Journal für die reine und angewandte Mathematik

gegründet 1826 von

August Leopold Crelle

fortgeführt von
C.W. Borchardt, K. Weierstrass, L. Kronecker, L. Fuchs,
K. Hensel, L. Schlesinger, H. Hasse, H. Rohrbach

gegenwärtig herausgegeben von

JRMAA8

Band 470

Walter de Gruyter · Berlin · New York 1996

⊚ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

Microform editions are available from: University Microfilms International,
300 N. Zeeb Road, Ann. Arbor, MI 48106, USA.

Printed in Germany.
Satz und Druck: Tutte Druckerei GmbH, Salzweg-Passau.
Buchbinderische Verarbeitung: Lüderitz & Bauer GmbH, Berlin.
ISSN 0075-4102

Inhalt

Indexed in Current Contents/PC & ES, CompuMath Citation Index, Science Citation Index (SCI), Research Alert, and MATH Database on STN International.
Covered by Current Mathematical Publications, Mathematical Reviews and Zentralblatt für Mathematik/Mathematics Abstracts.

Ausgabedatum des Bandes 470

31. Januar 1996

J. reine angew. Math. **470** (1996), 1—26

Journal für die reine und angewandte Mathematik

The field of moduli of abelian surfaces with complex multiplication

By *Naoki Murabayashi* at Yamagata

---

§ 1. Introduction

Let E be an elliptic curve defined over the complex number field $\mathbb{C}$ whose endomorphism ring $\mathrm{End}(E)$ is isomorphic to $\mathcal{O}_K$, the ring of integers of an imaginary quadratic field K. By the theory of complex multiplication, we have:

(1.1) The following statements are equivalent:

(i) E has a model defined over $\mathbb{Q}$.

(ii) The j-invariant j_E of E is contained in $\mathbb{Q}$.

(iii) The class number h_K of K is equal to one.

Independently Baker and Heegner-Stark solved the class number 1 problem in 1967:

(1.2) There are nine imaginary quadratic fields $\mathbb{Q}(\sqrt{d})$, $d<0$, having class number equal to 1:

$$-d = 3, 4, 7, 8, 11, 19, 43, 67, 163.$$

E has two representations. One of them is the algebraic one given by a Weierstrass equation

$$E: y^2 = 4x^3 - g_2 x - g_3,$$

where $g_2, g_3 \in \mathbb{C}$ such that $g_2^3 - 27g_3^2 \neq 0$. The other is the analytic one as a complex torus

$$E(\mathbb{C}) \cong \mathbb{C}/L_\tau, \quad L_\tau = \mathbb{Z} + \mathbb{Z}\tau,$$

where the complex points $E(\mathbb{C})$ is seen as a complex Lie group and $\tau \in \mathfrak{H}$, the complex upper half plane. So we can represent j_E in two manners:

$$j_E = 12^3 \frac{g_3^2}{g_2^3 - 27g_3^2}$$
$$= \frac{1}{q} + 744 + \sum_{n=1}^{\infty} c(n) q^n,$$

where $q = e^{2\pi i \tau}$, $c(n) \in \mathbb{Z}$ and we can compute $c(n)$. By computing an approximate value, we can exactly get the j-invariants of the elliptic curves of CM-type with the property in (1.1). This leads the following:

(1.3) We can exactly determine the set of $\mathbb{Q}$-rational CM-points (with respect to the maximal order) in $\mathscr{A}_1$, where $\mathscr{A}_1$ is the coarse moduli scheme of elliptic curves and isomorphic to the affine j-line $\mathbb{A}_j^1$.

In particular we have:

(1.4) For the elliptic curves as above, we can explicitly determine a defining equation of them.

Now we can present the following questions:

(Q : i, g) Generalize the statement (1.i) to the case of g-dimensional abelian varieties of CM-type ($1 \leqq i \leqq 4$).

For a g-dimensional abalian variety ($g \geqq 2$), its defining equations are very complex. So we can not expect a good answer for (Q : 4, g) ($g \geqq 2$). But for the case of $g = 2$ or 3, since any simple principally polarized abelian variety of dimension g is isomorphic to the Jacobian variety of a curve of genus g, we can try to determine a defining equation of the curve corresponding to the considered abelian variety. In particular in the case of $g = 2$ we have Igusa's j-invariants and their representations by theta series. Therefore, if the answer for (Q : 1, 2) and (Q : 2, 2) are known, we can find those for (Q : 3, 2) and (Q : 4, 2).

In this paper we will give the answer for (Q : 1, 2). More precisely our result (Theorem 4.12 in § 4) is the generalization of "(ii) $\Leftrightarrow$ (iii) in (1.1)" to the case of $g = 2$ under the restriction to CM by the maximal order.

§ 2. Preliminaries

In this section, we give a review of the theory of abelian varieties of CM-type and their field of moduli. By an algebraic number field, we always mean a subfield of $\mathbb{C}$ algebraic over $\mathbb{Q}$ of finite degree. A CM-field is a totally imaginary quadratic extension of a totally real algebraic number field. Let K be a CM-field with $[K:\mathbb{Q}] = 2n$ and F the maximal real subfield of K. Let $\mathcal{O}_K$ be the ring of integers in K. We consider a structure $P = (A, C, \theta)$ formed by an abelian variety A of dimension n defined over $\mathbb{C}$, a polarization C of A, and an injection θ of K into $\mathrm{End}^0(A) := \mathrm{End}(A) \otimes \mathbb{Q}$ such that $\theta^{-1}(\mathrm{End}(A)) = \mathcal{O}_K$. We always assume that $\theta(K)$ is stable under the involution of $\mathrm{End}^0(A)$ determined by C.

One can prove that there exists a subfield M_K of $\mathbb{C}$ which is uniquely characterized by the following condition:

An automorphism σ of $\mathbb{C}$ is the identity map on M_K if and only if there is an isomorphism λ of A to A^σ such that $\lambda(C) = C^\sigma$ and $\lambda \circ \theta(a) = \theta^\sigma(a) \circ \lambda$ for all $a \in K$.

M_K is called the field of moduli of P. Let

$$\Phi : K \to \mathrm{End}_{\mathbb{C}}(\mathrm{Lie}(A))$$

be the representation of K, through θ, on the Lie algebra of A and let $\sigma_1, \ldots, \sigma_n$ be the n injections of K into $\mathbb{C}$ which form Φ. Then one writes

$$\Phi \sim \sum_{i=1}^{n} \sigma_i .$$

An isomorphism $\tilde{\Phi}$ of $\mathbb{R}$-linear spaces,

$$\tilde{\Phi} : K \otimes_{\mathbb{Q}} \mathbb{R} \to \mathbb{C}^n, \quad \alpha \otimes a \mapsto (a\alpha^{\sigma_1}, \ldots, a\alpha^{\sigma_n})$$

can be defined. For each $\alpha \in K$, let

$$S_\Phi(\alpha) = S(\alpha) = \begin{pmatrix} \alpha^{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \alpha^{\sigma_n} \end{pmatrix}$$

be the diagonal matrix $\mathrm{diag}(\alpha^{\sigma_1}, \ldots, \alpha^{\sigma_n})$. Then there exist a fractional ideal $\mathfrak{a}$ of K and an analytic isomorphism

$$\lambda : \mathbb{C}^n/\tilde{\Phi}(\mathfrak{a}) \to A$$

such that the following diagram is commutative for all $\alpha \in \mathcal{O}_K$:

$$\begin{array}{ccc} \mathbb{C}^n/\tilde{\Phi}(\mathfrak{a}) & \xrightarrow{\lambda} & A \\ {\scriptstyle S_\Phi(\alpha)}\downarrow & & \downarrow{\scriptstyle \theta(\alpha)} \\ \mathbb{C}^n/\tilde{\Phi}(\mathfrak{a}) & \xrightarrow[\lambda]{} & A . \end{array}$$

Take a basic polar divisor in C and consider its Riemann form $E(x, y)$ on $\mathbb{C}^n$ with respect to λ. Then there is an element η of K such that

$$E(\tilde{\Phi}(x), \tilde{\Phi}(y)) = \mathrm{Tr}_{K/\mathbb{Q}}(\eta x y^\varrho) \quad ((x, y) \in K \times K),$$

$$\eta^\varrho = -\eta, \quad \mathrm{Im}(\eta^{\sigma_i}) > 0 \quad (i = 1, \ldots, n),$$

where ϱ denotes the complex conjugation in $\mathbb{C}$. Since η is obtained from a basic divisor, we have

$$\mathrm{Tr}_{K/\mathbb{Q}}(\eta \mathfrak{a} \mathfrak{a}^\varrho) = \mathbb{Z} .$$

P is said to be of type $(K, \Phi; \eta, \mathfrak{a})$ (with respect to λ).

Put

$$\tilde{C}_K := F_+ \times I_K / \{\{xx^\varrho, (x)\} \mid x \in K^\times\},$$

where F_+ denotes the group of totally positive elements of F and I_K denotes the group of fractional ideals in K. For an element $\{s, \mathfrak{b}\}$ of $F_+ \times I_K$, its image in $\tilde{C}_K$ is denoted by $(s, \mathfrak{b})$. If $P' = (A', C', \theta')$ is an another structure of type $(K, \Phi; \eta', \mathfrak{a}')$, then P is isomorphic to P' if and only if $(P' : P) := (\eta^{-1}\eta', \mathfrak{a}'^{-1}\mathfrak{a})$ is equal to 1 in $\tilde{C}_K$.

Let (K', Φ') be the reflex of (K, Φ). Put

$$I_0(\Phi') := \{\mathfrak{b} \in I_{K'} \mid \mathfrak{b}^{\Phi'} = (\alpha), \alpha\alpha^\varrho = N(\mathfrak{b}) \text{ for } \exists \alpha \in K^\times\},$$

where $\mathfrak{b}^{\Phi'}$ denotes the product $\prod_{i=1}^{m} \mathfrak{b}^{\tau_i}$ if $\Phi' \sim \sum_{i=1}^{m} \tau_i$. From the theory of complex multiplication, the following holds (see [5], p. 68):

Theorem 2.1. *M_K is an unramified abelian extension of K' corresponding to the ideal group $I_0(\Phi')$.*

From now on we assume that there exists a subfield D' (resp. D) of K' (resp. K) such that there is an isomorphism

$$\mathrm{Gal}(K'/D') \to \mathrm{Gal}(K/D), \quad \sigma_0 \mapsto [\sigma_0],$$

satisfying $\mathrm{tr}\,\Phi(a^{[\sigma_0]}) = (\mathrm{tr}\,\Phi(a))^{\sigma_0}$ for every $a \in K$. This condition was introduced in [5]. For $\sigma \in \mathrm{Aut}(\mathbb{C}/D')$, put

$$P_\sigma := (A^\sigma, C^\sigma, \theta_\sigma),$$

where

$$\theta_\sigma(a) := \theta(a^{[\sigma_0]^{-1}})^\sigma$$

and $\sigma_0 = \sigma|_{K'}$ denotes the restriction of σ to K'. Then P_σ is of the same type (K, Φ) so that

$$(P_\sigma : P) \in \tilde{C}_K$$

can be considered. If $\sigma|_{M_K}$ is the identity map, then $(P_\sigma : P) = 1$. It follows that for $\sigma \in \mathrm{Gal}(M_K/D')$, $(P_\sigma : P)$ is well defined. The following is Proposition 3 in [5]:

Proposition 2.2. *Let M_D be the field of moduli of a structure $(A, C, \theta|_D)$, where $\theta|_D$ is the restriction of θ to D. If (K, Φ) is primitive (i.e. A is simple), then*

$$\mathrm{Gal}(M_K/M_D) = \{\sigma \in \mathrm{Gal}(M_K/D') \mid (P_\sigma : P) = 1\}.$$

Take a basic polar divisor X contained in C and consider the isogeny φ_X of A onto its Picard variety:

$$\varphi_X : A \to \mathrm{Pic}^0(A), \quad v \mapsto [X_v - X].$$

Then there exists an ideal $\mathfrak{f}$ of $\mathcal{O}_F$ which is uniquely determined by the property

$$\mathrm{Ker}(\varphi_X) = \{y \in A \mid \theta(\mathfrak{f})y = 0\}. \tag{2.1}$$

It is known that $\mathfrak{f}\mathcal{O}_K = (\eta)\mathfrak{d}_K \mathfrak{a}\mathfrak{a}^{\varrho}$, where $\mathfrak{d}_K$ denotes the different of K over $\mathbb{Q}$ ([6], p. 118). We will write $\mathfrak{f} = \mathfrak{f}(P)$.

Proposition 2.3 ([5], (2.7)). *Assume that (K, Φ) is primitive. If M_D is linearly disjoint with K' over D', then*

$$\mathfrak{f}(P)^{\gamma} = \mathfrak{f}(P) \quad \forall \gamma \in \mathrm{Gal}(K/D).$$

§ 3. Arbitrary conjugations of abelian varieties of CM-type

In this section, we review an analytic description of conjugations of abelian varieties of CM-type under an arbitrary automorphism of $\mathbb{C}$. We refer to [1] and [3] for details.

Let

$$r_K : K_{\mathbb{A}}^{\times} \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$$

denote the reciprocity law of class field theory, where $K_{\mathbb{A}}^{\times}$ denotes the idèle group of K and K^{ab} denotes the maximal abelian extension of K in $\bar{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Let $\hat{\mathbb{Z}}$ denote the profinite completion of $\mathbb{Z}$ and let

$$\chi : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \hat{\mathbb{Z}}^{\times}$$

denote the cyclotomic character. Since the archimedean part of $K_{\mathbb{A}}^{\times}$ is connected in our case, a homomorphism

$$\bar{r}_K : (K \otimes \hat{\mathbb{Z}})^{\times} \to \mathrm{Gal}(K^{\mathrm{ab}}/K)$$

can be defined such that the diagram

$$\begin{array}{ccc} K_{\mathbb{A}}^{\times} & \xrightarrow{r_K} & \mathrm{Gal}(K^{\mathrm{ab}}/K) \\ {\scriptstyle \text{forget components at } \infty} \downarrow & & \downarrow {\scriptstyle \sigma \mapsto \sigma^{-1}} \\ (K \otimes \hat{\mathbb{Z}})^{\times} & \xrightarrow[\bar{r}_K]{} & \mathrm{Gal}(K^{\mathrm{ab}}/K) \end{array}$$

commutes.

Let s be an automorphism of $\mathbb{C}$ and put $\sigma := s|_{\bar{\mathbb{Q}}} \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Select analytic isomorphisms

$$\lambda : \mathbb{C}^n/\tilde{\Phi}(\mathfrak{a}) \to A(\mathbb{C}) \quad \text{and} \quad \mu : \mathbb{C}^n/\tilde{\Phi}^s(\mathfrak{b}) \to A^s(\mathbb{C}),$$

where $\Phi^s \sim \sum_{i=1}^{n} \sigma_i s$ if $\Phi \sim \sum_{i=1}^{n} \sigma_i$. Let A_{tor} and A^s_{tor} denote the torsion subgroups of $A(\mathbb{C})$ and $A^s(\mathbb{C})$, respectively. Then there exists a unique $g \in (K \otimes \hat{\mathbb{Z}})^{\times}$ rendering the diagram

$$\begin{array}{ccc} K/\mathfrak{a} & \xrightarrow{\lambda\circ\tilde{\Phi}} & A_{\mathrm{tor}} \\ {\scriptstyle x\mapsto gx}\downarrow & & \downarrow{\scriptstyle a\mapsto a^s} \\ K/\mathfrak{b} & \xrightarrow[\mu\circ\tilde{\Phi}^s]{} & A^s_{\mathrm{tor}} \end{array}$$

commutative. It can be shown that modulo $K^\times \subsetneqq (K\otimes\hat{\mathbb{Z}})^\times$, g depends only upon K, Φ, and σ. Set

$$g_K(\sigma,\Phi) := gK^\times \in (K\otimes\hat{\mathbb{Z}})^\times/K^\times.$$

We consider Tate's half-transfer construction. For each $\varphi\in\Phi$, choose an element $w_\varphi\in\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ such that

$$\varphi = w_\varphi|_K.$$

Put $w_{\varphi\varrho} := w_\varphi\varrho$. Then for each embedding $\tau: K\to\bar{\mathbb{Q}}$, we have chosen an element $w_\tau\in\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ such that

$$\tau = w_\tau|_K \quad\text{and}\quad w_{\tau\varrho} = w_\tau\varrho \quad\text{for all } \tau.$$

Since we denote the Galois action by superscript, the order of the multiplication in $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is opposite to that in [1] and [3]. Then, according to Tate, for each $\sigma\in\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ and CM-type Φ of K, there exists a unique $F_K(\sigma,\Phi)\in\mathrm{Gal}(K^{\mathrm{ab}}/K)$ such that

$$F_K(\sigma,\Phi) \overset{\mathrm{def}}{\equiv} \prod_{\tau\in\Phi}(w_\tau\sigma w_{\tau\sigma}^{-1}) \bmod \mathrm{Gal}(\bar{\mathbb{Q}}/K^{\mathrm{ab}})$$

independent of the choice of a lifting w_φ for each $\varphi\in\Phi$ and the choice of an ordering of the product. Tate called this construction the half-transfer because

$$F_K(\sigma,\Phi)\,F_K(\sigma,\Phi^\varrho) = \mathrm{Ver}_{K/\mathbb{Q}}(\sigma),$$

where $\mathrm{Ver}_{K/\mathbb{Q}}:\mathrm{Gal}(\mathbb{Q}^{\mathrm{ab}}/\mathbb{Q})\to\mathrm{Gal}(K^{\mathrm{ab}}/K)$ denotes the transfer homomorphism. Tate proved that there exists a unique $f_K(\sigma,\Phi)\in(K\otimes\hat{\mathbb{Z}})^\times/K^\times$ such that

$$\begin{aligned} \big((\varrho\otimes 1)f_K(\sigma,\Phi)\big)f_K(\sigma,\Phi) &\equiv \chi(\sigma) \bmod K^\times, \\ \bar{r}_K\big(f_K(\sigma,\Phi)\big) &= F_K(\sigma,\Phi). \end{aligned}$$

The following result is Theorem 3.1 of Chapter 7 in Lang [3]:

Theorem 3.1. *There exists an element $e\in(F\otimes\hat{\mathbb{Z}})^\times$ which is uniquely determined (modulo ±1) by K, Φ, σ independently of A such that*

$$e^2 = 1 \quad\textit{and}\quad g_K(\sigma,\Phi) = f_K(\sigma,\Phi)e.$$

In particular A^s is isomorphic to $\mathbb{C}^n/\tilde{\Phi}^s(f\mathfrak{a})$, where $f = f_K(\sigma,\Phi)$.

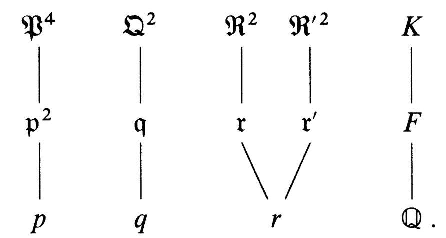
Remark 3.2. Tate conjectured that one can always take $e = 1$ and this was proved by Deligne in his letter to Tate, dated Oct. 8, 1981, using his theory of absolute Hodge cycles on abelian varieties. But full details of the proof are not yet extant.

§ 4. Main results and their proof

In this section, we restrict ourselves to the case in which A is a two-dimensional abelian variety. By changing the identification of K as a subfield of $\mathbb{C}$ if necessary, we may assume that Φ contains the identity. We note that the field of moduli of (A, C) coincides with the field of moduli $M_{\mathbb{Q}}$ of $(A, C, \theta|_{\mathbb{Q}})$. By 8.4, Example (2) in [6] (p. 73–p. 74), Proposition 5.17 in [4] (p. 131), and Theorem 2.1, we have

Lemma 4.1. *Assume that A is simple and the field of moduli of (A, C) is equal to $\mathbb{Q}$. Then, K is cyclic over $\mathbb{Q}$ and $I_0(\Phi') = I_K$. Conversely if K is cyclic over $\mathbb{Q}$, then A is simple.*

From now on, we assume that $K/\mathbb{Q}$ is cyclic. Hence we can write that $\Phi \sim 1 \oplus \sigma$ where σ is a generator of $\mathrm{Gal}(K/\mathbb{Q})$. Then $K' = K$, $\Phi' \sim 1 \oplus \sigma^3$. Since $K/\mathbb{Q}$ is quartic and cyclic, the following three diagrams show the only possibilities for the ramification of a prime in $K/\mathbb{Q}$:



type I type II type III

Let S_K (resp. T_K, U_K) be the set of prime numbers which ramify in $K/\mathbb{Q}$ as type I (resp. type II, type III).

Lemma 4.2. *Suppose $I_0(\Phi') = I_K$. Then $S_K = \{p\}$, $U_K = \emptyset$.*

Proof. Let $K = \mathbb{Q}(\zeta_5)$, where ζ_m denotes a primitive m-th root of unity. Since the class number h_K of K is equal to one in this case, $I_0(\Phi') = I_K$. On the other hand, by the theory of cyclotomic fields, we have $S_K = \{5\}$, $T_K = U_K = \emptyset$. So the lemma holds in this case.

Assume $K \neq \mathbb{Q}(\zeta_5)$. In this case the group of roots of unity in K coincides with $\{\pm 1\}$. Denote by E_K the unit group of K. By Dirichlet's unit theorem,

$$E_K = \{\pm 1\} \times \langle \varepsilon_0 \rangle \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}\,.$$

Since $\sigma^2 = \varrho$ on K, $\varepsilon_0^{\varrho} = \varepsilon_0$. So we have $E_K = E_F$. Let p be an element of S_K (S_K is nonempty by Minkowski's theorem). We have $\mathfrak{P}\mathfrak{P}^{\sigma^3} = \mathfrak{P}^2 = \mathfrak{p}\mathcal{O}_K$. Since $\mathfrak{P} \in I_0(\Phi')$, there exists $\alpha \in K^\times$ such that $\mathfrak{p}\mathcal{O}_K = (\alpha)$, $\alpha\alpha^{\varrho} = N(\mathfrak{P}) = p$. Then we have $(\alpha^{\varrho}) = (\alpha)$. Accordingly there exists $\varepsilon \in E_K = E_F$ such that $\varepsilon = \alpha/\alpha^{\varrho}$. Since $\varepsilon\varepsilon^{\varrho} = 1$ and $\varepsilon^{\varrho} = \varepsilon$, $\varepsilon^2 = 1$. So we get $\alpha^2 = \pm p$. Since F is the unique proper intermediate field in $K/\mathbb{Q}$, $F = \mathbb{Q}(\sqrt{p})$. Therefore $S_K = \{p\}$.

Assume that there exists an element r of U_K. By the same argument, we obtain that $\mathbb{Q}(\sqrt{r}) \subsetneqq K$. So $K = \mathbb{Q}(\sqrt{p}, \sqrt{r})$. This is a contradiction. Hence $U_K = \emptyset$. □

If $S_K = \{p\}$, then we have $F = \mathbb{Q}(\sqrt{p})$ with $p = 2$ or $p \equiv 1 \pmod 4$. From the genus theory of quadratic fields over $\mathbb{Q}$, it follows that the narrow class number h_F^+ of F is odd. Since $h_F^+ = h_F$ or $2h_F$, this implies that $h_F^+ = h_F$. Therefore we obtain that $N_{F/\mathbb{Q}}(\varepsilon_0) = -1$, where ε_0 is a fundamental unit of F, and h_F is odd. So the canonical map

$$C_F \to C_K, \quad \mathrm{cl}(\mathfrak{b}) \mapsto \mathrm{cl}(\mathfrak{b}\mathcal{O}_K)$$

is injective, where C_K (resp. C_F) denotes the ideal class group of K (resp. F).

Proposition 4.3. *Let p be a prime number such that $p \equiv 1 \pmod 4$. Let ε_0 be a fundamental unit of F such that $\varepsilon_0 > 0$. Then the following assertions hold*:

(a) *If $p \equiv 1 \pmod 8$, the extension $F(\sqrt{-\varepsilon_0\sqrt{p}})/F$ is ramified at both of the prime ideals of F lying above* (2)

(b) *If $p \equiv 5 \pmod 8$, the extension $F(\sqrt{-\varepsilon_0\sqrt{p}})/F$ is unramified at the prime ideal of F lying above* (2).

Proof. Assume that $p = 1 \pmod 8$. We have the decomposition $(2) = \mathfrak{q}\mathfrak{q}'$ in F, where $\mathfrak{q}$ and $\mathfrak{q}'$ are distinct prime ideals of F. We put $F_1 := F(\sqrt{-\varepsilon_0\sqrt{p}})$. To show that $\mathfrak{q}$ is ramified in F_1/F, suppose to the contrary that it is unramified; thus $\mathfrak{q}'$ is also unramified. By the theory of Kummer extensions the only primes of $\mathbb{Q}$ which ramify in $F_1/\mathbb{Q}$ are 2, p, and ∞, and we have assumed 2 is unramified, so the ramification occurs at p and ∞. Since p is prime to the degree of the extension $F_1/\mathbb{Q}$, the conductor $\mathfrak{c}(F_1/\mathbb{Q})$ of the abelian extension $F_1/\mathbb{Q}$ is $p\infty$. Therefore we have

$$F_1 \subseteqq \mathbb{Q}(\zeta_p).$$

Since $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ is a cyclic group of order $p-1$, for each positive divisor m of $p-1$, there exists a unique subgroup C_m of order m. Since $p \equiv 1 \pmod 8$,

$$(p-1)/4 \equiv 0 \pmod 2.$$

On the other hand, the order of $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/F_1)$ is equal to $(p-1)/4$. Therefore,

$$\mathrm{Gal}(\mathbb{Q}(\zeta_p)/F_1) \supseteqq C_2 = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}(\zeta_p)^+),$$

where $\mathbb{Q}(\zeta_p)^+$ denotes the maximal real subfield of $\mathbb{Q}(\zeta_p)$. So we have

$$F_1 \subseteqq \mathbb{Q}(\zeta_p)^+.$$

This is a contradiction.

Next, assume that $p \equiv 5 \pmod 8$. Then $\mathfrak{q} = (2)$ is a prime ideal of F. By the theory of Kummer extensions of prime degree, it holds that $\mathfrak{q}$ is unramified in F_1/F if and only if

(4.1) the equation: $X^2 \equiv -\varepsilon_0\sqrt{p} \pmod{\mathfrak{q}^2}$ has an integral solution in F.

Let

$$\psi : (\mathcal{O}_F/\mathfrak{q}^2)^\times \to (\mathcal{O}_F/\mathfrak{q})^\times$$

be the canonical projection. $(\mathcal{O}_F/\mathfrak{q})^\times$ is a cyclic group of order 3. For $a \pmod{\mathfrak{q}^2} \in (\mathcal{O}_F/\mathfrak{q}^2)^\times$, $a \pmod{\mathfrak{q}^2} \in \mathrm{Ker}(\psi)$ if and only if $x := (a-1)/2 \in \mathcal{O}_F$. So we have the isomorphism

$$\mathrm{Ker}(\psi) \to \mathcal{O}_F/\mathfrak{q} \cong (\mathbb{Z}/2\mathbb{Z})^2, \quad a \pmod{\mathfrak{q}^2} \mapsto x \pmod{\mathfrak{q}}.$$

Hence we have

$$(\mathcal{O}_F/\mathfrak{q}^2)^\times \cong (\mathbb{Z}/2\mathbb{Z})^2 \oplus \mathbb{Z}/3\mathbb{Z}.$$

So

$$(4.1) \text{ holds} \Leftrightarrow (-\varepsilon_0\sqrt{p})^3 \equiv 1 \pmod{\mathfrak{q}^2}.$$

Now we claim that $\mathfrak{q}$ ramifies in $F(\sqrt{\varepsilon_0})/F$. Assume that this is false. Let ∞_1 be the archimedean place of F corresponding to $F \subseteqq \mathbb{C}$ and let ∞_2 be the other one. Then $F(\sqrt{\varepsilon_0})/F$ is unramified at ∞_1 and ramified at ∞_2. So,

$$\mathfrak{c}\big(F(\sqrt{\varepsilon_0})/F\big) = \infty_2.$$

Therefore $F(\sqrt{\varepsilon_0})$ is included in the maximal ray class field H_{∞_2} over F with conductor ∞_2. In our case H_{∞_2} coincides with the Hilbert class field H_F of F. Hence $2|h_F$. This is a contradiction.

Similarly we get that $\mathfrak{q}$ ramifies in $F(\sqrt{-\varepsilon_0})/F$. These are equivalent to

$$\varepsilon_0^3 \not\equiv \pm 1 \pmod{\mathfrak{q}^2}.$$

Since $\varepsilon_0^3 \pmod{\mathfrak{q}^2} \in \mathrm{Ker}(\psi)$ and

$$\mathrm{Ker}(\psi) = \{\pm 1 \pmod{\mathfrak{q}^2}, \pm\sqrt{p} \pmod{\mathfrak{q}^2}\},$$

we have

$$\varepsilon_0^3 \equiv \sqrt{p} \quad \text{or} \quad -\sqrt{p} \pmod{\mathfrak{q}^2}.$$

Next we claim that $\mathfrak{q}$ ramifies in $F(\sqrt{\varepsilon_0\sqrt{p}})/F$. In fact, we assume that this is false. Then,

$$\mathfrak{c}\big(F(\sqrt{\varepsilon_0\sqrt{p}})/\mathbb{Q}\big) = p.$$

So we have

$$F(\sqrt{\varepsilon_0\sqrt{p}}) \subseteqq \mathbb{Q}(\zeta_p)^+.$$

Therefore, $(p-1)/2 \equiv 0 \pmod 4$. This contradicts to $p \equiv 5 \pmod 8$.

From this we obtain

$$\varepsilon_0\sqrt{p} \pmod{\mathfrak{q}^2} \notin ((\mathcal{O}_F/\mathfrak{q}^2)^\times)^2.$$

Now we assume that $\varepsilon_0^3 \equiv \sqrt{p} \pmod{\mathfrak{q}^2}$. Then, it holds that

$$\varepsilon_0 \sqrt{p} \equiv \varepsilon_0^4 \pmod{\mathfrak{q}^2} \in ((\mathcal{O}_F/\mathfrak{q}^2)^\times)^2$$

and this is a contradiction. So we have

$$\varepsilon_0^3 = -\sqrt{p} \pmod{\mathfrak{q}^2}.$$

Hence

$$(-\varepsilon_0 \sqrt{p})^3 \equiv 1 \pmod{\mathfrak{q}^2}. \quad \square$$

Proposition 4.4. *Let K be a quartic* CM*-field over $\mathbb{Q}$ and let F be the maximal real subfield of K. Put $T_K = \{q_1, \ldots, q_t\}$ $(t \geq 0)$. Then $K/\mathbb{Q}$ is cyclic, $S_K = \{p\}$, and $U_K = \emptyset$ if and only if K has one of the following representations*:

(I) $K = \mathbb{Q}(\sqrt{-q_1 \cdots q_t \varepsilon_0 \sqrt{p}})$, *where ε_0 is a fundamental unit of F such that $\varepsilon_0 > 0$. Moreover the prime numbers $p, q_1, \ldots, q_t$ satisfy one of the following conditions:*

(I_1) $p \equiv 1 \pmod 8$, $t \geqq 1$, $q_1 \cdots q_t \equiv 3 \pmod 4$, $\left(\frac{p}{q_i}\right) = -1$ $(i = 1, \ldots, t)$;

(I_2) $p \equiv 5 \pmod 8$. *Moreover if* $t \geqq 1$,

$$q_1 \cdots q_t \equiv 1 \pmod 4 \quad \textit{and} \quad \left(\frac{p}{q_i}\right) = -1 \quad (i = 1, \ldots, t);$$

(I_3) $p \equiv 5 \pmod 8$, $t \geqq 1$, $q_1 = 2$. *Moreover if* $t \geqq 2$,

$$\left(\frac{p}{q_i}\right) = -1 \quad (i = 2, \ldots, t);$$

(I_4) $p = 2$. *Moreover if* $t \geqq 1$,

$$q_i \equiv \pm 5 \pmod 8 \quad (i = 1, \ldots, t).$$

(II) $K = \mathbb{Q}(\sqrt{-q_2 \cdots q_t \varepsilon_0 \sqrt{p}})$:

$$p \equiv 5 \pmod 8,\ t \geqq 2,\ q_1 = 2,\ q_2 \cdots q_t \equiv 3 \pmod 4,\ \left(\frac{p}{q_i}\right) = -1 \quad (i = 2, \ldots, t).$$

Proof. ($\Rightarrow$). Let $K = \mathbb{Q}(\zeta_5)$. By a direct calculation, we can see that K has a representation as type (I_2). Assume that $K \neq \mathbb{Q}(\zeta_5)$. First we will show that

$$\langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t) \rangle \ (\subseteqq C_K) \cong (\mathbb{Z}/2\mathbb{Z})^t, \tag{4.2}$$

where $\mathfrak{Q}_i$ is the unique prime ideal of K lying above (q_i) $(i = 1, \ldots, t)$. Suppose that

$$\mathfrak{Q}_1^{\delta_1} \cdots \mathfrak{Q}_t^{\delta_t} = (\alpha), \quad \alpha \in K^\times, \quad \delta_i = 0 \text{ or } 1 \quad (i = 1, \ldots, t).$$

Then $(\alpha^2) = (q_1^{\delta_1} \cdots q_t^{\delta_t})$. So, $\alpha^2 = \pm \varepsilon_0^n q_1^{\delta_1} \cdots q_t^{\delta_t}$. We suppose that n is odd, i.e. $n = 2m+1$. Then we have

$$\left(\frac{\alpha}{\varepsilon_0^m}\right)^2 = \pm \varepsilon_0 q_1^{\delta_1} \cdots q_t^{\delta_t}.$$

Therefore we obtain $K = F(\sqrt{\pm \varepsilon_0 q_1^{\delta_1} \cdots q_t^{\delta_t}})$. Since none of $\pm \varepsilon_0 q_1^{\delta_1} \cdots q_t^{\delta_t}$ are totally negative, this contradicts the fact that K is totally imaginary. Hence we get that n is even, i.e. $n = 2m$. So,

$$\left(\frac{\alpha}{\varepsilon_0^m}\right)^2 = \pm q_1^{\delta_1} \cdots q_t^{\delta_t}.$$

Since $F = \mathbb{Q}(\sqrt{p})$ is the unique proper intermediate field of $K/\mathbb{Q}$, we especially have $\delta_1 = \cdots = \delta_t = 0$. Therefore we obtain (4.2). Put

$$C_0(K/\mathbb{Q}) = \{\mathrm{cl}(\mathfrak{b}) \mid \mathfrak{b} \in I_K \text{ s.t. } \mathfrak{b}^\sigma = \mathfrak{b}\} \quad (\subseteqq C_K),$$

where σ is a generator of $\mathrm{Gal}(K/\mathbb{Q})$. By (a.2) of Proposition A.1 in [5] (p. 81), $|C_0(K/\mathbb{Q})| = 2^t$. So we have

$$(4.3) \qquad C_0(K/\mathbb{Q}) = \langle \mathrm{cl}(\mathfrak{P}), \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle = \langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle,$$

where $\mathfrak{P}$ is the prime ideal of K lying above (p). Therefore it can be written that

$$\mathfrak{P} = \mathfrak{Q}_1^{\lambda_1} \cdots \mathfrak{Q}_t^{\lambda_t}(\alpha), \quad \alpha \in K^\times, \quad \lambda_i = 0 \text{ or } 1 \quad (i = 1, \ldots, t).$$

So,

$$(\sqrt{p}) = \mathfrak{p}\mathcal{O}_K = \mathfrak{P}^2 = (q_1^{\lambda_1} \cdots q_t^{\lambda_t} \alpha^2).$$

Hence we obtain

$$(q_1^{\lambda_1} \cdots q_t^{\lambda_t} \alpha)^2 = \pm \varepsilon_0^n q_1^{\lambda_1} \cdots q_t^{\lambda_t} \sqrt{p}.$$

As before we must have that $n = 2m+1$ and

$$\left(\frac{q_1^{\lambda_1} \cdots q_t^{\lambda_t} \alpha}{\varepsilon_0^m}\right)^2 = -\varepsilon_0 q_1^{\lambda_1} \cdots q_t^{\lambda_t} \sqrt{p}.$$

Therefore K can be written in the form

$$K = \mathbb{Q}(\sqrt{-q_1^{\lambda_1} \cdots q_t^{\lambda_t} \varepsilon_0 \sqrt{p}}).$$

From the theory of Kummer extensions, if $2 \notin \{q_1, \ldots, q_t\}$, we have

$$\lambda_1 = \cdots = \lambda_t = 1.$$

On the other hand, if $2 \in \{q_1, \ldots, q_t\}$ (i.e. $q_1 = 2$), we have

$$\lambda_1 = 0 \text{ or } 1, \quad \lambda_2 = \cdots = \lambda_t = 1.$$

If $p \equiv 1 \pmod 8$, then $2 \notin T_K$. In particular both of the prime ideals $\mathfrak{q}$ and $\mathfrak{q}'$ of F lying above (2) do not ramify in the extension $F(\sqrt{-q_1 \cdots q_t \varepsilon_0 \sqrt{p}})/F$. So the equation

$$X^2 \equiv -q_1 \cdots q_t \varepsilon_0 \sqrt{p} \pmod{\mathfrak{q}^2}$$

must have an integral solution in F. Since $\mathcal{O}_F/\mathfrak{q}^2 \cong \mathbb{Z}/4\mathbb{Z}$, $-q_1 \cdots q_t \varepsilon_0 \sqrt{p} \equiv 1 \pmod{\mathfrak{q}^2}$. On the other hand, from Proposition 4.3(a), it follows that $-\varepsilon_0 \sqrt{p} \equiv 3 \pmod{\mathfrak{q}^2}$. Therefore $q_1 \cdots q_t \equiv 3 \pmod{\mathfrak{q}^2}$. Hence we get that $t \geqq 1$ and $q_1 \cdots q_t \equiv 3 \pmod 4$. The other conditions in (I_1) follow from the fact that (q_i) remains prime in $\mathbb{Q}(\sqrt{p})$ $(i = 1, \ldots, t)$.

In the other cases we get the conditions by the same argument.

($\Leftarrow$). It can be easily shown that $\mathbb{Q}(\sqrt{-q_1 \cdots q_t \varepsilon_0 \sqrt{p}})$ is an imaginary cyclic quartic extension of $\mathbb{Q}$. In each case we can see that $S_K = \{p\}$, $T_K = \{q_1, \ldots, q_t\}$, and $U_K = \emptyset$ from the theory of Kummer extensions and Proposition 4.3. □

Proposition 4.5. *Let (K, Φ) be a* CM*-type such that $K/\mathbb{Q}$ is a cyclic quartic extension and $\Phi \sim 1 \oplus \sigma$, $\mathrm{Gal}(K/\mathbb{Q}) = \langle \sigma \rangle$. Let $T_K = \{q_1, \ldots, q_t\}$ $(t \geqq 0)$ and let $\mathfrak{Q}_i$ be the unique prime ideal of K lying above (q_i) $(i = 1, \ldots, t)$. Then the following are equivalent*:

(i) $I_0(\Phi') = I_K$.

(ii) $S_K = \{p\}$, $U_K = \emptyset$, $C_K = \langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle \oplus \bar{C}_F$, *where $\bar{C}_F$ is the image of the canonical map $C_F \to C_K$.*

(iii) $S_K = \{p\}$, $U_K = \emptyset$, $h_K = 2^t h_F$.

Proof. (i) $\Rightarrow$ (ii). From Lemma 4.2 we obtain that $S_K = \{p\}$ and $U_K = \emptyset$. Put

$$I_0(K/F) = P_K\{\mathfrak{b} \in I_K \mid \mathfrak{b}^\varrho = \mathfrak{b}\},$$

where P_K denotes the group of principal ideals in K. Then, $I_0(\Phi') \subseteqq I_0(K/F)$ (see the appendix in [5]). By the assumption,

$$C_K = C_0(K/F) = \langle \mathrm{cl}(\mathfrak{P}), \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t), \bar{C}_F\rangle,$$

where $\mathfrak{P}$ is the unique prime ideal of K lying above (p). From (4.3) and the fact that $|\bar{C}_F|$ is odd,

$$C_K = \langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle \oplus \bar{C}_F.$$

(ii) $\Rightarrow$ (i). For any $\mathfrak{b} \in I_F$, we put $\bar{\mathfrak{b}} := \mathfrak{b}\mathcal{O}_K \in I_K$. Then, $\bar{\mathfrak{b}}\bar{\mathfrak{b}}^{\sigma^3} = \mathfrak{b}\mathfrak{b}^\sigma \mathcal{O}_K = (N(\mathfrak{b}))$ and $N(\mathfrak{b})N(\mathfrak{b})^\varrho = N(\mathfrak{b})^2 = N(\bar{\mathfrak{b}})$. So we have $\bar{\mathfrak{b}} \in I_0(\Phi')$ for $\forall \mathfrak{b} \in I_F$. Since $\mathfrak{Q}_i\mathfrak{Q}_i^{\sigma^3} = \mathfrak{Q}_i^2 = (q_i)$ and $q_i q_i^\varrho = q_i^2 = N(\mathfrak{Q}_i)$ $(i = 1, \ldots, t)$, $\mathfrak{Q}_i \in I_0(\Phi')$. From the fact that $P_K \subseteqq I_0(\Phi')$, we obtain (i).

It is clear that (ii) $\Rightarrow$ (iii).

(iii) $\Rightarrow$ (ii). Assuming (iii), we have

$$C_K \supseteqq \langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle \oplus \bar{C}_F$$

and

$$\langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle \cong (\mathbb{Z}/2\mathbb{Z})^t, \quad \bar{C}_F \cong C_F.$$

Comparing the orders, we get (ii). □

Proposition 4.6. *Let (K, Φ) be as in Proposition 4.5. Let $P = (A, C, \theta)$ be of type $(K, \Phi; \eta, \mathfrak{a})$. Denote by M_K (resp. M) the field of moduli of (A, C, θ) (resp. that of (A, C)). Assume that $I_0(\Phi') = I_K$. Then:*

(a) $M_K = K$;

(b) $M = \mathbb{Q}$ *or* F.

Proof. The assertion (a) follows from Theorem 2.1. By the assumption we can put $D = D' = \mathbb{Q}$, D and D' being as in § 2. Since $C_K = C_0(K/F)$, it can be written that $\mathfrak{a} = (\alpha)\mathfrak{a}_0$, where $\alpha \in K^\times$ and $\mathfrak{a}_0 \in I_K$ such that $\mathfrak{a}_0^\varrho = \mathfrak{a}_0$. Then P is isomorphic to the structure of type $(K, \Phi; \alpha\alpha^\varrho\eta, \mathfrak{a}_0)$. From Proposition 4 in [5], we obtain $(P_\varrho : P) = 1$. By Proposition 2.2, $\varrho \in \mathrm{Gal}(K/M)$. Hence $M \subseteqq F$. So we get the assertion (b). □

Proposition 4.7. *The notation and assumptions being as in Proposition 4.6, suppose that $\mathfrak{f}(P)^\sigma = \mathfrak{f}(P)$. Let $T_K = \{q_1, \ldots, q_t\}$ $(t \geqq 0)$ and let $\mathfrak{Q}_i$ be the unique prime ideal of K lying above (q_i). Then $(P_\sigma : P)$ can be written uniquely in the form*

$$(P_\sigma : P) = (q_1^{\mu_1} \cdots q_t^{\mu_t}, \mathfrak{Q}_1^{\mu_t} \cdots \mathfrak{Q}_t^{\mu_t}), \quad \mu_i = 0 \text{ or } 1 \quad (i = 1, \ldots, t).$$

Proof. Take an element $\bar{\sigma} \in \mathrm{Aut}(\mathbb{C})$ such that $\bar{\sigma}|_K = \sigma$ and fix it. Let $P_{\bar{\sigma}}$ be of type $(K, \Phi; \eta', \mathfrak{a}')$. Put $(v, \mathfrak{b}) := (P_\sigma : P) \in \tilde{C}_K$. Therefore

$$N_{K/F}(\mathfrak{b}) = \mathfrak{b}\mathfrak{b}^\varrho = (\mathfrak{a}'^{-1}\mathfrak{a})(\mathfrak{a}'^{-1}\mathfrak{a})^\varrho = \mathfrak{a}\mathfrak{a}^\varrho(\mathfrak{a}'\mathfrak{a}'^\varrho)^{-1}. \tag{4.4}$$

On the other hand, it holds that $\mathfrak{f}(P_{\bar{\sigma}}) = \mathfrak{f}(P)$. In fact, putting $\mathfrak{f} = \mathfrak{f}(P)$ and taking a basic polar divisor X contained in C, we have

$$\begin{aligned}\{x^{\bar{\sigma}} \in A^{\bar{\sigma}} \mid \theta_{\bar{\sigma}}(\mathfrak{f})x^{\bar{\sigma}} = 0\} &= \{x^{\bar{\sigma}} \in A^{\bar{\sigma}} \mid \theta_{\bar{\sigma}}(\mathfrak{f}^\sigma)x^{\bar{\sigma}} = 0\}\\ &= \{x^{\bar{\sigma}} \in A^{\bar{\sigma}} \mid \theta^{\bar{\sigma}}(\mathfrak{f})x^{\bar{\sigma}} = 0\}\\ &= \{x \in A \mid \theta(\mathfrak{f})x = 0\}^{\bar{\sigma}}\\ &= (\mathrm{Ker}(\varphi_X))^{\bar{\sigma}}\\ &= \{y \in A \mid \varphi_X(y) = 0\}^{\bar{\sigma}}\\ &= \{y^{\bar{\sigma}} \in A^{\bar{\sigma}} \mid \varphi_{X^{\sigma}}(y^{\bar{\sigma}}) = 0\}\\ &= \mathrm{Ker}(\varphi_{X^{\sigma}}).\end{aligned}$$

Since $X^{\bar{\sigma}}$ is a basic polar divisor contained in $C^{\bar{\sigma}}$, we get the claim. Consequently $(\eta')\mathfrak{d}_K \mathfrak{a}'\mathfrak{a}'^{\varrho} = (\eta)\mathfrak{d}_K \mathfrak{a}\mathfrak{a}^{\varrho}$. So we have

$$\mathfrak{a}\mathfrak{a}^{\varrho}(\mathfrak{a}'\mathfrak{a}'^{\varrho})^{-1} = (\eta'\eta^{-1}) = (v). \tag{4.5}$$

By (4.4) and (4.5), we obtain $N_{K/F}(\mathfrak{b}) = (v)$. Since

$$\mathrm{Ker}(N_{K/F} : C_K \to C_F) = \langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle,$$

by multipling a suitable element $\{xx^{\varrho}, (x)\}$ $(x \in K^{\times})$ of $F_+ \times I_K$ we can uniquely write

$$\mathfrak{b} = \mathfrak{Q}_1^{\mu_1} \cdots \mathfrak{Q}_t^{\mu_t}, \quad \mu_i = 0 \text{ or } 1 \quad (i = 1, \ldots, t).$$

Then $(v) = N_{K/F}(\mathfrak{b}) = \mathfrak{b}^2 = (q_1^{\mu_1} \cdots q_t^{\mu_t})$. Since $v \in F_+$, it can be written that $v = \varepsilon_0^{2m} q_1^{\mu_1} \cdots q_t^{\mu_t}$. So

$$(v, \mathfrak{b})\big(\varepsilon_0^m(\varepsilon_0^m)^{\varrho}, (\varepsilon_0^m)\big)^{-1} = (q_1^{\mu_1} \cdots q_t^{\mu_t}, \mathfrak{Q}_1^{\mu_1} \cdots \mathfrak{Q}_t^{\mu_t}).$$

Therefore we can uniquely write

$$(P_{\sigma} : P) = (q_1^{\mu_1} \cdots q_t^{\mu_t}, \mathfrak{Q}_1^{\mu_1} \cdots \mathfrak{Q}_t^{\mu_t}). \quad \square$$

Proposition 4.8. *Let (K, Φ) be as in Proposition* 4.5.

(a) *Let $P_i = (A_i, C_i, \theta_i)$ be of type $(K, \Phi; \eta_i, \mathfrak{a}_i)$ with respect to λ_i $(i = 1, 2)$. Then A_1 is isomorphic to A_2 if and only if $\mathfrak{a}_2 = (\alpha)\mathfrak{a}_1$ for some $\alpha \in K^{\times}$.*

(b) *Let $P_3 = (A_3, C_3, \theta_3)$ be of type $(K, \Phi^{\sigma}; \eta_3, \mathfrak{a}_3)$. Then A_1 is isomorphic to A_3 if and only if $\mathfrak{a}_3 = (\beta)\mathfrak{a}_1^{\sigma^3}$ for some $\beta \in K^{\times}$.*

Proof. We shall prove assertion (a). A_1 is isomorphic to A_2 if and only if there exists a $\mathbb{C}$-linear isomorphism

$$\tilde{\iota} : \mathbb{C}^2 \to \mathbb{C}^2$$

such that $\tilde{\iota}(\tilde{\Phi}(\mathfrak{a}_1)) = \tilde{\Phi}(\mathfrak{a}_2)$. Set

$$\iota : A_1 \xrightarrow{\lambda_1^{-1}} \mathbb{C}^2/\tilde{\Phi}(\mathfrak{a}_1) \xrightarrow{\text{isomorphism induced by } \tilde{\iota}} \mathbb{C}^2/\tilde{\Phi}(\mathfrak{a}_2) \xrightarrow{\lambda_2} A_2$$

and define the $\mathbb{Q}$-algebra isomorphism

$$\iota^* : \mathrm{End}^0(A_2) \to \mathrm{End}^0(A_1), \quad f \mapsto \iota^{-1} \circ f \circ \iota.$$

So there exists $\bar{\iota} \in \mathrm{Gal}(K/\mathbb{Q})$ such that the diagram

$$\begin{array}{ccc} K & \xrightarrow{\bar{\iota}} & K \\ \theta_2 \downarrow & & \downarrow \theta_1 \\ \mathrm{End}^0(A_2) & \xrightarrow[\iota^*]{} & \mathrm{End}^0(A_1) \end{array}$$

commutes. We can write $\bar{\imath} = \sigma^{-i}$ $(0 \leqq i \leqq 3)$. There exists $\alpha \in K^{\times}$ such that $\tilde{\imath}(\tilde{\Phi}(1)) = \tilde{\Phi}(\alpha)$. Then for any $a \in K$,

$$\begin{aligned}\tilde{\imath}(\tilde{\Phi}(a)) = \tilde{\imath}(S_{\Phi}(a)(\tilde{\Phi}(1))) &= (\tilde{\imath} \circ S_{\Phi}(a))(\tilde{\Phi}(1)) \\ &= (S_{\Phi}(a^{\sigma^i}) \circ \tilde{\imath})(\tilde{\Phi}(1)) \\ &= S_{\Phi}(a^{\sigma^i})(\tilde{\Phi}(\alpha)) \\ &= \tilde{\Phi}(\alpha a^{\sigma^i}).\end{aligned}$$

Therefore, defining the $\mathbb{R}$-linear isomorphism

$$(\sigma^i \otimes 1)^{\sim} : \mathbb{C}^2 \to \mathbb{C}^2$$

by the following commutative diagram

$$\begin{array}{ccc} K \otimes_{\mathbb{Q}} \mathbb{R} & \xrightarrow{\sigma^i \otimes 1} & K \otimes_{\mathbb{Q}} \mathbb{R} \\ {\scriptstyle \tilde{\Phi}} \downarrow & & \downarrow {\scriptstyle \tilde{\Phi}} \\ \mathbb{C}^2 & \xrightarrow[(\sigma^i \otimes 1)^{\sim}]{} & \mathbb{C}^2, \end{array}$$

we have $\tilde{\imath} = S_{\Phi}(\alpha) \circ (\sigma^i \otimes 1)^{\sim}$. Then it holds that

$$\begin{aligned}\tilde{\imath} \text{ is } \mathbb{C}\text{-linear} &\Leftrightarrow (\sigma^i \otimes 1)^{\sim} \text{ is } \mathbb{C}\text{-linear} \\ &\Leftrightarrow i = 0.\end{aligned}$$

So assertion (a) has been proved.

We can prove assertion (b) by the same argument. □

From now on, we assume the situation in Proposition 4.7. Let $\bar{\sigma} \in \operatorname{Aut}(\mathbb{C})$ be as in the proof of Proposition 4.7. Put $\tilde{\sigma} := \bar{\sigma}|_{\bar{\mathbb{Q}}}$. From (a) of Proposition 4.8, it follows that

$$\begin{aligned}(4.6) \qquad A \cong A^{\bar{\sigma}} &\Leftrightarrow \mu_1 = \cdots = \mu_t = 0 \quad \text{in Proposition 4.7} \\ &\Leftrightarrow (P_{\sigma} : P) = 1.\end{aligned}$$

On the other hand, by Theorem 3.1 and (b) of Proposition 4.8, we obtain

$$\begin{aligned}(4.7) \qquad A \cong A^{\bar{\sigma}} &\Leftrightarrow \operatorname{cl}(f\mathfrak{a}) = \operatorname{cl}(\mathfrak{a}^{\sigma^3}), \\ &\qquad \text{where } f = f_K(\tilde{\sigma}, \Phi) \in (K \otimes \hat{\mathbb{Z}})^{\times}/K^{\times} \\ &\Leftrightarrow \operatorname{cl}(il(f)) = \operatorname{cl}(\mathfrak{a}^{-2}), \\ &\qquad \text{where } il(f) \text{ denotes the fractional ideal in } K \text{ associated to } f\end{aligned}$$

$$\Leftrightarrow \; r_K(f)|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^{-2}}\right),$$

where H_K denotes the Hilbert class field of K

and $\left(\frac{H_K/K}{\mathfrak{b}}\right)$ denotes the Artin symbol associated to $\mathfrak{b} \in I_K$

$$\Leftrightarrow \; \bar{r}_K(f)|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^{2}}\right)$$

$$\Leftrightarrow \; F_K(\tilde{\sigma}, \Phi)|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^{2}}\right)$$

$$\Leftrightarrow \; \tilde{\sigma}^2 \varrho^{-1}|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^{2}}\right).$$

We put $C_1 := \langle \mathrm{cl}(\mathfrak{Q}_1), \ldots, \mathrm{cl}(\mathfrak{Q}_t)\rangle$ and $C_2 := \bar{C}_F$. By the Artin map, we can identify $C_K = C_1 \oplus C_2$ with $\mathrm{Gal}(H_K/K)$. Let H_1 be the fixed field of C_2 and H_2 the fixed field of C_1. Since $K/\mathbb{Q}$ is Galois, $H_K/\mathbb{Q}$ is also Galois.

Lemma 4.9. *$H_1/\mathbb{Q}$ is abelian and $H_2/\mathbb{Q}$ is Galois.*

Proof. Set $\sigma' := \tilde{\sigma}|_{H_K}$ ($\in \mathrm{Gal}(H_K/\mathbb{Q})$). Then

$$\mathrm{Gal}(H_K/\mathbb{Q}) = \langle \sigma', \mathrm{Gal}(H_K/K)\rangle = \langle \sigma', C_1, C_2\rangle.$$

For any $\mathrm{cl}(\mathfrak{b}) \in C_K$,

$$\sigma'^{-1}\left(\frac{H_K/K}{\mathfrak{b}}\right)\sigma' = \left(\frac{H_K/K}{\mathfrak{b}^{\sigma}}\right).$$

Since $C_i^{\sigma} = C_i$, we get $\sigma'^{-1} C_i \sigma' = C_i$ ($i = 1, 2$). Therefore $H_i/\mathbb{Q}$ is Galois ($i = 1, 2$). It follows that

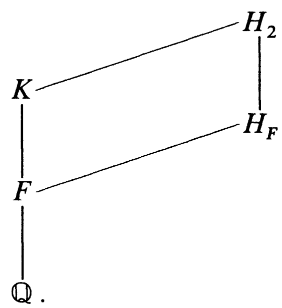
$$\mathrm{Gal}(H_1/\mathbb{Q}) \cong \langle \sigma', C_1, C_2\rangle / C_2.$$

For any $\mathrm{cl}(\mathfrak{b}_1) \in C_1$,

$$\begin{aligned}\sigma'^{-1}\left(\frac{H_K/K}{\mathfrak{b}_1}\right)\sigma'\left(\frac{H_K/K}{\mathfrak{b}_1}\right)^{-1} &= \left(\frac{H_K/K}{\mathfrak{b}_1^{\sigma}}\right)\left(\frac{H_K/K}{\mathfrak{b}_1^{-1}}\right)\\ &= \left(\frac{H_K/K}{\mathfrak{b}_1}\right)\left(\frac{H_K/K}{\mathfrak{b}_1^{-1}}\right)\\ &= 1.\end{aligned}$$

So $H_1/\mathbb{Q}$ is abelian. □

Now we have the following Hasse diagram:



In fact, since F is totally real, H_F is also totally real. So $H_F \cap K = F$. The following commutative diagram is well known:

$$
\begin{array}{ccc}
C_K & \xrightarrow{\text{Artin map}} & \mathrm{Gal}(H_K/K) \\
{\scriptstyle N_{K/F}}\downarrow & & \downarrow{\scriptstyle \text{restriction}} \\
C_F & \xrightarrow[\text{Artin map}]{} & \mathrm{Gal}(H_F/F)\,.
\end{array}
\tag{4.8}
$$

For any $\mathrm{cl}(\mathfrak{b}_1) \in C_1$, $N_{K/F}(\mathfrak{b}_1) \in P_F$. Consequently

$$\left(\frac{H_K/K}{\mathfrak{b}_1}\right)\bigg|_{H_F} = 1\,.$$

Therefore we have $H_F \subseteqq H_2$. Since $[H_2 : H_F] = 2$, $H_2 = \langle K, H_F \rangle$. Thus we get the diagram.

We can write $\mathrm{cl}(\mathfrak{a})$ uniquely in the form

$$\mathrm{cl}(\mathfrak{a}) = \mathrm{cl}(\mathfrak{a}_1)\,\mathrm{cl}(\bar{\mathfrak{a}}_2) \quad (\mathrm{cl}(\mathfrak{a}_1) \in C_1,\ \mathrm{cl}(\bar{\mathfrak{a}}_2) \in C_2,\ \bar{\mathfrak{a}}_2 = \mathfrak{a}_2 \mathcal{O}_K,\ \mathfrak{a}_2 \in I_F)\,,$$

where $\mathfrak{a}$ is the same ideal which occurs in the definition of $P = (A, C, \theta)$ being of type $(K, \Phi; \eta, \mathfrak{a})$. Then from the facts $H_K = \langle H_1, H_F \rangle$, $\mathrm{cl}(\mathfrak{a})^2 = \mathrm{cl}(\bar{\mathfrak{a}}_2^2)$ and the commutative diagram (4.8), it follows that

$$\tilde{\sigma}^2 \varrho^{-1}|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^2}\right)$$

if and only if the following two conditions are satisfied:

$$\tilde{\sigma}^2 \varrho^{-1}|_{H_1} = 1\,, \tag{4.9}$$

$$\tilde{\sigma}^2 \varrho^{-1}|_{H_F} = \left(\frac{H_F/F}{\mathfrak{a}_2^4}\right). \tag{4.10}$$

Proposition 4.10. *Under the assumptions in Proposition* 4.7, *condition* (4.10) *always holds.*

Proof. Since $\varrho|_{H_F} = 1$, $\tilde{\sigma}^2 \varrho^{-1}|_{H_F} = \tilde{\sigma}^2|_{H_F}$. So there exists a unique element $\mathrm{cl}(\mathfrak{b}) \in C_F$ such that

$$(\tilde{\sigma}|_{H_F})^2 = \left(\frac{H_F/F}{\mathfrak{b}}\right).$$

Put $\sigma_1 := \tilde{\sigma}|_{H_F}$ ($\in \mathrm{Gal}(H_F/\mathbb{Q})$). Therefore

$$\begin{aligned}\left(\frac{H_F/F}{\mathfrak{b}}\right) &= \sigma_1^2\\ &= \sigma_1^{-1}\sigma_1^2\sigma_1\\ &= \left(\frac{H_F/F}{\mathfrak{b}^\sigma}\right).\end{aligned}$$

Consequently we have $\mathrm{cl}(\mathfrak{b})^2 = 1$ in C_F. Since h_F is odd, $\mathrm{cl}(\mathfrak{b}) = 1$ in C_F. Hence $\sigma_1^2 = 1$.

On the other hand, from the assumption $\mathfrak{f}(P)^\sigma = \mathfrak{f}(P)$, it follows that

$$\mathrm{cl}(\mathfrak{d}_K \mathfrak{a}\mathfrak{a}^\varrho) \in C_0(K/\mathbb{Q}) = C_1.$$

Clearly $\mathfrak{d}_K^\sigma = \mathfrak{d}_K$. So we have

$$\mathrm{cl}(\mathfrak{a}\mathfrak{a}^\varrho) = \mathrm{cl}(\bar{\mathfrak{a}}_2)^2 \in C_1 \cap C_2 = \{1\}.$$

Therefore $\mathrm{cl}(\bar{\mathfrak{a}}_2) = 1$ in C_K. Hence we obtain $\mathrm{cl}(\mathfrak{a}_2) = 1$ in C_F. □

Proposition 4.11. *Assume the situation in Proposition* 4.7. *Let* $S_K = \{p\}$ *and* $T_K = \{q_1, \dots, q_t\}$ ($t \geqq 0$). *Then condition* (4.9) *holds if and only if* K *has the expression*

$$K = \mathbb{Q}(\sqrt{-q_1 \cdots q_t \varepsilon_0 \sqrt{p}}),$$

where ε_0 *is a fundamental unit of* F *such that* $\varepsilon_0 > 0$ *and the distinct prime numbers* $p, q_1, \dots, q_t$ *satisfy one of the following conditions*:

(I) $p \equiv 5 \pmod 8$. *Moreover if* $t \geqq 1$,

$$q_i \equiv 1 \pmod 4 \quad \textit{and} \quad \left(\frac{p}{q_i}\right) = -1 \quad (i = 1, \dots, t);$$

(II) $p \equiv 5 \pmod 8$, $t \geqq 1$, $q_1 = 2$. *Moreover if* $t \geqq 2$,

$$q_i \equiv 1 \pmod 4 \quad \textit{and} \quad \left(\frac{p}{q_i}\right) = -1 \quad (i = 2, \dots, t);$$

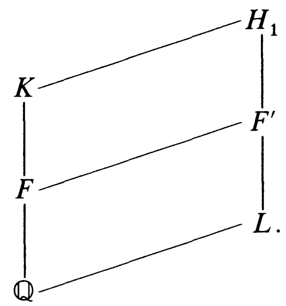
(III) $p = 2$. *Moreover if* $t \geqq 1$,

$$q_i \equiv 5 \pmod 8 \quad (i = 1, \dots, t).$$

Proof. We note that H_1 is a CM-field. Put $\sigma_2 := \tilde{\sigma}|_{H_1}$ $(\in \mathrm{Gal}(H_1/\mathbb{Q}))$. Then

$$\sigma_2^2(\varrho|_{H_1})^{-1} \in \mathrm{Gal}(H_1/K) \cong (\mathbb{Z}/2\mathbb{Z})^t.$$

Consequently $\sigma_2^4 = 1$. Therefore the order of σ_2 is equal to 4. Let L be the fixed field of $\langle\sigma_2\rangle$ and F' the fixed field of $\langle\sigma_2^2\rangle$. We have the following Hasse diagram:



So we obtain

$$\mathrm{Gal}(F'/\mathbb{Q}) \cong \mathrm{Gal}(F/\mathbb{Q}) \oplus \mathrm{Gal}(L/\mathbb{Q}) \cong (\mathbb{Z}/2\mathbb{Z})^{t+1}.$$

Clearly

$$\begin{aligned} (4.9) \text{ holds } &\Leftrightarrow \sigma_2^2 = \varrho|_{H_1} \\ &\Leftrightarrow F' \text{ is the maximal real subfield of } H_1. \end{aligned}$$

Now the proof will be divided into two parts: first, for $2 \notin T_K$, then for $2 \in T_K$.

First part: the case $2 \notin T_K$. Take a prime ideal $\mathfrak{P}$ of H_1 lying above (p) and fix it. Set

$$T := \{\xi \in \mathrm{Gal}(H_1/\mathbb{Q}) \mid x^\xi \equiv x \pmod{\mathfrak{P}}\ \forall x \in \mathcal{O}_{H_1}\},$$

the inertia group of $\mathfrak{P}$. Let L_1 be the fixed field of T which is called the inertia field of $\mathfrak{P}$. Since the ramification index of (p) in the extension $H_1/\mathbb{Q}$ is 4, we have $|T| = 4$.

Claim 1. $T \cong \mathbb{Z}/4\mathbb{Z}$.

Proof of Claim 1. Clearly

$$\mathrm{Gal}(H_1/\mathbb{Q}) = \mathrm{Gal}(H_1/K) \oplus \mathrm{Gal}(H_1/L), \quad \mathrm{Gal}(H_1/L) = \langle\sigma_2\rangle.$$

Now we assume that $T \cong (\mathbb{Z}/2\mathbb{Z})^2$. So T has two elements v_1 and v_2 of order 2. These can be written uniquely in the form:

$$v_1 = \tau_1\sigma_2^i, \quad v_2 = \tau_2\sigma_2^j \quad (i, j \in \{0, 2\}),$$

where τ_1, $\tau_2 \in \mathrm{Gal}(H_1/K)$. If $(i,j) \neq (2,2)$, then v_1 or $v_2 \in \mathrm{Gal}(H_1/K)$. If $(i,j) = (2,2)$, then $v_1^{-1}v_2 = \tau_1^{-1}\tau_2 \in \mathrm{Gal}(H_1/K)$. Therefore

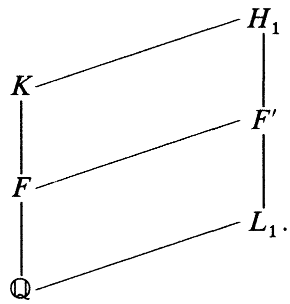
$$|T \cap \mathrm{Gal}(H_1/K)| \geqq 2.$$

So $\mathfrak{P}$ ramifies in H_1/K and this is a contradiction. □

By Claim 1,

$$T = \langle v \rangle, \quad v = \tau\sigma_2^i \quad (i = 1 \text{ or } 3,\ \tau \in \mathrm{Gal}(H_1/K)).$$

Since $v^2 = \sigma_2^2$, we have the diagram:



The set of prime numbers which ramify in $L_1/\mathbb{Q}$ is contained in T_K. Since $2 \notin T_K$ and $[L_1 : \mathbb{Q}] = 2^t$, we have

$$\mathfrak{c}_0(L_1/\mathbb{Q}) \mid q_1 \cdots q_t,$$

where $\mathfrak{c}_0(L_1/\mathbb{Q})$ denotes the finite part of the conductor of the abelian extension $L_1/\mathbb{Q}$. Therefore

$$L_1 \subseteqq \mathbb{Q}(\zeta_{q_1 \dots q_t}).$$

Put $n := q_1 \cdots q_t$. For any subfield M of $\bar{\mathbb{Q}}$, we set

$$V(M) := \{N \mid N \text{ is a quadratic field over } \mathbb{Q} \text{ s.t. } N \subseteqq M\}.$$

Clearly $V(L_1) \subseteqq V(\mathbb{Q}(\zeta_n))$. Comparing the orders, we have

$$V(L_1) = V(\mathbb{Q}(\zeta_n)).$$

On the other hand, since $\mathrm{Gal}(L_1/\mathbb{Q}) \cong (\mathbb{Z}/2\mathbb{Z})^t$,

$$L_1 = \langle N \mid N \in V(L_1) \rangle.$$

Therefore

$$\begin{aligned} F' \text{ is the maximal real subfield} &\Leftrightarrow L_1 \text{ is totally real} \\ &\Leftrightarrow N \text{ is real } \forall N \in V(L_1) \\ &\Leftrightarrow N \text{ is real } \forall N \in V(\mathbb{Q}(\zeta_n)) \\ &\Leftrightarrow \mathbb{Q}\left(\sqrt{(-1)^{\frac{q_i-1}{2}} q_i}\right) \text{ is real } (i = 1, \dots, t) \\ &\Leftrightarrow q_i \equiv 1 \ (\mathrm{mod}\ 4)\ (i = 1, \dots, t). \end{aligned}$$

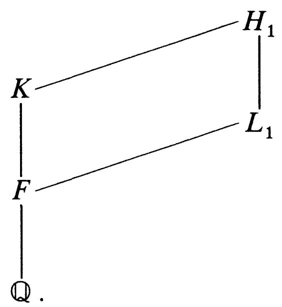
From Proposition 4.4 we obtain that

$$\text{(4.9) holds} \Leftrightarrow K = \mathbb{Q}(\sqrt{-q_1 \cdots q_t \varepsilon_0 \sqrt{p}}):$$
$$p, q_1, \ldots, q_t \text{ satisfy condition (I) or (III).}$$

Second part: the case $2 \in T_K$. Let $q_1 = 2$. Since (2) remains prime in $F = \mathbb{Q}(\sqrt{p})$, it holds that $p \equiv 5 \pmod 8$. Take a prime ideal $\tilde{\mathfrak{Q}}_1$ of H_1 lying above (2) and fix it. Let T be the inertia group of $\tilde{\mathfrak{Q}}_1$ (in $H_1/\mathbb{Q}$) and let L_1 be the fixed field of it. The ramification index of (2) in $H_1/\mathbb{Q}$ is 2, so we have $|T| = 2$. Therefore

$$T = \langle \nu \rangle, \quad \nu = \tau\sigma_2^2 \quad (\tau \in \mathrm{Gal}(H_1/K)).$$

So we have the diagram:



In this case K has an expression of type (I_3) or type (II) in Proposition 4.4.

Claim 2. If F' (which by definition is the fixed field of $\langle \sigma_2^2 \rangle$ in H_1) is the maximal real subfield of H_1, then K is not of type (II) in Proposition 4.4.

Proof of Claim 2. Assume that K has an expression of type (II). Put

$$\alpha := \sqrt{-q_2 \cdots q_t \varepsilon_0 \sqrt{p}} \in K.$$

Since $(\alpha, \tilde{\mathfrak{Q}}_1) = 1$ and

$$\alpha^\nu = \alpha^{\tau\sigma_2^2} = \alpha^\varrho = -\alpha,$$

we have $\tilde{\mathfrak{Q}}_1^2 \,\|\, (\alpha^\nu - \alpha)$. Therefore, letting

$$V_m := \{\xi \in \mathrm{Gal}(H_1/L_1) \,|\, x^\xi \equiv x \pmod{\tilde{\mathfrak{Q}}_1^{m+1}} \ \forall x \in \mathcal{O}_{H_1}\}$$

($m \geqq 0$) be the ramification group of degree m of $\tilde{\mathfrak{Q}}_1$ in $H_1/\mathbb{Q}$, we have $V_m = \{1\}$ if $m \geqq 2$. $V := V_1$ is simply called the ramification group of $\tilde{\mathfrak{Q}}_1$ and it is well known that T/V is a cyclic group whose order divides $N(\tilde{\mathfrak{Q}}_1) - 1$ which is odd. Since $|T| = 2$, we obtain

$$T = V = V_1 \supsetneqq V_2 = \{1\}.$$

Therefore, by Hasse's conductor formula, the 2-exponent of the conductor $\mathfrak{c}(H_1/\mathbb{Q})$ is 2. So we have

$$\mathfrak{c}(H_1/\mathbb{Q}) = 2^2 p q_2 \cdots q_t \infty .$$

Hence, $F' \subseteqq H_1 \subseteqq \mathbb{Q}(\zeta_{2^2 p q_2 \dots q_t})$. Comparing the orders, we have

$$V(F') = V(\mathbb{Q}(\zeta_{2^2 p q_2 \dots q_t})) .$$

Since $\mathbb{Q}(\sqrt{-1}) \subseteqq \mathbb{Q}(\zeta_{2^2 p q_2 \dots q_t})$, $\mathbb{Q}(\sqrt{-1}) \subseteqq F'$. So this contradicts to that F' is totally real. □

Now suppose K has an expression of type I_3 in Proposition 4.4. Set

$$\beta := \sqrt{-2 q_2 \cdots q_t \varepsilon_0 \sqrt{p}}$$

and let all notations be as in the diagram above. Since $\tilde{\mathfrak{Q}}_1 \| (\beta)$ and $\tilde{\mathfrak{Q}}_1^3 \| (\beta^\vee - \beta)$, it follows that

$$T = V = V_1 = V_2 \supsetneqq V_3 = \{1\} .$$

By Hasse's conductor formula,

$$\mathfrak{c}(H_1/\mathbb{Q}) = 2^3 p q_2 \cdots q_t \infty .$$

We note that

$$V(\mathbb{Q}(\zeta_{2^3})) = \{\mathbb{Q}(\sqrt{-1}), \mathbb{Q}(\sqrt{-2}), \mathbb{Q}(\sqrt{2})\} .$$

Claim 3. $L_1 \neq F'$.

Proof of Claim 3. Suppose that $L_1 = F'$. Then

$$\mathfrak{c}_0(F'/\mathbb{Q}) | p q_2 \cdots q_t$$

so

$$F' \subseteqq \mathbb{Q}(\zeta_{p q_2 \dots q_t}), \quad \text{and} \quad V(F') \subseteqq V(\mathbb{Q}(\zeta_{p q_2 \dots q_t})).$$

Since $|V(F')| = 2^{t+1} - 1$ and $|V(\mathbb{Q}(\zeta_{p q_2 \dots q_t}))| = 2^t - 1$, this is a contradiction.

Put

$$F'' := F' \cap L_1 .$$

By Claim 3, $[F' : F''] = [L_1 : F''] = 2$. So

$$\mathrm{Gal}(F''/\mathbb{Q}) \cong (\mathbb{Z}/2\mathbb{Z})^t .$$

Therefore we can get the following claim by the same method as in the first part.

Claim 4. F'' is totally real if and only if $q_i \equiv 1 \pmod 4$ $(i = 2, \dots, t)$. Moreover, if this holds, we have

$$V(F'') = \{\mathbb{Q}(\sqrt{m}) | m > 0, \, m \text{ divides } p q_2 \cdots q_t\} .$$

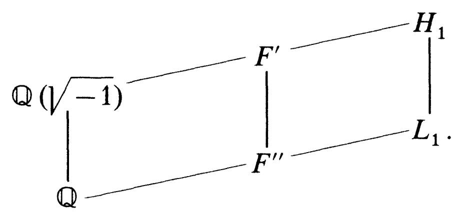
Claim 5. F' is the maximal real subfield if and only if F'' is totally real.

Proof of Claim 5. It is clear that the former property implies the latter.

Conversely suppose the latter. It holds that

$$F' = F''(\sqrt{-1}), \quad F''(\sqrt{-2}), \quad \text{or} \quad F''(\sqrt{2}).$$

Now suppose that $F' = F''(\sqrt{-1})$. Then we have the following diagram:



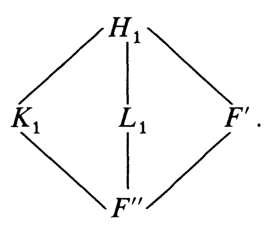
We have $\mathfrak{Q}_1 \parallel (1+\sqrt{-1})$. On the other hand, v does not fix $1+\sqrt{-1}$ because of $H_1 = L_1(1+\sqrt{-1})$. So $(1+\sqrt{-1})^v = 1-\sqrt{-1}$ and

$$\mathfrak{Q}_1^2 \parallel ((1+\sqrt{-1})^v - (1+\sqrt{-1})).$$

From Hasse's conductor formula, the 2-exponent of the conductor $\mathfrak{f}(H_1/\mathbb{Q})$ is 2 and this is a contradiction. Consequently we have

$$F' = F''(\sqrt{-2}) \quad \text{or} \quad F''(\sqrt{2}).$$

Set $K_1 := \langle K, F'' \rangle$. Then we have the diagram:



Since

$$K_1 = F''(\sqrt{-2q_2 \cdots q_t \varepsilon_0 \sqrt{p}}) = F''(\sqrt{-2\varepsilon_0 \sqrt{p}}),$$

it holds that

$$F' = F''(\sqrt{-2}) \Leftrightarrow L_1 = F''(\sqrt{\varepsilon_0 \sqrt{p}})$$

and

$$F' = F''(\sqrt{2}) \Leftrightarrow L_1 = F''(\sqrt{-\varepsilon_0 \sqrt{p}}).$$

Take a prime ideal $\mathfrak{r}$ of F'' lying above (2) and fix it. Then $\mathfrak{r}$ ramifies in either $F''(\sqrt{\varepsilon_0 \sqrt{p}})/F''$ or $F''(\sqrt{-\varepsilon_0 \sqrt{p}})/F''$. In fact, suppose that $\mathfrak{r}$ is unramified in both the extensions. Then $\mathfrak{r}$ is unramified in F'''/F'', where $F''' := F''(\sqrt{\varepsilon_0\sqrt{p}}, \sqrt{-\varepsilon_0\sqrt{p}})$. Since (2) is unramified in $F''/\mathbb{Q}$, (2) is also unramified in $F'''/\mathbb{Q}$. On the other hand, (2) ramifies in $F'''/\mathbb{Q}$ because of $F''' \supseteqq \mathbb{Q}(\sqrt{-1})$. This is a contradiction. Combing this with the two facts: $L_1 = F''(\sqrt{\varepsilon_0\sqrt{p}})$ or $F''(\sqrt{-\varepsilon_0\sqrt{p}})$; $\mathfrak{r}$ is unramified in L_1/F'' (by the definition of L_1), we obtain that

$$L_1 = F''(\sqrt{-\varepsilon_0\sqrt{p}}) \Leftrightarrow \mathfrak{r} \text{ is unramified in } F''(\sqrt{-\varepsilon_0\sqrt{p}})/F''.$$

Therefore

$$\begin{aligned} F' \text{ is the maximal real subfield} &\Leftrightarrow F' = F''(\sqrt{2}) \\ &\Leftrightarrow \mathfrak{r} \text{ is unramified in } F''(\sqrt{-\varepsilon_0\sqrt{p}})/F'' \\ &\Leftrightarrow \text{the prime ideal of } F \text{ lying above (2)} \\ &\qquad \text{is unramified in } F(\sqrt{-\varepsilon_0\sqrt{p}})/F. \end{aligned}$$

From (b) of Proposition 4.3, the last condition is true. So we get the assertion. □

From Claim 4, Claim 5, and Proposition 4.4, we have

$$\begin{aligned} (4.9) \text{ holds} \Leftrightarrow K = \mathbb{Q}(\sqrt{-q_1\cdots q_t\varepsilon_0\sqrt{p}})&: \\ p, q_1, \ldots, q_t \text{ satisfy condition (II)}&. \end{aligned}$$

This completes the proof of Proposition 4.11. □

From the propositions in this section and the facts stated in § 2, we get the main theorem.

Theorem 4.12. *Let K be a* CM*-field with* $[K:\mathbb{Q}] = 4$. *Let $P = (A, C, \theta)$ be a structure formed by a two-dimensional abelian variety A defined over $\mathbb{C}$, a polarization C of A, and an injection θ of K into* $\mathrm{End}^0(A)$ *such that* $\theta^{-1}(\mathrm{End}(A)) = \mathcal{O}_K$. *Assume that $\theta(K)$ is stable under the involution of* $\mathrm{End}^0(A)$ *determined by C. Let $\mathfrak{f}$ be the ideal of $\mathcal{O}_F$ which is uniquely determined by property* (2.1) *in* § 2.

