 6.3, p. 166 and Corollary 6.2, p. 165]. It follows in particular that $E \setminus \{0\}$ is homeomorphic to E and so is contractible. A simpler approach than this, still granted the theorem of Anderson and Kadec, is to invoke a theorem of Klee [**Kl**], which is somewhat simpler than the more general result of Anderson.

Proof of the Lemma. The space $E \setminus \{0\}$ is connected, so let $n > 0$ from here on.

Denote by ρ a metric on E that has the property that the ball $\{y \in E : \rho(y, x) < c\}$ is convex for every choice of $x \in E$ and all $r > 0$.

Denote by Σ the simplex in R^{n+1} determined by the origin and the unit vectors $e_j = (0, \dots, 1, \dots, 0)$, 1 in the j^{th} place. The boundary, $b\Sigma$, of Σ is topologically the $n-$sphere.

Consider a continuous map $\varphi : b\Sigma \to E \setminus \{0\}$. As $\varphi(b\Sigma)$ is compact, there is a $\delta > 0$ small enough that $\rho(\varphi(x), 0) > \delta$ for all $x \in b\Sigma$.

For a sufficiently fine simplicial refinement of the given triangulation of $b\Sigma$, say with vertices $\{x_1, \dots, x_r\}$ and with simplexes $T_1, \dots, T_s$, the unique continuous map $\psi : \Sigma \to E \setminus \{0\}$ that agrees with φ at each x_j and that is real affine on each T has the property that both of the sets $\varphi(T_j)$ and $\psi(T_j)$ are contained in some ball with respect to the metric ρ that does not contain 0. Consequently, φ and ψ are homotopic as maps into $E \setminus \{0\}$.

The range of ψ is contained in a finite dimensional real linear subspace of E, which we can take to have real dimension $d > n + 1$. As every map of the $n-$sphere into $\mathbb{R}^d \setminus \{0\}$ is homotopic to a constant, it follows that the original map φ is homotopic in $E \setminus \{0\}$ to a constant.

Remark. Granted that all of the homotopy groups $\pi_n(E \setminus \{0\})$ vanish, it follows that the space $E \setminus \{0\}$ is contractible. See [**BP**, Theorem 6.3, p. 79 and Corollary 6.5, p. 76].

Lemma 5.2. *If $H^1(\mathcal{M}, \mathcal{O}) = 0$ and $H^1(\mathcal{M}, \mathbb{Z}) = 0$, then the cohomology groups $H^1(\mathfrak{D}_P^+(\mathcal{M}), \mathbb{Z})$ and $H^1(\mathfrak{D}_P^+(\mathcal{M}, p), \mathbb{Z})$ vanish.*

The cohomology in question is taken in the sense of Čech theory.

Proof. This lemma depends on the following simple fact: Because the group $H^1(\mathcal{M}, \mathbb{Z})$ vanishes, it follows that every zero-free continuous $\mathbb{C}-$valued function on $\mathcal{M}$ is of the form e^h for some continuous function h. As every continuous logarithm of a holomorphic function is itself holomorphic, it follows that the space $\mathcal{O}^*(\mathcal{M})$ of zero-free holomorphic functions on $\mathcal{M}$ is connected and, indeed, is arcwise connected: If $f = e^h$ with $h \in \mathcal{O}(\mathcal{M})$, then $t \mapsto e^{th}$, $t \in [0,1]$ is a curve in $\mathcal{O}^*(\mathcal{M})$ that connects f to 1. (Note that $\mathcal{O}^*(\mathcal{M})$ is not open in $\mathcal{O}(\mathcal{M})$.)

Consider first the group $H^1(\mathfrak{D}_P^+(\mathcal{M}), \mathbb{Z})$.

If $\mathcal{O}(\mathcal{M}) = \mathbb{C}$, then $\mathfrak{D}_P^+(\mathcal{M}) = \{0\}$, and the assertion of the lemma is true. Thus, assume $\mathcal{O}(\mathcal{M})$ to be infinite dimensional. Let $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ be an open cover for $\mathfrak{D}_P^+(\mathcal{M})$, and let $\{m_{\alpha\beta}\}_{\alpha\beta \in A}$ be a 1-cocycle for this cover with values in $\mathbb{Z}$ so that for each choice of $\alpha, \beta \in A$, $m_{\alpha\beta}$ is a continuous $\mathbb{Z}$-valued function on $U_{\alpha\beta}$.

Because the map Div is continuous, if $\tilde{U}_\alpha = \mathrm{Div}^{-1}(U_\alpha)$, then $\tilde{\mathcal{U}} = \{\tilde{U}_\alpha\}_{\alpha \in A}$ is an open cover for the space $\mathcal{O}(\mathcal{M}) \setminus \{0\}$. With $\tilde{m}_{\alpha\beta} = m_{\alpha\beta} \circ \mathrm{Div}$, the family $\{\tilde{m}_{\alpha\beta}\}_{\alpha,\beta \in A}$ is a 1-cocycle for the covering $\tilde{\mathcal{U}}$ with values in $\mathbb{Z}$.

As the space $\mathcal{O}(\mathcal{M}) \setminus \{0\}$ has vanishing first integral cohomology, it follows that the cohomology group $H^1(\tilde{\mathcal{U}}, \mathbb{Z})$ vanishes, for the canonical map from this group into $H^1(\mathcal{O}(\mathcal{M}) \setminus \{0\}, \mathbb{Z})$ is injective. (See the discussion of Leray's theorem given in [**Gu**].) Consequently, the 1-cocycle $\{\tilde{m}_{\alpha\beta}\}_{\alpha,\beta \in A}$ is a 1-coboundary: There are continuous, $\mathbb{Z}-$valued functions $\tilde{n}_\alpha$ on $\tilde{U}_\alpha$ with $\tilde{m}_{\alpha\beta} = \tilde{n}_\alpha | U_{\alpha\beta} - \tilde{n}_\beta | U_{\alpha\beta}$.

For any divisor $D \in \mathfrak{D}^+(\mathcal{M})$, the fiber $\mathrm{Div}^{-1}(D)$ is of the form $\{gh : g \in \mathcal{O}(\mathcal{M}), g \text{ zero} - \text{free}\}$ for any fixed element h of the fiber. That the set of zero-free functions holomorphic on $\mathcal{M}$ is connected implies that the fiber $\mathrm{Div}^{-1}(D)$ is connected.

It follows that the continuous, $\mathbb{Z}-$valued functions $\tilde{n}_\alpha$ are constant on the fibers of Div and so are of the form $\tilde{n}_\alpha = n_\alpha \circ \mathrm{Div}$ for suitable continuous $\mathbb{Z}-$valued functions n_α on U_α. Plainly $n_\alpha | U_{\alpha\beta} - n_\beta | U_{\alpha\beta} = m_{\alpha\beta}$. That is, the cocycle $\{m_{\alpha\beta}\}_{\alpha\beta} \in A$ is a coboundary. It follows that the group $H^1(\mathfrak{D}_P^+(\mathcal{M}), \mathbb{Z})$ is trivial as claimed.

In its essence, the argument to prove that $H^1(\mathfrak{D}_P^+(\mathcal{M}, p), \mathbb{Z}) = 0$ is parallel to the preceding argument, but some preliminaries are necessary.

If the set $\mathcal{F}_p$ is defined by $\mathcal{F}_p = \{f \in \mathcal{O}(\mathcal{M}) : f(p) \notin \mathbb{C} \setminus (-\infty, 0]\}$, then $\mathrm{Div}\mathcal{F}_p = \mathfrak{D}_P^+(\mathcal{M}; p)$. This is so, for if $D \in \mathfrak{D}_P^+(\mathcal{M}; p)$, then $D = \mathrm{Div} f$ for an $f \in \mathcal{O}(\mathcal{M})$ with $f(p) \neq 0$, which yields that $D = \mathrm{Div}(f/f(p))$. The function $f/f(p)$ lies in $\mathcal{F}_p$. Conversely, it is evident that $\mathrm{Div}\mathcal{F}_p \subset \mathfrak{D}_P^+(\mathcal{M}; p)$.

The set $\mathcal{F}_p$ is open in $\mathcal{O}(\mathcal{M})$. It is also contractible. To see the latter point, define $H : [0,2] \times \mathcal{F}_p \to \mathcal{F}_p$ by

$$H(t,f)(z) = \begin{cases} f(p) e^{(1-t) \int_p^z \frac{df}{f}} & \text{when } t \in [0,1],\ f \in \mathcal{F}_p,\ z \in \mathcal{M} \\ \\ (2-t) f(p) + t - 1 & \text{when } t \in [1,2],\ f \in \mathcal{F}_p,\ z \in \mathcal{M}. \end{cases}$$

That the integral is well defined depends on the hypothesis that $H^1(\mathcal{M}, \mathbb{Z}) = 0$. The map H is continuous, and its range is contained in $\mathcal{F}_p$, $H(0, \cdot)$ is the identity on $\mathcal{F}_p$, and, finally, $H(2, \cdot)$ is the function identically 1.

As $\mathcal{F}_p$ is contractible, it follows that $H^1(\mathcal{F}_p, \mathbb{Z})$ vanishes.

If $D \in \mathfrak{D}_P^+(\mathcal{M}; p)$, then the fiber $(\mathrm{Div}^{-1} D) \cap \mathcal{F}_p$ be described as follows. Fix an $f_o \in \mathcal{O}(\mathcal{M})$ with $\mathrm{Div} f_o = D$. Without loss of generality, f_o can be chosen to take the value 1 at p. The condition that $\mathrm{Div} g = D$ is then the condition that for some $h \in \mathcal{O}^*(\mathcal{M})$, $g = hf_o$ The function g lies in $\mathcal{F}$ if and only if $h(p) \notin (-\infty, 0]$. It follows that the fiber $\mathrm{Div}^{-1}(D) \cap \mathcal{F}_p$ is (arcwise) connected: Fix $h \in \mathcal{O}^*(\mathcal{M})$ with $h(p) \notin (-\infty, 0]$. If $\lambda : [0,1] \to \mathcal{O}^*(\mathcal{M})$ is defined by

$$\lambda(t)(z) = f_o(z) e^{(1-t)\int_p^z \frac{dh}{h} + (1-t)L(h(p))}$$

where L denotes the branch of the logarithm defined on $\mathbb{C} \setminus (-\infty, 0]$ that vanishes at the point 1, then λ is a curve in $\mathcal{O}^*(\mathcal{M})$ that connects g to the function f_o. Note that as $\lambda(t)(p) = f_o(p) e^{(1-t)L(h(p))}$ for all t, the range of λ is contained in the set $\mathcal{F}_p$. Thus, as claimed, the fiber is connected.

The rest of the argument can proceed as in the case of $\mathfrak{D}_P^+(\mathcal{M})$: If $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ is an open cover of $\mathfrak{D}_P^+(\mathcal{M}; p)$, then the sets $\tilde{U}_\alpha = \mathcal{F}_p \cap \mathrm{Div}^{-1}(U_\alpha)$ constitute an open cover for $\mathcal{F}_p$. An integral 1-cocycle on $\mathfrak{D}_P^+(\mathcal{M}; p)$ induces a corresponding integral 1-cocycle on $\mathcal{F}_p$, which is a coboundary, for $H^1(\mathcal{F}_p, \mathbb{Z}) = 0$. The result follows as before.

Corollary 5.3. *If $H^1(\mathcal{M}, \mathcal{O})$ and $H^1(\mathcal{M}, \mathbb{Z})$ vanish, and if $p \in \mathcal{M}$, then there is a continuous map $\varsigma_p : \mathfrak{D}_P^+(\mathcal{M}; p) \to \mathcal{O}(\mathcal{M})$ with $\mathrm{Div} \circ \varsigma_p(D) = D$ for all $D \in \mathfrak{D}_P^+(\mathcal{M}; p)$.*

Proof. This is a consequence of the preceding lemma and Proposition 4.1.

Corollary 5.4. *If $H^1(\mathcal{M}, \mathcal{O})$ and $H^1(\mathcal{M}, \mathbb{Z})$ vanish, then the map $\mathrm{Div} : \mathcal{O}(\mathcal{M}) \setminus \{0\} \to \mathfrak{D}_P^+(\mathcal{M})$ is open.*

Proof. As usual, the case that $\mathcal{O}(\mathcal{M})$ consists only of constants requires separate comment, and is trivial. Therefore suppose that there is a nonconstant holomorphic function f on $\mathcal{M}$. Let $D_o \in \mathfrak{D}_P^+(\mathcal{M})$ be given. Fix $p \in \mathcal{M} \setminus \mathrm{supp} D_o$. There is an associated section $\varsigma_p : \mathfrak{D}_P^+(\mathcal{M}; p) \to \mathcal{O}(\mathcal{M})$.

Let $f_o \in \mathcal{O}(\mathcal{M})$ satisfy $\mathrm{Div}\, f_o = D_o$. We show that if W_o is a neighborhood of f_o in $\mathcal{O}(\mathcal{M}) \setminus \{0\}$, then $\mathrm{Div}\, W_o$ contains a neighborhood of D_o in $\mathfrak{D}_P^+(\mathcal{M})$. To do this, define $\sigma : \mathfrak{D}_P^+(\mathcal{M}; p) \to \mathcal{O}(\mathcal{M})$ by $\sigma(D) = \frac{f_o}{\varsigma_p(D_o)} \varsigma_p(D)$. This map is continuous, and it satisfies $\sigma(D_o) = f_o$. By continuity there is a neighborhood $\tilde{W}_o$ of D_o such that $\sigma(D) \in W_o$ if $D \in \tilde{W}_o$. This implies that the map Div carries W_o onto a set that contains the neighborhood $\tilde{W}_o$ of D_o.

Thus, as claimed, Div is an open map.

Corollary 5.5. *If $H^1(\mathcal{M}, \mathcal{O})$ and $H^1(\mathcal{M}, \mathbb{Z})$ vanish, then the topology on the space $\mathfrak{D}_P^+(\mathcal{M})$ is the quotient topology induced by the map $\mathrm{Div} : \mathcal{O}(\mathcal{M}) \setminus \{0\} \to \mathfrak{D}_P^+(\mathcal{M})$.*

This follows from the continuity and openness of the map Div. (See [**Ke**, p. 95].)

Corollary 5.6. *If $H^1(\mathcal{M}, \mathcal{O})$ and $H^1(\mathcal{M}, \mathbb{Z})$ vanish then the space $\mathfrak{D}_P^+(\mathcal{M})$ is metrizable.*

Remark. Granted the corollary, it is easy to see that *for every complex manifold $\mathcal{M}$, the space $\mathfrak{D}^+(\mathcal{M})$ of nonnegative divisors is metrizable.* Fix $\mathcal{M}$ and let $\{B_j\}_{j\in\mathbb{N}}$ be a countable collection of domains in $\mathcal{M}$ each of which is biholomorphic to a ball and the union of which covers $\mathcal{M}$. Define a map $\rho : \mathfrak{D}^+(\mathcal{M}) \to \Pi_{j\in\mathbb{N}}\mathfrak{D}^+(B_j)$ by $\rho D = \{D|B_j : j \in \mathbb{N}\}$. The map ρ is a homeomorphism from $\mathfrak{D}^+(\mathcal{M})$ onto the closed subset $\{\{D_j : j \in \mathbb{N}\} \in \Pi_{j\in\mathbb{N}}\mathfrak{D}^+(B_j) : D_j|B_{jk} = D_k|B_{jk} \text{ forall j, k} \in \mathbb{N}\}$ of $\Pi_{j\in\mathbb{N}}\mathfrak{D}^+(B_j)$. As the product is countable and as each $\mathfrak{D}^+(B_j) = \mathfrak{D}_P^+(B_j)$ is metrizable, the metrizability of $\mathfrak{D}^+(\mathcal{M})$ follows.

Proof of the corollary. Begin by recalling that the topology of a topological space Y is *regular* if given a point $y \in Y$ and a neighborhood U of y, there is a neighborhood V of y with $\bar{V} \subset U$.

The space $\mathcal{O}(\mathcal{M}) \setminus \{0\}$ is metrizable and separable, so its topology has a countable basis. As the map Div is open, it follows that the topology on $\mathfrak{D}_P^+(\mathcal{O})$ has a countable base.

Moreover, the topology of $\mathfrak{D}_P^+(\mathcal{M})$ is regular. For this, work with the relative weak* topology. If $D_o \in \mathfrak{D}_P^+(\mathcal{M})$, and if W is an open set containing D_o, then there is a collection $\{\alpha_1, \dots, \alpha_r\}$ of compactly supported smooth $(N-1, N-1)$-forms on $\mathcal{M}$ such that W contains the weak* neighborhood

$$W_1 = \left\{ D \in \mathfrak{D}_P^+(\mathcal{M}) : \left| \int_D \alpha_j - \int_{D_o} \alpha_j \right| < 1 \ \text{ for all } \ j = 1, ..., r \right\}$$

of D_o But then the closed neighborhood

$$\bar{W}_{\frac{1}{2}} = \left\{ D \in \mathfrak{D}_P^+(\mathcal{M}) : \left| \int_D \alpha_j - \int_{D_o} \alpha_j \right| \le \frac{1}{2} \ \text{ for all } \ j = 1, ..., r \right\}$$

is contained in W_1. Thus, the topology of $\mathfrak{D}_P^+(\mathcal{M})$ is regular as claimed.

The classical metrization theorem of Uryson and of Tikhonov [**Ke**, p. 125] implies that $\mathfrak{D}_P^+(\mathcal{M})$ is metrizable.

Recall now the notion of an ANR [**BP**]. A topological space Y is an ANR (*absolute neighborhood retract*) if it is metrizable and if whenever it is embedded topologically as a closed subset of a metric space, X, there is a neighborhood of X that retracts onto Y. We shall need three general facts about ANR's. One is a theorem of Hanner [**BP**, p. 69]: *A paracompact space that is locally an ANR is an ANR.* There is also the fact that *a retract of an ANR is an ANR* [**BP**, p. 68]. Finally, *an open set in a locally convex metrizable topological vector space is an ANR* [**BP**, p. 69].

The following corollary is now evident:

Corollary 5.7. *If $H^1(\mathcal{M},\mathcal{O})$ and $H^1(\mathcal{M},\mathbb{Z})$ vanish, then the space $\mathfrak{D}_P^+(\mathcal{M})$ is an ANR as is the space $\mathfrak{D}^+(\mathcal{M};p)$.*

Proof. If $\mathcal{O}(\mathcal{M})$ consists of the constants the result is clear.

If $\mathcal{O}(\mathcal{M})$ contains a nonconstant function, the case of $\mathfrak{D}_P^+(\mathcal{M})$ is a consequence of case of $\mathfrak{D}_P^+(\mathcal{M};p)$, for as noted above, a space that is locally an ANR is itself an ANR.

To treat the case of $\mathfrak{D}_P^+(\mathcal{M};p)$, notice that as the group $H^1(\mathfrak{D}_P^+(\mathcal{M};p),\mathbb{Z})$ vanishes, Proposition 4.1 applies to yield a map $\sigma : \mathfrak{D}_P^+(\mathcal{M};p) \to \mathcal{O}(\mathcal{M})$ with $\mathrm{Div}\circ\sigma(D) = D$ for every $D \in \mathfrak{D}_P^+(\mathcal{M};p)$. Consequently, the map $\sigma\circ\mathrm{Div} : \mathrm{Div}^{-1}(\mathfrak{D}^+(\mathcal{M};p)) \to \sigma(\mathfrak{D}^+(\mathcal{M};p))$ is a retraction of the open subset $\mathrm{Div}^{-1}(\mathfrak{D}^+(\mathcal{M};p))$ of $\mathcal{O}(\mathcal{M})$ onto the range of σ. It follows that $\mathfrak{D}^+(\mathcal{M};p)$ is an ANR as desired.

Corollary 5.8. *Let $\mathcal{M}$ be a complex manifold with $H^1(\mathcal{M},\mathcal{O}) = 0$ and $H^1(\mathcal{M},\mathbb{Z}) = 0$, and let $p \in \mathcal{M}$. For each $n = 0, 1, 2, \ldots,$ the Čech and the singular cohomology groups of $\mathfrak{D}_P^+(\mathcal{M})$ of dimension n with integral coefficients are isomorphic as are the corresponding groups of $\mathfrak{D}_P^+(\mathcal{M};p)$.*

Proof. It is known [**Bo**, p. 107], [**Ma**] that for topological spaces of type ANR the Čech and singular cohomology groups agree.

6. Liftings Over $\mathfrak{D}_P^+$.

If X is a topological space and $\mathcal{M}$ is a complex manifold, we denote by $\mathcal{C}_{X;\mathcal{O}(\mathcal{M})}$ and by $\mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})}$ the sheaves of germs of continuous $\mathcal{O}(\mathcal{M})$−valued functions and of continuous $\mathcal{O}^*(\mathcal{M})$−valued functions on X, respectively. These are sheaves of abelian groups.

Lemma 6.1. *Let $\mathcal{M}$ be a complex-analytic manifold of dimension $N \geq 1$ with $H^1(\mathcal{M},\mathbb{Z}) = 0$, and let X be a paracompact topological space.*

1) *Every continuous map $f : X \to \mathcal{O}^*(\mathcal{M})$ has a continuous logarithm if and only if the Čech cohomology group $H^1(X,\mathbb{Z}) = 0$ vanishes.*
2) *For $q \geq 1$ the Čech cohomology groups $H^q(X,\mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})})$ and $H^{q+1}(X, \mathbb{Z})$ are isomorphic.*

Proof. Define $E : \mathcal{C}_{X;\mathcal{O}(\mathcal{M})} \to \mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})}$ by $Eg = e^{2\pi i g}$. This is a homomorphism of sheaves of abelian groups.

The map E is surjective, that is, if $x_o \in X$ and $f_o : V \to \mathcal{O}^*(\mathcal{M})$ is a continuous map, V some neighborhood of x_o in X, then there is a neighborhood $W \subset V$ of x_o on which there is defined a continuous map $g : W \to \mathcal{O}(\mathcal{M})$ that satisfies $Eg(x) = f_o(x)$ for all $x \in W$. This is seen as follows.

Fix $\zeta \in \mathcal{M}$ and define a function $\epsilon : V \to \mathbb{C}^*$ by $\epsilon(x) = f_o(x)(\zeta)$, for all $x \in V$. (A word on notation may be in order here: For each $x \in V$, $f_o(x)$ is an element of $\mathcal{O}^*(\mathcal{M})$ and so has a value at the point $\zeta \in \mathcal{M}$, viz., $f_o(x)(\zeta)$.)

The evaluation function ϵ is continuous, so there is a neighborhood W of x_o, $W \subset V$, on which ϵ has a continuous logarithm: There is a continuous function $\lambda : W \to \mathbb{C}$ with $e^{\lambda(x)} = \epsilon(x)$ for all $x \in W$.

Define $\tilde{g} : W \to \mathcal{O}(\mathcal{M})$ by

$$\tilde{g}(x)(z) = \int_\zeta^z \frac{df_o(x)}{f_o(x)}$$

for all $x \in W$ and for all $z \in \mathcal{M}$. By hypothesis $H^1(\mathcal{M}, \mathbb{Z}) = 0$, so the integral does not depend on the choice of the path of integration from ζ to z. Hence $\tilde{g}$ is a well-defined, $\mathcal{O}(\mathcal{M})-$valued function on W; it depends continuously on $x \in W$.

The map $g : W \to \mathcal{O}(\mathcal{M})$ given by

$$g(x) = \frac{1}{2\pi i}(\tilde{g}(x) + \lambda(x)), \quad \text{for all } \ x \in W,$$

is continuous and satisfies $Eg = f_o$. Thus, E is surjective.

The kernel of the sheaf homomorphism E is the sheaf $\mathbb{Z}$ of germs of continuous integer-valued functions on X, so we infer that, under the hypothesis $H^1(\mathcal{M}, \mathbb{Z}) = 0$, there is an exact sheaf sequence

$$0 \to \mathbb{Z} \hookrightarrow \mathcal{C}_{X;\mathcal{O}(\mathcal{M})} \xrightarrow{E} \mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})} \to 0.$$

This gives rise to the associated exact cohomology sequence, which contains the segments

$$\Gamma(X, \mathcal{C}_{X;\mathcal{O}(\mathcal{M})}) \to \Gamma(X, \mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})}) \to H^1(X, \mathbb{Z}) \to H^1(X, \mathcal{C}_{X;\mathcal{O}(\mathcal{M})})$$

and

$$H^q(X, \mathcal{C}_{X;\mathcal{O}(\mathcal{M})}) \to H^q(X, \mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})}) \to H^{q+1}(X, \mathbb{Z}) \to H^{q+1}(X, \mathcal{C}_{X;\mathcal{O}(\mathcal{M})}).$$

Moreover the cohomology group $H^q(X, \mathcal{C}_{X;\mathcal{O}(\mathcal{M})})$ is zero for $q \geq 1$, since the sheaf $\mathcal{C}_{X;\mathcal{O}(\mathcal{M})}$ of germs of continuous $\mathcal{O}(\mathcal{M})-$valued functions on the paracompact space X is a *fine* sheaf, as it is because it is a sheaf of modules over the fine sheaf $\mathcal{C}_X$ of complex-valued continuous functions on X. Then the two statements of the lemma follow at once.

Lemma 6.2. *Let $\mathcal{M}$ be a complex-analytic manifold of dimension $N \geq 1$, such that $H^1(\mathcal{M}, \mathcal{O})$ and $H^1(\mathcal{M}, \mathbb{Z})$ vanish. Let X be a topological space that is paracompact and connected and has the further property that the Čech cohomology groups $H^1(X, \mathbb{Z})$ and $H^2(X, \mathbb{Z})$ vanish. Let $\mu : X \to \mathfrak{D}_P^+(\mathcal{M})$ be a continuous map. Then there exists a continuous map $\tilde{\mu} : X \to \mathcal{O}(\mathcal{M})$ such that* $\operatorname{Div} \tilde{\mu}(x) = \mu(x)$ *for all $x \in X$.*

Proof. The case that $\mathcal{O}(\mathcal{M}) = \mathbb{C}$ is clear, so we assume from here on that there is a nonconstant holomorphic function on $\mathcal{M}$.

To begin with, the Lemma is correct locally in X. Let $x \in X$. The point $\mu(x)$ lies in an open set $\mathfrak{D}_P^+(\mathcal{M}; p)$ for some choice of $p \in \mathcal{M}$. By

Corollary 5.3, there is a section $\varsigma_p : \mathfrak{D}^+_P(\mathcal{M};p) \to \mathcal{O}(\mathcal{M})$ of the map Div. By continuity, there is an open set V_x in X that contains x and such that $\mu(V_x) \subset \mathfrak{D}^+_P(\mathcal{M};p)$. As $\mathrm{Div} \circ \varsigma_p \circ \mu(x) = \mu(x)$, the Lemma is seen to be correct locally on X.

Let $\mathcal{V} = \{V_j\}_{j\in\mathbb{N}}$ be a locally finite open covering of X with the property that for each j there is a map $\mu_j : V_j \to \mathcal{O}(\mathcal{M})$ that satisfies $\mathrm{Div} \circ \mu_j = \mu$ in V_j. If $\mu_{jk} : V_{jk} \to \mathcal{O}(\mathcal{M})$ is defined by $\mu_{jk}(x)(z) = \frac{\mu_j(x)(z)}{\mu_k(x)(z)}$, then μ_{jk} is continuous, and $\{\mu_{jk}\}_{j,k\in\mathbb{N}}$ defines an element of the group of cocycles $\mathbb{Z}^1(\mathcal{V}, \mathcal{C}_{X;\mathcal{O}^*(\mathcal{M})})$. By hypothesis, $H^2(X,\mathbb{Z}) = 0$, so by Lemma 6.1, this cocycle is a coboundary: There are maps $\nu_j : V_j \to \mathcal{O}^*(\mathcal{M})$ with $\frac{\mu_j}{\mu_l} = \frac{\nu_j}{\nu_k}$ on V_{jk}. This implies that on V_{jk} the functions $\mu_j\nu_j^{-1}$ and $\mu_k\nu_k^{-1}$ agree, whence we can obtain a well defined map $\tilde{\mu} : X \to \mathcal{O}(\mathcal{M})$ by requiring that on V_j, $\tilde{\mu} = \mu_j\nu_j^{-1}$. This map $\tilde{\mu}$ satisfies $\mathrm{Div} \circ \tilde{\mu} = \mu$, so the Lemma is proved.

Lemma 6.3. *Let $\mathcal{M}$ be a complex manifold with $H^1(\mathcal{M},\mathcal{O}) = 0$ and $H^1(\mathcal{M}, Z) = 0$. For every nonnegative integer n the homotopy group $\pi_n(\mathfrak{D}^+_P(\mathcal{M}))$ is zero.*

Proof. The lemma is trivial if there are no nonconstant holomorphic functions on $\mathcal{M}$, in which case $\mathfrak{D}^+_P(\mathcal{M})$ contains only 0−the zero divisor.

Thus, assume that $\mathcal{O}(\mathcal{M})$ contains a nonconstant function.

That $\pi_0(\mathfrak{D}^+_P(\mathcal{M})) = 0$ means simply that the space $\mathfrak{D}^+_P(\mathcal{M})$ is arcwise connected, which it is.

In order to prove that $\pi_1(\mathfrak{D}^+_P(\mathcal{M})) = 0$, we begin by recalling that, as $H^1(\mathcal{M},\mathbb{Z}) = 0$, the space $\mathcal{O}^*(\mathcal{M})$ is arcwise connected.

We have to prove that each continuous map $\gamma : [0,1] \to \mathfrak{D}^+_P(\mathcal{M})$ with $\gamma(0) = \gamma(1)$ is homotopic to the constant map $t \mapsto \gamma(0)$. To do this, first lift γ to a closed curve μ in $\mathcal{O}(\mathcal{M})$. That is, μ is to be a closed curve that satisfies $\mathrm{Div} \circ \mu = \gamma$. Lemma 6.2 or, alternatively, Proposition 4.1, implies that there is a continuous map $\gamma^* : [0,1] \to \mathcal{O}(\mathcal{M})$ that satisfies $\mathrm{Div} \circ \gamma^* = \gamma$. In general, γ^* will not be a closed curve. However, the hypothesis that $H^1(\mathcal{M},\mathbb{Z}) = 0$ lets us modify γ^* to obtain a closed lifting, μ, of γ: The function $\frac{\gamma^*(0)}{\gamma^*(1)}$ is holomorphic and zero-free on $\mathcal{M}$ and so has a holomorphic logarithm, say λ. For μ take the function given by $\mu(t) = e^{t\lambda}\gamma^*(t)$. As $\mathrm{Div} \circ \mu = \gamma$, the range of μ is contained in the space $\mathcal{O}(\mathcal{M}) \setminus \{0\}$, which, by Lemma 5.1, has vanishing fundamental group, so the map μ is homotopic to the constant map $t \mapsto \mu(0)$: There is a map $\tilde{H} : [0,1] \times [0,1] \to \mathcal{O}(\mathcal{M}) \setminus \{0\}$ with $\tilde{H}(0,t) = \mu(t)$, $\tilde{H}(1,t) = \mu(0)$ for all $t \in [0,1]$ and $\tilde{H}(0,s) = H(1,s)$ for all s. Then the map $H = \mathrm{Div} \circ \tilde{H} : [0,1] \times [0,1] \to \mathfrak{D}^+_P(\mathcal{M})$ is a homotopy in $\mathfrak{D}^+_P(\mathcal{M})$ connecting γ with the constant map $t \mapsto \gamma(0)$.

Hence $\mathfrak{D}^+_P(\mathcal{M})$ is simply connected.

Next, assume that $n \geq 2$. There is a distinction between the case $n = 2$ and the case $n \geq 3$.

Consider first the case $n \geq 3$. We are to show that each continuous map $\phi : S^n \to \mathfrak{D}_P^+(\mathcal{M})$ is homotopic to a constant map. As the cohomology groups $H^1(S^n, \mathbb{Z})$ and $H^2(S^n, \mathbb{Z})$ vanish, Lemma 6.2 yields a continuous lift $\tilde{\varphi} : S^n \to \mathcal{O}(\mathcal{M})$ with $\operatorname{Div} \tilde{\varphi}(x) = \varphi(x)$ for all $x \in S^n$. The range of $\tilde{\varphi}$ is contained in the space $\mathcal{O}(\mathcal{M}) \setminus \{0\}$, which, by Lemma 5.1, has vanishing n^{th} homotopy group, so the map $\tilde{\varphi}$ is homotopic to a constant map: There is a map $\tilde{H} : [0,1] \times S^n \to \mathcal{O}(\mathcal{M}) \setminus \{0\}$ with $\tilde{H}(0,x) = \tilde{\varphi}(x)$ and $\tilde{H}(1,x) = f_o$ for all $x \in S^n$, for some fixed $f_o \in \mathcal{O}(\mathcal{M}) \setminus \{0\}$ that is independent of x. Then the map $H = \operatorname{Div} \circ \tilde{H} : [0,1] \times S^n \to \mathfrak{D}_P^+(\mathcal{M})$ is a homotopy in $\mathfrak{D}_P^+(\mathcal{M})$ connecting φ with a constant map.

The case $n = 2$ requires something more.

Let $\varphi : S^2 \to \mathfrak{D}_P^+(\mathcal{M})$ be a continuous map with S^2 realized in the usual way as

$$S^2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\right\}.$$

Denote by Σ^+ the subset $S^2 \setminus \{(-1,0,0)\}$ of S^2 and by Σ^- the subset $S^2 \setminus \{(1,0,0)\}$. Lemma 6.2 (considered for $X = \Sigma^+, \Sigma^-$ and $\mu = \varphi|\Sigma^+, \varphi|\Sigma^-$, respectively) yields continuous maps $\varphi^+ : \Sigma^+ \to \mathcal{O}(\mathcal{M})$ and $\varphi^- : \Sigma^- \to \mathcal{O}(\mathcal{M})$ with $\operatorname{Div} \circ \varphi^+ = \varphi$ on Σ^+ and $\operatorname{Div} \circ \varphi^- = \varphi$ on Σ^-. After multiplying φ^- by a function in $\mathcal{O}^*(\mathcal{M})$, we can suppose that $\varphi^+(0,0,1) = \varphi^-(0,0,1)$.

Set $\Sigma_{+-} = \Sigma^+ \cap \Sigma^-$.

Let z_o be a fixed point of $\mathcal{M}$, and define $h : \Sigma_{+-} \to \mathbb{C}^*$ by

$$h(x) = \frac{\varphi^+(x)(z_o)}{\varphi^-(x)(z_o)}, \quad \text{for all } x \in \Sigma_{+-}.$$

Define $f : \Sigma_{+-} \to \mathcal{O}^*(\mathcal{M})$ by

$$f(x) = \frac{1}{h(x)} \frac{\varphi^+(x)}{\varphi^-(x)} \quad \text{for all } x \in \Sigma_{+-}.$$

This is a continuous $\mathcal{O}^*(\mathcal{M})$-valued function with the properties that $f(x)(z_o) = 1$ for all $x \in \Sigma_{+-}$ and $f(0,0,1) \equiv 1$ on $\mathcal{M}$.

Define then $g : \Sigma_{+-} \to \mathcal{O}(\mathcal{M})$ by

$$g(x)(z) = \int_{z_o}^{z} \frac{df(x)}{f(x)} \quad \text{for all } x \in \Sigma_{+-}.$$

The integral is independent of the choice of the path of integration from z_o to $z \in \mathcal{M}$, for, if γ and γ' are two such integration paths, the map $\Sigma_{+-} \to \mathbb{C}$ given by $x \mapsto \int_{\gamma - \gamma'} \frac{df}{f(x)}$ for all $x \in \Sigma_{+-}$, being a continuous $2\pi i\mathbb{Z}-$valued map which is zero at the point $(0,0,1)$, is identically zero on Σ_{+-}. Consequently g is a well-defined continuous map from Σ_{+-} into $\mathcal{O}(\mathcal{M})$, and it satisfies $e^{g(x)(z)} = f(x)(z)$, for all $(x,z) \in \Sigma_{+-} \times \mathcal{M}$.

Since $H^1(S^2, \mathcal{C}_{S^2;\mathcal{O}(\mathcal{M})}) = 0$, there are continuous functions $g^\pm : \Sigma^\pm \to \mathcal{O}(\mathcal{M})$ with $g^+ - g^- = g$ on Σ_{+-}. Define $f^+ = e^{g^+}$, $f^- = e^{g^-}$ on Σ^+ and Σ^-, respectively. We have $\frac{f^+}{f^-} = f$ on Σ_{+-}.

Define $\tilde{\varphi}^+ = \frac{1}{f^+}\varphi^+$ on Σ^+ and $\tilde{\varphi}^- = \frac{1}{f^-}\varphi^-$ on Σ^-. Then $\tilde{\varphi}^+$ and $\tilde{\varphi}^-$ are continuous $\mathcal{O}(\mathcal{M})-$valued functions on Σ^+ and Σ^-, respectively. Moreover, $\tilde{\varphi}^+ = \tilde{\varphi}^-$ on Σ_{+-}. Accordingly, if we set $\tilde{\varphi} = \tilde{\varphi}^+$ on Σ^+ and $\tilde{\varphi} = \tilde{\varphi}^-$ on Σ^-, then $\tilde{\varphi}$ is a well-defined map from S^2 to $\mathcal{O}(\mathcal{M})$ with $\operatorname{Div} \circ \tilde{\varphi} = \varphi$.

As before, this implies that φ is homotopic to a constant, and the proof for the case $n = 2$ is complete.

The lemma is proved.

Now we are ready to prove our main theorems concerning $\mathfrak{D}_P^+(\mathcal{M})$.