Then A is simple and the field of moduli of (A, C) coincides with $\mathbb{Q}$ if and only if the following three conditions hold:

(a) *K has an expression of the form*

$$K = \mathbb{Q}(\sqrt{-q_1\cdots q_t\varepsilon_0\sqrt{p}}) \quad (t \geqq 0),$$

where ε_0 is a fundamental unit of $F = \mathbb{Q}(\sqrt{p})$ such that $\varepsilon_0 > 0$ and $p, q_1, \ldots, q_t$ are distinct prime numbers which satisfy one of the following conditions:

(a_1) $p \equiv 5 \pmod 8$. *Moreover if $t \geqq 1$,*

$$q_i \equiv 1 \pmod 4 \quad \textit{and} \quad \left(\frac{p}{q_i}\right) = -1 \quad (i = 1, \ldots, t);$$

(a_2) $p \equiv 5 \pmod 8$, $t \geqq 1$, $q_1 = 2$. *Moreover if* $t \geqq 2$,

$$q_i \equiv 1 \pmod 4 \quad \textit{and} \quad \left(\frac{p}{q_i}\right) = -1 \quad (i = 2, \ldots, t);$$

(a_3) $p = 2$. *Moreover if* $t \geqq 1$,

$$q_i \equiv 5 \pmod 8 \quad (i = 1, \ldots, t).$$

(b) $h_K = 2^t h_F$, *where* h_M *denotes the class number of an algebraic numer field* M.

(c) $\mathfrak{f}^\tau = \mathfrak{f}$ *for every* $\tau \in \mathrm{Gal}(F/\mathbb{Q})$.

Proof. ($\Rightarrow$). From Proposition 2.3, we get condition (c). By Lemma 4.1, we see that the situation in Proposition 4.7 holds. Combing Proposition 2.2 with (4.6) and (4.7), we obtain

$$\tilde{\sigma}^2 \varrho^{-1}|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^2}\right).$$

By Proposition 4.10, this is equivalent to condition (4.9). From Proposition 4.4, Proposition 4.5 and Proposition 4.11, we have conditions (a) and (b).

($\Leftarrow$). By Proposition 4.4 and Proposition 4.5, we obtain that K is cyclic over $\mathbb{Q}$ and $I_0(\Phi') = I_K$ from conditions (a) and (b). By Lemma 4.1, A is simple. Moreover by Theorem 2.1, $M_K = K$. Combing condition (c), we get the situation in Proposition 4.7. By Proposition 4.11, condition (a) implies

$$\tilde{\sigma}^2 \varrho^{-1}|_{H_K} = \left(\frac{H_K/K}{\mathfrak{a}^2}\right).$$

From (4.6) and (4.7), $(P_\sigma : P) = 1$ where σ is a generator of $\mathrm{Gal}(K/\mathbb{Q})$. By Proposition 2.2, $M_{\mathbb{Q}} = \mathbb{Q}$. □

Remark 4.13. There exist CM-fields K, which are cyclic and quartic over $\mathbb{Q}$ and satisfy $I_0(\Phi') = I_K$, but which do not have one of the forms given in Theorem 4.12; for such fields $M = F \neq \mathbb{Q}$. By Proposition 8 in [5], for such K, t must be greater or equal to 1. The CM-fields, being cyclic and quartic over $\mathbb{Q}$, with $I_0(\Phi') = I_K$ and $t = 1$ are all determined in [2]. These are exactly eight fields. Among of them, just the two fields

$$\mathbb{Q}(\sqrt{-3(2+\sqrt{2})}) = \mathbb{Q}(\sqrt{-3(1+\sqrt{2})\sqrt{2}})$$

and

$$\mathbb{Q}(\sqrt{-7(17+4\sqrt{17})}) = \mathbb{Q}(\sqrt{-7(4+\sqrt{17})\sqrt{17}})$$

satisfy none of the conditions (a_1), (a_2), and (a_3) in Theorem 4.12.

References

[1] *G. W. Anderson*, On a question arising from complex multiplication theory, Galois Representations and Arithmetic Algebraic Geometry, Adv. Stud. Pure Math. **12** (1987), 221–234.
[2] *K. Hardy*, *R. H. Hudson* and *D. Richman*, *K. S. Wiliams*, Determination of all imaginary quartic fields with class number 2, Trans. Amer. Math. Soc. **311** (No 1) (1989), 1–55.
[3] *S. Lang*, Complex multiplication, Springer-Verlag, New York 1983.
[4] *G. Shimura*, Introduction to the arithmetic theory of automorphic functions, Iwanami Shoten and Princeton University Press, 1971.
[5] *G. Shimura*, On abelian varieties with complex multiplication, Proc. London Math. Soc. **34** (1977), 65–86.
[6] *G. Shimura* and *Y. Taniyama*, Complex multiplication of abelian varieties and its application to number theory, Math. Soc. Japan, 1961.

Department of Mathematics, Faculty of Science, Yamagata University, Yamagata, Japan

Eingegangen 18. April 1994

J. reine angew. Math. **470** (1996), 27—50

Journal für die reine und angewandte Mathematik

On Sprindžuk's classification of transcendental numbers

By *Masaaki Amou**) at Kiryu

1. Introduction

In 1962 Sprindžuk [13] proposed a classification of the complex numbers, in connection with that of Mahler [9], by dividing them into four classes and he called the numbers in these classes $\tilde{A}$-, $\tilde{S}$-, $\tilde{T}$-, and $\tilde{U}$-numbers (cf. also [14], p. 140 and [15], p. 13). The purpose of this paper is to show the existence of numbers with certain properties related to this classification, and especially to show the existence of $\tilde{T}$-numbers. We first recall the definition of Sprindžuk's classification and the several known results concerning this classification, then we state our main results.

As usual, for both a polynomial $P \in \mathbb{Z}[x]$ and an algebraic number α, we denote by $H(P)$ and $H(\alpha)$ the *height* of P and that of α, respectively. Namely, $H(P)$ is the maximum of the absolute values of the coefficients of P and $H(\alpha)$ is the height of the minimal defining polynomial of α.

Let ω be a complex number, and let $w_d(\omega, h)$ be the function of positive integers d, h defined by Mahler [9] as

$$w_d(\omega, h) = \min |P(\omega)|,$$

where the minimum is taken over all polynomials $P \in \mathbb{Z}[x]$ with $\deg P \leqq d$, $H(P) \leqq h$ satisfying $P(\omega) \neq 0$. Sprindžuk's classification is based on the behaviour of $w_d(\omega, h)$ as well as Mahler's one, but in the former, the role of the variable d in the later is replaced by that of the variable h. More precisely Sprindžuk defines

$$v(\omega, h) = \limsup_{d \to \infty} \frac{\log\log(1/w_d(\omega, h))}{\log d}, \quad v(\omega) = \sup_{h \in \mathbb{N}} v(\omega, h).$$

*) This research was partly supported by the Grant-in-Aid for Encouragement of Young Scientists (No. 04740010), the Ministry of Education, Science and Culture, Japan.

Note that $v(\omega, h)$ is nondecreasing as a function of h. He calls $v(\omega)$ the *order* of ω. If it is finite, he also defines

$$t(\omega, h) = \limsup_{d \to \infty} \frac{-\log w_d(\omega, h)}{d^{v(\omega)}}, \quad t(\omega) = \limsup_{h \to \infty} \frac{t(\omega, h)}{\log h}.$$

He calls $t(\omega)$ the *type* of ω. In case $v(\omega) = \infty$, if there exists h such that $v(\omega, h) = \infty$, he denotes by $H_0 = H_0(\omega)$ the smallest h for which this holds, and otherwise he puts $H_0 = H_0(\omega) = \infty$. In this notation Sprindžuk defines the classes of $\tilde{A}$-, $\tilde{S}$-, $\tilde{T}$-, and $\tilde{U}$-numbers as follows:

$\tilde{A}$-numbers: $0 \leqq v(\omega) \leqq 1$; if $v(\omega) = 1$, then $t(\omega) = 0$,

$\tilde{S}$-numbers: $1 \leqq v(\omega) < \infty$; if $v(\omega) = 1$, then $t(\omega) > 0$,

$\tilde{T}$-numbers: $v(\omega) = \infty$; $H_0(\omega) = \infty$,

$\tilde{U}$-numbers: $v(\omega) = \infty$; $H_0(\omega) < \infty$.

Then Sprindžuk [13] proved the following fundamental result.

Theorem S. (i) *All algebraic numbers have orders not exceeding* 1 *and if an algebraic number has order* 1, *then it has type* 0.

(ii) *All transcendental numbers have order not less than* 1; *a real transcendental number of order* 1 *has type not less than* 1, *and a complex transcendental number of order* 1 *has type not less than* 1/2.

(iii) *Almost all real and almost all complex numbers belong to the class of* $\tilde{S}$*-numbers and have orders not exceeding* 2.

We know from (i) and (ii) that all algebraic numbers precisely consist of the class of $\tilde{A}$-numbers. Concerning (iii), Chudnovsky claimed the following best possible result (cf. [3], Chap. 2, Remark 7.5).

Theorem C. *Almost all real and almost all complex numbers belong to the class of* $\tilde{S}$*-numbers and have order* 1.

Remark. We give a proof of this result in the appendix because Chudnovsky only gave a sketch of the proof from which it seems difficult to complete the proof.

It is natural to ask the existence of $\tilde{S}$-numbers ω with $v(\omega) = v$ for any given $v > 1$, of $\tilde{U}$-numbers ω with $H_0(\omega) = h$ for any given positive integer h, and especially of $\tilde{T}$-numbers. One of our main results (Theorem 1 below) includes the affirmative answer to all the questions. Let $\mathbf{v} = \{v_h\}_{h \in \mathbb{N}}$ be a sequence of nonnegative real numbers (which may contain ∞) satisfying

$$(1.1) \qquad v_h = 0 \quad \text{for} \quad h < h_0, \quad 1 \leqq v_h \leqq v_{h+1} \leqq \infty \quad \text{for} \quad h \geqq h_0$$

with some positive integer h_0. For each such sequence $\mathbf{v}$, we define

$$\mathscr{R}(\mathbf{v}) = \{\omega \in \mathbb{R} \backslash \bar{\mathbb{Q}} \mid v(\omega, h) = v_h \text{ for all } h\},$$

and define the real number $m_{\mathbf{v}}$ and $M_{\mathbf{v}}$ as follows:

(i) If $h_0 > 1$, then $m_{\mathbf{v}} = h_0$ and $M_{\mathbf{v}} = h_0 + 1$.

(ii) If there exists a positive integer $h_0' \geqq 1$ such that $v_h = 1$ for $h < h_0'$ and $v_{h_0'} > 1$, then $m_{\mathbf{v}} = 1/(h_0' + 1)$ and $M_{\mathbf{v}} = 2$.

(iii) If $v_h = 1$ for all h, then $m_{\mathbf{v}} = 0$ and $M_{\mathbf{v}} = 2$.

Further, for any real numbers a, b with $0 \leqq a < b$, we define

$$\mathbf{E}(a, b) = \{x \in \mathbb{R} \mid a < |x| < b\}.$$

In this notation we have the following

Theorem 1. *For each sequence* $\mathbf{v} = \{v_h\}_{h \in \mathbb{N}}$ *satisfying* (1.1) *with some positive integer* h_0, $\mathscr{R}(\mathbf{v})$ *is contained and dense in* $\mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}})$.

Corollary. (i) *There exist* $\tilde{S}$*-numbers* ω *with* $v(\omega) = v$ *for any given* $v \geqq 1$.

(ii) *There exist* $\tilde{T}$*-numbers.*

(iii) *There exist* $\tilde{U}$*-numbers* ω *with* $H_0(\omega) = h$ *for any given positive integer* h.

We next state our second result. Let ω be a complex number, and let $\tilde{w}_d(\omega, h)$ be a function of positive integers d, h defined as

$$\tilde{w}_d(\omega, h) = \min |\omega - \alpha|,$$

where the minimum is taken over all algebraic numbers α with $\deg \alpha \leqq d$, $H(\alpha) = h$ satisfying $\alpha \neq \omega$. Then we define

$$\tilde{v}(\omega, h) = \limsup_{d \to \infty} \frac{\log \log (1/\tilde{w}_d(\omega, h))}{\log d}.$$

The function $\tilde{v}(\omega, h)$ of h describes how well ω can be approximated by algebraic numbers of height h. Let $\mathbf{v} = \{v_h\}_{h \in \mathbb{N}}$ be a sequence of nonnegative real numbers (which may contain ∞) satisfying

$$(1.2) \qquad v_h = 0 \quad \text{for} \quad h < h_0, \quad 1 \leqq v_h \leqq \infty \quad \text{for} \quad h \geqq h_0$$

with some positive integer h_0. Note that $\mathbf{v}$ need not be a nondecreasing sequence. For each such sequence $\mathbf{v}$, we define

$$\tilde{\mathscr{R}}(\mathbf{v}) = \{\omega \in \mathbb{R} \backslash \bar{\mathbb{Q}} \mid \tilde{v}(\omega, h) = v_h \text{ for all } h\}.$$

Further, for any $a \geqq 1$, we write simply $\mathbf{E}(a)$ instead of $\mathbf{E}(1/a, a)$ $(\mathbf{E}(1) = \emptyset)$. In this notation we have the following

Theorem 2. *For each sequence* $\mathbf{v} = \{v_h\}_{h \in \mathbb{N}}$ *satisfying* (1.2) *with some positive integer* h_0, $\tilde{\mathscr{R}}(\mathbf{v})$ *is contained and dense in* $\mathbf{E}(h_0 + 1) \setminus \mathbf{E}(h_0)$.

The plan of this paper is as follows. In the following three sections, we give preliminary results to prove our theorems. More precisely in § 2 we prove an effective version of a theorem of Kornblum [8] on the polynomial analogue of Dirichlet's theorem on primes in arithmetic progressions. In § 3, using this result, we prove Proposition 2 which gives a certain information on distribution of real algebraic numbers with given height. In § 4 we prove Lemma 2 which gives a lower bound for the distance between roots of polynomials. The proof of our theorems are divided into two parts, namely the proof of the inclusion relations and that of the denseness. In § 5 we prove the inclusion relations. In § 6, using Proposition 2 and Lemma 2, we prove Proposition 3. Then we deduce the denseness from it. Finally in § 7 we refer certain results on purely imaginary numbers analogous to Theorem 1 or Theorem 2, and propose a further problem.

The author thanks Professor K. Komatsu for referring him to the paper [6], and Professor M. Ishibashi for sending him the paper [7] before publication. He is also indebted to Professor I. Wakabayashi and the referee for valuable comments.

2. An effective version of a theorem of Kornblum

Let $\mathbb{F}_q[x]$ be the ring of polynomials over the finite field of q elements. In 1919 Kornblum [8] proved the polynomial analogue of Dirichlet's theorem on primes in arithmetic progressions as follows.

Theorem K. *Let* $\mathbf{A}$ *and* $\mathbf{M}$ *be relatively prime polynomials in* $\mathbb{F}_q[x]$. *Then there are infinitely many irreducible polynomials* $\mathbf{P} \in \mathbb{F}_q[x]$ *such that* $\mathbf{P} \equiv \mathbf{A} \pmod{\mathbf{M}}$.

The purpose of this section is to show the following effective version of Kornblum's theorem as a corollary to a recent result of Ishibashi [7] with the aid of Carlitz-Hayes cyclotomic theory of rational function fields over finite fields of constants (cf. Carlitz [2], Hayes [6]).

Proposition 1. *Let* $\mathbf{A}$ *and* $\mathbf{M}$ *be relatively prime polynomials in* $\mathbb{F}_q[x]$. *Then there exists a positive constant* c_1 *depending only on* q *such that, for each positive integer* d *greater than* $4 \deg \mathbf{M} + c_1 \log \max(2, \deg \mathbf{M})$, *there is an irreducible polynomial* $\mathbf{P} \in \mathbb{F}_q[x]$ *satisfying* $\deg \mathbf{P} = d$, $\mathbf{P} \equiv \mathbf{A} \pmod{\mathbf{M}}$.

We shall use the following corollary to prove Proposition 2 in the next section.

Corollary. *Let* $A \in \mathbb{Z}[x]$ *be a nonconstant polynomial, and let* h *be a positive integer with* $h \geqq H(A)$. *Assume that* $A(0) \neq 0$ *if* $h = 1$, *and that* $A(0)$ *is not divisible by some prime factor* p *of* h *if* $h > 1$. *Let* n *be a positive integer greater than* $\deg A$. *Then, there exists a positive constant* c_2 *depending only on* h *such that, for each positive integer* d *greater than* $4n + c_2 \log n$, *there is an irreducible polynomial* $P \in \mathbb{Z}[x]$ *satisfying*

$$\deg P = d, \quad H(P) = |a_d| = h\,, \quad \textit{and} \quad P \equiv A\,(\bmod\, x^n),$$

where a_d is the leading coefficient of P.

Proof. We put $p = 2$ if $h = 1$, and we choose a prime factor p of h which does not divide $A(0)$ if $h > 1$. Let $\bar{A}$ be the reduction modulo p of A. Since $\bar{A}$ has a nonzero constant term by the assumption, $\bar{A}$ and x^n are relatively prime in $\mathbb{F}_p[x]$. Hence, by setting $q = p$, $A = \bar{A}$, and $\mathbf{M} = x^n$ in the proposition, we have a positive constant c_1 depending only on p (and hence only on h) with the property stated there. Put $c_2 = c_1 + 2$. We show that this c_2 has the property stated in the corollary. Let d be an arbitrary positive integer greater than $4n + c_2 \log n\,(> 4n + c_1 \log n + 1)$. Then, by the proposition, we can take an irreducible polynomial $\mathbf{P} \in \mathbb{F}_p[x]$ such that $\deg \mathbf{P} = d'$, $\mathbf{P} \equiv \bar{A}\,(\bmod\, x^n)$, where $d' = d$ if $h > 1$, $d' = d - 1$ if $h > 1$. This implies the existence of an irreducible polynomial $Q \in \mathbb{Z}[x]$ such that

$$\deg Q = d', \quad H(Q) \leqq h\,, \quad \text{and} \quad Q \equiv A\,(\bmod\, x^n).$$

Put $P = Q$ if $h = 1$, $P = Q + hx^d$ if $h > 1$. We see that P satisfies the desired properties. This completes the proof of the corollary.

To prove Proposition 1, we first quote a result of Ishibashi [7]. Let k be an algebraic function field of one variable with constant field $\mathbb{F}_q$, and K be a finite Galois extension over k whose constant field has degree f_0 over $\mathbb{F}_q$. We denote by G the Galois group of the extension K/k, and by g_K the genus of K. For a prime divisor $\mathfrak{p}$ of k, we denote by $\left(\frac{K/k}{\mathfrak{p}}\right)$ the Artin symbol of $\mathfrak{p}$ in G (cf. Fried-Jarden [4], Chap. 5). Further, we denote by $\log_q x$ the logarithmic function of x with base q. In this notation Ishibashi [7] proved the following

Theorem I. *Let $\mathscr{C}$ be a conjugacy class in G. Then, for each positive number D greater than $2\log_q\{136(g_K + |G|)|G|\}$, there exists a finite prime divisor $\mathfrak{p}$ of k which is unramified in K such that $\left(\frac{K/k}{\mathfrak{p}}\right) = \mathscr{C}$, $D < \deg(\mathfrak{p}) < D + 2f_0$.*

We next summarize from the results of Hayes [6], §§2–4. In the following, for any polynomial $\mathbf{U} \in \mathbb{F}_q[x]$, we denote by $\hat{\mathbf{U}}$ the residue class modulo $\mathbf{M}$ which contains $\mathbf{U}$.

Theorem H. *Let $\mathbf{M} = a \prod_{i=1}^{l} \mathbf{P}_i^{n_i}$ be a nonconstant polynomial in $\mathbb{F}_q[x]$, where $a \in \mathbb{F}_q$ and each $\mathbf{P}_i$ is a monic irreducible polynomial. Then there exists an abelian extension K over $k = \mathbb{F}_q(x)$ with constant field $\mathbb{F}_q$ which satisfies the following properties:*

(i) *For a finite prime divisor (i.e., an irreducible polynomial) $\mathbf{P}$ of k, $\mathbf{P}$ is ramified in K if and only if $\mathbf{P}$ divides $\mathbf{M}$.*

(ii) *There exists an isomorphism φ from the group of units of $\mathbb{F}_q[x]/(\mathbf{M})$ to the Galois group $G = \mathrm{Gal}(K/k)$ such that, for any finite prime divisor $\mathbf{P}$ which does not divide $\mathbf{M}$, $\left(\frac{K/k}{\mathbf{P}}\right) = \varphi(\hat{\mathbf{P}})$.*

(iii) *Put* $a_i = q^{d_i n_i} - q^{d_i(n_i - 1)}$, *where* d_i *is the degree of* $\mathbf{P}_i$. *Then we have*

$$|G| = \prod_{i=1}^{l} a_i < q^{\deg \mathbf{M}}$$

and

$$\deg \mathfrak{D} = \sum_{i=1}^{l} \left\{ (n_i a_i - q^{d_i(n_i - 1)}) d_i + (q-2) \frac{a_i}{q-1} \right\},$$

where $\mathfrak{D}$ *is the different of* K/k.

Proof of Proposition 1. We assume that $\mathbf{M}$ can be factorized as in Theorem H. For this $\mathbf{M}$, let K be an abelian extension over $k = \mathbb{F}_q(x)$ which satisfies the properties stated in Theorem H. We now apply Theorem I to $\mathscr{C} = \varphi(\mathbf{A})$. Then, for each positive integer D greater than $D_0 = 2 \log_q \{136(g_K + |G|)|G|\}$, there exists a finite prime divisor $\mathbf{P}$ of k which is unramified in K such that $\left(\frac{K/k}{\mathbf{P}}\right) = \varphi(\hat{A})$, $D < \deg \mathbf{P} < D + 2$ (i.e., $\deg \mathbf{P} = D + 1$). From Theorem H (i) we deduce that $\mathbf{P}$ does not divide $\mathbf{M}$. Hence from Theorem H (ii) we deduce that $\varphi(\hat{\mathbf{P}}) = \varphi(\hat{\mathbf{A}})$, namely $\mathbf{P} \equiv \mathbf{A} \pmod{\mathbf{M}}$. Therefore, to complete the proof of the proposition, it remains only to estimate D_0 from above. By the Hurwitz formula, we have

$$2 g_K - 2 = -2|G| + \deg \mathfrak{D}.$$

Hence, using Theorem H (iii), we have

$$2 g_K < \sum_{i=1}^{l} (d_i n_i + 1) a_i < 2 q^{\deg \mathbf{M}} \deg \mathbf{M}.$$

This with the upper bound for $|G|$ in Theorem H (iii) implies the desired upper bound for D_0. Thus the proof of the proposition is completed.

3. Distribution of algebraic numbers with given height

In this section we prove Proposition 2 below which plays an essential role in the proof of our theorems. In the following, for any polynomial

$$P(x) = \sum_{i=0}^{d} a_i x^i \in \mathbb{Z}[x] \quad (a_d \neq 0),$$

we denote by $\hat{H}(P)$ the maximum of $|a_0|$ and $|a_d|$. Also, for any algebraic number α with minimal defining polynomial Ψ, we put $\hat{H}(\alpha) = \hat{H}(\Psi)$.

Proposition 2. *Let h be a positive integer, and let ξ be a real number in* $\mathbf{E}(h+1)$ *with* $|\xi| \neq 1$. *Then there exist a positive constant $c(\xi, h)$ depending only on ξ, h and an infinite set $\mathscr{N}(\xi, h)$ of positive integers associated to $c(\xi, h)$ such that, for each $d \in \mathscr{N}(\xi, h)$, there are at least $(2h+1)^{d/9}$ real algebraic numbers α satisfying* $\deg \alpha = d$, $H(\alpha) = \hat{H}(\alpha) = h$, *and*

$$|\xi - \alpha| < \exp(-c(\xi, h) d). \tag{3.1}$$

If a complex number α satisfies $P(\alpha) = 0$ for some nonzero polynomial $P \in \mathbb{Z}[x]$ of height h, then we have

$$\frac{1}{h+1} < |\alpha| < h+1$$

(cf. Shidlovskii [12], p. 30, Lemma 2). Hence we have the following corollary to the proposition which describes the topological closure of the set of all real algebraic numbers with a given height.

Corollary. *For each positive integer h, the set of all real algebraic numbers of height h is contained and dense in $\mathbf{E}(h+1)$.*

Proof of Proposition 2. We first prove the assertion in the case where

$$\frac{1}{h+1} < \xi < 1 .$$

To this aim, we expand 1 as a power series of ξ in the form

$$1 = \sum_{\nu=1}^{\infty} a_\nu \xi^\nu , \tag{3.2}$$

where a_ν's are the integers defined uniquely by the following inductive procedure: a_1 is the largest integer for which $1 > a_1 \xi$; a_2 is the largest integer for which $1 > a_1 \xi + a_2 \xi^2$, and so on. It is easily seen that $0 \leqq a_\nu \leqq h$ for all ν and $a_\nu \neq 0$ for infinitely many ν. For each n such that $a_n \neq 0$, we define a polynomial

$$P_n(x) = -1 + \sum_{\nu=1}^{n} a_\nu x^\nu ,$$

and consider the set of all polynomials $Q \in \mathbb{Z}[x]$ satisfying $\deg Q \leqq 2n$, $H(Q) \leqq h$, and $Q \equiv P_n (\operatorname{mod} x^{n+1})$. Obviously, this set contains $(2h+1)^n$ elements. By the corollary to Proposition 1, for each Q in this set and for each positive integer d greater than

$$4(2n+1) + c \log(2n+1)$$

with a certain positive constant c depending only on h, there exists an irreducible polynomial $P \in \mathbb{Z}[x]$ satisfying $\deg P = d$, $H(P) = |a_d| = h$, and $P \equiv Q (\operatorname{mod} x^{2n+1})$, where a_d is the leading coefficient of P. Hence, assuming $n > 4 + c \log(2n+1)$, we have a set of $(2h+1)^n$ primitive and irreducible polynomials $P \in \mathbb{Z}[x]$ satisfying $\deg P = 9n$. $H(P) = \hat{H}(P) = h$, and $P \equiv P_n (\operatorname{mod} x^{n+1})$.

We now claim that each P above has a real root α satisfying (3.1) with a certain positive constant $c(h, \xi)$ depending only on h, ξ and with $d = 9n$ provided that n is sufficiently large. To this aim, we wish to find a positive number ε as small as possible for which $P(\xi - \varepsilon) < 0 < P(\xi + \varepsilon)$. We shall use the equalities

$$P(\xi \pm \varepsilon) = P_n(\xi) + \{P_n(\xi \pm \varepsilon) - P_n(\xi)\} + \{P(\xi \pm \varepsilon) - P_n(\xi \pm \varepsilon)\} . \tag{3.3}$$

For any ε with $0 < \varepsilon < \xi$, using (3.3) with

$$P_n(\xi) < 0, \quad P_n(\xi - \varepsilon) - P_n(\xi) \leqq -a_1 \varepsilon \leqq -\varepsilon,$$

we have

$$P(\xi - \varepsilon) < -\varepsilon + |P(\xi - \varepsilon) - P_n(\xi - \varepsilon)| < -\varepsilon + \frac{h\xi^{n+1}}{1-\xi},$$

which implies that if

$$\frac{h\xi^{n+1}}{1-\xi} < \varepsilon < \xi,$$

then $P(\xi - \varepsilon) < 0$. On the other hand, for any ε with $0 < \varepsilon < 1 - \xi$, using (3.3) with

$$P_n(\xi) \geqq -\xi^n, \quad P_n(\xi + \varepsilon) - P_n(\xi) \geqq a_1 \varepsilon \geqq \varepsilon,$$

we have

$$P(\xi + \varepsilon) > -\xi^n + \varepsilon - \frac{h(\xi + \varepsilon)^{n+1}}{1 - \xi - \varepsilon},$$

which implies that if

$$\xi^n + \frac{2h\{(1+\xi)/2\}^{n+1}}{1-\xi} < \varepsilon < \frac{1-\xi}{2},$$

then $P(\xi + \varepsilon) > 0$. These estimates for ε imply our claim, and the assertion of the proposition in the present case holds.

Similarly we can prove the assertion in the case where $-1 < \xi < -1/(h+1)$ by expanding 1 as a power series of ξ in the form (3.2) with integral coefficients a_ν which satisfy $-h \leqq a_{2\nu-1} \leqq 0$, $0 \leqq a_{2\nu} \leqq h$ for $\nu \in \mathbb{N}$.

It remains the proof in the case where $1 < |\xi| < h + 1$. For any nonzero algebraic number α, we have $|\xi - \alpha^{-1}| = |\xi \alpha^{-1}(\xi^{-1} - \alpha)|$. Hence, by noticing $1/(h+1) < |\xi^{-1}| < 1$, $\deg \alpha^{-1} = \deg \alpha$, $H(\alpha^{-1}) = H(\alpha)$, and $\hat{H}(\alpha^{-1}) = \hat{H}(\alpha)$, the assertion in this case follows from the assertions in the former cases. This completes the proof of the proposition.

Definition. For $h \in \mathbb{N}$ and $\xi \in \mathbf{E}(h+1)$ with $|\xi| \neq 1$, we choose a positive constant $c(\xi, h)$ and an infinite set $\mathcal{N}(\xi, h)$ of positive integers satisfying the property stated in the proposition, and fix them. Then, for each $d \in \mathcal{N}(\xi, h)$, we define $\mathfrak{U}(\xi; d, h)$ to be the set of all real algebraic numbers α satisfying $\deg \alpha = d$, $H(\alpha) = \hat{H}(\alpha) = h$, and (3.1).

Note that $\mathfrak{U}(\xi; d, h)$ has at least $(2h+1)^{d/9}$ elements. We shall use these sets in § 6.

4. Distance between roots of polynomials

The main purpose of this section is to prove Lemma 2 below which gives a lower bound for the distance between roots of polynomials with integral coefficients. In the following, for any polynomial R with complex coefficients, we denote by $L(R)$ the *length* of R, i.e., the sum of the absolute values of the coefficients of R. To prove Lemma 2, we shall use the following facts on polynomials.

Lemma 1. *Let $R(x)$ be a nonzero polynomial with complex coefficients, which is factorized as $R(x) = c(x-\gamma_1)\cdots(x-\gamma_d)$, where γ_i's are not necessarily different. Then we have:*

(i) $|R(\omega)| \leqq L(R)\max(1,|\omega|)^d$ *for any complex number* ω.

(ii) $L(R/(x-\gamma)) \leqq dL(R)\max(1,|\gamma|)^{-1}$ *for any root* γ *of* R.

(iii) $|c|\max(1,|\gamma_1|)\cdots\max(1,|\gamma_d|) \leqq L(R)$.

Among these facts, (i) is trivial. The fact (ii) is a special case of Lemma H of Güting [5], and the fact (iii) is a result of Mahler [10].

Lemma 2. *Let P, Q be polynomials of degrees d, d' with integral coefficients, and let α, β be roots of P, Q with multiplicities s, s', respectively. Assume that $\alpha \neq \beta$. We have:*

(i) *If* $P(\beta) \neq 0$ *or* $Q(\alpha) \neq 0$, *then*

$$|\alpha-\beta| \geqq \max(d,d')^{-1} L(P)^{-(d'/ss')} L(Q)^{-(d/ss')}.$$

(ii) *If* $P(\beta) = 0$ *and* $Q(\alpha) = 0$, *then*

$$|\alpha-\beta| \geqq \max\left(d^{-(d/s)-1} L(P)^{-2(d-1)/s}, d'^{-(d'/s')-1} L(Q)^{-2(d'-1)/s'}\right).$$

Remark. The reader may compare this lemma with Theorem 6′ of Güting [5]. While Güting's theorem has good dependence on the heights of polynomials, our lemma has good dependence on the degrees of polynomials.

Proof of Lemma 2. For the proof of (i), assume that $P(\beta) \neq 0$. Let

$$P(x) = a\prod_{i=1}^{l}(x-\alpha_i)^{s_i} \quad (\alpha_1 = \alpha, s_1 = s)$$

and

$$\Psi(x) = b\prod_{j=1}^{q}(x-\beta_j) \quad (\beta_1 = \beta),$$

where Ψ is the minimal defining polynomial of β. Our starting point is the equality

$$|\alpha-\beta|^s=\frac{|P(\beta)|}{|a|\prod_{i=2}^{l}|\beta-\alpha_i|^{s_i}}.$$

Using Lemma 1 (i) and (ii), we have

$$|a|\prod_{i=2}^{l}|\beta-\alpha_i|^{s_i}\leqq d^sL(P)\max(1,|\beta|)^{d-s}\max(1,|\alpha|)^{-s}.$$

Since $P(\beta)\neq 0$, the two polynomials Ψ and P are relatively prime. This implies that their resultant

$$r(\Psi,P)=b^d\prod_{j=1}^{q}P(\beta_j)$$

is a nonzero integer. Hence, using Lemma 1 (i) for $|P(\beta_j)|$ with $j=2,\dots,q$, we have

$$|P(\beta)|\geqq b^{-d}L(P)^{1-q}\left(\prod_{j=2}^{q}\max(1,|\beta_j|)\right)^{-d}.$$

These estimates yield that

$$|\alpha-\beta|^s\geqq d^{-s}L(P)^{-q}\left(b\prod_{j=1}^{q}\max(1,|\beta_j|)\right)^{-d}.$$

Since Q is divisible by $\Psi^{s'}$ in $\mathbb{Z}[x]$, we have $q\leqq d'/s'$. Also, by Lemma 1 (iii), we have

$$\left(b\prod_{j=1}^{q}\max(1,|\beta_j|)\right)^{s'}\leqq L(Q).$$

Hence we obtain

$$|\alpha-\beta|\geqq d^{-1}L(P)^{-(d'/ss')}L(Q)^{-(d/ss')}.$$

The assertion of (i) now follows from the symmetry of the role of α and that of β.

For the proof of (ii), let us express P as in the proof of (i) but with $\alpha_1=\beta$, $\alpha_2=\alpha$, and $s_2=s$. In this case our starting point is the equality

$$|\alpha-\beta|^s=\frac{|P^{(s_1)}(\beta)/s_1!|}{|a|\prod_{i=3}^{l}|\beta-\alpha_i|^{s_i}}.$$

Using Lemma 1 (i) and (ii), we have

$$|a|\prod_{i=3}^{l}|\beta-\alpha_i|^{s_i}\leqq d^{s+s_1}L(P)\max(1,|\beta|)^{d-s-2s_1}\max(1,|\alpha|)^{-s}.$$

To estimate $|P^{(s_1)}(\beta)/s_1!|$ from below, we use the semidiscriminant

$$\mathrm{SD}(P) = a^{d-2} \prod_{i=1}^{l} \left(P^{(s_i)}(\alpha_i)/s_i!\right)$$

of P defined by Chudnovsky. Since $\mathrm{SD}(P)$ is a nonzero integer (cf. Chudnovsky [3], p. 35), using Lemma 1 (i) and (ii) for $|P^{(s_i)}(\alpha_i)/s_i!|$ with $i = 2, \ldots, l$, we have

$$|a|^{-(d-2)} d^{-(s_2 + \cdots + s_l)} L(P)^{-(l-1)} \left(\prod_{i=2}^{l} \max(1, |\alpha_i|)^{-(d-s_i)} \right)$$

as a lower bound for $|P^{(s_1)}(\beta)/s_1!|$. These estimates yield that

$$|\alpha - \beta|^s \geqq d^{-(d+s)} L(P)^{-l} \left(|a| \prod_{i=1}^{l} \max(1, |\alpha_i|) \right)^{-(d-2)}.$$

Therefore, using $l \leqq d$ and Lemma 1 (iii), we obtain

$$|\alpha - \beta| \geqq d^{-(d/s)-1} L(P)^{-2(d-1)/s}.$$

The assertion of (ii) now follows from the symmetry of the role of α and that of β. This completes the proof of the lemma.

For later convenience, we state the following corollary to the lemma.

Corollary. *With the notation as in the lemma, if $\alpha \neq \beta$, then we have*

$$\log|\alpha - \beta| \gg -\left((d + d')\log(d + d')\right)/\max(s, s'),$$

where the constant implied by the symbol $\gg$ depends only on the heights of P and Q.

At the end of this section, we quote a result of Chudnovsky (cf. [3], Chap. 2, Lemma 1.12) which shall be used in §6 and the appendix.

Lemma 3. *Let ω be a complex number. Let $P \in \mathbb{Z}[x]$ be a polynomial of degree d, and let α be a root of P nearest to ω. Then we have*

$$|P(\omega)| \geqq |\omega - \alpha|^s \{2d^2(d+1) H(P)^2\}^{-d+1},$$

where s is the multiplicity of α in P.

5. Proof of the inclusion relations

In this section we shall prove the inclusion relations stated in Theorems 1 and 2, namely:

(a) $\mathscr{R}(\mathbf{v}) \subset \mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}})$;

(b) $\tilde{\mathscr{R}}(\mathbf{v}) \subset \mathbf{E}(h_0 + 1) \backslash \mathbf{E}(h_0)$.

Lemma 4. *Let h be a positive integer, and let ω be a nonzero real number. We have:*

(i) *If $0 < |\omega| \leqq 1/(h+1)$, then $v(\omega, h) = 1$.*

(ii) *If ω is a transcendental number in $\mathbf{E}(h+1)$, then $v(\omega, h) \geqq 1$.*

(iii) *If $|\omega| \geqq h+1$, then $v(\omega, h) = 0$.*

Proof. We first prove (iii). Let P be a polynomial of degree $d \geqq 1$ and height at most h. Then we have

$$|\omega|^d - h\sum_{j=0}^{d-1} |\omega|^j = |\omega|^d - \frac{h|\omega|^d}{|\omega|-1} + \frac{h}{|\omega|-1}$$

as a lower bound for $|P(\omega)|$. Since $|\omega| \geqq h+1$, this bound is not smaller than $h/(|\omega|-1)$, which does not depend on d. This implies (iii).

For the proof of (i), let P be as above and denote by $\tilde{P}$ the reciprocal of P. Since $|1/\omega| \geqq h+1$, as in the proof of (iii), we have $h/(|1/\omega|-1)$ as a lower bound for $|\tilde{P}(1/\omega)|$. Hence we have $h|\omega|^{d+1}/(1-|\omega|)$ as a lower bound for $|P(\omega)|$. This implies that $v(\omega, h) \leqq 1$. On the other hand, we easily see that the opposite inequality $v(\omega, h) \geqq 1$ holds by considering a series of the values $|P(\omega)|$, where P ranges over all polynomials of the form $P(x) = x^d$. Therefore (i) holds.

Finally, we prove (ii) by using the pigeon-hole principle (cf. Baker [1], p. 87). Let $\mathfrak{P}^+(d, h)$ be the set of all polynomials of degree at most d with integral coefficients between 0 and h inclusive. Then $\mathfrak{P}^+(d, h)$ contains $(h+1)^{d+1}$ elements. Further, for each $P \in \mathfrak{P}^+(d, h)$, we have

$$|P(\omega)| \leqq (d+1)h\{\max(1, |\omega|)\}^d = c(\omega; d, h), \quad \text{say}.$$

Hence, if we devide the interval $[-c(\omega; d, h), c(\omega; d, h)]$ into $(h+1)^d$ disjoint subintervals each of length $2(d+1)h\{\max(1, |\omega|)/(h+1)\}^d$, then there exist two distinct polynomials P_1 and P_2 in $\mathfrak{P}^+(d, h)$ such that $P_1(\omega)$ and $P_2(\omega)$ belong to the same subinterval. Put $P = P_1 - P_2$. Since $P \in \mathbb{Z}[x]$ is a nonzero polynomial of degree at most d and height at most h which satisfies

$$|P(\omega)| \leqq 2(d+1)h\{\max(1, |\omega|)/(h+1)\}^d,$$

using $|\omega| < h+1$, we obtain the desired inequality. This completes the proof of the lemma.

The inclusion relation (a) obviously follows from this lemma. Moreover, we see from this lemma that if ω is an arbitrarily given real transcendental number, then a sequence $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ defined by $v_h = v(\omega, h)$ satisfies (1.1) with some positive integer h_0.

We now turn to the inclusion relation (b). It is a consequence of the following lemma which also shows that, if ω is an arbitrarily given transcendental number, then a sequence $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ defined by $v_h = \tilde{v}(\omega, h)$ satisfies (1.2) with some positive integer h_0.

Lemma 5. *Let h be a positive integer, and let ω be a real number which is different from ± 1, $h+1$, and $1/(h+1)$. We have:*

(i) *If ω is in $\mathbf{E}(h+1)$, then $\tilde{v}(\omega, h) \geqq 1$.*

(ii) *If ω is not in $\mathbf{E}(h+1)$, then $\tilde{v}(\omega, h) = 0$.*

Proof. Obvious from Proposition 2 and its corollary.

6. Proof of the denseness

In this section we prove Proposition 3 below and deduce the assertions of the denseness of our theorems from it. To state the proposition, we introduce two functions $\Phi(d)$ and $\Delta(d, h)$ as follows. First, $\Phi(d)$ is a nondecreasing positive function of positive integers d which satisfies

$$d(\log d)^4 = o\big(\Phi(d)\big), \quad \Phi(d) = o\big(d(\log d)^{4+\varepsilon}\big) \tag{6.1}$$

for any $\varepsilon > 0$, as d tends to infinity. Secondly, $\Delta(d, h)$ is a function of positive integers d, h such that, for each fixed h, $\Delta(d, h)$ is a nondecreasing positive function of d which satisfies

$$d(\log d)^3 = o\big(\Delta(d, h)\big) \tag{6.2}$$

as d tends to infinity. We note that this estimate need not hold uniformly in h.

Proposition 3. *Let l be a positive integer, and let $\mathbf{I}_0 \subset \mathbf{E}(l+1) \backslash \mathbf{E}(l)$ be a nonempty open interval whose endpoints are also contained in $\mathbf{E}(l+1) \backslash \mathbf{E}(l)$. In case $l = 1$, we assume that $\mathbf{I}_0$ does not contain ± 1. Let $\Phi(d)$ and $\Delta(d, h)$ be as above. Further, let $\{h_j\}_{j \in \mathbb{N}}$ be a sequence of positive integers greater than or equal to l taking each value $h \geqq l$ infinitely often. Then there exist an increasing and bounded sequence of real algebraic numbers $\{\gamma_j\}_{j \in \mathbb{N}}$ with $H(\gamma_j) = \hat{H}(\gamma_j) = h_j$ and a sequence of positive numbers $\{\lambda_j\}_{j \in \mathbb{N}}$ satisfying the following properties: If we put ω the limit of γ_j's, then, for each j, we have*

$$\frac{1}{4} \exp(-\Delta_j) < \omega - \gamma_j < \exp(-\Delta_j)\,, \tag{i_j}$$

where $\Delta_j = \Delta(d_j, h_j)$ with the degree d_j of γ_j.

(ii) *For any pair $(P, \alpha) \in \mathbb{Z}[x] \times \bar{\mathbb{Q}}$ with $P \neq 0$, $P(\alpha) = 0$, and $\alpha \neq \gamma_1, \gamma_2, \ldots$, we have*

$$|\omega - \alpha| \geqq \exp\big(-D(P, \alpha)\big)$$

with

$$D(P, \alpha) = \gamma_{H(P)} \Phi(d)/s\,, \tag{6.3}$$

where d and s denote the degree of P and the multiplicity of α in P, respectively.

Proof. We wish to construct γ_j's and λ_j's in order of $\gamma_1, \lambda_1, \gamma_2, \lambda_2, \ldots$ so that for each j, γ_j and λ_j satisfy six conditions (I_j)–(VI_j) below. Our method deeply depends on the one developed by Schmidt [11] for constructing certain numbers related to Mahler's classification, especially for constructing Mahler's T-numbers.

In the following we denote by (P, α) an arbitrary element of $\mathbb{Z}[x] \times \bar{\mathbb{Q}}$ with

$$P \neq 0, \quad P(\alpha) = 0 .$$

Also, for any pair (P, α) as above, we use the letters d and s for denoting the degree of P and the multiplicity of α in P, respectively. Further, for any real algebraic number γ, we denote by $\mathbf{I}(\gamma)$ the open interval consisting of all real numbers x which satisfy

$$\frac{1}{4}\exp(-\Delta(\deg\gamma, H(\gamma))) < x - \gamma < \frac{1}{2}\exp(-\Delta(\deg\gamma, H(\gamma))) .$$

Under the notation and the assumptions as in the proposition, the first three conditions are as follows:

(I_j) γ_j is a real algebraic number in $\mathbf{I}(\gamma_{j-1})$ with $H(\gamma_j) = \hat{H}(\gamma_j) = h_j$. Also we have $\mathbf{I}(\gamma_j) \subset \mathbf{I}_0$, $|\mathbf{I}(\gamma_j)| \leqq \frac{1}{2}|\mathbf{I}(\gamma_{j-1})|$. Here $\mathbf{I}(\gamma_{j-1})$ means $\mathbf{I}_0$ in the case where $j = 1$ and $|\cdot|$ denotes the length of intervals.

(II_j) For any pair (P, α) such that $H(P) \leqq j-1$, $\alpha \neq \gamma_1, \ldots, \gamma_j$, and

$$D(P, \alpha) > \Delta_{j-1} = \Delta(d_{j-1}, h_{j-1})$$

with degree d_{j-1} of γ_{j-1}, we have

$$|\gamma_j - \alpha| > 2\exp(-D(P, \alpha)) .$$

(III_j) For any pair (P, α) with $H(P) = j$, $\alpha \neq \gamma_j$, the same inequality as in (II_j) is satisfied.

Further conditions are more technical and necessary only to construct γ_j, λ_j with (I_j)–(III_j) successfully. Therefore, before stating such conditions, we explain how to deduce (i_j) and (ii) from the above conditions. We first show that (I_i) for $i \geqq j$ imply (i_j). In fact, the lower bound for $\omega - \gamma_j$ in (i_j) follows from $\gamma_{j+1} \in \mathbf{I}(\gamma_j)$ and from the fact that $\{\gamma_i\}_{i \in \mathbb{N}}$ is increasing. On the other hand, using $\gamma_{i+1} \in \mathbf{I}(\gamma_i)$ for $i \geqq j$, we have

$$\omega - \gamma_j = \sum_{i=j}^{\infty} (\gamma_{i+1} - \gamma_i) < \frac{1}{2} \sum_{i=j}^{\infty} \exp(-\Delta_i) .$$

Hence, using $|\mathbf{I}(\gamma_i)| \leqq \frac{1}{2}|\mathbf{I}(\gamma_{i-1})|$, i.e., $\exp(-\Delta_i) \leqq \frac{1}{2}\exp(-\Delta_{i-1})$ for $i \geqq j+1$, we obtain

$$\omega - \gamma_j < \frac{1}{2} \sum_{i=j}^{\infty} 2^{j-i} \exp(-\Delta_j) = \exp(-\Delta_j).$$

This gives the upper bound for $\omega - \gamma_j$ in (i_j).

We next claim that (I_i)–(III_i) for $i = 1, \ldots, j$ imply

$$|\gamma_j - \alpha| > \exp(-D(P, \alpha))$$

for all pairs (P, α) with $H(P) \leqq j$, $\alpha \neq \gamma_1, \ldots, \gamma_j$. The condition (ii) clearly follows from this claim. To prove the claim, in view of (II_j) and (III_j), it is enough to show the above inequalities for all pairs (P, α) with $H(P) \leqq j-1$, $\alpha \neq \gamma_1, \ldots, \gamma_j$, and $D(P, \alpha) \leqq \Delta_{j-1}$. Let (P, α) be such a pair and put $H(P) = h$. Following an argument of Schmidt [11], p. 284, we choose the smallest $k \geqq h$ such that $D(P, \alpha) \leqq \Delta_k$. Then, in either case where $k > h$ or $k = h$, we deduce from (I_i) for $i = k, \ldots, j$ and (II_k) or (III_k), respectively, that

$$|\gamma_j - \alpha| \geqq |\gamma_k - \alpha| - |\gamma_j - \gamma_k| > 2\exp(-D(P, \alpha)) - \exp(-\Delta_k).$$

Since $D(P, \alpha) \leqq \Delta_k$, we obtain the desired inequalities.

To state further conditions, we introduce a nondecreasing positive function $\varphi(d)$ of positive integers d with $\varphi(1) \geqq 1$ which satisfies

$$\log d = o(\varphi(d)), \quad d(\log d)^3 = o(\Phi(d)/\varphi(d)) \tag{6.4}$$

as d tends to infinity. Then we impose further three conditions as follows:

(IV_j) $1 \leqq \lambda_j \leqq d_j (\log d_j)^2$.

(V_j) For any pair (P, α) with $H(P) \leqq j$, $d = \deg P \leqq d_j / \log d_j$, we have

$$|\gamma_j - \alpha| > 4\exp(-\Delta_j / \varphi(d_j)).$$

(VI_j) The set $\mathbf{J}_j$, consisting of all real numbers $x \in \mathbf{I}_j = \mathbf{I}(\gamma_j)$ such that

$$|x - \alpha| > 3\exp(-D(P, \alpha)/\varphi(d))$$

holds for all pairs (P, α) with $H(P) \leqq j$, $\alpha \neq \gamma_1, \ldots, \gamma_j$, and $D(P, \alpha) > \Delta_j$ is a nonempty set.

In the following argument to construct γ_j and λ_j, for simplicity, we write $A \gg_j B$ for positive numbers A, B if there exists a positive constant c depending only on j, h_j such that $A \geqq cB$. We now start our inductive construction of γ_j's and λ_j's.

Construction of γ_1 and λ_1: Let ξ_1 be an arbitrary number in $\mathbf{I}_0$. Then, by Proposition 2, there exist infinitely many positive integers d_1 such that $\mathfrak{U}(\xi_1; d_1, h_1) \subset \mathbf{I}_0$, where $\mathfrak{U}(\xi_1; d_1, h_1)$ is the set defined in § 3 after the proof of Proposition 2. We show that there exist $\gamma_1 \in \mathfrak{U}(\xi_1; d_1, h_1)$ and λ_1 satisfying (I_1)–(V_1) whenever d_1 is sufficiently large. Put $\mathfrak{U}_1 = \mathfrak{U}(\xi_1; d_1, h_1)$ with sufficiently large d_1. Obviously, every $\gamma_1 \in \mathfrak{U}_1$ satisfies (I_1). There is no condition (II_1). We claim that there exists λ_1 with (IV_1) such that each $\gamma_1 \in \mathfrak{U}_1$ and λ_1 satisfy (III_1). In fact, for any $\gamma_1 \in \mathfrak{U}_1$ and any pair (P, α) with $H(P) = 1$, $\alpha \neq \gamma_1$, we have

$$\log|\gamma_1 - \alpha| \gg_1 -((d+d_1)\log(d+d_1))/s$$

by the corollary to Lemma 2. In view of (6.1) and (6.3), this implies our claim.

We next show the existence of $\gamma_1 \in \mathfrak{U}_1$ satisfying (V_1). Let $\gamma, \gamma' \in \mathfrak{U}_1$. By Lemma 2, the mutual distance of γ and γ' is larger than $\exp(-4d_1 \log d_1)$. On the other hand, in view of (6.1), (6.2), and (6.4), we have

$$d_1 \log d_1 = o\big(\varDelta_1/\varphi(d_1)\big)$$

as d_1 tends to infinity. Hence, for each pair (P, α), we have at most one element in $\mathfrak{U}_1$ whose distance from α is not larger than $4\exp\big(-\varDelta_1/\varphi(d_1)\big)$. Let A be the number of all elements in $\mathfrak{A}_1$, and let B be that of all pairs (P, α) with $H(P) = 1$, $\deg P \leqq d_1/\log d_1$. Since

$$A \geqq (2h_1+1)^{d_1/9}, \quad B \leqq (d_1/\log d_1)\, 3^{1+d_1/\log d_1},$$

the above observation implies the existence of $\gamma_1 \in \mathfrak{U}_1$ satisfying (V_1). Thus we have obtained γ_1 and λ_1 satisfying (I_1)–(V_1) with sufficiently large d_1.

We can also prove that there exists a positive constant C_1 depending only on φ, Φ, $\varDelta$, h_1 such that γ_1 and λ_1 above satisfy (VI_1) provided that $d_1 > C_1$. Let us postpone the proof.

Construction of γ_j and λ_j for $j > 1$: Assume that $j > 1$ and that

$$\gamma_1, \ldots, \gamma_{j-1}; \quad \lambda_1, \ldots, \lambda_{j-1}$$

with the desired properties have been constructed. In particular, we have a nonempty set $\mathbf{J}_{j-1}$ defined in (VI_{j-1}). Let ξ_j be an arbitrary number in $\mathbf{J}_{j-1}$. Since $\mathbf{J}_{j-1}$ is a subset of $\mathbf{I}_{j-1}$, by Proposition 2, there exist infinitely many positive integers d_j such that

$$\mathfrak{U}(\xi_j; d_j, h_j) \subset \mathbf{I}_{j-1}.$$

As in the step where $j = 1$, we show that there exist $\gamma_j \in \mathfrak{U}(\xi_j; d_j, h_j)$ and λ_j satisfying (I_j)–(V_j) whenever d_j is sufficiently large. Put $\mathfrak{U}_j = \mathfrak{U}(\xi_j; d_j, h_j)$ with sufficiently large d_j. Obviously, every $\gamma_j \in \mathfrak{U}_j$ satisfies (I_j). We claim that every $\gamma_j \in \mathfrak{U}_j$ satisfies (II_j). Let $\gamma_j \in \mathfrak{U}_j$, and let (P, α) be any pair with $H(P) \leqq j-1$, $\alpha \neq \gamma_1, \ldots, \gamma_j$, and $D(P, \alpha) > \varDelta_{j-1}$. Assume first that $D(P,\alpha)/\varphi(d) \leqq c(\xi_j, h_j)\, d_j$, where $c(\xi_j, h_j)$ is the constant appeared in the definition of $\mathfrak{U}_j$. Then, by the definition of $\mathbf{J}_{j-1}$ and that of $\mathfrak{U}_j$, we have

$$|\xi_j - \alpha| > 3\exp\big(-D(P,\alpha)/\varphi(d)\big), \quad |\gamma_j - \xi_j| < \exp\big(-c(\xi_j, h_j)\, d_j\big).$$

Hence, using $|\gamma_j - \alpha| \geqq |\xi_j - \alpha| - |\gamma_j - \xi_j|$, we have the desired inequality since $\varphi(d) \geqq 1$. Assume next that $D(P,\alpha)/\varphi(d) > c(\xi_j, h_j)\, d_j$. Then d is sufficiently large since d_j is sufficiently large. Applying the corollary to Lemma 2, we have

$$\log|\gamma_j-\alpha| \gg_j -\left(\left(d+\frac{D(P,\alpha)}{c(\xi_j,h_j)\varphi(d)}\right)\log\left(d+\frac{D(P,\alpha)}{c(\xi_j,h_j)\varphi(d)}\right)\right)/s.$$

In view of (6.1), (6.3), (6.4), and the above observation for d, this implies the desired inequality.

To show the existence of $\gamma_j \in \mathfrak{U}_j$ and λ_j satisfying the rest of the conditions (III_j)–(V_j), we can apply the same argument as in the case where $j=1$. Thus we have obtained $\gamma_j \in \mathfrak{U}(\xi_j; d_j, h_j)$ and λ_j satisfying (I_j)–(V_j) with sufficiently large d_j. We now turn to the final stage of the proof.

Nonemptiness of $\mathbf{J}_j$: Assume that γ_j and λ_j satisfy (I_j)–(V_j). We show that there is a positive constant C_j depending only on φ, Φ, Δ, j, h_j, $\lambda_1, \ldots, \lambda_{j-1}$ such that γ_j and λ_j satisfy (VI_j) provided that $d_j > C_j$. Let C_j be a sufficiently large constant depending only on the quantities listed above, and assume that $d_j > C_j$. We first claim that, for each $x \in \mathbf{I}_j$,

$$|x-\alpha| > 3\exp\left(-D(P,\alpha)/\varphi(d)\right)$$

holds for all pairs (P,α) with $H(P) \leqq j$, $\alpha \neq \gamma_1, \ldots, \gamma_j$, and

$$D(P,\alpha) > \Delta_j > D(P,\alpha)/\left(\varphi(d)\log d\right).$$

We consider two cases. Assume first that $d \leqq d_j/\log d_j$. Then, by (V_j) and $x \in \mathbf{I}_j$, we have

$$|x-\alpha| \geqq |\gamma_j-\alpha| - |x-\gamma_j| > 4\exp\left(-\Delta_j/\varphi(d_j)\right) - \frac{1}{2}\exp(-\Delta_j),$$

which with $\varphi(d_j) \geqq 1$ implies $|x-\alpha| > 3\exp\left(-\Delta_j/\varphi(d_j)\right)$. Hence, using $D(P,\alpha) > \Delta_j$, $d < d_j$, we obtain the desired inequalities. Assume next that $d > d_j/\log d_j$. Then, by the corollary to Lemma 2 and $x \in \mathbf{I}_j$, we have

$$\log|\gamma_j-\alpha| \gg_j -d(\log d)^2/s, \quad \log|x-\gamma_j| < -\Delta_j.$$

In view of $\Delta_j > D(P,\alpha)/\left(\varphi(d)\log d\right)$, (6.3), (6.4), and $\lambda_j \geqq 1$ in (IV_j) if $H(P)=j$, we deduce from the above inequalities that

$$|\gamma_j-\alpha| \geqq 2|x-\gamma_j|, \quad |\gamma_j-\alpha| > 6\exp\left(-D(P,\alpha)/\varphi(d)\right).$$

Hence we obtain

$$|x-\alpha| \geqq |\gamma_j-\alpha| - |x-\gamma_j| > 3\exp\left(-D(P,\alpha)/\varphi(d)\right),$$

as desired. Thus we have shown our claim.

Noticing the above claim, for each pair (P,α) with

$$H(P) \leqq j, \quad \alpha \neq \gamma_1, \ldots, \gamma_j, \quad D(P,\alpha) > \Delta_j, \quad \text{and} \quad D(P,\alpha)/\left(\varphi(d)\log d\right) \geqq \Delta_j,$$

we define an interval (which may be an empty set)

$$\mathfrak{X}(P,\alpha) = \{x \in \mathbf{I}_j \mid |x-\alpha| \leqq 3\exp\left(-D(P,\alpha)/\varphi(d)\right)\}. \tag{6.5}$$

Then we denote by $\tilde{\mathfrak{X}}$ the union of all these intervals and by $\tilde{\mathfrak{X}}_{d,h}$ the union of all these intervals with $\deg P = d$, $H(P) = h$. Since $\mathbf{J}_j = \mathbf{I}_j \backslash \tilde{\mathfrak{X}}$, our task is to show that the Lebesgue measure $\mu(\tilde{\mathfrak{X}})$ of $\tilde{\mathfrak{X}}$ is smaller than the length $|\mathbf{I}_j|$ of $\mathbf{I}_j$. We next claim that, for any pairs (P, α), (Q, β) with $\deg P = \deg Q = d$, $H(P) = H(Q) = h$, and $\alpha \neq \beta$, $\mathfrak{X}(P, \alpha)$ and $\mathfrak{X}(Q, \beta)$ are disjoint. In fact, by the corollary to Lemma 2, we have

$$\text{(6.6)} \qquad \log|\alpha - \beta| \gg_j -(d \log d)/\max(s, s'),$$

where s, s' are multiplicities of α, β in P, Q, respectively. In view of $D(P, \alpha) > \Delta_j$, (6.2), (6.3), and $\lambda_j \leqq d_j(\log d_j)^2$ in (IV$_j$) if $h = j$, we may assume that $d > c_j$ with a sufficiently large constant c_j which is determined by C_j. Hence (6.4)–(6.6) and $\lambda_j \geqq 1$ in (IV$_j$) if $h = j$ imply our claim. Therefore $\tilde{\mathfrak{X}}_{d,h}$ is the disjoint union of the intervals $\mathfrak{X}_1, \ldots, \mathfrak{X}_m$, say, where $\mathfrak{X}_i = \mathfrak{X}(P_i, \alpha_i)$ with $\deg P_i = d$, $H(P_i) = h$. We wish to show that

$$\text{(6.7)} \qquad \mu(\tilde{\mathfrak{X}}_{d,h}) = \sum_{i=1}^{m} |\mathfrak{X}_i| < 7 \exp(-\Delta_j \log d).$$

For $i = 1, \ldots, m$, let a_i and b_i be the left and the right end points of $\mathfrak{X}_i$, respectively. We may assume that $a_1 < a_2 < \cdots < a_m$. Then put $\mathfrak{Y}_i = [b_i, a_{i+1}]$ for $i = 1, \ldots, m-1$, and put $\mathbf{I}_j = [a_0, b_{m+1}]$. Again by (6.4)–(6.6), $d > c_j$, and $\lambda_j \geqq 1$ in (IV$_j$) if $h = j$, we have $|\mathfrak{X}_i| \leqq |\mathfrak{Y}_i|^{\log d}$ for $i = 1, \ldots, m-1$. Hence we obtain

$$\sum_{i=1}^{m-1} |\mathfrak{X}_i| \leqq \sum_{i=1}^{m-1} |\mathfrak{Y}_i|^{\log d} \leqq \left(\sum_{i=1}^{m-1} |\mathfrak{Y}_i| \right)^{\log d} < \exp(-\Delta_j \log d).$$

Also, by (6.5) and $D(P_m, \alpha_m)/\big(\varphi(d) \log d\big) \geqq \Delta_j$, we have

$$|\mathfrak{X}_m| \leqq 6 \exp\big(-D(P_m, \alpha_m)/\varphi(d)\big) \leqq 6 \exp(-\Delta_j \log d).$$

From these estimates we deduce (6.7). Consequently, using (6.7) with $d > c_j$, $h = 1, \ldots, j$, we obtain

$$\mu(\tilde{\mathfrak{X}}) < \sum_{h=1}^{j} \sum_{d > c_j} 7 \exp(-\Delta_j \log d) = 7j \sum_{d > c_j} d^{-\Delta_j},$$

which implies that $\mu(\tilde{\mathfrak{X}}) < |\mathbf{I}_j|$, as desired. This completes the proof of the nonemptiness of $\mathbf{J}_j$, and hence that of the proposition.

We now prove the denseness stated in Theorems 1 and 2, namely:

(a) $\mathscr{R}(\mathbf{v})$ is dense in $\mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}})$;

(b) $\tilde{\mathscr{R}}(\mathbf{v})$ is dense in $\mathbf{E}(h_0 + 1) \backslash \mathbf{E}(h_0)$.

Proof of (b). We apply Proposition 3 by taking $l = h_0$, $\mathbf{I}_0 \subset \mathbf{E}(h_0 + 1) \backslash \mathbf{E}(h_0)$ and assuming

$$\text{(6.8)} \qquad \lim_{d \to \infty} \frac{\log \Delta(d,h)}{\log d} = v_h$$

for each fixed h with $h \geqq h_0$. Then we have $\{\gamma_j)_{j \in \mathbb{N}}$, $\{\lambda_j\}_{j \in \mathbb{N}}$, and ω with the properties stated in the proposition. We wish to show that $\omega \in \tilde{\mathscr{R}}(\mathbf{v})$. In view of (i_j)'s and (ii), ω is a transcendental number. Since $\omega \in \mathbf{I}_0 \subset \mathbf{E}(h_0+1) \setminus \mathbf{E}(h_0)$, we see from Lemma 5 (ii) that $\tilde{v}(\omega,h) = 0$ for all $h < h_0$. We next show that $\tilde{v}(\omega,h) = v_h$ for all $h \geqq h_0$. We see from the upper bounds in (i_j)'s with (6.8) that $\tilde{v}(\omega,h) \geqq v_h$ for all $h \geqq h_0$. On the other hand, we see from the lower bounds in (i_j)'s with (6.8) and from (ii) with (6.1), (6.3) that $\tilde{v}(\omega,h) \leqq v_h$ for all $h \geqq h_0$. Thus we have shown that $\tilde{v}(\omega,h) = v_h$ for all h. Therefore $\omega \in \tilde{\mathscr{R}}(\mathbf{v})$, as desired. Since we can choose $\mathbf{I}_0$ with $\mathbf{I}_0 \subset \mathbf{E}(h_0+1) \setminus \mathbf{E}(h_0)$ arbitrarily, (b) holds.

Proof of (a). Since $\mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}})$ is the disjoint union of the sets

$$\mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}}) \cap (\mathbf{E}(k+1) \setminus \mathbf{E}(k))$$

for $k \in \mathbb{N}$, it is enough to show that $\mathscr{R}(\mathbf{v})$ is dense in each of these sets. So we choose a positive integer k arbitrarily such that $\mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}}) \cap (\mathbf{E}(k+1) \setminus \mathbf{E}(k))$ is nonempty. For such k, we can easily see that $k \geqq h_0$. We now apply Proposition 3 by taking

$$l = k, \quad \mathbf{I}_0 \subset \mathbf{E}(m_{\mathbf{v}}, M_{\mathbf{v}}) \cap (\mathbf{E}(k+1) \setminus \mathbf{E}(k))$$

and assuming (6.8) for each fixed h with $h \geqq 1$. Then we have $\{\gamma_j\}_{j \in \mathbb{N}}$, $\{\lambda_j\}_{j \in \mathbb{N}}$, and ω with the properties stated in the proposition. We wish to show that $\omega \in \mathscr{R}(\mathbf{v})$. As in the proof of (b), ω is a transcendental number. To show that $v(\omega,h) = v_h$ for all $h < k$, we consider two cases. We first assume that $h_0 > 1$. Then k must be equal to h_0. Hence, by Lemma 4 (iii), $v(\omega,h) = 0$ for all $h < k \ (= h_0)$. We next assume that $h_0 = 1$. If $k = 1$, then we have nothing to prove. If $k > 1$, then we have $|\omega| \leqq 1/k$. Hence, by Lemma 4 (i), $v(\omega,h) = 1$ for all $h < k$. By using $|\omega| \leqq 1/k$, we also have $k \leqq h_0'$ in the case where $m_{\mathbf{v}} = 1/(h_0'+1)$. Therefore we have shown the desired equalities.

We next show that $v(\omega,h) = v_h$ for an arbitrarily fixed $h \geqq k \ (\geqq h_0)$. To show that $v(\omega,h) \geqq v_h$, we consider only the suffices such that $h_j = h$. Then, in view of (i_j)'s, d_j tends to infinity as j tends to infinity. Let Ψ_j be the minimal defining polynomial of γ_j. Since $|\gamma| < h+1$ for any root γ of Ψ_j (cf. § 3), we have

$$|\Psi_j(\omega)| < h(|\omega| + h + 1)^{d_j - 1} |\omega - \gamma_j| .$$

Hence, by (6.2), we have $|\Psi_j(\omega)| < \exp(-\Delta_j/2)$ for all sufficiently large j. This with (6.8) implies that $v(\omega,h) \geqq v_h$. To show that $v(\omega,h) \leqq v_h$, we may assume that $v_h < \infty$. Let ε be an arbitrary positive number. Then, by (6.8) and (1.1), we have a positive constant $c(\varepsilon)$ such that

$$\text{(6.9)} \qquad \Delta(d,h') \leqq c(\varepsilon)\, d^{v_h + \varepsilon}$$

holds for all $d \in \mathbb{N}$ and all h' with $h' \leqq h$. Let P be an arbitrary polynomial with integral coefficients of degree at most d and height at most h, and let α be a root of P nearest to

ω. Our task is to estimate $|P(\omega)|$ from below by assuming that d is sufficiently large. By Lemma 3, we have

$$(6.10) \qquad |P(\omega)| \geqq |\omega - \alpha|^s d^{-4d},$$

where s is the multiplicity of α in P. We first assume that $\alpha = \gamma_j$ for some j. Then, combining (6.10) with the lower bound in (i_j), we have $-s\Delta(d_j, h_j) - 5d\log d$ as a lower bound for $\log|P(\omega)|$. Since γ_j is a root of P, we have $sd_j \leqq d$. Also, by $H(\gamma_j) = \hat{H}(\gamma_j) = h_j$, we have $h_j \leqq h$. Hence, by (6.9), we obtain

$$(6.11) \qquad \log|P(\omega)| \geqq -c(\varepsilon)\, d^{v_h + \varepsilon} - 5d\log d.$$

We next assume that $\alpha \neq \gamma_j$ for all j. Then, combining (6.10) with (ii), we have

$$(6.12) \qquad \log|P(\omega)| \geqq -\lambda_{H(P)}\Phi(d) - 4d\log d.$$

We now deduce the desired inequality from (6.11) and from (6.12) with (6.1). This completes the proof of (a), and that of Theorems 1 and 2.

7. Purely imaginary case and a further problem

We can obtain certain results for purely imaginary numbers analogous to Theorem 1 or Theorem 2. In this section we refer such results (Theorems 3 and 4 below) and propose a further problem. We first state the result analogous to Theorem 1. For each sequence $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ satisfying (1.1) with some positive integer h_0, we define

$$\mathscr{I}(\mathbf{v}) = \{\sqrt{-1}\,\omega \in \sqrt{-1}\,\mathbb{R}\backslash\bar{\mathbb{Q}} \mid v(\sqrt{-1}\,\omega, h) = v_h \text{ for all } h\},$$

and define the real numbers $m'_{\mathbf{v}}$ and $M'_{\mathbf{v}}$ as follows:

(i) If $h_0 > 1$, then $m'_{\mathbf{v}} = \sqrt{h_0}$ and $M'_{\mathbf{v}} = \sqrt{h_0 + 1}$.

(ii) If there exists a positive integer $h'_0 \geqq 1$ such that $v_h = 1$ for all $h < h'_0$ and $v_{h'_0} > 1$, then $m'_{\mathbf{v}} = 1/\sqrt{h'_0 + 1}$ and $M'_{\mathbf{v}} = \sqrt{2}$.

(iii) If $v_h = 1$ for all h, then $m'_{\mathbf{v}} = 0$ and $M'_{\mathbf{v}} = \sqrt{2}$.

Further, for any real numbers a, b with $0 \leqq a < b$, we define

$$\mathbf{E}'(a, b) = \{\sqrt{-1}\,x \in \sqrt{-1}\,\mathbb{R} \mid a < |x| < b\}.$$

In this notation we have the following

Theorem 3. *For each sequence* $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ *satisfying* (1.1) *with some positive integer* h_0, $\mathscr{I}(\mathbf{v})$ *is contained and dense in* $\mathbf{E}'(m'_{\mathbf{v}}, M'_{\mathbf{v}})$.

We next state the result analogous to Theorem 2. For each sequence $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ satisfying (1.2) with some positive integer h_0, we define

$$\tilde{\mathscr{I}}(\mathbf{v}) = \{\sqrt{-1}\,\omega \in \sqrt{-1}\,\mathbb{R}\backslash\mathbb{Q} \mid \tilde{v}(\sqrt{-1}\,\omega, h) = v_h \text{ for all } h\}.$$

Further, for any $a \geqq 1$, we write simply $\mathbf{E}'(a)$ instead of $\mathbf{E}'(1/\sqrt{a}, \sqrt{a})$ ($\mathbf{E}'(1) = \emptyset$). Then we have the following

Theorem 4. *For each sequence* $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ *satisfying* (1.2) *with some positive integer* h_0, $\tilde{\mathscr{I}}(\mathbf{v})$ *is contained and dense in* $\mathbf{E}'(h_0+1)\backslash\mathbf{E}'(h_0)$.

For the proof of these theorems, we need the following proposition which is an easy consequence of Proposition 2.

Proposition 2′. *Let h be a positive integer, and let η be a purely imaginary number in* $\mathbf{E}'(h+1)$ *with* $|\eta| \neq 1$. *Then there exist a positive constant $c(\eta, h)$ depending only on η, h and an infinite set $\mathscr{N}'(\eta, h)$ of positive integers associated to $c(\eta, h)$ such that, for each $d \in \mathscr{N}'(\eta, h)$, there are at least $(2h+1)^{d/18}$ purely imaginary algebraic numbers β satisfying* $\deg\beta = d$, $H(\beta) = \hat{H}(\beta) = h$, *and*

$$(7.1) \qquad |\eta - \beta| < \exp\left(-c(\eta, h)\,d\right).$$

Proof. Put $\xi = \eta^2$. Since $\xi \in \mathbf{E}(h+1)$ with $|\xi| \neq 1$, we can apply Proposition 2 to ξ. So we have a positive constant $c(\xi, h)$ and an infinite sequence $\mathscr{N}(\xi, h)$ of positive integers with the property stated in Proposition 2. Let $d' \in \mathscr{N}(\xi, h)$, and let α be a real algebraic number satisfying $\deg\alpha = d'$, $H(\alpha) = \hat{H}(\alpha) = h$, and

$$|\xi - \alpha| < \exp\left(-c(\xi, h)\,d'\right).$$

Assume that d' is sufficiently large, and put $d = 2d'$. Since $\xi < 0$, we have $\alpha < 0$. Hence the square roots of α are purely imaginary numbers. Let β be the square root of α which satisfies $\eta\beta < 0$. We claim that β is a purely imaginary algebraic number satisfying $\deg\beta = d$, $H(\beta) = \hat{H}(\beta) = h$, and (7.1) with some $c(\eta, h)$. Since $\alpha < 0$, α can not be a square in the number field $\mathbb{Q}(\alpha)$. Therefore, by a well known theorem of Abel and the fact that $\mathbb{Q}(\alpha, \beta) = \mathbb{Q}(\beta)$, β has degree d. So, if we denote by P the minimal defining polynomial of α, then that of β is $P(x^2)$. Hence we have $H(\beta) = \hat{H}(\beta) = h$. Finally, using the equality $|\eta - \beta| = |\xi - \alpha|/|\eta + \beta|$, we obtain the desired upper bound for $|\eta - \beta|$. Thus we have shown our claim. This implies the desired assertion, and the proof of the proposition is completed.

Corollary. *For each positive integer h, the set of all purely imaginary algebraic numbers of height h is contained and dense in* $\mathbf{E}'(h+1)$.

Proof. Let β be a purely imaginary algebraic number of height h, and let P be its minimal defining polynomial. Since we can represent P as $P(x) = P_1(x^2) + xP_2(x^2)$ with $P_1, P_2 \in \mathbb{Z}[x]$ and β is purely imaginary, we have $P_1(\beta^2) = P_2(\beta^2) = 0$. Hence, by the result quoted in § 3 before the corollary to Proposition 2, we have $\beta^2 \in \mathbf{E}(h+1)$, and so $\beta \in \mathbf{E}'(h+1)$. Combining this with Proposition 2′, we obtain the desired assertion. This completes the proof of the corollary.

The proof of Theorems 3 and 4 can be carried out by the similar way to that of Theorems 1 and 2 by using Proposition 2′ and Lemma 2. So we omit it.

Finally, we propose a problem which generalizes Theorems 1–4. For each sequence $\mathbf{v} = \{v_h\}_{h\in\mathbb{N}}$ satisfying (1.1) or (1.2) with some positive integer h_0 and for a root of unity ζ, we define

$$\mathscr{S}(\mathbf{v};\zeta) = \{\zeta\omega \in \zeta\mathbb{R}\backslash\bar{\mathbb{Q}} \mid v(\zeta\omega, h) = v_h \text{ for all } h\}$$

or

$$\tilde{\mathscr{S}}(\mathbf{v};\zeta) = \{\zeta\omega \in \zeta\mathbb{R}\backslash\bar{\mathbb{Q}} \mid \tilde{v}(\zeta\omega, h) = v_h \text{ for all } h\},$$

respectively. Then we propose the following

Problem. *What are the topological closures of $\mathscr{S}(\mathbf{v};\zeta)$ and $\tilde{\mathscr{S}}(\mathbf{v};\zeta)$?*

Since $\mathscr{R}(\mathbf{v}) = \mathscr{S}(\mathbf{v};\pm 1)$, $\tilde{\mathscr{R}}(\mathbf{v}) = \tilde{\mathscr{S}}(\mathbf{v};\pm 1)$, $\mathscr{I}(\mathbf{v}) = \mathscr{S}(\mathbf{v};\pm\sqrt{-1})$, and

$$\tilde{\mathscr{I}}(\mathbf{v}) = \tilde{\mathscr{S}}(\mathbf{v};\pm\sqrt{-1}),$$

Theorems 1–4 solve the problem in the case where $\zeta = \pm 1, \pm\sqrt{-1}$.

Appendix

In this appendix we prove Theorem C′ below which implies Theorem C stated in the introduction.

Theorem C′. *Let h be a positive integer. Then almost all real and almost all complex numbers ω satisfy*

$$(1) \qquad |P(\omega)| > \exp\left(-c(\omega,h)\, d \log((d+1)h)\right)$$

for all nonzero polynomials $P \in \mathbb{Z}[x]$ with $H(P) \leqq h$, where d denotes the degree of P and $c(\omega,h)$ is a positive constant depending only on ω and h.

Remark. Chudnovsky claimed a more general result as follows (cf. [3], Chap. 2, Remark 7.5): *Almost all real and almost all complex numbers ω satisfy*

$$|P(\omega)| > \exp\left(-6\, d \log(dh)\right)$$

for all nonzero polynomials $P \in \mathbb{Z}[x]$ with $dh \geqq c(\omega)$, where d and h denote the degree and the height of P, respectively, and $c(\omega)$ is a positive constant depending only on ω.

However we have not succeeded in proving this assertion.

For the proof of Theorem C′, we need the following simple result.

Lemma 6. *Let λ be a positive number. Let $P \in \mathbb{Z}[x]$ be a polynomial of degree d and height h, and let α be a nonzero root of P with multiplicity s. Assume that*

$$s \geqq \lambda \log((d+1)h).$$

Then we have

$$e^{-\frac{1}{\lambda}} \leqq |\alpha| \leqq e^{\frac{1}{\lambda}}.$$

Proof. This is a consequence of Lemma 1 (iii). In fact, from it we have

$$\max(1, |\alpha|)^s \leqq L(P) \leqq (d+1)h,$$

which with the assumption for s implies the desired upper bound for $|\alpha|$. If we consider the reciprocal $\tilde{P}$ of P, then we get the same upper bound for $|1/\alpha|$. This implies the desired lower bound for $|\alpha|$. Thus the proof of the lemma is completed.

Proof of Theorem C$'$. Since the proof for the assertion in $\mathbb{C}$ is the same as that for the assertion in $\mathbb{R}$, we prove only the assertion in $\mathbb{R}$. For any positive number γ, we denote by $\mathfrak{X}(\gamma)$ the set of all real numbers ω satisfying

$$|P(\omega)| \leqq \exp(-\gamma d \log((d+1)h))$$

for infinitely many polynomials $P \in \mathbb{Z}[x]$ with $H(P) \leqq h$, $P(0) \neq 0$, where d denotes the degree of P. Put

$$\mathfrak{X} = \bigcap_{\gamma > 0} \mathfrak{X}(\gamma), \quad \mathfrak{R} = \mathbb{R} \backslash \mathfrak{X}.$$

Since it is easily seen that each $\omega \in \mathfrak{R}$ satisfies (1) with some $c(\omega, h)$ for all nonzero polynomials $P \in \mathbb{Z}[x]$ with $H(P) \leqq h$, our task is to show that $\mathfrak{X}$ has measure 0. We note that if $\gamma < \gamma'$, then $\mathfrak{X}(\gamma) \supset \mathfrak{X}(\gamma')$. So it is enough to prove the following claim: For any $\varepsilon > 0$, there exists $\gamma > 0$ such that $\mu(\mathfrak{X}(\gamma)) < \varepsilon$, where μ denotes the Lebesgue measure on $\mathbb{R}$.

In the following, for simplicity, we denote by (P, α) an arbitrary element of $\mathbb{Z}[x] \times \bar{\mathbb{Q}}$ with $H(P) \leqq h$, $P(0) \neq 0$, and $P(\alpha) = 0$. Also, for any pair (P, α) as above, we use the letters d and s for denoting the degree of P and the multiplicity of α in P, respectively.