Theorem 6.4. *If $\mathcal{M}$ is a complex-analytic manifold of dimension $N \geq 1$ such that $H^1(\mathcal{M}, \mathbb{Z}) = 0$ and $H^1(\mathcal{M}, \mathcal{O}) = 0$, then there exists a continuous map $\varsigma : \mathfrak{D}_P^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ with the property that $\operatorname{Div} \varsigma(D) = D$ for all $D \in \mathfrak{D}_P^+(\mathcal{M})$.*

Proof. By Corollaries 5.6 and 5.7, $\mathfrak{D}_P^+$ is a metrizable ANR and so is also paracompact. Also, all the homotopy groups of $\mathfrak{D}_P^+(\mathcal{M})$ vanish. By a theorem of Hurewicz [**Hu**, p. 57], all the singular homology groups $H_n^s(\mathfrak{D}_P^+(\mathcal{M}))$ vanish, whence, by the universal coefficients theorem, the singular cohomology groups vanish. As the space is an ANR, this implies that the Čech cohomology groups vanish. The theorem now follows from Lemma 6.2.

Corollary 6.5. *If $\mathcal{M}$ is a complex-analytic manifold of dimension $N \geq 1$ such that $H^1(\mathcal{M}, \mathbb{Z}) = 0$ and $H^1(\mathcal{M}, \mathcal{O}) = 0$, then given an arbitrary topological space X and a continuous map $\psi \to \mathfrak{D}_P^+(\mathcal{M})$, there is a continuous map $\tilde{\psi} : X \to \mathcal{O}(\mathcal{M})$ with $\operatorname{Div} \circ \tilde{\psi} = \psi$.*

There is a result in the direction converse to Theorem 6.4 that requires fewer hypotheses on $\mathcal{M}$.

Theorem 6.6. *Let $\mathcal{M}$ be a complex-analytic manifold of dimension $N \geq 1$. If there exists a continuous map $\varsigma : \mathfrak{D}_P^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ with the property that $\operatorname{Div} h(D) = D$ for all $D \in \mathfrak{D}_P^+(\mathcal{M})$, then the space $\mathfrak{D}_P^+(\mathcal{M})$ is contractible, and every zero-free holomorphic function on $\mathcal{M}$ has a holomorphic logarithm.*

Remark. The condition that every-zero free holomorphic function on $\mathcal{M}$ have a logarithm is not, in general, equivalent to the condition that $H^1(\mathcal{M}, \mathbb{Z}) = 0$, i.e., to the condition that every zero-free *continuous* function have a continuous logarithm, though the equivalence is correct for Stein manifolds and, indeed, on all complex manifolds $\mathcal{M}$ with $H^1(\mathcal{M}, \mathcal{O}) = 0$. This follows by considering the cohomology sequence associated with the exact sheaf sequence $0 \to \mathbb{Z} \to \mathcal{O} \to \mathcal{O}^* \to 0$.

Proof. The existence of a continuous map $\varsigma : \mathfrak{D}_P^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ such that $\operatorname{Div} \varsigma(D) = D$ for all $D \in \mathfrak{D}_P^+(\mathcal{M})$ implies that $\mathfrak{D}_P^+(\mathcal{M})$ is an ANR, following the proof of Corollary 5.7.

Also, the existence of ς implies that every continuous map μ from any topological space X into $\mathfrak{D}_P^+(\mathcal{M})$ lifts to a continuous map $\tilde{\mu} : X \to \mathcal{O}(\mathcal{M})$ such that $\operatorname{Div} \circ \tilde{\mu} = \mu$.

Then all the homotopy groups of $\mathfrak{D}_P^+(\mathcal{M})$ vanish, following the part of the proof of Lemma 6.3 concerning the case $n \geq 3$. It is known [**Hu**, Th. 8.2, p. 218] that an ANR with vanishing homotopy groups is contractible.

It remains to prove that every zero-free holomorphic function on $\mathcal{M}$ has a logarithm.

Each continuous map $\alpha : S^1 \to \mathfrak{D}_P^+(\mathcal{M})$ lifts to a continuous map $\tilde{\alpha} : S^1 \to \mathcal{O}(\mathcal{M}) \setminus \{0\}$ such that $\operatorname{Div} \circ \tilde{\alpha} = \alpha$.

Given a function $g \in \mathcal{O}^*(\mathcal{M})$, choose a continuous map $H : [0,1] \to \mathcal{O}(\mathcal{M}) \setminus \{0\}$ with $H(0) = 1$ and $H(1) = g$. (Note: $H(0)$ is the function identically one.) Define $\varphi : [0,1] \to \mathfrak{D}_P^+(\mathcal{M})$ by $\varphi(t) = \operatorname{Div} H(t)$. This is a continuous path which is *closed*, for $\operatorname{Div} H(0) = \operatorname{Div} H(1) = 0$—the zero divisor.

This closed path φ lifts to a closed path $\Phi : [0,1] \to \mathcal{O}(\mathcal{M}) \setminus \{0\}$ with $\operatorname{Div} \Phi(t) = \varphi(t)$ for all $t \in [0,1]$. (That Φ is a closed path means that $\Phi(0) = \Phi(1)$.) There is a continuous map $\sigma : [0,1] \to \mathcal{O}^*(\mathcal{M})$ such that $H(t) = \sigma(t)\Phi(t)$ for all t. Then

$$g = \frac{H(1)}{H(0)} = \frac{\sigma(1)\Phi(1)}{\sigma(0)\Phi(0)} = \frac{\sigma(1)}{\sigma(0)}.$$

As σ is continuous and always zero-free, if c is a piecewise smooth closed path in $\mathcal{M}$, then $\frac{1}{2\pi i}\int_c \frac{d\sigma(t)}{\sigma(t)}$ is a continuous integer and so is constant. Thus,

$$\frac{1}{2\pi i}\int_c \frac{d\sigma(0)}{\sigma(0)} = \frac{1}{2\pi i}\int_c \frac{d\sigma(1)}{\sigma(1)}.$$

This implies that $\frac{1}{2\pi i}\int_c \frac{dg}{g} = 0$. As this is correct for all c, the function g has a logarithm.

The theorem is proved.

We give an additional result in the direction of Theorem 6.4 in which the geometric condition on $\mathcal{M}$ is replaced by a geometric condition on the space of parameters.

Theorem 6.7. *Let $\mathcal{M}$ be a complex manifold with $H^1(\mathcal{M}, \mathcal{O}) = 0$, and let X be a topological space with $H^1(X, \mathbb{Z}) = 0$. If $\mu : X \to \mathfrak{D}_P^+(\mathcal{M})$ is a continuous map, then there is a continuous map $\tilde{\mu} : X \to \mathcal{O}(\mathcal{M})$ with* $\operatorname{Div} \circ \tilde{\mu} = \mu$.

Proof. If $\Omega \subset \mathcal{M}$ is a domain biholomorphically equivalent to a ball, then as there is a continuous map $\varsigma_\Omega : \mathfrak{D}_P^+(\Omega) \to \mathcal{O}(\Omega)$ with $\operatorname{Div} \circ \varsigma_\Omega(D) = D$ for all $D \in \mathfrak{D}_P^+(\Omega)$, we can define $\tilde{\mu}_\Omega : X \to \mathcal{O}(\Omega)$ by $\tilde{\mu}_\Omega(x) = \varsigma_\Omega(\mu(x)|\Omega)$ to get a map with $\mathrm{Div} \circ \tilde{\mu}_\Omega(x) = \mu(x)|\Omega$ for all $x \in X$.

Consequently, there is a locally finite cover $\{U_j\}_{j\in\mathbb{N}}$ such that the intersections $U_{j_1\cdots j_r}$ are all contractible and such that for all $j \in \mathbb{N}$, there is a continuous function $F_j : X \times U_j \to \mathbb{C}$ with $F_j(x, \cdot) \in \mathcal{O}(U_j)$ for all $x \in X$ and with $\mathrm{Div}\, F_j(x, \cdot) = \mu(x)|U_j$ for all $x \in X$ and all $j \in \mathbb{N}$.

Define $\psi_{jk} : X \times U_{jk} \to \mathbb{C}$ by $\psi_{jk}(x, z) = \frac{F_j(x,z)}{F_k(x,z)}$. From here, the proof simply follows the last nine paragraphs of the proof of Proposition 4.1.

Corollary 6.8. *If $\mathcal{M}$ be a complex-analytic manifold such that $H^1(\mathcal{M}, \mathcal{O}) = 0$, then the following four conditions are equivalent:*

1) *$H^1(\mathcal{M}, \mathbb{Z}) = 0$;*
2) *The space $\mathfrak{D}_P^+(\mathcal{M})$ is simply connected;*
3) *The space $\mathfrak{D}_P^+(\mathcal{M})$ is contractible;*
4) *There exists a continuous map $h : \mathfrak{D}_P^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ with the property that $\mathrm{Div}\, h(D) = D$ for all $D \in \mathfrak{D}_P^+(\mathcal{M})$.*

Proof. That the condition 1) implies 4) is the content of Theorem 6.4. Theorem 6.6 yields that 4) implies 3). Plainly, 3) implies 2). What remains to be seen is that condition 2) implies condition 1).

To this end, we show that if $\mathfrak{D}_P^+(\mathcal{M})$ is simply connected, then each zero-free holomorphic function on $\mathcal{M}$ is an exponential. Accordingly, let $g \in \mathcal{O}(\mathcal{M})$ be zero free. Let $H : [0, 1] \to \mathcal{O}(\mathcal{M})$ be a continuous function with $H(0) = 1$, $H(1) = g$. The map $\varphi = \mathrm{Div} \circ H : [0, 1] \to \mathfrak{D}_P^+$ is a closed curve in $\mathfrak{D}_P^+$. As $\mathfrak{D}_P^+$ is contractible there is a continuous map $m : [0, 1] \times [0, 1] \to \mathfrak{D}_P^+(\mathcal{M})$ with $m(1, s) = D_1$ for all $s \in [0, 1]$, D_1 some fixed divisor, with $m(0, s) = \varphi(s)$ for all $s \in [0, 1]$ and with $m(t, 0) = m(t, 1)$ for all $t \in [0, 1]$. The preceding theorem provides a lift of the map m to a map $\tilde{m} : [0, 1] \times [0, 1] \to \mathcal{O}(\mathcal{M})$ with $\mathrm{Div} \circ \tilde{m} = m$. This implies that the map φ lifts to a *closed* map $\Phi : [0, 1] \to \mathcal{O}(\mathcal{M})$ with $\mathrm{Div} \circ \Phi = \varphi$. We are now in the situation encountered at the end of Theorem 6.6. The argument there shows that g is an exponential. It follows that $H^1(\mathcal{M}, \mathbb{Z}) = 0$.

The Corollary is proved.

Another corollary of Theorem 6.7 is the following:

Corollary 6.9. *If $\mathcal{M}$ is a complex manifold with $H^1(\mathcal{M}, \mathcal{O}) = 0$, then the homotopy groups $\pi_n(\mathfrak{D}_P^+(\mathcal{M}))$ vanish for $n = 0, 2, 3, \ldots$.*

7. Domains in Stein Manifolds.

Theorem 6.4 is established under two hypotheses: the analytic hypothesis that $H^1(\mathcal{M}, \mathcal{O})$ vanish and the geometric hypothesis that $H^1(\mathcal{M}, \mathbb{Z})$ vanish. Theorem 6.6 establishes the necessity of the geometric hypothesis. We shall show that for manifolds that are domains in Stein manifolds, the analytic hypothesis can be abandoned.

Let $\mathcal{M}$ be a domain in a Stein manifold. A theorem of Rossi [**Ro**] implies that $\mathcal{M}$ has an envelope of holomorphy, $\tilde{\mathcal{M}}$, which is itself a Stein manifold.

For each $g \in \mathcal{O}(\mathcal{M})$, let $\tilde{g} \in \mathcal{O}(\tilde{\mathcal{M}})$ denote the extension of g to $\tilde{\mathcal{M}}$. The map $g \mapsto \tilde{g}$ effects a topological isomorphism between the Fréchet spaces $\mathcal{O}(\mathcal{M})$ and $\mathcal{O}(\tilde{\mathcal{M}})$, which carries $\mathcal{O}^*(\mathcal{M})$ onto $\mathcal{O}^*(\tilde{\mathcal{M}})$.

We will need below the remark that if $H^1(\mathcal{M}, \mathbb{Z}) = 0$, then $H^1(\tilde{\mathcal{M}}, \mathbb{Z}) = 0$, too. This is evident: The vanishing of $H^1(\mathcal{M}, \mathbb{Z})$ implies that each zero-free $f \in \mathcal{O}(\mathcal{M})$ is of the form e^g, $g \in \mathcal{O}(\mathcal{M})$ whence each zero-free $\tilde{f} \in \mathcal{O}(\tilde{\mathcal{M}})$ is of form $e^{\tilde{g}}$, $\tilde{g} \in \mathcal{O}(\tilde{\mathcal{M}})$. As $\tilde{\mathcal{M}}$ is a Stein manifold, the latter condition implies that $H^1(\tilde{\mathcal{M}}, \mathbb{Z}) = 0$.

Lemma 7.1. *The map* $\psi : \mathfrak{D}_P^+(\mathcal{M}) \to \mathfrak{D}_P^+(\tilde{\mathcal{M}})$ *defined by by* $\psi(\operatorname{Div} g) = \operatorname{Div} \tilde{g}$*, is a homeomorphism.*

Proof. The map ψ is injective. If $f, g \in \mathcal{O}(\mathcal{M}) \setminus \{0\}$ and if $\operatorname{Div} \tilde{f} = \operatorname{Div} \tilde{g}$, then $\tilde{f}/\tilde{g} \in \mathcal{O}^*(\tilde{\mathcal{M}})$, whence $f/g \in \mathcal{O}^*(\mathcal{M})$.

The map ψ is plainly surjective.

The inverse $\psi^{-1} : \mathfrak{D}_P^+(\tilde{\mathcal{M}}) \to \mathfrak{D}_P^+(\mathcal{M})$ is continuous: If, for $\tilde{g}_o, \tilde{g}_n \in \mathcal{O}(\tilde{\mathcal{M}}) \setminus \{0\}$, $n = 1, 2, \ldots$, $\operatorname{Div} \tilde{g}_n \to \operatorname{Div} \tilde{g}_o$, then, with $g_o = \tilde{g}_o | \mathcal{M}$ and $g_n = \tilde{g}_n | \mathcal{M}$, $\operatorname{Div} g_n \to \operatorname{Div} g_o$.

It is less evident that ψ itself is continuous.

To prove the continuity of ψ, let g_n, $n = 1, 2, \ldots$ and g_o be elements of $\mathcal{O}(\mathcal{M}) \setminus \{0\}$ such that $\operatorname{Div} g_n \to \operatorname{Div} g_o$. Set $D_n = \operatorname{Div} g_n$ and $D_o = \operatorname{Div} g_o$. As $D_n \to D_o$, the set $\mathcal{F} = \{D_1, D_2, \ldots\}$ is a normal family of nonnegative divisors on $\mathcal{M}$.

Put $\tilde{\mathcal{F}} = \{\tilde{D}_1, \tilde{D}_2, \ldots\}$, where $\tilde{D}_n = \operatorname{Div} \tilde{g}_n$. The set $\tilde{\mathcal{F}}$ is a set of nonnegative divisors in the Stein manifold $\tilde{\mathcal{M}}$. According to results of Oka and Fujita - see [**Ba**] and the references it contains - the *domain of normality* of the family $\tilde{\mathcal{F}}$ is a domain of holomorphy in $\tilde{\mathcal{M}}$, say $\mathcal{M}_1$. (The domain of normality of a family of divisors is the largest domain on which the family is a normal family.) As $\mathcal{F}$ is a normal family on $\mathcal{M}$, we have $\mathcal{M} \subset \mathcal{M}_1 \subset \tilde{\mathcal{M}}$. This implies that $\mathcal{M}_1 = \tilde{\mathcal{M}}$, for $\tilde{\mathcal{M}}$ is the envelope of holomorphy of $\mathcal{M}$. That is, the family $\tilde{\mathcal{F}}$ is a normal family in $\tilde{\mathcal{M}}$. Consequently, the sequence $\{\tilde{D}_n\}_{n=1}^\infty$ has a convergent subsequence $\{\tilde{D}_{n_j}\}_{j=1}^\infty$. Let $\tilde{D}$ be the limit of this sequence. The divisor $\tilde{D}$ on $\tilde{\mathcal{M}}$ is a divisor that on $\mathcal{M}$ agrees with D_o. This implies that $\tilde{D} = \psi(D_o)$. We thus have that the sequence $\{\tilde{D}_n\}_{n=1}^\infty$ converges to $\psi(D_o)$. Accordingly, the map ψ is continuous.

The lemma is proved.

Theorem 7.2. *If* $\mathcal{M}$ *is a domain in a Stein manifold* $\mathcal{N}$*, and if* $H^1(\mathcal{M}, \mathbb{Z}) = 0$*, then there is a continuous map* $\varsigma : \mathfrak{D}_P^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ *with* $\operatorname{Div} \circ \varsigma(D) = D$ *for all* $D \in \mathfrak{D}_P^+(\mathcal{M})$.

Proof. Denote by $\tilde{\mathcal{M}}$ the envelope of holomorphy of $\mathcal{M}$. It is a Stein manifold and so satisfies $H^1(\tilde{\mathcal{M}}, \mathcal{O}) = 0$. We shall regard $\mathcal{M}$ as being a subset of $\tilde{\mathcal{M}}$. Let $\psi : \mathfrak{D}_P^+(\mathcal{M}) \to \mathfrak{D}_P^+(\tilde{\mathcal{M}})$ be the homeomorphism of Lemma 7.1. Theorem 6.4 applies to the manifold $\tilde{\mathcal{M}}$, so there is a map $\sigma : \mathfrak{D}_P^+(\tilde{\mathcal{M}})$ that satisfies $\mathrm{Div} \circ \sigma(D) = D$ for all $D \in \mathfrak{D}_P^+(\tilde{\mathcal{M}})$. The desired map ς is given by $\varsigma(D) = \sigma(\psi(D))|\mathcal{M}$.

The theorem is proved.

8. Line Bundles.

Some of the preceding results have analogues for holomorphic line bundles. Given a complex manifold $\mathcal{M}$ and a holomorphic line bundle $\mathcal{L} \xrightarrow{\pi} \mathcal{M}$ over $\mathcal{M}$, denote by $\mathcal{L}(\mathcal{M})$ the space of sections of $\mathcal{L}$ over $\mathcal{M}$. Given a holomorphic section, s, of $\mathcal{L}$, there is an associated divisor $\mathrm{Div}\ s \in \mathfrak{D}^+(\mathcal{M})$. If $\mathfrak{D}_{\mathcal{L}}^+(\mathcal{M})$ denotes the space of all these divisors, the map $\mathrm{Div} : \mathcal{L}(\mathcal{M}) \setminus \{0\} \to \mathfrak{D}^+(\mathcal{M})$ defined in this way is continuous as a consequence of the theorem of Andreotti and Norguet used above since the bundle $\mathcal{L}$ is locally trivial. As $\mathfrak{D}_{\mathcal{L}}^+(\mathcal{M})$ is a subset of the space $\mathfrak{D}^+(\mathcal{M})$, the natural topology on it is metrizable,

In this section we shall be concerned mainly with line bundles $\mathcal{L}$ for which the space $\mathcal{L}(\mathcal{M})$ is infinite dimensional over $\mathbb{C}$.

Lemma 8.1. *If $\mathcal{M}$ is a complex manifold for which $H^1(\mathcal{M}, \mathbb{Z})$ vanishes, then the group $H^1(\mathfrak{D}_{\mathcal{L}}^+(\mathcal{M}), \mathbb{Z})$ vanishes.*

Proof. The point is that for a given divisor $D \in \mathfrak{D}_{\mathcal{L}}^+(\mathcal{M})$, the fiber $\mathrm{Div}^{-1}(D)$ consists of all multiples of some fixed section $s \in \mathcal{L}(\mathcal{M})$ by zero-free holomorphic functions on $\mathcal{M}$. As we have seen earlier, the hypothesis that $H^1(\mathcal{M}, \mathbb{Z})$ vanish implies the connectedness of the space $\mathcal{O}^*(\mathcal{M})$. Granted this, the proof can proceed exactly along the lines of the first part of the proof of Lemma 5.2.

Theorem 8.2. *Let $\mathcal{M}$ be a complex manifold for which the groups $H^1(\mathcal{M}, \mathbb{Z})$ and $H^1(\mathcal{M}, \mathcal{O})$ both vanish. If $\mathcal{L} \xrightarrow{\pi} \mathcal{M}$ is a holomorphic line bundle for which the space $\mathcal{L}(\mathcal{M})$ is infinite dimensional, then there is a continuous map $\varsigma : \mathfrak{D}_{\mathcal{L}}^+(\mathcal{M}) \to \mathcal{L}(\mathcal{M})$ with* $\mathrm{Div} \circ \varsigma(D) = D$ *for all $D \in \mathfrak{D}_{\mathcal{L}}^+(\mathcal{M})$.*

Note. The hypotheses of this theorem involve the cohomology of $\mathcal{M}$ with values in the sheaf $\mathcal{O}_{\mathcal{M}}$, not with values in the sheaf $\mathcal{L}_{\mathcal{M}}$ of germs of sections of the line bundle $\mathcal{L}$ as might be expected.

Proof. Fix an open cover $\mathcal{U} = \{U_j\}_{j \in \mathbb{N}}$ of the manifold $\mathcal{M}$ by contractible Stein open sets U_j for which the intersections U_{jk} are also contractible.

For each j, let $r_j : \mathfrak{D}_{\mathcal{L}}^+(\mathcal{M}) \to \mathfrak{D}^+(U_j)$ be the restriction map given by $r_j(D) = D|U_j$. The hypotheses imply that the the bundle $\mathcal{L}$ is trivial over U_j; let $s_j \in \mathcal{L}(U_j)$ be a zero-free section. By Theorem 6.4, there is a continuous map $\sigma_j : \mathfrak{D}^+(U_j) \to \mathcal{O}(U_j)$ with $\mathrm{Div} \circ \sigma_j(D) = D$ for every $D \in \mathfrak{D}^+(U_j)$. Define then $\varsigma_j : \mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}) \to \mathcal{L}(U_j)$ by

$$\varsigma_j = (\sigma_j \circ r_j)s_j.$$

Thus, $\mathrm{Div} \circ \varsigma_j(D) = D|U_j$ for all $D \in \mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$.

Define $g_{jk}(D,z) = \frac{\varsigma_j(D)}{\varsigma_k(D)}(z)$ so that g_{jk} is a continuous, zero-free function on $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}) \times U_{jk}$ with the property that for each D, $g_{jk}(D,\cdot) \in \mathcal{O}(U_{jk})$.

If there are zero-free continuous functions g_j on $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}) \times U_j$ with $g_j(D,\cdot) \in \mathcal{O}(U_j)$ and $\frac{g_j}{g_k} = g_{jk}$ on $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}) \times U_{jk}$, then the the map $\varsigma : \mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}) \to \mathcal{L}(\mathcal{M})$ defined to be $g_j^{-1}\varsigma_j$ on U_j is a well defined continuous map from $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$ to $\mathcal{L}$ with the property we seek.

For each D, the functions $g_{jk}(D,\cdot)$ define a cohomology class in the group $H^1(\mathcal{M},\mathcal{O}^*)$. Attached to this cohomology class is the Chern class $c(D) \in H^2(\mathcal{M},\mathbb{Z})$. The exact cohomology sequence associated with the sheaf sequence $0 \to \mathbb{Z} \to \mathcal{O}_{\mathcal{M}} \to \mathcal{O}^*_{\mathcal{M}} \to 0$ shows that, under the hypothesis that $H^1(\mathcal{M},\mathcal{O})$ vanish, the cocycle $\{g_{jk}(D,\cdot)\}_{j,k\in\mathbb{N}}$ is trivial if and only if its Chern class is.

The Chern classes in question may be determined as follows: Since $H^1(\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}),\mathbb{Z})$ vanishes, the continuous zero-free functions g_{jk} have continuous logarithms, say λ_{jk} on U_{jk}. The Chern class $c(D)$ is the cohomology class in $H^2(\mathcal{M},\mathbb{Z})$ determined by the integral 2-cocycle $\{c_{jkl}(D)\}_{j,k,l\in\mathbb{N}}$ given by

$$c_{jkl}(D) = \frac{1}{2\pi i}\{\lambda_{jk} + \lambda_{kl} + \lambda_{lj}\}.$$

This is a continuous integer-valued function, so as $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$ is connected, it is independent of D.

Fix a divisor $D_o \in \mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$. By definition, there is a section $s_o \in \mathcal{L}(\mathcal{M})$ with $\mathrm{Div}\ s_o = D_o$. If we define $\kappa_k = \frac{\varsigma_j(D_o)}{s_o}$, then $\kappa_j \in \mathcal{O}^*(U_j)$, and $\frac{\kappa_j}{\kappa_k} = g_{jk}(D_o,\cdot)$. Thus, the cohomology class $\{c_{jkl}\}_{j,k,l\in\mathbb{N}}$ is trivial: There are integers $\{m_{jk}\}_{jk\in\mathbb{N}}$ with $m_{kl} - m_{jl} + m_{jk} = c_{jkl}$ on U_{jkl}.

Then the family $\{\log g_{jk}(D,\cdot) - m_{jk}\}_{j,k\in\mathbb{Z}}$ is a continuous family of 1-cocycles with values in $\mathcal{O}_{\mathcal{M}}$ for the covering $\mathcal{U}$. Again invoke Michael's selection theorem provides continuous functions g_j on $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M}) \times U_j$ with $g_j(D,z)$ depending holomorphically on $z \in U_j$ and with $g_{jk} = g_j - g_k$.

It follows that if $\varsigma(D)(z) = \frac{\varsigma_j(D)(z)}{g_j(D,z)}$, then ς is a well defined map from $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$ to $\mathcal{L}(\mathcal{M})$ with $\mathrm{Div}\varsigma(D) = D$ for all $D \in \mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$.

The Theorem is proved.

In contrast with the case of holomorphic functions-the case of the trivial bundle, there is no apparent extension of this result to the case of arbitrary domains in a Stein manifold. Whereas if D is a domain in a Stein manifold, the sheaf $\mathcal{O}_D$ of germs of holomorphic functions on D extends naturally to the sheaf $\mathcal{O}_{\tilde{D}}$ of germs of holomorphic functions on the envelope of holomorphy, $\tilde{D}$, of D, in general a locally free sheaf of rank 1 or, equivalently, a holomorphic line bundle, on D will not extend to a locally free sheaf on $\tilde{D}$.

As a particular instance in which the theorem applies, the following may be cited. Let Σ be a simply connected Stein manifold of dimension n. If one has a continuous family $\{D_x\}_{x\in X}$ of divisors each of which is known to be the divisor of a holomorphic n-form on Σ, then there is a corresponding continuous family $\{\Omega_x\}_{x\in X}$ of holomorphic n-forms on Σ with $\operatorname{Div}\Omega_x = D_x$.

As in the earlier-considered case of holomorphic functions, Theorem 8.2 yields some topological information about the space $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$.

Corollary 8.3. *If $\mathcal{M}$ is a complex manifold for which the groups $H^1(\mathcal{M},\mathbb{Z})$ and $H^1(\mathcal{M},\mathcal{O})$ vanish and if $\mathcal{L}\xrightarrow{\pi}\mathcal{M}$ is a holomorphic line bundle with $\mathcal{L}(\mathcal{M})$ infinite dimensional, then the space of divisors $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$ is a contractible ANR. In particular, all of its homotopy groups vanish.*

We conclude this section with some comments on the situation that obtains when the space of sections $\mathcal{L}(\mathcal{M})$ is finite dimensional. This can occur in either of two ways. It may be that $\mathcal{L}(\mathcal{M}) = \{0\}$. In this case $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$ is empty by definition.

Alternatively, $\dim_{\mathbb{C}}\mathcal{L}(\mathcal{M})$ may be a positive integer, say d. Then necessarily $\mathcal{O}(\mathcal{M})$ consists only of the constants: If s is a global section of $\mathcal{L}$ other than zero section, then $f\mapsto fs$ is a linear isomorphism of $\mathcal{O}(\mathcal{M})$ into $\mathcal{L}(\mathcal{M})$. If the latter space is finite dimensional, then necessarily $\mathcal{O}(\mathcal{M})$ reduces to the constants. (Recall here the standing convention that we are dealing *only* with connected manifolds.)

Thus, for $D\in\mathfrak{D}^+(\mathcal{L})$ the fiber $\pi^{-1}(D)$ is $\mathbb{C}^*$. Consequently, $\mathfrak{D}^+_{\mathcal{L}}(\mathcal{M})$ is topologically equivalent to the complex projective space $\mathbb{P}^{d-1}(\mathbb{C})$.

9. Lifting Normal Families.

In this section we remark that, in essence, Stoll's Theorem B has a solution whenever his Problem A has a general solution.

Theorem 9.1. *Let $\mathcal{M}$ be a complex-analytic manifold of dimension $N\geq 1$. Assume that there exists a continuous map $h:\mathfrak{D}^+_P(\mathcal{M})\to\mathcal{O}(\mathcal{M})$ with the property that $\operatorname{Div}h(D) = D$ for all $D\in\mathfrak{D}^+_P(\mathcal{M})$. If $\mathfrak{N}\subset\mathfrak{D}^+_P(\mathcal{M})$ is a compact set of nonnegative principal divisors on $\mathcal{M}$, there exists a compact set $K\subset\mathcal{M}$ and a compact set $\mathcal{K}$ of holomorphic functions on $\mathcal{M}$ with $\operatorname{Div}\mathcal{K} = \mathfrak{N}$ and such that, for each $f\in\mathcal{K}$, there is a point $z_f\in K$ with $f(z_f) = 1$.*

Proof. First, fix a compact set $X_o\subset\mathcal{M}$, and let $W_o\supset X_o$ be an open set with compact closure. Denote by $\mathfrak{N}_o$ the set of all divisors in $\mathfrak{N}$ the support of which is disjoint from W_o. This is plainly a closed subset of $\mathfrak{N}$, so it is a compact set of divisors. Let $h : \mathfrak{D}_P^+(\mathcal{M}) \to \mathcal{O}(\mathcal{M})$ be a continuous map as in the statement of the theorem. As the map is continuous, the set $h(\mathfrak{N}_0)$ is a compact subset of $\mathcal{O}(\mathcal{M})$, which we shall denote by $\mathcal{K}_o$.

Let z_o be a fixed point of the set K, and let ε_o denote the functional of evaluation at the point z_o, so that, for $f \in \mathcal{O}(\mathcal{M})$, $\varepsilon_o(f) = f(z_o)$. This is a continuous linear functional, and on the compact set $\mathfrak{N}_o$ it omits the value zero. Accordingly, there are positive numbers r_o and R_o such that $r_o < |\epsilon_o(f)| < R_o$ for every $f \in \mathcal{K}_o$. Define $\tilde{h}_o : \mathfrak{N} \to \mathcal{O}(\mathcal{M})$ by

$$\tilde{h}_o(D) = \frac{h(D)}{h(D)(z_o)}.$$

Then for all $D \in \mathfrak{N}_o$, we have $\tilde{h}_o(D)(z_o) = 1$. Moreover, the set $\tilde{\mathcal{K}}_o = \{\tilde{h}_o(D) : D \in \mathfrak{N}_o\}$ is a compact set in $\mathcal{O}(\mathcal{M})$.

To continue, denote by $\mathfrak{N}_1$ the subset of $\mathfrak{N}$ that consists of all the divisors in $\mathfrak{N}$ the supports of which meet $\overline{W}_o$. This is a compact set; it is not disjoint from $\mathfrak{N}_o$. Let $\{z_j\}_{j=1,2,\dots}$ be a countable dense set in $\mathcal{M}$, and for each j, let $\{V_{j,k}\}_{k=1,2,\dots}$ be a countable neighborhood basis for z_j that consists of relatively compact open sets. For each $j, k = 1, 2, \dots$, let

$$\mathfrak{N}_{j,k} = \left\{ D \in \mathfrak{N}_1 : (\operatorname{supp} D) \cap \bar{V}_{j,k} = \emptyset \right\}.$$

As $\bar{V}_{j,k}$ is compact, the set $\mathfrak{N}_{j,k}$ is open in $\mathfrak{N}_1$. We have that

$$\bigcup_{j,k=1,2,\dots} \mathfrak{N}_{j,k} = \mathfrak{N}_1,$$

so by compactness a finite number of the $\mathfrak{N}_{j,k}$ cover $\mathfrak{N}_1$. Choose such a finite set, say $\{\mathfrak{N}_{j_\nu,k_\nu}\}_{\nu=1,\dots,q}$. Let z_ν be the z_j associated with $\mathfrak{N}_{j_\nu,k_\nu}$. Each of the sets $\mathfrak{K}_\nu = \overline{\mathfrak{N}}_{j_\nu,k_\nu}, \nu = 1, \dots, q$, is compact. Consequently, each of the sets $\mathcal{K}_\nu = h(\mathfrak{K}_\nu)$ is a compact subset of $\mathcal{O}(\mathcal{M})$. For every ν, let $\tilde{\mathcal{K}}_\nu = \left\{ \frac{f}{f(z_\nu)} : f \in \mathcal{K}_\nu \right\}$. The set $\tilde{\mathcal{K}}_\nu$ is compact, $\operatorname{Div} \tilde{\mathcal{K}}_\nu = \mathfrak{K}_\nu$, and if $f \in \tilde{\mathcal{K}}_\nu$, then $f(z_\nu) = 1$.

The compact set K we seek in $\mathcal{M}$ is the set $\{z_o, z_1, \dots, z_q\}$, and the set $\mathcal{K}$ is the set $\cup_{\nu=1,\dots,q} \tilde{\mathcal{K}}_\nu$.

The proof of the Theorem is completed.

Appendix. Proof of Theorem 1.0.

Our object here is to give a proof of Theorem 1.0 along lines somewhat different from those occurring in Stoll's proof. In fact, we shall prove a formally different theorem:

Theorem 10.1. *If Ω is a domain in $\mathbb{C}^N$ that contains the closed unit ball $\bar{\mathbb{B}}_N$, then there is a map $h : \mathfrak{D}^+(\Omega; 0) \to \mathcal{O}(\mathbb{B}_N)$ that satisfies* $\operatorname{Div} \circ h(D) = D|\mathbb{B}_N$ *for all $D \in \mathfrak{D}^+(\Omega; 0)$ and that is continuous when the space $\mathfrak{D}^+(\Omega; 0)$ is endowed with the relative weak* topology and the space $\mathcal{O}(\mathbb{B}_N)$ is endowed with its usual topology of uniform convergence on compacta.*

This result is formally different from Theorem 1.0 in that here the space of divisors is taken to have the relative weak* topology whereas in Theorem 1.0, it is understood to be endowed with the topology introduced by Stoll. As convergence in Stoll's sense implies convergence in the weak* sense, as we noted in Section 5, the theorem just stated implies Theorem 1.0.

In the proof indicated below, it will be useful to know that on the space $\mathfrak{D}^+(\Omega)$ the relative weak* topology is identical with the relative strong topology. In Section 5 we proved this, but the first proof given there, that is, the proof of the equivalence of the relative weak* topology, the relative strong topology and Stoll's topology, depends on the work in Stoll's paper [**St**], which draws on Theorem 1.0. It was to avoid this circularity that we gave in Section 5 an independent proof of the equivalence of the relative weak* and the relative strong topologies on $\mathfrak{D}^+(\mathcal{M})$.

The proof we give for this follows well known lines; the whole point is to get solutions that vary continuously.

To begin the proof of Theorem 10.1, fix Ω, and fix $R_o > 0$ small enough that $\Omega \supset R_o\bar{\mathbb{B}}_N$. Define $\psi : (\mathbb{C}^N \setminus \{0\}) \times [0, 1] \to \mathbb{C}^N$ to be the real-analytic map given by

$$\psi(z, t) = (1 - t)z + tR_o \frac{z}{|z|}.$$

The partial map $\psi(\cdot, 0)$ is the identity, and $\psi(\cdot, 1)$ is the radial retraction of $\mathbb{C}^N \setminus \{0\}$ onto $R_o\mathbb{S}^{2N-1}$.

For an irreducible complex hypersurface, V, in Ω that does not pass through the origin, define a current $T_V \in \mathcal{D}_1(R_o\mathbb{B}_N)$ by the condition that if $\beta \in \mathcal{D}^{2N-1}(R_o\mathbb{B}_N)$, then

$$T_V(\beta) = \int_{V\times[0,1]} \psi^*\beta. \tag{4}$$

The integral is well defined, for since the support of β is a compact subset of $R_o\mathbb{B}_N$, the support of $\psi^*\beta$ is a compact subset of $(V \times [0, 1)$. The variety V has locally finite volume (of dimension $2N-2$) in Ω, so the real variety $V\times\mathbb{R}$ has locally finite volume (of dimension $2N - 1$) in $\Omega \times \mathbb{R}$. This implies that if $\gamma \in \mathcal{D}^{2N-1}(\Omega \times \mathbb{R})$, then $\int_{V\times\mathbb{R}} \gamma$ is defined and that the map $\gamma \mapsto \int_{V\times\mathbb{R}} \gamma$ is an element, $[V \times \mathbb{R}]$, of $\mathcal{D}_{2N-1}(\mathbb{C}^N \times \mathbb{R})$. It also implies that if γ is *any* $(2N - 1)-$form on $\mathbb{C}^N \times \mathbb{R}$ with locally bounded, measurable coefficients such that $\operatorname{supp}\gamma \cap (V \times \mathbb{R})$ is compact, then the integral $\int_{V\times\mathbb{R}} \gamma$ exists. In particular $T_V(\beta)$ is defined when $\beta \in \mathcal{D}^{2N-1}(R_o\mathbb{B}_N)$.

The current T_V satisfies

$$dT_V = -[V]|R_o\mathbb{B}_N \tag{5}$$

where, as usual, $[V]$ denotes the current of integration over the variety V. This fact is simply Stokes's theorem: By definition, $dT_V(\alpha) = T_V(d\alpha)$, which gives

$$T_V(d\alpha) = \int_{V\times[0,1]} \psi^* d\alpha = \int_{V\times[0,1]} d\psi^*\alpha,$$

and, by Stokes's theorem, this is

$$\int_{b(V\times[0,1])} \psi^*\alpha = \int_{b(V\times[0,1])} \psi^* d\alpha = -\int_V \alpha.$$

If $D = \sum_j m_j V_j$ is a nonnegative divisor on Ω with $V_1, V_2, \dots$ distinct, irreducible complex hypersurfaces in Ω none of which pass through the origin, define

$$T_D = \sum_j m_j T_{[V_j]}. \tag{6}$$

The family $V_j, j = 1, 2, \dots$, is locally finite in Ω, so the sum (6) is finite for each choice of D. We have

$$dT_D = -\sum_j m_j [V_j] | R_o\mathbb{B}_N.$$

Lemma 10.2. *The map $D \mapsto T_D$ is continuous from $\mathfrak{D}^+(\Omega; 0)$ to $\mathcal{D}_1(R_o\mathbb{B}_N)$ when the two spaces are given their respective weak* topologies.*

Proof. Given a net $\{D_\iota\}_{\iota\in I}$ in $\mathfrak{D}^+(\Omega; 0)$ that converges to $D_o \in \mathfrak{D}^+(\Omega; 0)$ in the sense that for each $\beta \in \mathcal{D}^{2N-2}(\Omega)$, $\lim_{\iota\in I} D_\iota(\beta) = D_o(\beta)$, we are to prove that for each $\beta \in \mathcal{D}^{2N-1}(R_o\mathbb{B}_N)$, $\lim T_{D_\iota}(\beta) = T_{D_o}(\beta)$, i.e., that if $D_\iota = \sum_j m_{\iota j} V_{\iota j}$ with $\{V_{\iota j}\}$ for fixed ι a locally finite family of irreducible complex hypersurfaces in Ω and if $D_o = \sum_j m_{oj} V_{oj}$ is the corresponding decomposition of D_o, then

$$\lim_\iota \sum_j m_{\iota j} \int_{V_{\iota j}\times[0,1]} \psi^*\beta = \sum_j m_{oj} \int_{V_{oj}\times[0,1]} \psi^*\beta \tag{7}$$

for each $\beta \in \mathcal{D}^{2N-1}(R_o\mathbb{B}_N)$.

To this end, note first that there is $\delta_o > 0$ sufficiently small that for $\iota \in I$ sufficiently large, $\mathrm{supp} D_\iota \cap \delta_o\mathbb{B}_N = \emptyset$.

As $D_\iota \to D_o$ in the weak* sense, the convergence also takes place in the sense of the strong topology, i.e., uniformly on bounded sets in $\mathcal{D}^{2N-1}(\Omega)$. (Recall the discussion at the end of Section 3.)

With $\beta \in \mathcal{D}^{2N-1}(\Omega)$, write $\beta = \sum_{|J|+|K|=2N-1} b_{JK} dz^J \wedge d\bar{z}^K$ for a suitable choice of functions $b_{JK} \in \mathcal{D}^0(R_o\mathbb{B}_N)$. The support of β is contained in the ball $(R_o - \delta_1)\mathbb{B}_N$ for some $\delta_1 > 0$. As $\psi(z,t) = (1-t)z + tR_o(\frac{z}{|z|}) \geq |z|$, it follows that for every $t \in [0,1]$ and for all J, K $\mathrm{supp}\, b_{JK}(\psi(z,t))$, qua function of z, is contained in the ball $(R_o - \delta_1)\mathbb{B}_N$. Moreover, if $\epsilon_o > 0$ is fixed, then the derivatives with respect to z of order no more than k of the functions $b_{JK}(z,t)$ are bounded uniformly in t (and in J and K) on the spherical region $\epsilon_o < |z| < R_o - \delta_1$. That is to say, the set $\{b_{JK}(\cdot, t)\}_{t\in[0,1]}$ is a bounded set in $\mathcal{D}^0(R_o\mathbb{B}_N \setminus \epsilon_o\mathbb{B}_N)$.

Write $\psi^*\beta = B' + B'' \wedge dt$ where $B' \in D^{2N-1}(R_o\mathbb{B}_N)$ does not contain the factor dt and where $B'' \in \mathcal{D}^{2N-2}(R_o\mathbb{B}_N)$ is also free of the factor dt.

Then

$$T_{D_\iota}(\beta) = \sum_j m_{\iota j} \int_0^1 \left(\int_{V_{\iota j}} B'' \right) dt.$$

The last paragraph implies that the family $B'' = \sum_{J,K} b''_{JK}(z,t)dz^J \wedge d\bar{z}^K$ of forms of degre $2N-2$ in z and $\bar{z}$ indexed by the parameter $t \in [0,1]$ is a bounded family in $\mathcal{D}^{2N-2}(R_o\mathbb{B}_N)$. Thus the convergence of $\sum_j m_{\iota j} \int_{V_{\iota j}} B''$ to $\sum_j m_{oj} \int_{V_{oj}} B''$ is uniform in t. Accordingly, we can integrate with respect to $t, t \in [0,1]$ to find that, as desired, $T_{D_\iota}(\beta) \to T_{D_o}(\beta)$.

The lemma is proved.

As noted above, if $D \in \mathfrak{D}^+(\Omega; 0)$, then the current $-T_D$ satisfies $d(-T_D) = [D]|(R_o\mathbb{B}_N)$. Set $-T_D = S_{0,1} + S_{1,0}$ with $S_{0,1} \in \mathcal{D}_{0,1}(R_o\mathbb{B}_N)$ and $S_{1,0} \in \mathcal{D}_{1,0}(R_o\mathbb{B}_N)$. As $d = \partial + \bar{\partial}$ and as $[D]$ is of bidegree $(1,1)$, it follows that the current $S_{0,1}$ satisfies the equation $\bar{\partial}S_{0,1} = 0$ on $R_o\mathbb{B}_N$. As a $\bar{\partial}-$closed current, it is $\bar{\partial}$-exact. We want to see that it is possible to choose a $\bar{\partial}-$primitive for $S_{0,1}$ that depends continuously on T_D and so continuously on D. In fact, this analysis will be carried out not on the full ball $R_o\mathbb{B}_N$ but rather on the smaller ball $\mathbb{B}_N$.

We shall need the observation that T_D is a current with measure coefficients, whence the same is true of S_{01}.

Recall the well known explicit solution for $\bar{\partial}$ on the ball that is given in detail, e.g., in [**Ru2**]. Let $\eta : \mathbb{B}_N \times \bar{\mathbb{B}}_N \to [0,1]$ be a function of class $\mathcal{C}^\infty$ that satisfies i) $\eta = 1$ near the diagonal $\Delta = \{(z,z) : z \in \mathbb{B}_N\}$ and ii) $\eta = 0$ on a neighborhood of $\mathbb{B}_N \times b\mathbb{B}_N$. The map $s : \mathbb{B}_N \times \bar{\mathbb{B}}_N \to \mathbb{C}^N$ is defined by

$$s(z,\zeta) = \eta(z,\zeta)(\zeta - z) + [1 - \eta(z,\zeta)](\bar{\zeta} - \bar{z}).$$

In terms of s the kernel K_s is defined by

$$K_s(z,\zeta) = \langle \zeta - z, s(z,\zeta)\rangle^{-N} \omega'(\bar{s}(z,\zeta)) \wedge \omega(\zeta).$$

In this definition, the following notation is used. $\omega(\zeta)$ denotes the holomorphic $N-$form $d\zeta_1 \wedge \cdots \wedge d\zeta_N$. If the vector $s(z,\zeta)$ has coordinates $(s_1, \ldots, s_N)$, then $\omega'(\bar{s}(z,\zeta))$ is the $(0, N-1)-$form given by

$$\omega'(\bar{s}(z,\zeta)) = \sum_{j=1}^N (-1)^{j-1} s_j \bar{\partial}_\zeta s_1 \wedge \cdots [j] \cdots \wedge \bar{\partial}_\zeta s_N.$$

Finally, $\langle \zeta - z, s(z,\zeta)\rangle = \sum_{j=1}^N (\zeta_j - z_j)s_j$. An application of the kernel K_s is that with it, one can solve $\bar{\partial}$ as follows. (See [**Ru2**, p. 351]). If

$\alpha = \sum_{j=1}^{N} a_j d\bar{\zeta}_j$ is a $\bar{\partial}-$closed form of bidegree $(0,1)$ on $\bar{\mathbb{B}}_N$ with coefficients of class $\mathcal{C}^1$ then the function u_α given for $z \in \mathbb{B}_N$ by

$$(8) \qquad u_\alpha(z) = \frac{1}{Nc_N} \int_{\mathbb{B}_N} K_s(z,\zeta) \wedge \alpha(\zeta)$$

is of class $\mathcal{C}^1$ on $\mathbb{B}_N$ and satisfies there $\bar{\partial} u_\alpha = \alpha$.

In this statement, c_N denotes the constant $\frac{1}{N!}(-1)^{N(N-1)/2}(2\pi i)^N$.

Write

$$K_s(z,\zeta) = \langle \zeta - z, s(\zeta,z)\rangle^{-N} \sum_{j=1}^{N} \Theta_j(z,\zeta)\omega_{[j]}(\bar{\zeta}) \wedge \omega(\zeta)$$

where by $\omega_{[j]}(\bar{\zeta})$ we understand the form obtained from $\omega(\bar{\zeta})$ by deleting the factor $d\bar{\zeta}_j$. Thus,

$$K_s(z,\zeta) \wedge \alpha = \langle \zeta - z, s(\zeta,z)\rangle^{-N} \sum_{j=1}^{N} (-1)^j \Theta_j(z,\zeta) a_j(\zeta) \omega(\bar{\zeta}) \wedge \omega(\zeta).$$

In this, Θ_j is a combination of the functions s_j and their first derivatives with respect to $\bar{\zeta}$.

As remarked above, the current $S_{0,1}$ has measure coefficients, say

$$S_{0,1} = \sum_{j=1}^{N} \mu_j d\bar{\zeta}_j.$$

Consider then the function

$$(9) \qquad U_D(z) = \frac{1}{nc'_n} \sum_{j=1}^{N} \int \frac{(-1)^j \Theta_j(z,\zeta)}{\langle \zeta - z, s(\zeta,z)\rangle^N} d\mu_j(\zeta).$$

Here the constant c'_N is the constant chosen so that $\frac{1}{c_N}\omega(\bar{\zeta}) \wedge \omega(\zeta) = \frac{1}{c'_N} d\mathcal{L}$ where we understand by $d\mathcal{L}$ Lebesgue measure on $\mathbb{C}^N$. Notice that the integration in (8) is supported in $\cup \operatorname{supp} \mu_j$. The function U_D satisfies $\bar{\partial} u_D = S_{0,1}$.

The last assertion requires a preliminary remark about regularity. The denominator $\langle \zeta - z, s(\zeta,z)\rangle$ has positive real part away from the diagonal, and near the diagonal, it agrees with the $|\zeta - z|^{2N}$. It follows that the term $\frac{\Theta_j(z,\zeta)}{\langle \zeta - z, s(\zeta,z)\rangle^N}$ is majorized by $\frac{\text{constant}}{|\zeta - z|^{2N-1}}$. As the convolution of a measure with compact support and a locally integrable function is locally integrable, it follows that the function U_D is locally integrable on $\mathbb{B}_N$. Thus, the equation $\bar{\partial} U_D = \ S_{0,1}$ is at least meaningful in the sense of distributions.

To prove that this equation is *correct*, we argue as follows. We are to see that if $\beta = \sum_{j=1}^{N} \beta_j \omega(\zeta) \wedge \omega_{[j]}(\bar{\zeta})$ is a smooth form on $\mathbb{B}_N$ with compact support, then

$$\int_{\mathbb{B}_N} U_D \bar\partial\beta = S_{0,1}(\beta) = \sum_{j=1}^N \int \beta_j(z) d\mu_j(z).$$

Introduce a smooth approximate identity $\{\chi_\epsilon\}_{\epsilon>0}$ with $\chi_\epsilon(z) = \epsilon^{-2N}\chi(\frac{z}{\epsilon^{2N}})$ for a nonnegative compactly supported smooth even function χ on $\mathbb{C}^n$ with integral one. Introduce also the convolution $S^\epsilon = \chi_\epsilon * S_{0,1}$.

This is a smooth form: $S^\epsilon = \sum_{j=1}^N \chi_\epsilon * \mu_j d\bar\zeta_j$, and it is $\bar\partial-$closed, for $S_{0,1}$ is $\bar\partial-$closed. Let u_{S^ϵ} be the solution of $\bar\partial u = S^\epsilon$ given by (8), so that, by definition,

$$u_{S^\epsilon}(z) = \frac{1}{Nc'_N}\int_{|\zeta|<1}\int \sum_{j=1}^N \frac{(-1)^j\Theta_j(z,\zeta)\chi_\epsilon(\zeta-\xi)}{\langle\zeta - z, s(\zeta,z)\rangle^N} d\mu_j(\xi) d\mathcal{L}(z).$$

For a smooth, compactly supported form β on $\mathbb{B}_N$

$$\begin{aligned}\int u_{S^\epsilon}\bar\partial\beta &= \int \bar\partial u_{S^\epsilon}\wedge\beta \\ &= \int S^\epsilon\wedge\beta \to S_{0,1}(\beta) = \sum_{j=1}^N\int\beta_j d\mu_j.\end{aligned}$$

Also, if b is the compactly supported smooth function on $\mathbb{B}_N$ such that

$$\int u_{S^\epsilon}\bar\partial\beta = \int \bar\partial u_{S^\epsilon}(z) b(z) d\mathcal{L}(z),$$

then

$$\begin{aligned}&\int u_{S^\epsilon}\bar\partial\beta \\ &= \int \frac{1}{Nc'_N}\sum_{j=1}^N\int\int \frac{(-1)^j\Theta_j(z,\zeta)\chi_\epsilon(\zeta-\xi)b(z)}{\langle\zeta - z, s(\zeta,z)\rangle^N} d\mu_j(\xi)d\mathcal{L}(\zeta)d\mathcal{L}(z) \\ &= \int \frac{1}{Nc'_N}\sum_{j=1}^N\int\int \frac{(-1)^j\Theta_j(z,\zeta)\chi_\epsilon(\zeta-\xi)b(z)}{\langle\zeta - z, s(\zeta,z)\rangle^N} d\mathcal{L}(\zeta)d\mathcal{L}(z)d\mu_j(\xi).\end{aligned}$$

As $\epsilon\to 0$, this tends to

$$\begin{aligned}&\frac{1}{Nc'_N}\sum_{j=1}^N\int\int\frac{(-1)^j\Theta_j(z,\xi)b(z)}{\langle(\xi - z, s(\xi,z))\rangle^N} d\mathcal{L}(z)d\mu_j(\xi) \\ &= \int U_D(z)b(z)d\mathcal{L}(z).\end{aligned}$$

This gives the desired equation

$$\int U_D \bar{\partial}\beta = S_{0,1}(\beta).$$

The integral (9) defines $U_D \in L^1_{loc}(\mathbb{B}_N)$, and U_D satisfies $\bar{\partial} U_D = S_{0,1}$, whence $\partial\bar{\partial} u_D = D$ on $\mathbb{B}_N$. The divisor $D|\mathbb{B}_N$ is principal: $D|\mathbb{B}_N = \mathrm{Div} f$ for some $f \in \mathcal{O}(\mathbb{B}_N)$, so $D|\mathbb{B}_N = \partial\bar{\partial} \log|f|$ also. Thus the function $w_D = U_D - log\,|f|$ is pluriharmonic on $\mathbb{B}_N$. This implies that U_D is pluriharmonic on $\mathbb{B}_N \setminus \mathrm{supp} D$. It follows further that $U_D = \log\,|fe^g|$ for some $g \in \mathcal{O}(\mathbb{B}_N)$, and thus, if we denote by V_D the pluriharmonic conjugate of U_D that vanishes at the origin, then $e^{U_D + iV_D} = fe^g$ is holomorphic and has divisor $D|\mathbb{B}_N$.

Finally, we must verify that the map $D \mapsto e^{U_D + iV_D}$ is continuous with respect to the weak* (or strong) topology on the space of divisors and the topology of uniform convergence on compacta on the space of holomorphic functions on $\mathbb{B}_N$. Granted that we are specifying that V_D is the pluriharmonic conjugate of U_D that vanishes at the origin, the function V_D depends continuously on U_D. Thus, it is enough to see that as $D_\iota \to D_o$, $U_{D_\iota}(z) \to U_{D_o}(z)$ uniformly on compacta in $\mathbb{B}_N \setminus \mathrm{supp}\, D_o$. For this, notice that if K is a compact subset of $\mathbb{B}_N \setminus \mathrm{supp}\, D_o$, then there is an open set $W \subset \mathbb{B}_N$ with $\mathrm{supp}\, D_\iota \cap W = \emptyset$ for all $\iota > \iota_o$. Since the convergence of D_ι to D_o is in the strong sense, it follows that the corresponding functions U_{D_ι} converge to U_{D_o} uniformly on K, and we are done.

References

[AN] A. Andreotti and F. Norguet, *Quelques proprietés de courants définies à l'aide de fonctions holomorphes*, Anais Acad. Brasileira de Cien., **37** (1965), 385-388.

[Ba] T.J. Barth, *Families of nonnegative divisors*, Trans. American Math. Soc., **131** (1968), 223-245.

[Bas] R. Basner, *Several dimensional properties of the spectrum of a uniform algebra*, Pacific J. Math., **74** (1978), 297-306.

[BP] C. Bessaga and A. Pełczyński, *Selected Topics in Infinite-Dimensional Topology*, PWN-Polish Scientific Publishers, Warsaw, 1975.

[Bo] K. Borsuk, *Theory of Retracts*, PWN-Polish Scientific Publishers, Warsaw, 1966.

[BT] R. Bott and L.W. Tu, *Differential Forms in Algebraic Topology*, Springer-Verlag, New York, Heidelberg Berlin, 1982.

[GS] I.M. Gel'fand and G.E. Shilov, *Generalized Functions,* I, Academic Press, London and New York, 1964.

[Gu] R.C. Gunning, *Introduction to Holomorphic Functions of Several Complex Variables*, III, Wadsworth Brooks Cole, Belmont, California, 1990.

[Hu] S.T. Hu, *Theory of Retracts*, Wayne State University Press, Detroit, 1965.

[Ke] J.L. Kelley, *General Topology*, Springer-Verlag, New York, Heidelberg, Berlin, 1961.

[Kl] V.L. Klee *A note on topological properties of normed spaces*, Proc. American Math. Soc., **7** (1956), 673-674.

[Ma] S. Mardesič, *Equivalence of singular and Čech homology for ANR-s. Application to unichoerence*, Fundamenta Mathematicae, **XLVI** (1958), 31-45.

[MG] G.L. McGrath, *Normal families on non-negative divisors on polycylinders*, Ph.D. Thesis, Notre Dame University, Notre Dame, Indiana, 1965.

[Mi] E. Michael, *Continuous selections*, I, Ann. Math., **63** (1956), 361-382.

[Na] R. Narasimhan, *The Levi problem for complex spaces*, Math. Ann., **142** (1961), 355-365.

[Ok] K. Oka, *Deuxième problème de Cousin*, J. Sci. Hiroshima Univ., **9** (1939), 7-19.

[Ro] H. Rossi, *On envelopes of holomorphy*, Comm. Pure Appl. Math., **16** (1963), 9-17.

[Ru1] W. Rudin, *Function Theory on Polydiscs*, Benjamin, New York and Amsterdam, 1969.

[Ru2] ———, *Function Theory on the Unit Ball of* $\mathbb{C}^N$, Springer-Verlag, New York, Heidelberg, Berlin, 1980.

[Si] Y.-T. Siu *Normal families of Cousin* I *and Cousin* II *data*, Tohoku Math. J., **21** (1969), 548-557.

[St] W. Stoll, *Normal families of nonnegative divisors*, Math. Zeit., **84** (1964), 154-218.

Received July 31, 1997 and revised March 17, 1998. The first author's research supported in part by 40% MURST funds. Research supported in part by Grant DMS-9322326 from the National Science Foundation. In addition, part of the work was carried out while the second author was a member of the Mathematical Sciences Research Institute at Berkeley.
Professor Lupacciolu met an untimely death on December 31, 1996. At that time, this work was essentially complete.

Universita della Basilicata
85100 Potenza
Italy

University of Washington
Seattle, WA 98195-4350
E-mail address: stout@math.washington.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

THE ESSENTIAL NORM OF A COMPOSITION OPERATOR ON BLOCH SPACES

Alfonso Montes-Rodríguez

We express the essential norm of a composition operator on the Bloch space and the little Bloch space as the asymptotic upper bound of a quantity involving the inducing map and the Pick-Schwarz Lemma. As a consequence, we obtain a new proof of a recently obtained characterization of the compact composition operators on Bloch spaces.

1. Introduction.

Let $\mathbb{D}$ denote the unit disk in the complex plane. A function f analytic on the unit disk is said to belong to the Bloch space $\mathcal{B}$ if

$$\sup_{\mathbb{D}}(1-|z|^2)|f'(z)| < \infty$$

and to the little Bloch space $\mathcal{B}_0$ if

$$\lim_{|z|\to 1^-}(1-|z|^2)|f'(z)| = 0.$$

It is well known and easy to prove that $\mathcal{B}$ is a Banach space under the norm

$$\|f\| = |f(0)| + \sup_{\mathbb{D}}(1-|z|^2)|f'(z)|$$

and that $\mathcal{B}_0$ is a closed subspace of $\mathcal{B}$. Good sources for results and references about Bloch functions are the papers of Anderson-Clunie-Pommerenke [**ACP**], Fernández [**Fe**], Pommerenke [**Po**], and the book of Zhu [**Zh**, Chapter 5].

If φ is an analytic function on $\mathbb{D}$ with $\varphi(\mathbb{D}) \subset \mathbb{D}$, then the equation $C_\varphi f = f \circ \varphi$ defines a composition operator C_φ on the space of all holomorphic functions on $\mathbb{D}$. The Pick-Schwarz Lemma (see [**CM**, p. 47], for instance) asserts that

$$\frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| \leq 1. \tag{1}$$

As noticed in [**MM**] this and the chain rule give an easy proof of the fact that C_φ acts boundedly on the Bloch space. In fact we have

$$\begin{aligned}(1-|z|^2)|(f\circ\varphi)'(z)| &= (1-|z|^2)|f'(\varphi(z))||\varphi'(z)|\\ &= \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|(1-|\varphi(z)|^2)|f'(\varphi(z))|\\ &\le \sup_{\mathbb{D}}(1-|\varphi(z)|^2)|f'(\varphi(z))|\\ &= \sup_{\varphi(\mathbb{D})}(1-|w|^2)|f'(w)|\\ &\le \sup_{\mathbb{D}}(1-|z|^2)|f'(z)|.\end{aligned}$$

In addition, if C_φ acts boundedly on $\mathcal{B}_0$ then φ must belong to $\mathcal{B}_0$. This follows from the fact that $C_\varphi z=\varphi$. Conversely, if $\varphi\in\mathcal{B}_0$, then from the estimates above it is easy to show that φ induces a continuous operator on $\mathcal{B}_0$ (see [**MM**]). The main goal of this paper is to compute the essential norm of C_φ in terms of an asymptotic bound involving the quantity

$$\frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|.$$

We recall that the essential norm of a continuous linear operator T is the distance from T to the compact operators, that is,

$$\|T\|_e=\inf\{\|T-K\| : K \text{ is compact}\}.$$

Notice that $\|T\|_e=0$ if and only if T is compact, so that estimates on $\|T\|_e$ lead to conditions for T to be compact. Thus we will obtain a different proof of a recent result of Madigan and Matheson [**MM**] in which they characterize those φ which induces compact composition operators on $\mathcal{B}$ and $\mathcal{B}_0$. The fundamental ideas of the proof are those used by J.H. Shapiro [**Sh1**] to obtain the essential norm of a composition operator on Hilbert spaces of analytic functions (Hardy and weighted Bergman spaces) in terms of natural counting functions associated with φ. However, since neither $\mathcal{B}$ nor $\mathcal{B}_0$ are Hilbert spaces our method differs in some interesting details from those of Shapiro.

Before going further, we want to say a word about the well-known heuristic principle which states that if a "big-oh" condition describes a class of bounded operators, then the corresponding "little-oh" condition picks out the subclass of compact operators. An excellent example of this principle in action can be seen in the paper of J.H. Shapiro [**Sh1**] mentioned above. The "big-oh" condition on Bloch spaces is given by (1). Madigan and Matheson were able to prove the "little-oh" condition, that is, that a composition operator C_φ on $\mathcal{B}_0$ is compact if and only if

$$\lim_{|z|\to 1^-} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| = 0.$$

They also obtained (with a different proof) that C_φ is compact on $\mathcal{B}$ if and only if for every $\varepsilon > 0$ there exists r, $0 < r < 1$, such that

$$\sup_{|\varphi(z)|>r} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| < \varepsilon.$$

As we will see later the conditions of compactness on $\mathcal{B}$ and $\mathcal{B}_0$ are actually the same. In fact, the essential norm of a composition operator is independent of the underlying space $\mathcal{B}$ or $\mathcal{B}_0$. This should not cause any surprise. The fact that $\mathcal{B}$ is isometrically isomorphic to the second dual of $\mathcal{B}_0$ and the inclusion $\mathcal{B}_0 \subset \mathcal{B}$ corresponds to the canonical imbedding of $\mathcal{B}_0$ into $\mathcal{B}_0^{\star\star}$ (see [**ACP**]) does not affect the computation of the essential norm. This is exactly what happens if we consider a bounded diagonal operator defined by a bounded sequence $\{a_n\}$ on the sequence spaces l^∞ and c_0, respectively. Then its essential norm equals $\limsup a_n$ and this quantity is independent of the underlying space. In fact the proof of the main result in the following section is done simultaneously for both $\mathcal{B}$ and $\mathcal{B}_0$.

Before proceeding further, the author would like to thank Nigel J. Kalton who indicated the proof of Proposition 2.3. The author would also like to thank Joel H. Shapiro for providing the proof of Theorem 2.5, some references and helpful comments.

2. Main result.

Main Theorem 2.1. *Suppose that C_φ defines a continuous operator on $\mathcal{B}$ (or on $\mathcal{B}_0$). Then*

$$\|C_\varphi\|_e = \lim_{s\to 1^-} \sup_{|\varphi(z)|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|. \tag{1}$$

In particular, C_φ is compact on $\mathcal{B}$ (or $\mathcal{B}_0$) if and only if

$$\lim_{s\to 1^-} \sup_{|\varphi(z)|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| = 0.$$

It is understood that if $\{z : |\varphi(z)| > s\}$ is the empty set for some $0 < s < 1$ the supremum equals zero. This happens when $\varphi(\mathbb{D})$ is a relatively compact subset of $\mathbb{D}$ and in this case it is easy to show that C_φ is a compact operator.

If φ has an angular derivative at a point $\xi \in \partial\mathbb{D}$, then we can apply the Julia Carathéodory Theorem (see [**Sh2**, p. 57]) and the Pick-Schwarz

Lemma to obtain

$$1 = \liminf_{z \to \xi} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| \le \lim_{s \to 1^-} \sup_{|\varphi(z)|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| \le 1.$$

Thus, as an immediate consequence of Theorem 2.1 we have $\|C_\varphi\|_e = 1$ whenever φ has a finite angular derivative.

Before proving Theorem 2.1 let us show that for the little Bloch space $\mathcal{B}_0$ there is an equivalent formula in terms of another quantity. This a simple consequence of the following proposition:

Proposition 2.2. *Suppose that C_φ defines a continuous operator on $\mathcal{B}_0$. Then*

$$\lim_{s \to 1^-} \sup_{|\varphi(z)|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| = \limsup_{|z| \to 1^-} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|. \tag{2}$$

Proof. As remarked in the introduction the fact that C_φ acts boundedly on $\mathcal{B}_0$ implies that $\varphi \in \mathcal{B}_0$. If $\varphi(\mathbb{D})$ is a relatively compact subset of $\mathbb{D}$, then both limits in (2) are zero and coincide. So we may suppose that $\varphi(\mathbb{D})$ is not a relatively compact subset of $\mathbb{D}$. Let $0 < s_n < 1$ be any increasing sequence tending to 1. We set $t_n = \inf\{t : |\varphi(z)| > s_n$ for some z with $|z| > t\}$. By continuity $\{t_n\}$ also tends to 1. Since $\{z : |z| > t_n\} = \{z : |\varphi(z)| > s_n$ and $|z| > t_n\} \cup \{z : |\varphi(z)| \le s_n$ and $|z| > t_n\}$ we find that the left hand side of (2) is less than or equal to the right hand side of (2). On the other hand, we can always find a sequence $\{z_n\}$ for which

$$\begin{aligned} \lim_{n \to \infty} \frac{1-|z_n|^2}{1-|\varphi(z_n)|^2}|\varphi'(z_n)| &= \lim_{s \to 1^-} \sup_{|z|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| \\ &= \limsup_{|z| \to 1^-} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|. \end{aligned} \tag{3}$$

Then either there is a subsequence $\{z_{n_k}\}$ such that $\{|\varphi(z_{n_k})|\} \to 1$ as $k \to \infty$, or for every positive integer n we have $|\varphi(z_n)| \le s_0$ for some $0 < s_0 < 1$. Clearly, in the former case both limits in (2) coincide. For the latter case we find that the limit in (3) is zero because $\varphi \in \mathcal{B}_0$. Since this limit is greater than or equal to the limit on the left hand side of (2), we find that they are the same again. The proof is now finished. □

Now we turn to the proof of our main result.

The lower estimate. First we show that:

$$\|C_\varphi\|_e \ge \lim_{s \to 1^-} \sup_{|\varphi(z)| \ge s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|. \tag{4}$$

Instead of the reproducing kernels used by Shapiro for the Hardy and Bergman spaces we will use the sequence $\{z^n\}_{n \ge 2}$. This sequence converges uniformly on compact subsets of the unit disk. An elementary computation shows that

$$\|z^n\| = \max_{\mathbb{D}}(1-|z|^2)|nz^{n-1}| = \frac{2n}{n+1}\left(\frac{n-1}{n+1}\right)^{(n-1)/2}.$$

Observe that for each $n \geq 2$ the above maximum is attained at any point on the circle centered at the origin and of radius $r_n = \left(\frac{n-1}{n+1}\right)^{1/2}$. These maxima form a decreasing sequence which tends to $2/e$.

Therefore, the sequence $\{z^n\}_{n\geq 2}$ is bounded away from zero. Now we consider the normalized sequence $\{f_n = \frac{z^n}{\|z^n\|}\}$ which also tends to zero uniformly on compact subsets of the unit disk. For each $n \geq 2$ we define the closed annulus $A_n = \{z \in \mathbb{D} : r_n \leq |z| \leq r_{n+1}\}$ and compute

$$\begin{aligned}\min_{A_n}(1-|z|^2)|f_n'(z)| &= (1-r_{n+1}^2)|f_n'(r_{n+1})| \\ &= \left(\frac{n+1}{n+2}\right)\left(\frac{n^2+n}{n^2+n-2}\right)^{(n-1)/2}. \qquad (5)\end{aligned}$$

Observe that these minima tend to 1 as $n \to \infty$ and for each $n \geq 2$ the minimum above is attained at any point of the circle centered at the origin and of radius r_{n+1}. For the moment fix any compact operator K on $\mathcal{B}_0$ or $\mathcal{B}$. The uniform convergence on compact subsets of the sequence $\{f_n\}$ to zero and the compactness of K imply that $\|Kf_n\| \to 0$. It is easy to show that if a bounded sequence that is contained in $\mathcal{B}_0$ converges uniformly on compact subsets of the unit disk, then it also converges weakly to zero in $\mathcal{B}_0$ as well as in $\mathcal{B}$. Thus

$$\begin{aligned}\|C_\varphi - K\| &\geq \limsup_n \|(C_\varphi - K)f_n\| \\ &\geq \limsup_n(\|C_\varphi f_n\| - \|Kf_n\|) \\ &= \limsup_n \|C_\varphi f_n\|.\end{aligned}$$

Upon taking the infimum of both sides of this inequality over all compact operators K, we obtain the lower estimate:

$$\begin{aligned}\|C_\varphi\|_e &\geq \limsup_n \|C_\varphi f_n\| \\ &= \limsup_n \sup_{\mathbb{D}}(1-|z|^2)|f_n'(\varphi(z))||\varphi'(z)| \\ &= \limsup_n \sup_{\mathbb{D}} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|(1-|\varphi(z)|^2)|f_n'(\varphi(z))|. \qquad (6)\end{aligned}$$

Now (6) is greater than or equal to

$$\limsup_n \sup_{\varphi(z)\in A_n} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|(1-|\varphi(z)|^2)|f_n'(\varphi(z))| \qquad (7)$$

and (7) is greater than or equal to

$$\limsup_n \sup_{\varphi(z)\in A_n} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)| \min_{\varphi(z)\in A_n}(1-|\varphi(z)|^2)|f_n'(\varphi(z))|. \tag{8}$$

If $\varphi(\mathbb{D})$ is a relatively compact subset of $\mathbb{D}$ both sides of (4) are zero and there is nothing to prove. Otherwise we find that $\min_{\varphi(z)\in A_n}(1-|\varphi(z)|^2)|f_n'(\varphi(z))|$ $= \min_{A_n}(1-|z|^2)|f_n'(z)|$ because the minimum in (5) is attained at any point on the circle centered at the origin and of radius r_{n+1}. Since these minima tend to 1 as $n \to \infty$, it follows that (8) is equal to

$$\limsup_n \sup_{\varphi(z)\in A_n} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|. \tag{9}$$

Finally, an easy exercise shows that (9) coincides with the right hand side of (4).

To obtain the upper estimate in the case of the Hardy and Bergman spaces, Shapiro [**Sh1**] used the operators P_n which take f to the nth partial sum of its Taylor series. On the Hardy space these operators satisfy: i) Each P_n is compact, ii) $(I-P_n)f$ tends to zero uniformly on compact subsets for any f in the Hardy space, and iii) for each n the norm in the Hardy space of $I-P_n$ equals 1. Although each P_n is also compact in the Bloch space, and $(I-P_n)f$ tends to zero uniformly on compact subsets for each function $f \in \mathcal{B}$, this sequence does not satisfy anything analogous to iii) above. In fact, $\|P_n\| \geq C\log n$ where C is a universal constant (see [**ACP**, p. 18-19]). Therefore, by the reverse triangle inequality $\|I-P_n\| \geq C\log n - 1$. One of the issues here is that in general it is not easy to compute exactly either the norms of Bloch functions, or the norms of operators defined on Bloch spaces.

To obtain the upper estimate we need the operators K_n, $n \geq 2$, which take each function $f(z)$ to $f(\frac{n-1}{n}z)$. Every operator K_n is compact on $\mathcal{B}$ (or $\mathcal{B}_0$). We also have that $(I-K_n)f$ tends to zero uniformly on compact subsets of the unit disk for every $f \in \mathcal{B}$, and (although we do not know if $\lim_{n\to\infty}\|I-K_n\| = 1$) we have the following proposition, whose proof is delayed.