Let γ and λ be positive numbers such that both λ and γ/λ are sufficiently large. For any $\omega \in \mathfrak{X}(\gamma)$, by Lemma 3 and the definition of $\mathfrak{X}(\gamma)$, there exist infinitely many pairs (P, α) satisfying

$$|\omega - \alpha| \leqq \exp\left(-\gamma' \frac{d}{s} \log((d+1)h)\right), \tag{2}$$

where $\gamma' = \gamma - 4$. We here denote by $\mathfrak{X}_1$ (resp. $\mathfrak{X}_2$) the set of all $\omega \in \mathfrak{X}(\gamma)$ satisfying (2) for infinitely many pairs (P, α) with

$$s < \lambda \log((d+1)h) \quad (\text{resp. } s \geqq \lambda \log((d+1)h)).$$

Then $\mathfrak{X}(\gamma) = \mathfrak{X}_1 \cup {}^{\phi}{}_2$. We first show that $\mu(\mathfrak{X}_1) = 0$. For any pair (P, α), we define

$$\mathfrak{I}(P,\alpha)=\left\{\omega\in\mathbb{R}\,\middle|\,|\omega-\alpha|<\exp\left(-\frac{\gamma'}{\lambda}d\right)\right\}.$$

Then, by the definition of $\mathfrak{X}_1$, each $\omega\in\mathfrak{X}_1$ belongs to infinitely many $\mathfrak{I}(P,\alpha)$'s. Since the number of all pairs (P,α) with $\deg P=d$, $H(P)\leqq h$ is less than $d(2h+1)^{d+1}$, we have

$$\mu(\mathfrak{X}_1)<\sum_{d\geqq d_0}2d(2h+1)^{d+1}\exp\left(-\frac{\gamma'}{\lambda}d\right)$$

for arbitrarily large d_0. This implies that $\mu(\mathfrak{X}_1)=0$. We next estimate $\mu(\mathfrak{X}_2)$ from above. For each $\omega\in\mathfrak{X}_2$, by the definition of $\mathfrak{X}_2$ and Lemma 6, there exist infinitely many pairs (P,α) satisfying

$$|\omega-\alpha|<\exp(-\gamma'\log((d+1)h)),\quad e^{-\frac{1}{\lambda}}\leqq|\alpha|\leqq e^{\frac{1}{\lambda}}.$$

Hence every $\omega\in\mathfrak{X}_2$ satisfies $e^{-(1/\lambda)}\leqq|\omega|\leqq e^{1/\lambda}$. This implies that

$$\mu(\mathfrak{X}_2)\leqq 2(e^{\frac{1}{\lambda}}-e^{-\frac{1}{\lambda}}),$$

which with the fact $\mu(\mathfrak{X}_1)=0$ proves our claim above. Therefore the proof of the theorem is completed.

References

[1] *A. Baker*, Transcendental Number Theory, Cambridge Univ. Press, 1979.

[2] *L. Carlitz*, A class of polynomials, Trans. Amer. Math. Soc. **43** (1938), 167–182.

[3] *G.V. Chudnovsky*, Contributions to the theory of transcendental numbers, Amer. Math. Soc., Math. surv. monogr. **19**, Providence, R. I., 1984.

[4] *M. Fried* and *M. Jarden*, Field Arithmetic, Springer, 1986.

[5] *R. Güting*, Polynomials with multiple zeros, Mathematika **14** (1967) 181–196.

[6] *D.R. Hayes*, Explicit class field theory for rational function fields, Trans. Amer. Math. Soc. **189** (1974), 77–91.

[7] *M. Ishibashi*, Bertrand-Tschebysheff theorem in function fields over finite fields, In: Proceedings of the symposium on analytic number theory and related topics, World Scientific (1993), 45–60.

[8] *H. Kornblum*, Über die Primfunktionen in einer arithmetischen Progression, Math. Z. **5** (1919), 100–111.

[9] *K. Mahler*, Zur Approximation der Exponentialfunktion und des Logarithmus, J. reine angew. Math. **166** (1931), 118–150.

[10] *K. Mahler*, An application of Jensen's formula to polynomials, Mathematika **7** (1960), 98–100.

[11] *W.M. Schmidt*, Mahler's *T*-numbers, Proc. Symp. Pure Math. **20** (1971), 275–286.

[12] *A.B. Shidlovskii*, Transcendental Numbers, de Gruyter Stud. Math. **12**, de Gruyter, 1989.

[13] *V.G. Sprindžuk*, On a classification of transcendental numbers (in Russian), Litovsk Mat. Sb. **2** (1962), 215–219.

[14] *V.G. Sprindžuk*, Mahler's problem in metric number theory, Amer. Math. Soc., Transl. math. monogr. **25**, Providence, R. I., 1969.

[15] *V.G. Sprindžuk*, Achievements and problems in Diophantine approximation theory, Uspekhi Mat. Nauk **35** (1980); English transl. in Russian Math. Surv. **35** (1980), 1–80.

Department of Mathematics, Gunma University, Tenjin-cho 1, Kiryu 376, Japan

Eingegangen 20. Januar 1995

J. reine angew. Math. **470** (1996), 51—88

Journal für die reine und angewandte Mathematik

PBW-bases of quantum groups

By *Claus Michael Ringel* at Bielefeld

1. Introduction

Let $\Delta = (a_{ij})_{ij}$ be the Cartan matrix of a finite-dimensional semisimple complex Lie algebra of rank n (see [H]), this is a symmetrizable matrix, and we denote by $(\varepsilon_i)_i$ the minimal symmetrization, thus ε_i are positive integers without a proper common divisor such that $\varepsilon_i a_{ij} = \varepsilon_j a_{ji}$.

Let $\mathbb{Q}(v)$ be the function field in one variable v over the field $\mathbb{Q}$ of rational numbers.

We denote by U^+ the $(+)$-part of the quantum group $U = U_q(\Delta)$ over $\mathbb{Q}(v)$ (as defined by Drinfeld [D] and Jimbo [J1] and modified by Lusztig [L2]), it is the free $\mathbb{Q}(v)$-algebra with generators $E_1, \dots, E_n$ and relations

$$\sum_{t=0}^{n(ij)} (-1)^t \begin{bmatrix} n(ij) \\ t \end{bmatrix}_{\varepsilon_i} E_i^t E_j E_i^{n(ij)-t} = 0,$$

for all $i \neq j$, where $n(ij) = -a_{ij} + 1$; we use the notation

$$[s] = \frac{v^s - v^{-s}}{v - v^{-1}} = v^{s-1} + v^{s-3} + \dots + v^{-s+1},$$

$$[s]! = \prod_{r=1}^{s} [r], \quad \text{and} \quad \begin{bmatrix} s \\ r \end{bmatrix} = \frac{[s]!}{[r]![s-r]!},$$

here, s, r are non-negative integers, and $r \leqq s$; also, given a polynomial f in the variable v and an integer a, we denote by f_a the polynomial obtained from f by replacing v by v^a, for example, $[s]_2 = \dfrac{v^{2s} - v^{-2s}}{v^2 - v^{-2}}$. The considerations in this paper will be restricted to the $(+)$-part U^+, but the reader should observe that one may use the well-known triangular decomposition $U = U^- \otimes U^0 \otimes U^+$, see [Ro], in order to obtain related results for the Borel part $U^0 \otimes U^+$ or even for U itself.

We denote by $\mathbf{e}_1, \dots, \mathbf{e}_n$ the standard basis of $\mathbb{Z}^n$. The Cartan matrix defines on $\mathbb{Z}^n$ a symmetric bilinear form by $(\mathbf{e}_i, \mathbf{e}_j) = \varepsilon_i a_{ij}$. We consider U^+ as a $\mathbb{Z}^n$-graded algebra by assigning to E_i the degree $\mathbf{e}_i$. Given a homogeneous element X of U^+, we denote its degree by $\mathbf{dim}\, X$.

We are going to present a sequence $X_1, \ldots, X_m$ of homogeneous elements of U^+ such that the monomials $X_1^{a_1} \cdots X_m^{a_m}$ form a $\mathbb{Q}(v)$-basis of U^+, and such that for all $i<j$

$$X_j X_i = v^{(\mathbf{dim} X_i, \mathbf{dim} X_j)} X_i X_j + \sum_{I(i,j)} c(a_{i+1}, \ldots, a_{j-1}) X_{i+1}^{a_{i+1}} \cdots X_{j-1}^{a_{j-1}}$$

with coefficients $c(a_{i+1}, \ldots, a_{j-1})$ in $\mathbb{Q}(v)$; here, the index set $I(i,j)$ is the set of sequences $(a_{i+1}, a_{i+2}, \ldots, a_{j-1})$ of natural numbers such that $\sum_{t=i+1}^{j-1} a_t \, \mathbf{dim} X_t = \mathbf{dim} X_i + \mathbf{dim} X_j$. We say that the sequence $X_1, \ldots, X_m$ *generates* a PBW-basis of U^+.

It follows that U^+ is an iterated skew polynomial ring over $\mathbb{Q}(v)$. To be more precise: Let U_j be the subalgebra of U^+ generated by $X_1, \ldots, X_j$. Thus $U_0 = \mathbb{Q}(v)$ and for $1 \leqq j \leqq m$, we have $U_j = U_{j-1}[X_j; \iota_j, \delta_j]$, with an automorphism ι_j and an ι_j-1-derivation δ_j of U_{j-1}. Note that the automorphism ι_j is given explicitly by

$$\iota_j(X_i) = v^{(\mathbf{dim} X_i, \mathbf{dim} X_j)} X_i \quad \text{for} \quad i<j,$$

and we will show that

$$\iota_j \delta_j = v^{(\mathbf{dim} X_j, \mathbf{dim} X_j)} \delta_j \iota_j .$$

The last assertion has the following consequence: we can apply a recent result of Goodearl and Letzter [GL] in order to see that all prime ideals of U^+ are completely prime. We are indebted to Alev and Goodearl for drawing our attention to this problem.

There do exist several investigations dealing with the construction of PBW-bases for U^+ (or related algebras), let us mention papers by Khoroshkin and Tolstoy [KT1], [KT2], [KT3], Levendorskii and Soibelman [LS], Lusztig [L1], [L4], Takeuchi [T], Xi [X], and Yamane [Y1], [Y2]. These investigations usually start with the Drinfeld-Jimbo presentation and use direct calculations, often involving a braid group operation. Here we want to show that the Hall algebra approach as introduced in [R3], [R5], [R7] (see also [L3] and especially the new paper by Green [Gr]) is very suitable to deal with the problem. As we will recall below, one may identify U^+ with the socalled twisted generic Hall algebra $\mathscr{H}_*(\vec{\Delta}) \otimes \mathbb{Q}(v)$. Using this identification, we obtain a special $\mathbb{Q}(v)$-basis of U^+, so that the basis elements may be considered not just as elements, but as algebraic objects with a rich structure: as modules over a finite-dimensional hereditary algebra Λ of finite representation type. Since the basis elements may be interpreted as Λ-modules, one can discuss their module theoretical, homological or geometrical properties: whether they are indecomposable, or multiplicityfree and so on. A slight modification of these basis elements will form the PBW-basis of interest; those elements which correspond to indecomposable modules (together with some ordering) will be a generating sequence for the PBW-basis. The Hall algebra approach allows to use the representation theory of finite-dimensional hereditary algebras in order to derive properties of U^+; in particular, the shape of the Auslander-Reiten quiver of Λ will be of importance. The Auslander-Reiten quiver of Λ encodes a lot of information about the PBW-basis which we will present. As examples, we will write down in full detail the rank 2 cases and the case $\mathbb{A}_n$. In the latter case, we provide an explicit comparison with the PBW-basis presented by Yamane [Y1], [Y2].

Before we use the Hall algebra approach, we want to exhibit our results on U^+ without any reference to the representation theory of finite-dimensional hereditary algebras. Here is the recipe for writing down explicitly a PBW-basis: a generating sequence will be indexed by the positive roots for Δ, thus we will consider suitable orderings of the set of positive roots.

We denote by Φ^+ the set of positive roots for Δ, and we assume that Φ^+ is embedded into $\mathbb{Z}^n$ so that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the simple roots and $(-,-)$ is the corresponding bilinear form.

For any positive root $\mathbf{a}$, let us define an element $X(\mathbf{a})$ of U^+. For the simple roots $\mathbf{e}_i$, take $X(\mathbf{e}_i) = E_i$, thus the generators $E_1, \ldots, E_n$ will belong to our basis, the remaining elements will be constructed inductively. Given a ring R, and $x_1, x_2, t_1, t_2 \in R$ with t_1, t_2 central in R; then the element $t_1 x_1 x_2 - t_2 x_2 x_1$ will be called a *skew commutator* of x_1 and x_2. We are going to construct the generating sequence $X_1, \ldots, X_m$ as iterated skew commutators, starting from $E_1, \ldots, E_n$. (This sequence may also be obtained using a braid group operation on U, as we will show at the end of the paper.)

First of all, we choose some orientation of the edges of the graph of Δ. Recall that one attaches a graph to Δ as follows: it has as vertices the integers $1, 2, \ldots, n$ and the edges are the subsets $\{i,j\}$ with $a_{ij} < 0$. To choose an orientation means to select for any edge $\{i,j\}$ one of the pairs (i,j) or (j,i); in case (i,j) is selected, we draw an arrow $i \to j$. Since the graph of Δ is a forest, it is sufficient to choose a total ordering $\prec$ on the set $\{1, 2, \ldots, n\}$ and to write $i \to j$ in case $\{i,j\}$ is an edge and $i \prec j$. Thus, we just may take the natural ordering of the integers $1 < 2 < \cdots < n$; in this case, $i \to j$ means $a_{ij} \neq 0$ and $i < j$. Usually, in Lie theory, constructions using Δ will not depend on the orientation of the edges. Thus, if we want to stress that we use the chosen orientation in an essential way, we will write $\vec{\Delta}$ instead of Δ.

Here is the first such instance: $\vec{\Delta}$ defines a (usually non-symmetric) bilinear form $\langle -,- \rangle$ on $\mathbb{Z}^n$ as follows: let $\langle \mathbf{e}_i, \mathbf{e}_i \rangle = \varepsilon_i$, and for $i \neq j$, let $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \varepsilon_i a_{ij}$ provided $i \to j$, and zero otherwise.

A pair $(\mathbf{b}, \mathbf{a})$ of positive roots will be called *$\vec{\Delta}$-orthogonal*, provided the following two conditions are satisfied:

$$\langle \mathbf{a}, \mathbf{b} \rangle = 0, \quad \text{and} \quad \langle \mathbf{b}, \mathbf{a} \rangle \leqq 0,$$

and we denote $r_{\mathbf{a}}^{\mathbf{b}} = -\langle \mathbf{b}, \mathbf{a} \rangle$. In the simply-laced cases ($\mathbb{A}_n$, $\mathbb{D}_n$, $\mathbb{E}_6$, $\mathbb{E}_7$, $\mathbb{E}_8$) we have $r_{\mathbf{a}}^{\mathbf{b}} \leqq 1$, in the cases $\mathbb{B}_n$, $\mathbb{C}_n$, $\mathbb{F}_4$, we have $r_{\mathbf{a}}^{\mathbf{b}} \leqq 2$, whereas for $\mathbb{G}_2$, we have $r_{\mathbf{a}}^{\mathbf{b}} \leqq 3$.

Suppose that $(\mathbf{b}, \mathbf{a})$ is a $\vec{\Delta}$-orthogonal pair, and that $X(\mathbf{a})$, $X(\mathbf{b})$ are already defined. If $r = r_{\mathbf{a}}^{\mathbf{b}} \geqq 1$, then also $\mathbf{a} + \mathbf{b}$ is a positive root and we define

$$X(\mathbf{a} + \mathbf{b}) = X(\mathbf{b})\,X(\mathbf{a}) - v^{-r}\,X(\mathbf{a})\,X(\mathbf{b}).$$

If $r = 2$, then one of $\mathbf{a}, \mathbf{b}$ is a short root, the other one is a long root. Consider first the case that $\mathbf{a}$ is a short root, thus $2\mathbf{a} + \mathbf{b}$ is a root and we define

$$X(2\mathbf{a}+\mathbf{b}) = \frac{1}{[2]}\left(X(\mathbf{a}+\mathbf{b})X(\mathbf{a}) - X(\mathbf{a})X(\mathbf{a}+\mathbf{b})\right).$$

Second, in case $\mathbf{a}$ is a long root, then $\mathbf{a}+2\mathbf{b}$ is a root and we define

$$X(\mathbf{a}+2\mathbf{b}) = \frac{1}{[2]}\left(X(\mathbf{b})X(\mathbf{a}+\mathbf{b}) - X(\mathbf{a}+\mathbf{b})X(\mathbf{b})\right).$$

It remains to consider the case $r = 3$, thus we deal with $\mathbb{G}_2$, and we assume that $\mathbf{a}$ is a short root and $\mathbf{b}$ is a long root. Then also $2\mathbf{a}+\mathbf{b}$, $3\mathbf{a}+\mathbf{b}$ and $3\mathbf{a}+2\mathbf{b}$ are positive roots, and we define

$$X(2\mathbf{a}+\mathbf{b}) = \frac{1}{[2]}\left(X(\mathbf{a}+\mathbf{b})X(\mathbf{a}) - v^{-1}X(\mathbf{a})X(\mathbf{a}+\mathbf{b})\right),$$

$$X(3\mathbf{a}+\mathbf{b}) = \frac{1}{[3]}\left(X(2\mathbf{a}+\mathbf{b})X(\mathbf{a}) - v\cdot X(\mathbf{a})X(2\mathbf{a}+\mathbf{b})\right),$$

$$X(3\mathbf{a}+2\mathbf{b}) = \frac{1}{[3]}\left(X(\mathbf{a}+\mathbf{b})X(2\mathbf{a}+\mathbf{b}) - v\cdot X(2\mathbf{a}+\mathbf{b})X(\mathbf{a}+\mathbf{b})\right).$$

We will show that these definitions are well-defined and we obtain in this way a set of elements $X(\mathbf{a})$ labelled by the positive roots $\mathbf{a}$.

Also, we may index the positive roots $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ such that $\langle \mathbf{a}_i, \mathbf{a}_j\rangle > 0$ implies $i \leqq j$. (We will see in Lemma 1 that such an ordering has the following additional property: If $\langle \mathbf{a}_i, \mathbf{a}_j\rangle < 0$, then $i > j$.) Then $X(\mathbf{a}_1), X(\mathbf{a}_2), \ldots, X(\mathbf{a}_m)$ generates a desired PBW-basis.

For any positive root $\mathbf{a}$, let $\varepsilon(\mathbf{a}) = \frac{1}{2}(\mathbf{a}, \mathbf{a})$; it is well-known that $\varepsilon(\mathbf{a})$ is a non-negative integer (it is equal to 1 in the simply-laced cases, it is 1 or 2 for $\mathbb{B}_n$, $\mathbb{C}_n$, $\mathbb{F}_4$, and 1, 2, or 3 in the case $\mathbb{G}_2$); of course, we have $\varepsilon(\mathbf{e}_i) = \varepsilon_i$. Consider the element

$$X(\mathbf{a})^{(t)} = \frac{1}{[t]!_{\varepsilon(a)}} X(\mathbf{a})^t,$$

these elements are called *divided powers*.

Let $\mathscr{A} = \mathbb{Z}[v, v^{-1}]$, and let $U^+_{\mathscr{A}}$ be the $\mathscr{A}$-subalgebra of U^+ generated by the elements $E_i^{(t)}$. We will see that $U^+_{\mathscr{A}}$ contains all the divided powers $X(\mathbf{a})^{(t)}$, thus $U^+_{\mathscr{A}}$ may be considered as an analogue of the Kostant $\mathbb{Z}$-form in classical Lie-theory.

The author is endebted to the referee for very useful comments concerning the presentation of the results.

2. Hall algebras

It has been shown in [R7] that U^+ may be identified with the twisted generic Hall algebra $\mathscr{H} = \mathscr{H}_*(\vec{\Delta}) \otimes \mathbb{Q}(v)$, where $\vec{\Delta}$ is obtained from Δ by choosing some orientation of the edges of the graph of Δ. (For a new, and much better proof we refer to Green [Gr].) For the convenience of the reader, we recall the main definitions.

Choose a k-species $\mathscr{S} = (F_i, {}_iM_j)$ of type $\vec{\Delta}$; we say that $\mathscr{S}$ is *reduced* provided $\varepsilon_i = \dim_k F_i$ for all i. (Readers not familiar with k-species may restrict their attention to the simply-laced cases $\mathbb{A}_n$, $\mathbb{D}_n$, $\mathbb{E}_6$, $\mathbb{E}_7$, $\mathbb{E}_8$. In these cases, we may consider instead of $\mathscr{S}$ a quiver of type Δ and its representations over k.) A reduced k-species $\mathscr{S}$ is given by field extensions F_i over k of degree ε_i, and, for $i \to j$ an F_i-F_j-bimodule ${}_iM_j$ of k-dimension $-\varepsilon_i a_{ij}$. Actually, in case k is a finite field, there is (up to isomorphism) just one reduced k-species of type $\vec{\Delta}$.

We consider representations of $\mathscr{S}$, a representation being of the form (V_i, f_{ij}), with V_i a finite-dimensional right F_i-vector space, and f_{ij}: $V_i \otimes {}_iM_j \to V_j$ an F_j-linear map, for any $i \to j$. Note that the representations of $\mathscr{S}$ are just the finite-dimensional right modules over the tensor algebra of $\mathscr{S}$.

Given a representation M of $\mathscr{S}$, we denote its isomorphism class by $[M]$ and by $\mathbf{dim}\, M$ its dimension vector, it is an element in the Grothendieck group $K_0(\mathscr{S})$ of all representations modulo exact sequences. We identify $\mathbb{Z}^n$ with $K_0(\mathscr{S})$ so that $\mathbf{dim}\, S_i = \mathbf{e}_i$, where S_i is the simple representation of $\mathscr{S}$ corresponding to the vertex i.

It is known [Ga], [DR1] that **dim** furnishes a bijection between the isomorphism classes of the indecomposable representations of $\mathscr{S}$ and the positive roots. For a positive root $\mathbf{a}$, we denote by $M(\mathbf{a}) = M_{\mathscr{S}}(\mathbf{a})$ an indecomposable representation of $\mathscr{S}$ with $\mathbf{dim}\, M(\mathbf{a}) = \mathbf{a}$. In particular, $S_i = M(\mathbf{e}_i)$ is the simple representation corresponding to the vertex i. Given a map α: $\Phi^+ \to \mathbb{N}_0$, we set

$$M(\alpha) = M_{\mathscr{S}}(\alpha) = \bigoplus_{\mathbf{a}} \alpha(\mathbf{a})\, M(\mathbf{a})$$

(for $t \in \mathbb{N}_0$, and any representation N, we denote by tN the direct sum of t copies of N). The theorem of Krull-Remak-Schmidt asserts that in this way, we obtain a bijection between the set of isomorphism classes of representations of $\mathscr{S}$ and the set $\mathscr{B}$ of maps α: $\Phi^+ \to \mathbb{N}_0$. For any $\alpha \in \mathscr{B}$, let $\mathbf{dim}\, \alpha = \sum_{\mathbf{a}} \alpha(\mathbf{a})\mathbf{a}$, thus $\mathbf{dim}\, M(\alpha) = \mathbf{dim}\, \alpha$.

Note that we will identify $\mathbf{a} \in \Phi^+$ with the corresponding characteristic function $\Phi^+ \to \mathbb{N}_0$; in particular, the simple roots $\mathbf{e}_i$ are considered as elements of $\mathscr{B}$. Given two functions $\Phi^+ \to \mathbb{N}_0$, we may add them, and we can take multiples by non-negative integers. However, since we identify positive roots with the corresponding characteristic function, we have to be careful about the addition: the addition inside $\mathbb{Z}^n$ will be denoted by $+$, that inside the set $\mathscr{B}$ will be denoted by $\oplus$. For $t \in \mathbb{N}_0$, t-fold multiple of $\alpha \in \mathscr{B}$ will be denoted by $t\alpha$ (for $t \geqq 2$, this uses the addition in $\mathscr{B}$; note that there are no proper multiples in Φ^+).

In order to define the multiplication of $\mathscr{H}$, we need Hall polynomials and the Euler form. Recall that we have defined a bilinear form $\langle -, - \rangle$ on $\mathbb{Z}^n$. This form is called the Euler form, since for representations M, N of $\mathscr{S}$, we have

$$\begin{aligned}\langle \mathbf{dim}\, M, \mathbf{dim}\, N \rangle &= \sum_{t \geqq 0} (-1)^t \dim_k \operatorname{Ext}^t(M, N) \\ &= \dim_k \operatorname{Hom}(M, N) - \dim_k \operatorname{Ext}^1(M, N),\end{aligned}$$

see [R1]. Note that $(-, -)$ is the symmetrization of $\langle -, - \rangle$.

We will have to consider polynomials in one variable q, with integral coefficients; we usually will consider the corresponding polynomial ring $\mathbb{Z}[q]$ as a subring of $\mathbb{Q}(v)$, where $q = v^2$.

We recall from [R3] that given three elements $\alpha, \beta, \gamma \in \mathscr{B}$, there exists a polynomial $\phi^{\beta}_{\alpha\gamma}$ in $\mathbb{Z}[q]$ such that for k a finite field, and $\mathscr{S}$ a reduced k-species of type $\vec{\Delta}$, the number $\phi^{\beta}_{\alpha\gamma}(|k|)$ is equal to the number of subrepresentations U of $M_{\mathscr{S}}(\beta)$ which are isomorphic to $M_{\mathscr{S}}(\gamma)$ such that $M_{\mathscr{S}}(\beta)/U$ is isomorphic to $M_{\mathscr{S}}(\alpha)$. The polynomials $\phi^{\beta}_{\alpha\gamma}$ are called Hall polynomials (some have been calculated explicitly, see [R4]). Sometimes, it will be convenient to write $\phi^{M(\beta)}_{M(\alpha)M(\gamma)}$ instead of $\phi^{\beta}_{\alpha\gamma}$.

We note the following: If the polynomial $\phi^{\beta}_{\alpha\gamma}$ is non-zero, then $\phi^{\beta}_{\alpha\gamma}(|k|) \neq 0$ for some finite field k. Consider now a reduced k-species $\mathscr{S}$ of type $\vec{\Delta}$, then $M_{\mathscr{S}}(\beta)$ has a submodule U which is isomorphic to $M_{\mathscr{S}}(\gamma)$ with $M_{\mathscr{S}}(\beta)/U$ isomorphic to $M_{\mathscr{S}}(\alpha)$. If U is a direct summand, then $M_{\mathscr{S}}(\beta) \cong M_{\mathscr{S}}(\alpha \oplus \gamma)$. If U is not a direct summand, then $\mathrm{Ext}^1(M_{\mathscr{S}}(\alpha), M_{\mathscr{S}}(\gamma)) \neq 0$.

More generally, we also will need the Hall polynomials $\phi^{\beta}_{\alpha_1,\dots,\alpha_n}$, where $\alpha_1, \dots, \alpha_n, \beta$ are elements of $\mathscr{B}$; again, these are polynomials in a variable q with integer coefficients and with the following property: for any finite field k, and $\mathscr{S}$ a reduced k-species of type $\vec{\Delta}$, the number $\phi^{\beta}_{\alpha_1,\dots,\alpha_n}(|k|)$ is the number of filtrations $M_{\mathscr{S}}(\beta) = N_0 \supseteqq N_1 \supseteqq \cdots \supseteqq N_t = 0$ such that N_{i-1}/N_i is isomorphic to $M_{\mathscr{S}}(\alpha_i)$, for $1 \leqq i \leqq t$.

By definition, $\mathscr{H}$ is the free $\mathbb{Q}(v)$-module with basis the set $\mathscr{B}$ of functions $\Phi^+ \to \mathbb{N}_0$ (or, equivalently, the set of isomorphism classes of representations of some fixed reduced k-species $\mathscr{S}$ of type $\vec{\Delta}$), with multiplication

$$\alpha * \gamma = v^{\langle \mathbf{dim}\,\alpha,\, \mathbf{dim}\,\gamma \rangle} \sum_{\beta \in \mathscr{B}} \phi^{\beta}_{\alpha\gamma}\, \beta,$$

for $\alpha, \gamma \in \mathscr{B}$.

Theorem. *There exists an isomorphism η: $U^+ \to \mathscr{H}$ of $\mathbb{Z}^n$-graded $\mathbb{Q}(v)$-algebras such that $\eta(E_i) = \mathbf{e}_i$.*

For a proof, see [R7], or, better [Gr].

In the last sections, we also will consider $\mathscr{H}_*(\vec{\Delta})$ itself; by definition, it is the $\mathscr{A}$-subalgebra generated by $\mathscr{B}$. Note that $\mathscr{B}$ is a free $\mathscr{A}$-basis of $\mathscr{H}_*(\vec{\Delta})$.

3. The basic formula

We will work with $\mathscr{H}$ instead of U^+. As we have seen, $\mathscr{H}$ is the free $\mathbb{Q}(v)$-module with basis the set $\mathscr{B} = \{\Phi^+ \to \mathbb{N}_0\}$, but sometimes it will be more convenient to work with a slightly modified basis:

We denote by dim: $\mathbb{Z}^n \to \mathbb{Z}$ the linear form given by $\dim \mathbf{e}_i = \varepsilon_i$. For $\alpha \in \mathscr{B}$, let $\dim \alpha = \dim(\mathbf{dim}\,\alpha)$; thus $\dim \alpha = \dim_k M_{\mathscr{S}}(\alpha)$, for any reduced k-species $\mathscr{S}$.

Also, let us recall from [R3] that given $\alpha \in \mathscr{B}$ and any reduced k-species $\mathscr{S}$ of type $\vec{\Delta}$, the k-dimension of the endomorphism ring of $M_{\mathscr{S}}(\alpha)$ is independent of $\mathscr{S}$ and is denoted by $\varepsilon(\alpha)$. Also observe that for a positive root $\mathbf{a}$, the k-dimension of the endomorphism ring of $M_{\mathscr{S}}(\mathbf{a})$ is just $\frac{1}{2}(\mathbf{a}, \mathbf{a})$, thus the two definitions of $\varepsilon(\mathbf{a})$ coincide.

Let

$$\langle\alpha\rangle = v^{-\dim\alpha + \varepsilon(\alpha)}\alpha,$$

the set of these elements $\langle\alpha\rangle$ with $\alpha \in \mathscr{B}$ is again a $\mathbb{Q}(v)$-basis of $\mathscr{H}$; this is the basis we are mainly interested in. Note that $\langle\mathbf{e}_i\rangle = \mathbf{e}_i$. The multiplication formula may be rewritten in this basis as follows:

$$\langle\alpha\rangle * \langle\gamma\rangle = v^{\varepsilon(\alpha)+\varepsilon(\gamma)+\langle\mathbf{dim}\,\alpha,\,\mathbf{dim}\,\gamma\rangle} \sum_{\beta\in\mathscr{B}} v^{-\varepsilon(\beta)}\phi^{\beta}_{\alpha\gamma}\langle\beta\rangle.$$

We have noted that for any reduced k-species $\mathscr{S}$, the k-dimension of the endomorphism ring of $M_{\mathscr{S}}(\alpha)$ is independent of $\mathscr{S}$. Similarly [R3], also the k-dimension of $\mathrm{Hom}(M_{\mathscr{S}}(\alpha), M_{\mathscr{S}}(\beta))$ is independen of $\mathscr{S}$ and will be denoted by $\varepsilon(\alpha, \beta)$. In the same way, the k-dimension of $\mathrm{Ext}^1(M_{\mathscr{S}}(\alpha), M_{\mathscr{S}}(\beta))$ is independent of $\mathscr{S}$ and will be denoted by $\zeta(\alpha, \beta)$. Note that $\langle\mathbf{dim}\,\alpha, \mathbf{dim}\,\beta\rangle = \varepsilon(\alpha,\beta) - \zeta(\alpha,\beta)$.

Proposition 1. *Let $\alpha_1, \ldots, \alpha_t \in \mathscr{B}$ and let us assume that for $i<j$, we have both $\zeta(\alpha_i, \alpha_j) = 0$ and $\varepsilon(\alpha_j, \alpha_i) = 0$. Then*

$$\left\langle \bigoplus_{i=1}^{t} \alpha_i \right\rangle = \langle\alpha_1\rangle * \cdots * \langle\alpha_t\rangle.$$

Proof. It is sufficient to prove the assertion for $t = 2$, the general case follows by induction. Thus, let $\alpha = \alpha_1$, and $\gamma = \alpha_2$. Since $\zeta(\alpha, \gamma) = 0$, the only β with $\phi^{\beta}_{\alpha\gamma} \neq 0$ is $\beta = \alpha \oplus \gamma$. And $\phi^{\alpha\oplus\gamma}_{\alpha\gamma} = 1$, since $\varepsilon(\gamma, \alpha) = 0$. Also,

$$\langle\mathbf{dim}\,\alpha, \mathbf{dim}\,\gamma\rangle = \varepsilon(\alpha,\gamma) - \zeta(\alpha,\gamma) = \varepsilon(\alpha,\gamma).$$

On the other hand, we have

$$\varepsilon(\alpha\oplus\gamma) = \varepsilon(\alpha) + \varepsilon(\gamma) + \varepsilon(\alpha,\gamma) + \varepsilon(\gamma,\alpha) = \varepsilon(\alpha) + \varepsilon(\gamma) + \varepsilon(\alpha,\gamma).$$

Altogether:

$$\begin{aligned}
\langle\alpha\rangle * \langle\gamma\rangle &= v^{-\dim\alpha + \varepsilon(\alpha) - \dim\gamma + \varepsilon(\gamma)}\alpha * \gamma \\
&= v^{-\dim\alpha\oplus\gamma + \varepsilon(\alpha) + \varepsilon(\gamma) + \langle\mathbf{dim}\,\alpha,\,\mathbf{dim}\,\gamma\rangle}\phi^{\alpha\oplus\gamma}_{\alpha\gamma}\,\alpha\oplus\gamma \\
&= v^{-\dim\alpha\oplus\gamma + \varepsilon(\alpha\oplus\gamma)}\alpha\oplus\gamma \\
&= \langle\alpha\oplus\gamma\rangle.
\end{aligned}$$

Theorem 1. *Let $\alpha, \gamma \in \mathscr{B}$ with $\varepsilon(\gamma, \alpha) = 0$ and $\zeta(\alpha, \gamma) = 0$. Then we have*

$$\langle\gamma\rangle * \langle\alpha\rangle = v^{(\mathbf{dim}\,\alpha,\,\mathbf{dim}\,\gamma)}\langle\alpha\rangle * \langle\gamma\rangle + \sum_{\beta\in J(\alpha,\gamma)} c_\beta\langle\beta\rangle,$$

with coefficients c_β in $\mathbb{Z}[v, v^{-1}]$, where $J(\alpha, \gamma)$ is the set of maps $\beta \in \mathscr{B}$ different from $\alpha \oplus \gamma$ such that $\phi^{\beta}_{\gamma,\alpha} \neq 0$.

Proof. The previous proposition shows that

$$\langle\alpha\rangle * \langle\gamma\rangle = \langle\alpha \oplus \gamma\rangle.$$

Thus, let us consider $\langle\gamma\rangle * \langle\alpha\rangle$, it can be written in the form $\sum_\beta c'_\beta \beta$, where β ranges over all elements from $\mathscr{B}$ such that $\phi^{\beta}_{\gamma,\alpha} \neq 0$, and clearly the coefficients c'_β belong to $\mathbb{Z}[v, v^{-1}]$. Of course, we have

$$c'_\beta = c_\beta v^{-\dim\beta + \varepsilon(\beta)}.$$

It remains to calculate the coefficient $c'_{\alpha\oplus\gamma}$. Let $\mathscr{S}$ be a reduced k-species. Let U be a submodule of $M_{\mathscr{S}}(\alpha \oplus \gamma)$ isomorphic to $M_{\mathscr{S}}(\alpha)$, with factor module isomorphic to $M_{\mathscr{S}}(\gamma)$. Clearly, U has to be a direct summand (see [R2], Lemma 2.3.1). Since $\operatorname{Hom}(M_{\mathscr{S}}(\gamma), M_{\mathscr{S}}(\alpha)) = 0$, the theorem of Krull-Remak-Schmidt asserts that U is the image of a homomorphism of the form $(1, f)$: $M_{\mathscr{S}}(\alpha) \to M_{\mathscr{S}}(\alpha) \oplus M_{\mathscr{S}}(\gamma)$, where f is a homomorphism $M_{\mathscr{S}}(\alpha) \to M_{\mathscr{S}}(\gamma)$. In this way, we obtain a bijection between the considered submodules U and the elements of $\operatorname{Hom}(M_{\mathscr{S}}(\alpha), M_{\mathscr{S}}(\gamma))$. Thus, we see that

$$\begin{aligned}\phi^{\alpha\oplus\gamma}_{\gamma\alpha} &= q^{\varepsilon(\alpha,\gamma)}\\ &= v^{2\cdot\varepsilon(\alpha,\gamma)}.\end{aligned}$$

Also,

$$\varepsilon(\alpha \oplus \gamma) = \varepsilon(\alpha) + \varepsilon(\gamma) + \varepsilon(\alpha, \gamma).$$

Finally, we note that

$$\begin{aligned}(\mathbf{dim}\,\alpha, \mathbf{dim}\,\gamma) &= \langle\mathbf{dim}\,\alpha, \mathbf{dim}\,\gamma\rangle + \langle\mathbf{dim}\,\gamma, \mathbf{dim}\,\alpha\rangle\\ &= \varepsilon(\alpha, \gamma) - \zeta(\gamma, \alpha).\end{aligned}$$

Altogether, we see that

$$\begin{aligned}c'_{\alpha\oplus\gamma} &= v^{-\dim\gamma + \varepsilon(\gamma) - \dim\alpha + \varepsilon(\alpha)}\, v^{\langle\mathbf{dim}\,\gamma, \mathbf{dim}\,\alpha\rangle}\, \phi^{\alpha\oplus\gamma}_{\gamma\alpha},\\ &= v^{-\dim\alpha\oplus\gamma + \varepsilon(\alpha\oplus\gamma) - \varepsilon(\alpha,\gamma)}\, v^{-\zeta(\gamma,\alpha)}\, v^{2\cdot\varepsilon(\alpha,\gamma)}\\ &= v^{-\dim\alpha\oplus\gamma + \varepsilon(\alpha\oplus\gamma)}\, v^{(\dim\alpha, \dim\gamma)}.\end{aligned}$$

This shows that

$$c'_{\alpha\oplus\gamma}\, \alpha \oplus \gamma = v^{(\mathbf{dim}\,\alpha, \mathbf{dim}\,\gamma)} \langle\alpha \oplus \gamma\rangle.$$

4. Ordering the positive roots

It is well-known that $\mathscr{S}$ is representation-directed, see [BGP], [DR2]; this means that there exists a total ordering of the positive roots, say $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ such that

$\mathrm{Hom}(M_{\mathscr{S}}(\mathbf{a}_i), M_{\mathscr{S}}(\mathbf{a}_j)) \neq 0$ (or equivalently, $\varepsilon(\mathbf{a}_i, \mathbf{a}_j) \neq 0$) implies that $i \leqq j$: such an ordering will be called $\vec{\Delta}$*-admissible*. The usual visualization using the Auslander-Reiten quiver will be recalled in Section 6.

Lemma 1. *A total ordering* $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ *of the positive roots is* $\vec{\Delta}$*-admissible if and only if the following property is satisfied*: $\langle \mathbf{a}_i, \mathbf{a}_j \rangle > 0$ *implies* $i \leqq j$.

Such an ordering has the additional property: $\langle \mathbf{a}_i, \mathbf{a}_j \rangle < 0$ *implies* $i > j$.

Proof. We write $M(\mathbf{a}_i)$ instead of $M_{\mathscr{S}}(\mathbf{a}_i)$.

First, let us observe that $\mathrm{Ext}^1(M(\mathbf{a}_i), M(\mathbf{a}_j)) \neq 0$ implies $i > j$. For, assume that there exists an exact sequence

$$0 \to M(\mathbf{a}_j) \to N \to M(\mathbf{a}_i) \to 0$$

which does not split. Let $M(\mathbf{a}_t)$ be an indecomposable direct summand of N. Then $\mathrm{Hom}(M(\mathbf{a}_j), M(\mathbf{a}_t)) \neq 0$ and $\mathrm{Hom}(M(\mathbf{a}_t), M(\mathbf{a}_i)) \neq 0$, thus $j \leqq t \leqq i$. We cannot have $\mathbf{a}_j = \mathbf{a}_i$, since $\mathrm{Ext}^1(M(\mathbf{a}_i), M(\mathbf{a}_i)) = 0$. Thus $i > j$. Of course, $i > j$ implies that $\mathrm{Hom}(M(\mathbf{a}_i), M(\mathbf{a}_j)) = 0$. In particular, the groups $\mathrm{Ext}^1(M(\mathbf{a}_i), M(\mathbf{a}_j))$ and $\mathrm{Hom}(M(\mathbf{a}_i), M(\mathbf{a}_j))$ cannot be non-zero at the same time.

As a consequence, we see: if $\langle \mathbf{a}_i, \mathbf{a}_j \rangle > 0$, then

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \dim_k \mathrm{Hom}(M(\mathbf{a}_i), M(\mathbf{a}_j)),$$

whereas, if $\langle \mathbf{a}_i, \mathbf{a}_j \rangle < 0$, then

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = -\dim_k \mathrm{Ext}^1(M(\mathbf{a}_i), M(\mathbf{a}_j)).$$

Assume that there is given an ordering $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ with the property that $\langle \mathbf{a}_i, \mathbf{a}_j \rangle > 0$ implies $i \leqq j$. If $\mathrm{Hom}(M(\mathbf{a}_i), M(\mathbf{a}_j)) \neq 0$, then $\mathrm{Ext}^1(M(\mathbf{a}_i), M(\mathbf{a}_j)) = 0$ and therefore $\langle \mathbf{a}_i, \mathbf{a}_j \rangle > 0$, thus, by assumption $i \leqq j$.

For the converse, let us assume that the ordering $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ is $\vec{\Delta}$-admissible. Let $\langle \mathbf{a}_i, \mathbf{a}_j \rangle > 0$. Then $\mathrm{Hom}(M(\mathbf{a}_i), M(\mathbf{a}_j)) = \langle \mathbf{a}_i, \mathbf{a}_j \rangle > 0$, and therefore $i \leqq j$. On the other hand, let $\langle \mathbf{a}_i, \mathbf{a}_j \rangle < 0$. Then $\mathrm{Ext}^1(M(\mathbf{a}_i), M(\mathbf{a}_j)) = -\langle \mathbf{a}_i, \mathbf{a}_j \rangle < 0$. As we have seen above, we have $i > j$. This completes the proof.

Let us fix some $\vec{\Delta}$-ordering $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$.

Proposition 1′. *For any* $\alpha \in \mathscr{B}$, *we have*

$$\langle \alpha \rangle = \langle \alpha(\mathbf{a}_1)\mathbf{a}_1 \rangle * \cdots * \langle \alpha(\mathbf{a}_m)\mathbf{a}_m \rangle.$$

Proof. For $i < j$, we have $\varepsilon(\mathbf{a}_j, \mathbf{a}_i) = 0$, by assumption, and we also have $\zeta(\mathbf{a}_i, \mathbf{a}_j) = 0$. Thus, the assertion is a direct consequence of Proposition 1.

Given an element w in $\mathscr{H}$, and $t \geqq 0$, then w^{*t} denotes its t-th power. For a positive root $\mathbf{a}$, we also consider a corresponding divided power of $\langle \mathbf{a} \rangle$:

$$\langle \mathbf{a} \rangle^{(*t)} = \frac{1}{[t]!_{\varepsilon(\mathbf{a})}} \langle \mathbf{a} \rangle^{*t}.$$

Lemma 2. *Let* $\mathbf{a}$ *be a positive root, and* $t \geqq 0$. *Then*

$$\langle t\mathbf{a} \rangle = \langle \mathbf{a} \rangle^{(*t)}.$$

Proof. Let $\mathscr{S}$ be a k-species, where k is a finite field. The number of filtrations

$$tM_{\mathscr{S}}(\mathbf{a}) = M_0 \supset M_1 \supset \cdots \supset M_t = 0$$

with factors isomorphic to $M_{\mathscr{S}}(\mathbf{a})$ is given by evaluating the following polynomial at $|k|$:

$$\frac{(v^{2\varepsilon(\mathbf{a})t} - 1)(v^{2\varepsilon(\mathbf{a})(t-1)} - 1) \cdots (v^{2\varepsilon(\mathbf{a})} - 1)}{(v^{2\varepsilon(\mathbf{a})} - 1)^t} = \left(v^{\binom{t}{2}}[t]!\right)_{\varepsilon(\mathbf{a})} = v^{\varepsilon(\mathbf{a})\binom{t}{2}} \cdot [t]!_{\varepsilon(\mathbf{a})}.$$

Let us express $\mathbf{a}^{*t}$ in the basis $\mathscr{B}$. Since $\zeta(\mathbf{a}, \mathbf{a}) = 0$, we see that $\mathbf{a}^{*t}$ is a multiple of $t\mathbf{a}$, namely

$$\mathbf{a}^{*t} = v^{\varepsilon(\mathbf{a})\binom{t}{2}} v^{\varepsilon(\mathbf{a})\binom{t}{2}} \cdot [t]!_{\varepsilon(\mathbf{a})} \cdot t\mathbf{a}.$$

It follows that

$$\begin{aligned}
\langle \mathbf{a} \rangle^{(*t)} &= \frac{1}{[t]!_{\varepsilon(\mathbf{a})}} \langle \mathbf{a} \rangle^{*t} \\
&= \frac{1}{[t]!_{\varepsilon(\mathbf{a})}} v^{-t\dim\mathbf{a} + t\varepsilon(\mathbf{a})} \mathbf{a}^{*t} \\
&= v^{-t\dim\mathbf{a} + t\varepsilon(\mathbf{a})} v^{\varepsilon(\mathbf{a})\binom{t}{2}} v^{\varepsilon(\mathbf{a})\binom{t}{2}} \cdot t\mathbf{a} \\
&= \langle t\mathbf{a} \rangle,
\end{aligned}$$

since $\varepsilon(t\mathbf{a}) = t^2 \varepsilon(\mathbf{a}) = \left(t + 2\binom{t}{2}\right)\varepsilon(\mathbf{a})$. This completes the proof.

We define

$$X_i = \langle \mathbf{a}_i \rangle.$$

Proposition 2. *Let* $\alpha \in \mathscr{B}$, *and set* $\alpha(i) = \alpha(\mathbf{a}_i)$. *Then*

$$\langle \alpha \rangle = X_1^{(*\alpha(1))} * \cdots * X_m^{(*\alpha(m))} = \left(\prod_{i=1}^m \frac{1}{[\alpha(i)]!_{\varepsilon(\mathbf{a}_i)}} \right) X_1^{*\alpha(1)} * \cdots * X_m^{*\alpha(m)}.$$

Proof. According to Proposition 1′, we have

$$\langle \alpha \rangle = \langle \alpha(1)\mathbf{a}_1 \rangle * \cdots * \langle \alpha(m)\mathbf{a}_m \rangle.$$

Lemma 2 asserts that

$$\langle \alpha(i)\mathbf{a}_i \rangle = X_i^{(*\alpha(i))}.$$

This shows the first equality. The second uses just the definition of divided powers.

Theorem 2. *The monomials* $X_1^{*\alpha(1)} * \cdots * X_m^{*\alpha(m)}$ *with* $\alpha(1), \ldots, \alpha(m) \in \mathbb{N}_0$ *form a* $\mathbb{Q}(v)$*-basis of* $\mathcal{H}$, *and for all* $i < j$

$$X_j * X_i = v^{(\mathbf{dim}\, X_i,\, \mathbf{dim}\, X_j)} X_i * X_j + \sum_{I(i,j)} c(a_{i+1}, \ldots, a_{j-1}) X_{i+1}^{*a_{i+1}} * \cdots * X_{j-1}^{*a_{j-1}}$$

with coefficients $c(a_{i+1}, \ldots, a_{j-1})$ *in* $\mathbb{Q}(v)$. *Here, the index set* $I(i,j)$ *is the set of sequences* $(a_{i+1}, a_{i+2}, \ldots, a_{j-1})$ *of natural numbers such that* $\sum_{t=i+1}^{j-1} a_t \mathbf{a}_t = \mathbf{a}_i + \mathbf{a}_j$.

Proof. Given $\alpha(1), \ldots, \alpha(m) \in \mathbb{N}_0$, define $\alpha \in \mathcal{B}$ by $\alpha(\mathbf{a}_i) = \alpha(i)$. According to Proposition 2, we have

$$X_1^{*\alpha(1)} * \cdots * X_m^{*\alpha(m)} = \left(\prod_{i=1}^{m} [\alpha(i)]!_{\varepsilon(\mathbf{a}_i)} \right) \langle \alpha \rangle,$$

thus the given monomials are non-zero scalar multiples of the elements of $\mathcal{B}$, and therefore form a $\mathbb{Q}(v)$-basis of $\mathcal{H}$.

Let $i < j$. We apply Theorem 1 to $\mathbf{a}_i$, $\mathbf{a}_j$. We have to show that for $\beta \in J(\mathbf{a}_i, \mathbf{a}_j)$, the element $\langle \beta \rangle$ is a scalar multiple of some monomial $X_{i+1}^{*a_{i+1}} * \cdots * X_{j-1}^{*a_{j-1}}$ with $\sum_{t=i+1}^{j-1} a_t \mathbf{a}_t = \mathbf{a}_i + \mathbf{a}_j$.

Let $\beta \in J(\mathbf{a}_i, \mathbf{a}_j)$, and let $\beta(t) = \beta(\mathbf{a}_t)$. Since $\phi^{\beta}_{\mathbf{a}_j \mathbf{a}_i} \neq 0$, there is an exact sequence

$$0 \longrightarrow M(\mathbf{a}_i) \xrightarrow{\ f\ } \bigoplus_{t=1}^{m} \beta(t) M(\mathbf{a}_t) \xrightarrow{\ g\ } M(\mathbf{a}_j) \longrightarrow 0,$$

and we write $f = (f_t)_t$ with $f_t \colon M(\mathbf{a}_i) \to \beta(t) M(\mathbf{a}_t)$. Note that the sequence does not split, since otherwise $\beta = \mathbf{a}_i \oplus \mathbf{a}_j$, contrary to the assumption $\beta \in J(\mathbf{a}_i, \mathbf{a}_j)$. Consider some t with $\beta(t) > 0$. We claim that then $f_t \neq 0$. Otherwise, the cokernel of f would split off $\beta(t)$ copies of $M(\mathbf{a}_t)$, and since the cokernel of f is indecomposable, this would mean that the sequence splits. Since $\operatorname{Hom}(M(\mathbf{a}_i), M(\mathbf{a}_t)) \neq 0$, it follows that $i \leqq t$. Also, we can exclude the case $i = t$, since in this case f_t, and therefore also f would be a split monomorphism. Altogether, we see that $i < t$. The dual arguments (applied to g) show that also $t < j$. According to Proposition 2, $\langle \beta \rangle$ is a scalar multiple of $X_{i+1}^{*a_{i+1}} * \cdots * X_{j-1}^{*a_{j-1}}$. Also, the exact sequence exhibited above shows that $\sum_{t=i+1}^{j-1} a_t \mathbf{a}_t = \mathbf{a}_i + \mathbf{a}_j$. This completes the proof.

5. The automorphisms and skew derivations

For any element $\mathbf{d} \in \mathbb{Z}^n$, there exists an automorphism $\iota_{\mathbf{d}}$ of $\mathscr{H}$ given by

$$\iota_{\mathbf{d}}(w) = v^{(\mathbf{dim}\, w, \mathbf{d})} w$$

for w a homogeneous element of $\mathscr{H}$ (of degree $\mathbf{dim}\, w$).

We also note the following rather obvious assertion:

Lemma 3. *Let R be a ring, let ι be an endomorphism of R. For $r \in R$, we define a map $\delta_r\colon R \to R$ by*

$$\delta_r(x) = rx - \iota(x)r \quad \textit{for} \quad x \in R.$$

Then δ_r is a ι-1-derivation.

Proof. The map δ_r clearly is additive. Also, for $x, y \in R$,

$$\begin{aligned}\delta_r(xy) &= rxy - \iota(xy)r \\ &= rxy - \iota(x)ry + \iota(x)ry - \iota(x)\iota(y)r \\ &= \big(rx - \iota(x)r\big)y + \iota(x)\big(ry - \iota(y)r\big) \\ &= \delta_r(x)\cdot y - \iota(x)\cdot\delta_r(y).\end{aligned}$$

One may call δ_r an *inner* ι-1-derivation.

Let K be a commutative ring, and assume that R is a K-algebra and that ι is K-linear. Then also δ_r is K-linear, for any $r \in R$. If R is generated as a K-algebra by $r_1, \ldots, r_n$, and δ is a ι-1-derivation, the values $\delta(r_i)$, with $1 \leqq i \leqq n$, determine δ uniquely.

Let $\mathscr{H}_j$ be the subalgebra of $\mathscr{H}$ generated by $X_1, \ldots, X_j$. Thus $\mathscr{H}_0 = \mathbb{Q}(v)$ and for $1 \leqq j \leqq m$, we have $\mathscr{H}_j = \mathscr{H}_{j-1}[X_j; \iota_j, \delta_j]$, with an automorphism ι_j and a ι_j-1-derivation δ_j of $\mathscr{H}_{j-1}$. Note that the automorphism ι_j of $\mathscr{H}_{j-1}$ is given explicitly by

$$\iota_j(X_i) = v^{(\mathbf{dim}\, X_i,\, \mathbf{dim}\, X_j)} X_i \quad \text{for} \quad i < j;$$

it is just the restriction of $\iota_{\mathbf{a}_j}$ to $\mathscr{H}_{j-1}$.

The ι_j-1-derivation δ_j is given by the formula

$$\begin{aligned}\delta_j(X_i) &= X_j * X_i - \iota_j(X_i) * X_j \\ &= \sum_{I(i,j)} c(a_{i+1}, \ldots, a_{j-1}) X_{i+1}^{*a_{i+1}} * \cdots * X_{j-1}^{*a_{j-1}}\end{aligned}$$

for $i < j$; in particular, δ_j is the restriction of the inner ι_j-1-derivation δ_{X_j} to $\mathscr{H}_{j-1}$.

Altogether, we see that $\mathscr{H}$ is an iterated skew polynomial ring over $\mathbb{Q}(v)$.

Theorem 3. *The automorphism ι_j and the ι_j-1-derivation δ_j of $\mathscr{H}_{j-1}$ satisfy the following relation*:

$$\iota_j \delta_j = v^{(\mathbf{a}_j, \mathbf{a}_j)} \delta_j \iota_j .$$

Proof. Let $i < j$. First of all, we have $\iota_j(X_i) = v^{(\mathbf{a}_i, \mathbf{a}_j)} X_i$, and therefore

$$\delta_j \iota_j(X_i) = v^{(\mathbf{a}_i, \mathbf{a}_j)} \delta_j(X_i) .$$

Let us denote $\mathbf{d} = \mathbf{a}_i + \mathbf{a}_j$. We know that $\delta_j(X_i)$ is a linear combination of monomials of the form $X_{i+1}^{*a_{i+1}} * \cdots * X_{j-1}^{*a_{j-1}}$ where $\sum_{t=i+1}^{j-1} a_t \mathbf{a}_t = \mathbf{a}_i + \mathbf{a}_j = \mathbf{d}$, thus $\delta_j(X_i)$ belongs to $\mathscr{H}_{\mathbf{d}}$. It follows that

$$\iota_j \delta_j(X_j) = v^{(\mathbf{d}, \mathbf{a}_j)} \delta_j(X_i) = v^{(\mathbf{a}_j, \mathbf{a}_j) + (\mathbf{a}_i, \mathbf{a}_j)} \delta_j(X_i) = v^{(\mathbf{a}_j, \mathbf{a}_j)} \delta_j \iota_j(X_i) ,$$

since $(\mathbf{d}, \mathbf{a}_j) = (\mathbf{a}_i + \mathbf{a}_j, \mathbf{a}_j) = (\mathbf{a}_i, \mathbf{a}_j) + (\mathbf{a}_j, \mathbf{a}_j)$.

As an application we obtain:

Corollary. *Any prime ideal of U^+ is completely prime.*

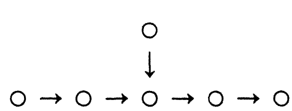
Proof. The corresponding assertion for $\mathscr{H}$ is a direct consequence of Theorem 2, Theorem 3 and a recent result of Goodearl and Letzter [GL], Theorem 2.3, see also [Go].

6. The Auslander-Reiten quiver

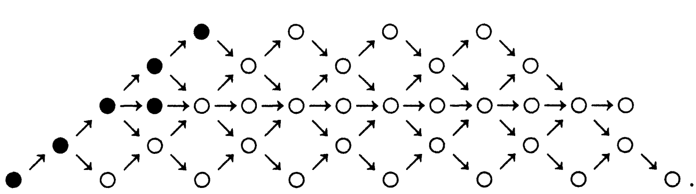
The Auslander-Reiten quiver is a convenient tool to visualize the module category of a finite-dimensional algebra. We want to point out that the Auslander-Reiten quiver is also extremely useful for dealing with the corresponding Hall algebras. Let us formulate a combinatorial version of those definitions which are needed.

As vertices of the Auslander-Reiten quiver Γ take the positive roots. If $\mathbf{a}, \mathbf{b}$ are positive roots, write $\mathbf{a} \to \mathbf{b}$ provided the following conditions are satisfied: first, $\mathbf{a} \neq \mathbf{b}$, second $\langle \mathbf{a}, \mathbf{b} \rangle > 0$, and third, if $\mathbf{c}$ is a positive root with $\langle \mathbf{a}, \mathbf{c} \rangle > 0$ and $\langle \mathbf{c}, \mathbf{b} \rangle > 0$, then $\mathbf{a} = \mathbf{c}$ or $\mathbf{c} = \mathbf{b}$. In general, if $\langle \mathbf{a}, \mathbf{b} \rangle > 0$, then there exists a path $\mathbf{a} = \mathbf{a}_0 \to \mathbf{a}_1 \to \cdots \to \mathbf{a}_t = \mathbf{b}$ (of length $t \geqq 0$) from $\mathbf{a}$ to $\mathbf{b}$.

For example, the Auslander-Reiten quiver for $\mathbb{E}_6$ with the following orientation



is of the form



We may endow Γ with a valuation by associating to any arrow $\mathbf{a} \to \mathbf{b}$ a pair of positive integers, namely $\left(\frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\varepsilon(b)}, \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\varepsilon(a)}\right)$. In case the valuation of an arrow is (1,1), one usually drops these numbers, this happens for all the arrows in the simply-laced cases; in the cases $\mathbb{B}_n$, $\mathbb{C}_n$, $\mathbb{F}_4$, some arrows will carry a valuation (1, 2) or (2, 1); for $\mathbb{G}_2$, all the arrows will carry a valuation (1, 3) or (3, 1). Note that the valuation of Γ allows to recover $\varepsilon(\mathbf{a})$ for any vertex $\mathbf{a}$.

The Auslander-Reiten quiver is usually considered as a translation quiver: some of the vertices are called projective, for any of the remaining vertices, say $\mathbf{a}$, there is defined a vertex $\tau\mathbf{a}$, such that there exists an arrow $\tau\mathbf{a} \to \mathbf{b}$ if and only if there exists an arrow $\mathbf{b} \to \mathbf{a}$. Here is the combinatorial recipe:

For any vertex i of Δ, let $\bar{\sigma}_i$ be the reflection in $\mathbb{Z}^n$ at $\mathbf{e}_i$ with respect to the symmetric bilinear form $(-,-)$. The orientation $\vec{\Delta}$ determines a unique Coxeter element C; for example, if we start with the natural ordering $1 < 2 < \cdots < n$ of the vertex set, then $C = \bar{\sigma}_1 \bar{\sigma}_2 \cdots \bar{\sigma}_n$. A positive root $\mathbf{a}$ will be said to be *projective*, provided $C(\mathbf{a})$ is no longer positive; for the remaining positive roots, let $\tau\mathbf{a} = C(\mathbf{a})$. Note that for any non-projective positive root $\mathbf{a}$, there are paths from $\tau\mathbf{a}$ to $\mathbf{a}$, and all are of length 2. (In the drawing of an Auslander-Reiten quiver, the translation τ usually will correspond to a shift from right to left; sometimes one connects $\mathbf{a}$ and $\tau\mathbf{a}$ by a dotted line; in the drawing above, the projective vertices have been denoted by $\bullet$.) The translation τ is very useful. For example, given positive roots $\mathbf{a}$, $\mathbf{b}$, then either $\mathbf{a}$ is projective and then $\zeta(\mathbf{a}, \mathbf{b}) = 0$, or else $\zeta(\mathbf{a}, \mathbf{b}) = \varepsilon(\mathbf{b}, \tau\mathbf{a})$. In particular, if $\zeta(\mathbf{a}, \mathbf{b}) \neq 0$, then there is a path from $\mathbf{b}$ to $\mathbf{a}$.

A path $\mathbf{a}_1 \to \mathbf{a}_2 \to \cdots \to \mathbf{a}_t$ in the Auslander-Reiten quiver is said to be *sectional*, provided we have $\mathbf{a}_{i-1} \neq \tau\mathbf{a}_{i+1}$, for all $1 < i < t$. Finally, two vertices $\mathbf{a}$, $\mathbf{b}$ are said to be *incomparable* provided there is no path from $\mathbf{a}$ to $\mathbf{b}$, and no path from $\mathbf{b}$ to $\mathbf{a}$.

What kind of information can be read off from the Auslander-Reiten quiver? First of all, the $\vec{\Delta}$-admissible orderings of the positive roots are just the refinements of the arrow relation: in order to deal with a PBW-basis, we may start with any total ordering $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ of the positive roots such that $\mathbf{a}_i \to \mathbf{a}_j$ implies $i < j$. Given positive roots $\mathbf{a}$, $\mathbf{b}$, there exists a $\vec{\Delta}$-admissible ordering $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ such that $\mathbf{a} = \mathbf{a}_i$, $\mathbf{b} = \mathbf{a}_j$ with $i < j$, if and only if there is no path from $\mathbf{b}$ to $\mathbf{a}$.

Also, *in case there is an arrow* $\mathbf{a} \to \mathbf{b}$ *in the Auslander-Reiten quiver, then*

$$\langle \mathbf{b} \rangle * \langle \mathbf{a} \rangle = v^{\max(\varepsilon(\mathbf{a}), \varepsilon(\mathbf{b}))} \langle \mathbf{a} \rangle * \langle \mathbf{b} \rangle .$$

Actually, the same formula is true in the general case that there exists a sectional path $\mathbf{a} = \mathbf{a}_1 \to \mathbf{a}_2 \cdots \to \mathbf{a}_t = \mathbf{b}$ (note that the existence of such a sectional path implies that we have $\operatorname{Ext}^1(M_{\mathcal{S}}(\mathbf{a}) \oplus M_{\mathcal{S}}(\mathbf{b}), M_{\mathcal{S}}(\mathbf{a}) \oplus M_{\mathcal{S}}(\mathbf{b})) = 0$). Of course, *in case the vertices* $\mathbf{a}$, $\mathbf{b}$ *are incomparable, then*

$$\langle \mathbf{a} \rangle * \langle \mathbf{b} \rangle = \langle \mathbf{b} \rangle * \langle \mathbf{a} \rangle .$$

As we have seen above, the main problem to be solved is an effective procedure for calculating the various skew derivations.

Lemma 4. *Let* **a**, **b** *be positive roots, and assume that there is no path from* **b** *to* **a**. *If* $\delta_{\langle \mathbf{b}\rangle}(\langle \mathbf{a}\rangle) \neq 0$, *then* $\zeta(\mathbf{b}, \mathbf{a}) \neq 0$, *and therefore, there is a path from* **a** *to* **b**.

Proof. Since there is no path from **b** to **a**, we know that $\zeta(\mathbf{a}, \mathbf{b}) = 0$, thus $\langle \mathbf{a}\rangle * \langle \mathbf{b}\rangle$ is a multiple of $\langle \mathbf{a} \oplus \mathbf{b}\rangle$. If we assume that $\zeta(\mathbf{b}, \mathbf{a}) = 0$, then also $\langle \mathbf{b}\rangle * \langle \mathbf{a}\rangle$ is a multiple of $\langle \mathbf{a} \oplus \mathbf{b}\rangle$, therefore $\delta_{\langle \mathbf{b}\rangle}(\langle \mathbf{a}\rangle) = 0$.

Thus, we have to consider pairs **a**, **b** with a path from **a** to **b**. If we express $\delta_{\langle \mathbf{b}\rangle}(\langle \mathbf{a}\rangle)$ as a linear combination

$$\sum_{\beta \in J(\mathbf{a},\mathbf{b})} c_\beta \langle \beta \rangle$$

as in Theorem 1, the index set $J(\mathbf{a}, \mathbf{b})$ will involve only $\beta \in \mathscr{B}$ such that $\beta(\mathbf{c}) \neq 0$ implies that there exists paths of length at least 1 from **a** to **c** and from **c** to **b**.

7. The rank 2 cases

Assume that we deal with a Cartan matrix of the form

$$\begin{pmatrix} 2 & a_{12} \\ a_{21} & 2 \end{pmatrix};$$

where one of the numbers $-a_{12}, -a_{21}$ is equal to 1, whereas the other is $r = -\varepsilon_1 a_{12} = -\varepsilon_2 a_{21} = 1, 2,$ or 3. We work with the orientation $1 \to 2$. Note that $\mathbf{d} = \mathbf{e}_1 + \mathbf{e}_2$ is a positive root and let us stress that $\mathbf{e}_i = \langle \mathbf{e}_i \rangle$.

We have

$$\mathbf{e}_2 * \mathbf{e}_1 = \mathbf{e}_1 \oplus \mathbf{e}_2,$$

whereas

$$\mathbf{e}_1 * \mathbf{e}_2 = v^{-r}(\mathbf{d} + \mathbf{e}_1 \oplus \mathbf{e}_2).$$

Also note that $\dim \mathbf{d} = \varepsilon_1 + \varepsilon_2$, and $\varepsilon(\mathbf{d}) = \varepsilon_1 + \varepsilon_2 - r$, thus $\langle \mathbf{d} \rangle = v^{-r}\mathbf{d}$. It follows that

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \langle \mathbf{d} \rangle = v^{-r}\mathbf{d} = \mathbf{e}_1 * \mathbf{e}_2 - v^{-r}\mathbf{e}_2 * \mathbf{e}_1.$$

Consider now the cases $r \geqq 2$, and let us assume that $\varepsilon_2 = 1$, thus $\mathbf{a}_2$ is a short root, whereas $\mathbf{a}_1$ is a long one. On the one hand, we have

$$\mathbf{e}_2 * \langle \mathbf{d} \rangle = \langle \mathbf{d} \oplus \mathbf{e}_2 \rangle = v^{-r+1}\mathbf{d} \oplus \mathbf{e}_2,$$

where we use that $\dim \mathbf{d} \oplus \mathbf{e}_2 = r + 2$ and $\varepsilon(\mathbf{d} \oplus \mathbf{e}_2) = 3$. On the other hand, we use that $\zeta(\mathbf{d}, \mathbf{e}_2) = r - 1$, and a rather easy calculation of Hall polynomials, in order to see that

$$\begin{aligned} \langle \mathbf{d} \rangle * \mathbf{e}_2 &= v^{-r}\mathbf{d} * \mathbf{e}_2 \\ &= v^{-2r+1}\big((q+1)(\mathbf{e}_1 + 2\mathbf{e}_2) + q(\mathbf{d} \oplus \mathbf{e}_2)\big) \\ &= v^{-2r+2}(v + v^{-1})(\mathbf{e}_1 + 2\mathbf{e}_2) + v^{-2r+3}\mathbf{d} \oplus \mathbf{e}_2. \end{aligned}$$

It follows that

$$\mathbf{e}_1 + 2\mathbf{e}_2 = \frac{v^{2r-2}}{[2]}\left(\langle \mathbf{d}\rangle * \mathbf{e}_2 - v^{-r+2}\cdot \mathbf{e}_2 * \langle \mathbf{d}\rangle\right).$$

After these preparations, we may consider in detail the different cases:

Case $\mathbb{A}_2$. In this case, we have $r = 1$, thus

$$\langle \mathbf{e}_1 + \mathbf{e}_2\rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-1}\cdot \mathbf{e}_2 * \mathbf{e}_1 .$$

The corresponding Auslander-Reiten quiver is of the form

$$\begin{array}{ccccc} & \nearrow & \mathbf{e}_1+\mathbf{e}_2 & \searrow & \\ \mathbf{e}_2 & & \cdots & & \mathbf{e}_1 . \end{array}$$

Case $\mathbb{B}_2$, with $\varepsilon_1 = 2$. The considerations above show that

$$\langle \mathbf{e}_1 + \mathbf{e}_2\rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-2}\cdot \mathbf{e}_2 * \mathbf{e}_1 ,$$

$$\langle \mathbf{e}_1 + 2\mathbf{e}_2\rangle = \frac{1}{[2]}\left(\langle \mathbf{e}_1 + \mathbf{e}_2\rangle * \mathbf{e}_2 - \mathbf{e}_2 * \langle \mathbf{e}_1 + \mathbf{e}_2\rangle\right);$$

here, we have used that $\dim \mathbf{e}_1 + 2\mathbf{e}_2 = 4$, whereas $\varepsilon(\mathbf{e}_1 + 2\mathbf{e}_2) = 2$, so that

$$\langle \mathbf{e}_1 + 2\mathbf{e}_2\rangle = v^{-4+2}(\mathbf{e}_1 + 2\mathbf{e}_2) = \frac{1}{[2]}\left(\langle \mathbf{d}\rangle * \mathbf{e}_2 - \mathbf{e}_2 * \langle \mathbf{d}\rangle\right).$$

In this case, we deal with the Auslander-Reiten quiver

$$\begin{array}{ccccccccc} & \nearrow & \mathbf{e}_1+2\mathbf{e}_2 & \searrow & & & \cdots & \nearrow & \mathbf{e}_1 \\ \mathbf{e}_2 & & \cdots & & \mathbf{e}_1+\mathbf{e}_2 . & & & & \end{array}$$

The arrows $\nearrow$ carry the valuation $(1, 2)$, the arrows $\searrow$ carry the valuation $(2, 1)$.

In Appendix 1, we will present the multiplication table for the elements

$$X_1 = \langle \mathbf{e}_2\rangle, \quad X_2 = \langle \mathbf{e}_1 + 2\mathbf{e}_2\rangle, \quad X_3 = \langle \mathbf{e}_1 + \mathbf{e}_2\rangle, \quad X_4 = \langle \mathbf{e}_1\rangle .$$

Case $\mathbb{B}_2$, with $\varepsilon_1 = 1$. Here, a similar calculation shows that

$$\langle \mathbf{e}_1 + \mathbf{e}_2\rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-2}\cdot \mathbf{e}_2 * \mathbf{e}_1 ,$$

$$\langle 2\mathbf{e}_1 + \mathbf{e}_2\rangle = \frac{1}{[2]}\left(\mathbf{e}_1 * \langle \mathbf{e}_1 + \mathbf{e}_2\rangle - \langle \mathbf{e}_1 + \mathbf{e}_2\rangle * \mathbf{e}_1\right).$$

The corresponding Auslander-Reiten quiver is of the form

$$\begin{array}{ccccccccc} & \nearrow & \mathbf{e}_1+\mathbf{e}_2 & \searrow & & & \cdots & \nearrow & \mathbf{e}_1 \\ \mathbf{e}_2 & & \cdots & & 2\mathbf{e}_1+\mathbf{e}_2 . & & & & \end{array}$$

Here, the arrows $\nearrow$ carry the valuation $(2,1)$, the arrows $\searrow$ carry the valuation $(1,2)$.

Case $\mathbb{G}_2$, with $\varepsilon_1 = 3$. The Auslander-Reiten quiver looks as follows:

$$\begin{array}{ccccccccccc} & & \mathbf{e}_1+3\mathbf{e}_2 & & \cdots & & 2\mathbf{e}_1+3\mathbf{e}_2 & & \cdots & & \mathbf{e}_1 \\ & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & \\ \mathbf{e}_2 & & \cdots & & \mathbf{e}_1+2\mathbf{e}_2 & & \cdots & & \mathbf{e}_1+\mathbf{e}_2\,. & & \end{array}$$

The arrows $\nearrow$ carry the valuation $(1,3)$, the arrows $\searrow$ carry the valuation $(3,1)$.

For any positive root $\mathbf{a}$, we can express $\langle \mathbf{a} \rangle$ as a skew commutator as follows:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-3} \mathbf{e}_2 * \mathbf{e}_1 ,$$

$$\langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = \frac{1}{[2]} (\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle * \mathbf{e}_2 - v^{-1} \mathbf{e}_2 * \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle),$$

$$\langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle = \frac{1}{[3]} (\langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle * \mathbf{e}_2 - v \cdot \mathbf{e}_2 * \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle),$$

$$\langle 2\mathbf{e}_1 + 3\mathbf{e}_2 \rangle = \frac{1}{[3]} (\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle * \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle - v \cdot \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle * \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle).$$

The first equality has been shown above. For the second equality, we only have to add to previous considerations that $\langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = v^{-5+1}(\mathbf{e}_1 + 2\mathbf{e}_2)$, since in this case, $\varepsilon(\mathbf{e}_1 + 2\mathbf{e}_2) = 1$.

The calculation of $\langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle$ proceeds as follows: We have

$$\mathbf{e}_2 * \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = \langle (\mathbf{e}_1 + 2\mathbf{e}_2) \oplus \mathbf{e}_2 \rangle = v^{-2}((\mathbf{e}_1 + 2\mathbf{e}_2) \oplus \mathbf{e}_2),$$

and, on the other hand,

$$\begin{aligned} \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle * \mathbf{e}_2 &= v^{-5}((q^2 + q + 1)(\mathbf{e}_1 + 3\mathbf{e}_2) + q^2((\mathbf{e}_1 + 2\mathbf{e}_2) \oplus \mathbf{e}_2)) \\ &= v^{-3}([3](\mathbf{e}_1 + 3\mathbf{e}_2)) + v^{-1}((\mathbf{e}_1 + 2\mathbf{e}_2) \oplus \mathbf{e}_2). \end{aligned}$$

Since $\mathbf{e}_1 + 3\mathbf{e}_2$ has dimension 6, and the dimension of its endomorphism ring is 3, we see that $\langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle = v^{-3}(\mathbf{e}_1 + 3\mathbf{e}_2)$.

A similar proof shows the last equality.

In Appendix 1, we will present the multiplication table for the elements

$$\begin{array}{lll} X_1 = \langle \mathbf{e}_2 \rangle, & X_2 = \langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle, & X_3 = \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle, \\ X_4 = \langle 2\mathbf{e}_1 + 3\mathbf{e}_2 \rangle, & X_5 = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle, & X_6 = \langle \mathbf{e}_1 \rangle. \end{array}$$

Case $\mathbb{G}_2$, with $\varepsilon_1 = 1$. The Auslander-Reiten quiver looks as follows:

$$\begin{array}{ccccccccccc} & \nearrow & \mathbf{e}_1+\mathbf{e}_2 & \searrow & \cdots & \nearrow & 2\mathbf{e}_1+\mathbf{e}_2 & \searrow & \cdots & \nearrow & \mathbf{e}_1 \\ \mathbf{e}_2 & & \cdots & & 3\mathbf{e}_1+2\mathbf{e}_2 & & \cdots & & 3\mathbf{e}_1+\mathbf{e}_2\,. & & \end{array}$$

This time, the arrows $\nearrow$ carry the valuation (3, 1), the arrows $\searrow$ carry the valuation (1, 3).

For any positive root $\mathbf{a}$, we can express $\langle \mathbf{a} \rangle$ as a skew commutator as follows:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-3}\mathbf{e}_2 * \mathbf{e}_1\,,$$

$$\langle 2\mathbf{e}_1 + \mathbf{e}_2 \rangle = \frac{1}{[2]}\left(\mathbf{e}_1 * \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle - v^{-1}\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle * \mathbf{e}_1\right),$$

$$\langle 3\mathbf{e}_1 + \mathbf{e}_2 \rangle = \frac{1}{[3]}\left(\mathbf{e}_1 * \langle 2\mathbf{e}_1 + \mathbf{e}_2 \rangle - v \cdot \langle 2\mathbf{e}_1 + \mathbf{e}_2 \rangle * \mathbf{e}_1\right),$$

$$\langle 3\mathbf{e}_1 + 2\mathbf{e}_2 \rangle = \frac{1}{[3]}\left(\langle 2\mathbf{e}_1 + \mathbf{e}_2 \rangle * \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle - v \cdot \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle * \langle 2\mathbf{e}_1 + \mathbf{e}_2 \rangle\right).$$

We are going to present the elements of the form $\langle \mathbf{e}_1 + t\mathbf{e}_2 \rangle$ with $1 \leqq t \leqq a_{12}$ as linear combinations of monomials; this will be needed at the end of the paper.

Case $\mathbb{A}_2$, thus $\varepsilon_1 = 1$, $\varepsilon_2 = 1$:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-1}\mathbf{e}_2 * \mathbf{e}_1\,.$$

Case $\mathbb{B}_2$, with $\varepsilon_1 = 2$, $\varepsilon_2 = 1$:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-2}\mathbf{e}_2 * \mathbf{e}_1\,.$$

Case $\mathbb{G}_2$, with $\varepsilon_1 = 3$, $\varepsilon_2 = 1$:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-3}\mathbf{e}_2 * \mathbf{e}_1\,.$$

Case $\mathbb{B}_2$, with $\varepsilon_1 = 1$, $\varepsilon_2 = 2$:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-2}\mathbf{e}_2 * \mathbf{e}_1\,,$$

$$\langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2^{(*2)} - v^{-1}\mathbf{e}_2 * \mathbf{e}_1 * \mathbf{e}_2 + v^{-2}\mathbf{e}_2^{(*2)} * \mathbf{e}_1\,.$$

Case $\mathbb{G}_2$, with $\varepsilon_1 = 1$, $\varepsilon_2 = 3$:

$$\langle \mathbf{e}_1 + \mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2 - v^{-3}\mathbf{e}_2 * \mathbf{e}_1\,,$$

$$\langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2^{(*2)} - v^{-2}\mathbf{e}_2 * \mathbf{e}_1 * \mathbf{e}_2 + v^{-4}\mathbf{e}_2^{(*2)} * \mathbf{e}_1\,,$$

$$\langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle = \mathbf{e}_1 * \mathbf{e}_2^{(*3)} - v^{-1}\mathbf{e}_2 * \mathbf{e}_1 * \mathbf{e}_2^{(*2)} + v^{-2}\mathbf{e}_2^{(*2)} * \mathbf{e}_1 * \mathbf{e}_2 - v^{-3}\mathbf{e}_2^{(*2)} * \mathbf{e}_1\,.$$

For arbitrary $\vec{\Delta}$, we will be interested in the elements of the form $\langle \mathbf{e}_j + t\mathbf{e}_i \rangle$ with $j \neq i$ and $0 \leqq t \leqq a_{ij}$. Here is the general rule for expressing them as linear combinations of monomials:

Proposition 3. *Let $0 \leqq t \leqq -a_{ij}$. If i is a sink for $\vec{\Delta}$, then*

$$\langle \mathbf{e}_j + t\mathbf{e}_i \rangle = \sum_{r+s=t} (-1)^r v^{-\varepsilon_i r(-a_{ij}-t+1)} E_i^{(r)} E_j E_i^{(s)};$$

if i is a source, then

$$\langle \mathbf{e}_j + t\mathbf{e}_i \rangle = \sum_{r+s=t} (-1)^r v^{-\varepsilon_i r(-a_{ij}-t+1)} E_i^{(s)} E_j E_i^{(r)}.$$

Proof. We may restrict to the case where i, j are the only vertices. If we assume that i is a sink, then we relabel the vertices as follows: $j = 1$, $i = 2$, and we use the previous consideration. The case when i is a source follows by duality.