Proposition 2.3. *There exists a sequence of convex combinations L_n of K_n ($L_n = \sum_{m\geq n} c_{n,m}K_m$ with $c_{m,n} > 0$ and $\sum_{m\geq n} c_{n,m} = 1$) such that $\lim_{n\to\infty}\|I-L_n\| = 1$.*

The upper estimate. The goal now is to show that

$$\|C_\varphi\|_e \leq \lim_{s\to 1^-} \sup_{|\varphi(z)|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|. \tag{10}$$

This will be accomplished by applying Proposition 2.3. Since each L_n is compact so is $C_\varphi L_n$. Therefore

$$\|C_\varphi\|_e \le \|C_\varphi - C_\varphi L_n\| = \|C_\varphi(I - L_n)\|.$$

On the other hand, we have

$$\begin{aligned}
&\|C_\varphi(I-L_n)\| \\
&= \sup_{\|f\|=1} \|C_\varphi(I-L_n)f\| \\
&= \sup_{\|f\|=1} \sup_{|z|<1} (1-|z|^2)\,|((I-L_n)f)'(\varphi(z))|\,|\varphi'(z)| \qquad (11)\\
&= \sup_{\|f\|=1} \sup_{|z|<1} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|\,(1-|\varphi(z)|^2)\,|((I-L_n)f)'(\varphi(z))|.
\end{aligned}$$

Now fix $0 < s < 1$. Then the right hand side of (11) is less than or equal to

$$\begin{aligned}
(12)\quad &\sup_{\|f\|=1} \sup_{|\varphi(z)|\le s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|\,(1-|\varphi(z)|^2)\,|((I-L_n)f)'(\varphi(z))| \\
&+ \sup_{\|f\|=1} \sup_{|\varphi(z)|> s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|\,(1-|\varphi(z)|^2)\,|((I-L_n)f)'(\varphi(z))|.
\end{aligned}$$

By applying the Pick-Schwarz Lemma in the first term, and the fact that for f in the unit ball

$$\begin{aligned}
&\sup_{|\varphi(z)|>s} (1-|\varphi(z)|^2)\,|((I-L_n)f)'(\varphi(z))| \\
&\le \sup_{|z|<1} (1-|z|^2)\,|((I-L_n)f)'(z)| \le \|I-L_n\|
\end{aligned}$$

to the second term, we find that (12) is less than or equal to

$$\begin{aligned}
(13)\qquad &\sup_{\|f\|=1} \sup_{|w|\le s} (1-|w|^2)\,|((I-L_n)f)'(w)| \\
&\quad + \|I-L_n\| \sup_{|\varphi(z)|>s} \frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|.
\end{aligned}$$

Let us prove that the first term in (13) tends to zero as $n \to \infty$. By the triangle inequality we have that the first term in (13) is less than or equal to

$$(14)\qquad \sum_{m\ge n} c_{n,m} \sup_{\|f\|=1} \sup_{|w|\le s} (1-|w|^2)|((I-K_m)f)'(w)|.$$

By the triangle inequality again we find that $(1-|w|^2)|((I-K_m)f)'(w)|$ is less than or equal to

$$\sup_{\|f\|=1}\sup_{|w|\leq s}(1-|w|^2)\left|f'(w)-f'\left(\left(1-\frac{1}{m}\right)w\right)\right| + \frac{1}{m}\sup_{\|f\|=1}\sup_{|w|\leq s}(1-|w|^2)\left|f'\left(\left(1-\frac{1}{m}\right)w\right)\right|. \tag{15}$$

By integrating f'' along the radial segment $[(1-1/m)w, w]$ it is easy to see that the first term in (15) is less than or equal to

$$\frac{1}{m}\sup_{\|f\|=1}\sup_{|w|\leq s}(1-|w|^2)|w||f''(\xi(w))|, \tag{16}$$

where $\xi(w)$ belongs to the radial segment $[(1-1/m)w, w]$ that is still contained in the closed disk of radius s. The Cauchy inequalities applied to a circle $C(\xi(w))$ centered at $\xi(w)$ and of any fix radius $0 < R < 1-s$ yields that (16) is less than or equal to

$$\frac{1}{mR}\sup_{\|f\|=1}\sup_{|w|\leq s}(1-|w|^2)|w|\max_{|z|=s+R}|f'(z)|. \tag{17}$$

On the other hand, on the unit ball of $\mathcal{B}$ (or $\mathcal{B}_0$) we have $\max_{|z|=s+R}|f'(z)| \leq \frac{1}{1-(s+R)^2}$. So we find that (17) is less than or equal to

$$\frac{1}{mR}\sup_{|w|\leq s}(1-|w|^2)|w|\frac{1}{1-(s+R)^2} \leq \frac{1}{mR}\frac{s}{1-(s+R)^2}.$$

Since the second term in (15) is less than $1/m$ we find that (15) is $\leq C/m$, where C only depends on s. Therefore, we find that (14) is less than or equal to

$$\sum_{m\geq n} c_{n,m}\frac{C}{m} \leq \sum_{m\geq n} c_{n,m}\frac{C}{n} = \frac{C}{n}$$

which tends to zero as $n \to \infty$. Hence, letting $n \to \infty$ in (13), applying Proposition 2.3 and putting everything together, the following inequality follows

$$\|C_\varphi\|_e \leq \sup_{|\varphi(z)|>s}\frac{1-|z|^2}{1-|\varphi(z)|^2}|\varphi'(z)|.$$

Since s was arbitrary inequality (10) holds.

Remarks. 1. By the triangle inequality we have $\|I-K_n\| \leq \|I\|+\|K_n\| = 2$. Therefore, if we use the sequence $\{K_n\}$ instead of the sequence $\{L_n\}$ in the proof of the upper estimate, then we obtain twice the upper estimate. However, that is enough to characterize the compact composition operators on Bloch spaces without requiring Proposition 2.3.

2. Note that it is not possible to have a sequence $\{L_n\}$ of convex combinations of $\{K_n\}_{m\geq n}$ such that $\lim \|I - L_n\| < 1$. For if this were the case, we could obtain an upper estimate strictly less than the lower estimate, a contradiction. Thus in order to prove Proposition 2.3 it is enough to construct a sequence $\{L_n\}$ of convex combinations of $\{K_m\}_{m\geq n}$ such that $\limsup_{n\to\infty} \|I - L_n\| \leq 1$.

The proof of Theorem 2.1 will be completed once we have proved Proposition 2.3. In order to do this we need some basic facts about Bloch spaces. Recall that dual space $\mathcal{B}_0^\star$ of $\mathcal{B}_0$ is isomorphic to the space $A^1(\mathbb{D})$ of analytic functions on the unit disk such that

$$\int_{\mathbb{D}} |g(z)|\, dA(z) < \infty$$

where $dA(z)$ is Lebesgue area measure on $\mathbb{D}$, normalized to have total mass 1, that is, $dA(z) = \frac{1}{\pi} dxdy = \frac{1}{\pi}\rho d\theta d\rho$ for $z = x + iy = \rho e^{i\theta}$. This duality is realized by the integral pairing

$$\langle f, g\rangle = \int_{\mathbb{D}} f(z)\overline{g(z)}\, dA(z)$$

(see [**Zh**, p. 87]). Let $0 < r < 1$ be fixed and let $K_r : \mathcal{B}_0 \to \mathcal{B}_0$ be the operator which assigns to each function f the function $f(rz)$. Now, for any $f(z) = \sum_{n=0}^{\infty} a_n z^n \in \mathcal{B}_0$ and any $g(z) = \sum_{n=0}^{\infty} b_n z^n \in A^1(\mathbb{D})$ a straightforward computation shows that

$$\langle f(rz), g(z)\rangle = \sum_{n=0}^{\infty} \frac{r^n}{n+1} a_n \bar{b}_n = \langle f(z), g(rz)\rangle.$$

Thus, the adjoint operator $K_r^\star : A^1(\mathbb{D}) \to A^1(\mathbb{D})$ acts in the same way as does K_r. We also have that the Bloch space $\mathcal{B}$ is the dual of $A^1(\mathbb{D})$ under the same integral pairing. Thus in a similar way, it can be shown that the bi-adjoint operator $K_r^{\star\star} : \mathcal{B} \to \mathcal{B}$ of K_r is the operator that assigns to each function $f(z)$ the function $f(rz)$. Thus, a little abuse of the language allows us denote $K_r^\star$ and $K_r^{\star\star}$ by K_r. With this we may observe that if we have constructed the sequence $\{L_n\}$ required by Proposition 2.3 for $\mathcal{B}_0$, then just considering the bi-adjoint sequence the result follows for the Bloch space $\mathcal{B}$. This is trivial because $L_n^{\star\star} = \left(\sum_{m\geq n} c_{n,m} K_n\right)^{\star\star} = \sum_{m\geq n} c_{n,m} K_n$ and $\|(I - L_n)^{\star\star}\| = \|I - L_n\|$.

To prove Proposition 2.3 we also need the following proposition about the compact operators K_r.

Proposition 2.4. *For any $g \in A^1(\mathbb{D})$ we have $\|K_r g - g\| \to 0$ as $r \to 1^-$.*

Proof. Let $\varepsilon > 0$ be fixed. By the continuity of the integral we can find an s, $1 > s > 0$, such that

$$\int_{|z|>s} |g(z)|\, dA(z) < \frac{\varepsilon}{3}.$$

Now $rs \to s$ and $1/r \to 1$ as $r \to 1$. Therefore, the change of variables $w = rz$ and the above display show that

$$\begin{aligned}\int_{|z|>s} |g(rz)|\, dA(z) &= \frac{1}{r}\int_{rs<|w|\le r} |g(w)|\, dA(w)\\ &\le \frac{1}{r}\int_{rs<|w|} |g(w)|\, dA(w) < \frac{\varepsilon}{3}\end{aligned}$$

for r near enough to 1. On the other hand, since $K_r g$ tends to g uniformly on compact subsets of the unit disk as $r \to 1^-$, we have $\max_{|z|\le s} |g(rz)-g(z)| < \varepsilon/3$ for r near enough to 1. Thus for r close to 1 we have

$$\begin{aligned}\|g(rz) - g(z)\| &= \int_{|z|\le s} |g(rz) - g(z)|\, dA(z) + \int_{|z|>s} |g(rz) - g(z)|\, dA(z)\\ &< \frac{\varepsilon}{3}s^2 + \int_{|z|>s} |g(z)|\, dA(z) + \int_{|z|>s} |g(rz)|\, dA(z)\\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}\\ &= \varepsilon.\end{aligned}$$

Since ε was arbitrary, the result follows. □

Given two Banach spaces X and Y we denote by $\mathcal{L}(X, Y)$ the Banach space of bounded operators from X into Y and by $\mathcal{K}(X, Y)$ the Banach space of compact operators from X into Y. We need a theorem of Mazur that asserts that if a sequence in a Banach space converges weakly, then some sequence of convex combinations converges in norm (see [**Di**, 15]). We begin with the following theorem, whose proof was provided by Joel H. Shapiro (alternatively, in the proof of Proposition 2.3, we can use Theorem 1 in [**Ka1**]).

Theorem 2.5. *Suppose X and Y are Banach spaces and $\{T_n\}$ is a sequence of compact linear operators from X to Y. Suppose further that for every $y^\star \in Y^\star$ and $x^{\star\star} \in X^{\star\star}$ we have: $\langle T_n^\star y^\star, x^{\star\star}\rangle \to 0$. Then there is a sequence $\{S_n\}$ of convex combinations of the original T_n such that $\|S_n\| \to 0$.*

Proof. Let Q denote the cartesian product of the closed unit ball of $Y^\star$ and the closed unit ball of $X^{\star\star}$, where each ball has its respective weak star topology. Thus Q is a compact Hausdorff space. For $T \in \mathcal{K}(X, Y)$ the function $\widehat{T^\star} : Q \to \mathbb{C}$ defined by:

$$\widehat{T^\star}((y^\star, x^{\star\star})) = \langle T^\star y^\star, x^{\star\star}\rangle \quad (= x^{\star\star}(T^\star(y^\star))) \qquad (x^{\star\star} \in X^{\star\star} \text{ and } y^\star \in Y^\star)$$

belongs to $C(Q)$ (see [**Ka1**, Lemma 1]), and the map $T^\star \to \widehat{T^\star}$ is an isometry taking a certain closed subspace of $\mathcal{K}(Y^\star, X^\star)$ (namely the weak-star continuous compacts) onto a closed subspace of $C(Q)$. By this correspondence and the Hahn-Banach theorem, $T_n^\star \to 0$ weakly in $\mathcal{L}(Y^\star, X^\star)$ if and only if $\widehat{T_n^\star}$ tends weakly in $C(Q)$.

By the Riesz Representation Theorem and the Lebesgue bounded convergence theorem, a sequence of functions in $C(Q)$ converges weakly to zero if and only if it converges pointwise to zero. But the hypothesis on $\{T_n^\star\}$ is just the statement that $\widehat{T_n^\star} \to 0$ pointwise on Q. In addition, it follows from the Uniform Boundedness Principle that $\sup_n \|T_n\| < \infty$, hence because $\|T_n\| = \|T_n^\star\|$, the sequence $\widehat{T_n^\star}$ is also bounded. Thus $T_n^\star \to 0$ weakly in $\mathcal{L}(Y^\star, X^\star)$ and so by Mazur's theorem, there is a sequence of convex combinations $\{S_n^\star\}$ of the original operators $\{T_n^\star\}$, such that $\|S_n^\star\| \to 0$. Thus also $\|S_n\| \to 0$, which is the desired result. □

To prove Proposition 2.3 we will use the fact that $\mathcal{B}_0$ is isomorphic to the sequence space c_0. For completeness we include a proof of this fact. Let us consider the function $\phi(r) = 1 - r^2$ defined on the interval $[0, 1]$ and let $h_\infty(\phi)$ be the Banach space of complex-valued functions, u harmonic in the unit disk with the norm

$$\|u\|_\phi = \sup_{\mathbb{D}} |u(z)|\phi(z)$$

and let $h_0(\phi)$ be the closed subspace of functions u for which $|u(z)|\phi(z) \to 0$ as $|z| \to 1^-$. The space $h_0(\phi)$ is isomorphic to the sequence space c_0 (see [**SW**, Theorem 7]). Finally, we denote by $H_0(\phi)$ the closed subspace of those functions in $h_0(\phi)$ that are analytic on the unit disk. Now, observe that $h_0(\phi)$ is self-conjugate, that is, $u \in h_0(\phi)$ if and only if its conjugate $\bar{u} \in h_0(\phi)$. This fact along with the Closed Graph Theorem implies that the Riesz projection $P : h_0(\phi) \to H_0(\phi)$ defined by

$$Pu = \frac{1}{2}(u + i\bar{u}) + \frac{1}{2}u(0)$$

is bounded. Thus we can express $h_0(\phi) = H_0(\phi) \oplus \ker P$. Now, a famous theorem of Pelczyński (see [**Pe**, Theorem 1]) asserts that if F is a complemented subspace of c_0, then either F is isomorphic to c_0 or F is of finite dimension. Since $H_0(\phi)$ is complemented in a space isomorphic to c_0, it follows that $H_0(\phi)$ is isomorphic to c_0. Finally, since $H_0(\phi)$ is isometrically isomorphic to $\mathcal{B}_0$ (consider the map $f \to f'$), it follows that $\mathcal{B}_0$ is isomorphic to c_0.

As mentioned, the following argument was indicated by N. J. Kalton. Some parts of this argument already appear in [**Ka1**] (see also [**Ka2**, Theorem 2.4] and [**HWW**, Chapter VI, Theorem 4.17 and Theorem 5.7]).

Proof of Proposition 2.3. As pointed out before it is enough to prove the result for the little Bloch space. By Remark 2 above it will be sufficient to show that for any $\varepsilon > 0$ there exists a convex linear combination L_n of $\{K_m\}_{m\geq n}$ with $\|I - L_n\| < 1+\varepsilon$. Once this is done the proof can be completed by a simple diagonal argument.

Since $\mathcal{B}_0$ is isomorphic to the sequence space c_0, James's Theorem (see [**LT**, p. 97]) can be applied to find that there exists a Banach subspace $X_0 \subset c_0$ such that the Banach-Mazur distance from $\mathcal{B}_0$ to X_0 is strictly less than $\sqrt{1+\varepsilon}$. That is, there is an isomorphism $T : \mathcal{B}_0 \to X_0$ such that $\|T\|\|T^{-1}\| < \sqrt{1+\varepsilon}$. We define $T_n = TK_nT^{-1} : X_0 \to X_0$. Upon applying Proposition 2.4 we find that

$$\lim_{n\to\infty} \|T_n^\star x^\star - x^\star\| = 0 \tag{18}$$

for each $x^\star \in X_0^\star$. If P_n is the sequence of coordinate projections on c_0, then we also have

$$\lim_{n\to\infty} \|P_n^\star x^\star - x^\star\| = 0 \tag{19}$$

for each $x^\star \in l^1 = c_0^\star$ the dual space of c_0. Now, if J denotes the inclusion from X_0 into c_0, then $JT_n - P_nJ \in K(X_0, c_0)$. Furthermore, by applying (18) and (19) the sequence $\langle (JT_n - P_nJ)^\star x^\star, y^{\star\star}\rangle$ tends to zero for $y^{\star\star} \in X_0^{\star\star}$ and $x^\star \in l^1$. Thus we may apply Theorem 2.5 to see that there exist a sequence of convex combinations of $\{JT_n - P_nJ\}$ that tends to zero in norm. This implies that there are sequences $\{T_n^c\}$ and $\{P_n^c\}$ of convex combinations of $\{T_m\}_{m\geq n}$ and $\{P_m\}_{m\geq n}$, respectively, such that $JT_n^c - P_n^cJ$ tends to zero in norm. Therefore, we have for all sufficiently large n:

$$\|I - T_n^c\| = \|J(I - T_n^c)\| \leq \|(I - P_n^c)J\| + \|JT_n^c - P_n^cJ\| \leq \sqrt{1+\varepsilon},$$

where we have used successively: The fact that $J : X_0 \to c_0$ is the inclusion map, the triangle inequality, and the inequality $\|(I - P_n^c)J\| \leq 1$. Finally, if we set $L_n = T^{-1}T_n^cT$, then

$$\|I - L_n\| = \|T^{-1}(I - T_n^c)T\| \leq \|T^{-1}\|\|I - T_n^c\|\|T\| < 1+\varepsilon$$

that is what we had to prove. The proof of Proposition 2.3, and therefore that of Theorem 2.1, is now completed. $\square$

Remark. A sequence of compact operators satisfying the required properties to get the upper estimate can also be obtained more directly by using the theory of M-ideals of compact operators (see [**Ka2**, Theorem 2.4] and [**HWW**, Chapter VI, Theorem 4.17]).

References

[ACP] J.M Anderson, J. Clunie and Ch. Pommerenke, *On Bloch functions and normal functions*, J. Reine Angew. Math., **270** (1974), 12-37.

[CM] C. Cowen and B. MacCluer, *Composition Operators on Spaces of Analytic Functions*, CRC Press, 1995.

[Di] J. Diestel, *Sequences and Series in Banach Spaces*, Springer-Verlag, Berlin-New York, 1984.

[Fe] J.L. Fernández, *On the coefficients of Bloch functions*, J. London Math. Soc., **29**(**2**) (1984), 94-102.

[HWW] P. Harmand, D. Werner and W. Werner, *M-ideals in Banach Spaces and Banach Algebras*, Springer-Verlag, Berlin-Heidelberg-New York, 1993.

[Ka1] N.J. Kalton, *Spaces of compact operators*, Math. Ann., **208**(**2**) (1974), 267-278.

[Ka2] ———, *M-ideals of compact operators*, Illinois J. Math., **37** (1993), 147-169.

[LT] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces* I, Springer, Berlin-Heidelberg-New York, 1977.

[MM] K. Madigan and A. Matheson, *Compact composition operators on the Bloch space*, Trans. Amer. Math. Soc., **347** (1995), 2679-2687.

[Pe] A. Pelczynski, *Projections in certain Banach spaces*, Studia Math., **19** (1960), 209-228.

[Po] Ch. Pommerenke, *On Bloch functions*, J. London Math. Soc., **2**(**2**), (1970), 689-685.

[Sh1] J.H. Shapiro, *The essential norm of a composition operator*, Annals of Math., **125** (1987), 375-404.

[Sh2] ———, *Composition Operators and Classical Function Theory*, Springer-Verlag, 1993.

[SW] A.L. Shields and D.L. Williams, *Bounded projections, duality, and multipliers in spaces of harmonic functions*, J. Reine Angew. Math., **299**/**300** (1978), 256-279.

[Zh] K. Zhu, *Operator Theory in Function Spaces*, Marcel Dekker, I.N.C., 1990.

Received May 30, 1997 and revised May 3, 1998.

Universidad de Sevilla
Sevilla 41080
Spain
E-mail address: amontes@cica.es

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

ON THE GEOMETRY OF TWO DIMENSIONAL PRYM VARIETIES

R. Smith and R. Varley

If Σ is a smooth genus two curve, $\Sigma \subset \mathrm{Pic}^1(\Sigma)$ the Abel embedding in the degree one Picard variety, $|2\Sigma|$ the projective space parametrizing divisors on $\mathrm{Pic}^1(\Sigma)$ linearly equivalent to 2Σ, and $\mathrm{Pic}^0(\Sigma)_2 = G \cong (\mathbb{Z}/2\mathbb{Z})^4$ the subgroup of points of order two in the Jacobian variety $J(\Sigma) = \mathrm{Pic}^0(\Sigma)$, then G acts on $|2\Sigma|$ and the quotient variety $|2\Sigma|/G$ parametrizes two fundamental moduli spaces associated with the curve Σ. Namely, Narasimhan-Ramanan's work implies an isomorphism of $|2\Sigma|/G$ with the space $\mathcal{M}$ of (S-equivalence classes of semi-stable, even) $\mathbb{P}^1$ bundles over Σ, and Verra has defined a precise birational correspondence between $|2\Sigma|/G$ and Beauville's compactification of $\mathcal{P}^{-1}(J(\Sigma))$ the fiber of the classical Prym map over $J(\Sigma)$. In this paper we give a new (birational) construction of the composed Narasimhan-Ramanan-Verra map $\alpha : \mathcal{M} \dashrightarrow \mathcal{P}^{-1}(J(\Sigma))$, defined purely in terms of the geometry of a (generic stable) $\mathbb{P}^1$ bundle $X \to \Sigma$ in $\mathcal{M}$, and also an explicit rational inverse map $\beta : \mathcal{P}^{-1}(J(\Sigma)) \dashrightarrow \mathcal{M}$. The map α may be viewed as an analog for Prym varieties of Andreotti's reconstruction of a curve C of genus g from the branch locus of the canonical map on the symmetric product $C^{(g-1)}$. The map β assigns to an étale double cover $\pi : \tilde{C} \to C$ in $\mathcal{P}^{-1}(J(\Sigma))$, where $\tilde{C}$ and C are curves of genera 5 and 3 respectively, the $\mathbb{P}^1$ bundle $\varphi : X \to \Sigma$, where $X = \{$divisors D in $\tilde{C}^{(4)} : \pi_*(D) \equiv \omega_C$, and $h^0(D)$ is even$\}$ and $\varphi : X \to \varphi(X) \cong \Sigma \subset \mathrm{Pic}^4(\tilde{C})$ is the Abel map.

Introduction.

To motivate the constructions in this paper, recall that Torelli's problem for a curve C usually means recovering C from the theta divisor Θ of the Jacobian variety $J(C)$. Andreotti showed it is equivalent to recover C from its symmetric product $C^{(g-1)}$ and, by pulling back the canonical line bundle K_Θ via the Abel map $C^{(g-1)} \to \Theta$, he recovered C from the branch locus of the associated canonical map $C^{(g-1)} \dashrightarrow \mathbb{P}^{g-1}$ ([**A**]). In the case of a Prym variety of a double cover of curves $\pi : \tilde{C} \to C$, the theta divisor Ξ of the Prym variety $P(\tilde{C}/C)$ does not always determine the double cover. In particular a double cover of a curve C of genus three depends on 6 parameters while the theta divisor of the two dimensional Prym variety (P, Ξ) depends on only 3. Hence the Prym theta divisor Ξ cannot determine uniquely the double cover $\pi : \tilde{C} \to C$, (although a prescription can be given by which Ξ determines *all* double covers with Prym variety P). If one takes the analog of Andreotti's approach however, starting from the Abel parametrization of the Prym theta divisor $\tilde{C}^{(2g-2)} \supset X \to \Xi$, (see Section 2 for the precise definition), it turns out there is sometimes more information in X than is contained in Ξ. The key is that the map $\varphi : X \to \Xi$ is not birational, but (generically) a $\mathbb{P}^1$ bundle. This is the approach taken by H. Yin in [**Y**], where he proves the following infinitesimal Torelli result: If C is nonhyperelliptic of genus three, then $\tilde{C}$ is birational to a component of the Hilbert scheme of sections of the $\mathbb{P}^1$ bundle $X \to \Xi$, and X has the same number of moduli as does the double cover $\pi : \tilde{C} \to C$, i.e. $T^1(\tilde{C}/C) \cong T^1(X)$. This means essentially that the map from the space of double covers $\tilde{C}/C$ to the space of all X's is a local isomorphism. In the present article we extend the results of Yin and the parallel with Andreotti's work, and link them with the results of Narasimhan-Ramanan and of Verra, on rank 2 vector bundles and the fibers of the Prym map over two dimensional abelian varieties. In particular we show (when $g(C) = 3$ and C is smooth, nonhyperelliptic), that X determines $\tilde{C}/C$ uniquely, and that the correspondence between X and $\tilde{C}/C$ is induced by Narasimhan-Ramanan's equivalence between semi-stable vector bundles on Ξ and divisors in $|2\Theta|$ on $\mathrm{Pic}^1(\Xi)$. We discuss this in more detail next.

i) In [**N-R**] Narasimhan and Ramanan give an isomorphism from the set $\mathcal{M}$ of S-equivalence classes of semi-stable rank 2 vector bundles with determinant $\mathcal{O}$ on a genus two curve Σ, to the three dimensional linear system $|2\Theta|$ on $\mathrm{Pic}^1(\Sigma)$, by associating to each rank 2 bundle $\mathcal{E}$ the set $D_{\mathcal{E}}$ of invertible sheaves ξ in $\mathrm{Pic}^1(\Sigma)$ such that $H^0(\mathcal{E} \otimes \xi) \neq 0$. They prove $D_{\mathcal{E}}$ is the support of a unique divisor $\tilde{C}_{\mathcal{E}}$ in $|2\Theta|$, (where $\Theta = \Sigma \subset \mathrm{Pic}^1(\Sigma)$ is defined by the natural embedding $p \to \mathcal{O}(p)$), and that the map $\mathcal{M} \to |2\Theta|$ taking $\mathcal{E}$ to $\tilde{C}_{\mathcal{E}}$ is bijective, (where a nonstable bundle S-equivalent to $\mu \oplus \mu^{-1}$ with μ in $\mathrm{Pic}^0(\Sigma)$ corresponds to the point $\Theta_\mu + \Theta_{\mu^{-1}}$ on the Kummer surface in $|2\Theta|$).

ii) With the same notation as above, Verra ([**V**]) describes the fibers of the extended Prym map $\mathcal{P} : \mathcal{R}_3 \to \mathcal{A}_2$, [or $\overline{P} : \overline{\mathcal{R}}_3 \to \mathcal{A}_2$ in Verra's notation], via a rational map $|2\Theta| \dashrightarrow \mathcal{P}^{-1}(J(\Sigma))$. Since every divisor $\tilde{C}$ in $|2\Theta|$ is invariant for the involution ι on $\mathrm{Pic}^1(\Sigma)$ taking ξ to $K_\Sigma - \xi$, he defines this map by taking $\tilde{C}$ in $|2\Theta|$ to $\tilde{C} \to C$ where $C = \tilde{C}/\iota$. He shows this map is defined except along the 16 conic "tropes" on the Kummer surface $\mathcal{K}$ in $|2\Theta|$, and carries the complement $|2\Theta| - \mathcal{K}$ onto the set of those irreducible stable double covers $(\tilde{C}/C)$ in $\mathcal{P}^{-1}(J(\Sigma))$ such that $\tilde{C}$ is not hyperelliptic. He shows also that if $\mathrm{Aut}(\Sigma) \cong \mathbb{Z}_2$, the fibers of the map $|2\Theta| \dashrightarrow \mathcal{P}^{-1}(J(\Sigma))$

are orbits of the action on $|2\Theta|$ by the group $G \cong (\mathbb{Z}_2)^4$ of points of order two in $\mathrm{Pic}^0(\Sigma)$. Finally he shows how to make the map $|2\Theta| \dashrightarrow \mathcal{P}^{-1}(J(\Sigma))$ a surjective morphism by blowing up first at the 16 "points of order two" on the Kummer surface (the singular points on the union of the tropes), then along the strict transform of the tropes, so that the G action on $|2\Theta|$ extends and induces an isomorphism from the G-quotient of the blowup to $\mathcal{P}^{-1}(J(\Sigma))$. Moreover, the exceptional plane $|\mathcal{O}_\Sigma(2\Sigma)|$ over a point of order two, the plane of bicanonical divisors on $\Sigma \cong \Theta$, parametrizes the hyperelliptic double covers $\tilde{C} \to C$ in $\mathcal{P}^{-1}(J(\Sigma))$, where each hyperelliptic curve $\tilde{C}$ arises as a double cover of Σ branched over a bicanonical divisor. (The involution defining the double cover $\tilde{C} \to C$ is the product of the hyperelliptic involution on $\tilde{C}$ and the involution induced by the double cover $\tilde{C} \to \Sigma$.)

Since two vector bundles $\mathcal{E}, \tilde{\mathcal{E}}$ determine isomorphic $\mathbb{P}^1$ bundles if and only if $\mathcal{E} \cong \tilde{\mathcal{E}} \otimes \mathcal{L}$ for some line bundle $\mathcal{L}$, and since $\det(\mathcal{E} \otimes \mathcal{L}) = \det(\mathcal{E}) \otimes \mathcal{L}^2$, two stable rank 2 vector bundles $\mathcal{E}$, $\tilde{\mathcal{E}}$ of determinant $\mathcal{O}$ on Σ determine isomorphic $\mathbb{P}^1$ bundles if and only if $\mathcal{E} \cong \tilde{\mathcal{E}} \otimes \mathcal{L}$ where $\mathcal{L}^2 \cong \mathcal{O}$. Moreover, this action of the group G on $\mathcal{M}$ corresponds to the action on $|2\Theta|$, so that stable $\mathbb{P}^1$ bundles on Σ correspond via Narasimhan-Ramanan and Verra bijectively to the set of irreducible stable double covers $(\tilde{C}/C)$ in $\mathcal{P}^{-1}(J(\Sigma))$ such that $\tilde{C}$ is not hyperelliptic.

iii) In the present article we show that the $\mathbb{P}^1$ bundle $\varphi : X \to \Sigma$ associated by the Abel parametrization to a double cover $\tilde{C} \to C$ of a smooth nonhyperelliptic genus three curve with Prym variety $J(\Sigma)$ determines $\tilde{C} \to C$ uniquely, that X arises from a stable rank 2 vector bundle on Σ with determinant $\mathcal{O}$, and that the association $(\tilde{C}/C) \mapsto X$ inverts the maps of Narasimhan-Ramanan and Verra. Since we give an explicit inverse we also recover, at least for smooth nonhyperelliptic $\tilde{C}$, that their correspondence is bijective. We also reconstruct from X a two parameter family of hyperelliptic double covers with the same Prym variety as $\tilde{C}/C$.

More precisely, by analogy with Andreotti's proof, we reconstruct the double cover $\pi : \tilde{C} \to C$, from the branch locus of the finite map $h : X \to \mathbb{P}^2$ associated to the line bundle $\mathcal{T}_\varphi \otimes \varphi^*(K_\Sigma)$, (the pulled back canonical bundle of the Prym theta divisor $\Xi \cong \Sigma$, twisted by the bundle $\mathcal{T}_\varphi$ of tangents "along the fibers" of φ). If $\tilde{C} \to C$ is an étale double cover of a smooth nonhyperelliptic genus three curve, with Prym variety $(P, \Xi) \cong (J(\Sigma), \Sigma) =$ the jacobian of a genus two curve Σ, we prove Yin's rank 2 vector bundle $\mathcal{E}$ on Σ which defines the Abel $\mathbb{P}^1$ bundle $X \to \Sigma$, can be taken to have determinant $\mathcal{O}$ and is stable, and that the divisor $\tilde{C}_{\mathcal{E}}$ in $|2\Theta|$ associated to $\mathcal{E}$ by Narasimhan-Ramanan is isomorphic to $\tilde{C}$. Conversely if $\mathcal{E}$ is any stable rank 2 vector bundle over Σ with determinant $\mathcal{O}$, and if the corresponding divisor $\tilde{C}_{\mathcal{E}} = \tilde{C}$ in $|2\Theta|$ is smooth, with natural involution $\xi \mapsto (K - \xi)$, we prove the associated Prym variety is isomorphic to $J(\Sigma)$, and the Abel $\mathbb{P}^1$ bundle $\varphi : X \to \Xi$ constructed from a Poincaré line bundle on $\mathrm{Pic}^4(\tilde{C}) \times \tilde{C}$, is isomorphic to $\mathbb{P}(\mathcal{E})$. This proves that the Prym construction inverts the Narasimhan-Ramanan correspondence at least for smooth divisors in $|2\Theta|$, and hence recovers for these divisors that their correspondence is bijective. Finally we show that the fibers of the map $h : X \to \mathbb{P}^2$ over lines yield hyperelliptic double covers $\pi : \tilde{C}_\lambda \to C_\lambda$ with the same Prym variety as $\pi : \tilde{C} \to C$.

1. Background: The Prym variety of a double cover $\pi : \tilde{C} \to C$.

Let $\tilde{C}/C$ denote an étale double cover $\pi : \tilde{C} \to C$ of smooth connected nonhyperelliptic curves, with $g(C) = 3$, and $g(\tilde{C}) = 5$, and let $P = P(\tilde{C}/C)$ denote the two dimensional Prym variety of $\tilde{C}/C$. Up to choice of origin, this p.p.a.v. (principally polarized abelian variety) can be constructed as follows (Mumford [**Mu1**, p. 342]): The double cover $\pi : \tilde{C} \to C$ determines a norm map $\mathrm{Nm} : \mathrm{Pic}^d(\tilde{C}) \to \mathrm{Pic}^d(C)$ for every d. For $d = 2g(C) - 2 = 4$, we consider $\mathrm{Nm} : \mathrm{Pic}^4(\tilde{C}) \to \mathrm{Pic}^4(C)$ and the inverse image $\mathrm{Nm}^{-1}(K_C)$ of the canonical line bundle K_C of C. This inverse image has exactly two connected components, $\mathrm{Nm}^{-1}(K_C) = P^+ \cup P^-$, according to the parity of the number of sections of the corresponding line bundles. In particular, we take $P = P^+ = \{$line bundles L on $\tilde{C}$ with an even number of sections, and such that $\mathrm{Nm}(L) = K_C\}$, to be the Prym variety $P(\tilde{C}/C)$ (except for a choice of origin). The theta divisor on P is given by the underlying reduced variety Ξ of the intersection divisor $P^+ \cap \tilde{\Theta} = 2\Xi$, where $\tilde{\Theta} = \{$effective line bundles of degree 4 on $\tilde{C}\}$, $\Xi = \{$line bundles L on $\tilde{C}$, with $h^0(L) > 0$ and even, such that $\mathrm{Nm}(L) = K_C\}$. Then (P, Ξ) is a two dimensional p.p.a.v., Ξ is isomorphic to a smooth genus two curve Σ, and (P, Ξ) is isomorphic to the Jacobian variety $(J(\Sigma), \Sigma)$ of Σ, except for lack of a choice of origin. The restriction of the canonical involution $\mu \mapsto (K_{\tilde{C}} - \mu)$ from $\mathrm{Pic}^4(\tilde{C})$ to $P \cong \mathrm{Pic}^1(\Sigma)$ preserves Σ and corresponds to the involution $\mu \mapsto (K_\Sigma - \mu)$ of $\mathrm{Pic}^1(\Sigma)$, and to the involution $-\mathrm{id}$ of $(J(\Sigma), \Sigma)$. (The natural involution of Pic^{2g-2} taking D to $\tilde{K} - D$ always restricts on P^+ to a translate of the natural involution $-\mathrm{id}$, hence on Σ to the canonical involution.)

2. The Abel parameterization $\varphi : X \to \Xi$ of a Prym theta divisor, and the restricted norm map $h : X \to |K_C| \cong \mathbb{P}^{2*}$.

Now recall the construction of the variety X parametrizing the Prym theta divisor Ξ. Let C be a smooth connected nonhyperelliptic curve of genus three and $\tilde{C} \to C$ an étale connected double cover, where $\tilde{C}$ has genus five, and $\Xi \subset P \subset \mathrm{Pic}^4(\tilde{C})$ the embedded Prym variety constructed above. The Abel-Jacobi map $\varphi : \tilde{C}^{(4)} \to \tilde{\Theta}$ is a resolution of singularities of $\tilde{\Theta}$, a local isomorphism over smooth points of $\tilde{\Theta}$, and with a $\mathbb{P}^1$ fiber over each point of the curve $\operatorname{sing}(\tilde{\Theta})$. Since $h^0(L) \geq 2$ for L in Ξ, by Riemann's singularity theorem $\Xi \subset \operatorname{sing}(\tilde{\Theta})$, and if we define $X = \varphi^{-1}(\Xi) \subset \tilde{C}^{(4)}$, then $\varphi : X \to \Xi$ is a $\mathbb{P}^1$ bundle over the genus two curve $\Xi = \Sigma$.

The other map of interest to us is the restriction h to X, of the finite norm map $\mathrm{Nm} : \tilde{C}^{(4)} \to C^{(4)}$. The map $h : X \to |K_C| \cong \mathbb{P}^{2*}$, takes a point $D = p_1 + p_2 + p_3 + p_4$ of X, to the canonical divisor $\overline{D} = \overline{p}_1 + \overline{p}_2 + \overline{p}_3 + \overline{p}_4$ on C, where $\pi(p_i) = \overline{p}_i$. The norm map $\mathrm{Nm} : \tilde{C}^{(4)} \to C^{(4)}$ has degree $2^4 = 16$, but is branched over any divisor $\overline{D}$ having multiplicities, i.e. in which some $\overline{p}_i$ equals some $\overline{p}_j$. We claim that h has degree 8 and is branched precisely over the dual curve $C^* \subset |K_C|$. By Mumford's theory in [**Mu1**, **Mu2**, **W**], exactly half the divisors of $\tilde{C}^{(4)}$ over a given point of $C^{(4)}$ have even parity, so h has degree 8. Since all effective canonical divisors of C are cut out on the canonical plane embedding of C in $|K_C|^*$ by lines, h is branched at most over C^* = lines cutting divisors on C with multiplicities. Consider a general such canonical divisor with multiplicities $\overline{D} = 2\overline{p} + \overline{r} + \overline{s}$ on C. The ramified divisors over it in $\tilde{C}^{(4)}$ have form $p + p' + r + s$, where p, p' are the preimages of p, and r, s are preimages of $\overline{r}$, $\overline{s}$, under the double cover $\pi : \tilde{C} \to C$. If r', s' are the other preimages of $\overline{r}$,$\overline{s}$, then by Mumford's theory, $p+p'+r+s$, and $p+p'+r+s'$, have opposite parity, as do $p+p'+r+s$, and $p+p'+r'+s$. Thus there are exactly two ramification divisors over $\overline{D}$ with even parity, which we may assume are $p + p' + r + s$, and $p + p' + r' + s'$. Hence over a smooth point on C^* in $|K_C|$, corresponding to the canonical divisor $\overline{D} = 2\overline{p} + \overline{r} + \overline{s}$ on C, there are exactly the two simple ramification points $D = p + p' + r + s$, and $D' = p + p' + r' + s'$, of $h : X \to |K_C|$.

Since the map $h : X \to |K_C|$ recovers the dual curve C^* as its branch locus, it recovers also C as the dual of the branch locus, when C is nonhyperelliptic. To recover $\tilde{C}/C$ from X we will show how to recover the map h from X, and then also the double cover $\pi : \tilde{C} \to C$ from h. Since Σ has genus two, any map of $\mathbb{P}^1$ to Σ is constant, so X contains a unique family of copies of $\mathbb{P}^1$ hence has only one structure of $\mathbb{P}^1$ bundle. Thus we can already recover Ξ and the $\mathbb{P}^1$ bundle map $\varphi : X \to \Xi$, from X. In fact, as Yin shows in [**Y**], $\varphi : X \to \Xi \subset P$ is equivalent to an Albanese map for X.

In passing, we calculate the genus of the inverse image curve over a general line H in $|K_C|$, by the map h. Since h has degree 8, the branch locus C^* has degree 12, and there are two ramification points of h over each general branch point, by Riemann-Hurwitz the genus of the preimage $h^{-1}(H)$ is $(1/2)[2 + 8(-2) + 24] = 5$.

3. The line bundle $\mathcal{O}_X(1) = h^*(\mathcal{O}(1))$.

Next we recover the line bundle giving rise to the restricted norm map h, in terms of the Abel map φ, as follows: Let $\mathcal{O}_X(1) = h^*(\mathcal{O}(1))$, be the pullback to X, of the line bundle $\mathcal{O}_{|K_C|}(1)$ on the canonical space $|K_C|$. We claim:

Proposition 1. *$\mathcal{O}_X(1) = \mathcal{T}_\varphi \otimes \varphi^*(K_\Xi)$, where $\mathcal{T}_\varphi$ is the relative tangent sheaf of φ (= the sheaf of tangents along the fibers of φ), and $\varphi^*(K_\Xi)$ is the pullback of the canonical sheaf on the genus two curve Ξ.*

Proof. Consider the line bundle $\mathcal{O}_X(1)$ restricted to one fiber of the $\mathbb{P}^1$ bundle $\varphi : X \to \Xi$. We will show this restriction has degree two by intersecting a divisor of $\mathcal{O}_X(1)$ with a fiber of φ. Consider first a divisor of $\mathcal{O}_{|K_C|}(1)$ of form $H_{\overline{p}}$, consisting of all canonical divisors on C containing a given point $\overline{p}$. This is represented by a line $H_{\overline{p}}$, tangent to C^*. The pullback of this divisor to X has two components, $h^*(H_{\overline{p}}) = D_p + D_{p'}$, represented by divisors on $\tilde{C}$ which contain respectively either p or p', the two preimages of $\overline{p}$ in $\tilde{C}$ under $\pi : \tilde{C} \to C$. Since C^* is irreducible of degree 12 when C is smooth, $H_{\overline{p}}$ is not contained in the branch locus C^* and h is unramified at a general point of each component of $h^*(H_{\overline{p}}) = D_p + D_{p'}$, so this fiber is reduced. Now each fiber of $\varphi : X \to \Xi$ is a pencil of divisors of degree 4 moving in a linear series on $\tilde{C}$, without fixed points, ($\tilde{C}$ is not trigonal since the singular locus of $\tilde{\Theta}$ contains the curve Ξ of genus two, whereas for $\tilde{C}$ trigonal the singular locus of $\tilde{\Theta}$ is two copies of $\tilde{C}$ by [**A-M**, p. 212]). Hence exactly one divisor in each pencil contains p, and exactly one contains p'. Thus in general $h^*(H_{\overline{p}})$ meets each fiber of φ twice, and the degree of the restriction of $\mathcal{O}_X(1)$ to each fiber is 2.

Now the restriction of $\mathcal{T}_\varphi$ to a fiber of φ, is the tangent sheaf of a projective line, hence also of degree 2. Thus the tensor product $\mathcal{O}_X(1) \otimes \mathcal{T}_\varphi^*$ has degree zero on the fibers, and thus is the pullback $\varphi^*(\mathcal{L})$ of a sheaf $\mathcal{L}$ from Ξ by the map $\varphi : X \to \Xi$. Since for each point p on $\tilde{C}$, D_p has been seen to be generically a section of φ, by Zariski's Main Theorem φ restricts on D_p to an isomorphism of D_p with Ξ, so we can compute $\mathcal{L}$ by restricting $\mathcal{O}_X(1) \otimes \mathcal{T}_\varphi^*$ to the section D_p. The tangent sheaf $\mathcal{T}_\varphi$ to the fibers of φ is the normal sheaf to a section D_p, so $\mathcal{T}_\varphi^*$ restricts on D_p to $\mathcal{O}(-D_p)|_{D_p}$, while $\mathcal{O}_X(1) = \mathcal{O}(D_p + D_{p'})$. Thus $\mathcal{O}_X(1) \otimes \mathcal{T}_\varphi^*$ restricts on D_p to $\mathcal{O}(D_{p'})|_{D_p}$. To compute this, we must intersect $D_{p'}$ with D_p. This intersection is made up of points of X representing divisors (on $\tilde{C}$) which contain both p and p'. Such a divisor must lie over a divisor on C of form $2\,\overline{p} + \overline{r} + \overline{s}$. Hence the divisor on $\tilde{C}$ is a ramification divisor of h of form $p + p' + r + s$, or $p+p'+r'+s'$, as we have computed before in Section 2. Thus the intersection divisor $D_p \cdot D_{p'}$ contains at most two points, and contains exactly two unless $r + s = r' + s'$, i.e. unless $s = r'$, $s' = r$. That would mean our ramification divisor $p + p' + r + s = p + p' + r + r'$ lies over the divisor $2\overline{p} + 2\overline{r}$, one of the 28 divisors cut by a bitangent line to the canonical plane quartic C. Thus for a general point p on $\tilde{C}$, $D_p \cdot D_{p'}$ contains exactly two points. To see that the two intersection points are simple, note that the fiber $h^{-1}(H)$ over a general line H has arithmetic genus 5 by the computation at the end of Section 2, so the reduced curve $h^{-1}(H_{\overline{p}}) = D_p \cup D_{p'}$ also has arithmetic genus 5, and is composed of two components, each isomorphic to the curve Ξ of genus two. Since the components meet twice, both intersections are transverse. Thus the divisor $D_p \cdot D_{p'}$ has degree 2 on D_p, and (as can be seen from the explicit description of its points above) is invariant for the involution ι on X induced by the involution ι of $\tilde{C}$ associated to $\pi : \tilde{C} \to C$.

Since the full inverse image of a canonical divisor $\overline{p} + \overline{q} + \overline{r} + \overline{s}$ on C under the unramified double cover π, is a canonical divisor $p + p' + q + q' + r + r' + s + s'$ on $\tilde{C}$, the involution ι on X takes $(p + q + r + s)$ to $p' + q' + r' + s' = K_{\tilde{C}} - (p+q+r+s)$. Thus ι corresponds to the canonical involution $\mu \mapsto (K_{\tilde{C}} - \mu)$ on $\mathrm{Pic}^4(\tilde{C})$, which on Ξ must be the hyperelliptic one, i.e. on $\Xi = \{$effective divisors of degree one on $\Xi\}$, the involution taking μ to $(K_\Xi - \mu)$.

Since the divisor cut by $D_{p'}$ on D_p is invariant under the involution of X, it corresponds under the isomorphism $\varphi : D_p \to \Xi$, to an invariant divisor for the canonical hyperelliptic involution on Ξ, hence to a canonical divisor on Ξ. Thus $\mathcal{O}_X(1) \otimes \mathcal{T}_\varphi^*$ restricts on D_p to $\varphi^*(K_\Xi)$, hence $\mathcal{O}_X(1) \otimes \mathcal{T}_\varphi^* \cong \varphi^*(K_\Xi)$ also on X, and thus $\mathcal{O}_X(1) = \mathcal{T}_\varphi \otimes \varphi^*(K_\Xi)$. $\square$

Corollary 1. *Since X determines the unique $\mathbb{P}^1$ bundle structure $\varphi : X \to \Xi$, it determines also $\mathcal{T}_\varphi$, $\varphi^*(K_\Xi)$, and hence $\mathcal{O}_X(1) = \mathcal{T}_\varphi \otimes \varphi^*(K_\Xi)$, the line bundle associated to the map $h : X \to |K_C|$.*

4. Completeness of the linear system defining h, and the recovery of $\pi : \tilde{C} \to C$ from X.

We claim the map $h : X \to |K_C|$ is given by a complete linear system, and hence the map h itself is determined by X. To see this, we push $\mathcal{O}_X(1)$ down to Ξ by φ, to get $\varphi_*(\mathcal{O}_X(1)) \cong \varphi_*(\mathcal{T}_\varphi \otimes \varphi^*(K_\Xi)) \cong \varphi_*(\mathcal{T}_\varphi) \otimes K_\Xi$. Then using Serre duality, $h^0(\varphi_*(\mathcal{O}_X(1))) = h^0(\varphi_*(\mathcal{T}_\varphi) \otimes K_\Xi) = h^1(\varphi_*(\mathcal{T}_\varphi)^*)$. In [**Y**], H. Yin shows that $\varphi_*(\mathcal{T}_\varphi)$ is canonically isomorphic to the self dual bundle $\mathrm{End}_0(\triangle)$ of trace zero endomorphisms of $\triangle$, where $\triangle$ is the direct image of a Poincaré line bundle as described below, and computes $h^1(\varphi_*(\mathcal{T}_\varphi)) = 3$. Hence $h^0(\mathcal{O}_X(1)) = h^0(\varphi_*(\mathcal{O}_X(1))) = h^0(\varphi_*(\mathcal{T}_\varphi) \otimes K_\Xi) = h^1(\varphi_*(\mathcal{T}_\varphi)^*) = h^1(\varphi_*(\mathcal{T}_\varphi)) = 3$. Thus X determines both the line bundle and the linear system defining the map $h : X \to |K_C|$, hence also h, its branch locus C^* and thus also the curve C. (Later we will show the bundle $\triangle$ is stable and can be taken to have determinant $\mathcal{O}$, so then $\triangle$ is "simple", i.e. $h^0(\mathrm{End}_0(\triangle)) = 0$, and Riemann Roch gives $\chi(\mathrm{End}_0(\triangle)) = \mathrm{degree}(\mathrm{End}_0(\triangle)) + \mathrm{rank}(\mathrm{End}_0(\triangle)) \cdot (1 - g) = 0 + 3(1 - 2) = -3$. Hence $h^1(\mathrm{End}_0(\triangle)) = 3 + h^0(\mathrm{End}_0(\triangle)) =$

3, giving another computation of $h^1(\varphi_*(\mathcal{T}_\varphi))$, modulo the isomorphism of $\varphi_*(\mathcal{T}_\varphi)$ with $\mathrm{End}_0(\triangle)$.)

To get the double cover $\tilde{C} \to C$, look again at the components of the inverse images of the lines $H_{\overline{p}}$ tangent to the branch locus C^*. These components are parametrized by the points of $\tilde{C}$. I.e. the map taking p on $\tilde{C}$ to the component D_p of divisors in $h^{-1}(H_{\overline{p}})$ containing p, is birational from $\tilde{C}$ to the curve of components of inverse images of tangent lines $H_{\overline{p}}$ to C^*, and this birational map takes the involution on $\tilde{C}$ to that on X. Thus this curve of components determines both the curve $\tilde{C}$, and the involution defining the double cover $\pi : \tilde{C} \to C$. Thus X determines $\pi : \tilde{C} \to C$.

Note. Since we showed in §3 that D_p is a section of $\varphi : X \to \Sigma$, the map taking p to D_p represents $\tilde{C}$ as a curve in the variety of sections of φ with self intersection equal to $D_p \cdot D_p = \deg(\mathcal{T}_\varphi|_{D_p}) = \deg(\mathcal{O}_X(1) \otimes \varphi^*(K^*_\Sigma))|_{D_p} = \deg(h|_{D_p}) + \deg(K^*_\Sigma) = 4 - 2 = 2$. Yin shows [**Y**, p. 23] that this representation of $\tilde{C}$ is birational onto a component Z of the Hilbert scheme S_2 of all sections of X of self intersection 2. From the next section it will follow that Z is the unique component of S_2.

5. Relationship between the correspondence $\tilde{C}/C \mapsto X$ and that of Narasimhan-Ramanan and Verra.

Now consider the moduli space $\mathcal{M}$ of S-equivalence classes of semi-stable rank 2 vector bundles of determinant $\mathcal{O}$, over the smooth curve Σ of genus 2, and the isomorphism $\mathcal{M} \cong |2\Theta|$, where $\Theta = \Sigma \subset \mathrm{Pic}^1(\Sigma)$. Recall that to the bundle $\mathcal{E}$ Narasimhan-Ramanan associate the subset $D_\mathcal{E} = \{\xi \text{ in } \mathrm{Pic}^1(\Sigma) : H^0(\mathcal{E} \otimes \xi) \neq 0\} \subset \mathrm{Pic}^1(\Sigma)$, the support of a unique divisor $\tilde{C}_\mathcal{E}$ linearly equivalent to 2Θ, and that a nonstable bundle S-equivalent to $\mu \oplus \mu^{-1}$ with μ in $\mathrm{Pic}^0(\Sigma)$ corresponds to the point $\Theta_\mu + \Theta_{\mu^{-1}}$ on the Kummer surface $\mathcal{K}$ in $|2\Theta|$. Then Verra's map $\tilde{C}_\mathcal{E} \mapsto (\tilde{C}_\mathcal{E}, \iota)$ carries the complement of $\mathcal{K}$ in $|2\Theta|$ onto the set of irreducible stable double covers $(\tilde{C}, \iota)$ in $\mathcal{P}^{-1}(J(\Sigma))$ such that $\tilde{C}$ is not hyperelliptic. Composing gives a rational map $\mathcal{M} \dashrightarrow \mathcal{P}^{-1}(J(\Sigma))$, taking stable bundles to irreducible stable double covers, which factors through the quotient of $\mathcal{M}$ by the group G of points of order 2 in $\mathrm{Pic}^0(\Sigma)$, i.e. by a rational map taking stable $\mathbb{P}^1$ bundles to irreducible stable double covers. We will show that the Abel parametrization studied by Yin gives a birational inverse to this map. I.e. let $\tilde{C}/C$ be a double cover of a smooth, nonhyperelliptic, genus three curve, and $\varphi : X \to \Xi$ the associated Abel parametrization of the theta divisor on the Prym variety $(P, \Xi) \cong (J(\Sigma), \Sigma) =$ the Jacobian of a smooth genus two curve Σ. If $\mathcal{E}$ is a rank 2 vector bundle of determinant $\mathcal{O}$ on Σ with $\mathbb{P}(\mathcal{E}) \cong X$, then $\mathcal{E}$ is stable and the associated divisor $\tilde{C}_\mathcal{E}$ in $|2\Xi|$ is isomorphic to $\tilde{C}$ (preserving the involution). Conversely, if $\mathcal{E}$ is a stable bundle on Σ with associated divisor

$\tilde{C}$ in $\mathrm{Pic}^1(\Sigma)$, where $\tilde{C}$ is smooth and nonhyperelliptic, then $\mathbb{P}(\mathcal{E}) \cong X$, the Abel parametrization associated to the Prym theta divisor = of $(\tilde{C}, \iota)$.

First we recall some basic facts about bundles:

1) Definition: For a rank 2 vector bundle $\mathcal{E}$, $\deg(\mathcal{E}) = \deg(\det(\mathcal{E}))$. $\mathcal{E}$ is "stable" (resp. semi-stable) if and only if for every sub-line bundle $\mathcal{L} \subset \mathcal{E}$, $\deg(\mathcal{L}) < (1/2)\deg(\mathcal{E})$, (resp. $\deg(\mathcal{L}) \leq (1/2)\deg(\mathcal{E})$). $\mathbb{P}(\mathcal{E})$ is stable if and only if $\mathcal{E}$ is stable.

2) If $\mathcal{E}$, $\tilde{\mathcal{E}}$ are rank 2 vector bundles on Σ, then $\mathbb{P}(\mathcal{E}) \cong \mathbb{P}(\tilde{\mathcal{E}})$ if and only if $\mathcal{E} \cong \tilde{\mathcal{E}} \otimes \mathcal{L}$ for some line bundle $\mathcal{L}$.

3) Every $\mathbb{P}^1$ bundle on Σ has form $\mathbb{P}(\mathcal{E})$ for some rank 2 vector bundle $\mathcal{E}$.

4) Since $\det(\mathcal{E} \otimes \mathcal{L}) = \det(\mathcal{E}) \otimes \mathcal{L}^2$, a $\mathbb{P}^1$ bundle $\varphi : X \to \Sigma$ has form $\mathbb{P}(\mathcal{E}) \cong X$ for some $\mathcal{E}$ with $\det(\mathcal{E}) \cong \mathcal{O}$ if and only if $X \cong \mathbb{P}(\tilde{\mathcal{E}})$ for some $\tilde{\mathcal{E}}$ with $\deg(\tilde{\mathcal{E}})$ even.

In Yin's treatment of the $\mathbb{P}^1$ bundle $\varphi : X \to \Xi \subset \mathrm{Pic}^4(\tilde{C})$, he considers a Poincaré bundle $\mathcal{L}$ on $\tilde{C} \times \mathrm{Pic}^4(\tilde{C})$, i.e. a line bundle $\mathcal{L}$ such that for each ν in $\mathrm{Pic}^4(\tilde{C})$, we have $\mathcal{L}|_{\tilde{C}\times\{\nu\}} \cong \nu$. If we denote the restriction to $\tilde{C} \times \Sigma$ also by $\mathcal{L}$, and the second projection by μ, he puts $\mathcal{E} = \mu_*(\mathcal{L})$ so that $\mathcal{E}$ is a rank two vector bundle on Σ such that $X \cong \mathbb{P}(\mathcal{E})$.

Lemma 1. *Given a stable rank* 2 *vector bundle* $\mathcal{E}$ *with* $\det(\mathcal{E}) \cong \mathcal{O}$*, on a smooth genus* 2 *curve* Σ*, and associated* $\mathbb{P}^1$ *bundle* $X = \mathbb{P}(\mathcal{E})$*, the natural map from* $\{$*sections* σ *of* $X \to \Sigma$*, such that* $(\sigma^2) = 2\}$ *to* $\{\xi$ *in* $\mathrm{Pic}^1(\Sigma) : H^0(\mathcal{E} \otimes \xi) \neq 0\}$ *is a bijection.*

Proof. A section $\sigma : \Sigma \to X$ is equivalent to a sub-line bundle $\mathcal{L} \subset \mathcal{E}$, where for each point p the fiber $\mathcal{L}_p$ defines $\sigma(p)$. The map from the set of sections to $\mathrm{Pic}(\Sigma)$ takes σ to the abstract (dual) line bundle $\mathcal{L}^*$, forgetting the inclusion $\mathcal{L} \subset \mathcal{E}$. We show next that if $(\sigma^2) = 2$, then $\deg(\mathcal{L}^*) = 1$. Since $(\sigma^2) = \deg(N_\sigma/X)$, we want to calculate the degree of the normal bundle to σ in X. Now $N_\sigma/X \cong \mathcal{H}\mathrm{om}(\mathcal{L}, \mathcal{E}/\mathcal{L}) \cong \mathcal{H}\mathrm{om}(\mathcal{L}, \mathcal{L}^*) \cong (\mathcal{L}^*)^2$, so $(\sigma^2) = 2\deg(\mathcal{L}^*)$, and $(\sigma^2) = 2$ implies $\deg(\mathcal{L}^*) = 1$. (We have used $\mathcal{O} \cong \det(\mathcal{E}) \cong \mathcal{L} \otimes (\mathcal{E}/\mathcal{L})$ to deduce $\mathcal{E}/\mathcal{L} \cong \mathcal{L}^*$.) Next, if $0 \to \mathcal{L} \to \mathcal{E}$ is exact, then tensoring with $\mathcal{L}^*$ gives $0 \to \mathcal{O} \to \mathcal{E} \otimes \mathcal{L}^*$ exact, so the image of 1 defines a nonzero section of the sheaf $\mathcal{E} \otimes \mathcal{L}^*$. Thus the forgetful map from $\{$sections σ of $X \to \Sigma$, such that $(\sigma^2) = 2\}$ to $\{\xi$ in $\mathrm{Pic}^1(\Sigma) : H^0(\mathcal{E} \otimes \xi) \neq 0\}$ is well defined.

To prove surjectivity, if ξ is in $\mathrm{Pic}^1(\Sigma)$, and $H^0(\mathcal{E} \otimes \xi) \neq 0$, there is a nonzero homomorphism $\mathcal{O} \to \mathcal{E} \otimes \xi$ sending 1 to a nonzero global section s of $\mathcal{E} \otimes \xi$, so multiplication by this section gives a nonzero homomorphism $\xi^* \to \mathcal{E}$. This is an injective sheaf map since $\mathcal{O}$ is a sheaf of integral domains, and defines an injection at least on those fibers of the line bundle ξ^* at points where s does not vanish. This defines a rational section $\Sigma \to X$ of the $\mathbb{P}^1$ bundle $X = \mathbb{P}(\mathcal{E})$ which extends uniquely to a regular section σ, hence defines a unique sub-line bundle $\mathcal{L} \subset \mathcal{E}$ such that $\xi^* \subset \mathcal{L} \subset \mathcal{E}$ and $\mathcal{E}/\mathcal{L}$ is a line bundle. Since $\mathcal{E}$ is stable, $-1 = \deg(\xi^*) \leq \deg(\mathcal{L}) < \deg(\mathcal{E})/2 = 0$. Thus $\deg(\mathcal{L}) = -1$, hence $\mathcal{L} = \xi^*$, and $\mathcal{L}^* \cong \xi$, proving surjectivity.

To prove injectivity, we must show for each line bundle ξ in $\mathrm{Pic}^1(\Sigma)$ there is at most one sub-line bundle of $\mathcal{E}$ which is the image of an inclusion $\xi^* \to \mathcal{E}$. It suffices to show that the space of homomorphisms $\mathrm{Hom}(\xi^*, \mathcal{E})$ is at most one dimensional. Adapting an argument of [**N-R**], if $\dim(\mathrm{Hom}(\xi^*, \mathcal{E})) \geq 2$, and if the evaluation map $\xi^* \otimes \mathrm{Hom}(\xi^*, \mathcal{E}) \to \mathcal{E}$, were injective on fibers, then we would have $\mathcal{E} \cong \xi^* \oplus \xi^*$, which is impossible since $\det(\mathcal{E}) \cong \mathcal{O}$. Thus for some p in Σ, there would be a nonzero sheaf homomorphism $\xi^* \to \mathcal{E}$ which is zero on the fiber ξ^*_p. But we have just argued in proving surjectivity that when $\mathcal{E}$ is stable, and $\deg(\xi) = 1$, any nonzero map $\xi^* \to \mathcal{E}$ has a line bundle quotient, hence is injective on all fibers. $\square$

Corollary 2. *If C is nonhyperelliptic, the $\mathbb{P}^1$ bundle $X \to \Xi$ arising from the Abel parametrization of the Prym theta divisor $\Xi(\tilde{C}/C)$ is stable.*

Proof. We showed in the note at the end of §4 that X has sections σ with $(\sigma^2) = 2$, and we want to deduce that if $X = \mathbb{P}(\mathcal{E})$, then $\deg(\mathcal{E})$ is even. If $\mathcal{L} \subset \mathcal{E}$ is a sub-line bundle corresponding to a section σ of $\mathcal{E}$, yielding an exact sequence $0 \to \mathcal{L} \to \mathcal{E} \to \mathcal{N} \to 0$, then $\det(\mathcal{E}) \cong \mathcal{L} \otimes \mathcal{N}$, and as above $(\sigma^2) = \deg(\mathrm{Hom}(\mathcal{L}, \mathcal{N})) = \deg(\mathcal{L}^* \otimes \mathcal{N}) = \deg(\mathcal{N}) - \deg(\mathcal{L})$. Hence $\deg(\mathcal{E}) = \deg(\mathcal{L} \otimes \mathcal{N}) = \deg(\mathcal{L}) + \deg(\mathcal{N}) = (\sigma^2) + 2\deg(\mathcal{L}) \equiv (\sigma^2) \bmod 2$, so $(\sigma^2) = 2$ implies $\deg(\mathcal{E})$ is even. Now the line bundle on $\tilde{C} \times \Sigma$ used in Yin's construction of $\mathcal{E}$ is only defined up to tensoring with the pullback of a line bundle $\mathcal{N}$ from Σ, so by the projection formula, $\mathcal{E}$ may be replaced by $\mathcal{E} \otimes \mathcal{N}$ where $\mathcal{N}^{-2} = \det(\mathcal{E})$, so we may assume $\det(\mathcal{E}) \cong \mathcal{O}$.

Next we show $\mathcal{E}$ is semi-stable. Any section σ of $\mathbb{P}(\mathcal{E}) \cong X$ of self intersection 2 defines a sub-line bundle $\xi^{-1} \subset \mathcal{E}$ and an exact sequence $0 \to \xi^{-1} \to \mathcal{E} \to \xi \to 0$, with $2 = (\sigma^2) = 2\deg(\xi)$, so that ξ belongs to $\mathrm{Pic}^1(\Sigma)$. If $\mathcal{E}$ is not semi-stable, and $\mathcal{L}$ is a line bundle of positive degree mapping into $\mathcal{E}$, it cannot map into ξ^{-1} hence maps nontrivially into ξ by composition, hence has degree exactly 1. Thus $\mathcal{L} \cong \xi$, and the sequence splits, i.e. $\mathcal{E} \cong \xi^{-1} \oplus \xi$. We will show this is impossible by showing it contradicts the representation in §4 of $\tilde{C}$ as a complete curve of sections of X of self intersection 2. I.e. if $\mathcal{E} \cong \xi^{-1} \oplus \xi$, then we can represent $X \cong \mathbb{P}(\xi^{-1} \oplus \xi) \cong \mathbb{P}(\mathcal{O} \oplus \xi^2)$, so that X is the completion of the line bundle ξ^2 by adding a section S_∞ at infinity, and the section of X corresponding to ξ^{-1} corresponds to the zero section S_0 of the line bundle ξ^2.

Claim. If $X \cong \mathbb{P}(\mathcal{O} \oplus \xi^2)$, then all sections of X of self intersection 2 correspond to elements of the vector space of regular sections of the line bundle ξ^2.

Proof. Regular sections of the line bundle give regular sections of the completion of the line bundle, i.e. of X. Conversely, we must show a section of

X misses the section S_∞ at infinity. A section of X is irreducible (since isomorphic to Σ), hence cannot contain the infinite section without equalling it. Since S_∞ has self intersection -2, it cannot equal a regular section of self intersection 2, so must meet it properly. Since S_0 and F (the zero section and a fiber), form a $\mathbb{Z}$ homology basis of X, a section σ of X is homologous to $aS_0 + bF$, where $F \cdot S_0 = 1$, and a, b are integers. Since $F \cdot \sigma = 1$, and $F \cdot F = 0$, we have $a = 1$, so $\sigma = S_0 + bF$, and $(\sigma^2) = 2 = 2 + 2b$. Thus $b = 0$, $\sigma = S_0$ in homology, hence $\sigma \cdot S_\infty = 0$. But since σ meets the infinite section properly, σ must be disjoint from S_∞. □

Since $\tilde{C}$ is complete, $\tilde{C}$ cannot map injectively into the affine space of all sections of $\mathbb{P}(\mathcal{O} \oplus \xi^2)$ of self intersection 2. Since by the argument in §4, $\tilde{C}$ does inject into the variety of sections of X of self intersection 2, thus $\mathcal{E}$ is not isomorphic to $\xi^{-1} \oplus \xi$, and hence $\mathcal{E}$ is semi-stable. Now that we know $\mathcal{E}$ is semi-stable, the Narasimhan-Ramanan curve $D_{\mathcal{E}} = \{\xi$ in $\mathrm{Pic}^1(\Sigma) : H^0(\mathcal{E} \otimes \xi) \neq 0\}$, is defined and equals the support of a divisor $\tilde{C}_{\mathcal{E}}$ in $|2\Theta|$. By Verra's explicit description of curves in $|2\Theta|$, $D_{\mathcal{E}}$ is either reduced and irreducible of arithmetic genus five, or a union of one or two copies of Σ. Moreover we have maps $\tilde{C} \to Z \to D_{\mathcal{E}}$, where Z is a curve of sections of $\varphi : X \to \Sigma$ of self intersection 2, and $\tilde{C} \to Z$ is bijective. By [**Y**, Lemma 4, p. 23], the composition $\tilde{C} \to D_{\mathcal{E}}$ is a nonconstant finite map from $\tilde{C}$ to $D_{\mathcal{E}}$. We argue next this map is generically injective. By [**Y**, Lemma 3, p. 23], Z is an irreducible component of the variety of sections of X of self intersection 2. If $p \neq q$ are distinct points of $\tilde{C}$ corresponding to distinct sections $\sigma_p \neq \sigma_q$ of X, but the same line bundle $\xi_p = \xi_q$ in $\mathrm{Pic}^1(\Sigma)$, then ξ_p^{-1} embeds onto more than one sub-line bundle of $\mathcal{E}$, so the vector space $\mathrm{Hom}(\xi_p^{-1}, \mathcal{E})$ has dimension ≥ 2. Thus by varying the embedding of ξ_p in $\mathcal{E}$, there is a $\mathbb{P}^1$ of sections of X passing through the point σ_p on Z, and all having self intersection 2.

Since the component Z is birational to $\tilde{C}$ of genus 5, this can only occur for a finite number of points p, so the map $\tilde{C} \to D_{\mathcal{E}}$ is generically injective as claimed. Then the divisor $\tilde{C}$ of $|2\Theta|$ supported on $D_{\mathcal{E}}$ is not one of the reducible divisors $\Theta_\mu + \Theta_{\mu^{-1}}$ on the Kummer surface, for μ in $\mathrm{Pic}^0(\Xi)$, which parametrize semi-stable, nonstable bundles of the form $\mu \oplus \mu^{-1}$. Hence $\mathcal{E}$ is in fact stable, and $D_{\mathcal{E}}$ is reduced irreducible of arithmetic genus five, and thus $\tilde{C} \to D_{\mathcal{E}}$ is an isomorphism. □

It now follows from Lemma 1 that Z is the only component of sections of X of self intersection 2, and both maps $\tilde{C} \to Z \to D_{\mathcal{E}}$ are isomorphisms. Consequently, the divisor $\tilde{C}_{\mathcal{E}}$ associated by [**N-R**] to the $\mathbb{P}^1$ bundle $X = \mathbb{P}(\mathcal{E})$ is isomorphic to the original smooth genus five curve $\tilde{C}$, hence we have proved

Corollary 3. *Yin's correspondence $\tilde{C}/C \mapsto X$ is right inverse to Narasimhan-Ramanan's correspondence $X = \mathbb{P}(\mathcal{E}) \mapsto \tilde{C}_{\mathcal{E}}$, (modulo the action of* *the group* $G \cong (\mathbb{Z}_2)^4$ *of points of order two in* $\mathrm{Pic}^4(\tilde{C})$). (*There are* 16 *choices of the vector bundle* $\mathcal{E}$ *on* Σ *defining the same* $\mathbb{P}^1$ *bundle* X*, and* 16 *divisors* $\tilde{C}_{\mathcal{E}}$ *in* $|2\Theta|$ *defining isomorphic curves, since* G *acts on* $|2\Theta|$.)

Since the map $\tilde{C}/C \mapsto X$ is right inverse to N-R's, and theirs is known to be an isomorphism between $\mathcal{M}$ and $|2\Theta|$, it follows that Yin's construction is also left inverse to N-R's, at least on an open subset of the set of stable bundles in $\mathcal{M}$. I.e. a stable bundle $\mathcal{E}$ yields an irreducible curve $\tilde{C}$, and if $\tilde{C}$ is smooth and nonhyperelliptic, and if X is the $\mathbb{P}^1$ bundle over Ξ defined in the Prym construction associated to $(\tilde{C}, \iota)$, then $X \cong \mathbb{P}(\mathcal{E})$.

Without assuming the result that the map of [**N-R**] is an isomorphism we can partially recover it as follows: Given a stable vector bundle $\mathcal{E}$ of determinant $\mathcal{O}$ on a smooth genus two curve Σ, such that the divisor $\tilde{C}_{\mathcal{E}}$ in $|2\Theta|$ defined by [**N-R**] is smooth, it follows from [**V**] that the canonical involution ι of $\mathrm{Pic}^1(\Sigma)$ is fix point free on $\tilde{C}_{\mathcal{E}}$, and that $\tilde{C}_{\mathcal{E}}$ and $\tilde{C}_{\mathcal{E}}/\iota$ are both nonhyperelliptic. If we put $X = \mathbb{P}(\mathcal{E})$ and denote the $\mathbb{P}^1$ bundle map by $\varphi : X \to \Sigma$, then it follows from Lemma 1 above that the curve Z of sections of $\varphi : X \to \Sigma$ of self intersection 2 is isomorphic to $\tilde{C}_{\mathcal{E}}$. Hence we may write $\tilde{C}_{\mathcal{E}}$ for Z and consider the tautological map $\gamma : \tilde{C}_{\mathcal{E}} \times \Sigma \to X$ given by evaluating a section at a point in Σ, i.e. $\gamma(\sigma, y) = \sigma(y)$. Thus the points y of Σ parametrize a family of maps $\gamma_y : \tilde{C}_{\mathcal{E}} \to \varphi^{-1}(y) \cong \mathbb{P}^1$. Since these maps are limits of such maps in cases where we know X arises from the Abel parametrization of the Prym theta divisor $\Xi \cong \Sigma$, and in those cases γ_y is a degree 4 map $\tilde{C} \to \mathbb{P}^1$, it follows that here also $\mathrm{degree}(\gamma_y) = 4$. If $\mathcal{L} = \gamma^*(\mathcal{O}_{\mathcal{E}}(1))$, the family of line bundles $\mathcal{L}_y$ on $\tilde{C}_{\mathcal{E}}$ for y in Σ thus defines a map $\Sigma \to \mathrm{Pic}^4(\tilde{C}_{\mathcal{E}})$ with target in $W^1_4(\tilde{C}_{\mathcal{E}}) = \mathrm{sing}(\Theta(\tilde{C}_{\mathcal{E}}))$. Since as above this map is a limit of embeddings of $\Sigma \cong \Xi = P^+ \cap \tilde{\Theta} \subset \mathrm{Pic}^4(\tilde{C})$ of Σ onto a Prym theta divisor, it follows that the limiting map is still nonconstant in our case. Since $\mathrm{sing}(\Theta(\tilde{C}_{\mathcal{E}}))$ has only two components, one of genus 4 and one of genus 2 [**ACGH**, p. 274], it follows that the map $\Sigma \to \mathrm{Pic}^4(\tilde{C}_{\mathcal{E}})$ is again an embedding onto the unique genus two component, i.e. onto the Prym theta divisor Ξ associated to $(\tilde{C}_{\mathcal{E}}, \iota)$. Thus $\mathcal{L} = \gamma^*(\mathcal{O}_{\mathcal{E}}(1))$ is the restriction to $\tilde{C}_{\mathcal{E}} \times \Sigma$ of a Poincaré line bundle for $\tilde{C}_{\mathcal{E}} \times \mathrm{Pic}^4(\tilde{C}_{\mathcal{E}})$. It follows then that pushing down this $\mathcal{L}$ gives the $\mathbb{P}^1$ bundle $X(\tilde{C}_{\mathcal{E}}, \iota)$ coming from the Abel parametrization of $\Sigma \cong \Xi(\tilde{C})$, so to show $\mathbb{P}(\mathcal{E}) \cong X(\tilde{C}_{\mathcal{E}}, \iota)$ we must compare the push down of $\mathcal{L}$ with the original bundle $\mathcal{E}$.

Lemma 2. *With notation as above,* $\mu_*(\mathcal{L}) \cong \mu_*(\gamma^*(\mathcal{O}_{\mathcal{E}}(1))) \cong \mathcal{E}$.

Proof. Look again at the maps $\gamma : \tilde{C}_{\mathcal{E}} \times \Sigma \to X = \mathbb{P}(\mathcal{E})$, $\varphi : X \to \Sigma$ and (their composition) the second projection $\varphi \circ \gamma = \mu : \tilde{C}_{\mathcal{E}} \times \Sigma \to \Sigma$. Then pull back the bundle $\mathcal{E}$ from Σ to $\varphi^*(\mathcal{E})$ on X, and consider the fundamental exact sequence $0 \to \mathcal{O}_{\mathcal{E}}(-1) \to \varphi^*(\mathcal{E}) \to \mathcal{O}_{\mathcal{E}}(1) \to 0$, where the quotient sheaf is dual to $\mathcal{O}_{\mathcal{E}}(-1)$ because $\det(\mathcal{E}) \cong \mathcal{O}$. Then pull back this sequence by γ to $\tilde{C} \times \Sigma$ getting $0 \to \gamma^*(\mathcal{O}_{\mathcal{E}}(-1)) \to \gamma^*\varphi^*(\mathcal{E}) \to \gamma^*(\mathcal{O}_{\mathcal{E}}(1)) \to 0$, and push that down to Σ by μ, getting a long exact sequence that begins with $0 \to \mu_*(\gamma^*\mathcal{O}_{\mathcal{E}}(-1)) \to \mu_*(\gamma^*\varphi^*(\mathcal{E})) \to \mu_*\gamma^*(\mathcal{O}_{\mathcal{E}}(1))$. Since the line bundle $\gamma^*(\mathcal{O}_{\mathcal{E}}(-1))$ restricts to have degree -4 on the fibers of μ, the leftmost sheaf is zero and we have an injection of bundles $\mu_*(\gamma^*\varphi^*(\mathcal{E})) \to \mu_*\gamma^*(\mathcal{O}_{\mathcal{E}}(1))$. Since $\gamma^*(\mathcal{O}_{\mathcal{E}}(1))$ restricts to line bundles of degree 4 with two sections on the fibers of μ, the second bundle is also a 2-plane bundle, and the map is an isomorphism. Since by the projection formula $\mu_*(\gamma^*\varphi^*(\mathcal{E})) \cong \mathcal{E}$, we have $\mathcal{E} \cong \mu_*(\gamma^*\varphi^*(\mathcal{E})) \cong \mu_*\gamma^*(\mathcal{O}_{\mathcal{E}}(1)) = \mu_*(\mathcal{L})$. $\square$

Remark. Since by hypothesis $\det(\mathcal{E}) \cong \mathcal{O}$, not only is $\mathbb{P}(\mathcal{E}) \cong X(\tilde{C}_{\mathcal{E}}, \iota)$, the Abel parametrization of the Prym theta divisor $\Xi(\tilde{C}_{\mathcal{E}}, \iota)$, but $\mathcal{E}$ is in fact one of the 16 normalized push downs of Poincaré line bundles associated to $X(\tilde{C}_{\mathcal{E}}, \iota)$ above.