A direct proof of Proposition 3 may be given along the line of the proof of the fundamental relations, see [R5].

8. The inductive construction of exceptional modules

Let Λ be a finite-dimensional hereditary k-algebra (for example, the tensor algebra of a k-species as above), let n be the number of isomorphism classes of simple Λ-modules.

Recall that a (finite-dimensional) Λ-module M is said to be *exceptional* provided its endomorphism ring is a division ring and $\mathrm{Ext}^1(M, M) = 0$. A set M_i $(i \in I)$ of modules is said to be *orthogonal* provided $\mathrm{Hom}(M_i, M_j) = 0$ for all $i \neq j$.

Proposition 4. *Let M be a non-simple exceptional Λ-module. Then there exist orthogonal exceptional modules M_1, M_2, and an exact sequence*

$$0 \to a_2 M_2 \to M \to a_1 M_1 \to 0,$$

with $a_1, a_2 \geqq 1$.

Remark. In the case of an algebraically closed field, the result is due to Schofield [S].

Crawley-Boevey recently has shown that a braid group operates transitively on the set of complete exceptional sequences. Ideas from [CB] and [R8], which have been developed in order to establish this result, will be used for a proof of Proposition 4.

We recall the following: A sequence $(M_1, \ldots, M_s)$ of exceptional modules is called *exceptional*, provided $\mathrm{Hom}(M_j, M_i) = 0 = \mathrm{Ext}^1(M_j, M_i)$ for any pair $i < j$. An exceptional sequence $(M_1, \ldots, M_s)$ is said to be *complete*, if $s = n$. (Note that a pair $(\mathbf{b}, \mathbf{a})$ of positive roots is $\vec{\Delta}$-orthogonal if and only if $(M(\mathbf{b}), M(\mathbf{a}))$ is an orthogonal exceptional sequence; this explains the order of the roots.) In our case of a finite-dimensional hereditary k-algebra, the indecomposable projective modules, the simple modules, as well as the indecomposable injective modules, always taken in a suitable order, yield examples of complete exceptional sequences.

Given a module N, we denote by $\mathscr{C}(N)$ the smallest subcategory containing N and being closed under extensions, kernels of epimorphisms and cokernels of monomorphisms. In case $(M_1, \ldots, M_s)$ is an exceptional sequence, then Crawley-Boevey [CB] has shown that $\mathscr{C}(M_1 \oplus \cdots \oplus M_s)$ is equivalent to the module category of a finite-dimensional hereditary algebra with precisely s isomorphism classes of simple modules.

Lemma 5. *Let M be a non-simple exceptional Λ-module. Then there exists a module N such that (M, N) or (N, M) is an exceptional sequence, and M considered as an object of $\mathscr{C}(M \oplus N)$ is not simple.*

Proof. Take a complete exceptional sequence $(N_1, \ldots, N_n)$ with $M = N_j$ for some j, and such that the length of $\bigoplus_{i=1}^{n} N_i$ is minimal.

Let M be an exceptional Λ-module with the following property: If (M, N) or (N, M) is an exceptional sequence, then M is simple in $\mathscr{C}(M \oplus N)$. Under this assumption, the reduction process as exhibited in [R8] shows that the sequence is orthogonal. Namely, in case we use transpositions, the modules N_i are not changed, only their indices are. We cannot use a proper reduction which does not involve M, since this would contradict our minimality assumption. Thus, assume that we make a proper reduction involving M. Up to duality, we have $\operatorname{Hom}(N_j, N_{j+1}) \neq 0$. By assumption we know that $M = N_j$ is a simple object in $\mathscr{C}(M \oplus N_{j+1})$, thus there is an exact sequence

$$0 \to tM \to N_{j+1} \to N' \to 0$$

with $\operatorname{Hom}(M, N') = 0$. The proper reduction replaces the pair (M, N_{j+1}) by the pair (N', M), thus we obtain a new exceptional sequence

$$(N_1, \ldots, N_{j-1}, N', M, N_{j+2}, \ldots, N_n)$$

of smaller length. This contradiction shows that our given sequence was orthogonal.

Note that given an orthogonal, complete exceptional sequence $(N_1, \ldots, N_n)$, all the modules N_i are simple [R8]. Thus, we see that M is simple.

Proof of Proposition 4. Assume that M is non-simple, and exceptional. According to Lemma 5, there exists an exceptional sequence (M, N) or (N, M) such that M is not simple in $\mathscr{C}(M \oplus N)$. Let M_1, M_2 be the two simple objects in this subcategory $\mathscr{C}(M \oplus N)$, with $\operatorname{Ext}^1(M_1, M_2) \neq 0$. Since M is not simple, in $\mathscr{C}(M \oplus N)$, there exists an exact sequence

$$0 \to a_2 M_2 \to M \to a_1 M_1 \to 0,$$

with $a_1, a_2 \geqq 1$. This completes the proof.

Consider now the special case when Λ is the tensor algebra of a k-species $\mathcal{S}$ of type $\vec{\Delta}$. Assume that there is given an exact sequence

$$0 \to a_2 M_2 \to M \to a_1 M_1 \to 0,$$

with indecomposable modules M, M_1, M_2, and a_1, $a_2 \geqq 1$. Let $M = M(\mathbf{a})$, $M_1 = M(\mathbf{a}_1)$, $M_2 = M(\mathbf{a}_2)$ with positive roots $\mathbf{a}, \mathbf{a}_1, \mathbf{a}_2$. We see that $\mathbf{a} = a_1 \mathbf{a}_1 + a_2 \mathbf{a}_2$. Note that the linear combinations of $\mathbf{a}_1, \mathbf{a}_2$ which are roots form a root system of rank 2, thus it is of type $\mathbb{A}_2$, $\mathbb{B}_2$, or $\mathbb{G}_2$; in particular, we see that $a_1 a_2 \leqq 3$. We are looking for conditions in order to have $a_1 = 1 = a_2$.

Recall that a root $\mathbf{a}$ is called *sincere*, provided $\mathbf{a} = \sum_{i=1}^{n} c_i \mathbf{e}_i$ with $c_i \neq 0$ for all i.

Proposition 5. *Let $\mathcal{S}$ be a k-species of type $\vec{\Delta}$, with Δ not of the form $\mathbb{A}_1$ or $\mathbb{G}_2$. Let $\mathbf{a}$ be a sincere positive root. In case Δ is of the form $\mathbb{C}_n$ for some $n \geqq 2$, assume in addition that $\mathbf{a}$ is short. Then there is an exact sequence*

$$0 \to M(\mathbf{a}_2) \to M(\mathbf{a}) \to M(\mathbf{a}_1) \to 0$$

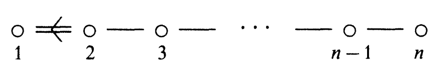
with $\vec{\Delta}$-orthogonal positive roots $\mathbf{a}_1, \mathbf{a}_2$.

Proof. According to Proposition 4, there is an exact sequence

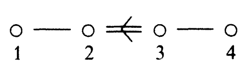
$$0 \to a_2 M(\mathbf{a}_2) \to M(\mathbf{a}) \to a_1 M(\mathbf{a}_1) \to 0,$$

with $\vec{\Delta}$-orthogonal positive roots $\mathbf{a}_1, \mathbf{a}_2$ and a_1, $a_2 \geqq 1$. In case the roots $\mathbf{a}_1$, $\mathbf{a}_2$ have the same length, the root system generated by $\mathbf{a}_1, \mathbf{a}_2$ is of type $\mathbb{A}_2$, thus $a_1 = 1 = a_2$. We assume that $\mathbf{a}_1$, $\mathbf{a}_2$ have different length. The existence of a sincere root implies that Δ is connected, thus Δ is of the form $\mathbb{B}_n$, $\mathbb{C}_n$ with $n \geqq 2$, or $\mathbb{F}_4$. If $\mathbf{a}$ is a short root, then $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, thus again $a_1 = 1 = a_2$. Thus, we can assume that $\mathbf{a}$ is a long root, and therefore only the cases $\mathbb{B}_n$, with $n \geqq 3$, and $\mathbb{F}_4$ remain.

In drawing the graph of Δ, we use the usual conventions. In the case $\mathbb{B}_n$, we label the vertices of the graph of Δ as follows:



with $a_{12} = -2$. Similarly, in the case $\mathbb{F}_4$, we label the vertices



with $a_{23} = -2$. We use an arbitrary ordering $\prec$ on the set of vertices $\{1, 2, \ldots, n\}$. Let $P(i)$ be the projective cover of S_i, for $1 \leqq i \leqq n$, and $Q(i)$ its injective envelope. Let $\mathbf{p}(i)$, $\mathbf{q}(i)$ be positive roots such that $M(\mathbf{p}(i)) = P(i)$, $M(\mathbf{q}(i)) = Q(i)$.

We deal with the Auslander-Reiten quiver Γ. We can write $\mathbf{a} = \tau^{-s}\mathbf{p}(i)$ for some $1 \leqq i \leqq n$ and some $s \geqq 0$. Since $\mathbf{a}$ is a long root, we have $i \geqq 2$ for $\mathbb{B}_n$, and $i \geqq 3$ for $\mathbb{F}_4$.

Consider first the case when $i < n$. Then there exists an arrow from $\mathbf{a}_2$ to $\mathbf{a}$ in Γ such that $\mathbf{a}_2$ lies in the τ-orbit of $\mathbf{p}(i+1)$; a corresponding non-zero map $M(\mathbf{a}_2) \to M(\mathbf{a})$ is injective and its cokernel is indecomposable, say of the form $M(\mathbf{a}_1)$, and $\mathbf{a}_1$ lies in the τ-orbit of $\mathbf{p}(n)$; here, we use that $\mathbf{a}$ is sincere, so that the "wing" with center $\mathbf{a}$ exists. In this way, we have obtained the two positive roots $\mathbf{a}_1$, $\mathbf{a}_2$ we were looking for.

This shows that we can assume that $i = n$. Case $\mathbb{B}_n$: Since $\mathbf{a}$ is sincere, there is a path $\mathbf{p}(1) \to \cdots \to \mathbf{a}$, and a path $\mathbf{a} \to \cdots \to \mathbf{q}(1)$, and an easy length consideration shows that both paths are sectional. In particular, the root $\mathbf{a}$ is uniquely determined. Up to duality, we can assume that $\mathbf{q}(1)$ is a simple root. It follows that we deal with the natural ordering $1 \prec 2 \prec \cdots \prec n$, since otherwise we would obtain some $i > 1$ with $\operatorname{Hom}(M, Q(i)) = 0$. As a consequence, $M(\mathbf{a}) = Q(n)$, and therefore $\mathbf{a} = 2\mathbf{e}_1 + \sum_{i=2}^{n} \mathbf{e}_i$. The case $n = 2$ has been excluded, thus $n \geqq 3$, and there is an exact sequence of the form

$$0 \to M\Big(\sum_{i=3}^{n} \mathbf{e}_i\Big) \to M \to M(2\mathbf{e}_1 + \mathbf{e}_2) \to 0\,.$$

Case $\mathbb{F}_4$. Consider the sectional paths $\mathbf{p}(1) \to \cdots \to \mathbf{b}_1$ and $\mathbf{b}_2 \to \cdots \to \mathbf{q}(1)$, where both $\mathbf{b}_1, \mathbf{b}_2$ belong to the τ-orbit of $\mathbf{p}(4)$. We have $\mathbf{b}_1 = \tau^2 \mathbf{b}_2$. Note that $\operatorname{Hom}(P(1), \tau M(\mathbf{b}_2)) = 0$, thus the two modules $M(\mathbf{b}_1)$, $M(\mathbf{b}_2)$ are the only modules $N = M(\mathbf{c})$ with $\mathbf{c}$ in the τ-orbit of $\mathbf{p}(4)$ which satisfy $\operatorname{Hom}(P(1), N) \neq 0$. This shows that $M(\mathbf{a})$ is one of these two modules. Up to duality, we can assume that $\mathbf{b}_1$ or $\tau^{-1}\mathbf{b}_1$ is equal to $\mathbf{q}(4)$. In both cases, there exists an exact sequence

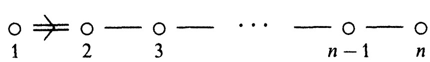
$$0 \to M(\tau^2 \mathbf{b}_1) \to M(\mathbf{b}_1) \to M(\mathbf{b}_2) \to 0\,,$$

thus, for $\mathbf{a} = \mathbf{b}_1$, we have found an exact sequence as required. It remains to consider the case $\mathbf{a} = \mathbf{b}_2$. Note that $\mathbf{b}_2 \neq \tau\mathbf{q}(4)$, since $\operatorname{Hom}(M(\tau\mathbf{q}(4)), M(\mathbf{q}(4))) = 0$, whereas we assume that $\mathbf{a}$ is sincere. It follows that $\mathbf{b}_2 = \mathbf{q}(4)$. But $\mathbf{q}(4)$ is sincere only in case we deal with the natural ordering $1 \prec 2 \prec 3 \prec 4$, and then we have the exact sequence

$$0 \to S_4 \to M(\mathbf{q}(4)) \to M(\mathbf{q}(3)) \to 0\,,$$

as required. This completes the proof.

Remark. In Proposition 5, we had to exclude the case $\mathbb{A}_1$, since otherwise $\mathbf{a}$ would be a simple root, thus $M(\mathbf{a})$ a simple representation. Also, we have excluded the case $\mathbb{G}_2$, since it has been discussed in detail before: the only $\bar{\Delta}$-orthogonal positive roots are the simple roots, and there are three sincere roots which are not of the form $\mathbf{e}_1 + \mathbf{e}_2$. Finally, consider the case $\mathbb{C}_n$:



with $a_{21} = -2$. The only long positive roots are given by $\mathbf{e}_1 + 2\sum_{i=2}^{t} \mathbf{e}_i$, where $1 \neq t \leqq n$. We see that there is just one sincere long positive root, and it cannot be written in the form $\mathbf{a} + \mathbf{b}$, where $\mathbf{a}, \mathbf{b}$ are long positive roots. It follows that it cannot be written in the form $\mathbf{a} + \mathbf{b}$, where $\mathbf{a}, \mathbf{b}$ are $\bar{\Delta}$-orthogonal positive roots.

9. Skew commutators

Let **a**, **b** be roots. The linear combinations of **a**, **b** which are roots form a root system of type $\mathbb{A}_1 \times \mathbb{A}_1$, $\mathbb{A}_2$, $\mathbb{B}_2$, or $\mathbb{G}_2$.

Assume that the pair $(\mathbf{a}_1, \mathbf{a}_2)$ is $\vec{\Delta}$-orthogonal. Let $r = -\langle \mathbf{a}_1, \mathbf{a}_2 \rangle$. Then $M(\mathbf{a}_1)$, $M(\mathbf{a}_2)$ are orthogonal exceptional modules. The subcategory

$$\mathscr{C} = \mathscr{C}(M(\mathbf{a}_1) \oplus M(\mathbf{a}_2))$$

is equivalent to the category of representations of a K-species $\mathscr{S}'$ of type $\vec{\Delta}'$, where Δ' is a Cartan matrix of rank 2. Here, we take $K = k$ in the simply-laced cases, and also in case at least one of the roots $\mathbf{a}_1$, $\mathbf{a}_2$ is a short root; in case both roots $\mathbf{a}_1$, $\mathbf{a}_2$ are long, let $K = \operatorname{End} M(\mathbf{a}_1)$. The last case can happen only for $r = 2$. Note that as a K-species, $\mathscr{S}'$ again is reduced. Let us fix an equivalence from the category of representations of $\mathscr{S}'$ onto $\mathscr{C}$, and denote it by ν. Since the objects $M(\mathbf{a}_1)$, $M(\mathbf{a}_2)$ are the simple objects in $\mathscr{C}$, they are the images of the simple representations S_1' and S_2' of $\mathscr{S}'$ under ν. We obtain an embedding of the Grothendieck group $K_0(\mathscr{S}') = \mathbb{Z}^2$ into $K_0(\mathscr{S})$, which again we denote by ν (with $\nu(\mathbf{e}_i) = \mathbf{a}_i$, for $i = 1, 2$). Under this embedding, the bilinear form $\langle -, - \rangle$ of $K_0(\mathscr{S})$ restricts to a scalar multiple of the corresponding bilinear form of $K_0(\mathscr{S}')$, since $\mathscr{C}$ is an extension closed full subcategory of the category of representations of $\mathscr{S}$. In fact, in case $K = k$, we obtain the corresponding bilinear form itself, in case $K = \operatorname{End} M(\mathbf{a}_1)$, we obtain the r-multiple: given two representations M, N of $\mathscr{S}'$, we have

$$\begin{aligned}\langle \nu M, \nu N \rangle &= \dim_k \operatorname{Hom}(\nu M, \nu N\rangle - \dim_k \operatorname{Ext}^1(\nu M, \nu N)\\ &= [K:k](\dim_K \operatorname{Hom}(M, N) - \dim_K \operatorname{Ext}^1(M, N))\\ &= [K:k]\langle M, N \rangle,\end{aligned}$$

and, in case $\mathbf{a}_1, \mathbf{a}_2$ both are long roots, then $[K:k] = 2 = r$.

Since Δ' is of rank 2, it is of the form $\mathbb{A}_1 \times \mathbb{A}_1$, $\mathbb{A}_2$, $\mathbb{B}_2$, or $\mathbb{G}_2$. In case $r = 0$, the only indecomposable objects in $\mathscr{C}$ are the two modules $M(\mathbf{a}_1)$, $M(\mathbf{a}_2)$, thus let us assume that $r > 0$. For the connected rank 2 cases, we have seen above how to express the indecomposable as skew commutators, and we claim that we obtain corresponding formulae when we replace $\mathbf{e}_i$ by $\mathbf{a}_i$; in case of two long roots $\mathbf{a}_1$, $\mathbf{a}_2$, we also have to replace v by v^2. It is sufficient to consider $\mathscr{H}' = \mathscr{H}_*(\vec{\Delta}') \otimes \mathbb{Q}(v)$, and the ring homomorphism $\mathscr{H}' \to \mathscr{H}$ which sends E_i to $\langle \mathbf{a}_i \rangle$, and v to $v^{[K:k]}$. In this way, we see that the recipe outlined in the introduction is based on our calculations in the rank 2 cases.

Of course, we may use the information provided by Proposition 5 in order to improve the inductive construction of the elements $X(\mathbf{a})$. For any $\vec{\Delta}$-orthogonal pair $(\mathbf{a}, \mathbf{b})$, with $r = r_{\mathbf{a}}^{\mathbf{b}} \geqq 1$, we have defined

$$X(\mathbf{a} + \mathbf{b}) = X(\mathbf{b})\, X(\mathbf{a}) - v^{-r} X(\mathbf{a})\, X(\mathbf{b}),$$

and we wonder which additional elements we have to take care off. We may assume that Δ is connected and not of type $\mathbb{G}_2$. In the simply-laced cases, we will have obtained a PBW-basis in this way, thus, we only have to consider the cases $\mathbb{B}_n$, $\mathbb{C}_n$ and $\mathbb{F}_4$. We work with the graph and the labelling as presented in the last section.

For $\mathbb{B}_n$, the only positive root to be considered in addition is $2\mathbf{e}_1 + \mathbf{e}_2$; for $\mathbb{C}_n$; we have to deal with all the additional roots $\mathbf{e}_1 + \sum_{i=2}^{t} \mathbf{e}_i$ for $2 \leqq i \leqq t$; and for $\mathbb{F}_4$, we have to take into account the roots $2\mathbf{e}_2 + \mathbf{e}_3$ and $2\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3$. For these roots $\mathbf{a}$, we have to construct $X(\mathbf{a})$ as outlined in the introduction, as the $\frac{1}{[2]}$-multiple of a proper commutator.

10. The $\mathscr{A}$-form $U_{\mathscr{A}}^+$

Recall that we denote $\mathscr{A} = \mathbb{Z}(v, v^{-1}]$ and that $U_{\mathscr{A}}^+$ is the $\mathscr{A}$-subalgebra of U^+ generated by the elements $E_i^{(t)}$.

Proposition 6. *The $\mathscr{A}$-algebra $\mathscr{H}_*(\vec{\Delta})$ is generated by the elements $\mathbf{e}_i^{(t)}$, with $1 \leqq i \leqq n$ and $t \geqq 1$.*

This is known, see [R6], [R7]. For the convenience of the reader, we outline below a direct argument.

Corollary. *The isomorphism $\eta\colon U^+ \to \mathscr{H}$ of $\mathbb{Q}(v)$-algebras defined by $\eta(E_i) = \mathbf{e}_i$ maps $U_{\mathscr{A}}^+$ onto $\mathscr{H}_*(\vec{\Delta})$.*

In order to present a proof of Proposition 6, we need the following lemma:

Lemma 6. *For any $\mathbf{d} = (d_1, \ldots, d_n) \in \mathbb{Z}^n$,*

$$\mathbf{e}_1^{(*d_1)} * \cdots * \mathbf{e}_n^{(*d_n)} = \sum_{\mathbf{dim}\,\alpha = \mathbf{d}} v^{-\zeta(\alpha,\alpha)} \langle \alpha \rangle .$$

Proof. We have $\mathbf{e}_i^{(*d_i)} = \langle d_i \mathbf{e}_i \rangle = v^{-d_i\varepsilon_i + d_i^2\varepsilon_i} d_i\mathbf{e}_i$, according to Lemma 2. Let us denote $\beta = \bigoplus_{i=1}^{n} d_i \mathbf{e}_i$. Thus $\dim \mathbf{d} = \dim \beta = \sum d_i \varepsilon_i$, and $\varepsilon(\beta) = \sum d_i^2 \varepsilon_i$. Also,

$$\zeta(\beta, \beta) = \sum_{i<j} \zeta(d_i\mathbf{e}_i, d_i\mathbf{e}_j),$$

since for $i \geqq j$, we have $\zeta(d_i\mathbf{e}_i, d_j\mathbf{e}_j) = 0$.

Any module M with dimension vector $\mathbf{d}$ has a unique filtration

$$M = M_0 \supseteqq M_1 \supseteqq \cdots \supseteqq M_n = 0$$

with factors M_{i-1}/M_i isomorphic to $d_i S_i$, since $\operatorname{Ext}^1(S_i, S_j) = 0$ for $i \geqq j$. This shows that the Hall polynomial $\phi^{\alpha}_{d_1\mathbf{e}_1, \ldots, d_n\mathbf{e}_n}$ is equal to 1, for any α with $\mathbf{dim}\,\alpha = \mathbf{d}$.

Also, for $i<j$, we have $\langle d_i\mathbf{e}_i, d_j\mathbf{e}_j\rangle = -\zeta(d_i\mathbf{e}_i, d_j\mathbf{e}_j)$, therefore

$$\begin{aligned}\mathbf{e}_1^{(*d_1)} * \cdots * \mathbf{e}_n^{(*d_n)} &= v^{-\sum d_i\varepsilon_i + \sum d_i^2\varepsilon_i} d_1\mathbf{e}_1 * \cdots * d_n\mathbf{e}_n \\ &= v^{-\sum d_i\varepsilon_i + \sum d_i^2\varepsilon_i} v^{-\sum\limits_{i<j}\langle d_i\mathbf{e}_i, d_j\mathbf{e}_j\rangle} \sum_{\mathbf{dim}\,\alpha=\mathbf{d}} [\alpha] \\ &= v^{-\dim\beta+\varepsilon(\beta)-\zeta(\beta,\beta)} \sum_{\mathbf{dim}\,\alpha=\mathbf{d}} [\alpha] \\ &= v^{-\dim\mathbf{d}+\langle\mathbf{d},\mathbf{d}\rangle} \sum_{\mathbf{dim}\,\alpha=\mathbf{d}} [\alpha]\,.\end{aligned}$$

We insert $[\alpha] = v^{\dim\alpha-\varepsilon(\alpha)}\langle\alpha\rangle$ and note that $\langle\mathbf{d},\mathbf{d}\rangle = \langle\mathbf{dim}\,\alpha, \mathbf{dim}\,\alpha\rangle = \varepsilon(\alpha)-\zeta(\alpha,\alpha)$. In this way, we obtain the formula as stated.

Proof of Proposition 6. According to Lemma 2, $\mathbf{e}_i^{(t)} = \langle\mathbf{e}_i\rangle^{(t)} = \langle t\mathbf{e}_i\rangle$, thus the elements $\mathbf{e}_i^{(t)}$ belong to $\mathscr{H}_*(\vec{\Delta})$.

Let $\mathscr{G}$ be the subring of $\mathscr{H}_*(\vec{\Delta})$ generated by the $\mathbf{e}_i^{(t)}$, with $1\leqq i\leqq n$ and $t\geqq 1$. Let us show that any $\alpha\in\mathscr{B}$ belongs to $\mathscr{G}$. We use induction on $\dim\alpha$. If there are at least two different positive roots $\mathbf{a}_i$ with $\alpha(\mathbf{a}_i)\neq 0$, then we use Proposition 1′ in order to see that $\langle\alpha\rangle = \langle\alpha(\mathbf{a}_1)\mathbf{a}_1\rangle * \cdots * \langle\alpha(\mathbf{a}_m)\mathbf{a}_m\rangle$. By assumption, we know that $\dim\alpha(\mathbf{a}_i)\mathbf{a}_i < \dim\alpha$, for all i; thus by induction, all the elements $\langle\alpha(\mathbf{a}_i)\mathbf{a}_i\rangle$ belong to $\mathscr{G}$. This shows that $\langle\alpha\rangle$ belongs to $\mathscr{G}$. It remains to be seen that for any positive root $\mathbf{a}$, and any $t\in\mathbb{N}_0$, the element $\langle t\mathbf{a}\rangle$ belongs to $\mathscr{G}$.

We apply Lemma 6 for $\mathbf{d} = t\mathbf{a}$. Note that $\zeta(t\mathbf{a}, t\mathbf{a}) = 0$, thus

$$\langle t\mathbf{a}\rangle = \mathbf{e}_1^{(*d_1)} * \cdots * \mathbf{e}_n^{(*d_n)} - \sum_{\substack{\mathbf{dim}\,\beta=\mathbf{d}\\ \beta\neq t\mathbf{a}}} v^{-\zeta(\beta,\beta)}\langle\beta\rangle\,.$$

If $\beta\neq t\mathbf{a}$ is given with $\mathbf{dim}\,\beta = \mathbf{d}$, then β cannot be a multiple of a root, thus there are at least two different roots $\mathbf{a}_i$ with $\beta(\mathbf{a}_i)\neq 0$, and therefore we know already that $\langle\beta\rangle$ belongs to $\mathscr{G}$. Of course, also $\mathbf{e}_1^{(*d_1)} * \cdots * \mathbf{e}_n^{(*d_n)}$ belongs to $\mathscr{G}$, therefore $\langle t\mathbf{a}\rangle$ belongs to $\mathscr{G}$. This completes the proof.

We obtain the following consequence, where $X_1, \ldots, X_m$ generates a PBW-basis of U^+ as constructed above.

Theorem 4. *The elements* $X_1^{(*\alpha(1))} * \cdots * X_m^{(*\alpha(m))}$ *with* $\alpha(1), \ldots, \alpha(m,)\in\mathbb{N}_0$ *form an* $\mathscr{A}$*-basis of* $U_{\mathscr{A}}^+$.

11. The subalgebra $\mathscr{H}\,|\,\mathscr{M}$

Let $\mathscr{S}$ be a reduced k-species of type $\vec{\Delta}$. We denote by rep-$\mathscr{S}$ the category of all representations of $\mathscr{S}$. Recall that a subcategory $\mathscr{M}$ of rep-$\mathscr{S}$ is said to be *closed under direct summands* provided for every module M in $\mathscr{M}$, all its direct summands belong to $\mathscr{M}$. A subcategory $\mathscr{M}$ of rep-$\mathscr{S}$ will be said to be *closed under potential extensions* provided for M_1, M_2 in $\mathscr{M}$, and $\phi^M_{M_1M_2}\neq 0$, also M belongs to $\mathscr{M}$. (Let us stress that it may happen that

$\phi^M_{M_1M_2} \neq 0$, whereas there is no exact sequence of the form $0 \to M_2 \to M \to M_1 \to 0$; the evaluation of the Hall-polynomial $\phi^M_{M_1M_2} \neq 0$ at $|k|$ may be zero: as an example, for $\mathbb{D}_4$, there are indecomposable modules M, M_1, M_2 with $\phi^M_{M_1M_2} = q - 2$.

The following is obvious from the definition of the Hall multiplication: *Let $\mathcal{M}$ be a subcategory of* rep-$\mathcal{S}$ *which is closed under potential extension. Let $\mathcal{H}|\mathcal{M}$ be the $\mathbb{Q}(v)$-subspace of $\mathcal{H}$ generated by the elements $[M]$ with $M \in \mathcal{M}$. Then $\mathcal{H}|\mathcal{M}$ is a subring. Similarly, let $\mathcal{H}_*(\vec{\Delta})|\mathcal{M}$ be the $\mathcal{A}$-submodule of $\mathcal{H}_*(\vec{\Delta})$ generated by the elements $[M]$ with $M \in \mathcal{M}$. Then $\mathcal{H}_*(\vec{\Delta})|\mathcal{M}$ is a subring.*

Proposition 7. *Let $\mathcal{M}$ be a subcategory of* rep-$\mathcal{S}$ *which is closed under direct summands and potential extensions. Let $\mathcal{N} \subseteq \mathcal{M}$ be a subcategory, and assume that any representation in $\mathcal{M}$ has a filtration with factors in $\mathcal{N}$. Then $\mathcal{H}|\mathcal{M}$ is generated as a $\mathbb{Q}(v)$-subalgebra by the elements $[N]$ with N in $\mathcal{N}$.*

The proof is similar to that of Proposition 4: Let $\mathcal{G}$ be the $\mathbb{Q}(v)$-subalgebra generated by the elements $[N]$ with N in $\mathcal{N}$. In order to show that any $[M]$, with M in $\mathcal{M}$ belongs to $\mathcal{G}$, we use induction. Let $M = M(\alpha)$ for some $\alpha \in \mathcal{B}$. Note that M is the direct sum of the modules $M(\alpha(\mathbf{a})\mathbf{a})$, with $\mathbf{a} \in \Phi^+$. If $\alpha(\mathbf{a}_i) \neq 0$ for at least two different positive roots $\mathbf{a}_i$, then all the modules $M(\alpha(\mathbf{a})\mathbf{a})$ have smaller length, and by induction $[M(\alpha(\mathbf{a})\mathbf{a})]$ belongs to $\mathcal{G}$. Proposition 1′ shows that also $[M]$ itself belongs to $\mathcal{G}$. It remains to consider the case $\alpha = t\mathbf{a}$ for some positive root $\mathbf{a}$ and some $t \geqq 1$. By assumption, $M = M(t\mathbf{a})$ has a filtration with factors in $\mathcal{N}$, thus, there are modules $N_1, \ldots, N_s$ with $\phi^M_{N_1,\ldots,N_s} \neq 0$. Write

$$[N_1] * \cdots * [N_s] = \sum_{\mathbf{dim}\,\beta = t\mathbf{a}} c_\beta \cdot [M(\beta)],$$

with coefficients $c_\beta \in \mathbb{Q}(v)$. Note that $[M(t\mathbf{a})]$ occurs with a non-zero factor $c_{t\mathbf{a}}$. Thus

$$[M] = c_{t\mathbf{a}}^{-1}\Big([N_1] * \cdots * [N_s] - \sum_{\substack{\mathbf{dim}\,\beta = t\mathbf{a} \\ \beta \neq t\mathbf{a}}} c_\beta \cdot [M(\beta)]\Big).$$

If $\beta \neq t\mathbf{a}$ is given with $\mathbf{dim}\,\beta = \mathbf{d}$, then there are at least two different roots $\mathbf{a}_i$ with $\beta(\mathbf{a}_i) \neq 0$, and therefore we know already that $[M(\beta)]$ belongs to $\mathcal{G}$. Altogether, this shows that $[M]$ belongs to $\mathcal{G}$.

For example, if i is a vertex for $\vec{\Delta}$, let rep-$\mathcal{S}\langle i\rangle$ be the subcategory of all representations which do not have S_i as a direct summand. By construction, this subcategory is closed under direct summands. If i is a sink or a source for $\vec{\Delta}$, then rep-$\mathcal{S}\langle i\rangle$ is also closed under potential extensions (note that for i a sink, S_i is projective; similarly, for i a source, S_i is injective). In these two cases, we will consider

$$\mathcal{H}\langle i\rangle = \mathcal{H}|\text{rep-}\mathcal{S}\langle i\rangle,$$

$$\mathcal{H}_*(\vec{\Delta})\langle i\rangle = \mathcal{H}_*(\vec{\Delta})|\text{rep-}\mathcal{S}\langle i\rangle.$$

The set of isomorphism classes of representations in rep-$\mathcal{S}\langle i\rangle$ will be denoted by $\mathcal{B}\langle i\rangle$, thus $\mathcal{H}\langle i\rangle$ is the free $\mathbb{Q}(v)$-module with basis $\mathcal{B}\langle i\rangle$.

Lemma 7. *Let i be a sink or a source. Let $\mathcal{N}$ be the set of indecomposable representations N with dimension vector $\mathbf{e}_j + t\mathbf{e}_i$ for some $j \neq i$, and some $t \geqq 0$. Then any representation in* rep-$\mathcal{S}\langle i\rangle$ *has a filtration with factors in $\mathcal{N}$.*

Proof. We consider the case where i is a sink. The other case follows by duality. Since i is a sink, S_i is projective. Thus, a representation M belongs to rep-$\mathcal{S}\langle i\rangle$ if and only if $\mathrm{Hom}(M, S_i) = 0$. In particular, we see that in this case the subcategory rep-$\mathcal{S}\langle i\rangle$ is closed under factor modules.

Let M be a representation in rep-$\mathcal{S}\langle i\rangle$. Let M' be the maximal submodule of M without composition factor of the form S_i. The composition factors of M' belong to $\mathcal{N}$, thus it remains to exhibit a filtration of M/M' with factors in $\mathcal{N}$. By definition of M', the socle of M/M' is a direct sum of copies of S_i. It follows that the socle of M/M' is contained in the radical of M/M', since $\mathrm{Hom}(M, S_i) = 0$. We can assume that $M/M' \neq 0$. Let M'' be a submodule of M containing M' such that $N = M''/M'$ is local and of Loewy length 2. Then clearly N belongs to $\mathcal{N}$. Also, M/M'' again belongs to rep-$\mathcal{S}\langle i\rangle$, thus by induction M/M'' has a filtration with factors in $\mathcal{N}$. This completes the proof.

Of course, if N is an indecomposable representations with dimension vector $\mathbf{e}_j + t\mathbf{e}_i$ for some $j \neq i$, and some $t \geqq 0$, then $t \leqq -a_{ij}$, and, conversely such indecomposable representations do exist.

Consider again the case where i is a sink. The bimodule ${}_jM_i$ has, as a right F_i-vector space, dimension $-a_{ij}$ (since its k-dimension is $-\varepsilon_i a_{ij}$). Let M' be an F_i-subspace of ${}_jM_i$ of codimension t. We construct a representation of $\mathcal{S}$ by attaching F_j to j, and ${}_jM_i/M'$ to i, and we use the projection map $F_j \otimes {}_jM_i \to {}_jM_i/M'$for the arrow $j \to i$.

The case when i is a source follows by duality; of course, we also can write down an explicit recipe: attach again F_j to j, an F_i-subspace M'' of $\mathrm{Hom}_{F_j}({}_iM_j, F_j)$ to i, and use as map $M'' \otimes {}_iM_j \to F_j$ the evaluation map.

Corollary. *Let i be a sink or a source. The $\mathbb{Q}(v)$-algebra $\mathcal{H}\langle i\rangle$ is generated by the elements $\mathbf{e}_j + t\mathbf{e}_i$, where $j \neq i$, and $0 \leqq t \leqq -a_{ij}$.*

12. Reflection functors

Let i be a vertex of $\vec{\Delta}$. Let $\sigma_i\vec{\Delta}$ be obtained from $\vec{\Delta}$ by changing the orientation of all arrows which have i as starting point or end point.

Let $\mathcal{S}$ be a reduced k-species of type $\vec{\Delta}$. Let $\sigma_i\mathcal{S}$ be the k-species obtained from $\mathcal{S}$ by replacing ${}_rM_s$ by its k-dual, if $r = i$ or $s = i$; note that $\sigma_i\mathcal{S}$ is a reduced k-species of type $\sigma_i\vec{\Delta}$.

Let us assume that i is a sink for $\vec{\Delta}$. We denote by σ_i^+, the Bernstein-Gelfand-Ponomarev reflection functor, see [BGP], [DR2]: it is an equivalence

$$\sigma_i^+ : \text{rep-}\mathcal{S}\langle i\rangle \to \text{rep-}\sigma_i\mathcal{S}\langle i\rangle.$$

Theorem 5. *Let i be a sink. The functor σ_i^+ yields an $\mathscr{A}$-algebra isomorphism*

$$\sigma_i \colon \mathscr{H}_*(\vec{\Delta})\langle i\rangle \to \mathscr{H}_*(\sigma_i\vec{\Delta})\langle i\rangle$$

where

$$\sigma_i\langle M\rangle = \langle \sigma_i^+ M\rangle .$$

Proof. The subcategories rep-$\mathscr{S}\langle i\rangle$, rep-$\sigma_i\mathscr{S}\langle i\rangle$, are closed under extensions. Let N_1, N_2 be in rep-$\mathscr{S}\langle i\rangle$. If M is a module with a submodule M_1 isomorphic to N_2, such that M/M_1 is isomorphic to N_1, then M belongs to rep-$\mathscr{S}\langle i\rangle$, and

$$\dim_k M = \dim_k N_1 + \dim_k N_2 ,$$

and

$$\dim_k \sigma_i^+ M = \dim_k \sigma_i^+ N_1 + \dim_k \sigma_i^+ N_2 .$$

Also, we may calculate the Hall-polynomials, as well as the bilinear form between modules in rep-$\mathscr{S}\langle i\rangle$ inside this subcategory, and therefore

$$\phi^M_{N_1N_2} = \phi^{\sigma_i^+M}_{\sigma_i^+N_1\sigma_i^+N_2} \quad \text{and} \quad \langle \mathbf{dim}\, N_1, \mathbf{dim}\, N_2\rangle = \langle \mathbf{dim}\,\sigma_i^+ N_1, \mathbf{dim}\,\sigma_i^+ N_2\rangle .$$

Recall that

$$\langle N_1\rangle * \langle N_2\rangle = \sum v^{c(M)}\phi^M_{N_1N_2}\langle M\rangle ,$$

with $c(M) = \varepsilon(N_1) + \varepsilon(N_2) + \langle \mathbf{dim}\, N_1, \mathbf{dim}\, N_2\rangle - \varepsilon(M)$. Similarly, we have

$$\langle \sigma_i^+ N_1\rangle * \langle \sigma_i^+ N_2\rangle = \sum v^{c(M)}\phi^M_{\sigma_i^+N_1\sigma_i^+N_2}\langle M\rangle ,$$

with the same function c. It follows that

$$\begin{aligned}
\sigma_i(\langle N_1\rangle * \langle N_2\rangle) &= \sigma_i\big(\sum v^{c(M)}\phi^M_{N_1N_2}\langle M\rangle\big) \\
&= \sum v^{c(M)}\phi^M_{N_1N_2}\langle \sigma_i^+ M\rangle \\
&= \sum v^{c(M)}\phi^{\sigma_i^+M}_{\sigma_i^+N_1\sigma_i^+N_2}\langle \sigma_i^+ M\rangle \\
&= \langle \sigma_i^+ N_1\rangle * \langle \sigma_i^+ N_2\rangle \\
&= \sigma_i\langle N_1\rangle * \sigma_i\langle N_2\rangle .
\end{aligned}$$

This shows that σ_i is a ring homomorphism.

Example. Let i be a sink. Let $j \neq i$, and $0 \leqq t \leqq -a_{ij}$. Then

$$\sigma_i\langle \mathbf{e}_j + t\mathbf{e}_i\rangle = \langle \mathbf{e}_j + (-a_{ij} - t)\,\mathbf{e}_i\rangle .$$

Proof. Recall that we denote by $\bar{\sigma}_i$ the reflection in $\mathbb{Z}^n$ at $\mathbf{e}_i$ with respect to the symmetric bilinear form $(-,-)$, thus $\bar{\sigma}_i(\mathbf{e}_j) = \mathbf{e}_j - \dfrac{2(\mathbf{e}_i, \mathbf{e}_j)}{(\mathbf{e}_i, \mathbf{e}_j)}\,\mathbf{e}_i$ and $(\mathbf{e}_i, \mathbf{e}_j) = \varepsilon_i a_{ij}$, whereas $(\mathbf{e}_i, \mathbf{e}_i) = 2\varepsilon_i$. This shows that $\bar{\sigma}_i(\mathbf{e}_j) = \mathbf{e}_j - a_{ij}\mathbf{e}_i$, and that $\bar{\sigma}_i(\mathbf{e}_i) = -\mathbf{e}_i$, thus

$\bar{\sigma}_i(\mathbf{e}_j+t\mathbf{e}_i)=\mathbf{e}_j+(-a_{ij}-t)\mathbf{e}_i$. On the other hand, for M an indecomposable representation of $\mathscr{S}$ different from S_i, we have $\mathbf{dim}\,\sigma_i^+(M)=\bar{\sigma}_i\,\mathbf{dim}\,M$, see [DR2].

Lusztig has proposed several braid group operations on U. In particular, consider for any i the following operator $T''_{i,1}$ defined for the canonical generators E_i, F_i, K_μ of U by

$$\begin{aligned}
T''_{i,1}(E_i) &= -F_iK_i^{\varepsilon_i},\\
T''_{i,1}(F_i) &= -K_i^{-\varepsilon_i}E_i,\\
T''_{i,1}(E_j) &= \sum_{r+s=-a_{ij}}(-1)^r v^{-\varepsilon_i r}E_i^{(s)}E_jE_i^{(r)} \quad \text{for } j\neq i,\\
T''_{i,1}(F_j) &= \sum_{r+s=-a_{ij}}(-1)^r v^{\varepsilon_i r}F_i^{(r)}F_jE_i^{(s)} \quad \text{for } j\neq i,\\
T''_{i,1}(K_\mu) &= K_{\bar{\sigma}_i(\mu)},
\end{aligned}$$

this defines an automorphism $T''_{i,1}$ of U.

Theorem 6. *Let i be a sink. The homomorphism σ_i is the restriction of $T''_{i,1}$ to $\mathscr{H}_*(\vec{\Delta})\langle i\rangle$.*

We show that $T''_{i,1}$ and σ_i have the same effect on a generating set of $\mathscr{H}\langle i\rangle$. We have seen above that the elements $\langle\mathbf{e}_j+t\mathbf{e}_i\rangle$ with $j\neq i$ and $0\leqq t\leqq a_{ij}$ form such a generating set, and that $\sigma_i\langle\mathbf{e}_j+t\mathbf{e}_i\rangle=\langle\mathbf{e}_j+(-a_{ij}-t)\mathbf{e}_i\rangle$.

On the other hand, Lusztig has shown in [L4], 37.2.5, that

$$T''_{i,1}(x_{i,j;1,t;-1})=x'_{i,j;1,-a_{ij}-t;-1}$$

where

$$\begin{aligned}
x_{i,j;1,t;-1} &= \sum_{r+s=t}(-1)^r v^{-\varepsilon_i r(-a_{ij}-t+1)}E_i^{(r)}E_jE_i^{(s)},\\
x'_{i,j;1,t;-1} &= \sum_{r+s=t}(-1)^r v^{-\varepsilon_i r(-a_{ij}-t+1)}E_i^{(s)}E_jE_i^{(r)}.
\end{aligned}$$

By Proposition 3, we know that $x_{i,j;1,t;-1}=\langle\mathbf{e}_j+t\mathbf{e}_i\rangle_{\vec{\Delta}}$, since i is a sink for $\vec{\Delta}$, and that $x'_{i,j;1,t;-1}=\langle\mathbf{e}_j+t\mathbf{e}_i\rangle_{\sigma_i\vec{\Delta}}$, since i a source for $\sigma_i\vec{\Delta}$. Here, we have added to $\langle\mathbf{e}_j+t\mathbf{e}_i\rangle$ the indices $\vec{\Delta}$ and $\sigma_i\vec{\Delta}$, respectively, in order to point out the relevant orientation. This completes the proof.

13. Construction of the PBW-basis, using a braid group operation

We recall from Lusztig [L4] that the operators $T''_{i,1}$, where i runs through the vertices of the graph of Δ, define a braid group operation on U. The considerations above allow to see that our generating sequences for PBW-bases can be obtained from the generators $E_1,\ldots,E_n$ using this braid group operation (in the simply-laced cases, a similar result was pointed out by Lusztig in [L3]).

Recall that a sequence $i_m,\ldots,i_1$ is called a *sink sequence* for $\vec{\Delta}$, provided i_m is a sink for $\vec{\Delta}$, and for any $1\leqq t<m$, the vertex i_t is a sink for the orientation $\sigma_{i_{t+1}}\cdots\sigma_{i_m}\vec{\Delta}$.

Consider a $\vec{\Delta}$-admissible ordering $\mathbf{a}_1, \ldots, \mathbf{a}_m$ of the positive roots. There exists a sequence $i_1, \ldots, i_m$ of vertices of $\vec{\Delta}$, with the following properties:

(1) The sequence $i_m, \ldots, i_1$ is a sink sequence for $\vec{\Delta}$.

(2) We have $\mathbf{a}_j = \bar{\sigma}_{i_1}\bar{\sigma}_{i_2}\cdots\bar{\sigma}_{i_{j-1}}(\mathbf{e}_{i_j})$.

(We may construct this sequence $i_1, \ldots, i_m$ as follows: Let C be the Coxeter transformation for $\vec{\Delta}$. For any j, there exist some power C^s such that $C^s(\mathbf{a}_j)$ is a positive root, but $C^{s+1}(\mathbf{a}_j)$ is not positive. Then there exists a unique vertex i_j such that $\langle C^s(\mathbf{a}_j), \mathbf{e}_{i_j}\rangle > 0$, this is the vertex we are interested in. In terms of representation theory: the representation $M(C^s(\mathbf{a}_j))$ is indecomposable projective, thus has a unique simple factor module, namely $M(\mathbf{e}_{i_j})$.)

We fix a $\vec{\Delta}$-admissible ordering $\mathbf{a}_1, \ldots, \mathbf{a}_m$ of the positive roots and we set $X_j = \langle \mathbf{a}_j \rangle$.

Theorem 7.

$$X_j = T''_{i_1,1} T''_{i_2,1} \cdots T''_{i_{j-1},1}(E_{i_j}).$$

Proof. Since we deal with a sink sequence, the operations $T''_{i,1}$ are given by σ_i, see Theorem 6, thus we can use the reflection functors σ_i^+, see Theorem 5. This shows that

$$T''_{i_1,1} T''_{i_2,1} \cdots T''_{i_{j-1},1}(E_{i_i}) = \langle \bar{\sigma}_{i_1}\bar{\sigma}_{i_2}\cdots\bar{\sigma}_{i_{j-1}}(S_{i_j})\rangle,$$

where S_{i_j} is the simple representation of $\bar{\sigma}_{i_{j-1}}\cdots\bar{\sigma}_{i_2}\bar{\sigma}_{i_1}\mathcal{S}$ corresponding to the vertex i_j. However,

$$\langle \bar{\sigma}_{i_1}\bar{\sigma}_{i_2}\cdots\bar{\sigma}_{i_{j-1}}(S_{i_j})\rangle = \langle \bar{\sigma}_{i_1}\bar{\sigma}_{i_2}\cdots\bar{\sigma}_{i_{j-1}}(\mathbf{e}_{i_j})\rangle = X_j.$$

This completes the proof.

Appendix 1. The rank 2 cases

For all the rank 2 cases, we are going to present the multiplication table for one of the generating sequences for a PBW-basis, explicitly.

Case $\mathbb{A}_2$. Let

$$S_1 = \langle \mathbf{e}_2 \rangle, \quad X_2 = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle, \quad X_3 = \langle \mathbf{e}_1 \rangle.$$

Then

$$X_2 * X_1 = vX_1 * X_2,$$

$$X_3 * X_2 = vX_2 * X_3,$$

$$X_3 * X_1 = v^{-1}X_1 * X_3 + X_2.$$

Actually, for arbitrary $\mathbb{A}_n$, an explicit presentation of $\mathcal{H}$ by generators and relations, using as generating set the generating sequence for a PBW-basis, will be given in Appendix 2.

Case $\mathbb{B}_2$, with $\varepsilon_1 = 2$. The elements

$$X_1 = \langle \mathbf{e}_2 \rangle, \quad X_2 = \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle, \quad X_3 = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle, \quad X_4 = \langle \mathbf{e}_1 \rangle$$

satisfy the following relations:

$$\begin{aligned}
X_2 * X_1 &= v^2 X_1 * X_2, \\
X_3 * X_2 &= v^2 X_2 * X_3, \\
X_4 * X_3 &= v^2 X_3 * X_4, \\
X_3 * X_1 &= X_1 * X_3 + [2] X_2, \\
X_4 * X_2 &= X_2 * X_4 + (v^2 - 1) X_3^{(*2)}, \\
X_4 * X_1 &= v^{-2} X_1 * X_4 + X_3 .
\end{aligned}$$

Proof. The vanishing of a skew commutator for X_i, X_{i+1} follows from the general considerations concerning Auslander-Reiten quivers. We have seen above how to write X_2 and X_3 as skew commutators. This yields the fourth and the sixth equality. It remains to show the fifth equality.

We have $X_4 = \langle \mathbf{e}_1 \rangle = \mathbf{e}_1$, and $X_2 = \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = v^{-4+2}(\mathbf{e}_1 + 2\mathbf{e}_2)$, thus

$$X_4 * X_2 = v^{-2} \mathbf{e}_1 * (\mathbf{e}_1 + 2\mathbf{e}_2).$$

Note that $\langle \mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = -2$. The Hall polynomial $\phi^{\mathbf{e}_1 \oplus (\mathbf{e}_1 + 2\mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2}$ is given by q^2. On the other hand, there are $q^2 + 1$ images of non-zero maps $M(\mathbf{e}_1 + 2\mathbf{e}_2) \to M(2(\mathbf{e}_1 + \mathbf{e}_2))$. The number of images of the form $M(\mathbf{e}_1 + \mathbf{e}_2)$ is $q + 1$, the remaining ones are of the form $M(\mathbf{e}_1 + 2\mathbf{e}_2)$. This shows that the Hall polynomial $\phi^{2(\mathbf{e}_1 + \mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2}$ is given by $q^2 - q = v^4 - v^2$. Thus, we see that

$$\begin{aligned}
X_4 * X_2 &= v^{-2} \mathbf{e}_1 * (\mathbf{e}_1 + 2\mathbf{e}_2) \\
&= v^{-4} v^4 \mathbf{e}_1 \oplus (\mathbf{e}_1 + 2\mathbf{e}_2) + v^{-4}(v^4 - v^2)\big(2(\mathbf{e}_1 + \mathbf{e}_2)\big) \\
&= X_2 * X_4 + v^{-4+2}(v^4 - v^2) X_3^{(*2)},
\end{aligned}$$

here, we use that $X_2 * X_4 = v^{-2}(\mathbf{e}_1 + 2\mathbf{e}_2) * \mathbf{e}_1 = \mathbf{e}_1 \oplus (\mathbf{e}_1 + 2\mathbf{e}_2)$ and that

$$X_3^{(*2)} = \langle 2(\mathbf{e}_1 + \mathbf{e}_2) \rangle = v^{-6+4}\big(2(\mathbf{e}_1 + \mathbf{e}_2)\big).$$

Case $\mathbb{G}_2$, with $\varepsilon_1 = 3$. We denote the elements as follows:

$$\begin{aligned}
&X_1 = \langle \mathbf{e}_2 \rangle, && X_2 = \langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle, && X_3 = \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle, \\
&X_4 = \langle 2\mathbf{e}_1 + 3\mathbf{e}_2 \rangle, && X_5 = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle, && X_6 = \langle \mathbf{e}_1 \rangle.
\end{aligned}$$

We have the following relations: Of course, as usual, we have

$$X_{i+1} * X_i = v^3 X_i * X_{i+1} \quad \text{for all} \quad 1 \leqq i \leqq 5.$$

In addition

$$\begin{aligned}
X_3 * X_1 &= \quad v X_1 * X_3 + [3] X_2, \\
X_4 * X_2 &= v^3 X_2 * X_4 + (v^6 - v^4 - v^2 + 1) X_3^{(*3)}, \\
X_5 * X_3 &= \quad v X_3 * X_5 + [3] X_4, \\
X_6 * X_4 &= v^3 X_4 * X_6 + (v^6 - v^4 - v^2 + 1) X_5^{(*3)},
\end{aligned}$$

$$\begin{aligned}
X_4 * X_1 &= \quad X_1 * X_4 + (v^3 - v^{-1}) X_3^{(*2)}, \\
X_5 * X_2 &= \quad X_2 * X_5 + (v^3 - v^{-1}) X_3^{(*2)}, \\
X_6 * X_3 &= \quad X_3 * X_6 + (v^3 - v^{-1}) X_5^{(*2)},
\end{aligned}$$

$$\begin{aligned}
X_5 * X_1 &= v^{-1} X_1 * X_5 + [2] X_3, \\
X_6 * X_2 &= v^{-3} X_2 * X_6 + (v^2 - 1) X_3 * X_5 + (v^3 - v - v^{-1}) X_4,
\end{aligned}$$

$$X_6 * X_1 = v^{-3} X_1 * X_6 + X_5 .$$

Consider $X_6 * X_3$. Note: $X_6 = \langle \mathbf{e}_1 \rangle = \mathbf{e}_1$. $X_3 = \langle \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = v^{-5+1}(\mathbf{e}_1 + 2\mathbf{e}_2)$, and $\langle \mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2 \rangle = -3$. The Hall polynomial $\phi^{\mathbf{e}_1 \oplus (\mathbf{e}_1 + 2\mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2}$ is given by q^3. On the other hand, there are $q^3 + 1$ images of non-zero maps $M(\mathbf{e}_1 + 2\mathbf{e}_2) \to M(2(\mathbf{e}_1 + \mathbf{e}_2))$. The number of images of the form $M(\mathbf{e}_1 + \mathbf{e}_2)$ is $q + 1$, the remaining ones are of the form $M(\mathbf{e}_1 + 2\mathbf{e}_2)$. This shows that the Hall polynomial $\phi^{2(\mathbf{e}_1 + \mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 2\mathbf{e}_2}$ is given by $q^3 - q = v^6 - v^2$. Thus, we see that

$$\begin{aligned}
X_6 * X_3 &= v^{-4} \mathbf{e}_1 * (\mathbf{e}_1 + 2\mathbf{e}_2) \\
&= v^{-7} v^6 \mathbf{e}_1 \oplus (\mathbf{e}_1 + 2\mathbf{e}_2) + v^{-7}(v^6 - v^2)\big(2(\mathbf{e}_1 + \mathbf{e}_2)\big) \\
&= X_3 * X_6 + (v^{-3} - v^{-1}) X_5^{(*2)},
\end{aligned}$$

here, we use that $X_3 * X_6 = v^{-4}(\mathbf{e}_1 + 2\mathbf{e}_2) * \mathbf{e}_1 = v^{-1} \mathbf{e}_1 \oplus (\mathbf{e}_1 + 2\mathbf{e}_2)$, and that $X_5^{(*2)} = \langle 2(\mathbf{e}_1 + \mathbf{e}_2) \rangle = v^{-8+4}\big(2(\mathbf{e}_1 + \mathbf{e}_2)\big)$. This proves the assertion concerning $X_6 * X_3$. The shift by τ yields a similar formula for $X_4 * X_1$, by duality, we obtain the corresponding result for $X_5 * X_2$.

Consider $X_6 * X_2$. We have $X_6 = \langle \mathbf{e}_1 \rangle = \mathbf{e}_1$, $X_2 = \langle \mathbf{e}_1 + 3\mathbf{e}_2 \rangle = v^{-6+3}(\mathbf{e}_1 + 3\mathbf{e}_2)$, and $\langle \mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2 \rangle = -6$. Therefore

$$\begin{aligned}
X_6 * X_2 &= v^{-3} \mathbf{e}_1 * (\mathbf{e}_1 + 3\mathbf{e}_2) \\
&= v^{-9}\big(c_1 \mathbf{e}_1 \oplus (\mathbf{e}_1 + 3\mathbf{e}_2) + c_2 (\mathbf{e}_1 + \mathbf{e}_2) \oplus (\mathbf{e}_1 + 2\mathbf{e}_2) + c_3 (2\mathbf{e}_1 + 3\mathbf{e}_2)\big),
\end{aligned}$$

where

$$\begin{aligned}
c_1 &= \phi^{\mathbf{e}_1 \oplus (\mathbf{e}_1 + 3\mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2}, \\
c_2 &= \phi^{(\mathbf{e}_1 + \mathbf{e}_2) \oplus (\mathbf{e}_1 + 2\mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2}, \\
c_3 &= \phi^{2\mathbf{e}_1 + 3\mathbf{e}_2}_{\mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2}
\end{aligned}$$

are the various Hall polynomials. Clearly, $c_1 = \phi^{\mathbf{e}_1 \oplus (\mathbf{e}_1 + 3\mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2} = q^3$. The last polynomial has been determined in [R4], it is $c_3 = \phi^{2\mathbf{e}_1 + 3\mathbf{e}_2}_{\mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2} = q^3 - q^2 - q$.

Thus, it remains to calculate c_2. Consider maps $f\colon M(\mathbf{e}_1 + 3\mathbf{e}_2) \to M(\mathbf{e}_1 + 2\mathbf{e}_2)$ and $f'\colon M(\mathbf{e}_1 + 3\mathbf{e}_2) \to M(\mathbf{e}_1 + \mathbf{e}_2)$. The corresponding map

$$\begin{bmatrix} f \\ f' \end{bmatrix} : M(\mathbf{e}_1 + 3\mathbf{e}_2) \to M(\mathbf{e}_1 + 2\mathbf{e}_2) \oplus M(\mathbf{e}_1 + \mathbf{e}_2)$$

is injective if and only if $f \neq 0$ and f' cannot be factored through f. Since End $(M(\mathbf{e}_1 + 3\mathbf{e}_2))$ operates transitively on Hom $(M(\mathbf{e}_1 + 3\mathbf{e}_2), M(\mathbf{e}_1 + 2\mathbf{e}_2))$, we can fix some projection $\pi\colon M(\mathbf{e}_1 + 3\mathbf{e}_2) \to M(\mathbf{e}_1 + 2\mathbf{e}_2)$ and we obtain all images of injective maps

$$M(\mathbf{e}_1 + 3\mathbf{a}_2) \to M(\mathbf{e}_1 + 2\mathbf{e}_2) \oplus M(\mathbf{e}_1 + \mathbf{e}_2)$$

by using only the maps $\begin{bmatrix} \pi \\ f' \end{bmatrix}$, where f' does not factor through π. Now assume there is given another map $f''\colon M(\mathbf{e}_1 + 3\mathbf{e}_2) \to M(\mathbf{e}_1 + \mathbf{e}_2)$ such that $\begin{bmatrix} \pi \\ f' \end{bmatrix}$ and $\begin{bmatrix} \pi \\ f'' \end{bmatrix}$ have the same image. The projectivity of $M(\mathbf{e}_1 + 3\mathbf{e}_2)$ shows that there exists an automorphism ϱ of $M(\mathbf{e}_1 + 3\mathbf{e}_2)$ such that $\pi\varrho = \pi$ and $f'\varrho = f''$. The first equality implies that $\varrho = 1$, thus $f' = f''$. Also, since π is surjective, the multiplication by π yields a bijection between

$$\mathrm{Hom}(M(\mathbf{e}_1 + 2\mathbf{e}_2), M(\mathbf{e}_1 + \mathbf{e}_2))$$

and the set of those elements of

$$\mathrm{Hom}(M(\mathbf{e}_1 + 3\mathbf{e}_2), M(\mathbf{e}_1 + \mathbf{e}_2))$$

which factor through π. This shows that

$$c_2 = \phi^{(\mathbf{e}_1 + \mathbf{e}_2) \oplus (\mathbf{e}_1 + 2\mathbf{e}_2)}_{\mathbf{e}_1, \mathbf{e}_1 + 3\mathbf{e}_2} = q^{\varepsilon(\mathbf{e}_1 + 3\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2)} - q^{\varepsilon(\mathbf{e}_1 + 2\mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2)} = q^3 - q^2.$$

Also, we note that

$$\begin{aligned} X_2 * X_6 &= \langle \mathbf{e}_1 \oplus (\mathbf{e}_1 + 3\mathbf{e}_2) \rangle = v^{-9+9} \mathbf{e}_1 \oplus (\mathbf{e}_1 + 3\mathbf{e}_2), \\ X_3 * X_5 &= \langle (\mathbf{e}_1 + \mathbf{e}_2) \oplus (\mathbf{e}_1 + 2\mathbf{e}_2) \rangle = v^{-9+4} (\mathbf{e}_1 + \mathbf{e}_2) \oplus (\mathbf{e}_1 + 2\mathbf{e}_2), \\ X_4 &= \langle 2\mathbf{e}_1 + 3\mathbf{e}_2 \rangle = v^{-9+3} (2\mathbf{e}_1 + 3\mathbf{e}_2). \end{aligned}$$

Therefore

$$\begin{aligned} X_6 * X_2 &= v^{-9} \big(v^6 \mathbf{e}_1 \oplus (\mathbf{e}_1 + 3\mathbf{e}_2) + (v^6 - v^4)(\mathbf{e}_1 + \mathbf{e}_2) \oplus (\mathbf{e}_1 + 2\mathbf{e}_2) \\ &\qquad + (v^6 - v^4 - v^2)(2\mathbf{e}_1 + 3\mathbf{e}_2) \big) \\ &= v^{-9} v^6 v^0 X_2 * X_6 + v^{-9} (v^6 - v^4) v^5 X_3 * X_5 + v^{-9} (v^6 - v^4 - v^2) v^6 X_4 . \end{aligned}$$

Consider $X_4 * X_2$. We have $X_4 = v^{-9+3}(2\mathbf{e}_1 + 3\mathbf{e}_2)$, $X_2 = v^{-6+3}(\mathbf{e}_1 + 3\mathbf{e}_2)$, and $\langle 2\mathbf{e}_1 + 3\mathbf{e}_2, \mathbf{e}_1 + 3\mathbf{e}_2\rangle = -3$. Thus

$$X_4 * X_2 = v^{-6-3-3}\big(c_1\big((\mathbf{e}_1 + 3\mathbf{e}_2) \oplus (2\mathbf{e}_1 + 3\mathbf{e}_2)\big) + c_2\big(3(\mathbf{e}_1 + 2\mathbf{e}_2)\big)\big)$$

with

$$c_1 = \phi^{(\mathbf{e}_1+3\mathbf{e}_2)\oplus(2\mathbf{e}_1+3\mathbf{e}_2)}_{2\mathbf{e}_1+3\mathbf{e}_2,\mathbf{e}_1+3\mathbf{e}_2},$$

$$c_2 = \phi^{3(\mathbf{e}_1+2\mathbf{e}_2)}_{2\mathbf{e}_1+3\mathbf{e}_2,\mathbf{e}_1+3\mathbf{e}_2}.$$

We calculate the Hall polynomials. We have $c_1 = \phi^{(\mathbf{e}_1+3\mathbf{e}_2)\oplus(2\mathbf{e}_1+3\mathbf{e}_2)}_{2\mathbf{e}_1+3\mathbf{e}_2,\mathbf{e}_1+3\mathbf{e}_2} = q^6 = v^{12}$.

On the other hand, given three maps f, f', f'': $M(\mathbf{e}_1 + 3\mathbf{e}_2) \to M(\mathbf{e}_1 + \mathbf{e}_2)$, the corresponding map

$$\begin{bmatrix} f \\ f' \\ f'' \end{bmatrix} : M(\mathbf{e}_1 + 3\mathbf{e}_2) \to 3M(\mathbf{e}_1 + \mathbf{e}_2)$$

is injective if and only if f, f', f'' are linearly independent over $k = \mathrm{End}\big(M(\mathbf{e}_1 + \mathbf{e}_2)\big)$. This shows that the number of injective maps $M(\mathbf{e}_1 + 3\mathbf{e}_2) \to 3M(\mathbf{e}_1 + \mathbf{e}_2)$ is given by the polynomial $(q^3 - 1)(q^3 - q)(q^3 - q^2)$. Of course, different triples will yield the same image if and only if they are obtained from each other by the multiplication using an automorphism of $M(\mathbf{e}_1 + 3\mathbf{e}_2)$, and the number of such automorphisms is given by the polynomial $q^3 - 1$. Note that the cokernel of any injective map $M(\mathbf{e}_1 + 3\mathbf{e}_2) \to 3M(\mathbf{e}_1 + \mathbf{e}_2)$ is of the form $M\big(3(\mathbf{e}_1 + 2\mathbf{e}_2)\big)$. Altogether we see that

$$c_2 = \phi^{3(\mathbf{e}_1+2\mathbf{e}_2)}_{2\mathbf{e}_1+3\mathbf{e}_2,\mathbf{e}_1+3\mathbf{e}_2} = (q^3 - q)(q^3 - q^2) = v^{12} - v^{10} - v^8 + v^6.$$

Thus:

$$\begin{aligned} X_4 * X_2 &= v^{-12}\big(v^{12}(\mathbf{e}_1 + 3\mathbf{e}_2) \oplus (2\mathbf{e}_1 + 3\mathbf{e}_2) + (v^{12} - v^{10} - v^8 + v^6)\big(3(\mathbf{e}_1 + 2\mathbf{e}_2)\big)\big) \\ &= v^3 X_2 * X_4 + (v^6 - v^4 - v^2 + 1)X_3^{(*3)}, \end{aligned}$$

since

$$X_2 * X_4 = \langle(\mathbf{e}_1 + 3\mathbf{e}_2) \oplus (2\mathbf{e}_1 + 3\mathbf{e}_2)\rangle = v^{-15+12}(\mathbf{e}_1 + 3\mathbf{e}_2) \oplus (2\mathbf{e}_1 + 3\mathbf{e}_2)$$

and $X_3^{(*3)} = \langle 3(\mathbf{e}_1 + 2\mathbf{e}_2)\rangle = v^{-15+9}\big(3(\mathbf{e}_1 + 2\mathbf{e}_2)\big)$. This completes the consideration of $X_4 * X_2$. The shift by τ^{-1} yields the corresponding result for $X_6 * X_2$.

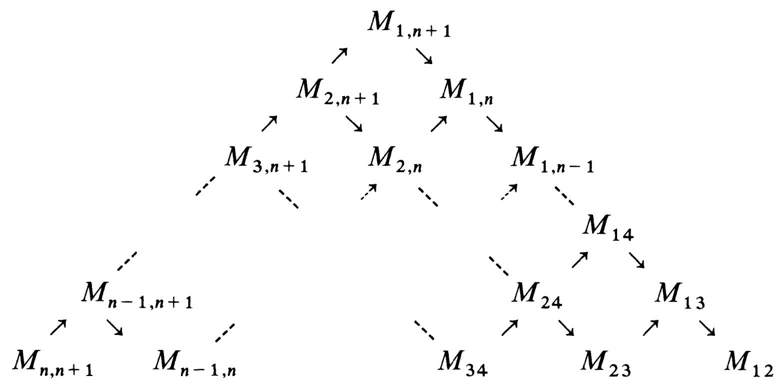
Appendix 2. The case $\mathbb{A}_n$

A PBW-basis for this case has been exhibited by Yamane [Y1], [Y2]. We will show that his basis can be derived easily from the considerations above. For the convenience of the reader, we choose an analogous indexing, and label the cases in the same way as he did.

We consider the following orientation:

$$\underset{1}{\circ} \longrightarrow \underset{2}{\circ} \longrightarrow \cdots \longrightarrow \underset{n-1}{\circ} \longrightarrow \underset{n}{\circ}.$$

Let $P(i)$ be the projective cover of S_i, for $1 \leqq i \leqq n$, and set $P(n+1) = 0$. Note that for $i < j$, there is an embedding $P(j) \subset P(i)$, and we denote by $M_{ij} = P(i)/P(j)$ the corresponding factor module. In this way, $S_i = M_{i,i+1}$, and M_{ij} is a serial module of length $j - i$. The Auslander-Reiten quiver has the following shape:



Some general features should be mentioned: If M_{rs}, M_{ij} are indecomposable modules, then $\dim \operatorname{Hom}(M_{rs}, M_{ij}) \leqq 1$, and $\dim \operatorname{Ext}^1(M_{rs}, M_{ij}) \leqq 1$. We write $\langle rs, ij \rangle = \langle \mathbf{dim}\, M_{rs}, \mathbf{dim}\, M_{ij} \rangle$, it follows that $-1 \leqq \langle rs, ij \rangle \leqq 1$, and for $(r,s) \neq (i,j)$, we have $-1 \leqq \langle rs, ij \rangle + \langle ij, rs \rangle \leqq 1$.

Given a pair of modules M_{ij}, M_{rs} with $i < r$ or with $i = r$, and $j < s$, there are six different cases to be considered.

Case (I): $i = r < j < s$. There exists an epimorphism $M_{rs} \to M_{ij}$. Thus

$$\langle rs, ij \rangle = 1, \quad \langle ij, rs \rangle = 0.$$

Case (II): $i < r < s < j$. There are no homomorphisms and no extensions between M_{rs} and M_{ij}, thus

$$\langle rs, ij \rangle = 0, \quad \langle ij, rs \rangle = 0.$$

Case (III): $i < r < j = s$. There exists an inclusion $M_{rs} \to M_{ij}$. Thus

$$\langle rs, ij \rangle = 1, \quad \langle ij, rs \rangle = 0.$$

Case (IV): $i < r < j < s$. There exists an exact sequence

$$0 \to M_{rs} \to M_{is} \oplus M_{rj} \to M_{ij} \to 0,$$

in particular, we have a non-zero map $M_{rs} \to M_{rj} \to M_{ij}$. Thus

$$\langle rs, ij \rangle = 1, \quad \langle ij, rs \rangle = -1.$$

Case (V): $i < j = r < s$. There exists an exact sequence

$$0 \to M_{rs} \to M_{is} \to M_{ij} \to 0,$$

and $\operatorname{Hom}(M_{rs}, M_{ij}) = 0$. Thus

$$\langle rs, ij\rangle = 0, \quad \langle ij, rs\rangle = -1.$$

Case (VI): $i < j < r < s$. There are no homomorphisms and no extensions between M_{rs} and M_{ij}, thus again

$$\langle rs, ij\rangle = 0, \quad \langle ij, rs\rangle = 0.$$

These considerations are sufficient for writing down the automorphism $\iota_{ij} = \iota_{\mathbf{dim}\, M_{ij}}$. The corresponding skew derivation $\delta_{ij} = \delta_{\langle M_{ij}\rangle}$ will be zero in the cases (I), (II), (III) and (VI), thus it remains to consider the cases (IV) and (V). Since in these cases $\dim \operatorname{Ext}^1(M_{ij}, M_{rs}) = 1$, we see that $\delta_{ij}(\langle M_{rs}\rangle)$ will be a multiple of an element of our basis.

Let $m(ij) = \dim M_{ij} = j - i$. Thus $\langle M_{ij}\rangle = v^{-m(ij)+1}[M_{ij}]$.

Case (IV). The exact sequence

$$0 \to M_{rs} \to M_{is} \oplus M_{rj} \to M_{ij} \to 0$$

shows that $\delta_{ij}(\langle M_{rs}\rangle)$ is a multiple of $\langle M_{is} \oplus M_{rj}\rangle$. We determine the corresponding coefficient. Obviously, the Hall polynomial is

$$\phi^{M_{is}\oplus M_{rj}}_{M_{ij}M_{rs}} = q - 1,$$

thus the coefficient of $[M_{is} \oplus M_{rj}]$ in $[M_{ij}] * [M_{rs}]$ is $v^{\langle ij, rs\rangle}(q-1) = v^{-1}(q-1)$. Let us note that there are no homomorphisms between M_{is} and M_{rj}, thus $\dim \operatorname{End}(M_{is} \oplus M_{rj}) = 2$; it follows that $\langle M_{is} \oplus M_{rj}\rangle = v^{-m(ij)-m(rs)+2}[M_{is} \oplus M_{rj}]$, and that $\langle M_{is} \oplus M_{rj}\rangle = \langle M_{is}\rangle * \langle M_{rj}\rangle$. On the other hand, we note that

$$\langle M_{ij}\rangle = v^{-m(ij)+1}[M_{ij}], \quad \langle M_{rs}\rangle = v^{-m(rs)+1}[M_{rs}].$$

Altogether, we see that

$$\delta_{ij}(\langle M_{rs}\rangle) = (v - v^{-1})\langle M_{is}\rangle * \langle M_{rj}\rangle.$$

Case (V). The exact sequence

$$0 \to M_{rs} \to M_{is} \to M_{ij} \to 0$$

shows that $\delta_{ij}(\langle M_{rs}\rangle)$ is a multiple of $\langle M_{is}\rangle$. Here, the Hall polynomial $\phi^{M_{is}}_{M_{ij}M_{rs}}$ is equal to 1, thus the coefficient of $[M_{is}]$ in $[M_{ij}] * [M_{rs}]$ is $v^{\langle ij, rs\rangle} = v^{-1}$. We have $\langle M_{ij}\rangle = v^{-m(ij)+1}$, $\langle M_{rs}\rangle = v^{-m(rs)+1}$ and $\langle M_{is}\rangle = v^{-m(ij)-m(rs)+1}$. Altogether, we see that

$$\delta_{ij}(\langle M_{rs}\rangle) = \langle M_{is}\rangle.$$

We use the notation $X_{ij} = \langle M_{ij}\rangle$. We have shown that

$$X_{ij} * X_{rs} = v \cdot X_{rs} * X_{ij}, \qquad \text{in case (I), (III)},$$

$$X_{ij} * X_{rs} = X_{rs} * X_{ij}, \qquad \text{(II), (VI)},$$

$$X_{ij} * X_{rs} = X_{rs} * X_{ij} + (v - v^{-1}) X_{is} * X_{rj}, \qquad \text{(IV)},$$

$$X_{ij} * X_{rs} = v^{-1} X_{rs} * X_{ij} + X_{is}, \qquad \text{(V)}.$$

Condition (V) asserts that we may define X_{ij} inductively by

$$X_{is} = X_{ij} * X_{js} - v^{-1} X_{js} * X_{ij}, \quad \text{for} \quad i < j < s,$$

starting with $X_{i,i+1} = \mathbf{e}_i$. These elements in U^+ have been presented already by Jimbo in [J2] (with v replaced by v^{-1}). In order to rewrite these generators and relations in the form presented by Yamane [Y1], [Y2], we have to adjoin a square root t of v (this element t is denoted by q in Yamane; but of course, in our presentation, we have $q = v^2 = t^4$; there is a good reason to stick to the notation $q = v^2$, since in our approach, this q is usually evaluated at prime powers, namely at the cardinality of some finite field).

Let $\mathbb{Q}(t)$ be the rational function field in one variable t, let $v = t^2$, and consider $\mathscr{H}_t = \mathscr{H} \otimes_{\mathbb{Q}(v)} \mathbb{Q}(t)$. We define

$$E_{ij} = t^{j-i-1} X_{ij} (= t^{-j+i+1} [M_{ij}]).$$

Then we obtain the following relations:

$$E_{ij} * E_{rs} = t^2 \cdot E_{rs} * E_{ij}, \qquad \text{in case (I), (III)},$$

$$E_{ij} * E_{rs} = E_{rs} * E_{ij}, \qquad \text{(II), (VI)},$$

$$E_{ij} * E_{rs} = E_{rs} * E_{ij} + (t^2 - t^{-2}) E_{is} * E_{rj}, \qquad \text{(IV)},$$

$$t^2 E_{ij} * E_{rs} = E_{rs} * E_{ij} + t \cdot E_{is}, \qquad \text{(V)}.$$

Proof. The extra factors t^{j-i-1}, t^{s-r-1} cancel in all but the last equality. The last equality can be written as follows: $E_{ij} * E_{rs} = v^{-1} \cdot E_{ji} * E_{rs} + t^{-1} E_{is}$, and this relation is directly derived from the relations before.

Our identification of U^+ and $\mathscr{H}$ yields an identification of $U_t^+ = U^+ \otimes_{\mathbb{Q}(v)} \mathbb{Q}(t)$ and $\mathscr{H}_t$. We have $E_{i,i+1} = E_i$, and the last relation above shows that for $j - i > 1$, we have $E_{ij} = t E_{i,j-1} * E_{j-1,j} - t^{-1} E_{j-1,j} E_{i,j-1}$. This shows that the elements E_{ij} of U_t^+ coincide with the elements e_{ij} as introduced by Yamane.

References

[BGP] *I.N. Bernstein, I.M. Gelfand, V.A. Ponomarev*, Coxeter functors and Gabriel's theorem, Uspechi Mat. Nauk **28** (1973); Russian Math. Surv. **28** (1973), 17–32.

[CB] *W.W. Crawley-Boevey*, Exceptional sequences of representations of quivers, Proc. ICRA VI, Carleton-Ottawa Math. LNS **14** (1992).

[DR1] *V. Dlab*, *C.M. Ringel*, On algebras of finite representation type, J. Algebra **33** (1975), 306–394.

[DR2] *V. Dlab*, *C.M. Ringel*, Indecomposable representations of graphs and algebras, Mem. Amer. Math. Soc. **173** (1976).

[D] *V.G. Drinfeld*, Hopf algebras and the Yang-Baxter equation, Soviet. Math. Dokl. **32** (1985), 254–258.

[Ga] *P. Gabriel*, Unzerlegbare Darstellungen I, Manuscr. Math. **6** (1972), 71–103.

[Go] *K.R. Goodearl*, Uniform ranks of prime factors of skew polynomial rings, Ring Theory, Proc. Ohio State – Denison Conference 1992, to appear.

[GL] *K.R. Goodearl*, and *E.S. Letzter*, Prime factor algebras of the coordinate ring of quantum matrices, Proc. Amer. Math. Soc., to appear.

[Gr] *J.A. Green*, Hall algebras, hereditary algebras and quantum groups, Invent. math., to appear.

[H] *J.E. Humphreys*, Introduction to Lie Algebras and Representation Theory, Springer Grad. Texts Math. **9** (1972).

[J1] *M. Jimbo*, A q-analogue of $U(\mathbf{g})$ and the Yang-Baxter equation, Letters Math. Phys. **10** (1985), 63–69.

[J2] *M. Jimbo*, A q-analogue of $U(gl(N+1))$, Hecke algebra, and the Yang-Baxter equation, Letters Math. Phys. **11** (1986), 247–252.

[KT1] *S.M. Khoroshkin*, *V.N. Tolstoy*, Universal R-matrix for quantized (super)algebras, Comm. Math. Phys. **141** (1991), 599–617.

[KT2] *S.M. Khoroshkin*, *V.N. Tolstoy*, The Cartan-Weyl basis and the universal R-matrix for .quantum Kac-Moody algebras and superalgebras, Proc. Second Wigner Symp. Goslar 1991, to appear.

[KT3] *S.M. Khoroshkin*, *V.N. Tolstoy*, Twisting of quantum (super)algebras, connection of Drinfeld's and Cartan-Weyl realizations for quantum affine algebras, Max-Planck-Institut Bonn, Preprint MPI/94-23.

[LS] *S. Levendorskii*, *Y. Soibelman*, Algebras of functions on compact quantum groups, Schubert cells and quantum tori, Comm. Math. Phys. **139** (1991), 141–170.

[L1] *G. Lusztig*, Finite dimensional Hopf algebras arising from quantum groups, J. Amer. Math. Soc. **3** (1990), 257–296.

[L2] *G. Lusztig*, Quantum groups at roots of 1, Geom. Dedicata (1990).

[L3] *G. Lusztig*, Canonical bases arising from quantized enveloping algebras, J. Amer. Math. Soc. **3** (1990), 447–498.

[L4] *G. Lusztig*, Introduction to quantum groups, Progr. Math. **110**, Birkhäuser, 1993.

[R1] *C.M. Ringel*, Representations of K-species and bimodules, J. Algebra **41** (1976), 269–302.

[R2] *C.M. Ringel*, Tame algebras and integral quadratic forms, Springer Lect. Notes Math. **1099** (1984).

[R3] *C.M. Ringel*, Hall algebras, In: Topics in Algebra, Banach Center Publ. **26** (1990), 433–447.

[R4] *C.M. Ringel*, Hall polynomials for the representation-finite hereditary algebras, Adv. Math. **84** (1990), 137–178.

[R5] *C.M. Ringel*, Hall algebras and quantum groups, Invent. math. **101** (1990), 583–592.

[R6] *C.M. Ringel*, From representations of quivers via Hall and Loewy algebras to quantum groups, In: Proc. Int. Conf. Algebra, Novosibirsk 1989, Contemp. Math. **131.2** (1992), 381–401.

[R7] *C.M. Ringel*, Hall algebras revisted, Israel Math. Conf. Proc. **7** (1993), 171–176.

[R8] *C.M. Ringel*, The braid group action on the set of exceptional sequences, Proc. Oberwolfach Conf. Abelian groups and modules 1993, Contemp. Math., to appear.

[Ro] *M. Rosso*, Finite dimensional representations of the quantum analog of the enveloping algebra of a complex simple Lie algebra, Comm. Math. Phys. **117** (1988), 581–593.

[S] *A. Schofield*, Semi-invariants of quivers, J. London Math. Soc. **43** (1991), 383–395.

[T] *M. Takeuchi*, The q-bracket product and quantum enveloping algebras of classical type, J. Math. Soc. Japan **42** (1990), 605–628.

[X] *N. Xi*, Root vectors in quantum groups, Preprint.

[Y1] *H. Yamane*, A Poincaré-Birkhoff-Witt theorem for the quantum group of type A_N, Proc. Japan Acad. **64** (A) (1988), 385–386.

[Y2] *H. Yamane*, A Poincaré-Birkhoff-Witt theorem for quantized universal enveloping algebras of type A_N, Publ. RIMS, Kyoto Univ. **25** (1989), 503–520.

Fakultät für Mathematik, Universität Bielefeld, D-33501 Bielefeld
e-mail: ringel@mathematik.uni-bielefeld.de

Eingegangen 1. Oktober 1993, in revidierter Fassung 24. Februar 1995

J. reine angew. Math. **470** (1996), 89—107

Journal für die reine und angewandte Mathematik

Restricted mean value property for positive functions

By *W. Hansen* and *N. Nadirashvili* at Bielefeld

0. Introduction

In the following U will denote a non-empty domain in the Euclidean space $\mathbb{R}^d$, $d \geqq 1$, and for the most part we shall restrict our attention to the case where $d \geqq 3$ or $\complement U$ is non-polar. Let λ be Lebesgue measure on $\mathbb{R}^d$ and for every $x \in \mathbb{R}^d$ and $r > 0$ let

$$B(x, r) = \{y \in \mathbb{R}^d : \|y - x\| < r\}, \quad \lambda_{B(x,r)} = \big(\lambda(B(x, r))\big)^{-1} 1_{B(x,r)} \lambda,$$

i.e., $\lambda_{B(x,r)}$ is the normed Lebesgue measure on the open ball $B(x, r)$.

It is well known that a locally bounded Lebesgue measurable function f on U is harmonic if and only if $\lambda_{B(x,r)}(f) = f(x)$ for every $x \in U$ and $r > 0$ such that $\overline{B(x, r)} \subset U$. The problem to what extent harmonicity of f is already a consequence of knowing that for every $x \in U$ there exists *one* radius $r > 0$ such that $B(x, r) \subset U$ and $\lambda_{B(x,r)}(f) = f(x)$ has a long history starting early this century. For a recent survey see [NV].

In the following let r always denote a strictly positive function on U such that $B(x, r(x)) \subset U$ for every $x \in U$. If $U = \mathbb{R}^d$ we assume that $r \leqq \|\cdot\| + M$ for some $M \in \mathbb{R}_+$. A real function f on U is called *r-median* if

$$\lambda_{B(x, r(x))}(f) = f(x)$$

for every $x \in U$ (where we implicitly assume that f is Lebesgue-integrable on each ball $B(x, r(x))$).

W. A. Veech [Ve 2] proved that for a bounded Lipschitz domain U an r-median function f on U is harmonic provided that $|f|$ is bounded by a harmonic function on U and r is bounded away from zero on compact subsets of U. In [Ve 3] he showed that an r-median function $f \geqq 0$ on a bounded Lipschitz domain U is harmonic if $\alpha\tilde{\varrho} \leqq r \leqq (1 - \alpha)\tilde{\varrho}$ for some $0 < \alpha < 1$ and a strictly positive function $\tilde{\varrho}$ on U such that $\tilde{\varrho} \leqq \operatorname{dist}(\cdot, \complement U)$ and $|\tilde{\varrho}(x) - \tilde{\varrho}(y)| \leqq \|x - y\|$ for all $x, y \in U$. Later A. Ancona pointed out that the restriction on the boundary of U in the proof of Veech can be avoided. A different proof for general domains U (and for slightly more general means) has been given by J. R. Baxter [Ba 2]. Whereas Veech and Baxter used (quite different) probabilistic tools, A. Cornea and J. Veselý recently gave a purely analytic proof for more or less the same result. In [Ve 3] and [CV] an essential step is the proof of a Harnack inequality for positive r-median functions. It implies that Choquet theory for compact convex sets is applicable and hence an investigation of extremal positive r-median functions is sufficient. A careful study of the discrete Martin kernel corresponding to the transition kernel $P:(x,A)\mapsto \lambda_{B(x,r(x))}(A)$ then allows these authors to show that every extremal positive r-median function is harmonic.

In [Ve 3] Veech conjectured that every positive r-median function on a bounded domain U is already harmonic if only r is locally bounded away from zero. By [Hu] this is true if the dimension d is equal to one. In [HN 6] we showed, however, that for $d \geqq 2$ and any $0<\alpha<1$ there always exist $\mathscr{C}^\infty$-functions r and f on U such that $0<r\leqq \alpha \operatorname{dist}(\cdot, \complement U)$ and f is a positive and r-median function, but not harmonic.

In this paper we intend to prove the following:

Theorem. *Let $0<\alpha\leqq 1$ such that $r(y)\geqq \alpha\big(r(x)-\|x-y\|\big)$ for every $x\in U$ and all $y\in B\big(x,r(x)\big)$. Then every positive r-median function on U is harmonic.*

By [HN 7] this is true if $d\leqq 2$ and $\complement U$ is polar. So we shall assume in the following that $d\geqq 3$ or that $\complement U$ is non-polar, i.e., that U is a Green domain.

In order to compare our assumption on r with the hypothesis of Veech, Baxter, Cornea, and Veselý let us define

$$\varrho(x) := \sup\{r(z)-\|x-z\| : z\in U\}\,.$$

If $\tilde\varrho$ is any real function on U such that $r\leqq\tilde\varrho$ and $|\tilde\varrho(x)-\tilde\varrho(y)|\leqq\|x-y\|$ then $\varrho\leqq\tilde\varrho$ (since $r(z)-\|x-z\|\leqq\tilde\varrho(z)-\|x-z\|\leqq\tilde\varrho(x)$). In particular,

$$\varrho\leqq \operatorname{dist}(\cdot,\complement U)\quad \text{if } U\neq\mathbb{R}^d,\quad \varrho\leqq\|\cdot\|+M\quad \text{if } U=\mathbb{R}^d.$$

In fact, ϱ is the smallest Lipschitz function (with Lipschitz constant 1) majorizing r. (Indeed, taking $z=x$ we see that $r\leqq\varrho$ and, given $x,y\in U$, we have

$$r(z)-\|x-z\|\leqq r(z)-\|y-z\|+\|x-y\|$$

for every $z\in U$, hence $\varrho(x)\leqq\varrho(y)+\|x-y\|$.) Moreover, obviously

$$r(y)\geqq\alpha\big(r(x)-\|x-y\|\big)$$

for $x\in U$ and $y\in B\big(x,r(x)\big)$ if and only if $r\geqq\alpha\varrho$.

So our geometric condition on r is equivalent to $\alpha\tilde\varrho\leqq r\leqq\tilde\varrho$ for some real function $0<\tilde\varrho\leqq\operatorname{dist}(\cdot,\complement U)$ on U satisfying $|\tilde\varrho(x)-\tilde\varrho(y)|\leqq\|x-y\|$ for all $x,y\in U$. I.e., we replace the inequality $r\leqq(1-\alpha)\tilde\varrho$ of Veech and others by the considerably weaker condition $r\leqq\tilde\varrho$.

As in [HN1], [HN2] the essential element in our proof will be the relation between supermedian functions for random walks given by the measures $P(x,\cdot) = \lambda_{B(x,r(x))}$ and positive solutions of Schrödinger equations $\Delta u - \delta 1_A \varrho^{-2} u = 0$ for small $\delta > 0$. It will allow us to show that (transfinite) sweeping by P on cubes Q in U is uniformly comparable to the hitting of Q by Brownian motion provided the diameter of Q is controlled from below by $\inf \varrho(Q)$. Expressed analytically, the discrete equilibrium potential e_Q^r and the (classical) equilibrium potential e_Q for these cubes then satisfy $e_Q^r \leqq e_Q \leqq c_0 e_Q^r$.

It will be convenient to use a Whitney type decomposition $\mathcal{Q}_\alpha$ of U into cubes Q such that the diameter of each Q is controlled by the values of $\alpha\varrho$ on Q. We shall see that the assumption $\alpha\varrho \leqq r$ implies that the discrete Green function g generated by the transition kernel P satisfies a uniform Harnack inequality on the cubes $Q \in \mathcal{Q}_\alpha$. The relation between the equilibrium potentials e_Q^r and e_Q then yields that for each $Q \in \mathcal{Q}_\alpha$ and every measure ν on $\bar{Q}$ the discrete potential g^ν is comparable to the equilibrium potential e_Q. This comparison enables us to prove in a very direct way that every positive r-median function is bounded by a harmonic function and hence harmonic by [HN1], [HN2].

Finally, let us note that in spite of some probabilistic terminology used above all our proofs will be purely analytic and that our method can be applied to more general means.

1. Sweeping on cubes by volume means

Let us fix a cube Q such that $\bar{Q} \subset U$ (for us any Borel set Q in $\mathbb{R}^d$ such that $\bar{Q}$ is a compact cube and Q contains the interior of $\bar{Q}$ will be a cube). Let e_Q denote the equilibrium potential for Q with respect to U, i.e., e_Q is the smallest continuous real potential on U such that $e_Q = 1$ on $\bar{Q}$. In probabilistic terms, $e_Q(x)$ is the probability that Brownian motion starting at $x \in U$ will hit Q before leaving U.

Let us suppose that r is a Borel measurable function and define a Markov kernel P on U by

$$P(x.\cdot) = \lambda_{B(x,r(x))} \quad \text{for } x \in U.$$

We shall say that a Borel measurable numerical function $s \geqq 0$ on U is *P-supermedian* if $Ps \leqq s$. Of course the equilibrium potential e_Q is P-supermedian. We now intend to define a discrete equilibrium potential e_Q^r associated with P. It will be a P-supermedian function such that $e_Q^r \leqq 1$ on U and $e_Q^r = 1$ on Q, but not necessarily the smallest one having these properties if r is not locally bounded away from zero.

Let us first choose a compactification Y of U as in [HN1]: We take a sequence (q_n) of continuous potentials on U having compact superharmonic support such that for every $x \in U$ and every neighborhood W of x there exist $n, m \in \mathbb{N}$ with

$$0 \leqq q_n - q_m \leqq 1_W, \quad (q_n - q_m)(x) > 0\,.$$

There exists a metrizable compactification Y of U such that the functions q_n, $n \in \mathbb{N}$, can be extended to continuous functions on Y separating the points of Y. If U is a bounded regular set or $U = \mathbb{R}^d$ ($d \geqq 3$) then U is the one-point compactification of Y.

We define kernels $S_{r,Q}$ and $T_{r,Q}$ on Y by

$$S_{r,Q}(x,\cdot) = \lambda_{B(x,r(x))}, \quad T_{r,Q} = 1_{\complement Q}\lambda_{B(x,r(x))} \quad \text{for } x \in U \setminus Q$$

and

$$S_{r,Q}(x,\cdot) = T_{r,Q}(x,\cdot) = \varepsilon_x \quad \text{for } x \in Q \cup (Y \setminus U).$$

For every $x \in U$ let $\mathscr{M}_{S_{r,Q}}(x)$ denote the smallest w^*-closed $S_{r,Q}$-stable set of positive measures on Y containing ε_x. By Propositions 3.2, 3.3, 3.4 in [HN1] there exists a unique P-invariant measure σ_x in $\mathscr{M}_{S_{r,Q}}(x)$ and we have $\sigma_x(U \setminus Q) = 0$ and $\sigma_x(v) \leqq v(x)$ for every l.s.c. function $v > -\infty$ on Y satisfying $S_{r,Q}v \leqq v$.

If r is locally bounded away from zero (and this will be the case in our applications later on) then simply

$$\mathscr{M}_{S_{r,Q}}(x) = \{S^m_{r,Q}(x,\cdot) : m = 0, 1, 2, \ldots, \infty\}, \quad \sigma_x = S^\infty_{r,Q}(x,\cdot)$$

where $S^\infty_{r,Q} = \lim_{m \to \infty} S^m_{r,Q}$ and the sequence $(S^m_{r,Q}(x, A))_{m \in \mathbb{N}}$ is increasing to $\sigma_x(A)$ for every Borel subset A of Q (cf. [HN5]). Similarly we have $\mathscr{M}_{T_{r,Q}}(x)$ and a $T_{r,Q}$-invariant measure $\tau_x \in \mathscr{M}_{T_{r,Q}}(x)$. By Lemma 3.6 in [HN1] we know that $1_{Y \setminus Q}\sigma_x = 1_{Y \setminus Q}\tau_x$. Again this is almost evident if r is locally bounded away from zero.

Clearly, σ_x is a probability measure since P is a Markov kernel. Thus

$$\sigma_x(Q) = 1 - \sigma_x(Y \setminus U) = 1 - \tau_x(Y \setminus U).$$

We finally define the discrete equilibrium potential e^r_Q on U by

$$e^r_Q(x) := \sigma_x(Q) \quad \text{for } x \in U,$$

i.e., $e^r_Q(x)$ is the probability that the random walk (transfinite unless r is locally bounded away from zero) given by P starting at x will hit Q.

Obviously, $0 \leqq e^r_Q \leqq 1$. Moreover, $e^r_Q = 1$ on Q since $\mathscr{M}_{S_{r,Q}}(x) = \{\varepsilon_x\}$ for $x \in Q$. Using the transfinite construction of the measures σ_y for $y \in Y$ (cf. [HN2], p. 163) it is easy to see that $(y, A) \mapsto \sigma_y(A)$ is a kernel and that for every (finite) measure μ on Y the measure $\int \sigma_y \, \mu(dy)$ is the $S_{r,Q}$-invariant measure contained in $\mathscr{M}_{S_{r,Q}}(\mu)$. This shows in particular that e^r_Q is Borel measurable. Moreover, it implies that

$$\sigma_x = \int \sigma_y \, \lambda_{B(x,r(x))}(dy) \quad \text{for } x \in U \setminus Q$$

since $\mathscr{M}_{S_{r,Q}}(x) = \{\varepsilon_x\} \cup \mathscr{M}_{S_{r,Q}}(\lambda_{B(x,r(x))})$ for $x \in U \setminus Q$. Thus

$$e^r_Q(x) = \lambda_{B(x,r(x))}(e^r_Q) = Pe^r_Q(x) \text{ for every } x \in U \setminus Q.$$

In particular, e^r_Q is P-supermedian since obviously $Pe^r_Q \leqq 1 = e^r_Q$ on Q.

1.1. Remark. It is easily seen that there exists a smallest P-supermedian function $\tilde{e}_Q^r$ such that $\tilde{e}_Q^r = 1$ on Q (take $f_0 = 1_Q, f_{n+1} = \sup(f_n, Pf_n)$, $\tilde{e}_Q^r = \lim_{n\to\infty} f_n$; cf. [BH], p. 42). Moreover, $P\tilde{e}_Q^r = \tilde{e}_Q^r$ on $U\setminus Q$ since the function s defined by $s = 1$ on Q and $s = P\tilde{e}_Q^r$ on $U\setminus Q$ is P-supermedian, and it is not difficult to show that

$$\tilde{e}_Q^r(x) = S_{r,Q}^\infty(x, Q).$$

This implies that $\tilde{e}_Q^r = e_Q^r$ if r is locally bounded away from zero whereas in general only $\tilde{e}_Q^r \leqq e_Q^r$ and we may have strict inequality. For example, if $r < \operatorname{dist}(\cdot, Q)$ on $U\setminus\bar{Q}$ then $\tilde{e}_Q^r = 0$ on $U\setminus\bar{Q}$ since $1_{\bar{Q}}$ is P-supermedian, but $e_Q^r > 0$ on U by the next theorem.

We want to compare e_Q^r and e_Q for $\delta(Q) \geqq \gamma \inf \varrho(Q)$. Since $\varrho \leqq \inf \varrho(Q) + \delta(Q)$ on Q, the lower bound for $\delta(Q)$ implies that $\varrho \leqq \left(1 + \frac{1}{\gamma}\right)\delta(Q)$ on Q. Given $\gamma > 0$ we first define

$$\beta = \left(8\left(1 + \frac{1}{\gamma}\right)\sqrt{d}\right)^{-d}.$$

Then by Proposition 2.4 in [HN1] there exists a (very small) constant $\delta > 0$ such that for every $x \in \mathbb{R}^d$, for every $r > 0$ and for every Borel subset A of $B := B(x, r)$ satisfying $A \cap B(x, r/2) = \emptyset$ or $\lambda_{B(x,r)}(A) \geqq \beta$ the following holds: If V is a Borel measurable function on B such that

$$0 \leqq V \leqq \delta 1_A \operatorname{dist}(\cdot, \complement B)^{-2}$$

and if u is a continuous solution of $\Delta u - Vu = 0$ on B such that $0 \leqq u \leqq 1$ then

$$u(x) \geqq \int_{\complement A} u \, d\lambda_{B(x, r(x))}$$

(note that $u(x) \leqq \lambda_{B(x, r(x))}(u)$ since u is subharmonic).

For every cube Q let x_Q be the center of Q and

$$Q^\gamma = x_Q + (1 + \gamma)(\mathring{Q} - x_Q).$$

So Q^γ is an open cube with $x_{Q^\gamma} = x_Q$ and $\overline{Q^\gamma} \subset U$. Let

$$Q_1 = [0, 1]^d, \quad V_1 = \delta\left(\left(1 + \frac{1}{\gamma}\right)\sqrt{d}\right)^{-2} 1_{Q_1}, \quad g_1 = H_{Q_1^\gamma}^{\Delta - V_1} 1,$$

i.e., g_1 is the continuous function on $\overline{Q_\gamma^1}$ satisfying $g_1 = 1$ on the boundary of Q_1^γ and $\Delta g_1 - V_1 g_1 = 0$ on Q_1^γ. Obviously, g_1 is subharmonic on Q_1^γ and not harmonic on Q_1. Hence $g_1 < 1$ on Q_1^γ. Define

$$a := \sup g_1(Q_1), \quad c_0(\gamma) := \frac{1}{1 - a}.$$

Then we have the following result which has some independent interest.

1.2. Theorem. *Let γ be a strictly positive real number. Then*

$$e_Q^r \leqq e_Q \leqq c_0(\gamma)\, e_Q^r$$

for every cube Q in U satisfying $\delta(Q) \geqq \gamma \inf \varrho(Q)$.

Proof. Fix a cube Q such that $\delta(Q) \geqq \inf \varrho(Q)$. Extend e_Q by 0 to a l.s.c. function on Y. Then $S_{r,Q}\, e_Q \leqq e_Q$, hence $\sigma_x(e_Q) \leqq e_Q(x)$ and

$$e_Q^r(x) = \sigma_x(Q) = \sigma_x(e_Q) \leqq e_Q(x)$$

for every $x \in U$ since $\sigma_x(U \backslash Q) = 0$ and $e_Q = 1$ on Q.

Let $x \in U$ such that $Q \cap B(x, r(x)/2) \neq \emptyset$. We claim that

$$\lambda(Q \cap B(x, r(x))) \geqq \beta \lambda(B(x, r(x)))\,.$$

Indeed, if $y \in Q \cap B(x, r(x)/2)$ then

$$r(x) \leqq \varrho(y) + \|x - y\| < \left(1 + \frac{1}{\gamma}\right) \delta(Q) + \frac{r(x)}{2}\,,$$

hence

$$\delta(Q) > \frac{1}{2\left(1 + \frac{1}{\gamma}\right)}\, r(x)\,.$$

Let

$$s(x) = \frac{r(x)}{4\left(1 + \frac{1}{\gamma}\right)\sqrt{d}}\,.$$

Then $2s(x)$ is smaller than the lateral length of Q, hence

$$\lambda(Q \cap B(y, s(x))) \geqq 2^{-d} \lambda(B(y, s(x)))\,.$$

Moreover, obviously $B(y, s(x)) \subset B(x, r(x))$. So we conclude that

$$\lambda(Q \cap B(x, r(x))) \geqq 2^{-d} \lambda(B(y, s(x))) = \beta \lambda(B(x, r(x)))\,.$$

Define

$$V := \delta 1_Q \varrho^{-2}$$

and let u denote the (unique) positive continuous function on U such that

$$u + Ku = 1$$

where

$$Ku := G^{uV\lambda} = \int u(y)\, V(y)\, G(\cdot, y)\, \lambda(dy)$$

and G denotes the Green function for U (see [HN1], [HN2]). Then evidently $u \leqq 1$ and $\Delta u - Vu = 0$ on U.

Moreover, $r(x) \leqq \varrho(y) + \|x - y\|$, hence

$$\operatorname{dist}(y, \complement B(x, r(x))) = r(x) - \|x - y\| \leqq \varrho(y)$$

for $y \in B(x, r(x))$, i.e., $V \leqq \delta 1_Q \operatorname{dist}(\cdot, \complement B(x, r(x)))^{-2}$. Thus

$$(*) \qquad u(x) \geqq \int_{\complement Q} u \, d\lambda_{B(x, r(x))} .$$

Since $\varrho \leqq \left(1 + \frac{1}{\gamma}\right) \delta(Q)$ on Q we have

$$\tilde{V} := \delta \left(1 + \frac{1}{\gamma}\right)^{-2} \delta(Q)^{-2} 1_Q \leqq \delta 1_Q \varrho^{-2} = V .$$

Let

$$g := H_{Q^\gamma}^{\Delta - \tilde{V}} 1 .$$

The scaling properties for the Schrödinger equation yield that

$$\sup g(Q) = \sup g_1(Q_1) = a .$$

On the other hand $g = 1 \geqq u$ on ∂Q^γ and

$$\Delta g - Vg = (\tilde{V} - V) g \leqq 0 \quad \text{on } Q^\gamma ,$$

i.e., g is $(\Delta - V)$-superharmonic on Q^γ. The boundary minimum principle for $(\Delta - V)$-superharmonic functions hence implies that $g - u \geqq 0$ on Q^γ. In particular,

$$\sup u(\bar{Q}) \leqq \sup g(\bar{Q}) = a .$$

There exists a continuous real potential q on U such that $\{e_Q \geqq \varepsilon q\}$ is a compact subset of U for each $\varepsilon > 0$. Fix $\varepsilon > 0$ and define a continuous real function v on U by

$$v = 1 - u - (1 - a) e_Q + \varepsilon q .$$

Then $v \geqq 0$ on $U \backslash \{e_Q \geqq \varepsilon q\}$, v is superharmonic on $U \backslash \bar{Q}$ and $v \geqq 1 - a - (1 - a) = 0$ on $\bar{Q}$. Hence the boundary minimum principle implies that $v \geqq 0$ on $U \backslash \bar{Q}$. This shows that

$$1 - u \geqq (1 - a) e_Q .$$

In order to finish the proof it suffices to establish that

$$e_Q^r \geqq 1 - u .$$

To that end we proceed similarly as for the proof of Proposition 2.6 in [HN2]. Fix $m \in \mathbb{N}$ and define a l.s.c. function φ_m on Y by

$$\varphi_m(x) = \liminf_{y \to x, y \in U} \left(u + \frac{q}{m} \right)(y) \quad \text{for } x \in Y.$$

Then $\varphi_m = u + q/m$ on U and hence by (*)

$$T_{r,Q}\varphi_m \leqq \varphi_m .$$

Fix $x_0 \in U$. Then the corresponding measure τ_{x_0} satisfies

$$\tau_{x_0}(\varphi_m) \leqq \varphi_m(x_0) = u(x_0) + \frac{q(x_0)}{m} .$$

As soon as we know that $\varphi_m \geqq 1$ on $Y \setminus U$ for every $m \in \mathbb{N}$ we conclude that $\tau_{x_0}(Y \setminus U) \leqq u(x_0)$, hence by our previous observations

$$e_Q^r(x_0) = \sigma_{x_0}(Q) = 1 - \tau_{x_0}(Y \setminus U) \geqq 1 - u(x_0)$$

and the proof will be finished.

So fix $m \in \mathbb{N}$. Let (U_n) be an exhaustion of U, i.e., the sets U_n are open, $\bar{U}_n \subset U_{n+1}$ for every $n \in \mathbb{N}$ and $\bigcup_{n=1}^{\infty} U_n = U$. There exist $x_n \in U \setminus U_n$, $n \in \mathbb{N}$, such that

$$\lim_{n \to \infty} \varphi_m(x_n) = \inf \varphi_m(Y \setminus U) .$$

If $\limsup_{n \to \infty} u(x_n) = 1$ then of course $\lim_{n \to \infty} \varphi_m(x_n) \geqq 1$ since $\varphi_m|_U \geqq u$. So suppose

$$\limsup_{n \to \infty} u(x_n) < 1 .$$

Then $\liminf_{n \to \infty} Ku(x_n) > 0$. Since $Ku \leqq 1$ and Ku is a continuous real potential on U which is harmonic on $U \setminus \bar{Q}$ we know that $Ku \leqq e_Q$. So our choice of q implies that $\lim_{n \to \infty} q(x_n) = \infty$ and hence $\lim_{n \to \infty} \varphi_m(x_n) = \infty$. Thus $\varphi_m \geqq 1$ on $Y \setminus U$ and the proof is complete. □

1.3. Proposition. *Let γ be a strictly positive real number. Then there exists a constant $c_1 > 0$ (depending only on γ and the dimension d) such that*

$$Pe_Q^r \leqq e_Q^r \leqq c_1 Pe_Q^r$$

for every cube Q in U satisfying $\delta(Q) \geqq \gamma \inf \varrho(Q)$.

Proof. We already noted that $Pe_Q^r \leqq e_Q^r$ on U and that $Pe_Q^r = e_Q^r$ on $U \setminus Q$. Since $e_Q^r = 1$ on Q, it hence suffices to find a constant $c_1 > 0$ such that $c_1 \lambda_{B(x,r(x))}(Q) \geqq 1$ for every $x \in Q$. So fix $x \in Q$. If $2r(x)$ is less than the lateral length of Q then obviously

$$\lambda_{B(x,r(x))}(Q) \geqq 2^{-d}.$$

If not, i.e., if $r(x) > \delta(Q)/(2\sqrt{d})$, then

$$\lambda_{B(x,r(x))}(Q) \geqq 2^{-d} \frac{\lambda\left(B\left(x, \frac{\delta(Q)}{2\sqrt{d}}\right)\right)}{\lambda(B(x,r(x)))} = \left(\frac{\delta(Q)}{4\sqrt{d}r(x)}\right)^d \geqq \left(\frac{1}{4\left(1+\frac{1}{\gamma}\right)\sqrt{d}}\right)^d$$

since $r(x) \leqq \varrho(x) \leqq \left(1 + \frac{1}{\gamma}\right)\delta(Q)$. □

2. Discrete Green function and P-potentials

As before let P denote the kernel given by

$$P(x, \cdot) = \lambda_{B(x,r(x))}$$

for $x \in U$ and let

$$p(x,y) := \frac{1}{\lambda(B(x,r(x)))} 1_{B(x,r(x))}(y)$$

for $x, y \in U$. Defining recursively

$$p_1 = p \quad \text{and} \quad p_{n+1}(x,y) = \int_U p_n(x,z)p(z,y)\,\lambda(dz)$$

for $x, y \in U$ we obviously have

$$P^n(x, \cdot) = p_n(x, \cdot)\,\lambda$$

for every $n \in \mathbb{N}$ and $x \in U$. By induction it is easily verified that

$$p_{n+1}(x,y) = \int_U p(x,z)p_n(z,y)\,\lambda(dz).$$

The (discrete) Green function associated with P is defined by

$$g(x,y) := \sum_{n=1}^{\infty} p_n(x,y) \quad \text{for } x, y \in U.$$

For every $y \in U$ the positive numerical function $g(\cdot, y)$ is P-supermedian since

$$P(g(\cdot, y))(x) = \int_U p(x,z) g(z,y) \lambda(dz)$$

$$= \sum_{n=1}^{\infty} \int_U p(x,z) p_n(z,y) \lambda(dz) = \sum_{n=1}^{\infty} p_{n+1}(x,y) \leqq g(x,y).$$

So for every measure v on U we obtain a P-supermedian function g^v on U defining

$$g^v(x) := \int g(x,y) v(dy) \quad \text{for } x \in U.$$

A P-supermedian function $s < \infty$ on U is called *P-potential* if $\lim_{n\to\infty} P^n s = 0$. It is evident that every P-supermedian minorant of a P-potential is a P-potential.

2.1. Lemma. *Let s be a (real) P-potential. Then Ps is a P-potential and*

$$Ps = g^{(s-Ps)\lambda}.$$

Proof. Of course Ps is P-supermedian and $Ps \leqq s < \infty$. So Ps is a P-potential. Moreover,

$$Ps = \sum_{k=1}^{n} P^k(s - Ps) + P^{n+1}s$$

for every $n \in \mathbb{N}$, hence

$$Ps = \sum_{k=1}^{\infty} P^k(s-Ps) = \sum_{k=1}^{\infty} \int p_k(\cdot, y)(s-Ps)(y) \lambda(dy) = g^{(s-Ps)\lambda}. \quad \square$$

For the following two statements assume that r is locally bounded away from zero.

2.2. Lemma. *Let s be a P-supermedian function on U such that $s \leqq q$ for some (classical) continuous real potential q on U. Then s is a P-potential.*

Proof. There exists a continuous real potential q' on U such that the set $\{q \geqq \varepsilon q'\}$ is compact for every $\varepsilon > 0$. Fix $\varepsilon > 0$ and let $\varphi = (q - \varepsilon q')^+$. Then φ is a continuous function with compact support in U, hence $\lim_{n\to\infty} P^n \varphi = 0$ by [HN 5]. Moreover, $s \leqq \varphi + \varepsilon q'$, hence

$$P^n s \leqq P^n \varphi + \varepsilon P^n q' \leqq P^n \varphi + \varepsilon q'$$

for every $n \in \mathbb{N}$ and we conclude that $\lim_{n\to\infty} P^n s = 0$. $\square$

2.3. Corollary. *For every cube Q such that $\bar{Q} \subset U$ the function e_Q^r is a P-potential and*

$$Pe_Q^r = g^{\sigma_Q^r}$$

where the measure $\sigma_Q^r := (e_Q^r - Pe_Q^r)\lambda$ is supported by Q.

Proof. e_Q^r is a P-supermedian function, $Pe_Q^r = e_Q^r$ on $U \setminus Q$ and e_Q^r is majorized by the (classical) equilibrium potential e_Q. So 2.3 is a consequence of 2.1 and 2.2. $\square$

3. Uniform Harnack inequalities for the discrete Green function

From now on we fix a real α such that $0 < \alpha \leqq 1$ and assume that

$$\alpha \varrho \leqq r .$$

Three elementary lemmas will lead to a uniform Harnack inequality for the functions $g(x, \cdot)$, $x \in U$.

3.1. Lemma. *For all $x, y, y' \in U$ such that $\|y - y'\| \leqq \alpha\varrho(y)/3$,*

$$\int g(x, \cdot)\, d\lambda_{B(y, \alpha\varrho(y)/3)} \leqq \left(\frac{4}{\alpha}\right)^d g(x, y') .$$

Proof. For every $z \in B(y, \alpha\varrho(y)/3)$

$$r(z) \geqq \alpha\varrho(z) \geqq \alpha(\varrho(y) - \|y - z\|) > \alpha(\varrho(y) - \alpha\varrho(y)/3) \geqq \frac{2\alpha}{3}\varrho(y)$$

and

$$r(z) \leqq \varrho(z) \leqq \varrho(y) + \alpha\varrho(y)/3 \leqq \frac{4}{3}\varrho(y) ,$$

hence

$$p(z, y') = \frac{1}{\lambda(B(z, r(z)))} \geqq \left(\frac{\alpha}{4}\right)^d \frac{1}{\lambda(B(y, \alpha\varrho(y)/3))} .$$

With $c' := \left(\frac{4}{\alpha}\right)^d$ this implies that

$$\int p_n(x, \cdot)\, d\lambda_{B(y, \alpha\varrho(y)/3)} \leqq c' \int_{B(y, \alpha\varrho(y)/3)} p_n(x, z) p(z, y')\, \lambda(dz) \leqq c' p_{n+1}(x, y')$$

for every $n \in \mathbb{N}$ and hence

$$\int g(x, \cdot)\, d\lambda_{B(y, \alpha\varrho(y)/3)} = \sum_{n=1}^{\infty} \int p_n(x, \cdot)\, d\lambda_{B(y, \alpha\varrho(y)/3)}$$

$$\leqq c' \sum_{n=1}^{\infty} p_{n+1}(x, y') \leqq c' g(x, y') . \quad \square$$

3.2. Lemma. *For all $x, y \in U$*

$$p(x, y) \leqq 2^d \int p(x, \cdot)\, d\lambda_{B(y, \alpha\varrho(y)/3)} .$$

Proof. If $y \notin B(x, r(x))$ then $p(x, y) = 0$. So assume that $y \in B(x, r(x))$. Since

$$p(x, \cdot) = p(x, y) = (\lambda(B(x, r(x))))^{-1}$$

on $B(x, r(x))$, it suffices to show that $1 \leqq 2^d \lambda_{B(y, \alpha\varrho(y)/3)}(B(x, r(x)))$.

If $x \in B(y, \alpha\varrho(y)/3)$, then $r(x) \geqq \frac{2}{3}\alpha\varrho(y)$, hence $B(y, \alpha\varrho(y)/3) \subset B(x, r(x))$ and $\lambda_{B(y,\alpha\varrho(y)/3)}(B(x, r(x))) = 1$. So suppose that $x \notin B(y, \alpha\varrho(y)/3)$ and let z be the point on the segment $[x, y]$ having distance $\alpha\varrho(y)/6$ from y. Then $B(z, \alpha\varrho(y)/6)$ is a subset of $B(x, r(x)) \cap B(y, \alpha\varrho(y)/3)$ and hence

$$\lambda_{B(y,\alpha\varrho(y)/3}(B(x, r(x))) \geqq \lambda_{B(y,\alpha\varrho(y)/3)}(B(z, \alpha\varrho(y)/6)) = 2^{-d}. \quad \square$$

3.3. Lemma. *For all $x, y \in U$,*

$$g(x, y) \leqq 2^d \int g(x, \cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)}.$$

Proof. By 3.2

$$p(x, y) \leqq 2^d \int p(x, \cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)}$$

and

$$\begin{aligned} p_{n+1}(x, y) &= \int_U p_n(x, z) p(z, y)\, \lambda(dz) \\ &\leqq 2^d \int_U p_n(x, z) \left(\int p(z, \cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)}\right) \lambda(dz) \\ &= 2^d \int \left(\int_U p_n(x, z) p(z, \cdot)\, \lambda(dz)\right) d\lambda_{B(y,\alpha\varrho(y)/3)} \\ &= 2^d \int p_{n+1}(x, \cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)} \end{aligned}$$

for every $n \in \mathbb{N}$. The proof is completed taking the sum over n. $\square$

Combining 3.1 and 3.3 and defining

$$c_2 := \left(\frac{8}{\alpha}\right)^d$$

we obtain the desired Harnack inequality:

3.4. Proposition. *For all $x, y, y' \in U$ such that $\|y - y'\| \leqq \alpha\varrho(y)/3$,*

$$g(x, y) \leqq c_2\, g(x, y').$$

Let us recall that for every Q with $\bar{Q} \subset U$ we know by (2.3) that e_Q^r is a P-potential and $Pe_Q^r = g^{\sigma_Q^r}$ with $\sigma_Q^r = (e_Q^r - Pe_Q^r)\lambda$ and $\sigma_Q^r(\complement Q) = 0$. For every measure ν on U let $\|\nu\|$ denote the total mass of ν. Proposition 3.4 has the following useful consequence.

3.5. Corollary. *Let Q be a cube in U such that $\delta(Q) \leqq \frac{\alpha}{3} \inf \varrho(Q)$ and let ν be a finite measure ν on $\bar{Q}$. Then g^ν is a P-potential and*

$$g^\nu \leqq c_2 \frac{\|\nu\|}{\|\sigma_Q^r\|}.$$

Proof. It suffices to show that $g^\nu \leqq c_2 \|\nu\| / \|\sigma_Q^r\|$ (and hence $g^\nu < \infty$). Fix $x \in U$. If $y, y' \in \bar{Q}$ then $\|y - y'\| \leqq \delta(Q) \leqq \frac{\alpha}{3}\varrho(y)$, hence $g(\cdot, y) \leqq c_2 g(\cdot, y')$. This shows that

$$\|\sigma_Q^r\| g(x,y) = \int g(x,y)\,\sigma_Q^r(dy') \leqq c_2 \int g(x,y')\,\sigma_Q^r(dy') = c_2 g^{\sigma_Q^r}(x) \leqq c_2$$

for every $y \in \bar{Q}$ and therefore

$$\|\sigma_Q^r\| g^\nu(x) = \|\sigma_Q^r\| \int g(x,y)\,\nu(dy) \leqq c_2 \|\nu\| . \quad \square$$

4. Decomposition $\mathscr{Q}_a$ of U into cubes

In order to combine the results of the preceding sections it will be useful to have a decomposition of U into cubes Q such that $\gamma \leqq \delta(Q)/\inf \varrho(Q) \leqq \frac{\alpha}{3}$ for some constant $\gamma > 0$.

If $U \neq \mathbb{R}^d$ let us start with a Whitney decomposition $\mathscr{Q}$ of U, i.e., $\mathscr{Q}$ is a partition of U into countably many disjoint (dyadic) cubes of the form

$$Q = \prod_{i=1}^{d} [m_i 2^{-m}, (m_i + 1) 2^{-m}[, \quad m \in \mathbb{Z},\ m_i \in \mathbb{Z}$$

such that

$$\frac{1}{3} < \frac{\delta(Q)}{\operatorname{dist}(Q, \complement U)} \leqq 1 .$$

In particular, $\delta(Q) > \frac{1}{3} \inf \varrho(Q)$ for every $Q \in \mathscr{Q}$ since $\varrho \leqq \operatorname{dist}(\cdot, \complement U)$.

If $U = \mathbb{R}^d$ $(d \geqq 3)$ we take $M = \varrho(0)$, $Z = \mathbb{Z}^d \cap \partial([-2,1]^d)$ and begin with

$$\mathscr{Q} := \{[-M, M[^d\} \cup \{2^m M(z + [0,1[^d) : m \in \mathbb{Z}^+, z \in Z\} .$$

If $m \in \mathbb{Z}^+$, $z \in Z$ and $Q = 2^m M(z + [0,1[^d)$ then $\delta(Q) = 2^m M\sqrt{d}$ and

$$\bar{Q} \cap [-2^m M, 2^m M]^d \neq \emptyset ,$$

hence $\inf \varrho(Q) \leqq \varrho(0) + 2^m M\sqrt{d} \leqq 2^{m+1} M\sqrt{d}$. Therefore $\delta(Q) \geqq \frac{1}{2} \inf \varrho(Q)$ for all $Q \in \mathscr{Q}$ (this is obvious for $Q = [-M, M[^d$).

In both cases we now construct $\mathscr{Q}_\alpha$ by successive subdivisions of the cubes in $\mathscr{Q}$ where a subdivision of a cube

$$Q = \prod_{i=1}^{d} [a_i, a_i + b[$$

consists in replacing it by the 2^d subcubes

$$\prod_{i=1}^{d} \left[a_i + \varepsilon_i \frac{b}{2}, a_i + (1+\varepsilon_i)\frac{b}{2} \right[$$

with $\varepsilon_i \in \{0,1\}$ for $1 \leqq i \leqq d$.

Given $Q_1 \in \mathcal{Q}$ we first take a subdivision of Q_1. For each resulting cube Q we evidently have $\delta(Q) = \frac{1}{2}\delta(Q_1)$ and $\inf \varrho(Q) \geqq \inf \varrho(Q_1)$, and if $\delta(Q) > \frac{\alpha}{3} \inf \varrho(Q)$ the cube Q will be subdivided. And so on. If $k \in \mathbb{N}$ such that $2^{-k} < \frac{\alpha}{3} \inf \varrho(Q_1)/\delta(Q_1)$ this procedure will stop after at most k subdivisions and we shall have obtained a partition of Q_1 into finitely many half-open cubes Q such that $\delta(Q) \leqq \frac{\alpha}{3} \inf \varrho(Q)$ and Q is contained in a cube Q' with $\delta(Q') = 2\,\delta(Q)$ and $\delta(Q')/\inf \varrho(Q') > \frac{\alpha}{3}$. This implies that

$$\inf \varrho(Q) \leqq \inf \varrho(Q') + \delta(Q) < \left(\frac{6}{\alpha} + 1\right)\delta(Q) \leqq \frac{7}{\alpha}\delta(Q).$$

Thus

$$\frac{\alpha}{7} < \frac{\delta(Q)}{\inf \varrho(Q)} \leqq \frac{\alpha}{3}.$$

Applying this construction to each $Q_1 \in \mathcal{Q}$ we get the partition $\mathcal{Q}_\alpha$.

5. Discrete Martin kernel versus classical Martin kernel

Let us fixe a cube $Q_0 \in \mathcal{Q}_a$ and let $\lambda_0 = 1_{Q_0}\lambda$. By 3.5

$$\int g(x,y)\,\lambda_0(dx) \leqq \frac{c_2}{\|\sigma_Q^r\|}\lambda(Q_0) < \infty$$

for every $Q \in \mathcal{Q}_\alpha$ and for every $y \in Q$. So we may define Martin kernels k and K on $U \times U$ by

$$k(x,y) = \frac{g(x,y)}{\lambda_0(g(\cdot,y))}, \quad K(x,y) = \frac{G(x,y)}{\lambda_0(G(\cdot,y))}$$

for $x, y \in U$. For every measure μ on U we define a P-supermedian function k^μ and a hyperharmonic function K^μ on U by

$$k^\mu(x) = \int k(x,y)\,\mu(dy), \quad K^\mu(x) = \int K(x,y)\,\mu(dy)$$

for $x \in U$. Of course,

$$\lambda_0(k^\mu) = \int \lambda_0(k(\cdot,y))\,\mu(dy) = \int 1\,d\mu = \|\mu\|$$

and similarly

$$\lambda_0(K^\mu) = \|\mu\|\,.$$

For every $Q \in \mathcal{Q}_\alpha$ let

$$\tau_Q = \left(\int G(x,\cdot)\,\lambda_0(dx)\right)\sigma_Q$$

(i.e., $\tau_Q = G^{\lambda_0}\sigma_Q$ because of the symmetry of G) so that

$$K^{\tau_Q} = G^{\sigma_Q} = e_Q\,.$$

Defining $\tau_Q^r = \left(\int g(x,\cdot)\,\lambda_0(dx)\right)\sigma_Q^r$ we have

$$k^{\tau_Q^r} = g^{\sigma_Q^r} = Pe_Q^r\,.$$

For every $Q \in \mathcal{Q}_\alpha$ and for all $y, y' \in \bar{Q}$ we have by 3.4 the inequality

$$g(x,y) \leqq c_2\, g(x,y')$$

for every $x \in U$, and hence

$$\lambda_0(g(\cdot,y)) \leqq c_2\,\lambda_0(g(\cdot,y'))\,.$$

So we obtain the following

5.1. Lemma. *For every $x \in U$, for every $Q \in \mathcal{Q}_\alpha$ and for all $y, y' \in \bar{Q}$,*

$$k(x,y) \leqq c_2^2\, k(x,y')\,.$$

By 1.2 and 1.3 there exists a constant $c > 0$ such that

$$Pe_Q^r \leqq e_Q \leqq c\,Pe_Q^r$$

for every $Q \in \mathcal{Q}_\alpha$. Let

$$C = c\,c_2^4$$

(and note that C depends only on α and the dimension d). The following theorem is the key to harmonic majorants for positive r-median functions.

5.2. Theorem. *For every $Q \in \mathcal{Q}_\alpha$ and for every measure ν on $\bar{Q}$ with $\|\nu\| = \|\tau_Q\|$,*

$$C^{-1} K^{\tau_Q} \leqq k^\nu \leqq C K^{\tau_Q}\,.$$

Proof. Lemma 5.1 implies that

$$c_2^{-2} k^{\tau_Q^r} \leqq \frac{\|\tau_Q^\tau\|}{\|\nu\|}\, k^\nu \leqq c_2^2\, k^{\tau_Q^r}$$

(see the proof of 3.5) where $k^{\tau_Q^r} = Pe_Q^r$, $Pe_Q^r \leqq e_Q \leqq c\,Pe_Q^r$ and $e_Q = K^{\tau_Q}$. Therefore

$$(c c_2^2)^{-1} K^{\tau_Q} \leqq \frac{\|\tau_Q^r\|}{\|\nu\|}\, k^\nu \leqq c_2^2\, K^{\tau_Q}\,.$$

Integrating with respect to λ_0 we obtain that

$$(cc_2^2)^{-1}\|\tau_Q\| \leqq \|\tau_Q^r\| \leqq c_2^2\|\tau_Q\|$$

where by assumption $\|\tau_Q\| = \|v\|$. Thus

$$(cc_2^4)^{-1}K^{\tau_Q} \leqq k^v \leqq cc_2^4 K^{\tau_Q}. \quad \square$$

6. Positive r-median functions

After the preparations in the preceding section it is now easy to establish the main result of this paper.

6.1. Theorem. *Let U be a non-empty domain in $\mathbb{R}^d$, $d \geqq 1$, let r be a strictly positive real function on U such that $B(x, r(x)) \subset U$ for every $x \in U$. If $U = \mathbb{R}^d$ assume that $r \leqq \|\cdot\| + M$ for some $M \in \mathbb{R}_+$. Let $\alpha \in]0,1]$ such that $r(y) \geqq \alpha(r(x) - \|x-y\|)$ for every $x \in U$ and all $y \in B(x, r(x))$. Then every r-median function $f \geqq 0$ on U is harmonic.*

Proof. We already noted in the introduction that because of the results in [HN 7] we only have to study the case where U is a Green domain. By [Ve 2] (cf. also Lemma 3.8 in [HN 1]) we may assume that r and hence the r-median function $f \geqq 0$ is Borel measurable. So we may use the notations and results of the preceding sections. Let us note that $\int_Q f\,d\lambda < \infty$ for every $Q \in \mathcal{Q}_\alpha$. Indeed, taking $x \in Q$ we have $Q \subset B(x, r(x))$, hence

$$\int_Q f\,d\lambda \leqq \int_{B(x,r(x))} f\,d\lambda = \lambda(B(x, r(x)))f(x) < \infty.$$

In particular, $\lambda_0(f) < \infty$.

Let q be a strictly positive continuous real potential on U. For every $n \in \mathbb{N}$ let

$$f_n := \inf(f, nq).$$

For the moment fix $n \in \mathbb{N}$. Since f_n is P-supermedian and bounded by the potential nq, we know by 2.2 that f_n is a P-potential and hence by 2.1

$$Pf_n = g^{\varphi_n \lambda} \quad \text{with } \varphi_n = f_n - Pf_n.$$

Then of course

$$Pf_n = k^{\mu_n} \quad \text{for } \mu_n := \left(\int g(x,\cdot)\,\lambda_0(dx)\right)\varphi_n \lambda$$

and $Pf_n \leqq f$. Define

$$v_n := \sum_{Q \in \mathcal{Q}_\alpha} \frac{\mu_n(Q)}{\|\tau_Q\|}\tau_Q.$$

By Theorem 5.2 we know that

$$C^{-1}\frac{\mu_n(Q)}{\|\tau_Q\|}K^{\tau_Q} \leqq k^{1_Q \mu_n} \leqq C\frac{\mu_n(Q)}{\|\tau_Q\|}K^{\tau_Q}$$

for every $Q \in \mathcal{Q}_\alpha$. Taking the sum we conclude that

$$C^{-1}K^{\nu_n} \leqq k^{\mu_n} \leqq CK^{\nu_n}.$$

Integrating the left inequality with respect to λ_0 we get

$$\|\nu_n\| = \lambda_0(K^{\nu_n}) \leqq C\lambda_0(k^{\mu_n}) \leqq C\lambda_0(f) < \infty.$$

Thus the sequence $(\|\nu_n\|)$ is bounded and there exists a subsequence (ν_n') of (ν_n) converging weakly to a finite measure ν on the (classical) Martin compactification $U \cup \partial^M U$ of U and

$$\lim_{n\to\infty} K^{\nu_n'} = K^{\nu}.$$

Evidently, the sequence (f_n) is increasing to f, hence the sequence (Pf_n) is increasing to $Pf = f$ and we obtain that

$$f \leqq CK^{\nu}.$$

By [HN1], [HN4] it remains to show that K^{ν} is harmonic, i.e., that $\nu(U) = 0$.

We just noted that $\lim_{n\to\infty} f_n = \lim_{n\to\infty} Pf_n = f$. Therefore $\lim_{n\to\infty} \varphi_n = 0$. Moreover, $0 \leqq \varphi_n \leqq f$ and by 3.5

$$\int g(x,y)\,\lambda_0(dx) \leqq \frac{c_2}{\|\sigma_Q^r\|}\,\lambda(Q_0)$$

for every $Q \in \mathcal{Q}_\alpha$ and $y \in Q$. So $\lim_{n\to\infty} \mu_n(Q) = 0$ for every $Q \in \mathcal{Q}_\alpha$ by Lebesgue's convergence theorem. If ψ is a continuous real function with compact support in U, then there are only finitely many $Q \in \mathcal{Q}_\alpha$ such that $\bar{Q} \cap \{\psi \neq 0\} \neq \emptyset$, and hence

$$\lim_{n\to\infty} \nu_n(\psi) = 0.$$

Thus $\nu(U) = 0$ and the proof is completed.

Let us note, however, that we can completely avoid to use the previous result on harmonically bounded functions: Knowing that for every positive r-median function there exists a positive harmonic function h on U such that $C^{-1}h \leqq f \leqq Ch$ it is obvious that positive r-median functions satisfy Harnack inequalities. So we may apply Choquet theory and it is sufficient to consider extremal r-median functions. But for any extremal r-median function $f > 0$ the inequality $C^{-1}h \leqq f$, $h > 0$ harmonic on U, implies that f is a multiple of h and hence harmonic. □

7. Generalizations

In view of [HN2], [HN4], [HN5] it is possible to obtain the preceding results for more general means:

Let U be a Green domain (with Green function G) in $\mathbb{R}^d$, $d \geqq 1$, let $p: U \times U \to \mathbb{R}^+$ be a Borel measurable function, and define

$$\mu_x = p(x, \cdot)\,\lambda$$

for $x \in U$. Let r and ϱ be as before, suppose that $\alpha\varrho \leqq r$ for some real $\alpha > 0$ and that for some strictly positive numbers a, b, β and some $\eta \in]0,1[$ the following holds for every $x \in U$:

(i) $\mu_x(h) = h(x)$ for every harmonic function $h \geqq 0$ on U,
$\mu_x(s) \leqq s(x)$ for every superharmonic function $s \geqq 0$ on U,

(ii) $p(x,y) \leqq a \int p(x,\cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)}$ for every $y \in U$,

(iii) $p(x,\cdot) \geqq \beta(\lambda(B(x,r(x))))^{-1}$ on $B(x,r(x))$ and
$(G^{\varepsilon_x} - G^{\mu_x})\,1_{U \setminus B(x,\eta r(x))} \leqq b\varrho^2 p(x,\cdot)$.

If $d \leqq 2$ we suppose in addition that

(iv) $\sup\{r(x)^{-2} \int (G^{\varepsilon_x} - G^{\mu_x})\, d\lambda : x \in U\} < \infty$.

A Lebesgue measurable function $f \geqq 0$ on U is said to be *μ-supermedian* (*μ-median* resp.) if

$$\int f\, d\mu_x \leqq f(x) \qquad \left(\int f\, d\mu_x = f(x) < \infty \text{ resp.}\right)$$

for every $x \in U$.

Properties (i), (iii), and (iv) are already familiar from [HN 2], [HN 4] (recall that (iv) holds if there exists $c > 0$ such that, for every $x \in U$, the measure μ_x is supported by a domain $V_x \subset B(x, cr(x))$). The additional property (ii) is a mild regularity condition replacing Lemma 3.2. Defining (p_n) recursively as in Section 2 and taking $g = \sum_{n=1}^{\infty} p_n$ the proof of Lemma 3.3 shows that

$$g(x,y) \leqq a \int g(x,\cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)}$$

for all $x, y \in U$. Moreover, the first inequality of (iii) implies that

$$\int g(x,\cdot)\, d\lambda_{B(y,\alpha\varrho(y)/3)} \leqq \frac{1}{\beta}\left(\frac{4}{\alpha}\right)^d g(x,y')$$

for all $x, y, y' \in U$ such that $\|y - y'\| \leqq \alpha\varrho(y)/3$. Indeed, the proof of Lemma 3.1 carries over to our more general situation if we replace the equality $p(z,y') = (\lambda(B(z,r(z))))^{-1}$ by the inequality $p(z,y') \geqq \beta(\lambda(B(z,r(z))))^{-1}$. Taking now $c_2 := (4/\alpha)^d a/\beta$ we thus get as before that

$$g(x,y) \leqq c_2 g(x,y')$$

for all $x, y, y' \in U$ such that $\|y - y'\| \leqq \alpha\varrho(y)/3$.

From Proposition 2.5 in [HN 2], Proposition 2.7 in [HN 4], and from [HN 5] we know that the results of Sections 1 and 2 hold as well in our more general situation. Proceeding as in Sections 5 and 6 we hence conclude that there exists a constant $C > 0$ such that for every μ-median function $f \geqq 0$ on U there exists a harmonic function $h \geqq 0$

on U with $C^{-1}h \leqq f \leqq Ch$. So the results of [HN2], [HN4] (or the reasoning used at the end of the proof of 6.1) imply that any such function f is harmonic. Thus we obtain the following result:

7.1. Theorem. *Every μ-median function $f \geqq 0$ on U is harmonic.*

7.2. Remark. Again it is clear that our method applies as well to more general elliptic operators.

References

[Ba1] *J. R. Baxter*, Restricted mean values and harmonic functions, Trans. Amer. Math. Soc. **167** (1972), 451–463.

[Ba2] *J. R. Baxter*, Harmonic functions and mass cancellation, Trans. Amer. Math. Soc. **245** (1978), 375–384.

[BH] *J. Bliedtner, W. Hansen*, Potential Theory – An Analytic and Probabilistic Approach to Balayage, Universitext, Springer, Berlin–Heidelberg–New York–Tokyo 1986.

[CV] *A. Cornea, J. Veselý*, Martin compactification for discrete potential theory, Potential Analysis **4** (1995), 547–569.

[HN1] *W. Hansen, N. Nadirashvili*, A converse to the mean value theorem for harmonic functions, Acta Math. **171** (1993), 139–163.

[HN2] *W. Hansen, N. Nadirashvili*, Mean values and harmonic functions, Math. Ann. **297** (1993), 157–170.

[HN3] *W. Hansen, N. Nadirashvili*, Littlewood's one circle problem, J. London Math. Soc. **50** (1994), 349–360.

[HN4] *W. Hansen, N. Nadirashvili*, Liouville's theorem and the restricted mean value property, J. Math. Pures Appl. **50** (1994), 249–260.

[HN5] *W. Hansen, N. Nadirashvili*, On the restricted mean value property for measurable functions, Classical and Modern Potential Theory and Applications, NATO ASI Series, eds. K. GowriSankaran et al., Kluwer (1994), 267–271.

[HN6] *W. Hansen, N. Nadirashvili*, On Veech's conjecture for harmonic functions, Ann. Sc. Norm. Sup. Pisa Cl. Sci. **22** (1995), 137–153.

[HN7] *W. Hansen, N. Nadirashvili*, Restricted mean value property on $\mathbb{R}^d$, $d \leqq 2$, Exposition. Math. **13** (1995), 93–95.

[Hu] *F. Huckemann*, On the 'one circle' problem for harmonic functions, J. London Math. Soc. (2) **29** (1954), 491–497.

[NV] *I. Netuka, J. Veselý*, Mean value property and harmonic functions, Classical and Modern Potential Theory and Applications, NATO ASI Series, eds. K. GowriSankaran et al., Kluwer (1994), 359–398.

[Su] *L. Susanka*, Potential theoretic and stopping time methods for converse mean value theorems, Thesis 1991.

[Ve1] *W. A. Veech*, A converse to Gauss' theorem, Bull. Amer. Math. Soc. **78** (1971), 444–446.

[Ve2] *W. A. Veech*, A zero-one law for a class of random walks and a converse to Gauss' mean value theorem, Ann. Math. (2) **97** (1973), 189–216.

[Ve3] *W. A. Veech*, A converse to the mean value theorem for harmonic functions, Amer. J. Math. **97** (1975), 1007–1027.

Fakultät für Mathematik, Universität Bielefeld, Universitätsstraße, D-33615 Bielefeld

e-mail: hansen@mathematik.uni-bielefeld.de

Eingegangen 3. April 1995

J. reine angew. Math. **470** (1996), 109—122

Journal für die reine und
angewandte Mathematik

A Barth-Lefschetz type theorem for branched coverings of Grassmannians

By *Meeyoung Kim* at Notre Dame

---

1. Introduction

In [L1], R. Lazarsfeld proved a generalization of Barth-Larsen theorem [B]: *Let X be an n-dimensional connected complex projective manifold and $f: X \to \mathbb{P}^n$ be a branched covering of degree d. Then the induced maps $f^*: H^i(\mathbb{P}^n, \mathbb{C}) \to H^i(X, \mathbb{C})$ are isomorphisms for $i \leqq n - d + 1$.*

There is a certain vector bundle $\mathscr{E}$ on $\mathbb{P}^n$ associated with f, which satisfies a strong positivity condition, i.e. $\mathscr{E}(-1)$ is spanned. In this proof, the ampleness of $\mathscr{E}$ plays an essential role in applying Sommese's Lefschetz-type theorem (cf. [S1], Proposition 1.17). Lazarsfeld also showed the homotopy version of the above theorem using the spannedness of $\mathscr{E}(-1)$ (cf. [L2], Theorem 5.1).

Once we replace $\mathbb{P}^n$ by the Grassmannian $\mathrm{Gr}(r, n)$, which is the set of $n - r$ dimensional linear subspaces in $\mathbb{C}^n$, in general the associated vector bundle $\mathscr{E}$ may be no longer even ample, but we can still expect it to be spanned. In Proposition (3.2), we prove that $\mathscr{E}$ is spanned. Not only can we expect spannedness, but also some further positivity of $\mathscr{E}$, namely k-ampleness for some k; k-ampleness is the generalization of ampleness introduced by Sommese [S1]:

Let E be a spanned vector bundle on a compact complex manifold X. E is k-ample if for each subvariety $Z \subset X$ such that $E|_Z$ has a trivial quotient bundle, it is true that $\dim Z \leqq k$, equivalently, the maximum of the dimensions of the fibres of the map $\mathbb{P}(E) \to \mathbb{P}^N$ associated with $H^0(\mathbb{P}(E), \xi_E)$ is less than or equal to k where ξ_E is the tautological line bundle on $\mathbb{P}(E)$.

Using this concept we get, for Grassmannians, a result analogous to Lazarsfeld's, see Corollary (4.3). The analogous homotopy version, which answers the question raised in [FL], p. 86, is Proposition (4.4).

The k-ampleness of $\mathscr{E}$ essentially guides the bound for the cohomology isomorphisms in (4.3) which is obtained by using Schubert calculus, appropriate Koszul complexes, and Kodaira-Griffiths vanishing theorems.

I am grateful to my advisor, A. J. Sommese, for suggesting and discussing this problem. I would also like to thank M. A. de Cataldo and J.-M. Hwang for helpful discussions.

Notation. Throughout the paper, we will use the following conventions:

(1.1) Little or no distinction is made between a vector bundle E and the associated sheaf of germs of holomorphic sections.

(1.2) $E^{\oplus n}$, $E^{\otimes n}$ and $E^{(n)}$ denote n-th direct sum, tensor power and symmetric power of a vector bundle E, respectively.

(1.3) Given a continuous map $f\colon Y \to Z$ and a sheaf of abelian groups $\mathscr{G}$ on Y, $f_{(q)}(\mathscr{G})$ denotes the q-th direct image sheaf.

(1.4) We say a vector bundle E is spanned on X if its associated sheaf is generated by its global sections at every point of X.

(1.5) We use the notation of Hodge-Pedoe [HP], XIV, when dealing with Schubert calculus.

In the next section, we recall some basic facts about Grassmannians and Grassmannian bundles that are needed in the sequel, but for which we could not find an adequate reference.

2. Grassmannians

Let X be a compact complex manifold, E be a vector bundle of rank e on X, $\mathrm{Gr}(s, E)$ denote the set of $e - s$ dimensional linear subspaces of the fibres of E. Let $\pi : \mathrm{Gr}(s, E) \to X$ be the projection induced by the bundle projection $E \to X$. There is the universal exact sequence:

$$0 \to F^{\vee} \to \pi^* E \to Q \to 0$$

on $\mathrm{Gr}(s, E)$ where $F^{\vee}$ is the tautological rank $(e - s)$-subbundle of $\pi^* E$. We denote by $\xi_{s,E}$, the tautological line bundle $\det Q$ on $\mathrm{Gr}(s, E)$. Indeed, the restriction of $\xi_{s,E}$ to each fibre of π is $\iota^* \mathcal{O}_{\mathbb{P}^N}(1)$, where $\mathrm{Gr}(s, e) \overset{\iota}{\hookrightarrow} \mathbb{P}^N$ is the Plücker embedding. In particular, $\mathrm{Gr}(1, E)$ is $\mathbb{P}(E)$ with projection $\pi : \mathbb{P}(E) \to X$ and $\xi_{1,E}$ is the tautological line bundle on $\mathbb{P}(E)$.

(2.1) For π: $\mathbb{P}(E) \to X$ and a coherent sheaf $\mathscr{G}$ on X, we have:

$$\pi_{(q)}(\pi^* \mathscr{G} \otimes \Omega^p_{\mathbb{P}(E)} \otimes \xi^m_{1,E}) = 0 \quad \text{where } p \geqq 0,\ q > 0, \quad \text{and} \quad m \geqq 1,$$

so that we deduce:

$$H^q(X, \mathscr{G} \otimes E^{(m)}) \cong H^q\big(\mathbb{P}(E), \pi^* \mathscr{G} \otimes \xi^m_{1,E}\big) \quad \text{for} \quad q \geqq 0,\ m \geqq 1.$$

(2.2) We know that $\pi_* \xi^k_{1,E} \cong E^{(k)}$ for $\mathrm{Gr}(1, E) = \mathbb{P}(E) \xrightarrow{\pi} X$ and $k \geqq 1$. If $s \geqq 2$, then this is no longer true. However the following holds:

$\pi_* \xi^k_{s,E}$ is a quotient bundle of the k-th symmetric power $(\wedge^s E)^{(k)}$ for $\pi : \mathrm{Gr}(s, E) \to X$. In particular, we see that $\pi_* \xi_{s,E} \cong \wedge^s E$.

The following generalization [SS], Theorem 5.45, of Le Potier's lemma (cf. [SS], Theorem 5.16) enables us to carry vanishing theorems for line bundles over to vector bundles.

(2.3) Theorem (Shiffman-Sommese). *Let $\pi : B \to X$ be a fibre bundle with compact fibre B_x, $x \in X$ where B_x, B and X are complex manifolds. Let $\mathscr{V}$ be a vector bundle on B_x. Suppose ξ is a vector bundle on B such that for all $x \in X$, there exists a neighborhood U of x and a local trivialization $\phi = (\pi, \phi_2) : \pi^{-1}(U) \to U \times B_x$ such that $\xi = \phi_2^* \mathscr{V}$ on $\pi^{-1}(U)$. Then $\pi_* \xi$ is a vector bundle on X with fibres $\pi_* \xi \otimes_{\mathcal{O}_X} k(x) = \Gamma(B_x, \xi|_{B_x})$ for every $x \in X$. If $H^q(B_x, \Omega^p_{B_x}(\mathscr{V})) = 0$ for $(p, q) \neq (0, 0)$ and for all $x \in X$, then*

$$H^q(X, \mathscr{G} \otimes \Omega^p_X \otimes \pi_* \xi) \cong H^q(B, \pi^* \mathscr{G} \otimes \Omega^p_B \otimes \xi)$$

for all $p, q \geqq 0$ and for all coherent sheaves $\mathscr{G}$ on X.

(2.4) Remark. By Le Potier vanishing theorem for Grasmannians (cf. [SS], Theorem 4.56):

$$H^q(\mathrm{Gr}(s, e), \Omega^p_{\mathrm{Gr}(s,e)} \otimes \mathcal{O}_{\mathrm{Gr}(s,e)}(1)) = 0, \quad \text{for } (p, q) \neq (0, 0),$$

we can apply Theorem (2.3) to the Grassmannian bundle $B = \mathrm{Gr}(s, E) \xrightarrow{\pi} X$. We get:

$$H^q(X, \mathscr{G} \otimes \Omega^p_X \otimes \wedge^s E) \cong H^q(X, \mathscr{G} \otimes \Omega^p_X \otimes \pi_* \xi_{s,E}) \cong H^q(\mathrm{Gr}(s, E), \pi^* \mathscr{G} \otimes \Omega^p_{\mathrm{Gr}(s,E)} \otimes \xi_{s,E})$$

for all $p, q \geqq 0$ and for any coherent sheaf $\mathscr{G}$ on X.

The following two propositions are Griffiths and Le Potier vanishing theorems (cf. [SS], Theorem 5.52 and Corollary 5.17); they are generalizations of Kodaira vanishing theorem which will be useful in the sequel of this paper.

(2.5) Proposition (Griffiths). *Let $\mathscr{G}$ be a spanned vector bundle on a compact complex manifold X. Suppose L is an ample line bundle on X. Then*

$$H^q(X, K_X \otimes \mathscr{G}^{(m)} \otimes \det \mathscr{G} \otimes L) = 0 \quad \textit{for} \quad q > 0, \quad m \geqq 0.$$

(2.6) Proposition (Le Potier). *Let E be an ample vector bundle of rank r on a compact complex manifold X. Then*

$$H^q(X, K_X \otimes E) = 0 \quad \textit{for} \quad q \geqq r.$$

Now let $V = \mathbb{C}^n \times \mathrm{Gr}(r, n)$ be the trivial n-bundle over $\mathrm{Gr}(r, n)$. On $\mathrm{Gr}(r, n)$ we have the universal exact sequence:

$$0 \to \mathscr{U}^\vee \to V \to \mathscr{Q} \to 0 \tag{2.7}$$

where $\mathcal{U}^\vee$ is the universal rank $n-r$ subbundle of V and Q is the universal rank r quotient bundle. From the above exact sequence, $\mathcal{U}$ and $\mathcal{Q}$ are immediately seen to be spanned.

(2.8) Since $\det\mathcal{Q} \cong \det\mathcal{U} \cong \mathcal{O}_{\mathrm{Gr}(r,n)}(1)$, we have

$$\wedge^{r-l}\mathcal{Q} \cong \wedge^{l}\mathcal{Q}^\vee(1), \quad \wedge^{n-r-l'}\mathcal{U} \cong \wedge^{l'}\mathcal{U}^\vee(1) \quad \text{for } l = 0, 1, \ldots, r \text{ and } l' = 0, 1, \ldots, n-r.$$

(2.9) The tangent bundle $T_{\mathrm{Gr}(r,n)} \cong \mathrm{Hom}_{\mathcal{O}_{\mathrm{Gr}(r,n)}}(\mathcal{U}^\vee, \mathcal{Q}) \cong \mathcal{U} \otimes \mathcal{Q}$, and the canonical bundle $K_{\mathrm{Gr}(r,n)} \cong \mathcal{O}_{\mathrm{Gr}(r,n)}(-n)$.

(2.10) Note that $h^0(\mathrm{Gr}(r,n), \mathcal{U}) = h^0(\mathrm{Gr}(r,n), \mathcal{Q}) = h^0(\mathrm{Gr}(r,n), V) = n$ because $\mathcal{U}^\vee$ and $\mathcal{Q}^\vee$ have no global sections and $h^1(\mathrm{Gr}(r,n), \mathcal{U}^\vee) = h^1(\mathrm{Gr}(r,n), \mathcal{Q}^\vee) = 0$ by (2.5).

With the notion of k-ampleness, we have:

(2.11) $\mathcal{U}$ and $\mathcal{Q}$ are $r(n-r)-r$ ample and $r(n-r)-(n-r)$ ample, respectively.

Proof. The statement for $\mathcal{U}$ follows by using the universal exact sequence (2.7), by the fact that the zero locus of a section of $\mathcal{Q}$ is the subgrassmannian $\mathrm{Gr}(r, n-1)$ and by the fact that for any subvariety $Z \subset \mathrm{Gr}(r,n)$ such that $\mathcal{U}|_Z$ has a trivial quotient bundle, there is a section $s \in H^0(\mathrm{Gr}(r,n), \mathcal{Q})$ which vanishes on Z. Similarly we can prove the statement for $\mathcal{Q}$. □

(2.12) $T_{\mathrm{Gr}(r,n)}$ is $r(n-r)-n+1$ ample.

Note that in the case of projective space, $r=1$, the tangent bundle is 0-ample, that is ample. For the k-ampleness of the tangent bundle of a homogeneous space G/P, see [G].

3. k-ampleness of $\mathcal{E}$

Let $f: X \to \mathrm{Gr}(r,n)$ be a degree d branched covering of the Grassmannian $\mathrm{Gr}(r,n)$ where X is a connected complex projective manifold. Then the flatness of f gives the vetor bundle $f_*\mathcal{O}_X = \mathcal{O}_{\mathrm{Gr}(r,n)} \oplus \mathcal{F}$ where $\mathcal{F}$ is the kernel of the trace map Tr in the split exact sequence:

$$0 \longrightarrow \mathcal{O}_{\mathrm{Gr}(r,n)} \overset{\mathrm{Tr}}{\underset{\leftarrow}{\longrightarrow}} f_*\mathcal{O}_X \longrightarrow f_*\mathcal{O}_X/\mathcal{O}_{\mathrm{Gr}(r,n)} \longrightarrow 0.$$

The rank $d-1$ vector bundle $\mathcal{E} = \mathcal{F}^\vee$ associated with f can be considered as a variety, $\mathcal{E} = \mathbf{Spec}((\mathrm{Sym})_{\mathrm{Gr}(r,n)}\mathcal{F})$. As in [L1], f factors through $\mathcal{E}$: $X \hookrightarrow \mathcal{E} \to \mathrm{Gr}(r,n)$ where $X \hookrightarrow \mathcal{E}$ is a closed embedding and $\mathcal{E} \to \mathrm{Gr}(r,n)$ is the bundle projection.

By taking the annihilators, we have the isomorphism $\mathrm{Gr}(r,n) \to \mathrm{Gr}(n-r,n)$. Therefore we may assume, without loss of generality, that $n \geqq 2r$. We may also assume that $2 \leqq r \leqq n-2$, otherwise the Grassmannian is the projective space and we are in the case of [L1].

We need the following canonical bundle formula to prove the main propositions in this section.

(3.1) Let $\mathscr{F}_1$ and $\mathscr{F}_2$ be two vector bundles on X of rank f_1 and f_2. For $i = 1, 2$ and any $0 \leqq r_i \leqq f_i$, let $\pi_i : \mathrm{Gr}(r_i, \mathscr{F}_i) \to X$ be the Grassmannian bundle projections, and let $B = \mathrm{Gr}(r_1, \mathscr{F}_1) \times_X \mathrm{Gr}(r_2, \mathscr{F}_2) \xrightarrow{\pi} X$ be the fibred product.

$$\begin{array}{ccc} B & \xrightarrow{p_1} & \mathrm{Gr}(r_1, \mathscr{F}_1) \\ {\scriptstyle p_2}\downarrow & \searrow \pi & \downarrow {\scriptstyle \pi_1} \\ \mathrm{Gr}(r_2, \mathscr{F}_2) & \xrightarrow[\pi_2]{} & X. \end{array}$$

Then $K_B \cong \pi^* K_X \otimes (\det \pi^* \mathscr{F}_1)^{r_1} \otimes (\det \pi^* \mathscr{F}_2)^{r_2} \otimes p_1^* \xi_1^{-f_1} \otimes p_2^* \xi_2^{-f_2}$ where ξ_1 and ξ_2 are the tautological line bundles of $\mathrm{Gr}(r_1, \mathscr{F}_1)$ and $\mathrm{Gr}(r_2, \mathscr{F}_2)$, respectively.

(3.2) Proposition. *Let $\mathscr{E}$ be a vector bundle on* $\mathrm{Gr}(r, n)$ *associated with a branched covering* $f: X \to \mathrm{Gr}(r, n)$. *Then $\mathscr{E}$ is spanned on* $\mathrm{Gr}(r, n)$.

Proof. Let $\mathfrak{B} = \mathscr{Q}^{\oplus n-r}$. Since a subgrassmannian $\mathrm{Gr}(r, n-1)$ can be realized as the zero locus of a section of $\mathscr{Q}$, for any point $x = \Omega_{0\,1 \ldots r-1} \in \mathrm{Gr}(r, n)$, there is a Koszul complex associated with $\mathfrak{B}$:

$$0 \to \wedge^{r(n-r)} \mathfrak{B}^\vee \to \wedge^{r(n-r)-1} \mathfrak{B}^\vee \to \cdots \to \mathfrak{B}^\vee \to \mathcal{O}_{\mathrm{Gr}(r,n)} \to \mathcal{O}_{\mathrm{Gr}(r,n)}/\mathscr{I}_{x,\mathrm{Gr}(r,n)} \to 0.$$

We tensor the above complex with $\mathscr{E}$ to get the following Koszul complex:

$$(3.2.1) \quad 0 \to \mathscr{E} \otimes \wedge^{r(n-r)} \mathfrak{B}^\vee \to \mathscr{E} \otimes \wedge^{r(n-r)-1} \mathfrak{B}^\vee \to \cdots \to \mathscr{E} \otimes \mathfrak{B}^\vee \to \mathscr{E} \to \mathscr{E}|_x \to 0.$$

Let $G = \mathrm{Gr}(r, n)$ from now on. Now we will prove that the following cohomology groups vanish for $i = 1, \ldots, r(n-r)$:

$$H^i(G, \mathscr{E} \otimes \wedge^i \mathfrak{B}^\vee) \cong H^{r(n-r)-i}(G, \mathscr{F} \otimes \wedge^i \mathfrak{B} \otimes \mathcal{O}_G(-n))^\vee = 0$$

by Serre duality and $K_G = \mathcal{O}_G(-n)$. Then by looking at the hypercohomology of (3.2.1), the above vanishings will tell us that $H^0(\mathrm{Gr}(r, n), \mathscr{E})$ generates each stalk $\mathscr{E}|_x$ and so $\mathscr{E}$ will be spanned.

Since $f_* \mathcal{O}_X = \mathcal{O}_G \oplus \mathscr{F}$, to show the above vanishings, it suffices to prove

$$(3.2.2) \qquad H^i(X, f^*(\wedge^i \mathfrak{B}^\vee(n)) \otimes K_X) = 0$$

by the Leray spectral sequence and Serre duality. Note that

$$(3.2.3) \qquad \wedge^i \mathfrak{B}^\vee \cong \bigoplus_{1 \leqq J \leqq n-r} \left(\bigotimes_{1 \leqq i_j \leqq r,\ \sum_{1 \leqq j \leqq J} i_j = i} \wedge^{i_j} \mathscr{Q}^\vee \right).$$

So for (3.2.2), it is enough to show that for $i = 1, \ldots, r(n-r)$,

$$(3.2.4) \qquad H^i\left(X, f^*\left(\left(\bigotimes_{1 \leqq i_j \leqq r,\ \sum_{1 \leqq j \leqq J} i_j = i} \wedge^{i_j} \mathscr{Q}^\vee\right)(n)\right) \otimes K_X\right) = 0$$

where each J is in the above range.

Case $J=1$. Let $\mathscr{G}=\mathrm{Gr}(r-i, f^*\mathscr{Q})$. For $1 \leqq i \leqq r(n-r)$, $k \geqq 1$,

$$\begin{aligned} H^k(X, f^*(\wedge^i \mathscr{Q}^\vee(n)) \otimes K_X) &\cong H^k(X, f^*(\wedge^{r-i}\mathscr{Q}(n-1)) \otimes K_X) \\ &\cong H^k(\mathscr{G}, \pi^*(K_X \otimes f^*(\mathscr{O}_G(n-1))) \otimes \xi_{r-i, f^*Q}) \\ &\cong H^k(\mathscr{G}, K_{\mathscr{G}} \otimes \xi_{r-i, f^*Q}^{r+1} \otimes \pi^* f^*(\mathscr{O}_G(n-1-r+i))) \\ &= 0 \end{aligned}$$

by (2.4), the canonical bundle formula (3.1) where $\pi: \mathscr{G} \to X$ is the bundle projection, and Nakai-Moishezon criterion ([H], p. 434, Theorem 5.1) and Kodaira vanishing theorem. In particular, for $k=i$ we have the desired vanishing (3.2.4).

Case $J=2$. We want to show that $H^i(X, f^*(\wedge^a \mathscr{Q}^\vee \otimes \wedge^b \mathscr{Q}^\vee(n)) \otimes K_X)=0$ where $a+b=i$, $1 \leqq a \leqq b \leqq r$ and $n-r \geqq 2$.