6. Recovering hyperelliptic double covers with the same Prym variety as $\pi : \tilde{C} \to C$, from X.

Next we show that X determines geometrically a family of hyperelliptic double covers $\tilde{C}_\lambda/C_\lambda$ parametrized by the dual projective plane $|K_C|^*$, and whose Prym varieties are all isomorphic to $P(\tilde{C}/C)$. These hyperelliptic double covers occur as fibers of the map h over lines H_λ in $|K_C|$. We know from our calculation at the end of Section 2 that these curves $h^{-1}(H_\lambda) = \tilde{C}_\lambda$ have genus 5. If H is transverse to the branch curve C^* then $\tilde{C}_\lambda$ is smooth, and we claim the involution on X restricts to a fix point free one on $\tilde{C}_\lambda$. Indeed, if $\overline{p} + \overline{q} + \overline{r} + \overline{s}$ is a canonical divisor on C without multiple points, then the π inverse image on $\tilde{C}$ consists of eight distinct points $p+p'+q+q'+r+r'+s+s'$ on $\tilde{C}$, so the involution ι on X takes $(p+q+r+s)$ for instance to $p'+q'+r'+s'$. In general for any D lying above $\overline{p}+\overline{q}+\overline{r}+\overline{s}$, D and $\iota(D)$ have disjoint supports, and in particular ι has no fixed points on X lying above $\overline{p}+\overline{q}+\overline{r}+\overline{s}$. Even if $D = 2\overline{p}+\overline{r}+\overline{s}$ is a general canonical divisor with multiple points, corresponding to a smooth point of C^*, ι does not fix any of the 6 points of X lying over D. Indeed the only way we can have $\iota(D) = D$, is for D to equal the union of two fibers of π, i.e. $D = p + p' + q + q'$. This occurs only over a canonical divisor of form $D = 2\overline{p} + 2\overline{q}$ on C, i.e. a divisor cut by one of the 28 bitangents to the canonical plane quartic C. Hence ι fixes at most 28 points of X, one over each node of the dual curve C^*. Thus if H_λ is a line in $|K_C|$ transverse to C^*, in particular not containing a node or cusp of C^*, then $\tilde{C}_\lambda$ is smooth and ι restricts to a fix point free involution of $\tilde{C}_\lambda$. We claim next that φ restricts on each such $\tilde{C}_\lambda$ to the “Abel-Prym” map $\varphi : \tilde{C}_\lambda \to \Xi \subset P$.

To compute the degree of φ on the curve $\tilde{C}_\lambda$ we must intersect $\tilde{C}_\lambda$ with a general $\mathbb{P}^1$ fiber of φ. Since the family $\{\tilde{C}_\lambda\}$ of inverse images of the lines H_λ in $|K_C|$ is a linear system on the smooth variety X, that intersection number is constant over the linear system, and we have just computed it to be 2 at a reducible curve of the form $D_p \cup D_{p'}$, the inverse image of a generic tangent line $H_{\overline{p}}$ to C^*. Thus $\varphi : \tilde{C}_\lambda \to \Xi \subset P$, is a degree two map from a smooth genus five curve with fix point free involution, to a two dimensional p.p.a.v., whose image is a one-cycle with homology class $2[\Xi]$. Moreover the involution induced on the source $\tilde{C}_\lambda$ by the involution on X, commutes with the natural involution $\mu \mapsto (K_{\tilde{C}} - \mu)$ on the target $\Xi \subset \mathrm{Pic}^4(\tilde{C})$. Hence, by Masiewicki's criterion [**Ma**], φ is the Abel-Prym map from the curve $\tilde{C}_\lambda$ to P, and since φ is 2:1 rather than injective, Masiewicki also implies that $\tilde{C}_\lambda$ is hyperelliptic. (Note that a special fiber $D_p \cup D_{p'}$ of h over a general tangent line $H_{\overline{p}}$ to C^* at a smooth point representing a canonical divisor $2\overline{p} + \overline{r} + \overline{s}$ on C with $\overline{p}, \overline{r}, \overline{s}$, distinct, is a hyperelliptic "Wirtinger" double cover of a nodal curve with normalization isomorphic to $\Sigma \cong \Xi$, cf [**B**, p. 175, Th. 5.4(i)].) This gives a parametrization of part of the hyperelliptic locus in the fiber of the Prym map over (P, Ξ) by the projective plane $|K_C|^*$. Moreover, since the hyperelliptic Wirtinger covers obtained vary with one parameter, the smooth hyperelliptic covers vary with two parameters, hence give the generic smooth hyperelliptic double cover in the fiber $\mathcal{P}^{-1}(J(\Sigma), \Sigma)$.

7. Remarks on some exceptional double covers.

Following the notation of the introduction, if $\mathcal{M}/G$ is the moduli space of S-equivalence classes of semi-stable even $\mathbb{P}^1$ bundles over the smooth genus two curve Σ with $\mathrm{Aut}(\Sigma) \cong \mathbb{Z}_2$, and $\mathcal{P}^{-1}(J(\Sigma))$ is the fiber of the Prym map over the Jacobian $J(\Sigma)$, we have shown that Yin's construction $(\tilde{C} \to C) \mapsto X$ in [**Y**] provides an explicit inverse $\mathcal{P}^{-1}(J(\Sigma)) \dashrightarrow \mathcal{M}/G$ to the Narasimhan-Ramanan-Verra map, at least for those classical étale double covers $\pi : \tilde{C} \to C$ such that $\tilde{C}$ is smooth and non hyperelliptic. One can ask what the construction yields for other double covers in the fiber $\mathcal{P}^{-1}(J(\Sigma))$ of Beauville's extended Prym map [**B**], for instance those corresponding to points on the two components of the exceptional divisor of the composition $\mathcal{P}^{-1}(J(\Sigma)) \cong \sigma_2(\sigma_1(\mathcal{M}/G))$ of the Narasimhan-Ramanan isomorphism with Verra's blowings up. General points of those two divisors correspond to étale double covers $\tilde{C} \to C$ where either C and $\tilde{C}$ are smooth and hyperelliptic, or $C = \Sigma \cup E$ is an "elliptic tail" (a copy of Σ joined at one point to an elliptic curve E) and $\tilde{C} = \Sigma_1 \cup \tilde{E} \cup \Sigma_2$ is two copies of Σ each joined at one point to an elliptic double cover $\tilde{E}$ of E. When $\tilde{C}$ is smooth and hyperelliptic, the Prym variety P can still be realized as those line bundles L in $\mathrm{Pic}^4(\tilde{C})$ with $\mathrm{Nm}(L) = \omega_C$ and $h^0(L)$ even, and thus we can define X again by restricting the Abel map over $\Xi \subset P$, $\tilde{C}^{(4)} \supset X \to \Xi \subset P \subset \mathrm{Pic}^4(\tilde{C})$. In this case X is always isomorphic to the trivial semi-stable $\mathbb{P}^1$ bundle $X \cong \Sigma \times |g^1_2(\tilde{C})| \cong \Sigma \times \mathbb{P}^1$, and hence X cannot recover uniquely the hyperelliptic double cover $\tilde{C} \to C$. Interestingly, the restricted norm map

$h : X \to |K_C|$ in this case is still finite and recovers at least the curve C from the branch locus, as in Andreotti's proof of Torelli for hyperelliptic curves. The reason h carries more information than X here is that the net $h^*(|K_C|^*)$ is not complete on X; in fact $h^0(X, \mathcal{T}_\varphi \otimes \varphi^*(K_\Sigma)) = 6$, so the map h is not determined by X in this case.

In the case where C has an elliptic tail, even the definition of the generalized Prym variety as a subvariety of $\mathrm{Pic}^4(\tilde{C})$ breaks down. I.e. Beauville realizes the generalized Prym variety naturally in $\mathrm{Pic}^{2g-2}(\tilde{C})$ (where $2g-1 =$ genus$(\tilde{C})$), only when the double cover satisfies condition $(*)$, ([**B**, p. 157]). When $(*)$ holds, the multidegree $\underline{d} = \underline{\deg}(\omega_C)$ is even, and P is then defined naturally in $\mathrm{Pic}^{\underline{d}/2}(\tilde{C})$, (Props. (3.10), (3.11) p. 162, and Rmk.(3.12) p. 163, which applies only to $(*)$ covers). When $C = \Sigma \cup E$, then $\underline{\deg}(\omega_C) = (3,1)$, $\underline{\deg}(\omega_{\tilde{C}}) = (3,2,3)$, the multidegree is no longer even, in particular $(*)$ fails, there is no natural component of the reducible variety $\mathrm{Pic}^4(\tilde{C})$ in which to locate the base curve Σ of the $\mathbb{P}^1$ bundle, and the construction given above of X breaks down. When the double cover is an étale "Wirtinger cover" $\tilde{C} = \Sigma_1 \cup \Sigma_2 \to \Sigma/p \sim q = C$ ([**V**, 3.15, p. 442]), then $\underline{\deg}(\omega) = 4, \underline{\deg}(\omega_{\tilde{C}}) = (4,4)$, hence evenness holds, and one can define a theta divisor $\tilde{\Theta}$ in $\mathrm{Pic}^{(2,2)}(\tilde{C})$ ([**B**], §2 p. 153, Prop. (2.2) p. 155). Since this double cover is étale over the node however, $(*)$ is violated, and the Prym variety is apparently not realized in $\mathrm{Pic}^{(2,2)}(\tilde{C})$, i.e. Rmk. (3.12), p. 163, [**B**], seems to fail.

As a smooth étale double cover $\pi : \tilde{C} \to C$ becomes either an elliptic tail $\tilde{C} = \Sigma_1 \cup \tilde{E} \cup \Sigma_2 \to \Sigma \cup_p E = C$, or a "Wirtinger cover" $\tilde{C} = \Sigma_1 \cup \Sigma_2 \to \Sigma/p \sim q = C$ (cf. [**V**], 3.15, 3.16, p. 442), (the two cases not satisfying Beauville's condition $(*)$), and hence $\mathrm{Pic}^4(\tilde{C})$ becomes reducible, it appears that the $\mathbb{P}^1$ bundles defined by the restricted Abel map $\tilde{C}^{(4)} \cup X \to \Xi \subset P \subset \mathrm{Pic}^4(\tilde{C})$ become, in some of the components of $\mathrm{Pic}^4(\tilde{C})$, several copies of restrictions of the classical Abel map $\Sigma^{(3)} \to \mathrm{Pic}^3(\Sigma)$, restricted over Abel curves in $\mathrm{Pic}^3(\Sigma)$, i.e. $\Sigma^{(3)} \supset X \to \Sigma + p + q \subset \mathrm{Pic}^3(\Sigma)$ or $\Sigma^{(3)} \supset X \to \Sigma + 2p \subset \mathrm{Pic}^3(\Sigma)$. If $p' = |g^1_2 - p|$, and $q' = |g^1_2 - q|$, then the curves $\Sigma/p \sim q \cong \Sigma/p' \sim q'$, and $\Sigma \cup_p E \cong \Sigma \cup_{p'} E$ are isomorphic in pairs, but the bundles $\Sigma^{(3)} \supset X \to \Sigma + p + q \subset \mathrm{Pic}^3(\Sigma)$, $\Sigma^{(3)} \supset X \to \Sigma + p' + q' \subset \mathrm{Pic}^3(\Sigma)$, $\Sigma^{(3)} \supset X \to \Sigma + 2p \subset \mathrm{Pic}^3(\Sigma)$, and $\Sigma^{(3)} \supset X \to \Sigma + 2p' \subset \mathrm{Pic}^3(\Sigma)$ although semi-stable, are at least not obviously isomorphic. Hence depending on the model chosen for the nodal curve C, one would obtain more than one possible limiting $\mathbb{P}^1$ bundle. This seems to match up with the fact that a general S-equivalence class of a nonstable, semi-stable, even $\mathbb{P}^1$ bundle over Σ, has two nonsplit representative isomorphism classes which are interchanged by the hyperelliptic involution ([**N-R**], Rmk. 1, p. 35). Note that the definition of the bundles $X \to \Sigma + p + q$ and $X \to \Sigma + 2p$ uses the information of the marked points p, or p and q on the curves C, but not the j-invariant of the elliptic tail. The maps h appear to be defined by complete linear series, with branch loci dual to either the nodal quartic $\Sigma/p \sim q$ or the quartic model of Σ with a cusp at p, so the j-invariant of the elliptic tail appears to be lost by the map h. A calculation shows that in each S-equivalence class corresponding to either a general elliptic tail or Wirtinger cover, there are exactly three nonisomorphic $\mathbb{P}^1$ bundles over Σ, one split and two nonsplit. One can construct a smooth, nonseparated moduli space containing both nonsplit bundles (see [**N-R**] p. 21, proof of Thm. 1), but there seems no reason to expect a natural construction which would pick out one of them. The split bundle is of course distinguished, but has too many deformations to be included naturally in a moduli space of isomorphism classes of general even $\mathbb{P}^1$ bundles over Σ, and we have not observed it arising even as a limit from the geometry of the Abel map construction in this paper.

References

[A] A. Andreotti, *On a theorem of Torelli*, Am. J. of Math., **80** (1958), 801-828.

[ACGH] E. Arbarello, M. Cornalba, P.A. Griffiths and J. Harris, *Geometry of Algebraic Curves, Grund. der Math. Wissen*, **267**, Springer, (1985).

[A-M] A. Andreotti and A. Mayer, *On period relations for abelian integrals on algebraic curves*, Ann. Scuola Norm. Sup. Pisa, **21** (1967), 189-238.

[B] A. Beauville, *Prym Varieties and the Schottky problem*, Invent. Math., **41** (1977), 149-196.

[Ma] L. Masiewicki, *Universal properties of Prym varieties with an application to algebraic curves of genus five*, Trans. Am. Math. Soc., **222** (1976), 221-240.

[Mu1] D. Mumford, *Prym Varieties* I, in 'Contributions to Analysis', Academic Press, New York, (1974), 325-350.

[Mu2] ———, *Theta characteristics of an algebraic curve*, Ann. scient. Ec. Norm Sup., **4.4** (1971), 181-192.

[N-R] M.S. Narasimhan and S. Ramanan, *Moduli of vector bundles on a compact Riemann surface*, Annals of Math., **89**(**2**) (1969), 14-51.

[V] A. Verra, *The fibre of the Prym map in genus three*, Math. Annalen, **276** (1987), 433-448.

[W] G. Welters, *Abel-Jacobi Isogenies for Certain Types of Fano Threefolds*, Mathematical Centre Tracts, Amsterdam, **141** (1981).

[Y] H. Yin, *Deformation of special subvarieties of divisors associated to double covers of genus three curves*, PhD thesis, University of Georgia, 1996.

Received July 21, 1997 and revised March 31, 1998.

University of Georgia
Athens, Georgia 30602
E-mail address: roy@math.uga.edu

University of Georgia
Athens, Georgia 30602

E-mail address: rvarley@math.uga.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

ON THE COMPLEXITY OF RATIONAL PUISEUX EXPANSIONS

P.G. Walsh

Dedicated to the memory of Professor Bernard Dwork

Duval defined and studied rational Puiseux expansions. In this paper we first prove that the existence of rational Puiseux expansions follows from the structure of algebraic extensions of a completion of the rational function field. We then describe a canonical system of rational Puiseux expansions, which are constructed in terms of the coefficients of classical Puiseux expansions. Using recent effective results on algebraic functions, we use this construction to prove that a system of rational Puiseux expansions exists whose height can be bounded in terms of the degrees and height of the polynomial determining the rational Puiseux expansions.

1. Introduction.

The layout of the paper is as follows. In Section 2 we make some definitions and state the main results of the paper. In Section 3 we discuss some of the theory of algebraic extensions of complete fields. In Section 4 some preliminary results are proved. In Section 5 the main theorems on the existence of rational Puiseux expansions are proved. Finally, in Section 6 we present an explicit construction of a system of rational Puiseux expansions, and thereby provide an effective version of the existence theorem.

2. Statement of main results.

Throughout this paper $F \in \mathbf{Q}[x, y]$ will denote a polynomial of degree $n > 0$ in y and of degree m in x. We will assume that $disc_y F \neq 0$, where $disc_y F$ is the discriminant of F, and F is regarded as a polynomial in y. Puiseux's theorem (see for example p. 118 of [**7**]) asserts the existence of n distinct formal series

$$y_i(x) = \sum_{k=f_i}^{\infty} a_{k,i} \left(x^{1/e_i}\right)^k \qquad (i = 1, ..., n)$$

such that

$$F(x,y) = v(x)\prod_{i=1}^{n}(y - y_i(x)), \tag{2.1}$$

where $v(x)$ is the leading coefficient of F when F is regarded as a polynomial in y, the $a_{k,i}$ are complex numbers, and f_i are integers defined by the condition $a_{f_i,i} \neq 0$. These are the n *Puiseux expansions* at $x = 0$ of the algebraic function y defined by $F(x,y) = 0$. For $i = 1, ..., n$, the integer e_i is the *ramification index* of the series y_i, and is minimal in the sense that if d is any positive divisor of e_i, then there is an index k not divisible by d such that $a_{k,i} \neq 0$. Throughout this paper we will make reference to

$$y(x) = \sum_{k=f}^{\infty} a_k \left(x^{1/e}\right)^k, \tag{2.2}$$

which is one of the n series given above.

Let ζ_e denote a primitive e-th root of unity. The *branch* of the series $y(x)$ is the set of series

$$B(y(x)) = \left\{ \sum_{k=f}^{\infty} a_k \left(\zeta_e^j x^{1/e}\right)^k ;\ j = 0, \ldots, e-1 \right\}.$$

Note that $B(y(x))$ contains precisely e distinct series. Let $L = \mathbf{Q}(a_f, a_{f+1}, \ldots)$, $s = [L : \mathbf{Q}]$, and $\sigma_1, \sigma_2, ..., \sigma_s$ be the s embeddings of L into $\overline{\mathbf{Q}}$, where $\overline{\mathbf{Q}}$ is an algebraic closure of $\mathbf{Q}$. The *conjugacy class* of $y(x)$ is

$$C(y(x)) = \left\{ \sum_{k=f}^{\infty} \sigma_i(a_k) \left(\zeta_e^j x^{1/e}\right)^k ;\ i = 1, \ldots, s\ ,\ j = 0, \ldots, e-1 \right\}.$$

Note that $C(y(x))$, and hence $B(y(x))$, consist entirely of Puiseux expansions at $x = 0$ of the algebraic function y defined by $F(x,y) = 0$.

In what follows, for any field E, $E((x)) = \{\sum_{k=l}^{\infty} c_k x^k\ ;\ l \in \mathbf{Z}\ ,\ c_k \in E \text{ for } k \geq l\}$ is the field of Laurent series in x with coefficients in E. The following is a consequence of the fact that the n Puiseux expansions of y are distinct. The details can be found in Lemma 1 and Lemma 2 of [**16**].

Proposition 2.1.

1. *The product* $\prod_{B(y(x))}(y - y_i(x))$ *is irreducible in* $\overline{\mathbf{Q}}((x))[y]$, *of degree* e *in* y.

2. *The product* $\prod_{C(y(x))}(y - y_i(x))$ *is irreducible in* $\mathbf{Q}((x))[y]$ *of degree* $e(s/s_0)$ *in* y, *where*
$s_0 = \#\left\{\sigma : L \hookrightarrow \overline{\mathbf{Q}};\ \exists\, t \in \mathbf{Z} \text{ such that } \sigma(a_k) = a_k \zeta_e^{tk} \text{ for all } k \geq f\right\}$.

The above result illustrates that the classical Puiseux expansions of y determine information about the factorization of F in $\overline{\mathbf{Q}}((x))[y]$, but fail to exhibit explicit information about the factorization of F in $\mathbf{Q}((x))[y]$. We will see that a rational Puiseux expansion associated to $y(x)$ has coefficients in a subfield of $L = \mathbf{Q}(a_f, a_{f+1}, \dots)$ which is of degree precisely s/s_0 over $\mathbf{Q}$. In this way, the rational Puiseux expansions not only contain all of the information of the classical Puiseux expansions, but also give information about the factorization of F in $\mathbf{Q}((x))[y]$ (see Theorem 2 of [**4**]).

We now proceed to define rational Puiseux expansions. A *parametrization* of the branch $B(y(x))$, determined by $y(x)$, with $y(x)$ as in (2.2), is the pair $(x, y) = (Z^e, \sum_{k=f}^{\infty} a_k Z^k)$. Parametrizations of $B(y(x))$ determined by two different elements in $B(y(x))$ are *equivalent parametrizations.* A *generalized parametrization* of $B(y(x))$, determined by $y(x)$, is $(x, y) = (\lambda Z^e, \sum_{k=f}^{\infty} \alpha_k Z^k)$, where $\lambda \in \overline{\mathbf{Q}}$, $\alpha_k = a_k(\lambda^{1/e})^k$ for $k \geq f$, and $\lambda^{1/e}$ is a fixed e-th root of λ. Two generalized parametrizations of $B(y(x))$ are *equivalent* if they are determined by two elements of $B(y(x))$. If $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, then σ acts on a generalized parametrization by $\sigma(\lambda Z^e, \sum_{k=f}^{\infty} \alpha_k Z^k) = (\sigma(\lambda) Z^e, \sum_{k=f}^{\infty} \sigma(\alpha_k) Z^k)$.

Definition. Assume that the algebraic function y has g branches. A *system of rational Puiseux expansions* of the algebraic function y is a set of g pairwise inequivalent generalized parametrizations of the branches of y with the property of being invariant under the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. Elements in such a system are called *rational Puiseux expansions.*

By invariance under the Galois action, we mean that under the action of any element in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, the images of two distinct generalized parametrizations are distinct.

For the sake of illustration we provide an example, given in [**4**]. Let

$$F(x, y) = (x^2 + y^2)^3 - 4x^2y^2.$$

In this case the algebraic function y has the 6 Puiseux expansions

$$\begin{aligned} y_1(x) &= 1/2x^2 + \cdots, \quad y_2(x) = -1/2x^2 + \cdots, \\ y_3(x) &= \sqrt{2}x^{1/2} + \cdots, \quad y_4(x) = i\sqrt{2}x^{1/2} + \cdots, \\ y_5(x) &= -\sqrt{2}x^{1/2} \cdots, \quad y_6(x) = -i\sqrt{2}x^{1/2} + \cdots. \end{aligned}$$

In this case the branches (and conjugacy classes) of the algebraic function y are

$$\{y_1(x)\}, \{y_2(x)\}, \{y_3(x), y_5(x)\}, \{y_4(x), y_6(x)\},$$

and a corresponding set of rational Puiseux expansions is the set of pairs

$$\{(T, 1/2T^2 + \cdots), (T, -1/2T^2 + \cdots), (1/2T^2, T + \cdots), (-1/2T^2, T + \cdots)\}.$$

In this example we see that the coefficients of the rational Puiseux expansions all lie in $\mathbf{Q}$. By Theorem 2.2 below this is not surprising since, for example for $y_3(x)$, $e = 2$, $s = 2$, and $s_0 = 2$, showing that the coefficients of the corresponding rational Puiseux expansion lie in an extension of $\mathbf{Q}$ of degree $s/s_0 = 2/2 = 1$. To see that $s_0 = 2$, we use the fact that all 6 of the Puiseux expansions have different leading terms, which, by the proof of Theorem 2.2, implies that the value l in Theorem 2.2 can be taken to be $f = 1$.

Note that rational Puiseux expansions are in one to one correspondence with the branches of the algebraic function y. Moreover, the classical Puiseux expansions are easily recovered from a system of rational Puiseux expansions. In particular, if $(\lambda Z^e, \sum_{k=f}^{\infty} \alpha_k Z^k)$ is a generalized parametrization representing a branch B, then B consists of the e series

$$\left\{ \sum_{k=f}^{\infty} \alpha_k (\lambda^{-1/e} \zeta^j)^k (x^{1/e})^k; \;\; j = 0, ..., e-1 \right\},$$

for some e-th root of unity ζ and some e-th root, $\lambda^{-1/e}$, of λ^{-1}.

The following result was proved in [**4**]. We provide an alternative proof based on the structure of algebraic extensions of complete fields.

Theorem 2.1. *Let $F(x,y) \in \mathbf{Q}[x,y]$ with $disc_y F \neq 0$, and let y denote the algebraic function defined by $F(x,y) = 0$. Then there exists a system of rational Puiseux expansions of y.*

The following result shows how a rational Puiseux expansion gives explicit information about the factorization of F in $\mathbf{Q}((x))[y]$.

Theorem 2.2. *Let $y(x)$ be as in (2.2), and let $(\lambda Z^e, \sum_{k=f}^{\infty} \alpha_k Z^k)$ be a rational Puiseux expansion parametrizing $B(y(x))$. Let $L = \mathbf{Q}(a_f, a_{f+1}, \cdots)$, $s = [L : \mathbf{Q}]$, and*

$$s_0 = \# \left\{ \sigma : L \hookrightarrow \overline{\mathbf{Q}};\ \exists\, t \in \mathbf{Z} \;\; \textit{such that} \;\; \sigma(a_k) = a_k \zeta_e^{tk} \;\; \textit{for all} \;\; k \geq f \right\}.$$

Then $\mathbf{Q}(\lambda, \alpha_f, \alpha_{f+1}, \cdots)$ is a subfield of L of degree s/s_0 over $\mathbf{Q}$, and hence

$$\#C(y(x)) = e \cdot [\mathbf{Q}(\lambda, \alpha_f, \alpha_{f+1}, \dots) : \mathbf{Q}].$$

Moreover, $\mathbf{Q}(\lambda, \alpha_f, \alpha_{f+1}, \dots) = \mathbf{Q}(\lambda, \alpha_f, \alpha_{f+1}, \dots, \alpha_l)$ for some $l \leq 4mn^2$.

It is noteworthy that these results can be extended to any base field K of characteristic zero, although for the sake of simplicity we will consider only the case that $K = \mathbf{Q}$.

By Proposition 2.1, $\#C(y(x))$ is the degree of the irreducible factor of F in $\mathbf{Q}((x))[y]$ with $y - y(x)$ as a factor. Thus, F is irreducible in $\mathbf{Q}((x))[y]$ precisely if $\#C(y(x)) = n$. In this way, along with Theorem 2.2, a system of rational Puiseux expansions gives information about the factorization of F in $\mathbf{Q}((x))[y]$.

In the final section of this paper we will describe an explicit system of rational Puiseux expansions. From this we deduce the following effective version of Theorem 2.1. We first require some definitions.

If $P \neq 0$ is a (possibly multivariate) polynomial with integer coefficients, the *content* of P is the greatest common divisor of the coefficients of P. If $P \neq 0$ has rational coefficients, the *denominator* of P is the unique positive integer denom(P) such that denom$(P) \cdot P$ has integer coefficients and content equal to one. If P has integer coefficients, then the *height* of P, denoted $ht(P)$, is the maximum of the absolute values of the coefficients of P. More generally, if P has rational coefficients, then the height of P is the maximum of denom(P) and $ht(\text{denom}(P) \cdot P)$. Given an algebraic number α, P_α will denote the defining polynomial of α, that is, the unique irreducible univariate polynomial $P(x)$, with integer coefficients and content equal to one, such that $P(\alpha) = 0$. The *height* of α, denoted $ht(\alpha)$, is defined to be $ht(P_\alpha)$.

Definition. Let $S = \left\{(\lambda_i Z^e, \sum_{k=f_i}^{\infty} \alpha_{i,k} Z^k);\ i = 1, \cdots, r\right\}$ be a system of rational Puiseux expansions. The *height* of S, denoted $ht(S)$, is the maximum of $ht(\lambda_1), ..., ht(\lambda_r)$.

Because of the fact that the classical Puiseux expansions are canonical, a system of rational Puiseux expansions is completely determined by the choice of the parameters $\lambda_1, ..., \lambda_r$. This remark not only provides justification for the previous definition, but also motivates the problem of determining a system of rational Puiseux expansions of small height.

Theorem 2.3. *Let $F(x, y) \in \mathbf{Q}[x, y]$ be of degree n in y, m in x, and let $h = ht(F)$. Assume that $disc_y F \neq 0$, and let y denote the algebraic function defined by $F(x, y) = 0$. Then there exists a positive constant C, with $C < 2500$, and a system S of rational Puiseux expansions of y with*

$$ht(S) < (2^n m n^{\log n} h)^{Cm^2 n^{10}}. \tag{2.3}$$

It seems difficult to prove such an effective result directly from the construction described in [**4**]. From the point of view of a bit-complexity analysis, the system of rational Puiseux expansions in the proof of Theorem

2.3 would be easier to compute than the one described in [**4**]. In fact, using the main result of either [**2**] or [**15**], it can be shown that the system of rational Puiseux expansions described in this paper can be computed in polynomial-time in the bit complexity of F, which does not seem to be the case for the system described in [**4**]. More precisely, because of the potential for large coefficient growth, a theoretical bit-complexity analysis of the construction of rational Puiseux expansions by the method in [**4**] yields an exponential running time in the number of bit operations. On the other hand, the computation described therein would probably be more practical than computing the system described here because of the fact that all of the computations in the algorithm of [**4**] are done in the inertial subfield instead of the extension field generated by the coefficients of the classical Puiseux expansions, as is done here. Therefore, the main purpose of our result is to prove the existence of at least one parameter λ with 'small' height.

3. Algebraic extensions of $\mathbf{Q}((x))$.

The purpose of this section is to present some of the basic theory of algebraic extension of complete fields. No new results are presented here.

Throughout this section the field of Laurent series with rational coefficients

$$\mathbf{Q}((x)) = \left\{ \sum_{k=f}^{\infty} a_k x^k;\ f \in \mathbf{Z}, a_k \in \mathbf{Q} \right\}$$

will be the base field whose extensions we will be interested in studying. Let $v = |\ |_x$ denote the valuation on $\mathbf{Q}(x)$ at $x = 0$, given by $v(0) = 0$ and $v(G/H) = 2^{-f}$, where $G \neq 0$ and $G/H = x^f(G_1/H_1)$ with $G_1(0) \neq 0$, $H_1(0) \neq 0$. Then $\mathbf{Q}((x))$ is the completion of $\mathbf{Q}(x)$ with respect to v, and $v(\sum_{k=f}^{\infty} a_k x^k) = 2^{-f}$, where $a_f \neq 0$. The integer f is the *order* of $\sum_{k=f}^{\infty} a_k x^k$, written as $f = \mathrm{ord}_x \sum_{k=f}^{\infty} a_k x^k$.

Let E denote an algebraic extension of $\mathbf{Q}((x))$ of degree n. Then (for example, see Theorem 7 in Chapter 2 of [**1**]) there is precisely one extension v_1 of v to E, given by

$$v_1(\alpha) = v(\mathcal{N}(\alpha))^{1/n},$$

where $\mathcal{N}$ is the norm from E to $\mathbf{Q}((x))$. The function ord_x is similarly extended to E by

$$\mathrm{ord}_x(\alpha) = (1/n)\mathrm{ord}_x(\mathcal{N}(\alpha)),$$

for $\alpha \in E$. The *ring of integers* of E is the set

$$\mathcal{O} = \{\alpha \in E;\ v_1(\alpha) \leq 1\},$$

and the unique maximal ideal consisting of the nonunits of $\mathcal{O}$ is

$$\mathcal{P} = \{\alpha \in E;\ v_1(\alpha) < 1\}.$$

The field $\mathcal{O}/\mathcal{P}$ is the *residue class field* of E. We remind the reader here of some of the basic properties of the residue class field.

Let $E_1 \subseteq E_2$ be algebraic extensions of $\mathbf{Q}((x))$, then (for example see Theorem 3 in Chapter 3 of [**1**]) the residue class field of E_1 is isomorphic to a subfield of the residue class field of E_2.

If K is an algebraic extension of $\mathbf{Q}$, and $e \geq 1$ is a positive integer, then the residue class field of $K((x^{1/e}))$ is K.

We now turn our attention to the structure of E. More precisely, by Puiseux's theorem, we may assume that E is of the form

$$E = \mathbf{Q}((x))(y_0),$$

where $y_0 = y(x)$ is given in (2.2).

Let K denote the residue class field of E, then $\delta = \delta(E/\mathbf{Q}((x))) = [K : \mathbf{Q}]$ is the *inertia* of E, and $K((x))$ is the *inertial subfield of* E (the inertia is normally denoted by f which we reserve for the order of a Laurent series). $K((x))$ is the maximal subfield of E of the form $L((x))$, with L an algebraic number field. The ring of integers of $K((x))$ is denoted $K[[x]]$. Note that K is the residue class field of $K((x))$.

An algebraic extension E of $\mathbf{Q}((x))$ is called *unramified* if the inertia of E over $\mathbf{Q}((x))$ equals the degree of E over $\mathbf{Q}((x))$, that is, $\delta(E/\mathbf{Q}((x))) = [E : \mathbf{Q}((x))]$. It is easy to see that such extensions are of the form $L((x))$ for some algebraic number field L. Thus, in the notation of the preceding paragraph, we have that $K((x))$ is the *maximal unramified extension of* $\mathbf{Q}((x))$ in E.

Again let E denote an algebraic extension of $\mathbf{Q}((x))$, and let ord_x denote the order function on E defined in Section 2. The set $G_E = \{\mathrm{ord}_x(\alpha);\ \alpha \in E\}$ is called the *value group* of E, and it is easy to see that the index $[G_E : \mathbf{Z}]$ is an integer, called the *ramification index* of L over $\mathbf{Q}((x))$, and denoted $e(E/\mathbf{Q}((x)))$. By Theorem 10 in Chapter 3 of [**1**] we have the relation

$$[E : \mathbf{Q}((x))] = e(E/\mathbf{Q}((x))) \cdot \delta(E/\mathbf{Q}((x))).$$

Thus, E is unramified precisely when $e(E/\mathbf{Q}((x))) = 1$. On the other hand, E is said to be *totally ramified* if $e(E/\mathbf{Q}((x))) = [E : \mathbf{Q}((x))]$. It is evident from the definitions of ramification and inertia that if $\mathbf{Q}((x)) \subseteq E_1 \subseteq E_2$ are algebraic extensions, then

$$e(E_2/\mathbf{Q}((x))) = e(E_2/E_1) \cdot e(E_1/\mathbf{Q}((x))),$$
$$\delta(E_2/\mathbf{Q}((x))) = \delta(E_2/E_1) \cdot \delta(E_1/\mathbf{Q}((x))).$$

Thus, $E = \mathbf{Q}((x))(y_0)$ is a totally ramified extension of its inertial subfield $K((x))$. It follows (for example see Proposition 3.4.3 of [**18**]) that

$$E = K((x))(w),$$

where w is a solution of

$$x^{e_1} = \alpha,$$

where $\alpha \in K((x))$, and $e_1 = e(E/K((x))) = [E : K((x))]$. It will be shown that the integer e_1 is equal to the ramification index e of the series $y_0 = y(x)$, justifying the terminology.

4. Preliminary Results.

In this section we prove some results required for the proofs of the main theorems. The following result is well known, the reader is referred to p. 70 of [**1**].

Lemma 4.1. *Let L be an algebraic number field, and let*

$$\alpha = 1 + a_1x + a_2x^2 + \dots$$

be a nonzero element in $1 + xL[[x]]$. If $e \geq 1$ is a positive integer, then the equation

$$Z^e = \alpha \tag{4.1}$$

has a solution in $L((x))$.

We will assume henceforth that F is an irreducible element in $\mathbf{Q}((x))[y]$, and y_0 will denote a root of F, given by $y_0 = y(x)$ in (2.2). Let $E = \mathbf{Q}((x))(y_0)$, then by the remarks in the previous section,

$$E = K((x))(T),$$

where K is the residue class field of E, and $T^{e_1} \in K((x))$, with $e_1 = [E : K((x))]$.

Lemma 4.2. *Let e denote the ramification index of y_0, then $e = e_1$. Moreover, $E = K((x))(T)$ where T satisfies the relation $T^e = \mu x$ for some $\mu \in K$.*

Proof. Write $T^{e_1} \in K((x))$ as

$$T^{e_1} = \alpha_t x^t + \alpha_{t+1}x^{t+1} + \dots = \alpha_t x^t(1 + P(x)),$$

where $t \in \mathbf{Z}$, $P(x) \in xK[[x]]$, and $\alpha_i \in K$ for $i \geq t$. By Lemma 4.1, it follows that T can be chosen so that $T^{e_1} = \beta x^t$ for some $\beta \in K$.

If $d = gcd(t, e_1)$, then E is a field extension of $K(\beta^{1/d})((x))$ of degree e_1/d for some d-th root of β. Thus, by the remarks in the previous section about $K((x))$ being the maximal unramified extension of $\mathbf{Q}((x))$ in E, it follows that $\beta^{1/d} \in K$, and hence that $T^{e_1/d} \in K((x))$. Thus, $d = 1$ since T is of degree e_1 over $K((x))$, and there are integers r and s such that $1 = rt - e_1 s$. By replacing T by T^r/x^s, it follows that $T^{e_1} = \mu x$, with $\mu = \beta^r$.