Now consider the following Koszul complex associated with the universal exact sequence (for all the Koszul complexes throughout this section, see [Sc]; which deals with the more general *Eagon-Northcott* complex):

$$(3.2.5) \qquad 0 \to \wedge^a \mathscr{Q}^\vee \to \mathscr{O}_G^{\oplus\binom{n}{a}} \to \mathscr{O}_G^{\oplus\binom{n}{a-1}} \otimes \mathscr{U} \to \cdots \to \mathscr{O}_G^{\oplus\binom{n}{1}} \otimes \mathscr{U}^{(a-1)} \to \mathscr{U}^{(a)} \to 0.$$

Pull back by f and tensor with $f^*(\wedge^b \mathscr{Q}^\vee(n)) \otimes K_X$ to get:

$$\begin{aligned} (3.2.6) \quad 0 &\to f^*(\wedge^a \mathscr{Q}^\vee \otimes \wedge^b \mathscr{Q}^\vee(n)) \otimes K_X \to [f^*(\wedge^b \mathscr{Q}^\vee(n)) \otimes K_X]^{\oplus\binom{n}{a}} \to \cdots \\ &\to [f^*(\mathscr{U}^{(a-1)} \otimes \wedge^b \mathscr{Q}^\vee(n)) \otimes K_X]^{\oplus\binom{n}{1}} \to f^*(\mathscr{U}^{(a)} \otimes \wedge^b \mathscr{Q}^\vee(n)) \otimes K_X \to 0. \end{aligned}$$

Again by using the hypercohomology of the above complex, it is enough to prove the following:

$$H^{i-l}(X, f^*(\mathscr{U}^{(l)} \otimes \wedge^b \mathscr{Q}^\vee(n)) \otimes K_X) \cong H^{i-l}(X, f^*(\mathscr{U}^{(l)} \otimes \wedge^{r-b}\mathscr{Q}(n-1)) \otimes K_X) = 0$$

for $0 \leqq l \leqq a$. By applying (3.1) to the following diagram,

$$\begin{array}{ccc} B & \xrightarrow{p_1} & \mathrm{Gr}(r-b, f^*\mathscr{Q}) =: \mathscr{G} \\ {\scriptstyle p_2}\downarrow & \searrow{\scriptstyle \pi} & \downarrow{\scriptstyle \pi_1} \\ \mathbb{P}(f^*\mathscr{U}) & \xrightarrow[\pi_2]{} & X, \end{array}$$

we get for $j \geqq 1$,

$$\begin{aligned} &H^j(X, f^*(\mathscr{U}^{(l)} \otimes \wedge^{r-b}\mathscr{Q}(n-1)) \otimes K_X) \\ &\cong H^j(\mathscr{G}, \pi_1^*(K_X \otimes f^*(\mathscr{U}^{(l)}(n-1))) \otimes \xi_1) \\ &\cong H^j(B, \pi^*(K_X \otimes f^*(\mathscr{O}_G(n-1))) \otimes \xi^l \otimes p_1^*\xi_1) \\ &\cong H^j(B, K_B \otimes p_1^*\xi_1^{r+1} \otimes p_2^*\xi_2^{n-r} \otimes \xi^l \otimes \pi^*(f^*\mathscr{O}_G(n-r+b-2))) \\ &= 0 \end{aligned}$$

by noting that $B = \mathbb{P}(\pi_1^* f^* \mathscr{U})$ and applying (2.1), (2.4), Nakai-Moishezon criterion and Kodaira vanishing theorem (note that f is finite) where ξ_1 (resp. ξ_2) is the tautological line bundle via π_1 (resp. π_2) and ξ is the one via p_1.

Case $J = 3$. By recalling (3.2.4), we want to show

$$H^i\big(X, f^*(\wedge^a \mathscr{Q}^\vee \otimes \wedge^b \mathscr{Q}^\vee \otimes \wedge^c \mathscr{Q}^\vee(n)) \otimes K_X\big) = 0$$

for $a + b + c = i$, $1 \leqq a \leqq b \leqq c \leqq r$ and $n - r \geqq 3$. Similar to the case of $J = 2$, after tensoring (3.2.6) with $f^*(\wedge^c \mathscr{Q}^\vee)$, it is enough to prove

$$H^{i-l}\big(X, f^*(\mathscr{U}^{(l)} \otimes \wedge^b \mathscr{Q}^\vee \otimes \wedge^c \mathscr{Q}^\vee(n)) \otimes K_X\big) = 0 \quad \text{for} \quad 0 \leqq l \leqq a.$$

For the above desired vanishings, by replacing a by b in (3.2.5), pulling back by f and tensoring with $f^*(\mathscr{U}^{(l)} \otimes \wedge^c \mathscr{Q}^\vee(n)) \otimes K_X$, for $0 \leqq l \leqq a$ and $0 \leqq k \leqq b$, we want to show the vanishings of

$$H^{i-l-k}\big(X, f^*(\mathscr{U}^{(l)} \otimes \mathscr{U}^{(k)} \otimes \wedge^c \mathscr{Q}^\vee(n)) \otimes K_X\big)$$

which is a subgroup of $H^{i-l-k}\big(X, f^*((\mathscr{U} \oplus \mathscr{U})^{(l+k)} \otimes \wedge^{r-c} \mathscr{Q}(n-1)) \otimes K_X\big)$. Using, as in the case of $J = 2$, the following diagram

$$\begin{array}{ccc} B & \xrightarrow{p_1} & \mathrm{Gr}(r-c, f^*\mathscr{Q}) =: \mathscr{G} \\ {\scriptstyle p_2}\downarrow & \searrow {\scriptstyle \pi} & \downarrow {\scriptstyle \pi_1} \\ \mathbb{P}(f^*(\mathscr{U} \oplus \mathscr{U})) & \xrightarrow[\pi_2]{} & X, \end{array}$$

we see that the cohomology group above is identical with

$$\begin{aligned} & H^{i-l-k}\big(\mathscr{G}, \pi_1^* f^*((\mathscr{U} \oplus \mathscr{U})^{(l+k)}(n-1)) \otimes \pi_1^* K_X \otimes \xi_1\big) \\ & \cong H^{i-l-k}\big(B, K_B \otimes p_1^* \xi_1^{r+1} \otimes p_2^* \xi_2^{2n-2r} \otimes \xi^{l+k} \otimes \pi^* f^* \mathscr{O}_G(n-r+c-3)\big) \\ & = 0 \end{aligned}$$

where ξ_1 (resp. ξ_2) is the tautological line bundle via π_1 (resp. π_2) and ξ is the one via p_1.

For the remaining cases, $J \geqq 4$, we complete the proof of the vanishings (3.2.2) as in the previous cases. □

(3.3) Remark. Let Q_n be an n-dimensional smooth Quadric in $\mathbb{P}^{n+1}$ and $\mathscr{E}$ be the vector bundle associated with a covering map of Q_n. Then we can directly show that $\mathscr{E}$ is 0-regular (cf. [M], p. 99), hence $\mathscr{E}$ is spanned.

Now for (3.4) below, we consider a branched covering $f: X \to \mathrm{Gr}(2, n)$ and the vector bundle $\mathscr{E}$ associated with f.

(3.4) Proposition. *On* $\mathrm{Gr}(2, n)$, $n \geqq 4$, $\mathscr{E}$ *is* $(n-2)$*-ample.*

Proof. It is enough to show that $\mathscr{E} \otimes \mathscr{Q}^\vee$ is spanned, for then we infer the proposition by the natural surjective map $(\mathscr{E} \otimes \mathscr{Q}^\vee) \otimes \mathscr{Q} \to \mathscr{E} \to 0$ and the $(n-2)$-ampleness of $\mathscr{Q}$ by (2.11) (cf. [S1], Corollary 1.10).

Tensor (3.2.1) with $\mathscr{Q}^\vee$ of $\mathrm{Gr}(2, n)$ to get:

$$0 \to \mathscr{E} \otimes \wedge^{2(n-2)} \mathfrak{B}^\vee \otimes \mathscr{Q}^\vee \to \mathscr{E} \otimes \wedge^{2(n-2)-1} \mathfrak{B}^\vee \otimes \mathscr{Q}^\vee \to \cdots \to \mathscr{E} \otimes \mathscr{Q}^\vee \to \mathscr{E} \otimes \mathscr{Q}^\vee|_x \to 0.$$

Let $G = \mathrm{Gr}(2, n)$. By looking at the hypercohomology of the above complex and by an argument similar to the one in the proof of (3.2), it suffices to show that for $i = 1, \ldots, 2(n-2)$, the following cohomology groups vanish:

$$(*) \qquad H^i\left(G, \mathscr{E} \otimes \left(\bigotimes_{1 \leqq i_j \leqq 2 \ \sum_{1 \leqq j \leqq J} i_j = i} \wedge^{i_j} \mathscr{Q}^\vee\right) \otimes \mathscr{Q}^\vee\right)$$

where J is an integer as in (3.2.3). Note that $(*)$ is a subspace of

$$(\#) \qquad H^i\left(X, f^*\left(\left(\bigotimes_{1 \leqq i_j \leqq 2 \ \sum_{1 \leqq j \leqq J} i_j = i} \wedge^{i_j} \mathscr{Q}^\vee\right) \otimes \mathscr{Q}^\vee(n)\right) \otimes K_X\right)$$

by Leray spectral sequence and Serre duality. We will show that $(*)$ vanishes in the case of $J = n-2$ and $(\#)$ vanishes otherwise. We need to keep in mind $\mathscr{Q}^\vee(1) = \mathscr{Q}$.

Case $J \neq n-2$. If $i = 1$, then

$$H^1(X, f^*(\mathscr{Q}^\vee \otimes \mathscr{Q}^\vee(n)) \otimes K_X) \cong H^1(X, f^*(\mathscr{Q}^{(2)}(n-2)) \otimes K_X) = 0$$

by the fact that $\mathscr{Q} \otimes \mathscr{Q} = \mathscr{Q}^{(2)} \oplus \det \mathscr{Q}$ and Griffiths vanishing theorem (2.5).

If $i = 2$, we need to show the following cohomology groups vanish:

(i) $H^2(X, f^*(\mathscr{Q}^\vee \otimes \mathscr{Q}^\vee \otimes \mathscr{Q}^\vee(n-1)) \otimes K_X)$ if $J = 2$ (which implies $n \geqq 5$),

(ii) $H^2(X, f^*(\mathscr{Q}^\vee(n-1)) \otimes K_X)$ otherwise.

Consider the following exact sequence by pulling back the universal exact sequence (2.7) by f and tensoring with $f^*(\mathcal{O}_G(n-1)) \otimes K_X$:

$$(3.4.1) \qquad 0 \to f^*(\mathscr{Q}^\vee(n-1)) \otimes K_X \to [f^*(\mathcal{O}_G(n-1)) \otimes K_X]^{\oplus n} \to f^*(\mathscr{U}(n-2)) \otimes K_X \to 0.$$

Since, for $k \geqq 1$, the k-th cohomology of the middle vector bundle vanishes, (ii) is isomorphic to $H^1(X, f^*(\mathscr{U}(n-1)) \otimes K_X)$ and it vanishes by (2.5). Repeating this argument, we see that the cohomology group in (i) is isomorphic to

$$\begin{aligned} &H^1(X, f^*(\mathscr{Q}^\vee \otimes \mathscr{Q}^\vee \otimes \mathscr{U}(n)) \otimes K_X) \\ &\cong H^1(X, f^*(\mathscr{U}(n-1)) \otimes K_X) \oplus H^1(X, f^*(\mathscr{Q}^{(2)} \otimes \mathscr{U}(n-2)) \otimes K_X) \\ &= 0 \end{aligned}$$

by applying the canonical bundle formula (3.1), and Nakai-Moishezon criterion and Kodaira (or Griffiths) vanishing theorem (cf. the proof of (3.2)).

If $i = 3$, we need to show the following groups vanish:

(i)′ $H^3(X, f^*(\mathcal{Q}^\vee \otimes \mathcal{Q}^\vee \otimes \mathcal{Q}^\vee \otimes \mathcal{Q}^\vee(n-1)) \otimes K_X)$ if $J = 3$ (which implies $n \geqq 6$),

(ii)′ $H^3(X, f^*(\mathcal{Q}^\vee \otimes \mathcal{Q}^\vee(n-1)) \otimes K_X)$ otherwise.

By tensoring (3.4.1) with $f^*\mathcal{Q}^\vee$ and by iterating the similar argument as in the case of $i = 2$, (ii)′ turns out to be

$$\begin{aligned} &H^1(X, f^*(\mathcal{U} \otimes \mathcal{U}(n-3)) \otimes K_X) \\ &\cong H^1(X, f^*(\mathcal{U}^{(2)}(n-3)) \otimes K_X) \oplus H^1(X, f^*(\wedge^2\mathcal{U}(n-3)) \otimes K_X) \\ &= 0 \end{aligned}$$

by (3.1) and (2.5). In a similar way we can show that (i)′ vanishes.

If $i \geqq 4$, by iterating this process, we complete the vanishing of (#).

Case $J = n - 2$. The desired vanishing is

$$H^{n-2}(G, \mathscr{E} \otimes \mathcal{Q}^{\vee \otimes n-2} \otimes \mathcal{Q}^\vee) \cong H^{n-2}(G, \mathscr{F} \otimes \mathcal{Q}^{\otimes n-2} \otimes \mathcal{Q}(-n)) = 0.$$

Since $f_* \mathcal{O}_X = \mathcal{O}_G \oplus \mathscr{F}$, we will show that (a) and (b) below are isomorphic:

(a) $$\begin{aligned} H^{n-2}(G, \mathcal{Q}^{\otimes n-1}(-n)) &\cong H^{n-2}(G, \mathcal{Q}^{\vee \otimes n-1}(n) \otimes K_G) \\ &\cong H^{n-2}(G, \mathcal{Q}^{\otimes n-1}(1) \otimes K_G), \end{aligned}$$

(b) $$H^{n-2}(X, f^*(\mathcal{Q}^{\vee \otimes n-1}(n)) \otimes K_X) \cong H^{n-2}(X, f^*(\mathcal{Q}^{\otimes n-1}(1)) \otimes K_X),$$

that will give us the desired vanishing.

Note that (a) is isomorphic to

(a′) $$H^{n-2}(G, \mathcal{Q}^{\otimes n-3} \otimes \mathcal{Q}^{(2)}(1) \otimes K_G)$$

since $\mathcal{Q} \otimes \mathcal{Q} = \mathcal{Q}^{(2)} \oplus \det \mathcal{Q}$ and $H^{n-2}(G, \mathcal{Q}^{\otimes n-3} \otimes (\det \mathcal{Q})(1) \otimes K_G) = 0$ by (2.5). Similarly, (b) is isomorphic to

(b′) $$H^{n-2}(X, f^*(\mathcal{Q}^{\otimes n-3} \otimes \mathcal{Q}^{(2)}(1)) \otimes K_X).$$

We will show that (a′) and (b′), both are isomorphic to $\mathbb{C}^n$.

Consider the following Koszul complex:

$$0 \to \wedge^2 \mathcal{Q} \otimes \mathcal{Q} \to \mathcal{Q} \otimes \mathcal{Q}^{(2)} \to \mathcal{Q}^{(3)} \to 0 .$$

Pull back by f and tensor with $f^*(\mathcal{Q}^{\otimes n-4}(1)) \otimes K_X$ to get:

$$0 \to f^*(\mathcal{Q}^{\otimes n-3}(2)) \otimes K_X \to f^*(\mathcal{Q}^{\otimes n-3} \otimes \mathcal{Q}^{(2)}(1)) \otimes K_X \to f^*(\mathcal{Q}^{\otimes n-4} \otimes \mathcal{Q}^{(3)}(1)) \otimes K_X \to 0.$$

Since $H^k(X, f^*(\mathcal{Q}^{\otimes n-3}(2)) \otimes K_X) = 0$ for $k \geqq 1$ by (2.5), so in particular,

$$H^{n-2}(X, f^*(\mathcal{Q}^{\otimes n-3} \otimes \mathcal{Q}^{(2)}(1)) \otimes K_X) \cong H^{n-2}(X, f^*(\mathcal{Q}^{\otimes n-4} \otimes \mathcal{Q}^{(3)}(1)) \otimes K_X).$$

Let us assume $n \geqq 5$. Then by an argument similar to the previous one, using the following Koszul complex:

$$0 \to \wedge^2 \mathcal{Q} \otimes \mathcal{Q}^{(2)} \to \mathcal{Q} \otimes \mathcal{Q}^{(3)} \to \mathcal{Q}^{(4)} \to 0$$

which gives

$$0 \to f^*(\mathcal{Q}^{\otimes n-3} \otimes \mathcal{Q}^{(2)}(2)) \otimes K_X \to f^*(\mathcal{Q}^{\otimes n-4} \otimes \mathcal{Q}^{(3)}(1)) \otimes K_X$$
$$\to f^*(\mathcal{Q}^{\otimes n-5} \otimes \mathcal{Q}^{(4)}(1)) \otimes K_X \to 0,$$

we have $H^{n-2}(X, f^*(\mathcal{Q}^{\otimes n-4} \otimes \mathcal{Q}^{(3)}(1)) \otimes K_X) \cong H^{n-2}(X, f^*(\mathcal{Q}^{\otimes n-5} \otimes \mathcal{Q}^{(4)}(1)) \otimes K_X)$. By iterating this process, we see that (b′) is isomorphic to

(b″) $$H^{n-2}(X, f^*(\mathcal{Q}^{(n-1)}(1)) \otimes K_X).$$

Now consider the Koszul complex associated with the universal exact sequence of $\mathrm{Gr}(2, n)$ (cf. (3.2.5)):

$$0 \to f^*(\mathcal{O}_G^{\oplus\binom{n}{n-1}}(1)) \otimes K_X \to f^*(\mathcal{O}_G^{\oplus\binom{n}{n-2}} \otimes \mathcal{Q}(1)) \otimes K_X \to \cdots$$
$$\to f^*(\mathcal{O}_G^{\oplus\binom{n}{1}} \otimes \mathcal{Q}^{(n-2)}(1)) \otimes K_X \to f^*(\mathcal{Q}^{(n-1)}(1)) \otimes K_X \to 0.$$

Then since

(3.4.2) $$H^k(X, f^*(\mathcal{Q}^{(l)}(1)) \otimes K_X) = 0 \quad \text{for} \quad l = 0, 1, \ldots, n-3 \quad \text{and} \quad k \geqq l+1$$

by Le Potier vanishing theorem (2.6), we have

$$H^{n-2}(X, f^*(\mathcal{Q}^{(n-1)}(1)) \otimes K_X) \cong [H^{n-2}(X, f^*(\mathcal{Q}^{(n-2)}(1)) \otimes K_X)]^{\oplus n}.$$

Consider the following Koszul complex associated with the universal exact sequence:

$$0 \to f^*\mathcal{O}_G \otimes K_X \to f^*(\mathcal{O}_G^{\oplus\binom{n}{n-2}}(1)) \otimes K_X \to \cdots$$
$$\to f^*(\mathcal{O}_G^{\oplus\binom{n}{1}} \otimes \mathcal{Q}^{(n-3)}(1)) \otimes K_X \to f^*(\mathcal{Q}^{(n-2)}(1)) \otimes K_X \to 0.$$

By looking at the hypercohomology of the above complex and at the vanishings (3.4.2) we see that

$$H^{n-2}(X, f^*(\mathcal{Q}^{(n-2)}(1)) \otimes K_X) \cong H^{2(n-2)}(X, f^*\mathcal{O}_G \otimes K_X) \cong H^0(X, f^*\mathcal{O}_G) = \mathbb{C}.$$

Hence (b′) is isomorphic to $\mathbb{C}^n$. Using the same complexes (without pulling back by f) we see that (a′) is also isomorphic to $\mathbb{C}^n$, hence we have all the desired vanishings. Therefore $\mathscr{E} \otimes \mathscr{Q}^\vee$ is spanned on $\mathrm{Gr}(2, n)$. □

(3.5) Proposition. *On* $\mathrm{Gr}(r, n)$ *for* $n \geqq 2r$, $\mathscr{E}$ *is* $r(n-r-1)$*-ample.*

Proof. Since $\mathscr{E}$ is spanned, $\xi_{\mathscr{E}}$ defines a morphism $\mathbb{P}(\mathscr{E}) \xrightarrow{\phi} \mathbb{P}^N$; let F be an irreducible component of maximal dimension of $\phi^{-1}(p)$, where $\phi^{-1}(p)$ is a fibre of ϕ of maximal dimension. Since F cannot meet the fibres of $\mathbb{P}(\mathscr{E}) \xrightarrow{\pi} \mathrm{Gr}(r, n)$, i.e. projective spaces $\mathbb{P}^{d-2}$, along a positive dimensional scheme, so $F \xrightarrow{\pi} \pi(F)$ is finite. We will consider the codimension of $V = \pi(F) \subset \mathrm{Gr}(r, n)$. We need to keep in mind that the preimage by f of a generic subgrassmannian is connected by Faltings' [F], Satz 5, Korollar, or Sommese-Van de Ven's connectedness result [SV], Theorem 2.2 (note that we can consider any generic subgrassmannian to be sitting insde of a higher dimensional generic subgrassmannian satisfying the bound for the connectedness result, which makes the preimage of any generic subgrassmannian connected), and it is smooth for a generic subgrassmannian by [H], III, Theorem 10.8.

We claim that the codimension of V in $\mathrm{Gr}(r, n)$ is at least r; this would imply directly (3.5). Suppose that codim $V = r - 1$ then

$$V \sim \sum \alpha_{a_1 \ldots a_r} \Omega_{a_1 \ldots a_r}$$

where $\alpha_{\underline{a}}$ is an integer, $0 \leqq a_1 \leqq a_2 \leqq \cdots \leqq a_r \leqq n-1$ and $\sum_{i=1}^{r} a_i = r(n-r) + \frac{1}{2}(r-2)(r-1)$. Fix generic $\mathrm{Gr}(r, r+2)$ $(\sim \Omega_{23\ldots r+1})$ and $\mathrm{Gr}(n-r, n-r+2)$ $(\sim \Omega_{01\ldots r-3\,n-2\,n-1})$. By standard Schubert calculus [HP], p. 327, we can check that V intersects either the generic $\mathrm{Gr}(r+2)$ or the generic $\mathrm{Gr}(n-r, n-r+2)$ in $\mathrm{Gr}(r, n)$.

If V meets $\mathrm{Gr}(r, r+2)$ then $\dim V \cap \mathrm{Gr}(r, r+2) \geqq r+1$. But this is contradiction since $\mathscr{E}|_{\mathrm{Gr}(r,r+2)}$ is r-ample for a generic $\mathrm{Gr}(r, r+2)$ by Proposition (3.4). In the same way, the other case gives a contradiction.

If codim $V < r - 1$ then the same argument gives a contradiction as above. □

(3.6) Remark. If $n < 2r$ then by using the natural isomorphism $\mathrm{Gr}(r, n) \to \mathrm{Gr}(n-r, n)$ we derive that $\mathscr{E}$ is $(r-1)$ $(n-r)$-ample.

(3.7) Remark. k-ampleness of $\mathscr{E}$ on $\mathrm{Gr}(r, n)$ in (3.5) is not sharp compared with (3.4). We want to leave this as a conjecture:

Conjecture. $\mathscr{E}$ is $r(n-r) - \max\{r, n-r\}$-ample.

Note that the conjecture is consistent with Lazarsfeld's result for the coverings of the projective spaces.

4. Barth-Lefschetz type theorem for coverings of Grassmannians

The goal of this last section is to show (4.3) and (4.4) using [L1], Theorem 2.1, and the following Sommese's Lefschetz type theorem.

(4.1) *Let E be a k-ample vector bundle of rank e on a connected projective manifold X of complex dimension n, then the map obtained by cupping with the top Chern class of $c_e(E)$:*

$$H^a(X, \mathbb{C}) \xrightarrow{\bigcup c_e(E)} H^a(X, \mathbb{C})$$

is injective if $a \leqq n - e - k$ and surjective if $a \geqq n - e + k$.

For the proof, see [S1], Proposition 1.17. (4.1) is a generalization of Kleiman, Bloch and Gieseker's result for ample vector bundles (see [K] and [BG]).

(4.2) Theorem. *Let Y be a connected complex projective manifold of dimension n and $E \xrightarrow{\pi} Y$ be a k-ample vector bundle of rank e. Suppose X is a connected complex projective manifold of dimension n embedded in the total space of E. Let $f: X \hookrightarrow E \to Y$ be the composition. Then $f^*: H^i(Y, \mathbb{C}) \to H^i(X, \mathbb{C})$ is an isomorphism for $i \leqq n - e - k$.*

Proof. Consider the diagram:

$$\begin{array}{ccc} X & \overset{j}{\hookrightarrow} & \mathbb{P}(E^\vee \oplus 1) \\ & {\scriptstyle f}\searrow \quad \swarrow{\scriptstyle \pi} & \\ & Y & \end{array}$$

where j is the composition $X \hookrightarrow E \subseteqq \mathbb{P}(E^\vee \oplus 1)$. As in the proof of Lazarsfeld [L1], Sec. 2, we see that $j^*(\xi) = 0$ and $j^*(\eta_X) = (\deg f)\, c_e(f^*E)$ where $\xi = c_1(\xi)$ is the first Chern class of ξ, the tautological line bundle of $\mathbb{P}(E^\vee \oplus 1)$, and $\eta_X \in H^{2e}(\mathbb{P}(E^\vee \oplus 1), \mathbb{C})$ is the cohomology class defined by X. Follow the same steps as in [L1] by replacing the ampleness of E by its k-ampleness. Then to conclude, it is enough to show that $f^*: H^l(Y, \mathbb{C}) \to H^l(X, \mathbb{C})$ is surjective for $l \geqq n + e + k$. We have

$$\begin{array}{ccccc} H^{l-2e}(X, \mathbb{C}) & \xrightarrow{j_*} & H^l(\mathbb{P}(E^\vee \oplus 1), \mathbb{C}) & \xleftarrow{\pi^*} & H^l(Y, \mathbb{C}) \\ \bigcup j^*(\eta_X) \searrow & & \downarrow j^* & & \swarrow f^* \\ & & H^l(X, \mathbb{C}) & & \end{array}$$

where j_* is the Gysin map. $\bigcup j^*(\eta_X)$ is surjective for $l \geqq n + e + k$ by (5.1). So is j^* in the same range. Since $j^*(\xi) = 0$ and $H^*(\mathbb{P}(E^\vee \oplus 1), \mathbb{C})$ is a $H^*(Y, \mathbb{C})$-algebra generated by ξ, f^* is surjective for $l \geqq n + e + k$. □

Using (3.4), (3.5) and (4.2), the following Barth-Lefschetz-Lazarsfeld type theorem is obtained:

(4.3) Corollary. *If $f: X \to \mathrm{Gr}(r, n)$ is a finite surjective morphism of degree d, then $f^*: H^i(\mathrm{Gr}(r, n), \mathbb{C}) \to H^i(X, \mathbb{C})$ is an isomorphism for $i \leqq r - d + 1$. If $r = 2$, then f^* is an isomorphism for $i \leqq n - d - 1$.*

Note. If $n \leqq 2r$, then since $\mathscr{E}$ is actually $(r-1)(n-r)$-ample (see (3.6)) we can write $n-r-d+1$ as the bound for cohomology isomorphisms in (4.3).

(4.4) Proposition. *If X is a connected complex projective manifold, $f: X \to \mathrm{Gr}(r, n)$ is a finite surjective morphism of degree d and $n \geqq 2r$, then for any fixed point $x \in X$, $f_*: \pi_i(X, x) \to \pi_i(\mathrm{Gr}(r, n), f(x))$ is an isomorphism for $i \leqq r-d+1$. If $f: X \to \mathrm{Gr}(2, n)$ is a double covering, then $f_*: \pi_i(X, x) \to \pi_i(\mathrm{Gr}(2, n), f(x))$ is an isomorphism for $i \leqq n-3$.*

Proof. We have

$$\begin{array}{ccccc} p^*\mathscr{E}|_X & \to & p^*\mathscr{E} & \to & \mathscr{E} \\ \downarrow & & \downarrow & & \downarrow p \\ X & \underset{i}{\hookrightarrow} & \mathscr{E} & \underset{p}{\to} & \mathrm{Gr}(r, n) \end{array}$$

where f factors through the embedding i and the bundle projection p. Note that since $\mathscr{E}$ is spanned, so are $p^*\mathscr{E}$ and $p^*\mathscr{E}|_X$. By Bertini's theorem for vector bundles, we may suppose that $X \cap \mathrm{Gr}(r, n) \subset X$, the zero locus of s is non-singular for $s \in \Gamma(X, p^*\mathscr{E}|_X)$.

To deduce the desired isomorphisms of homotopy groups, we will use the following Barth-Lefschetz theorems by Sommese and Sommese-Van de Ven.

(4.4.1) *If B is a submanifold of the Grassmannian $\mathrm{Gr}(r, n)$, then for $x \in B$*

$$\pi_j(\mathrm{Gr}(r, n), B, x) = 0 \quad \text{for} \quad j \leqq 2\dim B - 2r(n-r) + n.$$

(4.4.1) is a corollary of Sommese's homotopy results for homogeneous complex manifolds [S2], Proposition 3.4, which revolve round the theme of the k-ampleness of the tangent bundle of a homogeneous space.

(4.4.2) *Let E be a spanned k-ample vector bundle on a compact complex manifold W and Z be the zero locus of a holomorphic section of E. Fix $x \in X$. Then*

$$\pi_j(W, Z, x) = 0 \quad \text{for} \quad j \leqq \dim W - \operatorname{rank} E - k.$$

(See [S3] or [SV], Remark 3.2.1.)

Now fix $x \in Z = X \cap \mathrm{Gr}(r, n) = s^{-1}(0)$ and consider the homotopy sequences of the pairs (X, Z) and $(\mathrm{Gr}(r, n), Z)$:

$$\begin{array}{ccccc} \pi_i(Z, x) & \to & \pi_i(X, x) & \to & \pi_i(X, Z, x) \\ \cong\downarrow & & \downarrow f_* & & \downarrow f_* \\ \pi_i(Z, x) & \to & \pi_i(\mathrm{Gr}(r, n), x) & \to & \pi_i(\mathrm{Gr}(r, n), Z, x). \end{array}$$

By (4.4.1) and (4.4.2), $\pi_j(\mathrm{Gr}(r, n), Z, x) = 0$ for $j \leqq n-2d+2$ and $\pi_i(X, Z, x) = 0$ for $i \leqq r-d+1$ since $\mathscr{E}$ is $r(n-r-1)$-ample. Note that the first above bound for relative homotopy groups is strictly bigger than the second one. So from the diagram of exact sequence for homotopy groups, we conclude the first part of (4.4). In particular, if $r = 2$

and $d = 2$, we get the second part of (4.4) by looking at the above diagram and using $(n-2)$-ampleness of $\mathscr{E}$. □

(4.5) Remark. Corollary (4.3) can be also obtained as a corollary of Proposition (4.4) if $r \geqq 3$.

References

[B] *W. Barth*, Transplanting cohomology classes in complex-projective space, Amer. J. Math. **92** (1970), 951–967.

[BG] *S. Bloch* and *D. Gieseker*, The positivity of the Chern classes of an ample vector bundle, Invent. Math. **12** (1971), 112–117.

[F] *G. Faltings*, Formale Geometrie und homogene Räume, Invent. Math. **64** (1981), 123–165.

[FL] *W. Fulton* and *R. Lazarsfeld*, Connectivity and its applications in algebraic geometry, Proc. of the first midwest algebraic geometry conference, Springer Lect. Notes Math. **862** (1981), 26–92.

[G] *N. Goldstein*, Ampleness and connectedness in complex G/P, Trans. Amer. Math. Soc. **274** (1982), 361–373.

[H] *R. Hartshorne*, Algebraic geometry, Grad. Texts Math. **52**, Berlin–Heidelberg–New York 1977.

[HP] *W. V. D. Hodge* and *D. Pedoe*, Methods of algebraic geometry, Vol. **2**, Cambridge University Press, 1952.

[K] *S. L. Kleiman*, Ample vector bundles on algebraic surfaces, Proc. Amer. Math. Soc. **21** (1969), 673–676.

[L1] *R. Lazarsfeld*, A Barth-type theorem for branched coverings of projective space, Math. Ann. **249** (1980), 153–162.

[L2] *R. Lazarsfeld*, Thesis, Brown University, 1980.

[M] *D. Mumford*, Lectures on curves on an algebraic surface, Ann. Math. Stud. **59**, Princeton University Press, 1966.

[Sc] *F.-O. Schreyer*, Syzygies of canonical curves and special linear series, Math. Ann. **275** (1986), 105–137.

[SS] *B. Shiffman* and *A. J. Sommese*, Vanishing theorems on complex manifolds, Progr. Math. **56**, Birkhäuser 1985.

[S1] *A. J. Sommese*, Submanifolds of abelian varieties, Math. Ann. **233** (1978), 229–256.

[S2] *A. J. Sommese*, Complex subspaces of homogeneous complex manifolds II–Homotopy results, Nagoya Math. J. **86** (1982), 101–129.

[S3] *A. J. Sommese*, A convexity theorem, Proc. Symp. Pure Math. **40** (1983), 497–505.

[SV] *A. J. Sommese* and *A. Van de Ven*, Homotopy groups of pullbacks of varieties, Nagoya Math. J. **102** (1986), 79–90.

Department of Mathematics, 370 CCMB University of Notre Dame, Notre Dame, IN 46556-5683, USA
e-mail: Meeyoung.Kim.2@nd.edu

Eingegangen 7. Dezember 1994, in revidierter Fassung 23. Juni 1995

J. reine angew. Math. **470** (1996), 123—142

Journal für die reine und angewandte Mathematik

The Jeffrey-Kirwan localization theorem and residue operations in equivariant cohomology

By *Victor Guillemin*[1]) and *Jaap Kalkman*[2]) at Cambridge

1. Introduction

Let G be a compact connected Lie group and M a compact Hamiltonian G-space with symplectic form σ and moment map, $\mu : M \to \mathfrak{g}^*$. If zero is a regular value of μ, the action of G on the level set, $Z = \mu^{-1}(0)$, is locally free and the reduced space

$$M_{\text{red}} = Z/G$$

is a compact symplectic orbifold. Let $\iota : Z \to M$ be the inclusion map and $p : Z \to M_{\text{red}}$ the projection of Z onto M_{red}. p induces a bijective map

$$(1.1) \qquad p^{\#} : H^*(M_{\text{red}}, \mathbb{C}) \to H^*_G(Z, \mathbb{C})$$

(from ordinary cohomology into equivariant cohomology) and by composing the inverse of this map with ι^* one gets a surjective map

$$r : H^*_G(M, \mathbb{C}) \to H^*(M_{\text{red}}, \mathbb{C}) .$$

We will refer to this from now on as the *Kirwan* map. (Its surjectivity was proved by Kirwan in [Ki].)

Now let $d = \dim M_{\text{red}}$, and let $\alpha \in \Omega_G(M)$ be an equivariantly closed d-form. The localization theorem which we will be concerned with in this paper is a formula for evaluating the integral

$$(1.2) \qquad \int_{M_{\text{red}}} r(\alpha)$$

[1]) Supported by NSF grant DMS 890771.
[2]) Supported by a NATO science fellowship.

in terms of the restriction of α to the fixed point sets of certain one-dimensional subgroups of G. For the case of the circle this formula can be found in [Ka], and the main purpose of this paper is to reformulate the results of [Ka] in a way that makes sense for arbitrary groups. First, however, we will try to summarize in a few words the content of the Jeffrey-Kirwan theorem: This theorem expresses the integral

$$\int_{M_{\mathrm{red}}} r(\alpha) \exp i\sigma_{\mathrm{red}} \tag{1.3}$$

in terms of the restriction of α to the fixed point set of the Cartan subgroup, T, of G. If the fixed points of T are isolated, it says that the integral (1.3) is equal to the sum of residues at the fixed points:

$$\sum_p{}' (\mathrm{res})_{x=0} \frac{e^{i\mu(p)(x)} \pi(x) \iota_p^* \alpha(x)}{\prod \lambda_i(p)(x)} \tag{1.4}$$

the $\lambda_i(p)$'s being the weights of the isotropy representation of T at p and $\pi(x) = \prod \alpha_i(x)$ being the product of the roots of G [3]). The "prime" next to the summation sign means that the sum is over a subset of the fixed point set which we will describe more carefully below (see § 3), and there is also some ambiguity (about which we will have more to say below) in the definition of "$(\mathrm{res})_{x=0}$". For non-isolated fixed points their formula is more complicated, but in particular it says that the integral (1.3) is equal to:

$$\int r_T[(\alpha \exp i\tilde{\sigma})^{\#} \pi(x)]$$

(integration being over the T-reduced space, $\mu_T^{-1}(0)/T$). Here $\tilde{\sigma}$ is the equivariant symplectic form $\sigma + \mu(x)$, $(\alpha \exp i\tilde{\sigma})^{\#}$ is the image of $\alpha \exp i\tilde{\sigma}$ in $H_T(M)$, and r_T is the Kirwan map associated with the action of T on M. This shows that the *non-abelian* localization problem for G can be reduced to an abelian localization problem. In view of this result, we will assume from now on that $T = G$.

We will now give a brief summary of the contents of this paper: The set of critical points of the moment map μ can be written as a finite union of symplectic submanifolds of M:

$$M_{\mathrm{crit}} = \bigcup M_i\,, \tag{1.5}$$

each M_i being the fixed point set of a one-dimensional subgroup T_i of T (see [GS]). We will define in § 2 a residue operation

$$(\mathrm{res})_i : H_T^*(M) \to H_{T/T_i}^{*-d_i}(M_i) \tag{1.6}$$

where $d_i = \operatorname{codim} M_i - 2$. (This will involve some non-intrinsic choices: in particular, a splitting of T into a product of T_i and a complementary torus isomorphic to T/T_i. We

[3]) See [JK]. Another nice account of these results can be found in [Du].

will prove, however, that the definition itself is intrinsic.) Our main result then says that, for an equivariantly closed form, $\alpha \in \Omega_G^d(M)$, $d = \dim M_{\mathrm{red}}$,

$$\int_{M_{\mathrm{red}}} r[\alpha] = \sum{}' \int_{(M_i)_{\mathrm{red}}} r_i[(\mathrm{res})_i(\alpha)] \tag{1.7}$$

where r_i is the Kirwan map associated with the action of T/T_i on M_i. (The "prime" means as above that not all the i's occur in the summation. For details see theorem 3.1.) From this formula one gets, by induction, an expression for (1.2) as a sum of iterated residues of α over certain connected components of the fixed point set of T. In § 4 we will carry out the details of this induction and obtain explicit formulas for these iterated residues. (With some effort these formulas can also be obtained from the contour integral approach of [JK]. For details see [Ka2].)

If ξ is a regular value of the moment map and M_ξ is the reduced space $\mu^{-1}(\xi)/T$ one can define, in analogy with (1.1), a Kirwan map

$$r_\xi : H_T^*(M) \to H^*(M_\xi)$$

and, in analogy with (1.2) one can consider the expression

$$\int_{M_\xi} r_\xi[\alpha]. \tag{1.8}$$

It is not hard to see that this expression is constant on every connected component of the set of regular values. More explicitly the set Δ^0 of regular values of μ is a disjoint union of open convex polytopes

$$\Delta^0 = \Delta_1^0 \cup \cdots \cup \Delta_k^0$$

and (1.8) is constant on each Δ_i^0. In § 5 we will show that as one crosses a common codimension-one face F of two adjacent Δ_i^0's, the change in (1.8) is given by:

$$\sum{}'' \int_{(M_i)_\eta} (r_i)_\eta [(\mathrm{res})_i(\alpha)]$$

the sum being over the set of i's for which $\mu^{-1}(\eta)$ intersects M_i. (Added in proof: We were recently informed by Shaun Martin of a result of him which is quite similar to this.)

2. The formula (1.7) in the circle group case

We will describe in this section how the proof of (1.7) goes in the case $G = S^1$. Most of what we are about to relate has nothing to do with symplectic geometry; so, for the moment, the manifold M can be any compact S^1-manifold with boundary. We will, however, assume that M is oriented and that the action of S^1 on the boundary of M is locally free. Let $Z = \partial M$, let $X = Z/S^1$, and let ι and p be the inclusion of Z into M and the projection of Z onto X. Then, defining the Kirwan map, as above, by

$$r = (p^\#)^{-1} \iota^*, \tag{2.1}$$

one can ask if, in this more general setting, there is a localization formula for the integral

$$\int_X r[\alpha].$$

We will see shortly that the answer is yes, however, to see why, one needs to understand better the map (1.9) (or, what amounts to the same thing, the map

$$(p^\#)^{-1}: H^*_{S^1}(Z) \to H^*(X), \tag{2.2}$$

since $\iota^\#$ is just "restriction"). Let v be the infinitesimal generator of the action of S^1 and recall that the Cartan model $(\tilde{\Omega}, \tilde{d})$ for the equivariant DeRham theory of Z is the complex:

$$\tilde{\Omega} = \Omega^*(Z)^{S^1} \otimes \mathbb{C}[x] \tag{2.3}$$

where $\Omega^*(Z)^{S^1}$ is the algebra of S^1 invariant DeRham forms, $\mathbb{C}[x]$ is the polynomial ring in one variable and $\tilde{d}$ is the exterior differentiation operator

$$\tilde{d} = d \otimes 1 + \iota(v) \otimes x. \tag{2.4}$$

By assigning to x the formal degree two, $\tilde{d}$ becomes an operator of degree one and (as is proved for instance in [Ca]) computes the equivariant cohomology of Z. Thus, to prove that the inverse (2.2) exists one has to prove the following:

Proposition 2.1. *For every $\alpha \in \tilde{\Omega}^k$ with $\tilde{d}\alpha = 0$, there exists $v \in \tilde{\Omega}^{k-1}$ and $\gamma \in \Omega^k(X)$ such that*

$$\alpha = \tilde{d}v + p^*\gamma. \tag{2.5}$$

(This, of course, doesn't quite prove the invertibility of $p^\#$. It just proves the "ontoness"; however the "one-one-ness" is easy.)

Let us first prove (2.5) when $k > \dim X$ in which case (2.5) reduces to $\alpha = \tilde{d}v$. Let θ be an S^1-invariant DeRham one-form on Z with the property $\iota(v)\theta = 1$. Then $\tilde{d}\theta = x + \theta$, so if $\tilde{d}\alpha = 0$, we can formally set,

$$v = \frac{\theta\alpha}{x + d\theta} \tag{2.6}$$

and hence (formally)

$$\tilde{d}v = \frac{(x + d\theta)\alpha}{x + d\theta} = \alpha. \tag{2.7}$$

Notice, however, that the right hand side of (2.6) is

$$\frac{\theta}{x}\alpha \sum \frac{(-d\theta)^i}{x^i}$$

and if $\alpha = \sum \alpha_r x^r$ with $\alpha_r \in \Omega^*(Z)^{S^1}$ this can also be written as

$$\sum_{i,r} \theta\alpha_r(-d\theta)^i x^{r-i-1}\,. \tag{2.8}$$

Suppose now that $\deg \alpha \geqq \dim X$. Then $\deg \alpha_r + 2r \geqq \dim X$ and since $\theta\alpha_r(-d\theta)^i$ is zero if $\deg \alpha_r + 2i \geqq \dim Z$ the only non-zero terms in this sum are those for which $i < r$. Hence (2.8) is a polynomial in x; or, in other words, an element of $\tilde{\Omega}$. This proves that α is $\tilde{d}$-exact as claimed.

The argument we just sketched is due to Berline and Vergne (see [BV]). It breaks down if the degree of α is less than the dimension of Z; but we claim that, in this case, one can still deduce from it the weaker result (2.5). Namely if v_+ is the sum of the terms of degree $\geqq 0$ in (2.8) and β the coefficient of the term of degree -1 one gets

$$\alpha = \tilde{d}v_+ + \imath(v)\beta \tag{2.9}$$

by comparing terms of non-negative degree on both sides of (1.15). However, $\imath(v)\beta$ is horizontal and S^1-invariant; so it is of the form $p^*\gamma$ for a (unique) $\gamma \in \Omega^*(X)$, and (1.17) reduces to

$$\alpha = \tilde{d}v_+ + p^*\gamma$$

which, with v_+ replaced by v, proves the proposition. In addition it shows that at the level of forms, $(p^\#)^{-1}$ is given by

$$\alpha \to (\mathrm{res})_{x=0}\, \imath(v) \frac{\theta\alpha}{x + d\theta}\,. \tag{2.10}$$

Two other descriptions of this map are useful: On the one hand one can think of θ as a connection form with respect to the fibration, p, in which case the form

$$\imath(v)(\theta\alpha)(x + d\theta)^{-1}$$

is just the horizontal component of $\alpha(x + d\theta)^{-1}$; so (2.10) is the map

$$\alpha \to (\mathrm{res})_{x=0} \left(\frac{\alpha}{x + d\theta}\right)_{\text{horizontal}} \tag{2.11}$$

On the other hand, the right hand expression in (2.10) can also be written as

$$(\mathrm{res})_{x=0}\, p_* \frac{\theta\alpha}{x + d\theta} \tag{2.12}$$

where p_* is the Gysin map (or fiber integration map) on equivariant forms.

To get a localization theorem from this result, let M_i, $i = 1, \ldots, N$, be the connected components of the fixed point set of S^1 and let U_i be a tubular neighborhood of M_i such that $U_i \cap U_j = \emptyset$ for $i \neq j$ and $U_i \cap \partial M = \emptyset$. If α is a closed equivariant form on M of degree

$< M$, then, by applying Stokes' theorem to the identity (2.7) (which is valid on all of $M - \bigcup M_i$), one gets

$$0 = \int_M \alpha = \int_{\partial M} \frac{\theta\alpha}{x + d\emptyset} - \sum_k \int_{\partial U_k} \frac{\theta\alpha}{x + d\theta}$$

which can also be written

$$\int_X p_*\left(\frac{\theta\alpha}{x + d\theta}\right) = \sum_k \int_{\partial U_k} \frac{\theta\alpha}{x + d\theta} \tag{2.13}$$

and as one shrinks the radii of the tubular neighborhoods, U_i, to zero the expression on the right becomes localized on the components of the fixed point set. If M_k is oriented one can show, in fact, that as the radius of U_k shrinks to zero, the k-th term on the right tends to

$$\int_{M_k} \frac{\iota_k^* \alpha}{e(x, \nu_k)} \tag{2.14}$$

where ι_k is the inclusion map of M_k into M and $e(x, \nu_k)$ the equivariant Euler class of the oriented normal bundle, ν_k, of M_k[4]. In particular, if $\deg \alpha = \dim X$, one can take residues of both sides of (2.14), and one obtains

$$\int_X r(\alpha) = \sum_k \int_{M_k} (\mathrm{res})_k \alpha \tag{2.15}$$

with

$$(\mathrm{res})_k \alpha = (\mathrm{res})_{x=0} \frac{\iota_k^* \alpha}{e(x, \nu_k)}. \tag{2.16}$$

This is the residue formula (1.7) in the circle group case.

Remarks. (1) The multi-dimensional generalization of (2.16), which we will describe in § 3, is quite different from the definition of residue used by Jeffrey and Kirwan in their proof of (1.4). (Their definition involves a rather complicated contour integral, whereas ours involves a sequence of iterated residues of the same type as (2.16).) For the case of isolated fixed points the equivalence of their definition with ours is proved in [Ka 2].

(2) Our theorem also differs from theirs in another respect: Our definition of the superscript "prime" in (1.4) is not the same as theirs.

(3) In (2.13) and (2.15) one can allow M to be an orbifold-with-boundary since the only non-trivial ingredient involved in the proof of these two identities is Stokes' theorem.

(4) Here is another description of the mapping (2.11). In the finite sum

$$(\alpha)_{\mathrm{hor}} = \sum (\alpha_i)_{\mathrm{hor}} x^i$$

---

[4]) For a proof of this result see, for instance, [AB] or [BV].

substitute the curvature form, $-d\theta$, for the indeterminant, x. This version of (2.11) is known as the "Cartan" map. (See [Ca].)

(5) The symplectic version of (2.15) (for a compact Hamiltonian S^1-space with moment map, μ) can be derived from (2.15) by applying this formula to the manifold-with-boundary

$$M_+ = \{m \in M, \mu(m) \geqq 0\}$$

or to the manifold-with-boundary

$$M_- = \{m \in M, \mu(m) \leqq 0\} .$$

Note that μ is constant on each of the connected components M_i of the fixed point set. By applying (2.15) to M_+ one gets a residue formula involving the M_i's on which μ is positive, and by applying it to M_-, a residue formula involving the M_i's on which μ is negative.

Acknowledgements. We would like to thank Eugene Lerman and Shlomo Sternberg for many helpful suggestions and comments, and we would also like to thank Lisa Jeffrey for several illuminating discussions about localization and residues. Some of the results described here were presented for the first time in a course we taught at M.I.T. in the spring of 1994, and we are grateful to our auditors (in particular, Ana Canas da Silva, Eckhard Meinrenken and Chris Woodward) for proposing a number of improvements in our exposition. Finally, we would like to thank Sylvie Bessette for a beautiful job of editing this manuscript.

3. The definition of $(\mathrm{res})_i$ and the proof of (1.7)

Consider one of the summands in (1.5), e.g., M_1, and recall that there exists a circular subgroup T_1 of T which acts trivially on M_1. Let H be a codimension-one subtorus of T whose intersection with T_1 is finite, and let $\mathfrak{t}_1$ and $\mathfrak{h}$ be the Lie algebras of these two groups. We will choose a basis $x, y_1, \ldots, y_{n-1}$ of $\mathfrak{t}^*$ such that x is an integer basis of $\mathfrak{t}_1^*$ and $y_1, \ldots, y_{n-1}$ a basis of $\mathfrak{h}^*$. Let ν be the normal bundle of M_1 in M and (to simplify slightly the computations below) assume that ν splits into a direct sum of T-equivariant line bundles

$$\nu = \mathbf{L}_1 \oplus \cdots \oplus \mathbf{L}_l .$$

(This involves no essential loss of generality. For justification of this "splitting principle" see [BT], p. 275.) Then the equivariant Euler class, $e_T(\nu)$, of ν is a product:

$$\prod_{i=1}^{l} (m_i x + f_i^j y_j + \mu_i) \tag{3.1}$$

where $m_i \in \mathbb{Z}$, $i = 1, \ldots, l$, is the weight of the representation of T_1 on $\mathbf{L}_i$ and $f_i^j y_j + \mu_i$ is a Cartan representative of the H-invariant Chern class of $\mathbf{L}_i$. (Here we are using the Einstein summation convention with $f_i^j \in \Omega^0(M_1)$ and $\mu_i \in \Omega^0(M_1)$.) Now let

$$\alpha = \sum \alpha_{i,J} x^i y^J , \quad \alpha_{i,J} \in \Omega^*(M)$$

be an equivariantly closed element of $\Omega_T(M)$. Then, formally,

$$(\mathrm{res})_1(\alpha) \stackrel{\mathrm{def}}{=} (\mathrm{res})_{x=0} \frac{\iota^*\alpha}{e_T(\nu)}$$

where $\iota: M_1 \to M$ is the inclusion map. To give a rigorous meaning to the right hand side we will write

$$\frac{\iota^*\alpha}{e_T(\nu)} = \frac{1}{\prod m_i} x^{-l} \prod_{i=1}^{l} \left(1 + \frac{\mu_i + f_i^j y_j}{m_i x}\right)^{-1}$$

and expand the right hand side into a product of power series.

$$\frac{1}{\prod m_i} (\textstyle\sum \alpha_{i,J} x^i y^J) \displaystyle\prod_{i=1}^{l} \sum_{k_i=1}^{\infty} (-1)^{k_i} \left(\frac{\mu_i + f_i^j y_j}{m_i}\right)^{k_i} x^{-k_i - l}. \tag{3.2}$$

By multiplying out these power series term by term, one can rewrite this product as a sum

$$\sum_{j=m_0}^{-\infty} \beta_j x^j \tag{3.3}$$

where β_j is a polynomial in $y_1, \ldots, y_{n-1}$ of the form

$$\beta_j = \sum \beta_{j,J} y^J, \quad \beta_{j,J} \in \Omega^*(M_1). \tag{3.4}$$

Moreover, formally

$$d_T \frac{\iota^*\alpha}{e_T(\nu)} = 0$$

and since T_1 acts trivially on M_1, this implies

$$d_H \beta_j = 0 \tag{3.5}$$

for all j (in particular for $j = -1$). Now set

$$(\mathrm{res})_1(\alpha) = \beta_{-1}. \tag{3.6}$$

This definition would appear to depend on the choice of the splitting, $T = T_1 \times H$; but we will now show that it doesn't: If $T = T_1 \times H'$ is another splitting, then

$$y_i' = y_i$$

and

$$x' = x + \sum a_j y_j, \quad a_j \in \mathbb{R},$$

and with this change of variables (3.3) becomes

$$\sum_{i=m_0}^{-\infty} \beta_i (x' - \textstyle\sum a_j y_j)^i = \displaystyle\sum_{i=m_0}^{-\infty} \beta_i' (x')^i \tag{3.7}$$

where the β_i''s are computed from the β_i's as follows:

(1) If $i \geqq 0$, replace $(x' - \sum a_j y_j)^i$ by the finite sum

$$\sum_{r=0}^{i} (-1)^r \binom{i}{r} (\sum a_j y_j)^r (x')^{i-r} .$$

(2) If $i = -m < 0$, replace $(x' - \sum a_j y_j)^{-m}$ by the power series

$$\sum_{k=0}^{\infty} \binom{k+m-1}{k} (\sum a_j y_j)^k (x')^{-m-k} .$$

Note that the first substitution has no effect whatsoever on the negative terms in the sum (3.7), and the second substitution doesn't change $\beta_{-1}(y)$ (though it does change the other negative β_i's). Thus

$$\beta_{-1}(y) = \beta'_{-1}(y)$$

as claimed.

Another fact worth remarking is that (3.6) lowers the degree of α by the amount: $\operatorname{codim} M_1 - 2$. (N.B. β_{-1} is two degrees higher than $\beta_{-1}x^{-1}$. This accounts for the "-2".)

Having defined $(\mathrm{res})_1$ (and hence, in effect, $(\mathrm{res})_i$ for all i) we will turn to the proof of (1.7). We will begin with a few combinatorial preliminaries: Recall ([A] or [GS]) that the image, Δ, of the moment map is a convex polytope. Moreover, its interior contains the set, Δ^0, of regular values of μ, and this set is a disjoint union of convex polytopes

$$\Delta^0 = \Delta_1^0 \cup \cdots \cup \Delta_k^0 . \tag{3.8}$$

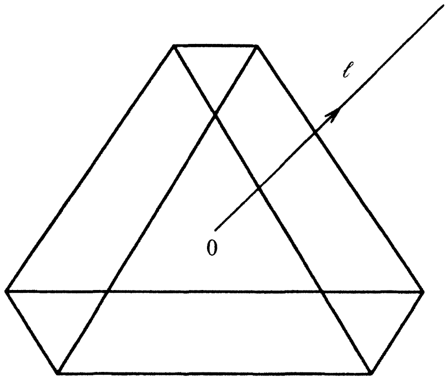
By assumption the origin is a regular value of μ, so it is contained in one of these Δ_i^0's. Let θ be a non-zero element of the weight lattice of $\mathfrak{t}^*$, and consider the ray through the origin:

$$\ell = \{t\theta, 0 \leqq t < \infty\} . \tag{3.9}$$

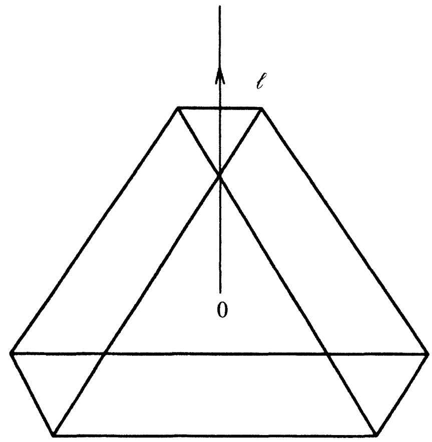
If one chooses θ appropriately, this ray won't intersect any of the walls of the Δ_i^0's of codimension greater than one, and will intersect the codimension-one walls transversally. (See the figure on page 132.)

Since θ belongs to the weight lattice of $\mathfrak{t}^*$, its annihilator

$$\mathfrak{h} = \{v \in \mathfrak{t}, \langle\theta, v\rangle = 0\}$$



Allowed



Forbidden

is the Lie algebra of a codimension-one subtorus, H, of T; and the assumptions on ℓ imply that the moment map, μ, is transverse to ℓ and that H acts in a locally free fashion on $\mu^{-1}(\ell)$. In particular, $\mu^{-1}(\ell)$ is a manifold-with-boundary and the quotient space

$$W_\ell \overset{\text{def}}{=} \mu^{-1}(\ell)/H$$

is an orbifold-with-boundary whose boundary is

$$\partial W_\ell = \mu^{-1}(0)/H\,.$$

In addition there is a residual action of the circle group, $S = T/H$, on W_ℓ, and this action is locally free on the boundary. Consider now the Kirwan map

$$r : H_T(M) \to H(M_{\text{red}})\,.$$

This factors into Kirwan maps

$$\sigma : H_T(M) \to H_S(W_\ell)$$

and

$$r_\ell : H_S(W_\ell) \to H(M_{\text{red}})\,.$$

We will use this factorization and the theorem (2.15) to prove

Theorem 3.1. *For* $[\alpha] \in H_T^d(M)$, $d = \dim M_{\text{red}}$,

$$\int_{M_{\text{red}}} r[\alpha] = \sum{}' \int_{(M_i)_{\text{red}}} r_i[(\text{res})_i(\alpha)]\,, \tag{3.10}$$

an "i" occurring in the sum on the right if and only if $M_i \cap \mu^{-1}(\ell)$ *is non-empty.*

Thus the summands in (3.10) correspond to the points of intersection of ℓ with the $(n-1)$-dimensional walls of the polyhedral complex (3.8). (Note, by the way, that ℓ is not unique. There are many rays with the properties above, and each such ray gives a different formula for the left hand side of (3.10).)

Proof of (3.10). The fixed point set of S in W_ℓ is the union of the sets

$$(\mu^{-1}(p_i) \cap M_i)/H = (W_\ell)_i \tag{3.11}$$

where the p_i's are the points of intersection of ℓ with the $(n-1)$-dimensional walls of the complex (3.8). However, (3.11) is just the reduced space associated with the action of T/T_i on M_i. Therefore, in view of (2.15), we need only check that if we take the residue $(\text{res})_i$ of α, upstairs on M, and then apply to it the Kirwan map, r_i, we get the same element of $H^*((W_\ell)_i)$ as that which we would have gotten by applying the residue operation, downstairs on W_ℓ, to the class, $\sigma(\alpha) \in H_S(W_\ell)$. This, however, follows easily from the following result:

Proposition 3.1. *The map $r_i : H_T(M_i) \to H_S((W_\ell)_i)$ takes the T-equivariant Euler class of the normal bundle of M_i onto the S-equivariant Euler class of the normal bundle of $(W_\ell)_i$.*

Proof. Let $Z_i = \mu^{-1}(p_i) \cap M_i$; and let π be the projection of Z_i onto $(W_\ell)_i$ and $\iota : Z_i \to M_i$ the inclusion map. Then, if ν is the normal bundle of M_i in M and ν_ℓ the normal bundle of $(W_\ell)_i$ in W_ℓ,

$$\iota^* \nu = \pi^* \nu_\ell \, .$$

Hence the proposition follows from the fact that, as a map from oriented real vector bundles to cohomology, the Euler class is functorial.

A few cautionary remarks about orientations. (1) Already, in (2.15), there is a small orientation problem (which we conveniently chose to ignore). Namely, for the left hand side of (2.15) to make sense, we need to assign an orientation to X. Since M is oriented and $X = \partial M / S^1$, this amounts to assigning an orientation to S^1, or equivalently, to making a choice between the two basis vectors, $\pm \frac{\partial}{\partial \theta}$, of the Lie algebra of S^1. At the root of this problem is the fact that the "x" in (2.16) is not an abstract indeterminant but is an element of the dual of the Lie algebra of S^1, and hence, one should choose x and $\frac{\partial}{\partial \theta}$ to have consistent orientations.

(2) The formula (3.10) may appear to be immune to orientation problems since M_{red} and $(M_i)_{\text{red}}$ are symplectic orbifolds. Nonetheless, for the same reason as above, one has to choose the indeterminant, x, figuring in the definition of $(\text{res})_i$ so that its orientation is consistent with the orientation of the ray, ℓ, at all points of intersection, p_i, of ℓ with the walls of Δ_0^i.

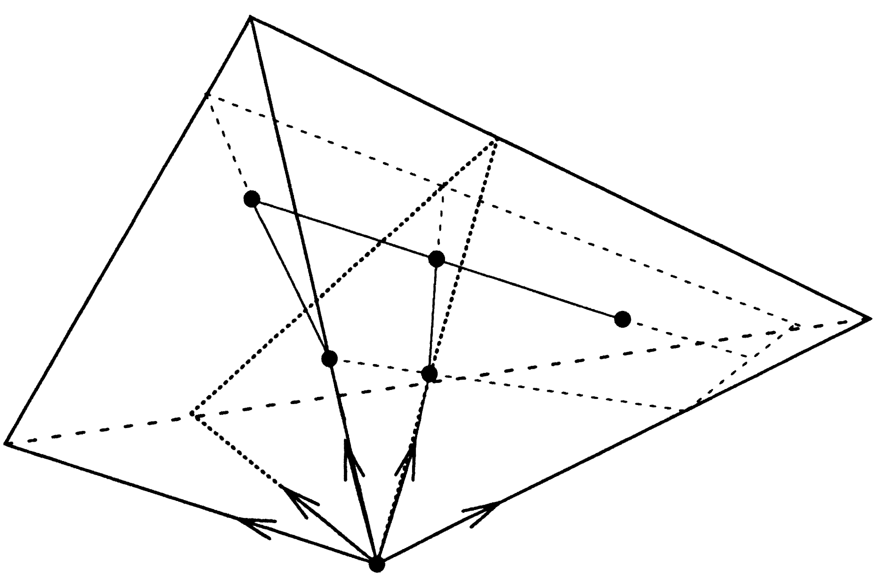
(3) By induction one gets from (3.10) a formula for the left hand side of (3.10) in terms of the restrictions of α to the components of the fixed point set of T. (For details see the next section.) A consistent choice of orientations at all stages of this induction seems to be equivalent to the choice of a polarization of the weights of the isotropy representations in the Jeffrey-Kirwan theorem.

4. The iterated residues

The formula (3.10) given one an inductive method for computing the integral (1.2) in terms of the restrictions of α to the components of the fixed point set of T. In this section we will carry out the details of this induction and obtain a formula for (1.2) in terms of iterated residues. We will begin by describing some of the combinatorial features of this formula: As in § 2 let Δ be the image of the moment map, and let

$$\Delta^0 = \Delta_1^0 \cup \cdots \cup \Delta_k^0 \tag{4.1}$$

be its set of regular values. Draw a line, ℓ, from the origin to the boundary of Δ which avoids all the codimension-two walls of the Δ_i^0's, and then, for each intersection, p_j, of ℓ with a codimension-one wall, W_j, draw a line, ℓ_j, from p_j to the boundary of $\Delta \cap W_j$ which avoids all the codimension-three walls of the Δ_i^0's. Now, from the intersection of the ℓ_j's with the codimension-two walls of the Δ_j^0's draw lines ... (it is clear how to continue). We will call the object obtained by this iteration process a *dendrite*[5]). Thus a dendrite will consist of a main branch (the line ℓ), secondary branches (the lines ℓ_i), branches going out of these secondary branches, and so on. The branches that are created at the $(k+1)$-st stage of this construction will lie on the $(n-k)$-skeleton of (4.1) and at the last stage they will lie on the edges of the Δ_i^0's and terminate in vertices. (See the figure below.) In particular, a dendrite D will be a (non-disjoint) union of *paths*, a path being a closed polygonal curve on D starting at the origin and terminating in a vertex. (In the figure D consists of two paths both terminating in the same vertex.)



[5]) Definition (from Webster's New Collegiate Dictionary): "dendrite. 1. A branching tree-like figure produced in a mineral by a foreign mineral. 2. A crystalline arboreal form." This, we feel, is a fairly appropriate description of the object above.

In our localization formula the only paths that will make contributions are paths of the following type: Suppose ℓ (the main branch of D) intersects the $(n-1)$-skeleton of (4.1) in a point p_1. Then, since p_1 is a critical value of μ, it must lie on the image of one of the M_i's on the list (1.5), and, to simplify our notation let's rebaptize:

$$M_i = M^{(1)}$$

and

$$T_i = T^{(1)}$$

where T_i is the one-dimensional subgroup of T leaving M_i fixed. Now repeat this construction with M replaced by $M^{(1)}$ and ℓ by the branch of D going out of p_1. By iteration one gets a sequence of branches of D,

$$\ell, \ell^{(1)}, \ldots, \ell^{(n)}, \tag{4.2}$$

a sequence of points on these branches,

$$p_0 = 0, p_1, p_2, \ldots, p_n \tag{4.3}$$

with $\ell^{(n)} = \{p_n\}$, a sequence of symplectic manifolds,

$$M = M^{(0)} \supset M^{(1)} \supset \cdots \supset M^{(n)} = F, \tag{4.4}$$

and a sequence of subtori of T

$$T^{(0)} = \{0\} \subset T^{(1)} \subset \cdots \subset T^{(n)} = T \tag{4.5}$$

(with $\dim T^{(i)} = i$) such that $M^{(i)}$ is a connected component of the fixed point set of $T^{(i)}$ and p_i is contained in $\mu(M^{(i)})$. (In particular, F is a connected component of the fixed point set of T whose image is p_n.) This construction gives rise to a path, P, consisting of the line segments joining 0 to p_1, p_1 to p_2, etc. We will call a path of this type an *admissible path* (and regard (4.4) and (4.5) as part of its defining data).

Now fix D and consider all admissible paths, P, joining the origin to the vertices of D. Each of these will make a contribution, $r(P)$, to the localization formula, and the integral (1.2) will be the sum of these contributions; so the main task we are faced with is to compute the $r(P)$'s. We will start by fixing some notation: Let P be defined by (4.2)–(4.5), and let $x_1, \ldots, x_n$ be a basis of $\mathfrak{t}^*$ such that $x_1^*, \ldots, x_i^*$ is a basis of the integer lattice in the Lie algebra of $T^{(i)}$. For each $M^{(i)}$ in the chain (4.4), let ν_i be the restriction to F of the normal bundle of $M^{(i)}$ in $M^{(i-1)}$. Then the normal bundle of F in M is the sum

$$\nu = \nu_1 \oplus \cdots \oplus \nu_n$$

and hence the equivariant Euler class of ν is:

$$\varepsilon_T(\nu) = \prod_{i=1}^{n} \varepsilon_T(\nu_i)\,. \tag{4.6}$$

Let

$$(4.7) \qquad \alpha_i, \quad i = 1, \ldots, N$$

be the weights of the representation of T on the typical fiber of v, and order these weights so that

$$\alpha_i, \quad N_{r-1} < i \leqq N_r$$

are the weights of the representation of T on the typical fiber of v_r (where

$$0 = N_0 < N_1 < \cdots < N_n = N).$$

Since $T^{(r-1)}$ acts trivially on v_r,

$$\alpha_j(x) = m_j x_r + \sum_{l > r} m_{jl} x_l, \quad m_j \neq 0,$$

when $N_{r-1} < j \leqq N_r$. Note also that

$$(4.8) \qquad \varepsilon_T(v_r) = \sum_{N_{r-1} < j \leqq N_r} (\alpha_j(x) + \mu_j), \quad \mu_j \in H^2(F).$$

Now let α be a closed equivariant form on M of degree d, where $d = \dim M - 2\dim T$, and let ι_F be the inclusion of F into M. Then by (3.10) the contribution, $r(P)$, to the localization formula is the integral over F of:

$$(4.9) \qquad (\mathrm{res})_{x_n = 0}\left(\frac{1}{\varepsilon_T(v_n)} (\mathrm{res})_{x_{n-1} = 0}\left(\frac{1}{\varepsilon_T(v_{n-1})} \cdots (\mathrm{res})_{x_1 = 0} \frac{\iota_F^* \alpha}{\varepsilon_T(v_1)}\right)\right)$$

which, in view of (4.6) can be written more compactly as

$$(4.10) \qquad (\mathrm{res})_{x_n = 0} \cdots (\mathrm{res})_{x_1 = 0} \frac{\iota_F^* \alpha}{\varepsilon_T(v)}.$$

This is the iterated residue we referred to at the end of section 1. The gist of what we have proved so far is the following

Theorem 4.1. *For a fixed dendrite D the integral* (1.2) *is equal to the sum over all paths P belonging to D of the residues* (4.10).

Finally we will describe how to go about evaluating these residues explicitly. For this we will need

Lemma 4.1. *Let A be a graded commutative algebra over $\mathbb{C}$ and let $f = f(x)$ be a polynomial in x with coefficients in A. Then, for indeterminants $y_1, \ldots, y_n$,*

$$(4.11) \qquad (\mathrm{res})_{x = \infty} \frac{f(x)}{(x - y_1) \cdots (x - y_n)} = \sum_i \frac{f(y_i)}{\prod_{j \neq i} (y_i - y_j)}.$$

More generally,

$$\text{(4.12)} \qquad (\text{res})_{x=\infty} \frac{f(x)}{(x-y_1)^{k_1}\cdots(x-y_n)^{k_n}} = \sum_i f_i^{(k_i-1)}(y_i)$$

where

$$f_i = \frac{1}{(k_i-1)!}\frac{f(x)}{\prod_{j\neq i}(x-y_j)}.$$

Proof. Without loss of generality we can take $A = \mathbb{C}$ in which case it suffices to prove (4.11) with the y_i's replaced by a "generic" set of complex numbers, $c_1, \ldots, c_n$ with $c_i \neq c_j$ for $i \neq j$. But in this case $(\text{res})_{x=\infty}$ can be replaced by the contour integral

$$\frac{1}{2\pi i}\int_\gamma \frac{f(x)}{(x-c_1)\cdots(x-c_n)}\,dx$$

over a contour in the complex plane which bounds a region containing all the c_i's (and, in this case, (4.11) is a standard identity in complex variable theory). Similar remarks apply to (4.12).

Since the left hand side of (4.11) is a polynomial in the indeterminants $y_1, \ldots, y_n$ this proves in particular:

Corollary 4.1. *For all m the expression*

$$\text{(4.13)} \qquad \sum_m \frac{y_i^m}{\prod_{j\neq i} y_i - y_j}$$

is a polynomial in $y_1, \ldots, y_n$.

Remark. In the applications below, the algebra A will be the tensor product

$$\Omega(F)\otimes\mathbb{C}[x_1, \ldots, x_n].$$

Before attempting to compute (4.10) in general we will first consider some special cases.

Case 1. F is a point p and the weights of the isotropy representation of T on the tangent space at p have the following genericity property: *Any subset of n distinct weights are a basis of* $\mathfrak{t}^*$. Under this assumption one gets the following formula for (4.10) by applying inductively the identity (4.11):

Theorem 4.2. *For every sequence*

$$J = (j_1, \ldots, j_{n-1})$$

with $j_r \leqq N_r$ *and* $j_r \neq j_s$ *for* $r \neq s$, *let* $Q_J \in \mathbb{R}^n$ *be the unique solution of the system of n linear equations*

$$(4.14)\qquad \alpha_j(x_1, \ldots, x_n) = 0, \quad j \in J, \qquad x_n = 1$$

and let A_J *be the determinant of the coefficient matrix of this system. Then* (4.10) *is equal to*

$$(4.15)\qquad \sum_J (\iota_p^* \alpha)(Q_J) A_J^{-1} \Big(\prod_{j \notin J} \alpha_j(Q_J) \Big)^{-1}.$$

Proof. The iterated residue

$$(\mathrm{res})_{x_{n-1}} \cdots (\mathrm{res})_{x_1} \frac{\alpha(x)}{\varepsilon_T(v)}$$

is equal, by (4.8) and (4.11), to

$$\sum_J \alpha(Q_J x_n) A_J^{-1} \Big(\prod_{j \neq J} \alpha_j(Q_J x_n) \Big)^{-1}$$

in view of the fact that, by eliminating $x_1, \ldots, x_{n-1}$ from the equations

$$\alpha_j(x_1, \ldots, x_n) = 0, \quad j \in J,$$

one ends up with the solution

$$(x_1, \ldots, x_n) = x_n Q_J.$$

Notice, however, that $\alpha(Q_J x)$ is a homogeneous expression in x of degree

$$\frac{d}{2} = \dim M/2 - n$$

whereas

$$\Big(\prod_{j \neq J} \alpha_j(Q_J x) \Big)^{-1}$$

is of degree $\frac{d}{2} + 1$, hence, if one factors out redundant powers of x_n, one obtains

$$\frac{1}{x_n} \sum_J \alpha(Q_J) A_J^{-1} \Big(\prod_{j \neq J} \alpha_j(Q_J) \Big)^{-1}$$

whose residue is clearly (4.14).

Remark. If one reorders the sequence $(j_1, \ldots, j_n)$ so that the j_r's still satisfy $j_r \leqq N_r$ the new sequence gives the same contribution to (4.14) as the old sequence, hence if n_J is the number of such reorderings (4.14) can be rewritten:

$$\sum_J \iota_p^* \alpha(Q_J) n_J A_J^{-1} \left(\prod_{j \neq J} \alpha_j(Q_J) \right)^{-1} \tag{4.16}$$

the sum now being over *unordered* sequences.

Case 2. $F = \{p\}$, but the α_i's no longer satisfy the genericity condition described above. We won't discuss this case in detail. The iterated residue (4.10) can be computed exactly as above, but now the denominators will involve multiple factors, and, hence, by (4.12) the answer will involve derivatives of α on the right hand side and be slightly more complicated than (4.15) (but not much).

Case 3. The α_i's satisfy the same genericity condition as in case 1, but $\dim F > 0$. In this case the expression whose iterated residue we have to compute is

$$\frac{\iota_F^* \alpha}{\prod (\alpha_i(x) + \mu_i)} \tag{4.17}$$

where the μ_i's are two-forms. Replacing the μ_i's by indeterminants, y_i, we get the formal expression

$$(\mathrm{res})_{x_n} \cdots (\mathrm{res})_{x_1} \frac{\iota_F^* \alpha}{\prod (\alpha_i(x) + y_i)} \tag{4.18}$$

and, by essentially the same argument as above, this is equal to:

$$\sum_J (\iota_F^* \alpha)(\bar{Q}_J) A_J^{-1} \left(\prod_{j \neq J} \alpha_j(\bar{Q}_J) \right)^{-1} \tag{4.19}$$

summed over all sequences, $J = (j_1, \ldots, j_n)$ with $j_r \neq j_s$ for $r \neq s$, and $j_r \leqq N_r$. Here however, $\bar{Q}_J$, is the solution in terms of y_j, $j \in J$, of the system of equations

$$\alpha_j(x) = y_j, \quad j \in J \tag{4.20}$$

and, hence, is an n-tuple of linear functionals in the y_j's. (As above, A_J is the determinant of the coefficient matrix of the system (4.20). However, this is now an $n \times n$ matrix rather than an $(n-1) \times (n-1)$ matrix.) The way we have written it, (4.18) appears to be a *rational* function of the y_i's; however, we claim:

Proposition 4.1. (4.18) *is a polynomial in* $y_1, \ldots, y_N$.

Proof. Induction via the identity (4.13).

Thus to compute the iterated residue of (4.17) we merely have to substitute the μ_i's into (4.19) in place of the y_i's.

The general case. This can be handled in the same way as above, except that (4.19) is a little more complicated. (In particular, it involves some derivatives of α with respect to $x_1, \ldots, x_n$.)

5. The localization formula for the reduced spaces M_ξ

For a regular value ξ of the moment map μ one defines the reduced space M_ξ to be the quotient $\mu^{-1}(\xi)/T$; and for this quotient one has a Kirwan map

$$r_\xi = (p_\xi^\#)^{-1} \iota_\xi^\#$$

where p_ξ is the projection of $\mu^{-1}(\xi)$ onto M_ξ and ι_ξ the inclusion of $\mu^{-1}(\xi)$ into M. Now let α be a closed T-equivariant form on M of degree d, where $d = \dim M - 2\dim T$, and for each ξ consider the integral

$$\int_{M_\xi} r_\xi[\alpha]. \tag{5.1}$$

We claim that as ξ varies inside a fixed connected component Δ_i^0 of the set Δ^0 of regular values of μ, the integral (4.1) stays constant. To prove this let $U_i = \mu^{-1}(\Delta_i^0)$, and note that the action of T on U_i is locally free, and hence the projection

$$\pi : U_i \to U_i/T$$

induces a bijective map on cohomology,

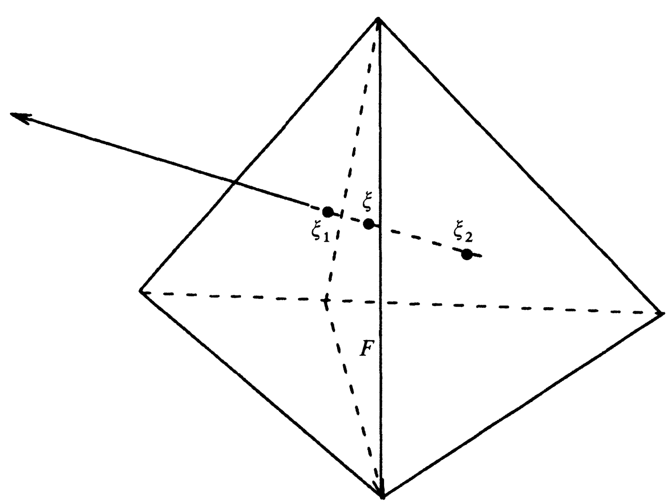
$$\pi^\# : H^*(U_i/T) \to H_T^*(U_i).$$

Letting j_ξ be the inclusion of M_ξ into U_i/T one can rewrite (5.1) as

$$\int_{M_\xi} j^{\#\xi}(\pi^\#)^{-1}(\alpha_i) \tag{5.2}$$

α_i being the restriction of α to U_i and this is clearly independent of ξ (since the homology class $(j_\xi)_\#[M_\xi]$ is independent of ξ).

What happens, however, if ξ crosses a common codimension-one wall F of two adjacent Δ_i^0's? The answer is easy to read off from the formula (3.10). Namely, let Δ_1^0 and Δ_2^0 be the two polyhedrons in the figure below and ξ_1 and ξ_2 points inside them:



Let ℓ be the oriented ray joining ξ_1 to ξ_2 and let ξ be the point where it intersects F. If one uses the formula (3.10) to compute (5.1) at ξ_1 one must include the residue terms associated with the intersection at ξ, but if one uses this formula to compute (5.1) at ξ_2 one can neglect these terms. Thus the change in (5.1) as one goes from Δ_1^0 to Δ_2^0 is

$$\sum{}'' \int_{(M_i)_\xi} (r_i)_\xi (\mathrm{res})_i[\alpha] \tag{5.3}$$

the sum being over all i for which M_i intersects $\mu^{-1}(\xi)$.

We will describe a couple of applications of this result:

(1) Let $\tilde{\sigma}$ be the equivariant symplectic form

$$\sigma + \mu(x) = \sigma + \sum_{i=1}^{n} \mu_i x_i ,$$

the μ_i's being the coordinates of the moment map, and let $w = (w_1, \ldots, w_n)$ be an n-tuple of complex numbers. Consider the expression (5.1) with α equal to the equivariant form

$$\frac{1}{d!}(\tilde{\sigma} - \sum w_i x_i)^d \tag{5.4}$$

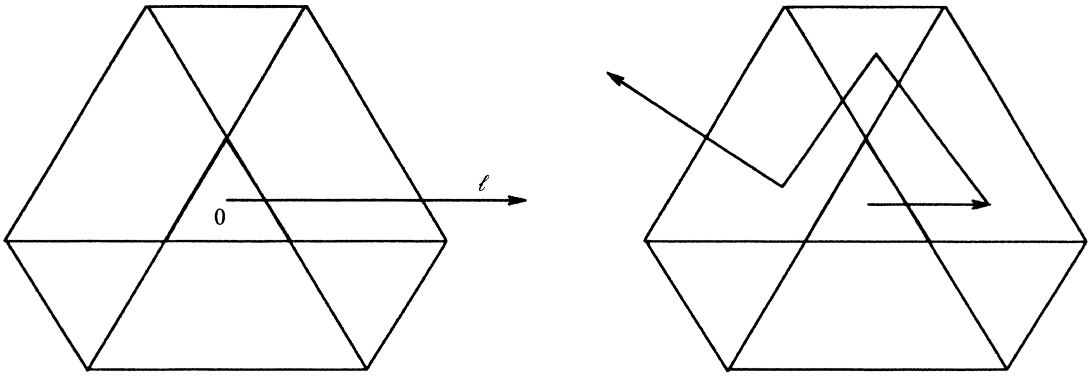
where $d = \dim M/2 - n$. By the result above this expression is equal to a constant

$$\gamma_i(w) \tag{5.5}$$

on each of the polytopes Δ_i^0, but this constant, of course, depends on w; and in fact it is clear from (4.4) that it is a polynomial in w of degree d. By definition this is the *Duistermaat-Heckman polynomial* associated with Δ_i^0, and, from (5.3), one gets an explicit formula for how this polynomial changes as one crosses an $(n-1)$-dimensional wall between two adjacent Δ_i^0's. (For a somewhat less precise result of this kind, see [GLS], theorem 5.1.)

(2) From (5.3) one gets a slightly more flexible method for computing the integral (1.2) than that which we described in § 2. Namely let $i_0, \ldots, i_m$ be a sequence such that $\Delta_{i_0}^0$ contains the origin, $\Delta_{i_r}^0$ and $\Delta_{i_{r+1}}^0$ have an $(n-1)$-dimensional wall in common, and $\Delta_{i_m}^0$ has an $(n-1)$-dimensional wall on the other side of which is the region exterior to Δ. From (4.3) one gets a formula for the (constant) value of the function (5.1) on the polytope $\Delta_{i_r}^0$ in terms of the value of this function on $\Delta_{i_{r+1}}^0$; and, since the value of (5.1) is zero when ξ is in the exterior of Δ, one gets by recursion, a formula for (1.2) (as a sum of expressions of the type (5.3)) of which the formula we described in § 3 is a special case.

We can illustrate the two methods schematically by the two figures below:



In the first figure we have joined the origin to the exterior of Δ by a single ray and in the second figure by a sequence of such rays.

References

[A] *M. Atiyah*, Convexity and commuting Hamiltonians, Bull. Lond. Math. Soc. **14** (1982), 1–15.

[AB] *M. Atiyah* and *R. Bott*, The moment map and equivariant cohomology, Topology **23** (1984), 1–28.

[BV] *N. Berline* and *M. Vergne*, Classes caractéristiques équivariantes, C.R. Acad. Sci. Paris **295** (1982), 539–541.

[BT] *R. Bott* and *L. Tu*, Differential forms in algebraic topology, Springer-Verlag, New York 1982.

[Ca] *H. Cartan*, Transgression dans un groupe de Lie et dans un espace fibré principal, Colloque de Topologie, CBRM Bruxelles (1950), 57–71.

[Du] *J.J. Duistermaat*, Equivariant cohomology and stationary phase, preprint 817, University of Utrecht, August 1993.

[GLS] *V. Guillemin*, *E. Lerman* and *S. Sternberg*, On the Kostant multiplicity formula, J. Geom. Phys. **5**:4 (1988), 721–750.

[GS] *V. Guillemin*, *S. Sternberg*, Convexity properties of the moment map, Invent. Math. **67** (1982), 491–513.

[JK] *L. Jeffrey*, *F. Kirwan*, Localization for non-abelian group actions, Topology, to appear.

[Ka] *J. Kalkman*, Cohomology rings of symplectic quotients, J. reine angew. Math. **458** (1995), 37–52.

[Ka2] *J. Kalkman*, Residues in non-abelian localization, hep-th / 9407113.

[Ki] *F. Kirwan*, Cohomology of quotients in symplectic and algebraic geometry, Math. Notes **31**, Princeton Univ. Press, Princeton NJ 1984.

[W] *E. Witten*, Two dimensional gauge theories revisited, J. Geom. Phys. **9** (1992), 303–368.

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Eingegangen 10. Oktober 1994, in revidierter Fassung 27. Juni 1995

J. reine angew. Math. **470** (1996), 143—152

Journal für die reine und angewandte Mathematik

A remark on Hodge cycles on moduli spaces of rank 2 bundles over curves

By *Indranil Biswas* at Saint-Martin d'Hères

1. Introduction

Let k be an algebraically closed field of characteristic zero and of finite transcendence degree over $\mathbb{Q}$. Let X/k be an irreducible smooth projective curve of genus $g \geqq 2$. Let $\mathcal{N}$ be the moduli space of stable vector bundles of rank 2 and fixed determinant with degree one over X. It is known that $\mathcal{N}/k$ is an irreducible smooth projective variety of dimension $3g-3$ ([S], Ch' VI A, Part 1) ($\mathcal{N}$ is nonempty as $g \geqq 2$). Consider the complex $\Omega^{\cdot}_{\mathcal{N}}$ in which $\Omega^{i}_{\mathcal{N}}$ is the sheaf of algebraic differential i-forms, and the homomorphisms are given by the exterior derivation. The n-th hypercohomology $\mathbb{H}^n(\Omega^{\cdot}_{\mathcal{N}})$, of the complex $\Omega^{\cdot}_{\mathcal{N}}$, with respect to the Zariski topology of $\mathcal{N}$ is denoted by $H^n_{\mathrm{DR}}(\mathcal{N})$ ([D], p.16–17). Let $\sigma : k \to \mathbb{C}$ be an embedding. Define $\mathcal{N}_{\mathbb{C}} := \mathcal{N} \times_{\mathrm{Spec}\,k} \mathrm{Spec}\,\mathbb{C}$. (It turns out that $\mathcal{N}_{\mathbb{C}}$ is the moduli space of stable bundles of rank two and fixed determinant with degree one over the Riemann surface $X_{\mathbb{C}} := X \times_{\mathrm{Spec}\,k} \mathrm{Spec}\,\mathbb{C}$ (Proposition 2.3).) The complex $\Omega^{\cdot}_{\mathcal{N}_{\mathbb{C}}}$ is quasi-isomorphic to the constant sheaf $\mathbb{C}$. In particular, $H^n_{\mathrm{DR}}(\mathcal{N}_{\mathbb{C}}) = H^n(\mathcal{N}_{\mathbb{C}}, \mathbb{C})$. Using the projection $p_1 : \mathcal{N}_{\mathbb{C}} \to \mathcal{N}$, we obtain a morphism of complexes $p_1^* \Omega^{\cdot}_{\mathcal{N}} \to \Omega^{\cdot}_{\mathcal{N}_{\mathbb{C}}}$. The induced homomorphism

$$\phi^n_k : H^n_{\mathrm{DR}}(\mathcal{N}) \otimes_{k,\sigma} \mathbb{C} \to H^n(\mathcal{N}_{\mathbb{C}}, \mathbb{C}),$$

is an isomorphism ([D], Theorem 1.4). A cohomology class $\theta \in H^{2i}_{\mathrm{DR}}(\mathcal{N})$ is called a *Hodge cycle relative to* σ *on* $\mathcal{N}$ if

$$\phi^{2i}_k(\theta) \in H^{2i}(\mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^i \mathbb{Q}) \cap H^{i,i}(\mathcal{N}_{\mathbb{C}}).$$

($H^n_{\mathrm{DR}}(\mathcal{N}_{\mathbb{C}}) = \sum_j H^{j,n-j}(\mathcal{N}_{\mathbb{C}})$ is the Hodge decomposition.) The class θ is called an *absolute Hodge cycle* if it is a Hodge cycle relative to any embedding of k in $\mathbb{C}$.

Our aim here is to prove that any Hodge cycle on $\mathcal{N}$ relative to σ is an absolute Hodge cycle (Theorem 4.5). Since, after choosing a point of X, the variety $\mathcal{N}$ can be identified with the moduli space of stable bundles of rank two and fixed determinant with any odd degree over X, the same statement holds for moduli spaces of rank two bundles with fixed odd degree determinant.

The moduli of rank one bundles over a curve, i.e. it's Jacobian is an abelian variety. Deligne proved that for an abelian variety defined over a field of finite transcendence degree over $\mathbb{Q}$, any Hodge cycle is an absolute Hodge cycle ([D], Main Theorem 2.11). We reduce our problem to the case of the Jacobian. A recent result on generators of the cohomology group of $\mathcal{N}_{\mathbb{C}}$ which was proved independently by several mathematicians – V. Baranovsky [Ba]; A. King and P. Newstead [KN]; B. Siebert and G. Tian [ST] and D. Zagier [Z] – forms a key input for the reduction.

It is easy to see that for any cycle Z on a smooth projective variety X defined over k the cycle class $[Z]$ is an absolute Hodge cycle ([D], p. 28, Example 2.1 (a)) on X. So if the Hodge conjecture is valid for X then any Hodge cycle on it is an absolute Hodge cycle. The Hodge conjecture for a smooth moduli space of vector bundles over the general curve is now known. For the case of rank two bundles it was proved in [BKN]. For higher rank it was proved in [BN]. Thus any Hodge cycle on the smooth moduli space of vector bundles over the general curve is an absolute Hodge cycle.

In a work in progress using the correspondence developed in [Bi] we hope to prove that any Hodge cycle on a smooth moduli space of vector bundles (of any rank) over a curve is actually an absolute Hodge cycle.

Acknowledgment. The author is grateful to M. S. Narasimhan for posing the question addressed here, and to N. Nitsure for some very useful conversations. Thanks are also due to the referee for pointing out several references.

2. Change of base field

We continue with the notations of the introduction. Let $E_k \to X$ be a stable vector bundle of rank 2 and degree 1. Given a homomorphism $\sigma : k \to \mathbb{C}$, we have the bundle

$$E_{\mathbb{C}} \to X_{\mathbb{C}} := X \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$$

where $E_{\mathbb{C}}$ is the pull-back of E_k using the projection $p_1 : X_{\mathbb{C}} \to X$. Let us recall the Proposition 3 (p. 97) of [L].

Lemma 2.1 ([L], Proposition 3). *The bundle $E_{\mathbb{C}}$ on $X_{\mathbb{C}}$ is a stable bundle.*

Remark 2.2. Using the argument in the proof of [L], Proposition 3, and the fact that a stable bundle is simple, i.e. endomorphisms of a stable bundle are scalars, it can be shown that for two stable bundles E_k and E'_k over X, if the corresponding bundles on $X_{\mathbb{C}}$ namely, $E_{\mathbb{C}}$ and $E'_{\mathbb{C}}$, are isomorphic then E_k is isomorphic to E'_k.

Let $\mathcal{M}$ (resp. $\mathcal{M}_{\mathbb{C}}$) be the moduli space of stable vector bundles of rank 2 and degree one over X (resp. $X_{\mathbb{C}}$).

Fix a homomorphism $\sigma : k \to \mathbb{C}$ (throughout the rest of this section). Choose and fix a point $z \in X$. Let $z_{\mathbb{C}}$ be the corresponding point in X. Define $\mathcal{N}$ (resp. $\mathcal{N}_{\mathbb{C}}$) to be the moduli space of stable bundles of the type $V \to X$ (resp. $W \to X_{\mathbb{C}}$) with

$$\operatorname{rank}(V) = 2 \text{ (resp. } \operatorname{rank}(W) = 2) \quad \text{and} \quad \overset{2}{\wedge} V = \mathcal{O}(z) \text{ (resp. } \overset{2}{\wedge} W = \mathcal{O}(z_{\mathbb{C}})).$$

Let J^1 (resp. $J^1_{\mathbb{C}}$) be the component of the Picard group of X (resp. $X_{\mathbb{C}}$) consisting of degree one line-bundles. The morphism $\det : \mathcal{M} \to J^1$ (resp. $\det : \mathcal{M}_{\mathbb{C}} \to J^1_{\mathbb{C}}$) defined by $E \mapsto \overset{2}{\wedge} E$ is a smooth morphism; and $\mathcal{N} = \det^{-1}(\mathcal{O}(z))$ (resp. $\mathcal{N}_{\mathbb{C}} = \det^{-1}(\mathcal{O}(z_{\mathbb{C}}))$).

The moduli space $\mathcal{M}_{\mathbb{C}}$ represents the following functor: To a variety $S/\mathbb{C}$ associate the isomorphism classes of families of stable bundles of rank 2 and degree 1 on $X_{\mathbb{C}}$, parametrized by S, modulo the equivalence relation given by tensoring with line bundles pulled back from S([N1], Theorem 5.12).

Let $\mathcal{E}$ be a universal family over $X \times \mathcal{M}$. From the openness of the subset corresponding to the stable bundles in any given family of bundles ([M], Theorem 2.8), we get that the subset of $\mathcal{M} \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ given by the points on which the restriction of the family $\mathcal{E}_{\mathbb{C}} := \mathcal{E} \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ is a stable bundle, is an open set. So, from Lemma 2.1 it follows that $\mathcal{E}_{\mathbb{C}}$ over $(X \times \mathcal{M}) \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ is a family of stable bundles. Hence there is a (unique) morphism $\varrho_{k,\mathbb{C}} : \mathcal{M} \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C} \to \mathcal{M}_{\mathbb{C}}$ representing this family $\mathcal{E}_{\mathbb{C}}$. Now both $\mathcal{M}$ and $\mathcal{M}_{\mathbb{C}}$ are smooth irreducible varietes of dimension $4g - 3$; and it is easy to see that at any point of $\mathcal{M} \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ the differential of the morphism $\varrho_{k,\mathbb{C}}$ is an isomorphism. So the differential of $\varrho_{k,\mathbb{C}}$ is an isomorphism everywhere. Since $\mathcal{N}_{\mathbb{C}}$ is simply connected, the restriction of $\varrho_{k,\mathbb{C}}$ to $\mathcal{N} \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ must be an isomorphism onto $\mathcal{N}_{\mathbb{C}}$. Now, since $J_{\mathbb{C}} = J \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ we have

Proposition 2.3. *The morphism $\varrho_{k,\mathbb{C}}$ induces an isomorphism between $\mathcal{N} \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$ and $\mathcal{N}_{\mathbb{C}}$. And, more generally, $\varrho_{k,\mathbb{C}}$ is an isomorphism.*

3. Cohomology of the moduli spaces

3 a. Cohomology of $\mathcal{N}_{\mathbb{C}}$. Given an embedding $\sigma : k \to \mathbb{C}$, from Proposition 2.4 we see that the variety $\mathcal{N}_{\mathbb{C}}$ is the moduli space of stable bundles of fixed determinant over the Riemann surface $X_{\mathbb{C}} := X \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$.

Let $\mathcal{E}_k \to X \times \mathcal{N}$ be a universal bundle. So

$$\mathcal{E}_{\mathbb{C}} := \mathcal{E}_k \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$$

is a universal bundle on $X_{\mathbb{C}} \times \mathcal{N}_{\mathbb{C}}$. Since any two universal bundles differ by tensoring with a line bundle, the endomorphism bundle $\operatorname{End}(\mathcal{E}_k)$ (and similarly $\operatorname{End}(\mathcal{E}_{\mathbb{C}})$) does not depend upon the choice of the universal bundle. Let

$$c_2 \in H^4(X_{\mathbb{C}} \times \mathcal{N}_{\mathbb{C}}, \mathbb{Q})$$

be the 2-nd Chern class of the bundle $\operatorname{End}(\mathcal{E}_{\mathbb{C}})$.

Let $1 \in H^0(X_\mathbb{C}, \mathbb{Q})$ and $[X_\mathbb{C}] \in H^2(X_\mathbb{C}, \mathbb{Q})$ be the positive generators. Consider the Künneth decomposition

$$c_2 := \alpha \otimes [X_\mathbb{C}] + \psi + \beta \otimes 1$$

of $c_2 \in H^4(X_\mathbb{C} \times \mathcal{N}_\mathbb{C}, \mathbb{Q})$ as in [N2], [ST] (α, β and ψ here differ from those in [N2] and [ST] by multiplication with rational numbers). Using the dual pairing between homology and cohomology, the cohomology class ψ can be thought of as an homomorphism from $H_1(X_\mathbb{C}, \mathbb{Q})$ into $H^3(\mathcal{N}_\mathbb{C}, \mathbb{Q})$. The classes α, β and the image of the homomorphism ψ together generate the cohomology algebra $H^*(\mathcal{N}_\mathbb{C}, \mathbb{Q})$ ([N2]).

For $r \geqq 0$, the homomorphism ψ induces the following composition of maps:

$$\text{(3.1)} \qquad \overset{r}{\wedge} \psi : \overset{r}{\wedge} H_1(X_\mathbb{C}, \mathbb{Q}) \to \overset{r}{\wedge} H^3(\mathcal{N}_\mathbb{C}, \mathbb{Q}) \to H^{3r}(\mathcal{N}_\mathbb{C}, \mathbb{Q}).$$

The cup-product on $H^1(X_\mathbb{C}, \mathbb{Q})$ gives a symplectic form $\omega \in \overset{2}{\wedge} H_1(X_\mathbb{C}, \mathbb{Q})$. Denote

$$\gamma := (\overset{2}{\wedge} \psi)(\omega) \in H^6(\mathcal{N}_\mathbb{C}, \mathbb{Q}).$$

It is known that the homomorphism ψ is an isomorphism ([MN]), Proposition 1).

Define $V_\mathbb{Q} := H_1(X_\mathbb{C}, \mathbb{Q})$ to be the $\mathbb{Q}$-vector space. For $i, j, k, l \geqq 0$, let $\alpha^i \beta^j \gamma^k \overset{l}{\wedge} \psi$ denotes the pair $(\overset{l}{\wedge} V_\mathbb{Q}, F_{i,j,k,l})$ consisting of the vector space $\overset{l}{\wedge} V_\mathbb{Q}$ and the homomorphism

$$\text{(3.2)} \qquad F_{i,j,k,l} : \overset{l}{\wedge} V_\mathbb{Q} \to H^{2i+4j+6k+3l}(\mathcal{N}_\mathbb{C}, \mathbb{Q})$$

defined by $v \mapsto \alpha^i \beta^j \gamma^k (\overset{l}{\wedge} \psi)(v)$, where $\overset{l}{\wedge} \psi$ is defined in 3.1.

Let us recall Proposition 5.1 and Proposition 5.2 of [ST]. (The combination of Proposition 5.1 and Proposition 5.2 of [ST] is equivalent to the Proposition 3.3 of [KN].)

Proposition 3.3 ([ST], Propositions 5.1, 5.2). *The homomorphism*

$$\left(\bigoplus_{a,b,l}^{a+b+2l<g-1} \alpha^a \beta^b \overset{2l}{\wedge} \psi \right) \oplus \left(\bigoplus_{a,b,c,l}^{a+b+c+2l=g-1} \alpha^a \beta^b \gamma^c \overset{2l}{\wedge} \psi \right)$$

surjects onto the even cohomology $H^{\mathrm{ev}}(\mathcal{N}_\mathbb{C}, \mathbb{Q})$. Moreover, the homomorphism obtained by restricting of the above homomorphism imposing the extra condition that

$$2a + 4b + 6c + 6l \leqq 3g - 3,$$

is an isomorphism onto $\bigoplus_{2i \leqq 3g-3} H^{2i}(\mathcal{N}_\mathbb{C}, \mathbb{Q})$.

Remark 3.4. The dimension of the variety $\mathcal{N}_\mathbb{C}$ is $3g - 3$. So the second part of the Proposition 3.3 (which is Proposition 5.2 of [ST]) says that the above homomorphism is an isomorphism up to half the real dimension of $\mathcal{N}_\mathbb{C}$. The class $\alpha \in H^2(\mathcal{N}_\mathbb{C}, \mathbb{Q})$ gives a polarization on $\mathcal{N}_\mathbb{C}$. Hence from the Hard Lefschetz Theorem, [GH], p.122, we know that for $i \leqq 3g-3$, the homomorphism $H^i(\mathcal{N}_{\mathbb{C}}, \mathbb{Q}) \to H^{6g-6-i}(\mathcal{N}_{\mathbb{C}}, \mathbb{Q})$ given by cupping with α^{3g-3-i} is an isomorphism. So if we choose a $\mathbb{Q}$-basis of $\bigoplus_{2i \leqq 3g-3} H^{2i}(\mathcal{N}_{\mathbb{C}}, \mathbb{Q})$ then, using the cup product with powers of α, we obtain a $\mathbb{Q}$-basis for $H^{ev}(\mathcal{N}_{\mathbb{C}}, \mathbb{Q})$.

Now we will reformulate Proposition 3.3 making it suitable for our use in the next section.

Define $\alpha' := (2\pi\sqrt{-1})\alpha$ and $\beta' := (2\pi\sqrt{-1})^2\beta$ and $\gamma' := (2\pi\sqrt{-1})^3\gamma$ and

$$\psi' := (2\pi\sqrt{-1})\psi .$$

Choose and fix a $\mu \in \mathbb{C}$ such that $\mu^2 = 2\pi\sqrt{-1}$. So ψ' gives a homomorphism from $H_1(X_{\mathbb{C}}, \mu\mathbb{Q})$ to $H^3(\mathcal{N}_{\mathbb{C}}, \mu^3\mathbb{Q})$ – which in turn, as in (3.1), induces a homomorphism

$$\stackrel{r}{\wedge} \psi' : \stackrel{r}{\wedge} H_1(X_{\mathbb{C}}, \mu\mathbb{Q}) \to H^{3r}(\mathcal{N}_{\mathbb{C}}, \mu^{3r}\mathbb{Q}) . \tag{3.5}$$

Define $V'_{\mathbb{Q}} := H_1(X_{\mathbb{C}}, \mu\mathbb{Q})$. Imitating the earlier construction, for any $i, j, k, l \geqq 0$ let $\alpha'^i \beta'^j \gamma'^k \stackrel{l}{\wedge} \psi'$ be the pair $(\stackrel{l}{\wedge} V'_{\mathbb{Q}}, F'_{i,j,k,l})$ consisting of the vector space $\stackrel{l}{\wedge} V'_{\mathbb{Q}}$ and the homomorphism, $F'_{i,j,k,l}$, of it into $H^n(\mathcal{N}_{\mathbb{C}}, \mu^n\mathbb{Q})$ with $n = 2i + 4j + 6k + 3l$, defined by replacing α, β, γ and ψ in (3.2) by α', β' etc.

We have the following corollary of Proposition 3.3.

Corollary 3.6. *The homomorphism*

$$\left(\bigoplus_{a,b,l}^{a+b+2l<g-1} \alpha'^a \beta'^b \stackrel{2l}{\wedge} \psi' \right) \oplus \left(\bigoplus_{a,b,c,l}^{a+b+c+2l=g-1} \alpha'^a \beta'^b \gamma'^c \stackrel{2l}{\wedge} \psi' \right)$$

surjects onto $\bigoplus_n H^{2n}(\mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^n\mathbb{Q})$. *Moreover, the restriction of the above homomorphism imposing the extra condition that* $2a + 4b + 6c + 6l \leqq 3g - 3$, *is an isomorphism onto* $\bigoplus_{2n \leqq 3g-3} H^{2n}(\mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^n\mathbb{Q})$.

As in Remark 3.4, the Hard Lefschetz Theorem implies that the homomorphism

$$\bigcup \alpha^{3g-3-2i} : H^{2i}(\mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^i\mathbb{Q}) \to H^{6g-6-2i}(\mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^{3g-3-i}\mathbb{Q})$$

is an isomorphism when $2i \leqq 3g - 3$.

3b. The de Rham cohomology of $\mathcal{N}$. In the introduction we briefly defined the de Rham cohomology of a variety. Denote the bundle $\mathrm{End}(\mathscr{E}_k)$, where $\mathscr{E}_k$ is the universal bundle on $X \times \mathcal{N}$, by W_k. Let

$$C_2(W_k) \in H^4_{\mathrm{DR}}(X \times \mathcal{N})$$

be the 2-nd Chern class of the bundle W_k. (We will interchangeably consider the i-th Chern class as a co-dimension i cycle modulo rational equivalence.) Recall the definition of Chern classes [Gr], 2.2, [D], p. 21. The first Chern class of a line bundle L on a variety Y/k is defined to be the image of the element in $H^1(Y, \mathcal{O}_Y^*)$ defining L, under the homomorphism of hypercohomologies induced by the following morphism of complexes on Y

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{O}_Y^* & \longrightarrow & 0 & \longrightarrow & \cdots \\ \downarrow & & \downarrow \mathrm{dlog} & & \downarrow & & \\ \mathcal{O}_Y & \xrightarrow{d} & \Omega_Y^1 & \xrightarrow{d} & \Omega_Y^2 & \xrightarrow{d} & \cdots \end{array}$$

where $\mathrm{dlog}(f) := df/f$. Chern classes of a vector bundle are defined using the splitting principle. Note that for a bundle V on a smooth variety Y/k and an embedding $\sigma: k \to \mathbb{C}$, the j-th (rational) Chern class $c_j(V_{\mathbb{C}}) \in H^{2j}(Y_{\mathbb{C}}, \mathbb{Q})$ of the bundle

$$V_{\mathbb{C}} := V \times_{\mathrm{Spec}\, k} \mathrm{Spec}\, \mathbb{C} \to Y_{\mathbb{C}} := Y \times_{\mathrm{Spec}\, k} \mathrm{Spec}\, \mathbb{C}$$

over the manifold $Y_{\mathbb{C}}$, coincides with $C_j(V) \otimes (2\pi\sqrt{-1})^{-i}$ under the identification

$$H^{2j}_{\mathrm{DR}}(Y) \otimes_{k,\sigma} \mathbb{C} = H^{2j}(Y_{\mathbb{C}}, \mathbb{C}).$$

Denote by α_k the element $p_{2*}(C_2(W_k)) \in H^2_{\mathrm{DR}}(\mathcal{N})$, where p_{2*} is the Gysin homomorphism. Define $\beta_k \in H^4_{\mathrm{DR}}(\mathcal{N})$ to be the pull-back $g^*(C_2(W_k))$, where $g: \mathcal{N} \to X \times \mathcal{N}$ is the embedding defined by $y \mapsto (z, y)$; z being the base point in X. Define

$$\gamma_k := p_{2*}(C_2(W_k) \cup C_2(W_k)) - 2\alpha_k \cup \beta_k. \tag{3.7}$$

Let $\chi: H^*_{\mathrm{DR}}(\mathcal{N}) \otimes_{k,\sigma} \mathbb{C} \to H(\mathcal{N}_{\mathbb{C}}, \mathbb{C})$ be the canonical isomorphism.

Proposition 3.8. *$\chi(\alpha_k \otimes 1) = \alpha$, $\chi(\beta_k \otimes 1) = \beta'$ and $\chi(\gamma_k \otimes 1) = \gamma'$.*

Proof. Chern classes commute with pull-back of bundles. Hence the class

$$C_2(W_k) \in H^4_{\mathrm{DR}}(X \times \mathcal{N})$$

corresponds to $(2\pi\sqrt{-1})^2 c_2 \in H^4(X_{\mathbb{C}} \times \mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^2 \mathbb{Q})$. Also the Gysin map commutes with base change. Hence, after examining the constants in the definition of α' and β' we conclude that $\chi(\alpha_k \otimes 1) = \alpha'$ and $\chi(\beta_k \otimes 1) = \beta'$. In order to prove $(\gamma_k \otimes 1) = \gamma'$ it is enough to show that γ' satisfies a similar identity as (3.7). More precisely, if

$$p_*: H^{j+2}(X_{\mathbb{C}} \times \mathcal{N}_{\mathbb{C}}, \mathbb{C}) \to H^j(\mathcal{N}_{\mathbb{C}}, \mathbb{C}), \quad j \geqq 0,$$

is the homomorphism given by integration along fibers multiplied by $(2\pi\sqrt{-1})^{-1}$, then we need to show that

$$\gamma' = p_*((2\pi\sqrt{-1})^2.c_2 \cup (2\pi\sqrt{-1})^2.c_2) - 2.\alpha' \cup \beta'.$$

After expressing the Künneth components of c_2 in terms of a symplectic basis of $H^1(X_{\mathbb{C}}, \mathbb{Q})$ it is a routine calculation to verify the above equality. □

Consider $\psi_k : H^1_{DR}(X) \to H^3_{DR}(\mathcal{N})$ defined by $\theta \mapsto p_{2*}(p_1^*(\theta) \cup C_2(W_k))$, where p_1 is the projection onto X. For integers $a, b, c, d \geqq 0$, as in Section 3a, let $\alpha_\beta^a \beta_k^b \gamma_k^c \overset{d}{\wedge} \psi_k$ denote the pair

$$(\overset{d}{\wedge} H^1_{DR}(X),\ F^k_{a,b,c,d})$$

consisting of the k-vector space $\overset{d}{\wedge} H^1_{DR}(X)$ and the homomorphism $F^k_{a,b,c,d}$ of it into $H^{2a+4b+6c+3d}_{DR}(\mathcal{N})$, which is defined by replacing α', β' etc. in the definition of $F'_{a,b,c,d}$ (in Section 3a) by α_k, β_k etc. (We identify $H^1(X_\mathbb{C}, \mathbb{Q})$ with $H_1(X_\mathbb{C}, \mathbb{Q})$ using the cup product (which is non-degenerate) in $H^1(X_\mathbb{C}, \mathbb{Q})$.)

Let $\chi' : H^*_{DR}(\mathcal{N}) \to H^*(\mathcal{N}_\mathbb{C}, \mathbb{C})$ be the k-linear injective homomorphism defined by $v \mapsto \chi(v \times 1)$. So $\chi = \chi' \otimes \mathrm{Id}$. Similarly, define $\tau : H^1_{DR}(X) \to H^1(X_\mathbb{C}, \mathbb{C})$ to be the canonical injection. As before, $H^1(X_\mathbb{C}, \mathbb{C})$ is identified with $H_1(X_\mathbb{C}, \mathbb{Q})$ using the cup product. With this notation it is easy to see that the following diagram commutes:

$$(3.9)\qquad \begin{array}{cccccc} F^k_{a,b,c,d} & : & \overset{d}{\wedge} H^1_{DR}(X) & \longrightarrow & H^{2a+4b+6c+3d}_{DR}(\mathcal{N}) \\ & & \downarrow \overset{d}{\wedge} & & \downarrow \chi' \\ F'_{a,b,c,d} \otimes_\mathbb{Q} \mathbb{Q} & : & \overset{d}{\wedge} H^1(X_\mathbb{C}, \mathbb{C}) & \longrightarrow & H^{2a+4b+6c+3d}(\mathcal{N}_\mathbb{C}, \mathbb{C})\,. \end{array}$$

Now Corollary 3.6 and the commutativity of (3.9) together imply

Proposition 3.10. *The homomorphism*

$$\left(\bigoplus_{a,b,l}^{a+b+2l<g-1} \alpha_k^a \beta_k^b \overset{2l}{\wedge} \psi_k\right) \oplus \left(\bigoplus_{a,b,c,l}^{a+b+c+2l=g-1} \alpha_k^a \beta_k^b \gamma_k^c \overset{2l}{\wedge} \psi_k\right)$$

surjects onto $\bigoplus_n H^{2n}_{DR}(\mathcal{N})$. *Moreover, the restriction of the above homomorphism imposing the extra condition that* $2a + 4b + 6c + 6l \leqq 3g - 3$, *is an isomorphism onto*

$$\bigoplus_{2n \leqq 3g-3} H^{2n}(\mathcal{N}_\mathbb{C}, (2\pi\sqrt{-1})^n \mathbb{Q})\,.$$

The Hard Lefschetz Theorem holds for the de Rham cohomology. So, in particular, for $2i \leqq 3g - 3$, the homomorphism

$$\bigcup \alpha_k^{3g-3-2i} : H^{2i}_{DR}(\mathcal{N}) \to H^{6g-6-2i}_{DR}(\mathcal{N})$$

is actually an isomorphism.

4. Absolute Hodge cycles on $\mathcal{N}$

We use the notations in the introduction. Let $\theta \in H^{2i}_{DR}(\mathcal{N})$ be a Hodge cycle relative to an embedding σ of k in $\mathbb{C}$. Denote the isomorphism from $H^n_{DR}(\mathcal{N}) \otimes_{k,\sigma} \mathbb{C}$, $n \geqq 0$, onto $H^n(\mathcal{N}_\mathbb{C}, \mathbb{C})$, given by the obvious map, by δ. So the condition that θ is a Hodge cycle relative to σ implies that $\delta(\theta) \in H^{2i}(\mathcal{N}_\mathbb{C}, (2\pi\sqrt{-1})^i \mathbb{Q})$. Now, if $2i \leqq 3g - 3$ then from the second part of the Corollary 3.6 the cohomology class $\delta(\theta)$ can be uniquely expressed as follows:

$$\delta(\theta) = \sum_{\lambda} \alpha'^{a_\lambda} \beta'^{b_\lambda} \gamma'^{c_\lambda} (\bigwedge^{d_\lambda} \psi')(m_\lambda), \tag{4.1}$$

where each d_λ is a non negative even integer, $m_\lambda \in \bigwedge^{d_\lambda} H_1(X_{\mathbb{C}}, \mu\mathbb{Q})$, and the indices $\{a_\lambda, \ldots, d_\lambda\}$ are chosen from the set of indices prescribed by the first part of the Corollary 3.6. Moreover if $2i > 3g-3$ then from the remark following Corollary 3.6, $\delta(\theta)$ can again be expressed uniquely as in (4.1) with the indices chosen from a different set of indices. So the (unique) expression (4.1) for $\delta(\theta)$ holds for all $i \geqq 0$. Note that d_λ is always an even integer.

Using the non-degenerate form on $H_1(X_{\mathbb{C}}, \mathbb{Q})$ given by the cap product, the $\mathbb{Q}$-vector space $H_1(X_{\mathbb{C}}, \mathbb{Q})$ is canonically identified with $H^1(X_{\mathbb{C}}, \mathbb{Q})$ – which in turn induces an isomorphism between $H_1(X_{\mathbb{C}}, \mu\mathbb{Q})$ and $H^1(X_{\mathbb{C}}, \mu\mathbb{Q})$. Using this isomorphism, the element $m_\lambda \in \bigwedge^{d_\lambda} H_1(X_{\mathbb{C}}, \mu\mathbb{Q})$ can be thought of as an element of $\bigwedge^{d_\lambda} H^1(X_{\mathbb{C}}, \mu\mathbb{Q})$. The cohomology algebra $\bigoplus_j H^j(J_{\mathbb{C}}, \mu^j\mathbb{Q})$ of the Jacobian of the curve $X_{\mathbb{C}}$ is canonically isomorphic to the exterior algebra $\bigwedge H^1(X_{\mathbb{C}}, \mu\mathbb{Q})$. In particular, $\bigwedge^{d_\lambda} H^1(X_{\mathbb{C}}, \mu\mathbb{Q})$ is isomorphic to

$$H^{d_\lambda}(J_{\mathbb{C}}, (2\pi\sqrt{-1})^{d_\lambda/2}\mathbb{Q}).$$

Let

$$\bar{m}_\lambda \in H^{d_\lambda}(J_{\mathbb{C}}, (2\pi\sqrt{-1})^{d_\lambda/2}\mathbb{Q})$$

be the image of m_λ under this isomorphism.

Let J_k be the Jacobian of the curve X. So $J_{\mathbb{C}} = J_k \times_{\operatorname{Spec} k} \operatorname{Spec} \mathbb{C}$. We have the isomorphism

$$f_j : H^j_{\mathrm{DR}}(J_k) \otimes_{k,\sigma} \mathbb{C} \to H^j(J_{\mathbb{C}}, \mathbb{C})$$

induced by the projection $J_{\mathbb{C}} \to J_k$.

Lemma 4.2. *The cohomology class $\bar{m}_\lambda$ is the image of an absolute Hodge cycle on J_k. In other words, there is an absolute Hodge cycle ω_λ on J_k such that $f_{d_\lambda}(\omega_\lambda \otimes 1) = \bar{m}_\lambda$.*

Proof. In view of the main theorem (Theorem 2.11 of [D]), it is enough to show that there is a Hodge cycle ω_λ on J_k relative to σ such that $f_{d_\lambda}(\omega_\lambda \otimes 1) = \bar{m}_\lambda$. From Proposition 3.10 and the remark following it, it follows that the class $\theta \in H^{2i}_{\mathrm{DR}}(\mathcal{N})$ can uniquely be expressed as

$$\theta = \sum_{v} \alpha_k^{p_v} \beta_k^{q_v} \gamma_k^{r_\lambda} (\bigwedge^{s_v} \psi_k)(u_v) \tag{4.3}$$

where $u_v \in \bigwedge^{s_v} H^1_{\mathrm{DR}}(X)$, and the indices (p_v, q_v, r_v, s_v) are chosen from the set of indices prescribed by the second part of Proposition 3.10 and the remark following the proposition. If we consider θ as an element of $H^{2i}(\mathcal{N}_{\mathbb{C}}, \mathbb{C})$ (using δ) and, similarly, consider u_v as an element of $\bigwedge^{s_v} H^1(X_{\mathbb{C}}, \mathbb{C})$, then from the commutativity of (3.9) it follows that (4.3) is an expression for $\delta(\theta)$. Now by the uniqueness of the expression (4.1) it follows that (4.3) must coincide with (4.1). In particular, any m_λ is of the form $(\overset{d_\lambda}{\wedge}\tau)(m'_\lambda)$ with

$$m'_\lambda \in \overset{d_\lambda}{\wedge} H^1_{\mathrm{DR}}(X)$$

(τ was defined in Section 3b). Since the cohomology algebra $H^*_{\mathrm{DR}}(J_k)$ is the exterior algebra $\wedge H^1_{\mathrm{DR}}(X)$ any $\bar{m}_\lambda$ must be of the form $f_{d_\lambda}(\omega_\lambda \otimes 1)$. By definition,

$$\bar{m}_\lambda \in H^{d_\lambda}(J_{\mathbb{C}}, (2\pi\sqrt{-1})^{d_\lambda/2}\mathbb{Q}).$$

So, in order to prove that ω_λ is a Hodge cycle on J_k relative to σ it is enough to show that $\bar{m}_\lambda \in H^{d_\lambda/2, d_\lambda/2}(J_{\mathbb{C}})$. This is equivalent to showing that m_λ is of type $(d_\lambda/2,\ d_\lambda/2)$ with respect to the type decomposition of $\wedge H^1(X_{\mathbb{C}}, \mathbb{C})$ induced by the Hodge decomposition of $H^1(X_{\mathbb{C}}, \mathbb{C})$. Note that for any $e \in \overset{p,q}{\wedge} H^1(X_{\mathbb{C}}, \mathbb{C})$ and $a, b, c \geqq 0$,

$$\text{(4.4)} \quad (\alpha'^a \beta'^b \gamma'^c \overset{p+q}{\wedge} \psi' \otimes \mathbb{C})(e) \in H^{a+2b+3c+2p+q,\, a+2b+3c+p+2q}(\mathcal{N}_{\mathbb{C}}).$$

Hence the condition that $\delta(\theta) \in H^{i,i}(\mathcal{N}_{\mathbb{C}})$ and the injectivity of the homomorphism in the second part of Proposition 3.10 (and the remark following it) combine together to imply that any m_λ is of the type $(d_\lambda/2,\ d_\lambda/2)$. This completes the proof. □

Recall the definition of an absolute Hodge cycle given in the introduction (see also [D]).

Theorem 4.5. *Any Hodge cycle on $\mathcal{N}$ relative to an embedding σ of k in $\mathbb{C}$, is an absolute Hodge cycle.*

Proof. As before, let $\theta \in H^{2i}_{\mathrm{DR}}(\mathcal{N})$ be a Hodge cycle relative to an embedding σ of k in $\mathbb{C}$. The decomposition (4.3) is canonical in the sense that it does not depend on the embedding σ. The decomposition (4.1) depends, at least a priori, on σ. The classes α_k, β_k and γ_k in $H^*_{\mathrm{DR}}(\mathcal{N})$ are cycle classes; in particular they are absolute Hodge cycles ([D], Example 2.1 (a)). Also any product of absolute Hodge cycles is again an absolute Hodge cycle. Any ω_λ is an absolute Hodge cycle on J_k (Lemma 4.2). So, (4.4) and the fact that the image of the homomorphism $\overset{2r}{\wedge}\psi'$, defined in (3.5), is in $H^{6r}(\mathcal{N}_{\mathbb{C}}, (2\pi\sqrt{-1})^{3r}\mathbb{Q})$, and the commutativity of (3.9) together imply that any class $\alpha_k^{p_\nu}\beta_k^{q_\nu}\gamma_k^{r_\nu}(\overset{s_\nu}{\wedge}\psi_k)(u_\nu)$ occurring in (4.3) is an absolute Hodge cycle on $\mathcal{N}$. This completes the proof. □

References

[BKN] *V. Balaji, A. King, P. Newstead*, Algebraic cohomology of the moduli space of rank 2 vector bundles on a curve, Preprint, alg-geom/9502019.

[Ba] *V. Baranovsky*, Cohomology ring of the moduli space of stable vector bundles with odd determinant, Izv. Russ. Acad. Nauk. **58** no. 4 (1994), 204–210.

[BN] *I. Biswas, M.S. Narasimhan*, Hodge classes of moduli spaces of parabolic bundles over the general curve, Preprint (1994).

[Bi] *I. Biswas*, A correspondence between moduli spaces of vector bundles over a curve, Preprint 1995.

[D] *P. Deligne* (Notes by J. Milne), Hodge cycles on abelian varieties, Lect. Notes Math. **900** (1982), 9–100.

[GH] *P. Griffiths, J. Harris*, Principles of algebraic geometry, Wiley-Interscience Publication, John Wiley & Sons, 1978.

[Gr] *A. Grothendieck*, La théorie des classes de Chern, Bull. Soc. Math. Fr. **86** (1958), 137–154.

[HN] *G. Harder, M.S. Narasimhan*, On the cohomology groups of moduli spaces of vector bundles over curves, Math. Ann. **212** (1975), 215–248.

[KN] *A. King, P. Newstead*, On the cohomology ring of the moduli space of rank 2 vector bundles on a curve, Preprint, alg-geom/9502018.

[L] *S. Langton*, Valuative criteria for families of vector bundles on algebraic varieties, Ann. Math. **101** (1975), 88–110.

[M] *M. Maruyama*, Openness of a family of torsion free sheaves, J. Math. Kyoto Univ. **16** (1976), 627–637.

[MN] *D. Mumford, P. Newstead*, Periods of a moduli space of bundles on curves, Amer. J. Math. **90** (1968), 1200–1208.

[N1] *P. Newstead*, Introduction to moduli problems and orbit spaces, Tata Inst. Lect. Notes **51** (1978).

[N2] *P. Newstead*, Characteristic classes of stable bundles of rank 2 over an algebraic curve, Trans. AMS **169** (1972), 337–345.

[S] *C. Seshadri* (rédigé par J.-M. Drezet), Fibrés vectoriels sur les courbes algébriques, Astérisque **96** (1982).

[ST] *B. Siebert, G. Tian*, Recursive relations for the cohomology ring of moduli spaces of stable bundles, Turkish J. Math. **19** (1995), 131–144.

[Z] *D. Zagier*, On the cohomology of moduli spaces of rank two vector bundles over curves.

Institut Fourier, Université de Grenoble, 100 rue des Maths, B.P. 74, F-38402 Saint-Martin d'Hères cedex

e-mail: biswas@puccini.ujf-grenoble.fr

Eingegangen 20. Januar 1995, in revidierter Fassung 1. Juli 1995

J. reine angew. Math. 470 (1996), 153—180

Journal für die reine und angewandte Mathematik

Décomposition d'un groupe en produit libre ou somme amalgamée

Par *T. Delzant* à Strasbourg

0.1. Produits libres. Le théorème de Grusko dit que le rang (nombre minimal de générateurs) d'un produit libre est additif:

$$\operatorname{rang}(A * B) = \operatorname{rang}(A) + \operatorname{rang}(B).$$

Le point de départ de cet article est qu'un fait analogue se produit pour les relations.

Soit G un groupe de présentation finie; on définit $T(G)$ comme étant le nombre minimum de relations triangulaires d'une présentation triangulaire de G, c'est-à-dire d'une présentation dont toutes les relations sont de longueur deux ou trois.

Théorème I. *T est additive pour les produits libres*:

$$T(A * B) = T(A) + T(B).$$

Ne prendre en compte que des présentations triangulaires n'est pas restrictif : si G est défini par une présentation quelconque $(h^\alpha, R_i)_{1 \leqq \alpha \leqq n; 1 \leqq i \leqq k}$, un argument de triangulation bien connu montre qu'il admet une présentation triangulaire ayant $\sum (|R_i| - 2)$ relations. Ainsi, $T(G)$ est le minimum, sur toutes les présentations de G de la somme $\sum_{1 \leqq i \leqq k} (|R_i| - 2)$.

Question. On peut définir pour tout groupe G et tout entier i un entier $S_i(G)$, en disant que $S_i(G) \leqq k$ si et seulement si il existe un polyèdre simplicial P i-connexe muni d'une action de G, dont la restriction au 0-squelette est libre, et dont le nombre de faces modulo l'action de G n'excède pas k. On vérifie que $S_1(G)$ est le rang de G, et $S_2(G)$ n'est autre que $T(G)$; le théorème de Grusko dit que S_1 est additif pour les produits libres, et le théorème I qu'il en va de même pour S_2; qu'en est-il de S_i?

Notons que c'est cette définition simpliciale de T que nous utiliserons par la suite.

0.2. Somme amalgamées, graphes de groupes. Contrairement au théorème de Grusko, le théorème I admet une généralisation naturelle au cas des sommes amalgamées, HNN-extensions, et plus généralement graphes de groupes au sens de Serre [S].

Soit G un groupe, et $(C_1, \dots, C_n)$ une famille de ses sous-groupes. On définit un invariant $T(G; C_1, \dots, C_n)$ du couple $(G; C_1, \dots, C_n)$ en disant que $T(G; C_1, \dots, C_n) \leqq k$ si et seulement si il existe un polyèdre simplicial simplement connexe de dimension deux P, muni d'une action de G vérifiant les deux propriétés suivantes:

(a) Le nombre de faces de P modulo l'action de G n'excède pas k.

(b) Les stabilisateurs des sommets de P sont conjugués à $C_1, \dots, C_n$, et chacun de ces groupes fixe un sommet de P.

L'objet de cet article est d'étudier le domaine de validité de l'inégalité:

$$(*) \qquad T(A; C) + T(B; C) \leqq T(A *_C B)$$

et de ses généralisations possibles, puis d'en exhiber quelques applications.

On dit que deux couples $(G; C_1, \dots, C_n)$ et $(H; D_1, \dots, D_n)$ sont *isomorphes* si il existe une permutation σ de $\{1, \dots, n\}$ et un isomorphisme de G sur H envoyant C_i sur *un conjugué* de $D_{\sigma(i)}$.

Dans ce qui suit, on fixe un couple $(G; C_1, \dots, C_n)$ et on suppose tous les groupes C_i finiment engendrés. On note $b_1(G; C_1, \dots, C_n)$ le premier nombre de Betti de G relativement à $(C_1, \dots, C_n)$ modulo 2, c'est-à-dire le rang du Z_2-espace vectoriel des homomorphismes de G à valeur dans Z_2 qui s'annulent sur les C_i.

Une *décomposition* de $(G; C_1, \dots, C_n)$ en *somme amalgamée*, est une écriture $G = A *_C B$ de sorte qu'il existe une partition $\{1, \dots, n\} = I_A \cup I_B$ telle que $C_i \subset A$ si $i \in I_A$ et $C_i \subset B$ sinon.

Plus généralement, nous souhaitons étudier des décompositions de G en graphes de groupes au sens de J.-P. Serre [S]. Rappelons (voir [S] ou [D-D]) qu'un graphe de groupes est la donnée d'un graphe abstrait X, de deux familles G_s et C_y de groupes respectivement indexées par les sommets et les arêtes de X, et pour chaque arête orientée y d'une injection $o: C_y \mapsto G_{o(y)}$, si $o(y)$ désigne l'origine de y. A un tel graphe est associé un groupe appelé *groupe fondamental* au sens de Serre; on le note $\pi_1(X)$.

Une *décomposition* de $(G; C_1, \dots, C_n)$ en *graphe de groupes*, est la donnée d'un graphe de groupes $(X, (G_s)_{s \in X^0}, (C_y)_{y \in X^1})$ de groupe fondamental isomorphe à G de sorte qu'il existe une partition $\{1, \dots, n\} = \bigcup I_s$ telle que $C_i \subset G_s$ pour tout $i \in I_s$.

Pour étudier le domaine de validité de $(*)$, il est nécessaire d'introduire des hypothèses de points fixes et minimalité.

Hypothèses de minimalité. La décomposition $G = A *_C B$ (resp. l'arête orientée y de la décomposition $G = \pi_1(X)$) du groupe équipé $(G; C_1, \dots, C_n)$ est dite *triviale* si $A = C$ ou $B = C$ (resp. l'origine de y est distincte de son extrémité et $o(C_y) = G_{o(y)}$).

La décomposition $G = A *_C B$ (resp. l'arête orientée y de X) est dite *réduite* si l'on ne peut décomposer les groupes équipés $(A; (C_i)_{i \in I_A})$ et $(B; (C_i)_{i \in I_B})$ (resp. le groupe équipé

$(G_{o(y)}; (C_z)_{o(y)=o(z)}, (C_i)_{i \in I_o(y)}))$ en graphe de groupes sans arêtes triviales avec un sommet marqué C (Figure I).

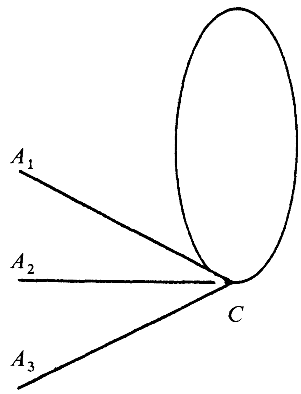


Figure I

Soit y une arête orientée non réduite. Une *réduction élémentaire* de $G = \pi_1(X)$ est une nouvelle décomposition $G = \pi_1(X')$ de $(G; C_1, \ldots, C_n)$ en graphe de groupes obtenue en remplaçant l'origine de y par le graphe de groupe correspondant et en supprimant l'arête triviale du graphe ainsi obtenue (Figure II). Par exemple, si $A = A' *_{C'} C$, alors $G = A' *_{C'} B = (A' *_{C'} C) *_C B = A *_C B$ est une réduction de $G = A *_C B$. Plus généralement, une *réduction* de X s'obtient en répétant cette opération à plusieurs reprises.

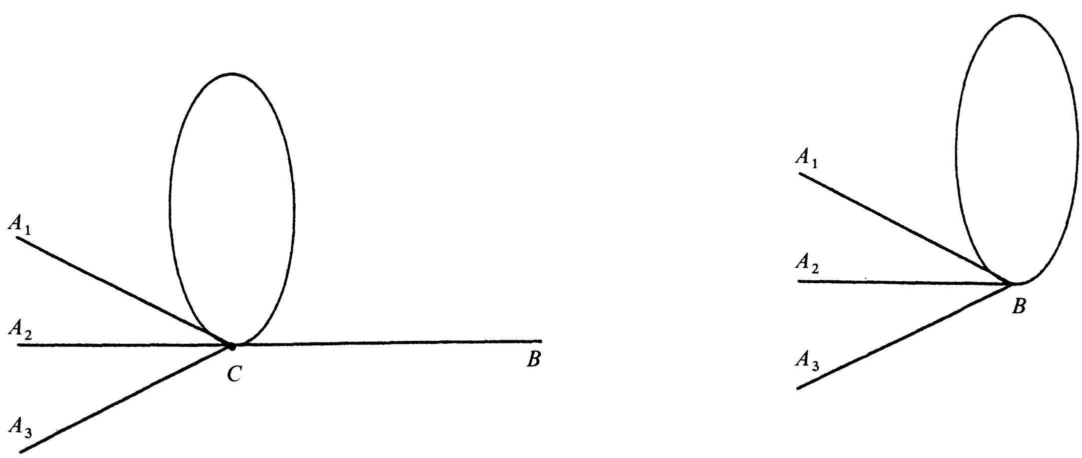


Figure II

Hypopthèses de points fixes. Soit T un arbre; on note $\bar{T}$ sa complétion, c'est-à-dire la réunion de T et de son bord idéal δT qui est l'ensemble des classes d'équivalence des demi-droites de T, où deux demi-droites sont dites équivalentes si elles coïncident en dehors d'un ensemble compact.

Le sous-groupe C de $(G; C_1, \ldots, C_n)$ est dit *rigide* (resp. *semi-rigide*) si le triplet $(G; C; C_1, \ldots, C_n)$ satisfait à l'une des propriétés de point fixe suivantes:

Rigide. Si G agit sans inversion dans un arbre T, si tous les groupes C_i fixent un sommet de T et si C contient le stabilisateur d'une arête non triviale, alors C fixe un sommet de T.

Semi-rigide. Si G agit sans inversion dans un arbre T, si tous les groupes C_i fixent un sommet de T et si C contient le stabilisateur d'une arête non triviale alors C fixe un point de la complétion $\bar{T}$.

La décomposition $G = A *_C B$ (resp. $G = \pi_1(X)$) est dite rigide ou semi-rigide si le groupe C (resp. C_y pour toute arête y) l'est.

Remarque. Les propriétés de rigidité ou semi-rigidité sont en particulier vérifiées si le groupe C satisfait lui même une propriété de point fixe indépendament du fait qu'il soit contenue dans G. Ainsi, notre définition de «semi-rigide» contient la notion de «petite» au sens de M. Bestvina et M. Feighn [B-F]. Par exemple si le groupe C ne contient pas le groupe libre à deux générateurs, il est «petit» et de ce fait satisfait la propriété de point fixe asymptotique. L'interêt principal de notre définition est de regrouper la définition de [B-F], et «rigide», qui est une hypothèse très naturelle en topologie de petite dimension (I.6).

Exemple. Un autre exemple particulièrement important est le suivant: soit $G = A *_C B$, et supposons que ni A ni B ne se décomposent comme somme amalgamée ou HNN-extension au-dessus d'un sous groupe strict de C; alors la décomposition $G = A *_C B$ est réduite et rigide; en particulier une décomposition d'un groupe en produit libre est réduite et rigide. Ce même phénomène de minimalité apparait en topologie de dimension deux ou trois quand l'on découpe une surface le long d'une courbe fermée simple non homotope à zéro ou une variété de dimension trois irréductible le long d'une surface incompressible (I.6).

Théorème II. *Soit $G = A *_C B$ une décomposition semi-rigide de $(G; C_1, \ldots, C_n)$. Supposons que $I_A \cup I_B$ est une partition de l'ensemble $\{1, \ldots, n\}$ pour laquelle $C_i \subset A$, $i \in I_A$ et $C_i \subset B$, $i \in I_B$.*

(1) *Si cette décomposition est réduite*:

$$T(A; C, (C_i)_{i \in I_A}) + T(B; C, (C_i)_{i \in I_B}) \leqq T(G; C_1, \ldots, C_n).$$

(2) *Si cette décomposition n'est pas réduite, il existe des sous-groupes $(D_1, \ldots, D_k)$ de C pour lesquels soit*

$$T(A; (C_i)_{i \in I_A}) + T(B; (D_j)_{1 \leqq j \leqq k}, (C_i)_{i \in I_B}) \leqq T(G; C_1, \ldots, C_n)$$

soit

$$T(A; (C_i)_{i \in I_A}, (D_j)_{1 \leqq j \leqq k}) + T(B; (C_i)_{i \in I_B}) \leqq T(G; C_1, \ldots, C_n).$$

(3) *Plus généralement, soit $G = \pi_1(X)$ une décomposition semi-rigide de $(G; C_1, \ldots, C_n)$. Il existe une réduction $(X'; (H_s), (D_y))$ de X pour laquelle*:

$$\sum T(H_s; (D_y)_{o(y)=s}, (C_i)_{i \in I_s}) \leqq T(G; C_1, \ldots, C_n).$$

Quel que soit le couple $(G; C)$, une décomposition de la forme $G = A *_C B$ (resp. $G = A *_C$) de (G, C) est rigide; compte-tenu des propriétés élémentaires de $T(.;.)$ on obtient (voir § I.4):

Corollaire I. *$T(G; C)$ est additif: si $G = A *_C B$,*

$$T(G; C) = T(A; C) + T(B; C).$$

Le corollaire suivant permet de donner une réponse à une question traditionelle de la théorie combinatoire des groupes, au moins pour les cas des décompositions semi-rigides : si un groupe finiment engendré G est une somme amalgameé, $G = A *_C B$, les groupes A et B sont aussi finiment engendrés; peut-on donner une borne sur leurs rangs qui ne dépende que des groupes G et C? (Cette question est notament étudiée dans [L-S], p. 91, [Z], [P-R-Z], [K-Z].) Des propriétés élémentaires de $T(.;.)$ et du (2) du th. II, on déduit (voir I.2):

Corollaire II. *Si $G = A *_C B$ est une décomposition semi-rigide du groupe G (par exemple si C ne contient pas le groupe libre à deux générateurs),*

$$\operatorname{rang}(A) + \operatorname{rang}(B) \leqq 2\operatorname{rang}(C) + 12\,T(G) + b_1(G).$$

La sous-additivité de la fonction $T(.;.)$ peut aussi (voir I.5) être mise à profit pour préciser le théorème de M.J. Dunwoody sur le problème d'accessibilité de C.T.C. Wall [W], [D], [D-D], ainsi que ses généralisations dûes à M. Bestvina et M. Feighn [B-F], M.J. Dunwoody et R.A. Fenn [D-F].

Corollaire III. *Soit $G = \pi_1(X)$ une décomposition semi-rigide de $(G; C_1, \ldots, C_n)$ en graphe de groupes sans arêtes triviales. Le nombre des sommets de X est majoré par $4\big(T(G; C_1, \ldots, C_n) + 2b_1(G; C_1, \ldots, C_n) + n\big)$.*

L'analogie de ces résultats avec des énoncés bien connus de la topologie de petite dimension n'est pas accidentelle : rappelons (voir le § I.6) que si M est une variété de dimension trois P^2-irréductible et S une surface incompressible (non nécessairement connexe mais dont deux composantes ne sont pas parallèles), alors la décomposition de $\pi_1(M)$ obtenue en découpant M le long de S est rigide. Soit $t(M)$ le nombre minimal de 3-simplexes d'une triangulation de M. Il est facile de voir que $T(\pi_1(M)) \leqq 2t(M)$; le corollaire III est donc une généralisation *algébrique* du théoreme de Kneser et Haken qui montre que le nombre maximal de composantes connexes de $M - S$ est majoré par $(2b_1(M) + 6t(M))$ ([He], p. 140; [J], Th. III.20).

Pour illustrer le corollaire I, comparons-le au théorème de Schubert sur l'additivité du genre des noeuds ([B-Z]) : grâce à la prop. I.1, si $T(A; C) = b_1(A; C) = 0$ alors $A = C$; par exemple, si $G = (A_i)_{1 \leqq i \leqq n} *_C$ est obtenue en amalgamant n groupes le long d'un même groupe C et si pour tout i, $A_i \neq C$, alors $n \leqq T(G; C) + b_1(G; C)$ (c'est essentiellement le même argument qui permet de déduire le cor. III du th. II). Ainsi, la sous-additivité de $T(A; C)$ (donc celle de $T(A, C) + b_1(A; C)$), fournit immédiatement un équivalent *algébrique* du résultat principal de [D-F] sur la finitude des sommes de noeuds en grande dimension.

Appelons *décomposition terminale* d'un groupe G une décomposition de ce groupe en graphe de groupes à stabilisateurs d'arêtes *finis*, sans arêtes triviales et à stabilisateurs de sommets indécomposables en somme amalgamée ou HNN-extension au dessus d'un groupe fini. Dans le langage de [D-D], l'arbre de Serre associé a une telle décomposition est *terminal*, et réciproquement tout arbre terminal détermine une telle décomposition. L'existence d'une décomposition terminale d'un groupe de présentation finie est (assurée par) le théorème d'accessibilité de M.J. Dunwoody. En général, il n'y a pas unicité; voir les résultats de W. Dicks et M.J. Dunwoody, D. Kramtsov et F. Herrlich ([D-D], prop. IV, 7.4, [K], [He]) sur ce problème, posé par P. Scott et C. Wall ([S-W], p. 193). Cependant, (voir le lemme I.2.1) il est aisé de voir qu'étant donné un groupe fini Φ, il n'y a qu'un nombre fini de classes d'isomorphismes de paires $(G; F)$ avec F isomorphe à Φ, et $T(G; F) + b_1(G; F)$ donné. D'autre part une décomposition terminale est réduite par définition; on en déduit (voir I.2):

Corollaire IV. (a) *Soit G un groupe de présentation finie. A isomorphisme près il n'y a qu'un nombre fini de décompositions réduites de G comme somme amalgamée $G = A *_F B$ ou* HNN-*extension $G = A *_F$ avec F fini.*

(b) *A isomorphisme près, il n'y a qu'un nombre fini de décompositions terminales de G.*

Un phénomène analogue, appelé «superaccessibilité», a été mis en évidence par Z. Sela [Se] dans son étude des décompositions *acylindriques* de certains groupes; cependant une décomposition terminale peut très bien être cylindrique au sens de [Se].

Plan. Cet article est divisé en trois partie. Dans la première sont établies les propriétés élémentaires de $T(.;.)$. Admettant le théorème II dans le cas rigide, on en déduit le cas général (semi-rigide) et les corollaires de l'introduction. Dans la seconde partie est démontré un lemme technique important (II.1.1). Celui-ci permet d'obtenir aisément le théorème I, en utilisant un théorème d'unicité de Kuròs; en particulier un lecteur seulement intéressé par ce théorème I pourra se contenter de lire le paragraphe II. La troisième partie consiste à montrer que le raisonnement qui permet d'obtenir le th. I se généralise au cas des graphes de groupes *rigides*; la difficulté essentielle est qu'on ne possède pas de théorème d'unicité pour de telles décompositions.

Remerciements. Je remercie M. Lustig pour d'utiles conversations sur cet article, G. Levitt qui m'a fait remarquer que le théorème d'accessibilité de [D] et [B-F] devait résulter d'une version anterieure de cette étude, E. Ghys qui m'a incité à traiter les exemples venant de la dimension 3, F. Laudenbach, V. Kharlamov et E. Giroux pour leurs encouragements.

Mais plus encore, il me faut indiquer que c'est sur l'insistance judicieuse du «referee» que j'ai opté pour une démonstration géométrique du théorème II alors qu'une démonstration algébrique m'avait paru plus naturelle dans un premier temps. En particulier la démonstration topologique du lemme II.1.1, central dans cet article, est celle suggérée par le referee. La première version présentée était algébrique, en terme de générateurs et relations, et évidement moins parlante : qu'il en soit chaleureusement remercié.

I. Préliminaires

Dans ce paragraphe on établit, en utilisant les propriétés bien connues des *orbièdres* pour lesquelles nous renvoyons à l'article fondamental de A. Haefliger [Ha], certaines propriétés élémentaires utiles de $T(.;.)$, puis on en déduit les corollaires I, II et IV de l'introduction. Ces résultats élémentaires ne sont utiles qu'à la démonstration du théorème II; ainsi leur étude peut être omise par un lecteur seulement intéressé par le théorème I.

Notations. On fixe un groupe G et une famille $(C_1, \dots, C_n)$ de sous-groupes. Si la famille des C_i est vide, on convient de poser $n = 1$ et $C_1 = \{1\}$. Rappelons que si $T(G; C_1, \dots, C_n) = t$ il existe un polyèdre simplicial simplement connexe P muni d'une action de G ayant t orbites de faces et dont les stabilisateurs de sommets sont exactement les groupes $C_1, \dots, C_n$ et leurs conjugués. Soit P' la première subdivision barycentrique de P, de sorte que G agit *sans inversion* sur P' au sens de [Ha]. Le quotient $Q = P'/G$ est un polyèdre simplicial de groupes ayant G comme groupe fondamental au sens de la théorie des orbièdres; ce polyèdre est developpable, et son revêtement universel est P'.

I.1 Propriétés. Dans ce qui suit, (VC_i) désigne un sous-groupe de rang minimal contenant tous les C_i.

Lemme I.1.1. (1) $\operatorname{rang}(G) \leqq \operatorname{rang}(VC_i) + 12.T(G; C_1, \dots, C_n) + b_1(G; C_1, \dots, C_n)$.

(2) $T(G; C_1, \dots, C_n) = 0$ *si et seulement si il existe une action de G dans un arbre et dont les stabilisateurs de sommets sont (exactement) les groupes* $C_1, \dots, C_n$ *et leurs conjugués.*

(3) *Si* $T(G; C_1, \dots, C_n) = 0 = b_1(G; C_1, \dots, C_n)$, *G est isomorphe au groupe fondamental d'un graphe de groupes dont les stabilisateurs de sommets sont les* C_i, *et le graphe sous-jacent un arbre.*

Remarque I.1.2. Soulignons un détail technique important dans la définition de T: on n'exige pas que l'action de G sur P soit *sans inversion* au sens de [Ha].

Démonstration. Démontrons le point (1). Soit T un arbre maximal tracé dans le 1-squelette de l'orbièdre Q. Grâce a la définition de Haefliger ([Ha], p. 514) de $\pi_1^{\mathrm{orb}}(Q, T)(= G)$, on se rend compte que G est engendré par les groupes C_i (et donc (VC_i)) et les arêtes de $Q - T$. Mais le nombre d'arêtes de $Q - T$ qui ne sont pas adjacentes à une face de Q n'excède pas $b_1(G; C_1, \dots, C_n) + n$: en effet les arêtes adjacentes à aucunes faces de $Q - T$ qui n'aboutissent pas à l'un des sommets marqués C_i définissent des homomorphismes vers Z_2 indépendants et nuls sur les C_i; ainsi, comme le nombre d'arêtes d'un triangle subdivisé est 12, le nombre total d'arêtes de $Q - T$ n'excède pas $12\,T(G; C_1, \dots, C_n) + b_1(G; C_1, \dots, C_n)$, d'où (1).

Le point (2) résulte de la définition de T et du fait qu'un polyèdre simplicial simplement connexe de dimension 2 sans face est un arbre.

Pour voir (3), on peut raisonner ainsi : soit T l'arbre obtenu grâce à (2). Comme tous les C_i fixent un point dans T et comme $b_1(G; C_1, \dots, C_n) = 0$, G agit sans inversion dans cet arbre. Soit $X = T/G$ le graphe de groupe quotient : c'est un arbre car $b_1(G; C_1, \dots, C_n) = 0$ et tous les C_i fixent un points dans T. D'autre part si l'un des C_i, disons C_1 fixe deux sommets distincts s, s' de T, ce groupe fixe aussi une arête issue de s : la première arête du segement $[s, s']$. On peut alors supprimer l'arête correspondante du graphe X pour obtenir un graphe ayant moins d'arêtes mais les mêmes stabilisateurs de sommets; on recommence ainsi jusqu'a obtenir une description de G ayant exactement n sommets de stabilisateurs C_i. □

Lemme I.1.3. (1) *Si il existe un x tel que $C' \subset xCx^{-1}$, alors*

$$T(A; C, C', C_1, \ldots, C_n) = T(A; C, C_1, \ldots, C_n).$$

(2) *Si C et C' sont deux sous-groupes de G pour lesquels il existe x et y tels que $xCx^{-1} \subset C'$ et $yC'y^{-1} \subset C$, alors*

$$T(A; C, C_1, \ldots, C_n) = T(A; C', C_1, \ldots, C_n).$$

Pour ne pas alourdir les notations, on traite le cas particulier où les groupes C_i sont absents, le cas général résultant d'un raisonement analogue.

Notons que le point (2) résulte du (1) grâce à la suite d'inégalités:

$$T(G; C') \geqq T(G; xCx^{-1}, C') = T(G; C, C') \geqq T(G; C, yC'y^{-1}) \geqq T(G; C).$$

Pour le (1), on montre successivement deux inégalités.

Tout d'abord, vérifions que $T(A; C, C') \leqq T(A; C)$. Soit P un A-polyèdre ayant $T(A; C)$ faces modulo l'action de G, une orbite $\omega = A/C$ de sommet de stabilisateur C. Soit P^* le polyèdre obtenu à partir de P en ajoutant à chaque élément $s \in A/C$, $s = g.C$, de ω des arêtes en bijection avec $gCg^{-1}/gxC'(gx)^{-1}$ et une extrémité libre; le polyèdre P^* est doué d'une action de A, ayant 2 orbites de sommets de stabilisateurs $C, xC'x^{-1}$.

L'inégalité inverse est plus délicate. Soit P un A-polyèdre simplement connexe, s un sommet de stabilisateur C et s' de stabilisateur $C' \subset C$. On joint s et s' par un chemin l. Soit Q le polyèdre obtenu en identifiant s' à tous les points de Cs', $g.s'$ à gCs' pour $g \in A$, l'arête l à Cl, ainsi que $g.l$ à gCl pour $g \in A$. On note π la projection de P sur Q. Montrons que Q est simplement connexe : soit γ un lacet de Q d'origine $\pi(s')$; ce lacet se relève dans P en un chemin $\tilde{\gamma}$ joignant s' à l'un de ses translatés $c_0 s'$ par C. Ainsi $c_0 \bar{l}.\tilde{\gamma} l$ est un chemin fermé dans P; celui-ci est homotope à zéro. Donc son image dans Q, qui vaut $\bar{l}\gamma l$ l'est aussi. Comme $\bar{l}$ est l'inverse de l dans Q, γ est bien homotope à zéro. □

Lemme I.1.4. (1) *Si $(C, C_1, \ldots, C_j) \subset A$, si $(C, C_{j+1}, \ldots, C_n) \subset B$, et si $G = A *_C B$ alors*

$$T(G; C, C_1, \ldots, C_j, C_{j+1}, \ldots, C_n) \leqq T(A; C, C_1, \ldots, C_j) + T(B; C, C_{j+1}, \ldots, C_n).$$

(2) *Si $D, D', C_1, \ldots, C_n \subset A$, si D et D' sont isomorphes, si $G = A *_{tDt^{-1} = D'}$ et si C est le sous-groupe engendré par t et D,*

$$T(G; D, D', C_1, \ldots, C_j) \leqq T(A; D, D', C_1, \ldots, C_j)$$

et

$$T(G; C, C_1, \dots, C_j) \leqq T(A; D, D', C_1, \dots, C_j).$$

(3) *Si* $X = (G_s, C_y)$ *est un graphe de groupe pour lequel chaque sommet s est équipé d'une famille de sous-groupes* $(C_i)_{i \in I_s}$ *et si* $G = \pi_1(X)$, *on a*:

$$T(G; C_1, \dots, C_n) \leqq \sum T(G_s, (o(C_y))_{o(y)=s}, (C_i)_{i \in I_s}).$$

Pour le (1) (resp. le (2)), on remarque d'abord que si Q_A et Q_B sont deux orbièdres *développables* de groupe fondamental A et B et si l'un des sommets a de Q_A a un stabilisateur C isomorphe a celui C de l'un des sommets b de Q_B (resp. si deux sommets a et a' de Q_A ont des stabilisateurs D et D' isomorphes), l'orbièdre Q obtenu en recollant Q_A et Q_B grâce à une arête marquée C (resp. l'orbièdre Q obtenu en ajoutant une arête d'extrémité a et b (resp. a et a') et marquée D), alors $\pi_1(Q) = A *_C B$ (resp. $A *_{tDt^{-1} = D'}$).

Notons que Q est développable puisque pour chacun de ses sommets, l'application canonique du groupe de ce sommet G_s dans $\pi_1^{\text{orb}}(Q)$, composée de l'injection (Q_A et Q_B sont développables) de G_s dans A ou B et de A ou B dans $A *_C B$ (resp. $A *_{tDt^{-1} = D'}$) est injective (voir [Ha], p. 516, Th. 4.1).

Si l'on applique cette construction en partant des polyèdres Q_A et Q_B fournis par la définition de $T(.;.)$, on remarque que le revêtement universel de Q est la première subdivision d'un G-polyèdre P ayant le nombre requis de sommets pour prouver (1) et (2).

Pour la deuxième inégalité de (2), on remarque que l'orbièdre Q^* obtenu à partir de Q en remplaçant son sommet marqué D par un sommet marqué C et en supprimant l'arête marquée D dont les deux extrémités sont ce sommet a même groupe fondamental que Q et même nombre de faces. Le cas des graphes de groupes (3) résulte de (1) et (2) par récurrence sur le nombre d'arêtes. □

I.2. Démonstration des corollaires I, II et IV de l'introduction. Dans ce sous-paragraphe, on explique comment les propriétés élémentaires de $T(.;.)$ permettent de montrer les principaux corollaires au théorème II annoncés dans l'introduction.

Corollaire I. On applique le théorème II, (2) à la décomposition $A *_C B$ de $(G; C)$. Si cette décomposition est réduite, il n'y a rien à démontrer. Sinon d'après le (2), il existe des sous-groupes D_i de C pour lesquels $T(A; C) + T(B; C, D_1, \dots, D_n) \leqq T(G; C)$. Comme tous les D_i sont contenus dans C, le lemme I.1.3 (1) montre que $T(B; C, D_1, \dots, D_n) = T(B; C)$ ce qu'il fallait montrer.

Corollaire II. Si la décomposition étudiée est réduite, il n'y a rien à démontrer compte-tenu du lemme I.1.1 (1); sinon en raisonnant comme précédement, on voit qu'il existe des sous-groupes $D_1, \dots, D_n$ de C pour lesquels $T(A; C) + T(B; D_1, \dots, D_n) \leqq T(G)$; comme tous les D_i sont contenus dans C, $\operatorname{rang}(VD_i) \leqq \operatorname{rang} C$. On applique le lemme I.1 (1) pour conclure.

Corollaire IV. Dans la description $G = \pi_1^{\text{orb}}(Q)$ de G comme groupe fondamental d'un certain orbièdre, les stabilisateurs des faces et arêtes de Q sont des sous-groupes des

C_i. Soit y une arête de Q adjacente à aucune face et séparante. Si C_i stabilise un sommet à droite et à gauche de l'arête y, il doit stabiliser y, et Q à même groupe fondamental que l'orbièdre obtenu en supprimant cette arête et en identifiant ses extrémités. En particulier quitte à effectuer cette opération un certain nombre de fois, on voit qu'on peut remplacer Q par un polyèdre Q' dont le nombre d'arêtes séparantes et adjacentes à aucune faces est majoré par n. Comme chaque face de Q a 3 cotés et comme Q a au plus $6T(G; C_1, \ldots, C_n)$ faces, le nombre d'arêtes total est borné par $3T(G; C_1, \ldots, C_n) + b_1(G; C_1, \ldots, C_n) + n$. Ainsi, si l'on suppose de plus que les C_i sont finis et fixés, il n'y a qu'un nombre fini de donnée combinatoires pour l'orbièdre Q'. En particulier:

Lemme I.2.1. *Soient $C_1, \ldots, C_n$ des groupes finis fixés et k un entier; il n'y a qu'un nombre fini de classes d'isomorphismes de paires $(G; C_1, \ldots, C_n)$ telles que $T(G; C_1, \ldots, C_n) + b_1(G; C_1, \ldots, C_n) < k$.*

Montrons donc le (a) du corollaire IV dans le cas d'une somme amalgamée. Compte tenu du théorème II, si $G = A *_F B$ est réduite, $T(A; F) + b_1(A, F) \leqq T(G) + b_1(G)$. Ainsi, compte tenu du lemme I.2.1, il suffit de voir qu'il n'y a qu'un nombre fini de classes d'isomorphismes de groupes finis F possibles. Il résulte du théorème d'accessibilité de Dunwoody [D] qu'il existe une décomposition terminale de G. On en fixe une X : pour montrer (a), il suffit donc de voir que F est l'un des stabilisateurs d'arête de X puisque ceux-ci sont en nombre finis. Mais les groupes A et B admettent des décompositions terminales X_A et X_B. Le groupe F étant fini, il fixe un sommet de X_A et un sommet de X_B, et est strictement contenu dans les stabilisateurs de ces sommets, sinon la décomposition $G = A *_F B$ ne serait pas réduite. Si X' est le graphe obtenu en recollant X_A et X_B le long d'une arête joignant ces deux sommets marquée F, X' est une autre décomposition terminale de G. Grâce à [D-D], proposition IV, 7.2, (i), les stabilisateurs d'arêtes de X et X' sont les mêmes à conjugaison près dans G; donc F est l'un des stabilisateurs d'arête de X. Pour montrer (b), il suffit d'appliquer (a) en raisonnant par récurrence sur le nombre d'arêtes d'une décomposition terminale du groupe étudié.

I.3. Décomposition rigides et réduites. On ne rappelle pas les diverses définitions de rigidité et semi-rigidité pour lesquelles on renvoit à l'introduction. Le lemme suivant sera surtout utile pour ramener l'étude des décompositions semi-rigides à celle des décompositions rigides.

Lemme I.3.1. (1) *Si $C' \subset C$ et si la décomposition $G = A *_C B$ (resp. $G = A *_C$) est rigide (resp. semi-rigide), toute décomposition $G = A' *_{C'} B'$ (resp. $G = A' *_{C'}$) l'est.*

(2) *Si la décomposition $G = A *_C B$ (resp. $G = A *_C$) est semi-rigide mais pas rigide, alors C se décompose $C = D *_{tdt^{-1} = \psi(d)}$ pour un certain $\psi : D \to D$; de plus il existe des décompositions de G, A et B, comme* HNN-*extension au dessus de D : $G = G' *_{tDt^{-1} = D'}$, $A = A' *_{tDt^{-1} = D'}$ et $B = B' *_{tDt^{-1} = D'}$, avec $G' = A' *_D B'$ (resp. $G = G' *_{tDt^{-1} = D'}$, $A = A' *_{tDt^{-1} = D'}$ avec $G' = A' *_{tDt^{-1} = D'}$). Si la décomposition $G = A *_C B$ est réduite, il en va de même de $G' = A' *_D B'$.*

La propriété de point fixe impliquée dans la définition de la rigidité est satisfaite par passage à un sous-groupe, c'est ce que dit (1).