To see that $e = e_1$, first notice that $T = (\mu x)^{1/e_1} \in E$, while $E = \mathbf{Q}((x))(y_0)$ contains only elements which are a sum of terms involving the formal variable $x^{1/e}$. Therefore, $e_1 > e$ is impossible, and the result will follow by proving that e divides e_1. Let p be a prime dividing e, and let r be the power of p which properly divides e. By the definition of the ramification index of y_0, there is a positive integer k, not divisible by p, such that the coefficient a_k of $x^{k/e}$ in y_0 is nonzero. Since y_0 can be written in the form $\sum \alpha_k((\mu x)^{1/e_1})^k$, it follows that p^r must divide e_1. Since this holds for any prime dividing e, it follows that e divides e_1. $\square$

Let all notation be as above: $y_0 = y(x)$ is as in (2.2), e is the ramification index of y_0, $f = \mathrm{ord}_x y_0$, $E = \mathbf{Q}((x))(y_0)$, K is the residue class field of E, and T is an element in E such that $E = K((x))(T)$ and $T^e = \mu x$ for some $\mu \in K$. Note that $E = K((T))$, a fact to be used later in this section. For $k \geq f$ define α_k by

$$\alpha_k = a_k(\mu^{-1/e})^k,$$

where $\mu^{1/e}$ is an e-th root of μ such that $\mu^{1/e}x^{1/e} \in E$ ($\mu^{1/e}x^{1/e} = T$, and T is defined by this). The author would like to express his gratitude to Professor Bernard Dwork for suggestions in [**5**] leading to the results of this section.

Lemma 4.3. *Let* $L = \mathbf{Q}(a_f, a_{f+1}, \dots)$ *and* $K_0 = \mathbf{Q}(\mu, \alpha_f, \alpha_{f+1}, \dots)$. *Then* $K = K_0$, $L = K(\mu^{1/e})$, *and* $[K(\mu^{1/e}) : K] = s_0$.

Proof. For brevity let $K_0 = \mathbf{Q}(\mu, \alpha_f, \alpha_{f+1}, \dots)$ and $K_1 = \mathbf{Q}(a_f, a_{f+1}, \dots)$. Since $y_0 \in E = K((x))(T)$, y_0 can be written in the form

$$y_0 = \sum_{k=f_1}^{\infty} \beta_k T^k,$$

where $\beta_k \in K$ for $k \geq f_1$. But, as Lemma 4.2 shows, $T = \mu^{1/e}x^{1/e}$, so that

$$y_0 = \sum_{k=f_1}^{\infty} \beta_k (\mu^{1/e})^k (x^{1/e})^k = \sum_{k=f}^{\infty} a_k (x^{1/e})^k.$$

Therefore $f = f_1$ and $\beta_k = \alpha_k$ for $k \geq f$. This shows that $K_0 \subseteq K$, and moreover that $y_0 \in K_0((x))(T)$, from which it follows that $K((x))(T) = K_0((x))(T)$. Also, $K_0 \subseteq K$ forces $[E : K_0((x))] \geq e$. On the other hand, $T^e \in K_0((x))$ and $E = K_0((x))(T)$, so that $[E : K_0((x))] \leq e$, and hence $[E : K_0((x))] = e$. But now this and $K_0 \subseteq K$ imply that $K((x)) = K_0((x))$, from which it follows that $K = K_0$.

To show that $K_1 = K(\mu^{1/e})$ it will be shown that they are both equal to the residue class field, say K_2, of the field $E_1 = E(x^{1/e})$. First observe that

$E_1 = \mathbf{Q}((x^{1/e}))(y_0)$, so that $a_k \in K_2$ for all $k \geq f$, and hence $K_1 \subseteq K_2$. Now $y_0 \in K_1(x^{1/e})$, therefore $E_1 = \mathbf{Q}((x^{1/e}))(y_0) \subseteq K_1(x^{1/e})$. By the remarks in the previous section it follows that $K_2 \subseteq K_1$, and hence K_1 is the residue class field of $E(x^{1/e})$. Now observe that $T = \mu^{1/e}x^{1/e} \in K(\mu^{1/e})((x))(x^{1/e})$, so that $E = K((x))(T) \subseteq K(\mu^{1/e})((x))(x^{1/e})$, and hence $E_1 \subseteq K(\mu^{1/e})((x^{1/e}))$. Since the residue class field of $K(\mu^{1/e})((x^{1/e}))$ is $K(\mu^{1/e})$, we have that $K_2 \subseteq K(\mu^{1/e})$. But $T = \mu^{1/e}x^{1/e} \in E$, so that $\mu^{1/e} \in E_1$, and hence $K(\mu^{1/e})(x^{1/e}) \subseteq E_1$. It follows that $K(\mu^{1/e}) \subseteq K_2$, and so $K(\mu^{1/e})$ is the residue class field of E_1, which is the required result.

By part 2 of Proposition 2.1, we have that y_0 is of degree es/s_0 over $\mathbf{Q}((x))$, and so $[E : \mathbf{Q}((x))] = es/s_0$. By Lemma 4.2, $[E : K((x))] = e$, and so $[K((x)) : \mathbf{Q}((x))] = [K : \mathbf{Q}] = s/s_0$. The last statement now follows from the fact that $[K(\mu^{1/e}) : \mathbf{Q}] = s$. $\square$

5. Proof of Theorem 2.1 and 2.2.

Proof of Theorem 2.1. Assume that F factors into irreducibles in $\mathbf{Q}((x))[y]$ as $F = F_1F_2\cdots F_r$, and that y_i is an algebraic function defined by $F_i(x, y_i) = 0$ for $i = 1, \dots r$. If U_i is a system of rational Puiseux expansions of y_i for $i = 1, \dots r$, then it follows that $U = \bigcup_{i=1}^r U_i$ is a system of rational Puiseux expansions of the algebraic function y. Thus, it is sufficient to consider the case that F is an irreducible element in $\mathbf{Q}((x))[y]$. By Proposition 2.1, the algebraic function y defined by $F(x, y) = 0$ has only one conjugacy class of Puiseux expansions at $x = 0$. Let $y_0 = y(x)$, given in (2.2), be one these expansions, and define E to be the algebraic extension $E = \mathbf{Q}((x))(y_0)$ of $\mathbf{Q}((x))$. By Lemma 4.2 there is an element $T \in \mathbf{Q}((x))(y_0)$ such that $E = K((x))(T)$, where K is the residue class field of E, and such that $T^e = \mu x$ for some $\mu \in K$. It follows that there are $\alpha_k \in K$ for $k \geq f$ such that $y_0 = \sum_{k=f}^{\infty} \alpha_k T^k$, and there is an e-th root, $\mu^{1/e}$ of μ, and corresponding e-th root, $\lambda^{1/e}$ of λ, such that $\alpha_k = a_k(\lambda^{1/e})^k$, for $k \geq f$, where $\lambda = \mu^{-1}$. Define

$$(x(T), y(T)) = \left(\lambda T^e, \sum_{k=f}^{\infty} \alpha_k T^k\right).$$

Then $(x(T), y(T))$ is a generalized parametrization of the branch containing y_0. Let $\mathcal{T}$ denote the orbit of $(x(T), y(T))$ under the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, then $\mathcal{T}$ is a set of generalized parametrizations of the branches of Puiseux expansions in the conjugacy class of y_0, and is invariant under the action of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. It suffices to prove that no two elements of $\mathcal{T}$ are equivalent.

If two elements of $\mathcal{T}$ are equivalent, then there are embeddings $\sigma, \gamma \in \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ and an integer i, with $0 \leq i \leq e-1$, such that

$$\sigma(\alpha_k)(\sigma(\lambda^{-1})^{1/e})^k = \gamma(\alpha_k)(\gamma(\lambda^{-1})^{1/e})^k \zeta_e^{ik}$$

for all $k \geq f$, where ζ_e is an e-th root of unity.

Regard σ and γ as embeddings of $K((x))$ into an algebraic closure in the obvious way. Recall that T is defined by $T = (\lambda^{-1})^{1/e}x^{1/e}$. Since $e = [E : K((x))]$, for each of σ and γ there are e distinct and unique embeddings of E into this algebraic closure, defined by $\overline{\sigma_j}(T) = \sigma(\lambda^{-1})^{1/e}x^{1/e}\zeta_e^j$ and $\overline{\gamma_j}(T) = \gamma(\lambda^{-1})^{1/e}x^{1/e}\zeta_e^j$ for $j = 0, \dots, e-1$. Consider the action of $\overline{\sigma_{e-i}}$ on y_0, where i is as above. Then

$$\overline{\sigma_{e-i}}(y_0) = \sum_{k=f}^{\infty} \sigma(\alpha_k)\overline{\sigma_{e-i}}(T)^k = \sum_{k=f}^{\infty} \sigma(\alpha_k)(\sigma(\lambda^{-1})^{1/e})^k \zeta_e^{-ik} x^{k/e}$$

while

$$\overline{\gamma_0}(y_0) = \sum_{k=f}^{\infty} \gamma(\alpha_k)\overline{\gamma_0}(T)^k = \sum_{k=f}^{\infty} \gamma(\alpha_k)(\gamma(\lambda^{-1})^{1/e})^k x^{k/e}.$$

Since y_0 is a primitive element of E, it follows that $\overline{\sigma_{e-i}} = \overline{\gamma_0}$. But if σ and γ are not equal, then the two sets of embeddings $\{\overline{\sigma_j};\ j = 0, \dots, e-1\}$ and $\{\overline{\gamma_j};\ j = 0, \dots, e-1\}$ are disjoint. Therefore, $\sigma = \gamma$, and hence $i = 0$, showing that no two elements of $\mathcal{T}$ are equivalent. □

Proof of Theorem 2.2. In order to prove the first part of the theorem, it is enough to show that $\mathbf{Q}(\lambda, \alpha_f, \dots)$ is the residue class field K of $E = \mathbf{Q}((x))(y_0)$. Since $y_0 \in \mathbf{Q}(\lambda, \alpha_f, \dots)((x))(Z)$, it follows that $E \subset \mathbf{Q}(\lambda, \alpha_f, \dots)((x))(Z)$. Recall that E is of degree e over $K((x))$, where e is the ramification index of y_0. As $\mathbf{Q}(\lambda, \alpha_f, \dots)((x))(Z)$ is of degree e over $\mathbf{Q}(\lambda, \alpha_f, \dots)((x))$, it follows that $K \subset \mathbf{Q}(\lambda, \alpha_f, \dots)$. Recall that the order of the conjugacy class $C(y_0)$ of y_0 is $e(s/s_0)$, and so because $(\lambda Z^e, \sum_{k=f}^{\infty} \alpha_k Z^k)$ is a rational Puiseux expansion parametrizing $B(y_0)$, its orbit under $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ has s/s_0 elements. Therefore, $[\mathbf{Q}(\lambda, \alpha_f, \dots) : \mathbf{Q}] = s/s_0 = [K : \mathbf{Q}]$ and hence $K = \mathbf{Q}(\lambda, \alpha_f, \dots)$.

The proof of the last statement in Theorem 2.2 is somewhat more involved. Let $\overline{y_0}$ denote the *singular part* of y_0, defined as the initial partial sum

$$\overline{y_0} = \sum_{k=f}^{l} a_k (x^{1/e})^k \tag{5.1}$$

of y_0, where l is minimal with the property that $\overline{y_0}$ is not equal to an initial partial sum of any other Puiseux expansion of the algebraic function y at $x = 0$ defined by $F(x, y) = 0$. We will need an upper bound for l. Define $H(t, z) = t^r F(t^e, zt^f)$, where r is chosen so that H is a polynomial not divisible by t. Then the series

$$z(t) = \sum_{k=0}^{\infty} a_{k+f} t^k$$

is a root of $H(t,z)$. Moreover, it is easy to see that $\deg_z H = n$ and $\deg_t H \leq 2mn$. Therefore, by Lemma 3 of [**3**], it follows that

$$l \leq 2mn(2n-1), \tag{5.2}$$

and that

$$\mathbf{Q}(a_f, a_{f+1}, \cdots) = \mathbf{Q}(a_f, a_{f+1}, \cdots, a_l).$$

It is now enough to prove that

$$\mathbf{Q}((x))(y_0) = \mathbf{Q}((x))(\overline{y_0}),$$

for the result follows immediately from the fact, which is proved in a similar manner as Lemma 4.3, that the residue class field of $\mathbf{Q}((x))(\overline{y_0})$ is $\mathbf{Q}(\alpha_f, \alpha_{f+1}, \cdots, \alpha_l)$.

Let $K' = \mathbf{Q}(\alpha_f, \alpha_{f+1}, \cdots, \alpha_l)$, then

$$\mathbf{Q}((x))(\overline{y_0}) \subseteq K'((x))(T) \subseteq K((x))(T) = \mathbf{Q}((x))(y_0),$$

so it is enough to prove that $\overline{y_0}$ is of degree at least es/s_0 over $\mathbf{Q}((x))$.

Let $G(y)$ be the defining polynomial of $\overline{y_0}$ over $\mathbf{Q}((x))$. For $j = 0, 1, \ldots, e-1$ and an embedding σ of L into $\overline{\mathbf{Q}}$, define the series

$$\overline{y_0}(x, \sigma, j) = \sum_{k=f}^{l} \sigma(a_k)(\zeta_e^j x^{1/e})^k.$$

Then it is easy to see that each $\overline{y_0}(x, \sigma, j)$ is a root of $G(y) = 0$. By using the exact argument given in the proof of Lemma 2 of [**16**], applied to $\overline{y_0}$, it follows that the number of distinct $\overline{y_0}(x, \sigma, j)$ is precisely es/s_0, and the result follows. □

6. Proof of the Theorem 2.3.

In this section we prove Theorem 2.3 by constructing an explicit system of rational Puiseux expansions. The proof of the bound in (2.3) is enabled by a recent effective result on algebraic functions proved in [**6**]. The author would like to greatly thank Professor Hendrik Lenstra for his suggestions in [**11**] on this work.

Let $F(x,y) \in \mathbf{Q}[x,y]$ be as in (2.1), with $n = \deg_y F$, $m = \deg_x F$, and $h = ht(F)$, and let

$$y(x) = \sum_{i=1}^{\infty} A_i x^{f_i/e_i} \tag{6.1}$$

be a Puiseux expansion at $x = 0$ of the algebraic function y defined by $F(x,y) = 0$, where for each $i \geq 1$, $A_i \neq 0$ and the greatest common divisor of f_i and e_i is 1.

In order to construct a system of rational Puiseux expansions of F, it suffices to consider the case that F is irreducible in $\mathbf{Q}((x))[y]$, for in general one would construct one system for each irreducible factor of F in $\mathbf{Q}((x))[y]$ and then take the union of these as a system of rational Puiseux expansions for F. We also make the assumption that F has integer coefficients, for this causes no loss of generality since one may replace F by $\mathrm{denom}(F) \cdot F$.

Let

$$y_l(x) = \sum_{i=1}^{l} A_i x^{f_i/e_i} \tag{6.2}$$

denote the singular part of $y(x)$, as defined in the proof of Theorem 2.2. Note that by (5.2),

$$l \leq 4mn^2,$$

and that $\mathbf{Q}(A_1, A_2, \dots) = \mathbf{Q}(A_1, A_2, \dots, A_l)$. Let $e = lcm(e_1, \dots, e_l)$, which by Theorem 6.1 of [**9**], is the ramification index of $y(x)$. For $i \geq 1$, let $c_i = e/e_i$, then from the definition of e and the fact that $gcd(e_i, f_i) = 1$, it follows that $gcd(c_1 f_1, \dots, c_l f_l, e) = 1$. Hence, there exist integers $n_1, \dots, n_l$ such that $0 \leq n_i \leq e - 1$ for each $i \geq 1$ and

$$\sum_{i=1}^{l} n_i c_i f_i \equiv -1 \pmod{e}.$$

Define algebraic numbers μ and λ by

$$\begin{aligned} \mu &= \prod_{i=1}^{l} A_i^{n_i}, \\ \lambda &= \mu^e, \end{aligned} \tag{6.3}$$

and for $i \geq 1$ put

$$\beta_i = A_i \mu^{c_i f_i}.$$

Let $K = \mathbf{Q}(\lambda, \beta_1, \beta_2, \dots)$, and $r = [K : \mathbf{Q}]$. Let $\sigma_1, \dots, \sigma_r$ denote the embeddings of K into $\overline{\mathbf{Q}}$. For $1 \leq j \leq r$ put

$$(x_{\sigma_j}(T), y_{\sigma_j}(T)) = \left(\sigma_j(\lambda) T^e, \sum_{i=1}^{\infty} \sigma_j(\beta_i T^{c_i f_i}) \right).$$

Lemma 6.1. *The set $S = \left\{ (x_{\sigma_j}(T), y_{\sigma_j}(T));\ \ j = 1, \dots, r \right\}$ is a system of rational Puiseux expansions of the algebraic function y defined by $F(x, y) = 0$.*

Proof. It is clear that S consists of generalized parametrizations of the branches of the algebraic function y at $x = 0$, and that to each branch of y

there is at least one parametrization of that branch in S. It suffices to show that no two elements of S are equivalent, in other words, to show that no two elements of S represent the same branch.

For a fixed embedding σ_j of K into $\overline{\mathbf{Q}}$ let $E(\sigma_j)$ denote the set of σ_k such that $(x_{\sigma_j}(T), y_{\sigma_j}(T))$ and $(x_{\sigma_k}(T), y_{\sigma_k}(T))$ are equivalent. By an argument identical to that given in the proof of Lemma 2 of [**16**], it can be seen that there is an extension σ_j' of σ_j to $K(\mu)$ such that

$$E(\sigma_j) = \left\{\sigma_j' \circ \theta;\ \ \theta \in E(\mathbf{1}_K)\right\}.$$

Thus, it suffices to prove that $E(\mathbf{1}_K) = \{\mathbf{1}_K\}$. Assume that without loss of generality that $\sigma_1 \in E(\mathbf{1}_K)$, and let σ_1' denote an extension of σ_1 to $K(\mu)$. The result will follow by showing that σ_1' is the identity map on K.

We are assuming that $(\lambda T^e, \sum_{i=1}^{\infty} \beta_i T^{c_i f_i})$ and $(\sigma_1(\lambda) T^e, \sum_{i=1}^{\infty} \sigma_1(\beta_i) T^{c_i f_i})$ represent the same branch. This implies that there is an integer j with $0 \le j \le e-1$ such that for all $i \ge 1$

$$\sigma_1(\beta_i)\sigma_1'(\mu)^{-c_i f_i} = \beta_i \mu^{-c_i f_i} \zeta_e^{j c_i f_i}.$$

From the definition of β_i we deduce that for all $i \ge 1$

$$\sigma_1'(A_i) = A_i \zeta_e^{j c_i f_i}.$$

It follows from the choice of the integers $n_1, \dots, n_l$ that

$$\sigma_1'(\mu) = \prod_{i=1}^{l} \sigma_1'(A_i)^{n_i} = \prod_{i=1}^{l} A_i^{n_i} \zeta_e^{j n_i c_i f_i} = \mu \zeta_e^{-j},$$

and hence also that $\sigma_1(\lambda) = \lambda$. Moreover, for $i \ge 1$,

$$\sigma_1'(\beta_i) = \sigma_1'(A_i)\sigma_1'(\mu)^{c_i f_i} = A_i \zeta_e^{j c_i f_i} \mu^{c_i f_i} \zeta_e^{-j c_i f_i} = \beta_i.$$

Recall that $K = \mathbf{Q}(\lambda, \beta_1, \beta_2, \dots)$, and so σ_1' is the identity on K, which is what we needed to prove. $\square$

To complete the proof of Theorem 2.3 we need to estimate the height of λ, where λ is given in (6.3). This is accomplished by proving a series of preliminary results.

We define quantities A and B which will be used frequently in what follows. Let

$$A = \max_{1 \le i \le l} \|A_i\|,$$

where for an algebraic number α with algebraic conjugates $\alpha = \alpha^{(1)}, \dots, \alpha^{(q)}$, the *house* of α is

$$\|\alpha\| = \max_{1 \le j \le q} |\alpha^{(j)}|.$$

Also, let

$$B = \max_{1 \le i \le l} \operatorname{denom}(A_i),$$

where for an algebraic number α, $\operatorname{denom}(\alpha)$ is the *denominator* of α, and is defined as the least positive integer v such that $v\alpha$ is an algebraic integer.

Lemma 6.2. *There is an algebraic integer α such that $\mathbf{Q}(\alpha) = \mathbf{Q}(A_1, A_2, \dots, A_l)$ and*

$$ht(\alpha) \le (8mn^4AB)^n.$$

Proof. Let $B_1, \dots, B_l$ be positive integers no larger than B such that for each $i = 1, \dots, l$, B_iA_i is an algebraic integer. Since $\deg(A_i) \le n$ for each $1 \le i \le l$, the proof on p. 139 of [14] shows that there are positive integers $t_1, \dots, t_{l-1}$, with $1 \le t_i \le n^2$ for each i, such that

$$\alpha = B_1A_1 + t_1B_2A_2 + \cdots + t_{l-1}B_lA_l$$

is a primitive element of $\mathbf{Q}(A_1, \dots, A_l)$. Moreover, α defined this way is an algebraic integer. We have then that

$$\|\alpha\| = \|B_1A_1 + t_1B_2A_2 + \cdots + t_{l-1}B_lA_l\| \le ln^2BA \le 4mn^4AB.$$

But since α is an algebraic integer, Lemma A.2 of [**13**] shows that

$$ht(\alpha) \le (2\|\alpha\|)^n,$$

from which the result follows. □

Lemma 6.3. *Let α be as in the previous result. For each i with $1 \le i \le l$ there exists a polynomial $P_i(x) \in \mathbf{Q}[x]$ of degree at most $n-1$ such that $A_i = P_i(\alpha)$ and*

$$ht(P_i) \le (8mn^4AB)^{2n^2}.$$

Proof. Fix i in the range $1 \le i \le l$. Define

$$A_i^* = \operatorname{denom}(A_i) \cdot A_i.$$

Then because A_i^* is an algebraic integer,

$$ht(P_{A_i^*}) \le ht\left(\prod_{j=1}^{n}(x + \operatorname{denom}(A_i) \cdot \|A_i\|)\right) \le (2BA)^n.$$

Also, because $A_i^* \in \mathbf{Q}(\alpha)$, we have that $A_i^* = p_i(\alpha)$, where $p_i(x) \in \mathbf{Q}[x]$ is some polynomial of degree no greater than $n-1$.

Let D denote a positive integer and let β denote an algebraic integer. For a polynomial $P \in (1/D)\mathbf{Z}[\beta][x]$, we define its *height* to be the largest of the absolute values of the rational integers appearing in $D \cdot Q$, and denote this height by $P_{\max}$, where it is understood that this height is dependent upon the fixed values D and β.

In the proof of Proposition 3.11 of [**10**], it was shown that a monic factor of degree m of a polynomial Q in $(1/D)\mathbf{Z}[\beta][x]$ has height bounded by

$$Q_{\max}\left(2(\deg(Q)+1)(c^3)(c-1)^{(c-1)}\left(\binom{2m}{m}\right)\right)^{1/2} ht(P_\beta)^{2(c-1)},$$

where c is the degree of β over $\mathbf{Q}$.

Using this together with the result of Lemma 6.2, it follows that $x-p_i(\alpha)$ is a factor of $P_{A_i^*}(x)$ in $\mathbf{Q}[\alpha][x]$ such that

$$\begin{aligned} ht(p_i(x)) &\leq ht(P_{A_i^*})[2(n+1)^2n^3n^n\cdot 2]^{1/2}(n+1)^{n-1}ht(P_\alpha)^{2n-2} \\ &\leq (2BA)^n(2n)^{\frac{n+5}{2}}(n+1)^{n-1}(8mn^4AB)^{2n^2-2n}\cdot B \\ &\leq (8mn^4AB)^{2n^2-n}. \end{aligned}$$

The result now follows by defining $P_i(x)=(1/\mathrm{denom}(A_i))p_i(x)$. $\square$

Lemma 6.4. *There is a polynomial $P\in\mathbf{Q}[x]$ with $\deg P\leq n-1$ such that $\lambda=P(\alpha)$ and*

$$ht(P)\leq (8mn^4AB)^{8mn^6}.$$

Proof. From the definition of λ we have that

$$\lambda=\prod_{i=1}^{l}P_i(\alpha)^{n_ie},$$

where the P_i are as in Lemma 6.3. Let $P^*(x)=\prod_{i=1}^{l}P_i(x)^{n_ie}$, then $\lambda=P^*(\alpha)$ and $\deg P^*\leq 4mn^3(n-1)^2$. Since α is of degree at most n, then reducing the higher powers of α appearing in $P^*(\alpha)$ results in a representation $\lambda=P(\alpha)$ for some polynomial $P\in\mathbf{Q}[x]$ of degree at most $n-1$. Using the estimates obtained so far, we can deduce an upper bound for $ht(P)$. For an element $\beta\in\mathbf{Q}[\alpha]$, with $\beta=p(\alpha)$ and $p\in\mathbf{Q}[x]$ of degree less than n, let $(\beta)_{\max}=ht(p)$. By an inductive argument it is easy to see that $(\alpha^t)_{\max}\leq ht(\alpha)(ht(\alpha)+1)^{t-n}$ for all $t\geq n$. Therefore, since each $n_i\leq n-1$, we have by Lemma 6.3 that

$$\begin{aligned} ht(P) &\leq ht(P^*)\sum_{i=0}^{\deg P^*}(\alpha^i)_{\max} \\ &\leq (ht(P^*))(\deg P^*+1)\cdot \max\nolimits_i\left((\alpha^i)_{\max}\right) \\ &\leq (4mn^5)\left(\max_{1\leq i\leq l} ht(P_i)\right)^{4mn^3(n-1)} ht((1+\cdots+x^{n-1})^{4mn^5})ht(P_\alpha)^{4mn^5} \\ &\leq (4mn^5)(8mn^4AB)^{8mn^5(n-1)}n^{4mn^5}ht(P_\alpha)^{4mn^5} \\ &\leq (8mn^4AB)^{8mn^6}. \end{aligned}$$

Completion of the proof of Theorem 2.3. We begin by computing an estimate for $\|\lambda\|$. Let $G = \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. With P as in Lemma 6.4 we have

$$\begin{aligned}\|\lambda\| &= \max_{\sigma\in G}\|\sigma(P(\alpha))\| \\ &= \max_{\sigma\in G}\|P(\sigma(\alpha))\| \\ &\leq ht(P)\cdot n\cdot\|\alpha\|^n.\end{aligned}$$

Let

$$\lambda^* = \prod_{i=1}^{l}(B_iA_i)^{n_ie},$$

then λ^* is an algebraic integer and

$$\|\lambda^*\| = \prod_{i=1}^{l}B_i^{n_ie}\|\lambda\|.$$

By Lemma A.2 of [**13**],

$$ht(\lambda^*) \leq (2\|\lambda^*\|)^n.$$

Therefore,

$$ht(\lambda^*) \leq 2^nB^{4mn^5}ht(P)^nn^n\|\alpha\|^n.$$

We now remark that if β is an algebraic number and $\nu = \mathrm{denom}(\beta)$, then β is a root of the polynomial $P_{\nu\beta}(\nu x)$, which has integer coefficients, leading coefficient equal to $\nu^{\deg(\beta)}$, and the same degree as P_β. It follows that if $lc(P_\beta)$ represents the leading coefficient of P_β, then $lc(P_\beta) \leq \mathrm{denom}(\beta)^{\deg(\beta)}$. Therefore,

$$\begin{aligned}ht(P_\lambda) &\leq lc(P_\lambda)\cdot ht(P_{\lambda^*}) \\ &\leq (\mathrm{denom}(\lambda))^n\cdot ht(P_{\lambda^*}) \\ &\leq B^{4mn^5}2^nB^{4mn^5}ht(P)^nn^n\|\alpha\|^n \\ &\leq B^{8mn^5}(2n)^n(8mn^4AB)^{n^2}(8mn^4AB)^{8mn^7} \\ &\leq (8mn^4AB)^{9mn^7}.\end{aligned}$$

Regard $y(x)$ as in (2.2), and let

$$y_{l_1} = \sum_{k=f}^{l_1}a_k(x^{1/e})^k$$

denote the singular part of $y(x)$. Then by (5.2), $l_1 \leq 4mn^2$. By Lemma 2.a of [**12**], it follows that

$$\|a_k\| \leq 2(h+1)[(h(2mn+1)(n+1))^{2n}]^{2mn^2+k},$$

for $k \geq 0$. Since $A = \max_{f\leq k\leq l_1}\|a_k\|$, it follows that

$$A \leq 2(h+1)[(h(2mn+1)(n+1))]^{12mn^3}.$$

Since $y(x)$ satisfies $F(x, y(x)) = 0$, it follows that for $b_k = a_{k+f}$, the series $z(t) = \sum_{k=0}^{\infty} b_k t^k$, satisfies $H(t, z(t)) = 0$, where $H(t, z) = t^r F(t^e, zt^{-f})$, and r is chosen to be the positive integer so that $H(t, z)$ is a polynomial, and $H(0, z) \neq 0$. Note that $f \leq m$ and $r \leq 2mn$, from which it follows that $\deg_t H \leq 2mn$, $\deg_z H \leq n$ and $ht(H) = ht(F)$. Applying the main result of [**6**] to the series $z(t)$, we obtain

$$\operatorname{denom}(b_k) \leq h(4.8n^{4+2.74\log n} 2^{2n} h^2 (2mn+1)^2)^{n(k+2nm)},$$

for all $k \geq 0$. We recall that $B = \max_{f \leq k \leq l_1} \operatorname{denom}(a_k)$, and so from the definition of the b_k and the bound $l_1 \leq 4mn^2$, it follows that

$$B \leq (4.8n^{4+2.74\log n} 2^{2n} h^2 (2mn+1)^2)^{6mn^3},$$

from which Theorem 2.3 follows.

References

[1] E. Artin, *Algebraic Numbers and Algebraic Functions*, Gordon and Breach, New York, 1967.

[2] A.L. Chistov, *Polynomial complexity of the Newton-Puiseux algorithm*, Lecture Notes in Computer Science, **233** (1986), 247-255.

[3] J. Coates, *Construction of rational functions on a curve*, Proc. Camb. Phil. Soc., **68** (1970), 105-123.

[4] D. Duval, *Rational Puiseux expansions*, Compositio Mathematica, **70** (1989), 119-154.

[5] B.M. Dwork, Personal communication, 1992.

[6] B.M. Dwork and A.J. van der Poorten, *The Eisenstein constant*, Duke Math. J., **65**(**1**) (1992), 23-43.

[7] M. Eichler, *Introduction to the Theory of Algebraic Numbers and Functions*, Academic Press, London, 1966.

[8] D.L. Hilliker and E.G. Straus, *Determination of bounds for the solutions to those binary diophantine equations that satisfy the hypotheses of Runge's theorem*, Trans. Amer. Math. Soc., **280** (1983), 637-657.

[9] H.T. Kung and J.F. Traub, *All algebraic functions can be computed fast*, J. Assoc. Comput. Mach., **25** (1978), 246-260.

[10] A. Lenstra, *Factoring polynomials over algebraic number fields*, Proc. EuroCal. 1983, Lecture Notes in Computer Science, **162** (1983), 245-254.

[11] H.W. Lenstra, Jr., Personal communication, 1994.

[12] W.M. Schmidt, *Eisenstein's theorem on power series expansions of algebraic functions*, Acta Arith., **56** (1990), 161-179.

[13] T.N. Shorey and R. Tijdeman, *Exponential Diophantine Equations*, Cambridge University Press, Cambridge, 1986.

[14] B.L. van der Waerden, *Modern Algebra*, Frederick Ungar, seventh edition, New York, 1970.

[15] P.G. Walsh, *A polynomial-time complexity bound for the computation of the singular part of an algebraic function*, Submitted to Mathematics of Computation, 1994.

[16] ———, *Irreducibility testing over local fields*, Submitted to Mathematics of Computation, 1994.

[17] ———, *The computation of Puiseux expansions and a quantitative version of Runge's theorem on diophantine equations*, Doctoral Thesis, University of Waterloo, 1994.

[18] E. Weiss, *Algebraic Number Theory*, McGraw-Hill, New York, 1963.

Received August 13, 1997 and revised May 20, 1998.

University of Ottawa
Ottawa, Ontario K1N-6N5
Canada
E-mail address: gwalsh@mathstat.uottawa.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 188, No. 2, 1999

A NEW CLASS OF "BOUNDARY REGULAR" MICRODIFFERENTIAL SYSTEMS

Giuseppe Zampieri

We give a new criterion for the propagation up to the boundary of the analytic singularities of the solutions of microdifferential systems. The class of systems we are able to treat is larger than in D'Ancona-Tose-Zampieri, 1990; namely the condition of transversal ellipticity is here replaced by the non-microcharacteristicity only for the conormal to the boundary. The method also is far different. It is perhaps the most effective application of the theory of the second microlocalization at the boundary by Uchida-Zampieri, 1990.

The microlocal theory of boundary value problems originated from the works by Kataoka and Schapira in the early 80's. In this frame the propagation of the singularities is now almost completely understood. Among other contributions we quote: Schapira, 1986, Kataoka, 1980, Schapira-Zampieri, 1987. This new contribution covers one of the few problems not yet explained at least in the case of transversal bicharacteristics.

Let M be a real analytic manifold, X a complexification of M, S a real analytic hypersurface of M, $M^\pm$ the two open components of $M \setminus S$ (in a neighborhood of a point $x \in S$). Let $T^*X \overset{\pi}{\to} X$ be the cotangent bundle to X endowed with the canonical 2-form $\sigma = \sigma^{\mathbb{R}} + \sqrt{-1}\sigma^{\mathbb{I}}$, and $T^*_M X \overset{\pi_M}{\to} M$ the conormal bundle to M in X. (Sometimes, if no confusion may arise, we write π instead of π_M.) The latter is $\mathbb{R}$-Lagrangian (i.e. $\sigma^{\mathbb{R}}|_{TT^*_M X} = 0$) and $\mathbb{I}$-symplectic (i.e. $\sigma^{\mathbb{I}}|_{TT^*_M X}$ is non-degenerate). In particular if $H = H^{\mathbb{R}} + \sqrt{-1}H^{\mathbb{I}}$ is the Hamiltonian isomorphism, then we have three identifications:

$$H : T^*T^*X \to TT^*X$$
$$H^{\mathbb{R}} : T^*T^*X \to TT^*X,$$
$$H^{\mathbb{I}} : T^*T^*_M X \to TT^*_M X.$$

We shall deal with the sheaves of Sato's microfunctions $\mathcal{C}_{M|X}$, $\mathcal{C}_{S|X}$ and the complexes of microfunctions at the boundary $\mathcal{C}_{M^\pm|X}$. Let V be a smooth involutive submanifold of $\dot{T}^*_M X (:\overset{\text{def.}}{=} T^*_M X \setminus T^*_X X)$, and $\tilde{V}$ the $\mathbb{R}$-Lagrangian submanifold obtained as the union of the complexifications of the bicharacteristic leaves of V. Assume there are real analytic functions r and s on $T^*_M X$ such that

$$s|_V = 0,\ r|_{S\times_M T^*_M X} = 0, \quad \text{and } \{s, r\} \equiv 1. \tag{1}$$

Let $\tilde{W}$ be the union of the integral leaves of $H^{\mathbb{R}}_{\Re e r^{\mathbb{C}}}$ issued from $\tilde{V} \cap \{\Re e r^{\mathbb{C}} = 0\}$; this also is an $\mathbb{R}$-Lagrangian submanifold. Let $\mathcal{M}$ be a coherent $\mathcal{E}_X$-module (i.e. a microdifferential system) in a neighborhood of a point $p \in S \times_M V$ and denote by $\mathrm{char}(\mathcal{M})$ the characteristic variety of $\mathcal{M}$. We note that, since $\mathcal{C}_{M^\pm|X}|_{T^*_M X}$ are concentrated in degree 0, they are endowed in a natural way with a structure of $\mathcal{E}_X$-modules. Let $x = \pi(p)$, recall the identification $\pi^*_M (= {}^t\pi'_M) : T^*_x M \to T^*_p T^*_M X$, and take $\theta \in T^*_S M$.

Theorem 1. *Assume*

$$\pm H^{\mathbb{I}}(\pi^*_M \theta) \notin C_p(\mathrm{char}(\mathcal{M}), \tilde{V}), \tag{2}$$

$$\pm H^{\mathbb{I}}(\pi^*_M \theta) \notin C_p(\mathrm{char}(\mathcal{M}), \tilde{W}), \tag{3}$$

where $C(\cdot,\cdot)$ is the Withney normal cone (cf. [**K-S 2**]*). Then*

$$\Gamma_{\pi^{-1}(S)} \mathrm{Hom}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M^\pm|X})_p = 0. \tag{4}$$

Observe now that we have an identification $T^*_x M \hookrightarrow T^*_x X \hookrightarrow T^*_p T^* X$ where the first embedding is obtained by means of the complex structure of X and the second by means of π^*. Let $V^{\mathbb{C}}$ be the complexification of V in T^*X. As an application of Theorem 1 we get the boundary version of the microlocal Holmgren's Theorem by Bony [**B**]:

Example 1. Assume

$$\mathbb{C}H(\pi^*\theta) \cap C_p(\mathrm{char}(\mathcal{M}), V^{\mathbb{C}}) = \{0\}.$$

(That is assume the embedding $S \hookrightarrow M$ be *non-microcharacteristic* for μ.) Then (4) follows. To see how this follows from Theorem 1, one just needs to remark that the assumption of Example 1 obviously implies (2) and (3).

Example 2. We take coordinates $z \in X$, $x \in M$, $(z;\zeta) \in T^*X$, $(x; \sqrt{-1}\eta) \in T^*_M X$, $z = x + \sqrt{-1}y$, $\zeta = \xi + \sqrt{-1}\eta$, write $z = (z_1, z', z'')$, $\zeta = (\zeta_1, \zeta', \zeta'')$, and assume

$$S = \{x \in M : x_1 = 0\},\ V = \{\eta_1 = 0, \eta' = 0\},\ p = (0; \sqrt{-1}\mathrm{d}x_n).$$

Then (2) is equivalent to

$$|\eta_1| \leq c[|\xi_1| + |\zeta'| + |\xi''| + |y''|] \quad \forall (z;\zeta) \in \mathrm{char}\mathcal{M}, \tag{5}$$

and (3) is equivalent to

$$|\eta_1| \leq c[|x_1| + |\zeta'| + |\xi''| + |y''|] \quad (z;\zeta) \in \mathrm{char}\mathcal{M}. \tag{6}$$

Let us consider the case $\mathcal{M} = \frac{\mathcal{E}_X}{\mathcal{E}_X P}$ where $P = P(x, D)$ is a differential operator with principal symbol $\sigma(P) = \zeta_1^2 + a(z'', \zeta') + b(z'', \zeta', \zeta'')$ with a, b

real on $T^*_M X$ homogeneous of order 2, and with $b|_{T^*_M X} \leq 0$. For S and V defined a above, (2) and (3) hold. In fact

$$\begin{cases} \Re\mathrm{e}\sigma(P) \leq \xi_1^2 - \eta_1^2 + |a(z'', \zeta')| \leq \xi_1^2 - \eta_1^2 + c|\zeta'|^2 & \text{if } \xi'' = 0,\ y'' = 0 \\ \Im\mathrm{m}\sigma(P) = 2\xi_1\eta_1 & \text{if } \xi' = 0,\ \xi'' = 0,\ y'' = 0. \end{cases}$$

Thus if $\sigma(P) = 0$, $\xi' = \xi'' = y'' = 0$ then either $\eta_1 = 0$ or $\xi_1 = 0$ whence $|\eta_1| \leq c|\eta'|$. By an easy variant of the local Bochner's tube theorem this implies

$$|\eta_1| \leq c[|\zeta'| + |\xi''| + |y''|] \quad \text{if } \sigma(P) = 0.$$

Thus for instance in $\mathbb{R}^4$ and with S defined by $x_1 = 0$, the operator

$$P_1 = D_1^2 \pm D_2^2 + D_3^2 + x_3^2 D_4^2,$$

verifies (4) at $p = (0; \pm\sqrt{-1}\mathrm{d}x_4)$,

$$P_2 = D_1^2 \pm x_3^m D_2^2 + x_3^2 D_3^2 + (x_3^2 + x_4^2) D_4^2,$$

at any $p = (0; \sqrt{-1}\eta)$ with $\eta_1 = 0$, $\eta_2 = 0$, and finally

$$P_3 = D_1^2 + x_3^2 D_2^2 + x_3^2 D_3^2 + (x_2^2 + x_3^2 + x_4^2) D_4^2,$$

at any $p = (0; \sqrt{-1}\eta)$ with $\eta_1 = 0$. (The V's we may use here are $V = \{\eta_1 = 0, \eta_2 = 0\}$ as for P_1, P_2 and $V = \{\eta_1 = 0\}$ as for P_3 respectively.) In particular the two traces on S of a real analytic solution u of $P_3 u = 0$ on $M^\pm$ are real analytic at 0.

Remark. In [**D'A-T-Z**, Corollary 1.2] one enconters the same statement as in Theorem 1 but with (2) replaced by:

(2-bis) $$\dot{T}_V T^*_M X \cap C_p(\mathrm{char}(\mathcal{M}), \tilde{V}) = \emptyset.$$

Note that (2-bis) implies (2) because $\mathrm{H}^{\mathbb{I}}(\pi^*\theta)$ belongs to $\dot{T}_V T^*_M X$ due to (1). But the converse is false as for instance for the above symbol $\zeta_1^2 + a + b$ (with $b|_{T^*_M X} \leq 0$) which fulfills (2-bis) only when $a|_{T^*_M X} < 0$ (and (2) for any $a \gtrless 0$).

Proof of Theorem 1. We use the trick of the adjunction of an auxillary variable due to M. Kashiwara. We put $\hat{M} = M \times \mathbb{R}$, $\hat{S} = S \times \mathbb{R}$, $\hat{X} = X \times \mathbb{C}$, $\hat{M}^\pm = M^\pm \times \mathbb{R}$, denote by t (resp. τ) the new variable in $\mathbb{R}$, (resp. $\mathbb{C}$), denote by $j : X \hookrightarrow \hat{X}$ the embedding, and pick up $\hat{p} \in p \times \Big(\{0\} \times_{\mathbb{R}} \dot{T}^*_{\mathbb{R}}\mathbb{C}\Big)$. Then from the exact sequence:

$$0 \to \mathcal{O}_{\hat{X}} \xrightarrow{\tau} \mathcal{O}_{\hat{X}} \to j_*\mathcal{O}_X \to 0,$$

we get, by applying the functor $\mu\mathrm{hom}(\mathbb{Z}_{\hat{M}^\pm}, \cdot)$, a new exact sequence

(7) $$0 \to (\mathcal{C}_{M^\pm|X})_p \xrightarrow{\otimes\delta_t} (\mathcal{C}_{\hat{M}^\pm|\hat{X}})_{\hat{p}} \xrightarrow{t} (\mathcal{C}_{\hat{M}^\pm|\hat{X}})_{\hat{p}}.$$

Therefore by the injectivity of the morphism $\otimes\delta_t$ on the left of (7) we can treat our problem at $\hat{p} \in \hat{V} = V \times \dot{T}^*_{\mathbb{R}}\mathbb{C}$, or else assume from the beginning V regular (involutive) i.e. suppose that the 1–form does not vanish on TV. We can then find complex symplectic homogeneous coordinates $(z, \zeta) = (x + \sqrt{-1}y; \xi + \sqrt{-1}\eta) \in T^*X$, $(x; \sqrt{-1}\eta) \in T^*_M X$ such that $r = x_1,\ s = \eta_1,\ V = \{(x; \sqrt{-1}\eta) \in T^*_M X; \eta_1 = \eta' = 0\}$. We put

$$X = \mathbb{C} \times X' \times X'', \quad M = \mathbb{R} \times M' \times M'',$$
$$S = \{0\} \times M' \times M'', \quad M^{\pm} = \mathbb{R}^{\pm} \times M' \times M'',$$
$$(8) \quad M_1 = \mathbb{R} \times X' \times M'', \quad S_1 = \{0\} \times X' \times M'' \quad M_1^{\pm} = \mathbb{R}^{\pm} \times X' \times M''.$$

We identify $\mathbb{C} \xrightarrow{\sim} \mathbb{R}^2, z_1 \mapsto (x_1, y_1)$, and complexify (x_1, y_1) to $(x_1^{\mathbb{C}}, y_1^{\mathbb{C}}) \in \mathbb{C}^2$. We set

$$\tilde{X} = \mathbb{C}^2 \times X' \times X'', \quad \tilde{M} = \mathbb{R}^2 \times X' \times M'',$$
$$\tilde{S} = (\{0\} \times \mathbb{R}) \times X' \times M'', \quad \tilde{M}^{\pm} = (\mathbb{R}^{\pm} \times \mathbb{R}) \times X' \times M''.$$

Note that (identifying $\tilde{M}$, $\tilde{S}$, $\tilde{M}^{\pm}$ to subsets of X), we have

$$V = M \times_{\tilde{M}} T^*_{\tilde{M}} X, \quad \tilde{V} = T^*_{\tilde{M}} X, \quad \tilde{W} = T^*_{\tilde{S}} X.$$

We shall deal with the sheaves (resp. complexes of sheaves) of usual (resp. "boundary") microfunctions $\mathcal{C}_{S|X}, \mathcal{C}_{M|X}, \mathcal{C}_{\tilde{S}|\tilde{X}}, \mathcal{C}_{\tilde{M}|\tilde{X}}$ (resp. $\mathcal{C}_{M^{\pm}|X}, \mathcal{C}_{\tilde{M}^{\pm}|\tilde{X}}$). Let $T^*X \overset{{}^t j'}{\longleftarrow} X \times_{\tilde{X}} T^*\tilde{X} \overset{j_\pi}{\hookrightarrow} T^*\tilde{X}$, be the mappings canonically associated to the embedding $j : X \hookrightarrow \tilde{X}$. Let $\mathcal{M}$ be a coherent $\mathcal{E}_X$-module (i.e. a microdifferential system) on X, and $\mathcal{O}_{\bar{\mathbb{C}}}$ (the module associated to) the Cauchy-Riemann equation $\bar{\partial}_{z_1}$. The proof of Theorem 1 will require several steps.

Proposition 2. (2) *and* (3) *imply that the natural morphisms*

$$(9) \qquad \mathrm{R}\mathcal{H}om_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M_1|X}) \xrightarrow{\sim} \mathrm{R}\Gamma_{\pi^{-1}(M_1)} \mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{M}|\tilde{X}})[+1]$$
$$\mathrm{R}\mathcal{H}om_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{S_1|X}) \xrightarrow{\sim} \mathrm{R}\Gamma_{\pi^{-1}(S_1)} \mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{S}|\tilde{X}})[+1],$$

are isomorphisms.

(Remark that ${}^t j'$ is injective over $j_\pi^{-1}(\mathrm{char}(\mathcal{O}_{\bar{\mathbb{C}}}))$. For this reason we neglect the functor $\mathrm{R}^t j'_* j_\pi^{-1}$ in the terms on the right side of the above isomorphisms. We shall often act similarly in the following.)

Proof. We consider the commuting diagrams:

$$\begin{array}{ccc} X & \hookrightarrow & \tilde{X} \\ \uparrow & & \uparrow \\ M_1 & \hookrightarrow & \tilde{M}, \end{array} \qquad \begin{array}{ccc} X & \hookrightarrow & \tilde{X} \\ \uparrow & & \uparrow \\ S_1 & \hookrightarrow & \tilde{S}. \end{array}$$

According to [**K-S 1**, Th. 2.3.1], what we need to prove is that the embedding $M_1 \hookrightarrow \tilde{M}$ (resp. $S_1 \hookrightarrow \tilde{S}$) is *microhyperbolic* for the system $\mathcal{O}_{\bar{\mathbb{C}}} \otimes \mathcal{M}$. (As for the additional condition (2.3.1) of loc. cit., this is always satisfied in a suitable neighborhood U of p (and with $\mathcal{M}$ still being the induced system of $\mathcal{O}_{\bar{\mathbb{C}}} \otimes \mathcal{M}|_U$ on ${}^t f' f_\pi^{-1}(U)$).) *Microhyperbolicity* means that in the identification:

$$(10) \qquad T_x^*\tilde{M} \hookrightarrow T_x^*\tilde{M} \oplus (T_{\tilde{M}}^*\tilde{X})_x \simeq T_x^*\tilde{X} \underset{\pi^*}{\hookrightarrow} T_p^*T^*\tilde{X},$$

(which follows from the fact that $\mathbb{R}^2 \times M''$ is totally real in $\mathbb{C}^2 \times X''$), we have

$$(11) \qquad \mathrm{H}^{\mathbb{R}}\left(\pi^*\left(T_{M_1}^*\tilde{M}\right)_x\right) \cap C_p\left(\mathrm{char}\left(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}_{z_1}}\right), T_{\tilde{M}}^*\tilde{X}\right) = \{0\},$$

and

$$(12) \qquad \mathrm{H}^{\mathbb{R}}\left(\pi^*\left(T_{S_1}^*\tilde{S}\right)_x\right) \cap C_p\left(\mathrm{char}\left(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}_{z_1}}\right), T_{\tilde{S}}^*\tilde{X}\right) = \{0\}$$

respectively. Let $(x_1^{\mathbb{C}}, y_1^{\mathbb{C}}; \xi_1^{\mathbb{C}}, \eta_1^{\mathbb{C}})$ be coordinates in $T^*\mathbb{C}^2_{(x_1^{\mathbb{C}}, y_1^{\mathbb{C}})}$; then (11) and (12) are equivalent, for $\sigma(P)(x_1^{\mathbb{C}}, \xi_1^{\mathbb{C}}, z', \xi', z'', \xi'') = 0$ and $\xi_1^{\mathbb{C}} + \sqrt{-1}\eta_1^{\mathbb{C}} = 0$, to:

$$(13) \qquad |\Re\mathrm{e}\eta_1^{\mathbb{C}}| \leq c[|\Re\mathrm{e}\xi_1^{\mathbb{C}}| + |\Im\mathrm{m}x_1^{\mathbb{C}}| + |\zeta'| + |\xi''| + |y''|],$$

and

$$(14) \qquad |\Re\mathrm{e}\eta_1^{\mathbb{C}}| \leq c[|x_1^{\mathbb{C}}| + |\zeta'| + |\xi''| + |y''|]$$

respectively. But by the substitution $\Re\mathrm{e}\eta_1^{\mathbb{C}} = -\Im\mathrm{m}\xi_1^{\mathbb{C}}$, (13) and (14) are immediate consequences of (2) and (3) respectively. □

Proposition 3. *Assume* (2) *and* (3). *Then the natural morphism*

$$(15) \qquad \mathrm{R}\mathcal{H}om_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M_1^{\pm}|X}) \xrightarrow{\sim} R\Gamma_{\pi^{-1}(M_1)}\mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{M}^{\pm}|\tilde{X}})$$

is an isomorphism.

Proof. (Again we neglect here the functor $\mathrm{R}^t j'_* j_\pi^{-1}$ in the right of (15).) The morphism $\mathcal{C}_{A|X} \to \mathrm{R}^t j'_* j_\pi^{-1}\mathrm{R}\Gamma_{\pi^{-1}(A)}\mathcal{C}_{\tilde{A}|\tilde{X}}$ ($A = M_1^{\pm}$, M_1, S_1) induces the vertical arrows in the following commuting diagram in the category $D^b(T^*X)$:

(16)

$$\begin{array}{ccc}
\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{S_1|X}) & \to & \mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M_1|X}) \dots \\
\downarrow & & \downarrow \\
\mathrm{R}\Gamma_{\pi^{-1}(S_1)}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{S}|\tilde{X}}) & \to & \mathrm{R}\Gamma_{\pi^{-1}(M_1)}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{M}|\tilde{X}}) \dots
\end{array}$$

$$\begin{array}{c}
\cdots \to \bigoplus_{\pm}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M_1^{\pm}|X}) \\
\downarrow \\
\dots \bigoplus_{\pm} \mathrm{R}\Gamma_{\pi^{-1}(M_1)}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{M}^{\pm}|\tilde{X}}).
\end{array}$$

By Proposition 3 the two first vertical arrows are isomorphisms. Hence the third is an isomorphism too. □

For V_1 defined in $T^*_M X$ by $\eta' = 0$, let us recall the complex by [**U-Z**] of 2-hyperfunctions at the boundary along V_1:

$$\mathcal{B}^{2,V_1}_{M^\pm|X} = \mathrm{R}\Gamma_{\pi^{-1}(M)}(\mathcal{C}_{M_1^\pm|X})[d], \tag{17}$$

$(d = \operatorname{codim} V_1)$. We put

$\tilde{M}_2 = \mathbb{R}^2 \times M' \times M''$, $\tilde{S}_2 = (\{0\} \times \mathbb{R}) \times M' \times M''$, $\tilde{M}_2^\pm = (\mathbb{R}^\pm \times \mathbb{R}) \times M' \times M''$,

and

$\tilde{M}_3 = (\mathbb{R} \times \mathbb{C}) \times X' \times M''$, $\tilde{S}_3 = (\{0\} \times \mathbb{C}) \times X' \times M''$, $\tilde{M}_3^\pm = (\mathbb{R}^\pm \times \mathbb{C}) \times X' \times M''$.

Along with V_1 we also consider in $T^*_{\tilde{M}_2}\tilde{X}$, $V_2 = \{\eta' = 0\}$, $V_3 = \{\Im\mathrm{m}\eta_1^{\mathbb{C}} = \eta' = 0\}$. We define similarly to (17):

$$\begin{aligned} \mathcal{B}^{2,V_2}_{\tilde{M}_2^\pm|\tilde{X}} &= \mathrm{R}\Gamma_{\pi^{-1}(\tilde{M}_2)}(\mathcal{C}_{\tilde{M}^\pm|\tilde{X}})[d], \\ \mathcal{B}^{2,V_3}_{\tilde{M}_2^\pm|\tilde{X}} &= \mathrm{R}\Gamma_{\pi^{-1}(\tilde{M}_2)}(\mathcal{C}_{\tilde{M}_3^\pm|\tilde{X}})[d+1]. \end{aligned} \tag{18}$$

According to [**U-Z**, Th. 2.6], $\mathcal{B}^{2,V_1}_{M^\pm|X}|_{V_1}$ and $\mathcal{B}^{2,V_i}_{\tilde{M}_2^\pm|\tilde{X}}|_{V_i}$, $i = 2, 3$ are all concentrated in degree 0, (whence they are naturally endowed with a structure of $\mathcal{E}_X$ or $\mathcal{E}_{\tilde{X}}$–modules). We also recall the complexes of usual 2–hyperfunctions by Kashiwara ([**K**]):

$$\mathcal{B}^{2,V_1}_{M|X},\ \mathcal{B}^{2,V_i}_{\tilde{M}|\tilde{X}}\,(i = 2,3),\ \mathcal{B}^{2,V_1}_{S|X},\ \mathcal{B}^{2,V_i}_{\tilde{S}_2|\tilde{X}}\,(i = 2,3), \tag{19}$$

defined similarly to (17), (18). It is classical that they are all concentrated in degree 0. We apply $\mathrm{R}\Gamma_{\pi^{-1}(M)}(\cdot)[d]$ to (9), (15) and get

$$\begin{aligned} \mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{B}^{2,V_1}_{M|X}) &\xrightarrow{\sim} \mathrm{R}\Gamma_{\pi^{-1}(M)}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\tilde{\mathbb{C}}}, \mathcal{B}^{2,V_2}_{\tilde{M}_2|\tilde{X}}) \\ \mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{B}^{2,V_1}_{S|X}) &\xrightarrow{\sim} \mathrm{R}\Gamma_{\pi^{-1}(S)}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\tilde{\mathbb{C}}}, \mathcal{B}^{2,V_2}_{\tilde{S}_2|\tilde{X}}) \\ \mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{B}^{2,V_1}_{M^\pm|X}) &\xrightarrow{\sim} \mathrm{R}\Gamma_{\pi^{-1}(M)}\mathrm{R}\mathcal{H}\mathrm{om}_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\tilde{\mathbb{C}}}, \mathcal{B}^{2,V_2}_{\tilde{M}_2^\pm|\tilde{X}}). \end{aligned} \tag{20}$$

The natural (restriction) morphism $\mathbb{Z}_{M_1} \to \mathbb{Z}_M$, resp. $\mathbb{Z}_{\tilde{M}_3} \to \mathbb{Z}_{\tilde{M}}$, induces a morphism

$$\mathcal{C}_{M|X}|_{V_1} \to \mathcal{B}^{2,V_1}_{M|X}, \quad \text{resp. } \mathcal{B}^{2,V_2}_{\tilde{M}_2|\tilde{X}}\Big|_{V_3} \to \mathcal{B}^{2,V_3}_{\tilde{M}_2|\tilde{X}}. \tag{21}$$

It is classical (cf. [**K**]) that the first is injective. We show now:

Proposition 4. *The morphism*

$$\mathcal{B}^{2,V_2}_{\tilde{M}_2|\tilde{X}}\Big|_{V_3} \to \mathcal{B}^{2,V_3}_{\tilde{M}_2|\tilde{X}} \tag{22}$$

is injective.

Proof. Fix $p = (x_o; \sqrt{-1}\eta'' \mathrm{d}x'') \in V_3$ and let Z_2, resp. Z_3, describe the family of closed convex subsets of $\mathbb{R}^{n+1}_{(\Im\mathrm{m} x_1^{\mathbb{C}}, \Im\mathrm{m} y_1^{\mathbb{C}}, y', y'')}$ such that

$$\begin{cases} Z_2 \subset \{y| < y'', \eta'' > \geq \epsilon(|y''| + |\Im\mathrm{m} x_1^{\mathbb{C}}| + |\Im\mathrm{m} y_1^{\mathbb{C}}|)\}, \\ Z_2 \cap \{y'' = 0,\, \Im\mathrm{m} x_1^{\mathbb{C}} = 0,\, \Im\mathrm{m} y_1^{\mathbb{C}} = 0\} \subset \{0\}, \end{cases}$$

resp.

$$\begin{cases} Z_3 \subset \{y| < y'', \eta'' > \geq \epsilon(|\Im\mathrm{m} x_1^{\mathbb{C}}| + |y''|)\}, \\ Z_3 \cap \{\Im\mathrm{m} x_1^{\mathbb{C}} = 0,\, y'' = 0\} \subset \{0\}. \end{cases}$$

Thus the arrow in (22) can be represented, between the stalks at p, by:

$$\varinjlim_{B, Z_2} \mathrm{H}^n_{\tilde{M}_2 + \sqrt{-1} Z_2}(B, \mathcal{O}_X) \to \varinjlim_{B, Z_3} \mathrm{H}^n_{\tilde{M}_2 + \sqrt{-1} Z_3}(B, \mathcal{O}_X),$$

for B describing a fundamental system of neighborhoods of $x_o = \pi(p)$. Now for any B (convex) and for any Z_2, Z_3, there exist $Z_2' \supset Z_2$ such that $Z_3 \setminus Z_2' \subset\subset B$. If then $K_2 = \bar{B} \cap (\tilde{M}_2 + \sqrt{-1} Z_2')$, $K_3 = \bar{B} \cap (\tilde{M}_2 + \sqrt{-1} Z_3)$, we have $K_3 \setminus K_2 = (\tilde{M}_2 + \sqrt{-1}(Z_3 \setminus Z_2')) \cap B$ and therefore

$$\mathrm{H}^{n-1}_{\tilde{M}_2 + \sqrt{-1}(Z_3 \setminus Z_2')}(B, \mathcal{O}_X) = \mathrm{H}^{n-1}_{K_3 \setminus K_2}(B, \mathcal{O}_X) = 0,$$

by a celebrated theorem due to M. Kashiwara. $\square$

Note that the first morphism in (21) is a particular case of the second. Hence Proposition 5 provides also a proof of the injectivity of the former.

The natural morphisms $\mathbb{Z}_{M_1^\pm} \to \mathbb{Z}_{M^\pm}$, resp. $\mathbb{Z}_{\tilde{M}_3^\pm} \to \mathbb{Z}_{\tilde{M}^\pm}$, in turn induce morphisms:

$$\mathcal{C}_{M^\pm|X}|_{V_1} \to \mathcal{B}^{2,V_1}_{M^\pm|X} \quad \text{resp. } \mathcal{B}^{2,V_2}_{\tilde{M}_2^\pm|\tilde{X}}\Big|_{V_3} \to \mathcal{B}^{2,V_3}_{\tilde{M}\pm_2|\tilde{X}}. \tag{23}$$

Neither of them is injective ([**U-Z**, Remark 2.7]). Nevertheless they can be injective when restricted to solutions of non-characteristic systems. Let $V_4 = {}^t j'^{-1}(V) \cap T^*_{\tilde{M}}\tilde{X}$ (i.e. V_4 is the submanifold of $T^*_{\tilde{M}_2}\tilde{X}$ defined by $\Im\mathrm{m}\xi_1^{\mathbb{C}} = \Im\mathrm{m}\eta_1^{\mathbb{C}} = \eta' = 0$); note that $V_4 = V_2 \cap \mathrm{char}(\mathcal{O}_{\bar{\mathbb{C}}_{z_1}}) = V_3 \cap \mathrm{char}(\mathcal{O}_{\bar{\mathbb{C}}_{z_1}})$. We have:

Proposition 5. *Let $S^{\mathbb{C}} \hookrightarrow X$ be non-characteristic for $\mathcal{M}$, and consider the sequence of morphisms:*

$$\begin{aligned} \mathrm{Hom}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M^\pm|X})|_V &\to \mathrm{Hom}_{\mathcal{E}_X}\left(\mathcal{M}, \mathcal{B}^{2,V_1}_{M^\pm|X}\right)\Big|_V \\ &\xrightarrow{\sim} \Gamma_{\pi^{-1}(M)}\mathrm{Hom}_{\mathcal{E}_{\tilde{X}}}\left(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}_{z_1}}, \mathcal{B}^{2,V_2}_{\tilde{M}_2^\pm|\tilde{X}}\right)\Big|_{V_4} \\ &\hookrightarrow \mathrm{Hom}_{\mathcal{E}_{\tilde{X}}}\left(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}_{z_1}}, \mathcal{B}^{2,V_2}_{\tilde{M}_2^\pm|\tilde{X}}\right)\Big|_{V_4} \\ &\to \mathrm{Hom}_{\mathcal{E}_{\tilde{X}}}\left(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}_{z_1}} \mathcal{B}^{2,V_3}_{\tilde{M}_2^\pm|\tilde{X}}\right)\Big|_{V_4}, \end{aligned} \tag{24}$$

with the first and the fourth arrow induced by (23), *the second by* (20), *and the third being the natural identification. Then the composition of the morphisms in* (23) *is injective.*

Remark 6. In particular the first morphism in (24) is injective. Our proof will show that this is in fact injective on the whole V_1 (not only on V) according to [**U-Z**, Th. 2.8]. However the full generalization of this statement (in analogy with Proposition 4), i.e. the injectivity of the last morphism in (24) is not clear to us because of the lack of a 2-microlocal version of the *watermelon-cut* Theorem (cf. [**S**]).

Proof. We consider

$$\begin{array}{ccc} \mathcal{B}^{2,V_2}_{\tilde{M}_2^{\pm}|\tilde{X}}\Big|_{\tilde{S}_2\times_{\tilde{M}_3}T^*_{\tilde{M}_3}\tilde{X}} & \to & \mathcal{B}^{2,V_3}_{\tilde{M}_2^{\pm}|\tilde{X}}\Big|_{\tilde{S}_2\times_{\tilde{M}_3}T^*_{\tilde{M}_3}\tilde{X}} \\ \downarrow & & \downarrow \\ \mathrm{R}\Gamma_{\tilde{F}^{\pm}}\mathcal{B}^{2,V_2}_{\tilde{S}_2|\tilde{X}}\Big|_{\tilde{S}_2\times_{\tilde{M}_3}T^*_{\tilde{M}_3}\tilde{X}}[1] & \to & \mathrm{R}\Gamma_{\tilde{F}^{\pm}}(\mathcal{B}^{2,V_3}_{\tilde{S}_2|\tilde{X}})\Big|_{\tilde{S}_2\times_{\tilde{M}_3}T^*_{\tilde{M}_3}\tilde{X}}[+1], \end{array} \tag{25}$$

where $\tilde{F}^{\pm} = (\tilde{S}_2 \times_{\tilde{M}_2} T^*_{\tilde{M}_2}\tilde{X}) \pm \mathbb{R}^+\theta$ with θ the exterior conormal to $\tilde{M}^+$ in $\tilde{M}$. Remark that the vertical arrows of (25) are induced by the natural morphisms $\mathcal{B}^{2,V_i}_{\tilde{M}_2^{\pm}|\tilde{X}} \to \mathcal{B}^{2,V_i}_{\tilde{S}_2|\tilde{X}}[1]$ which factorize through $\mathrm{R}\Gamma_{\tilde{F}^{\pm}}\mathcal{B}^{2,V_i}_{\tilde{S}_2|\tilde{X}}[1]$ (due to $\mathrm{supp}(\mathcal{B}^{2,V_i}_{\tilde{M}_2^{\pm}|\tilde{X}}) \cap \mathrm{supp}(\mathcal{B}^{2,V_i}_{\tilde{S}_2|\tilde{X}}) \subset \tilde{F}^{\pm}$). If we apply $\mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}(\mathcal{M} \otimes \mathcal{O}_{\bar{\mathbb{C}}_{z_1}}, \cdot)$ to (25) and take the 0-th cohomology, the arrow on the bottom becomes injective. In fact let $\tilde{Y} = \tilde{S}_2^{\mathbb{C}}$ be the complexification of S_2 (i.e. $\tilde{Y} = \mathbb{C}_{y_1^{\mathbb{C}}} \times X' \times X''$), denote by $k : \tilde{Y} \to \tilde{X}$ the natural embedding, and let $V_i' = {}^t k' k_\pi^{-1}(V_i)$. By the aid of division formulas for $\mathcal{B}^{2,V_i}_{\tilde{S}_2|\tilde{X}}$, the above injectivity is reduced to the injectivity of

$$\mathcal{B}^{2,V_2'}_{\tilde{S}_2|\tilde{Y}}\Big|_{V_3'} \hookrightarrow \mathcal{B}^{2,V_3'}_{\tilde{S}_2|\tilde{Y}}.$$

But this is, under different notations, the same statement as in Proposition 4. We consider now:

$$\begin{array}{ccc} \mathcal{C}_{M^{\pm}|X}|_{S\times_{M_1}T^*_{M_1}X} & \to & \mathcal{B}^{2,V_1}_{M^{\pm}|X}\Big|_{S\times_{M_1}T^*_{M_1}X} \\ \downarrow & & \downarrow \\ \mathrm{R}\Gamma_{F^{\pm}}(\mathcal{C}_{S|X})|_{S\times_{M_1}T^*_{M_1}X}[1] & \to & \mathrm{R}\Gamma_{F^{\pm}}(\mathcal{B}^{2,V_1}_{S|X})\Big|_{S\times_{M_1}T^*_{M_1}X}[1], \end{array} \tag{26}$$

with $F^{\pm} = S \times_M T^*_M X \pm \mathbb{R}^+\theta$. The arrow in the bottom is injective, over solutions of $\mathcal{M}$, for the same argument as for (25). Concerning the first vertical arrow, this is represented at each point $p \in S \times_M T^*_M X$ by

$$(\mathcal{C}_{M^\pm|X})_p \simeq \left(\frac{\mathcal{C}_{\bar{M}^\pm|X}}{C_{S|X}}\right)_p \hookrightarrow \mathcal{H}^1_{F^\pm}(\mathcal{C}_{S|X})_p,$$

(where $\mathcal{C}_{\bar{M}^\pm|X}$ are the Kataoka's microfunctions along the closed half-spaces $\bar{M}^\pm$) whose injectivity is immediately proved by the aid of a Legendre transformation (cf. [**Kat**] and [**S**]).

We are ready to conclude. We apply $\mathrm{R}\mathcal{H}om_{\mathcal{E}_X}(\mathcal{M},\cdot)$ to (26) and $\mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}(\mathcal{M}\otimes\mathcal{O}_{\bar{\mathbb{C}}_{z_1}},\cdot)$ to (25) respectively (and neglect $\mathrm{R}^t j'_* j_\pi^{-1}$). We glue the diagrams so obtained by means of the second and third of (20) and by the natural morphism $\mathrm{R}\Gamma_{\pi^{-1}(M)}(\cdot)\to\cdot$. We thus obtain a long diagram with the first vertical and all the bottom horizontal arrows injective over the 0-th cohomology. The composition of the upper horizontal arrows (which is precisely the sequence of morphisms in (24)) is therefore also injective. $\square$

End of proof of Theorem 1. Let $\pm\theta$ be the exterior conormals to $\tilde{M}^\pm$ in $\tilde{M}$ identified to vectors $\mathrm{H}^{\mathbb{R}}(\pm\pi^*\theta)$ of $T_pT^*\tilde{X}$ (cf. (10)). Then clearly

$$\mathrm{H}^{\mathbb{R}}(\pm\pi^*(\theta)) \notin C(\mathrm{char}(\mathcal{O}_{\bar{\mathbb{C}}}), \tilde{V}_3). \tag{27}$$

Let $\mathrm{SS}(\mathbb{Z}_{\tilde{M}_3^\pm})$ denote the microsupport of $\mathbb{Z}_{\tilde{M}_3^\pm}$ in the sense of [**K-S 2**]. One easily checks that

$$\mathrm{SS}(\mathbb{Z}_{\tilde{M}_3^\pm}) = \left(\overline{\tilde{M}_3^\pm} \times T^*_{\tilde{M}_3}\tilde{X}\right) \pm \mathbb{R}^+\theta.$$

It is also easy to see that (27) implies

$$-\mathrm{H}^{\mathbb{R}}(\pm\pi^*\theta) \notin C\left(\mathrm{char}(\mathcal{O}_{\bar{\mathbb{C}}}), \mathrm{SS}(\mathbb{Z}_{\tilde{M}_3^\pm})\right).$$

It follows, merely by definition of SS:

$$\mathrm{R}\Gamma_{\pi^{-1}(\tilde{S}_3)}\mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}(\mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{C}_{\tilde{M}_3^\pm|\tilde{X}}) = 0,$$

and thus, by applying $\mathrm{R}\Gamma_{\pi^{-1}(\tilde{M}_2)}(\cdot)[d+1]$:

$$\mathrm{R}\Gamma_{\pi^{-1}(\tilde{S}_2)}\mathrm{R}\mathcal{H}om_{\mathcal{E}_{\tilde{X}}}\left(\mathcal{O}_{\bar{\mathbb{C}}}, \mathcal{B}^{2,V_3}_{\tilde{M}_2^\pm|\tilde{X}}\right) = 0. \tag{28}$$

In conclusion we have

$$\begin{aligned}\Gamma_{\pi^{-1}(S)}\mathrm{Hom}_{\mathcal{E}_X}\left(\mathcal{M},\mathcal{C}_{M^\pm|X}\right)|_V &\hookrightarrow \Gamma_{\pi^{-1}(S)}\mathrm{Hom}_{\mathcal{E}_X}\left(\mathcal{M},\mathcal{B}^{2,V_1}_{M^\pm|X}\right)\Big|_V\\ &\xrightarrow{\sim} \Gamma_{\pi^{-1}(S)}\mathrm{Hom}_{\mathcal{E}_{\tilde{X}}}\left(\mathcal{M}\otimes\mathcal{O}_{\bar{\mathbb{C}}_{z_1}},\mathcal{B}^{2,V_2}_{\tilde{M}_2^\pm|\tilde{X}}\right)\Big|_{V_4}\\ &\to \Gamma_{\pi^{-1}(S)}\mathrm{Hom}_{\mathcal{E}_{\tilde{X}}}\left(\mathcal{M}\otimes\mathcal{O}_{\bar{\mathbb{C}}_{z_1}},\mathcal{B}^{2,V_3}_{\tilde{M}_2^\pm|\tilde{X}}\right)\Big|_{V_4}\\ &= 0,\end{aligned}$$

(where the first "$\hookrightarrow$" follows from Remark 6, the second "$\overset{\sim}{\to}$" from (20), the third "$\to$" from (23), and the last "$=$" from (28) respectively). On the other hand the composition of "$\hookrightarrow$", "$\overset{\sim}{\to}$" and "$\to$" is injective by Proposition 5; hence $\Gamma_{\pi^{-1}(S)}\mathrm{Hom}_{\mathcal{E}_X}(\mathcal{M}, \mathcal{C}_{M^\pm|X})|_V = 0$. The proof is complete. $\square$

Aknowledgements. The author is grateful to an anonimous referee for pointing out the problem in the second part of Remark 6 which lead to Proposition 5.

References

[B] J.M. Bony, *Extension du Théorème de Holmgren*, Sém. Goulaouic-Schwartz, (1975-76), Exp. 17.

[D'A-T-Z] P. D'Ancona, N. Tose and G. Zampieri, *Propagation of singularities up to the boundary along leaves*, Comm. in Partial Differential Equations, **15**(**4**) (1990), 453-460.

[K] M. Kashiwara, *Talks in Nice*, (1972).

[K-S 1] M. Kashiwara and P. Schapira, *Microhyperbolic systems*, Acta Math., **142** (1979), 1-55.

[K-S 2] ———, *Microlocal study of sheaves*, Astérisque, **128** (1985).

[Kat] K. Kataoka, *Microlocal theory of boundary value problems* I, J. Fac. Sci. Univ. Tokyo Sect. IA Math., **35** (1980), 355-399; II, ibid., **28** (1981), 31-56.

[S-K-K] M. Sato, M. Kashiwara and T. Kawai, *Hyperfunctions and pseudodifferential equations*, Springer Lecture Notes in Math., **287** (1973), 265-529.

[S] P. Schapira, *Front d'onde analytique au bord* I *and* II, C.R. Acad. Sci. Paris, **302**(**10**) (1986), 383-386; Sém. E.D.P. École Polyt., Exp., **13** (1986).

[S-Z] P. Schapira and G. Zampieri, *Regularity at the boundary for systems of microdifferential equations*, Pitman Res. Notes in Math., **158** (1987), 186-201.

[Sj] J. Sjöstrand, *Singularités analytiques microlocales*, Astérisque, **95** (1982).

[U-Z] M. Uchida and G. Zampieri, *Second microlocalization at the boundary and microhyperbolicity*, Publications of the Research Institute for Mathematical Sciences, Kyoto University, **26** (1990), 205-219.

Received April 22, 1997 and revised May 27, 1998.

Dip. Mat. Università
v. Belzoni 7, 35131 Padova
Italy
E-mail address: zampieri@math.unipd.it

CONTENTS

Volume 188, no. 1 and no. 2

B. **Corbas** and G.D. Williams: *Congruence of two-dimensional subspaces in $M_2(K)$ (characteristic 2)* 237

U. **Keich**: *The entropy distance between the Wiener and stationary Gaussian measures* 109

Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 188 No. 2 April 1999

0030-8730(199904)188:2;1-Z