Le point (2) est plus délicat : on explique le cas d'une somme amalgamée, le cas d'une HNN-extension étant analogue. Si la décomposition $G = A *_C B$ est semi-rigide mais pas rigide, il existe un arbre T muni d'une action de G pour laquelle tous les C_i fixent un point, C fixe un point à l'infini ω, et C contient un stabilisateur d'arête. Quitte à écraser toutes les arêtes de T dont le stabilisateur n'est pas conjugué à un sous-groupe de C, on peut supposer que C contient tous les stabilisateurs d'arêtes à conjugation près, et fixe un point ω à l'infini dans $\bar{T}$. Soit $c_1, \dots, c_n$ une famille de générateurs de C. Chacun fixe un axe $[w_i, \omega]$, avec $w_i \in \delta T$. Quitte à remplacer c_i par son inverse, on peut supposer cet axe fixe point par point, ou que c_i y agit comme translation de longueur positive l_i. Soit a un point commun à tous ces axes; la demi-droite $[a, \omega]$ est positivement invariante par tous les c_i. Quitte à changer c_i en $c_i c_j^{-1}$, et à répéter plusieurs fois cette opération, on se ramène au cas ou tous les l_i sont nuls sauf un. Ainsi, C est engendré par des éléments $d_1, \dots, d_p$ qui fixent a et un élément t qui agit par translation sur $[a, \omega]$; soit D_1 le sous-groupe de C stabilisateur de a; $t^{-1} D_1 t$ est un sous groupe de D_1. Il existe donc un homomorphisme $\psi : D_1 \to D_1$ tel que pour tout d de D_1 on ait $\psi(d) = t^{-1} d t$.

Donc le groupe C se décompose comme HNN-extension $D_1 *_{tdt^{-1} = \psi(d)}$.

On considère le graphe X, quotient de T par l'action de G. L'image de la première arête e de $[a, \omega]$ ne disconnecte pas le graphe X, car l'image de $[a, ta]$ est un lacet dans ce graphe; la théorie de Bass-Serre produit un groupe G' – le groupe fondamental du complémentaire de cette arête – et une décomposition $G = G' *_D$ (avec $D = G_e$ stabilisateur de e). Comme tous les stabilisateurs d'arêtes sont contenus dans des conjugués de C, on peut supposer $D \subset C$. Mais alors, D fixe $[u, \omega]$ pour un certain u, et t est une translation de $[b, \omega] \subset [a, \omega]$ pour b suffisament proche de ω. En partant du point b au lieu de a dans ce qui précède, on obtient un groupe D, un homomorphisme $\psi : D \mapsto D$, qui est un isomorphisme sur un sous groupe D' de D et des décompositions $G = G' *_{d = t\psi(d)'t^{-1}}$ et $C = C' *_{d = t\psi(d)'t^{-1}}$ au dessus du même sous-groupe D.

Le même argument appliqué à A et B, montre que $A = A' *_{d = t\psi(d)'t^{-1}}$, $B = B' *_{d = t\psi(d)'t^{-1}}$ et $G' = A' *_D B'$. C'est ce que dit (2).

Si la décomposition ainsi décrite de G' n'est pas réduite, alors on a par exemple $A' = A'' *_{D''} D$ (le cas d'une décomposition en graphe de groupes se traiterait de façon analogue); donc, comme $D' \subset D$ le groupe A se décompose

$$A = A' *_{d = t\psi(d)'t^{-1}} = (A'' *_{D''} D) *_d = t\psi(d)'t^{-1} = A'' *_{D''} C$$

et la décomposition initiale de G n'est pas non plus réduite. □

I.4. Sous-additivité de T et applications. Dans ce sous-paragraphe, on montre comment les lemmes I.1.3 et I.1.4 permettent de ramener la démonstration du théorème II au cas rigide. On explique aussi comment le (2) du théorème II résulte du (3).

Lemme I.4.1 *Pour démontrer le théorème* II, *il suffit de le montrer dans le cas «rigide».*

Démonstration. Pour simplifier les notations, on montre se résultat en l'absence des groupes C_i. Grâce aux propriétés de sous-additivité de T (lemme I.1.4), une récurrence ramène le problème au cas d'une seule arête : $G = A *_C B$, ou $G = A *_{tCt^{-1}=C'}$. On n'écrit que le cas d'une somme amalgamée; le cas général résulterait d'un raisonnement analogue. On fait une récurrence sur le premier nombre de Betti de C. Supposons que la décomposition étudiée soit semi-rigide mais pas rigide. Grâce au lemme I.3.1 (2), $C = D *_{tdt^{-1}=\psi(d)}$ pour un certain $\psi : D \to D$; $G = G' *_D$, $A = A' *_D$ et $B = B' *_D$, de sorte que $G' = A' *_D B'$.

Pour simplifier, supposons cette décomposition de G' réduite. Par hypothèse de récurrence, comme $b_1(D) = b_1(C) - 1$, $T(G'; D) \geqq T(A'; D) + T(B'; D)$. On peut aussi appliquer l'hypothèse de récurrence à la décomposition $G = G' *_{tDt^{-1}=D'}$, et on obtient

$$T(G'; D, D') \leqq T(G)$$

mais D' est contenu dans D, donc

$$T(G'; D) \leqq T(G)$$

d'où

$$T(G) \geqq T(A'; D) + T(B'; D).$$

Comme $D' \subset D$, $T(A'; D, D') = T(A'; D)$ et d'après le lemme I.1.4 (2), $T(A; C) \leqq T(A'; D, D')$. De même, $T(B; C) \leqq T(B'; D, D')$. Ainsi, on a:

$$T(A; C) + T(B; C) \leqq T(A'; D) + T(B'; D) \leqq T(G'; D) \leqq T(G).$$

Ce que nous souhaitions démontrer. Le cas où la décomposition $A' *_D B'$ n'est pas réduite résulterait du même raisonnement, en remplaçant $G' = A' *_D B'$ par une réduction. □

Lemme I.4.2. *La partie* (2) *du th.* II *résulte de la partie* (3).

On garde les notations de l'introduction : supposons que la décomposition $G = A *_C B$ ne soit pas réduite. Une réduction X' de cette décomposition s'obtient en deux étapes : on part d'une décomposition $A = \pi_1(X_A)$ de A dont C est l'un des stabilisateurs de sommets et $(H_s)_{s \in X_A}$ les autres dont toutes les arêtes aboutissant au point de stabilisateur C, et on note D_i les stabilisateurs de ces arêtes; X' est alors obtenu en attachant un sommet marqué B à X_A et une arête marquée C puis en supprimant l'arête triviale ainsi construite (Figure II). Le th. II (3) dit qu'alors:

$$\sum_s T\big(H_s; (C_l)_{l \in I_s}, (D_k)_{k \in K_s}\big) + T\big(B; D_1, \ldots, D_n, (C_i)_{i \in I_B}\big) \leqq T(G; C_1, \ldots, C_n).$$

En utilisant le lemme I.1.4 (3), on voit que:

$$T\big(A; C, (C_i)_{i \in I_A}\big) \leqq \sum T\big(H_s; (C_l)_{l \in I_s}, (D_k)_{k \in K_s}\big) + T(C; C, D_1, \ldots, D_n).$$

Mais d'après le lemme I.1.3, $T(C; C, \ldots) = T(C; C) = 0$ donc:

$$T\big(A; (C_i)_{i \in I_A}\big) + T\big(B, (D_j)_{1 \leqq j \leqq k}, (C_i)_{i \in I_B}\big) \leqq T(G; C_1, \ldots, C_n)$$

comme promis.

I.5. Sommets essentiels, accessibilité. Soit $G = \pi_1(X, s_0)$ une décomposition de $(G; C_1, \ldots, C_n)$ en graphe de groupes. Rappelons que l'ensemble $\{1, \ldots, n\}$ admet une partition $\bigcup_s I_s$, pour laquelle $C_i \subset G_s$, $i \in I_s$. Dans ce paragraphe, on définit la notion de sommet *essentiel*, et on montre comment le corollaire III résulte du théorème II.

Définitions. Un sommet s de valence un est dit *inessentiel* si I_s est vide et si son stabilisateur est égal à celui de l'arête dont il est issu. Un sommet s de valence deux de X (ayant deux arêtes adjacentes y et z) est dit *inessentiel* si I_s est vide, et G_s se décompose $G_s = C_y *_C C_z$; plus généralement, un sommet de valence k, origine de $y_1, \ldots, y_k$ est dit *inessentiel* si I_s est vide, si s n'est pas le point base s_0, et si son stabilisateur est isomorphe à un graphe de groupes dont le graphe sous-jacent est un arbre ayant k sommets de stabilisateurs égaux aux C_y.

Un sommet *essentiel* est un sommet qui n'est pas inessentiel.

Remarque I.5.1. Grâce au lemme I.1.1, on voit qu'un sommet s est inessentiel si et seulement si I_s est vide et si désignant par y les arêtes issues de s, on a : $T(G_s; (C_y)_{o(y)=s}) = b_1(G_s; (C_y)_{o(y)=s}) = 0$. En particulier, le nombre de sommets essentiels du graphe X n'excède pas $\sum_{s/I_s=\emptyset} T(G_s; (C_y)_{o(y)=s}) + b_1(G_s; (C_y)_{o(y)=s}) + n$.

Lemme I.5.2. *Si la décomposition est semi-rigide, sans arête triviale, le nombre de sommets inessentiels de valence deux n'excède pas la moitié du nombre de sommets total; il n'y a pas de sommets inessentiels de valence un.*

Si s est un sommet inessentiel de valence un, il est égal au stabilisateur de l'arête dont il est issu, en particulier celle-ci est triviale. Soit s un sommet de valence deux inessentiel. Ecrivons $G_s = C_y *_{C_x} C_z$ pour un certain groupe C_x. Enlevons ce sommet et remplaçons les deux arêtes y et z d'origine s par une nouvelle arête x allant de $e(y)$ à $e(z)$. On note X' le graphe ainsi obtenu. Montrons que $s' = o(x)$ et $s'' = e(x)$ sont des sommets essentiels de X'. Supposons au contraire s' inessentiel. On peut décomposer $G_{s'} = C_{y'} *_C C_x$; mais $C_x \subset C_y \subset G_{s'}$; cela contredit la (semi-)rigidité de y : si $C_y \neq C_x$, on peut construire une action de G sur un arbre tel que C contient un stabilisateur d'arête, mais ne fixe aucun point : pour faire cela, il suffit d'oter le sommet s' de x' et de faire agir G sur l'arbre de Bass-Serre du graphe ainsi obtenu. Donc $C_y = C_x$, ainsi $G_s = C_y *_{C_x} C_z = C_z$, et l'arête z était triviale. □

On obtient ainsi le corollaire III de l'introduction.

Corollaire I.5.3. *Le nombre de sommets d'une décomposition sans arête triviale et semi-rigide de $(G; C_1, \ldots, C_n)$ comme graphe de groupe n'excède pas*

$$4\big(T(G; C_1, \ldots, C_n) + 2b_1(G; C_1, \ldots, C_n) + n\big).$$

Si tous les sommets sont essentiels, il est majoré par $T(G; C_1, \ldots, C_n) + b_1(G; C_1, \ldots, C_n) + n$.

Démonstration. Soit X' la réduction de X fournie par le théorème II. Grâce à la remarque I.5.1 et aux propriétés de sous-additivité de $T(.,.)$, on voit que X' a au moins autant de sommets essentiels que X. D'autre part, jointe à la remarque I.5.1, l'inégalité exprimée par le théorème II (3) montre que le nombre de sommets essentiels de X' (et donc celui de X) n'excéde pas $T(G; C_1, \dots, C_n) + b_1(G; C_1, \dots, C_n) + n$.

Soient e le nombre de sommets essentiels, i_2 le nombre de sommets inessentiels de valence deux, e_2 de sommets essentiels de valence deux, s_1 le nombre de sommets de valence un, et s_3 le nombre de sommets de valence au moins égale à trois. Un calcul de caractéristique d'Euler montre que $s_3 \leqq s_1 + 2(b_1(X) - 1)$. Grâce au lemme précédent,

$$s_1 + i_2 + e_2 + s_3 \leqq 2(s_1 + e_2 + s_3) \leqq 2\big(2s_1 + e_2 + 2b_1(G, C_1, \dots, C_n)\big).$$

Comme tous les sommets de valence un sont essentiels, on obtient comme promis

$$\begin{aligned} s_1 + i_2 + e_2 + s_3 &\leqq 2\big(2e + 2b_1(G, C_1, \dots, C_n)\big) \\ &\leqq 4\big(T(G, C_1, \dots, C_n) + 2b_1(G, C_1, \dots, C_n) + n\big). \quad \square \end{aligned}$$

Le lemme suivant explique comment on peut se débarasser des sommets inessentiels par réduction.

Lemme I.5.3. *Soit $G = \pi(X)$ une décomposition de $(G; C_1, \dots, C_n)$; il existe une réduction de X sans sommets inessentiels.*

La démonstration se fait par récurrence sur le nombre d'arêtes. Si s est un tel sommet, et si $y_1, \dots, y_n$ sont les arêtes issues de s, G_s s'écrit comme groupe fondamental d'un arbre de groupe T_s ayant les C_{y_i} comme stabilisateurs de sommets. La réduction de X obtenue en remplaçant s par T_s et en supprimant les arêtes triviales ainsi crées a une arête de moins que X, sauf si X avait s comme seul sommet, auquel cas I_s est non vide (l'ensemble $\{1, \dots, n\}$ a au moins 1 élément) et s est essentiel.

I.6. Topologie de petite dimension.

Proposition I.6.1. *Soit M une surface (resp. une variété de dimension 3 irréductible) et C une courbe fermée simple non homotope à zéro (resp. F une surface incompressible). La décomposition de $\pi_1(M)$ obtenue en découpant M le long de C (resp. F) est rigide et réduite.*

Pour démontrer ce fait, nous utiliserons deux résultats : le premier est dû à Hempel [H], p. 89, th. 10.2; le second est expliqué dans [J], lemme III.9, th. III.10, où il est attribué à J. Stallings et F. Waldhausen:

Lemme I.6.2. *Soit M une variété P^2-irréductible de dimension 3 et F une composante du bord. Si $\pi_1(M) = \pi_1(F)$ est infini, alors $M = F \times [0,1]$.*

Lemme I.6.3. *Soit N variété de dimension trois, F une composante incompressible du bord; si $(\pi_1(N); \pi_1(F))$ se décompose comme somme amalgamée ou HNN-extension, $\pi_1(N) = A *_C B$ ou $A *_C$ avec $\pi_1(F) \subset A$, il existe une surface incompressible localement séparante Σ telle que $\pi_1(\Sigma) \subset C$ et $F \cap \Sigma = \emptyset$.*

Démonstration de la proposition I.6.1. On ne détaille que le cas d'une variété fermée de dimension trois; le cas de la dimension deux, ou d'une variété à bord est laissé au lecteur. On souhaite démontrer que si la surface F est incompressible la décomposition de

$\pi_1(M)$ qu'elle définit est rigide et réduite. Pour cela il suffit de montrer que $\pi_1(M)$ ne se décompose pas au-dessus d'un sous-groupe de $\pi_1(F)$.

Si c'était le cas, le lemme I.6.2 appliqué à $M = N$ montre que ce groupe se décompose aussi le long d'un sous-groupe de $\pi_1(\Sigma)$ de $\pi_1(F)$ qui est un groupe de surface fermée. Si cette surface est une sphère ou un plan projectif, M n'est pas irréductible. Donc il s'agit d'une surface de groupe fondamental infini, sous-groupe de $\pi_1(F)$: la dualité de Poincaré montre qu'il y est d'indice fini. Si $\pi_1(M)$ agit dans un arbre où $\pi_1(\Sigma)$ fixe un sommet, un argument classique de barycentre montre alors que $\pi_1(F)$ y fixe aussi un sommet. La décomposition est donc rigide. Elle est réduite par un argument analogue : la surface F sépare M en deux composantes, dont la première, disons N, satisfait $\pi_1(N) = A$. Supposons que A se décompose $A' *_C \pi_1(F)$ (une décomposition en graphe de groupes se traiterait d'une façon analogue); I.6.2 (appliqué à N) montre qu'il existe une décomposition de N le long d'une surface Σ, telle que $\pi_1(\Sigma) \subset C \subset \pi_1(F)$ et telle que la variété comprise entre Σ et F ait même groupe fondamental que F; en appliquant I.6.3, on voit que Σ est connexe et $\pi_1(F) = \pi_1(\Sigma) = C$.

Remarque. D'après ce raisonnement, si M est une variété de dimension trois P^2-irréductible, une décomposition réduite de $\pi_1(M)$ correspond exactement à une surface incompressible (non nécessairement connexe) dont deux composantes connexes ne sont pas parallèles. Ceci justifie l'interêt de cette notion.

II. Une construction topologique

La construction présentée dans ce paragraphe s'inspire deux idées fondamentales de la topologie de petite dimension et la théorie combinatoire des groupes. La première, due à J. Stallings, est la solution à la conjecture de Kneser : si le groupe fondamental d'une variété de dimension trois est un produit libre cette variété est une somme connexe. Plus généralement, si $G = \pi(X)$ est une décomposition du groupe fondamental d'une variété de dimension trois, il existe une décomposition géométrique de ce groupe fondamental obtenue en découpant la variété le long d'une surface incompressible Σ et telle que les composantes du complément de Σ fixent des sommets de X. La seconde, due à M.J. Dunwoody, est celle de résolution : si un groupe de présentation finie opère dans un arbre T, il opère dans un autre arbre $\hat{T}$ dont on sait contrôler le nombre d'orbites de sommets, et tel que les stabilisateurs de sommets et d'arêtes de $\hat{T}$ soient contenus dans ceux de T.

Après avoir énoncé le lemme II.1, nous montrons son efficacité en déduisant le théorème I de l'introduction; la troisième partie du paragraphe est consacrée à la démonstration de II.1. Nous utiliserons librement la notion de groupoïde fondamental d'un graphe de groupe pour laquelle nous renvoyons à [Se].

II.1. Lemme technique principal.

Lemme II.1.1. *Soit* $(G; C_1, \dots, C_n)$ *un groupe équipé d'une famille de sous-groupes,* $(X; (G_s)_{s \in X^0}, (C_y)_{y \in X^1})$ *un graphe de groupe,* $\{1, \dots, n\} = \bigcup_{s \in X^0} I_s$ *une partition de l'ensemble* $\{1, \dots, n\}$ *et* $\lambda : G \mapsto \pi_1(X, s_0)$ *un isomorphisme tel que si* $i \in I_s$ $\lambda(C_i)$ *soit conjugué à un sous-groupe de* G_s.

Il existe un graphe de groupes $(\hat{X}; (\hat{G}_s)_{s\in\hat{X}^0}, (C_y)_{y\in\hat{X}^1})$, un sommet $\hat{s}_0$ de $\hat{X}$, un isomorphisme $\hat{\lambda} : G \mapsto \pi_1(\hat{X}, \hat{s}_0)$, une partition $\bigcup_{s\in\hat{X}} \hat{I}_s$ de l'ensemble $\{1, \dots, n\}$ et un homomorphisme de groupoïdes : $J : \pi(\hat{X}) \mapsto \pi(X)$ tel que $\lambda = J \circ \lambda$, et satisfaisant les propriétés suivantes:

$$(*) \qquad \sum T(\hat{G}_s; (C_y)_{o(y=s)}, (C_i)_{i\in\hat{I}_s}) \leqq T(G; C_1, \dots, C_n).$$

$(**)$ *Il existe une application j de l'ensemble des sommets de $\hat{X}$ dans celui des sommets de X telle que, pour tout sommet s de $\hat{X}$, $J(\hat{G}_s)$ est contenu dans $G_{j(s)}$.*

$(***)$ *Pour toute arête y de $\hat{X}$, d'origine $o(y)$ et d'extrémité $e(y)$, $J(y)$ est une arête de X joignant $j(o(y))$ à $j(e(y))$. En particulier, comme J est un homomorphisme de groupoïdes, $J(\hat{C}_y) \subset C_{J(y)}$.*

$(****)$ *Si $i \in \hat{I}_s$, $\hat{\lambda}(C_i) \subset \hat{G}_s$.*

Dans toute la suite $I : \pi_1(X, s_0) \mapsto \pi_1(\hat{X}, \hat{s}_0)$ est défini par $I = \hat{\lambda} \circ \lambda^{-1}$, de sorte que $J \circ I = \mathrm{Id}$.

II.2. Démonstration du théorème I. Bien que le théorème I soit une conséquence formelle du théorème II, il est agréable d'en donner une démonstration autonome basée sur l'énoncé du paragraphe précédent et sur le théorème d'unicité de Kuròs que l'on peut énoncer ainsi:

Théorème II.2.1 (voir [S-W] Thm. 3.5). *Soit G un groupe de rang fini. Il existe des groupes $G_1, \dots, G_r$ non libres et indécomposables en produit libre et un entier s tels que:*

$G = G_1 * \cdots * G_r * L_s$, où L_s est le groupe libre de rang s.

Si $G = G'_1 * \cdots * G'_{r'} * L_{s'}$ est une autre décomposition satisfaisant les mêmes propriétés, $r = r'$, $s = s'$ et il existe une permutation s de $\{1, \dots, r\}$ telle que $G_i = G'_{\sigma(i)}$.

Pour démontrer le théorème I de l'introduction on peut raisonner ainsi.

Soit $G = A * B$ un produit libre. Soient $A = A_1 * \cdots * A_n * L_p$ et $B = B_1 * \cdots * B_m * L_q$ les décompositions de A et B fournies par le théorème de Kuròs; $G = A_1 * \cdots * A_n * B_1 * \cdots * B_m * L_{p+q}$ est une décomposition de Kuròs de G isomorphe à $G = G_1 * \cdots * G_r * L_s$. En particulier, pour démontrer le théorème I, il suffit de voir que $T(G) = \sum T(G_i) (= \sum T(A_i) + \sum T(B_j))$.

On applique le lemme II.1.1 à la décomposition de G en graphe de groupes, $G = \pi_1(X)$, où X est un graphe à r sommets marqués G_i et $r + s - 1$ arêtes à stabilisateurs triviaux.

Soit $\hat{X} : (\hat{G}_s, \hat{C}_y)$ le graphe fourni par le lemme II.1 : tous les stabilisateurs des arêtes de $\hat{X}$ sont triviaux d'après la propriété $(***)$; ainsi G est isomorphe au produit libre de ses stabilisateurs de sommets et d'un groupe libre de rang égal à $b_1(\hat{X})$.

Comme G_s est indécomposable en produit libre et est non libre, il satisfait la propriété de point fixe suivante:

Si G_s agit dans un arbre $\hat{T}$ et si les stabilisateurs des arêtes de $\hat{T}$ sous cette action sont triviaux, G_s fixe un sommet de $\hat{T}$.

Ainsi, en faisant agir G_s dans l'arbre de Bass-Serre de $\hat{X}$, on voit que $I(G_s)$ doit être contenu dans un conjugué de l'un des $\hat{G}_\sigma$. On a donc $I(G_s) \subset m\hat{G}_\sigma m^{-1}$. En appliquant J, on voit que $G_s \subset nJ(G_\sigma)n^{-1} \subset nG_{f(\sigma)}n^{-1}$. Or G_s ne peut être contenu dans l'un de ses conjugués, ni dans l'un des conjugués de G_t, $t \neq s$ puisque ces groupes fixent des points différents dans l'arbre de Serre de X; il en résulte que $G_s = mJ(\hat{G}_\sigma)m^{-1} = G_s$, et J, dont la restriction à chaque $\hat{G}_s$ est injective puisque λ et λ' sont des isomorphismes, définit un isomorphisme de $\hat{G}_\sigma$ et G_s. Comme par construction $\sum_\sigma T(\hat{G}_\sigma) \leqq T(G)$, on a obtient $\sum T(G_s) \leqq T(G)$ ce qui est le résultat souhaité. □

Démonstration du lemme II.1.1. Pour démontrer le lemme II.1, notre stratégie est simple : nous allons montrer que le graphe de groupe fourni par la «résolution» (voir [D]) de M.J. Dunwoody convient.

Soit P un polyèdre simplicial simplement connexe de dimension deux muni d'une action de G, et dont les stabilisateurs de sommets sont les C_i et leurs conjugués, et ayant $T(G; C_1, \dots, C_n)$ faces modulo l'action de G.

Soient $p_1, \dots, p_k$ des représentants des orbites des sommets de P, de stabilisateurs égaux à l'un des C_i; quitte à renuméroter la famille des C_i et à répéter certains de ses termes ce qui ne change pas la valeur de $T(;)$, on peut supposer que le nombre k d'orbites de points fixes est égal au nombre des C_i, et que C_i est le stabilisateur de p_i.

Notons qu'en général $k > n$ ce qui posera un petit problème par la suite, mais l'on peut supposer, quitte à renuméroter, que pour $j > k$ il existe $i \leqq n$ tel que C_i soit (conjugué) C_j.

Chaque sommet p de P s'écrit donc $p = g.p_i$, et g est unique dans le quotient G/C_i. Pour chacun des indices i, on choisit dans l'arbre de Bass-Serre T du graphe X un point fixe t_i par C_i.

Nous allons construire successivement plusieurs objets géométriques illustrés sur la Figure III.

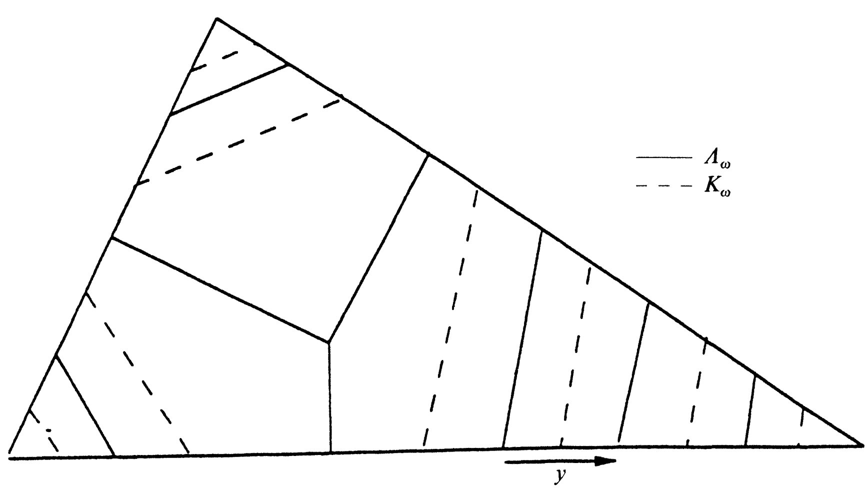


Figure III

L'application ϱ. La formule suivante définit une application ϱ équivariante du 0-squelette de P dans T: si $= gp_i$, $\varrho(p) = g.t_i$.

On prolonge ϱ au 1-squelette de P en imposant à l'image par ϱ d'une arête $[p, q]$ de p d'être l'unique segment joignant $\varrho(p)$ à $\varrho(q)$.

La lamination Λ. Chaque face Δ de P est alors équipée d'une *lamination* Λ_Δ (la terminologie est celle de Thurston pour les surfaces) avec une singularité en son centre : $\Lambda_\Delta \cap \delta\Delta$ est constitué des points du coté de Δ dont l'image par ϱ est un sommet de T : ainsi si m désigne la mesure naturelle sur P^1 tirée en arrière de la mesure simpliciale de T par l'application ϱ, Λ est constitué des points situés à une distance entière d'un sommet.

A l'exception d'un point sur chacun des cotés, chaque point de $\delta\Delta$ a exactement un vis-à-vis, c'est-à-dire un point ayant même image par ϱ dans T. On joint ces deux points par un segment quand ils appartiennent à $\Lambda_\Delta \cap \delta\Delta$.

Les trois points de $\delta\Delta$ ayant même image sont reliés par un tripode.

Il est aisé de faire en sorte que la lamination ainsi construite le soit de façon équivariante.

Le graphe K. On définit un graphe K (en pointillé dans la Figure III) sur P : dans chaque triangle Δ, K est obtenu en joignant les milieux des points de Λ à leurs vis-à-vis. Si une arête de P n'est adjacente à aucune face, la restriction de K à cette arête est aussi constitué des milieux des points de Λ.

Lemme II.3.2. *Le graphe K est transversalement orienté.*

Comme G agit sans inversion sur T, on peut équiper chaque arête de cet arbre d'une orientation, le résultat étant G invariant. Une telle orientation définit une orientation transverse de K (Figure III). □

L'arbre $\hat{T}$. En recopiant l'argument de Dunwoody [D], on définit un arbre : les sommets de $\hat{T}$ sont les composantes connexes de $P - K$; les arêtes sont les composantes connexes de K, les relations d'incidences sont les relations évidentes.

Notons que l'action de G dans $\hat{T}$ est sans inversion, puisque les arêtes de $\hat{T}$ sont orientées comme transversales aux composantes de K; cela permet de construire:

Le graphe de groupe $\hat{X}$. Par définition, c'est le quotient de $\hat{T}$ par l'action de G.

Par construction, l'application ϱ induit une application simpliciale G-équivariante de $\hat{T}$ dans T de façon tautologique : en effet deux points de Λ (ou K) ayant même image par ϱ définissent le même sommet de $\hat{T}$. En particulier les stabilisateurs des sommets et arêtes de $\hat{T}$ sont contenus dans ceux des sommets et arêtes de T qui sont leurs images par ϱ.

Par construction $\hat{X}$ définit une décomposition de $(G; C_1, \ldots, C_k)$ et l'homomorphisme ϱ passe au quotient en un homomorphisme de groupoïde J qui satisfait aux propriétés (**) et (***) et (****) requises par l'énoncé du lemme II.1.

Il reste à vérifier la propriété la plus intéressante de $\hat{X}$, c'est-à-dire l'inégalité (*) portant sur l'invariant T.

Tout d'abord le lemme II.3.3 permet d'attirer l'attention du lecteur sur une petite difficulté. Rappelons que l'ensemble des indices $\{1, \ldots, k\}$ numérote l'ensemble des orbites de sommets de P, et que cette numérotation est choisie de sorte que le stabilisateur de la i-ème orbite soit C_i si $i \leqq n$; de plus si $j > n$ il existe un i de sorte que $C_j = C_j$. Par construction il existe une partition $\bigcup \hat{J}_s$ de l'ensemble $\{1, \ldots, k\}$ telle que pour tout $j \in \hat{J}_s$, C_j fixe le sommet s dans $\hat{T}$. On note $\hat{I}_s$ la partition induite sur $\{1, \ldots, n\} \subset \{1, \ldots, k\}$.

Lemme II.3.3. *Pour tout sommet s de $\hat{X}$,*

$$T\big(\hat{G}_s; (\hat{C}_y)_{o(y)=s}, (C_j)_{j \in \hat{J}_s}\big) = T\big(\hat{G}_s; (\hat{C}_y)_{o(y)=s}, (C_i)_{i \in \hat{I}_s}\big).$$

Soit $j > n$; on sait que $C_j = C_i$ pour un certain indice $i \leqq n$. On distingue deux cas. Supposons tout d'abord $j \in \hat{J}_s$ et $i \in \hat{I}_s$. Dans ce cas, $T(\hat{G}_s; C_i, C_j, \ldots)$ à cause de la définition de $T(.;.)$. Si au contraire, $i \in \hat{I}_s$ et $j \in \hat{J}_t$, $t \neq s$, $C_i = C_j$ se doit de fixer deux sommets distincts, s et t. Ainsi, C_j est contenu dans le stabilisateur d'une arête issue de t, $C_j \subset x\hat{C}_y x^{-1}$, et $T(\hat{G}_t; \ldots, C_{j-1}, C_j, C_{j+1}, \ldots, C_y) = T(\hat{G}_s; C_{j-1}, C_j j+1, C_y)$ d'après le lemme I.1.3 (1), puisque C_j est contenu dans un conjugué de C_y. On supprime ainsi successivement tous les C_j, $j > n$ sans changer la valeur de T. □

Pour achever de démontrer le lemme II.1, il suffit donc d'obtenir:

Lemme II.3.4. $\sum_{s \in \hat{X}} T\big(\hat{G}_s; (\hat{C}_y)_{o(y)=s}, (C_j)_{j \in \hat{J}_s}\big) \leqq T(G; C_1, \ldots, C_n)$.

On considère le polyèdre $\bar{P}$ obtenu à partir de P en contractant chaque composante du graphe K en un point; triangle par triangle, le résultat $\bar{P}$ obtenu se lit sur la Figure IV.

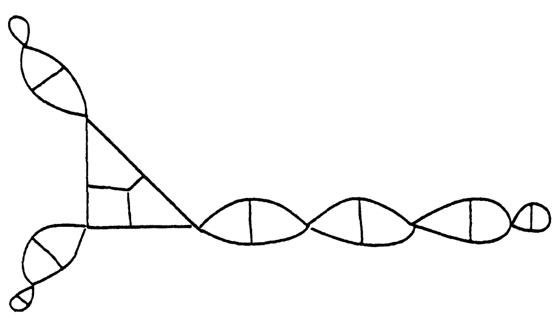


Figure IV

Les fibres de l'application $P \mapsto \bar{P}$ étant connexes, $\bar{P}$ est un G-polyèdre simplement connexe, constitué de digones et de triangles.

Les sommets de $\bar{P}$ sont d'une part les sommets de P, d'autre part les composantes connexes de Λ. Ainsi, modulo l'action de G les stabilisateurs des sommets de $\bar{P}$ sont les groupes $(C_i)_{1 \leqq i \leqq k}$, et les $\hat{C}_y$, où y décrit l'ensemble des arêtes de $\hat{X}$.

Soit Q le polyèdre simplicial de dimension deux obtenu à partir de $\bar{P}$ en y supprimant les intérieurs des digones et en identifiant les deux arêtes de chaque digones par un homéomorphisme fixant leurs sommets et préservant la mesure induite par ϱ. La Figure V montre le résultat obtenu face par faces. Comme les fibres de l'application $\bar{P} \mapsto Q$ sont connexes, Q est un polyèdre simplicial simplement connexe de dimension 2 ayant même-0-squelette que $\bar{P}$.

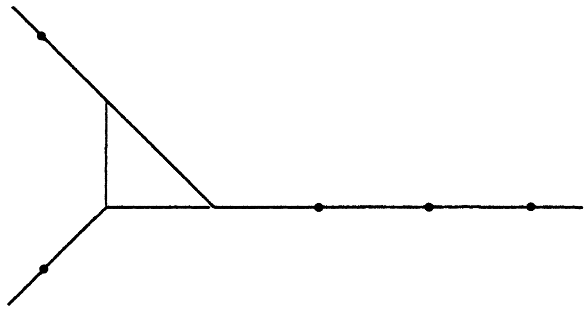


Figure V

Soit y une arête de l'arbre $\hat{T}$, c'est-à-dire une composante de K; y définit un point $f(y)$ dans Q, et $Q - f(y)$ à deux composantes connexes, qui sont simplement connexes d'après le théorème de Van-Kampen. Si s est un sommet de $\hat{T}$, on définit Q_s comme étant l'adhérence dans Q de la composante connexe de $Q - \{f(y)/o(y) = s\}$.

Les polyèdres Q_s sont simplement connexes, et Q est obtenu en recollant les Q_s grâce aux indications donnés par l'arbre $\hat{T}$: si s et s' sont les deux sommets d'une même arête, $Q_s \cap Q_{s'}$ est le point $f(y)$; sinon, $Q_s \cap Q_{s'}$ est vide.

De cette construction, et de la définition de $\hat{G}_s$, il résulte que si s est un sommet de $\hat{X}$ représenté par un point abusivement noté s de $\hat{T}$, si $y_1, \ldots, y_{k_s}$ sont les arêtes de $\hat{T}$ issues de s, et si $\hat{I}_s$ désigne l'ensemble des i tels que le sommet p_i soit dans Q_s, Q_s est un $\hat{G}_s$-polyèdre simplement connexe ayant $k_s + \# \hat{I}_s$ orbites de sommets de stabilisateurs respectifs $\hat{C}_{y_i}$, et C_i, $i \in \hat{I}_s$.

Pour chaque sommet s du graphe $\hat{X}$ on choisit un représentant (aussi noté s) dans $\hat{T}$. Ainsi Q est la réunion des $G.Q_s$, et si $s = s'$ $G.Q_s \cap G.Q_{s'}$ est constitué de points, donc ne contient pas de simplexes.

Le nombre de simplexes de Q modulo l'action de G est donc la somme $\sum_{s \in \hat{X}} T(\hat{G}_s; (\hat{C}_y)_{o(y)=s}, (C_j)_{j \in \hat{I}_s})$. Comme ce nombre est égal à celui des simplexes de P modulo l'action de G, puisque ceux-ci sont en correspondance bijective, on obtient l'inégalité souhaitée. □

Remarque II.3.5. Supposons que G soit le groupe fondamental d'une surface S ou d'une variété de dimension trois V, et que P soit le revêtement universel de S ou du deux-squelette Σ d'une triangulation d'une variété de dimension trois irréductible. Dans ce cas les composantes connexes de K sont des courbes fermées simples tracées dans S, ou les traces dans Σ d'une surface plongée dans V, et la décomposition de $\pi_1(S)$ ou $\pi_1(V)$

décrite par $\hat{T}$ est géométrique, c'est-à-dire qu'elle provient du découpage de S ou V le long d'une famille de courbe ou d'une surface. Ainsi la construction proposée est une généralisation de la construction de Stallings dans sa solution à la conjecture de Kneser.

III. Démonstration du théorème II

Partant d'un groupe G muni d'une telle décomposition $G = \pi_1(X, s_0)$, et d'un polyèdre simplement connexe P muni d'une action de G ayant T faces modulo G, nous avons construit une nouvelle décomposition $G = \pi_1(\hat{X}, \hat{s}_0)$ satisfaisant l'inégalité qu'exprime le théorème II. C'est en étudiant les liens entre X et $\hat{X}$ que nous obtiendrons le théorème II : malheureusement l'argument expliqué au II.2 pour démontrer le théorème I repose de façon essentielle sur le théorème d'unicité de Kurös (Th. II.2.11), et c'est ce point qu'il faudrat contourner.

III.1. Un lemme de point fixe. On ne répète pas les notations du paragraphe II. Chacun des G_s agit sur $\hat{T}$ via I; dans ce paragraphe, on met en évidence quelques propriétés de cette action.

L'hypothèse de rigidité est une propriété de point fixe; montrons qu'elle entraine:

Lemme III.1.1. *Supposons que la décomposition $G = \pi_1(X)$ soit rigide. Pour toute arête y, $I(C_y)$ fixe un sommet de $\hat{T}$.*

Démonstration. Par construction, chacun des C_i fixe un point dans $\hat{T}$. De plus pour toute arête y de $\hat{X}$, $J(\hat{C}_y) \subset C_{J(y)}$ et J est surjective. En appliquant I, on obtient $\hat{C}_y \subset I(C_y)$. Donc l'image par I de C_y contient un stabilisateur d'arête de $\hat{T}$; la définition de «rigide» dit qu'alors $I(C_y)$ fixe un sommet dans cet arbre. □

Vient alors l'important lemme d'accessibilité.

Lemme III.1.2. *Si $G = \pi(X)$ est une décomposition rigide de $(G; C_1, \ldots, C_n)$ dont les sommets sont tous esentiels. Le nombre de ses sommets n'excède pas $3T(G; C_1, \ldots, C_n) + 4b_1(G; C_1; \ldots, C_n) + n$.*

Remarquons que la borne de cet énoncé intermédiaire n'est pas optimale.

Démonstration. Pour démontrer ce lemme, on peut recopier l'argument de Dunwoody [D] ou raisonner ainsi.

Considérons le graphe de groupe $\hat{X}^*$ obtenu a partir de $\hat{X}$ en y supprimant les sommets inessentiels de valence 1 et les arêtes dont ils sont issus. Ces deux objets ont même groupe fondamental et on a: $\sum_{s \in \hat{X}^*} T(\hat{G}_s; (\hat{C}_y), (C_i)) = \sum_{s \in \hat{X}} T(\hat{G}_s; (\hat{C}_y), (C_i))$.

Soit s_i le nombre de sommets de valence i de $\hat{X}^*$.

Comme $\hat{X}^*$ n'a pas de sommets inessentiels de valence 1, ceux-ci sont tous essentiels, et d'après la remarque I.5.2,

$$(*)\qquad s_1 \leqq b_1(G; C_1, \ldots, C_n) + T(G; C_1, \ldots, C_n).$$

Soit s_i le nombre de sommets de valence i de $\hat{X}$. Le nombre de sommets de valence au moins égale à trois $\sum_{i \geqq 3} s_i$ est majoré par

$$\sum_{i \geqq 3} (i-2)s_i = 2b_1(\hat{X}) - 2 + 2s_1$$
$$\leqq 2b_1(G; C_1, \ldots, C_n) - 2 + b_1(G; C_1, \ldots, C_n) + T(G; C_1, \ldots, C_n) \leqq 3b_1 + T$$

(d'après $(*)$).

Le nombre de sommets de $\hat{X}^*$ tels que $\hat{I}_s \neq \emptyset$ est majoré par n. Celui tel que $T(\hat{G}_s; (\hat{C}_y)) \neq 0$ ou $b_1(\hat{G}_s; (\hat{C}_y)) \neq 0$ est majoré par $T(G; C_1, \ldots, C_n) + b_1(G; C_1, \ldots, C_n)$. Ainsi, si le nombre de sommets de X excède le nombre proposé dans le lemme, il existe un sommet s dans X tel que pour tout $\sigma \in j^{-1}(s)$, σ soit un sommet de valence deux satisfaisant $\hat{I}_\sigma = 0$ et $b_1(\hat{G}_\sigma; \hat{C}_y, \hat{C}_z) = T(\hat{G}_\sigma; \hat{C}_y, \hat{C}_z) = 0$ autrement dit est inessentiel: $\hat{G}_\sigma = \hat{C}_y *_{D_\sigma} \hat{C}_z$.

Soit $\tilde{\hat{X}}$ le graphe obtenu à partir de $\hat{X}^*$ en y otant les sommets de $j^{-1}(s)$ et leur deux arêtes adjacentes, et en les remplaçant par une seule arête marquée D_σ. Le graphe obtenu, noté $\tilde{\hat{X}}$ a même groupe fondamental que $\hat{X}$.

On fait agir le groupe G_s sur l'arbre de Bass-Serre $\tilde{\hat{T}}$ de ce graphe via l'homomorphisme I, et l'on considère le quotient comme description de G_s comme groupe fondamental d'un certain graphe de groupes que l'on note Γ.

En raisonnant comme au lemme III.1.1, on voit que pour chaque arête y issue de s, C_y fixe un sommet dans le graphe Γ. Comme modulo les C_y, le groupe G_s admet un nombre fini de générateurs, c'est-à-dire qu'il existe un nombre fini d'éléments qui joints aux C_y engendrent G_s, ce graphe à même groupe fondamental qu'un sous-graphe fini Γ_1. Notons que Γ_1 est un arbre car $b_1(G_s; (C_y)_{o(y)=s}) = 0$. Soit Γ_2 le graphe obtenu a partir de Γ_1 en y supprimant les arêtes triviales : par construction les stabilisateurs D_t des sommets de Γ_2 sont de la forme $J(m\hat{G}_{\sigma'}m^{-1}) \cap G_s$ avec $j(\sigma') \neq \sigma$. Un tel groupe fixant s et ms' dans l'arbre de Bass-Serre de X doit être contenu dans le stabilisateur d'une arête de X issue de s; ainsi, $D_t \subset aC_ya^{-1}$.

Notons que comme aucune arête de Γ_2 n'est triviale, si pour deux sommets t et t' de Γ_2 on a : $D_t \subset aD_{t'}a^{-1}$ alors $t = t'$ et $D_t = aD_{t'}a^{-1}$; sinon D_t serait contenu donc égal au stabilisateur d'une arête issue de t.

Soient $(D_1, \ldots, D_l)$ les stabilisateurs des sommets de Γ_2. Pour obtenir la contradiction souhaitée, il suffit de voir que s est inessentiel.

Montrons que $T(G_s; (C_y)_{o(y)=s}) = T(G_s; D_1, \ldots, D_l) = 0$. La seconde égalité provient du fait que G_s est écrit comme arbre de groupes ayant les D_i comme stabilisateurs et du lemme I.1.1. Pour la première, remarquons les deux faits suivants:

$$\forall y, \quad \exists t, \quad \exists a \quad / \quad C_y \subset aD_ta^{-1},$$
$$\forall t \quad \exists y, \quad \exists b \quad / \quad D_t \subset bC_yb^{-1}.$$

Ainsi, en appliquant plusieurs fois le lemme I.1.3 (2) on obtient:

$$T(G_s; (C_y)_{o(y)=s}) = T(G_s; (C_y)_{o(y)=s}, (D_t))$$

d'après la seconde ligne puis

$$T(G_s; (C_y)_{o(y)=s}, (D_t)) = T(G_s; D_t) = 0$$

d'après la première.

Comme on a pris soin à ce que $b_1(G_s; (C_y)_{o(y)=s}) = 0$, on voit que s est inessentiel. □

Raffinement. Un raffinement élémentaire de la décomposition $G = \pi(X)$ est une nouvelle décomposition de $(G; C_1, \ldots, C_n)$ obtenue en choisant un sommet s de X, en décomposant $(G_s; (C_i)_{i \in I_s}, (C_y)_{o(y)=s})$ en amalgame ou HNN, $A_s *_{C_s} B_s$ ou $A_s *_{C_s}$, et en remplaçant dans X le sommet s par deux sommets de stabilisateurs A_s et B_s (ou A_s) et une arête les joignants (Figure VI). On obtient ainsi un graphe de groupe ayant même groupe fondamental. On dit qu'une décomposition $G = \pi(Y)$ est une raffinement de la décomposition $G = \pi(X)$ si l'on obtient Y à partir d'une suite de raffinements élémentaires. Le lemme suivant, où l'on a mis des notations évidentes, résulte immédiatement du lemme I.1.4 pour le (a) et de la définition de réduction (élémentaire) pour le (b).

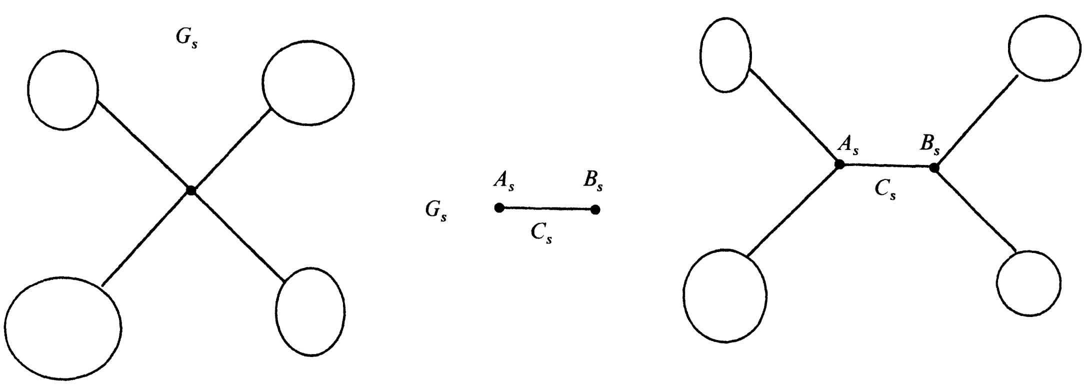


Figure VI

Lemme III.1.3. *Si Y est un raffinement de X,*

(a) $\sum_{s \in X^0} T(G_s; (C_y)_{o(y)=s}, (C_i)_{i \in I_s}) \leqq \sum_{s \in Y^0} T(H_s; (D_y)_{o(y)=s}, (C_i)_{i \in J_s})$,

(b) *si Y' est une réduction de Y, il existe une réduction de X' telle que Y' soit un raffinement de X'.* □

Maximalité. Grâce au lemme d'accessibilité III.1.2, on obtient l'existence d'un raffinement *maximal* à une décomposition rigide d'un groupe en graphe de groupe : c'est une décomposition rigide de G ayant un nombre maximal de sommets, tous essentiels.

On vérifie immédiatement que si la décomposition $G = \pi_1(X)$ est maximale, et si X' est une réduction de X, alors $G = \pi_1(X')$ est encore maximale. Ces notions de raffinement et de maximalité sont utiles à cause des deux lemmes suivants.

Lemme III.1.4. *Soit $G = \pi_1(X)$ une décomposition rigide sans sommets inessentiels et maximale de $(G; C_1, \ldots, C_n)$.*

(1) *Si x est une arête orientée non réduite, $\bar{x}$ est réduite.*

(2) *Si X_s est une décomposition de $(G_s; (C_y)_{o(y)=s}, (C_i)_{i \in I_s})$ dont les stabilisateurs d'arêtes sont conjugés à des sous-groupes de C_y, $y \in X^1$, alors X_s est un arbre et a exactement un sommet essentiel.*

(3) *Soit X_s comme dans* (2), *et G_s' le stabilisateur du sommet essentiel s de X_s. Il existe un sous-ensemble Y_s de l'ensemble des arêtes issue de s dans X, une décomposition X_s' de $(G_s; (C_y)_{o(y)=s}, (C_i)_{i \in I_s})$ dont les sommets sont indexés par $\{s\} \bigcup Y_s$, dont l'unique sommet essentiel, noté s, a toujours G_s' comme stabilisateur, et dont les autres stabilisateurs sont les (C_y) pour $y \in Y_s$ (Figure* VII).

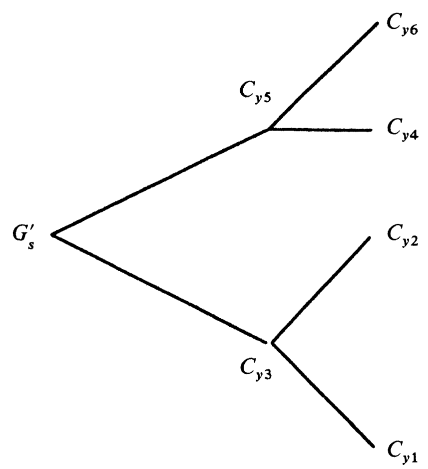


Figure VII

(1) Supposons que l'arête x a deux sommets distincts, s et t, et supposons que $(G_s; (C_i)_{i \in I_s}, (C_z)_{o(z)=s})$ ainsi que $(G_t; (C_i)_{i \in I_t}, (C_z)_{o(z)=t})$ se décomposent en $H_s *_{C_y} C_x$ et $C_x *_{C_z} H_t$ (ceci est un peu simplificateur, il faudrait écrire plutôt un graphe de groupe de ayant un stabilisateur de sommet égal à C_x mais cela ne change rien au raisonnement); on obtient alors de façon évidente une nouvelle décomposition de G ayant un sommet de plus de stabilisateur C_y : celui-ci doit être inessentiel, c'est-à-dire : $C_x = C_y *_F C_z$. En raisonnant comme au I.5.2, on construit alors une action de $(G; C_1, \ldots, C_n)$ dans un arbre ou C_x ne fixe aucun sommet, ce qui est une contradiction et démontre (1).

Le point (2) se montre en remplaçant le sommet s par le graphe X_s, ce qui contredit la maximalité sauf si les nouveaux sommets introduits sont inessentiels.

Le point (3) se montre en trois étapes illustrées sur les Figures VIII.

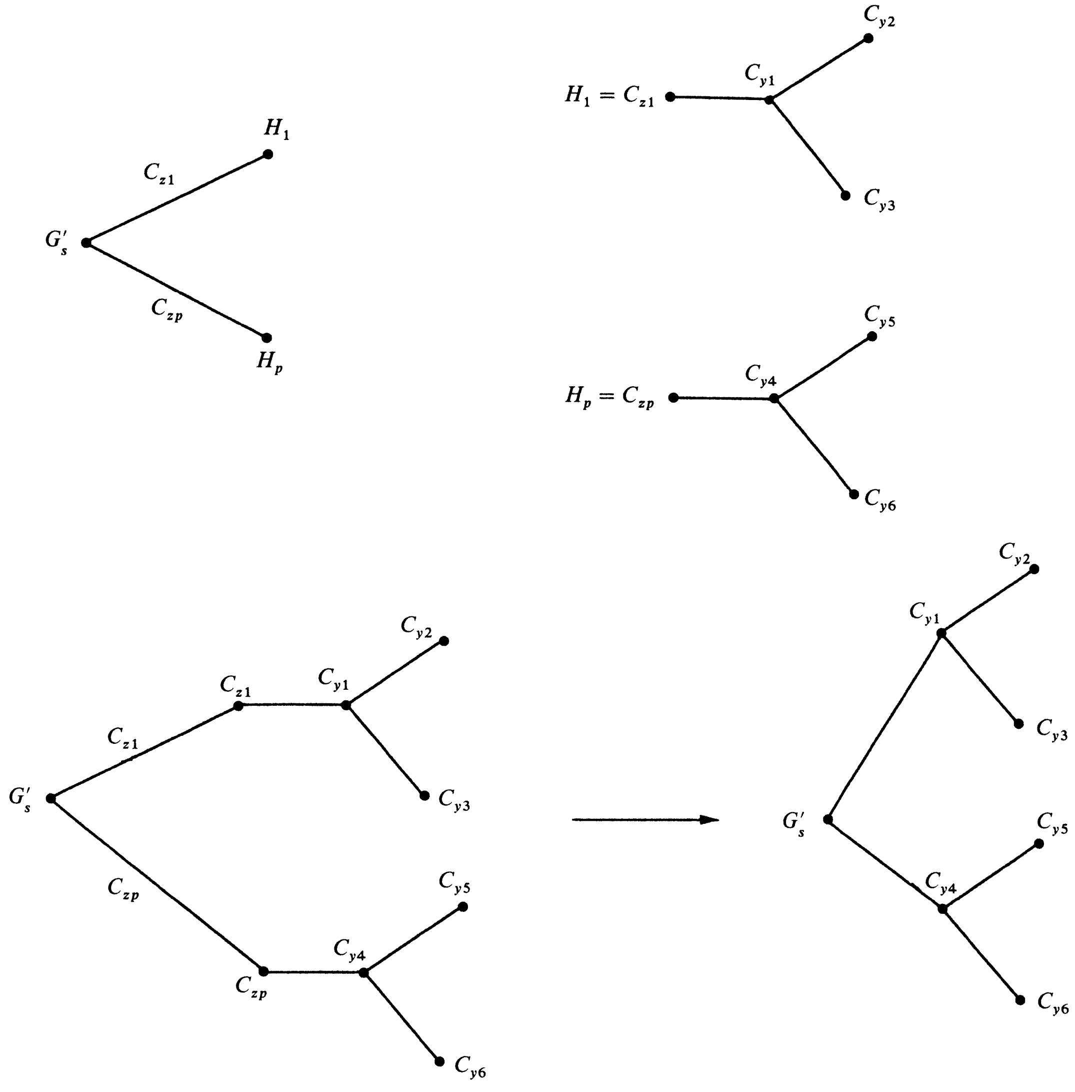


Figure VIII

Soient $z_1, \ldots, z_p$ les arêtes de X_s issues de s stabilisateurs C_z, et $Z_1, \ldots, Z_p$ les graphes des groupes complémentaires de ces arêtes dans X_s indexes de façon évidentes. On considère alors le graphe Y_s ayant $p+1$ sommets de stabilisateurs G'_s et $H_i = \pi_1(Z_i)$ et p arêtes $z_1, \ldots, z_n$ de stabilisateurs C_{z_i}.

D'après le (2), ces sommets sont inessentiels, autrement dit, il existe une partition $Z_s \bigcup_{1 \leq i \leq k} Y_i$ de l'ensembles des arêtes issue de s telle que:

– $C_y \subset G'_s$ si $y \in Z_s$ et $C_i \subset G'_s$ si $i \in I_s$.

– H_i est le groupe fondamental d'un arbre de groupe T_i ayant un sommet de stabilisateur C_{z_i} et $\# Y_i$ sommets de stabilisateurs C_y, $y \in Y_i$.

On obtient le graphe X_s' en *réduisant* les arêtes $\bar{y}_i$ grâce aux graphes T_i ainsi construits. Les sommets de X_s' ont pour stabilisateurs G_s' ainsi que les C_y, $y \in \bigcup Y_i = Y_s$. □

Lemme III.1.5. (a) *Si la décomposition $G = \pi(X)$ est réduite, rigide sans sommets inessentiels et maximale, il existe une section i de l'application j telle que pour tout sommet s, induise un isomorphisme de $\hat{G}_{i(s)}$ et G_s'.*

(b) *Si la décomposition $G = \pi(X)$ est rigide maximale et sans sommets inessentiels, elle admet une réduction X': $(G_s'; (C_y'), (C_i)_{i \in I_s})$ pour laquelle il existe une section i de l'application j telle que pour tout sommet s, J induise un isomorphisme de $\hat{G}_{i(s)}$ et G_s.*

On fait agir G sur l'arbre $\hat{T}$ via I, et on restreint cette action à G_s; le quotient $G_s \backslash \hat{T}$ est une description de G_s comme groupe fondamental d'un graphe de groupes. Comme G_s est finiment engendré modulo les C_y (pour $o(y) = s$), ce graphe a même groupe fondamental qu'un sous-graphe fini X_s, et l'on choisit X_s ayant un nombre minimal d'arêtes. Grâce au lemme III.1.1, chacun des C_y fixe un point dans $\hat{T}$. On peut donc appliquer le lemme III.1.4.2) qui montre que X_s a exactement un sommet essentiel. Soit G_s' le stabilisateur de ce sommet.

Par construction, $I(G_s')$ fixe un sommet t de $\hat{T}$, et le lemme II.1.1 montre que $J(\hat{G}_t) \subset G_{j(t)}$. Si $t \neq s$, G_s' est donc contenu dans le stabilisateur d'une arête issue de s et G_s' est inessentiel. Ainsi, $J(\hat{G}_t) \subset G_s'$ et $I(G_s') \subset \hat{G}_t$ avec $j(t) = s$; on pose $i(s) = t$, et I induit un isomorphisme de G_s' et $\hat{G}_{i(s)}$.

Montrons qu'il existe une réduction X' de X dont l'ensemble des stabilisateurs de sommets est l'ensemble des G_s', et dont les stabilisateurs d'arêtes sont contenus dans les $(C_y)_{o(y)=s}$. Il suffit de remplacer successivement chacun des sommets par le graphe Y_s obtenu au (3) du lemme III.1.4, et d'effectuer les réductions correspondantes, ce qui est loisible à cause du point (1) du même lemme. □

Montrons enfin que si i est l'application construite au lemme précédent on a:

Lemme III.1.6. (a) *Si X est réduite,*

$$T(G_s; (C_i)_{i \in I_s}, (C_y)_{o(y)=s}) \leqq T(\hat{G}_{i(s)}; (C_i)_{i \in \hat{I}_{i(s)}}, (\hat{C}_y)_{o(y)=i(s)}).$$

(b) *Sinon, la réduction X' construite au lemme* III.1.5 *satisfait*:

$$T(G_s'; (C_i)_{i \in I_s}, (C_y')_{o(y)=s}) \leqq T(\hat{G}_{i(s)}; (C_i)_{i \in \hat{I}_{i(s)}}, (\hat{C}_y)_{o(y)=i(s)}).$$

On montre (b) qui contient (a). De fait pour tout arête $y = [s, s']$ issue de s, $I(C_y')$ fixe une arête de $\hat{T}$ (deux sommets distincts) et est donc contenu dans un conjugué de $\hat{C}_z$ pour une certaine arête z issue de $i(s)$.

D'autre part la construction de X' montre que C_y' contient le stabilisateur $(J(\hat{C}_z))$ sous l'action de G_s' d'une arête issue de $i(s)$.

Ainsi, pour toute arête z issue de $i(s)$, il existe une arête y issue de s et un $a \in G'_s$ telle que $\hat{C}_z \subset aI(C'_y)a^{-1}$ et réciproquement pour toute arête issue de $i(s)$, il existe un $b \in \hat{G}_{i(s)}$ et une arête z telle que $I(C'_y) \subset b\hat{C}_z b^{-1}$.

•

De même, si $i \in \hat{I}_s$, $C_i \subset \mathrm{hat}\, G_s$, donc $J(C_i) \subset G'_s$, et soit i in G'_s soit C_i fixe au moins deux points dans l'arbre de X' et est donc contenu dans l'un des C'_y pour $o(y) = s$; réciproquement, si $i \in I_s$ soit $i \in \hat{I}_s$ soit $I(C_i)$ est contenu dans le stabilisateur d'une arête issue de s dans $\hat{X}$.

Ceci permet d'appliquer le lemme I.1.2 (2), et l'on obtient:

$$T\big(G'_s; (C'_s)_{o(y)=s}, (C_i)_{i \in I_s}\big) = T\big(\hat{G}_s; (I(C'_y))_{o(y)=s}, (I(C_i))_{i \in I_s}\big) = T\big(\hat{G}_s; (\hat{C}_z), (C_i)_{i \in \hat{I}_s}\big). \qquad \square$$

III.3. Conclusion. On voit maintenant la démonstration du théorème II : grâce à l'étude faite au premier chapitre, on sait que pour montrer le théorème II, il suffit d'étudier le cas d'une décomposition rigide (lemme I.4.1) dont tous les sommets sont essentiels (lemme I.5.4). Soit X un tel graphe. Grâce au lemme III.1.3, on peut supposer X maximal. Soit $\hat{X}$ le graphe construit en II.1.1 à partir de X. Par construction:

$$\sum T\big(\hat{G}_s; (\hat{C}_y)_{o(y)=s}, (C_i)_{i \in \hat{I}_s}\big) \leqq T(G; C_1, \ldots, C_n).$$

Les lemmes III.1.4 et III.1.5 permettent alors de construire une réduction X' de X pour laquelle on a

$$\sum T\big(G'_s; (C'_s)_{o(y)=s}, (C_i)_{i \in I_s}\big) \leqq T\big(\hat{G}_s; (\hat{C}_y)_{o(y)=s}, (C_i)_{i \in \hat{I}_s}\big) \leqq T(G; C_1, \ldots, C_n).$$

C'est ce que promet le théorème II (3); si X était réduite, $X' = X$.

La partie (2) du théorème II, qui contient (1) résulte alors de I.4.2.

Bibliographie

[B-F] *M. Bestvina, M. Feighn*, Bounding the complexity of simplicial group actions on trees, Invent. Math. **103** (1991), 449–469.

[B-Z] *G. Burde, H. Zieschang*, Knots, de Gruyter Stud. Math. **5**, Berlin 1985.

[D-D] *W. Dicks, M.J. Dunwoody*, Groups acting on graphs, Cambridge stud. adv. math. **17**, Cambridge.

[D] *M.J. Dunwoody*, The accessibility of finitely presented groups, Invent. Math. **81** (1985), 449–457.

[D-F] *M.J. Dunwoody, R.A. Fenn*, On the finiteness of higher knot sums, Topology **26**, No. 3 (1987), 337–343.

[Ha] *A. Haefliger*, Complexes of groups and orbihedra, in: Group theory from a geometrical viewpoint, E. Ghys, A. Haefliger, A. Verjovski Editors, World Scientific, 1991.

[H] *J. Hempel*, 3-manifolds, Ann. Math. Stud. **86**, Princeton University Press, 1976.

[He] *F. Herrlich*, Graph of groups with isomorphic fundamental groups, Arch. Math. **51**, No. 3 (1988), 232–237.

[J] *W. Jaco*, Lectures on three manifold topology, CBMS Reg. Conf. Ser. **43**, AMS (1980).

[K-Z] *R. Kaufmann, H. Zieschang*, On the rank of NEC groups, in: Discrete groups and geometry, Harvey and Maclachlan Editors, Lond. Math. Soc. Lect. Note Ser. **173** (1992), 137–147.

[K] *D.G. Kramtsov*, Finite graphs of groups with isomorphic fundamental groups, Algebra i Logica **30**, No. 5 (1991), 595–623.

[L-S] *R.C. Lyndon, P.E. Schupp*, Combinatorial group theory, Springer, 1977.

[P-R-Z] *N. Peczinski, G. Rosenberger, H. Zieschang*, Über Erzeugende ebener diskontinuierlicher Gruppen, Inv. Math. **29** (1975), 161–180.

[S-W] *P. Scott, C.T.C. Wall*, Topological methods in group theory, in: Homological group theory, C.T.C. Wall, Editor, London Math. Soc. Lect. Note **36** (1979), 137–173.

[Se] *Z. Sela*, Acylindrical accessibility for groups, preprint MPI 1992.

[S] *J.-P. Serre*, Arbres, amalgames, SL_2, Astérisque **46** (1977).

[W] *C.T.C. Wall*, Pair of relative cohomological dimension one, J. Pure Appl. Algebra **1** (1971), 141–154.

[Z] *H. Zieschang*, Über die Nielsensche Kürzungsmethode in freien Produkten mit Amalgam, Invent. Math. **190** (1970), 4–37.

UFR de Mathématique et d'Informatique, Université Louis Pasteur, 7, rue René Descartes, F-67084 Strassbourg
e-mail: delzant@math.u-strasbg.fr

Eingegangen 7. Juli 1994, in revidierter Fassung 3. Juli 1995

J. reine angew. Math. **470** (1996), 181—215

Journal für die reine und angewandte Mathematik

The stabilisation trick for coactions

By *Siegfried Echterhoff* at Paderborn and *Iain Raeburn* at Newcastle

---

Crossed products by coactions of a locally compact group G were introduced to allow the application of duality techniques to crossed products by actions of nonabelian groups. When G is abelian, a coaction δ of G on a C^*-algebra A is given by an action α of the dual group $\hat{G}$ on A, and the crossed product $A \times_\delta G$ by the coaction is isomorphic to the usual crossed product $A \times_\alpha \hat{G}$. Thus in general one expects crossed products by coactions to have important features in common with ordinary crossed products, especially those features which are apparantly true only for abelian groups.

In analysing an ordinary crossed product $A \times_\alpha H$, a standard technique involves choosing a normal subgroup N for which N and H/N are more manageable than H, and decomposing $A \times_\alpha H$ as an iterated crossed product $(A \times_\alpha N) \times H/N$. The action of H/N on $A \times_\alpha N$ is often twisted by a cocycle, and there are two different kinds of twisted crossed product for which one always has such a decomposition (see [G] and [PaR]). A corresponding theory for coactions is potentially very useful, since crossed products by coactions of abelian groups are familiar objects, and one can certainly have N and G/N abelian but G nonabelian. One notion of twisted crossed products by coactions was developed in [PR], and a decomposition theorem obtained which is analogous to that of Green for actions [G], Proposition 1. In particular, the twisted crossed product $A \times_{\delta, W} G$ is the quotient of $A \times_\delta G$ by a "twisting ideal" I_W, and, as in [G], its properties can often be deduced from those of $A \times_\delta G$.

The effective use of these decompositions of crossed products has been limited by the need to extend the established theory to the appropriate class of twisted crossed products. The stabilisation trick of [PaR], which says that every twisted action is Morita equivalent to an ordinary action (cf. [Kal]), provides an attractive route for extending theory cheaply. In [PR], the question of finding a similar stabilisation trick for twisted coactions was raised, but left open: the version in [PaR] seemed badly suited to the twisted actions of Green on which [PR] was based.

Recently the first author has given a version of the stabilisation trick which is compatible with Green's theory [Ech]. The construction in [Ech] is deeper than that of [PaR], and depends crucially on Green's imprimitivity theorem; however, the Morita equivalence thus obtained is compatible with induction of representations and the Mackey machine

[Ech], § 3. Here we shall show that a similar approach, based on Mansfield's imprimitivity theorem for crossed products by coactions, gives a stabilisation trick for twisted coactions which is compatible with Mansfield's theory of induced representations [M].

Our main theorem says that for every twisted coaction (δ, W) of $(G, G/N)$ on a C^*-algebra A, there is an ordinary coaction ε of N on B, say, which is Morita equivalent to (δ, W). This implies, for example, that the twisted crossed product $A \times_{\delta, W} G$ is Morita equivalent to $B \times_{\varepsilon} N$. After a short section on preliminaries, in which we set up terminology and recall the basic properties of twisted coactions, we discuss the relevant notion of Morita equivalence in § 2, and prove our main theorem in § 3. The idea is to prove that the equivalence in Mansfield's imprimitivity theorem is equivariant, and then the one we want will be the quotient of this corresponding to a twisting ideal. In our last section, we give some applications. We prove first that twisted coactions are Morita equivalent if and only if the dual actions are Morita equivalent in the sense of Combes [Com], and then that our stabilisation theorem is compatible with the induction process of [M].

Morita equivalence for coactions and twisted coactions was first discussed by Bui [Bui], who based his theory on results of Baaj and Skandalis about coactions on Hilbert modules [BS]. The resulting theory is asymmetrical: (A, δ_A) is Morita equivalent to (B, δ_B) if there is a coaction δ_X on a Hilbert B-module X which is compatible with δ_B, such that A can be naturally identified with $\mathscr{K}(X)$ and δ_A with a canonical coaction on $\mathscr{K}(X)$ induced by δ_X. In [ER] we gave a symmetric version of the theory, based on imprimitivity bimodules, which we feel is more elegant and easier to work with. However, since Mansfield constructed a Hilbert module X and identified $\mathscr{K}(X)$ rather than directly building an imprimitivity bimodule, here we have had to use the approach of [Bui] alongside our new one. While every Morita equivalence in Bui's sense is certainly one of ours, the converse is not so clear. We shall discuss this point in § 2, since we shall want to use results from [ER], and then prove that equivalences pass to quotients by appropriately induced ideals. In an appendix we shall prove that, at least for nondegenerate coactions, the two notions of equivalence coincide.

Acknowledgements. This research was supported by grants from the Australian Research Council and the University of Paderborn.

§ 1. Preliminaries

Conventions. Throughout, G will be a locally compact group, and N a closed amenable normal subgroup. We denote by $C_r^*(G)$ the reduced group C^*-algebra, acting concretely on $L^2(G)$ via the (integrated form of) the left regular representation λ^G. Since N is amenable, the representation $s \mapsto \lambda_{sN}^{G/N}$ of $C^*(G)$ on $L^2(G/N)$ factors through the reduced algebra $C_r^*(G)$, and induces a surjection q of $C_r^*(G)$ onto $C_r^*(G/N)$ (see, e.g., [M], Lemma 3). We also have $C_r^*(N) \cong C^*(N)$, the full group C^*-algebra of N.

Tensor products of C^*-algebras will always be spatial, so that, for example, if A acts concretely on a Hilbert space $\mathscr{H}$, then $A \otimes C_r^*(G)$ acts on $\mathscr{H} \otimes L^2(G) \cong L^2(G, \mathscr{H})$. If we have a nondegenerate homomorphism ϕ of one C^*-algebra A into the multiplier algebra $\mathscr{M}(B)$ of another, we shall without comment extend ϕ to a strictly continuous homomorphism of $\mathscr{M}(A)$ into $\mathscr{M}(B)$, and denote the extension by ϕ also. We write 1 for the identity of an algebra or the identity operator, and id_A for the identity homomorphism on an algebra A.

We shall denote by M the representation of $C_0(G)$ as multiplication operators on $L^2(G)$, and write either M_f or $M(f)$ for multiplication by the function f. We write w_G for the strictly continuous function $w_G : s \mapsto \lambda_s^G$ in

$$\mathcal{M}(C_0(G, C_r^*(G))) = \mathcal{M}(C_0(G) \otimes C_r^*(G)),$$

and observe that $M \otimes \mathrm{id}_G(w_G)$ is the unitary operator W_G on $L^2(G) \otimes L^2(G) \cong L^2(G \times G)$ given by $(W_G\xi)(s, t) := \xi(s, s^{-1}t)$. This operator arises frequently because it implements the equivalence between $\lambda^G \otimes \lambda^G$ and $\lambda^G \otimes 1$:

$$W_G(\lambda_s^G \otimes 1) W_G^* = \lambda_s^G \otimes \lambda_s^G.$$

This equation implies that the integrated form of $\lambda^G \otimes \lambda^G$ factors through the regular representation, and hence gives a nondegenerate homomorphism δ_G of $C_r^*(G)$ into the subalgebra $\mathcal{M}(C_r^*(G) \otimes C_r^*(G))$ of $B(L^2(G \times G))$, which is called the *comultiplication on* $C_r^*(G)$, and satisfies $(\delta_G \otimes \mathrm{id}) \circ \delta_G = (\mathrm{id} \otimes \delta_G) \circ \delta_G$.

Coactions, twisted coactions, and crossed products. Our conventions concerning coactions and crossed products will be those of [LPRS] and [PR]. Thus a *coaction* on a C^*-algebra A is an injective and nondegenerate homomorphism $\delta : A \to \mathcal{M}(A \otimes C_r^*(G))$ such that

(1) $\delta(a)(1 \otimes z), (1 \otimes z)\delta(a) \in A \otimes C_r^*(G)$ for $a \in A$ and $z \in C_r^*(G)$;

(2) $(\delta \otimes \mathrm{id}_G) \circ \delta = (\mathrm{id}_A \otimes \delta_G) \circ \delta$.

The coaction is *nondegenerate* if in addition

(3) $\delta(A)(1 \otimes C_r^*(G))$ is dense in $A \otimes C_r^*(G)$.

(In general, if an algebra C acts on a Banach space X, then $C \cdot X$, or just CX, will stand for the closed linear span of the set $\{c \cdot x : c \in C,\ x \in X\}$.) Note that nondegeneracy is automatic if G is amenable [L], [K] or discrete [BS], 7.15. The *crossed product* $A \times_\delta G$ is defined in terms of a faithful representation of A on $\mathcal{H}$ as the closed span in $B(\mathcal{H} \otimes L^2(G))$ of

$$\{\delta(a)(1 \otimes M_f) : a \in A, f \in C_0(G)\}.$$

This is independent of the choice of representation of A, and indeed the triple

$$(A \times_\delta G, \delta, 1 \otimes M)$$

has a universal property with respect to *covariant representations*: pairs (π, μ) of nondegenerate representations $\pi : A \to B(\mathcal{H})$, $\mu : C_0(G) \to B(\mathcal{H})$ satisfying

$$\pi \otimes \mathrm{id}_G(\delta(a)) = \mu \otimes \mathrm{id}_G(w_G)(\pi(a) \otimes 1)\mu \otimes \mathrm{id}_G(w_G^*) \quad \text{for } a \in A.$$

Given such a pair, we write $\pi \times \mu$ for the representation of $A \times_\delta G$ such that

$$\pi \times \mu\big(\delta(a)(1 \otimes M_f)\big) = \pi(a)\,\mu(f),$$

and we sometimes write j_A, $j_{C(G)}$ for the embeddings δ, $1 \otimes M$ of A, $C_0(G)$ in $\mathcal{M}(A \times_\delta G)$ (see [PR], §1, for a discussion of these points, and detailed references to [LPRS]). It follows easily from the universal property that there is a dual action $\hat{\delta}$ of G on $A \times_\delta G$ such that $\hat{\delta}_s\big(\delta(a)(1 \otimes M(f))\big) = \delta(a)\big(1 \otimes M(\sigma_s(f))\big)$, where $\sigma_s(f)(t) := f(ts)$; indeed, $\hat{\delta}$ is implemented spatially on $\mathcal{H} \otimes L^2(G)$ by the representation $1 \otimes \varrho$, where ϱ is the right regular representation of G.

Example 1.1. Suppose $\alpha : G \to \operatorname{Aut} B$ is an action of G, and i_B, i_G are the canonical embeddings of B, G in the multiplier algebra $\mathcal{M}(B \times_{\alpha,r} G)$ of the reduced crossed product, viewed concretely on $L^2(G, \mathcal{H})$, say. Then the pair $(i_B \otimes 1, i_G \otimes \lambda^G)$ is a covariant representation of (B, G, α) which induces a nondegenerate coaction

$$\hat{\alpha} : B \times_{\alpha,r} G \to \mathcal{M}\big((B \times_{\alpha,r} G) \otimes C_r^*(G)\big)$$

called the *dual coaction* (see [R], §3, for a discussion of why we use the reduced crossed product here). The duality theorem of Imai and Takai [IT] asserts that the second dual system $\big((B \times_{\alpha,r} G) \times_{\hat{\alpha}} G, \hat{\hat{\alpha}}\big)$ is isomorphic to $\big(B \otimes \mathcal{K}(L^2(G)), \alpha \otimes \operatorname{Ad} \varrho\big)$.

We can restrict a coaction δ of G to a coaction of G/N by composing with the canonical surjection q of $C_r^*(G)$ onto $C_r^*(G/N)$: the restriction $\delta|$ is by definition $(\mathrm{id}_A \otimes q) \circ \delta$ (cf. [M], Lemma 4). Following [PR], we say the coaction δ is given by a twist on G/N if $\delta|$ is implemented by a unitary $W \in \mathcal{M}\big(A \otimes C_r^*(G/N)\big)$, which is both a corepresentation of G/N (Condition (1) below), and satisfies an additional consistency condition ((3) below). To be more precise, we need to introduce the flip isomorphism $\sigma = \sigma_{C,D} : c \otimes d \mapsto d \otimes c$ of $C \otimes D$ onto $D \otimes C$.

Definition 1.2. A *twisted coaction* of $(G, G/N)$ on A is a pair (δ, W) where δ is a coaction of G and $W \in \mathcal{UM}\big(A \otimes C_r^*(G/N)\big)$ satisfies

(1) $(W \otimes 1)\big(\mathrm{id}_A \otimes \sigma_{G/N, G/N}(W \otimes 1)\big) = (\mathrm{id}_A \otimes \delta_{G/N})(W)$;

(2) $\delta|(a) = W(a \otimes 1)W^*$ for all $a \in A$; and

(3) $\delta \otimes \mathrm{id}_{G/N}(W) = \mathrm{id}_A \otimes \sigma_{G/N, G}(W \otimes 1_{C_r^*(G)})$.

A covariant representation (π, μ) of (A, G, δ) *preserves the twist* W if

$$\mu \otimes \mathrm{id}_{G/N}(w_{G/N}) = \pi \otimes \mathrm{id}_{G/N}(W),$$

and the *twisted crossed product* $A \times_{\delta,W} G$ is the quotient of $A \times_\delta G$ by the *twisting ideal*

$$I_W := \bigcap \{\ker \pi \times \mu : (\pi, \mu) \text{ is covariant and preserves } W\}.$$

For $n \in N$, the dual action $\hat{\delta}_n$ leaves I_W invariant, and hence induces a *dual action* of N on $A \times_{\delta, W} G$ [PR], §4.

We shall be interested in two main classes of twisted coactions, both discussed in [PR].

Example 1.3. If ε is a coaction of N on A, then the composition $\operatorname{Inf}\varepsilon$ of ε with the natural embedding $\mathrm{id}_A \otimes \lambda^G|_N$ of $A \otimes C^*(N)$ in $\mathcal{M}(A \otimes C_r^*(G))$ is a coaction of G on A, and $1 \otimes 1$ is a twist for $\operatorname{Inf}\varepsilon$ [PR], Example 2.4. The twisted crossed product $A \times_{\operatorname{Inf}\varepsilon, 1\otimes 1} G$ is naturally isomorphic to $A \times_\varepsilon N$ [PR], Example 2.14. Note that $\operatorname{Inf}\varepsilon$ is always nondegenerate: since N is amenable, ε is nondegenerate by [L], Lemma 3.8. Using

$$\lambda^G|_N(C^*(N))\, C_r^*(G) = C_r^*(G)$$

we obtain $\operatorname{Inf}\varepsilon(A)(1 \otimes C_r^*(G)) = (\mathrm{id}_A \otimes \lambda^G|_N)(\varepsilon(A)(1 \otimes C^*(N)))(1 \otimes C_r^*(G))$, which is dense in $A \otimes C_r^*(G)$ since $\varepsilon(A)(1 \otimes C^*(N))$ is dense in $A \otimes C^*(N)$.

Example 1.4. Suppose δ is a coaction of G on A, and let $B = A \times_{\delta|} G/N$. Then $(\pi, \mu) := ((j_A \otimes \mathrm{id}_G) \circ \delta, j_{C(G/N)} \otimes 1)$ is a covariant representation of $(A, G/N, \delta|)$ whose integrated form $\delta_B = \pi \times \mu : B \to \mathcal{M}(B \otimes C_r^*(G))$ is a coaction of G on B. If

$$W_B = (j_{C(G/N)} \otimes \mathrm{id}_{G/N})(w_{G/N}) \in \mathcal{M}(B \otimes C_r^*(G/N)),$$

then (δ_B, W_B) is a twisted coaction of $(G, G/N)$ on B [PR], Lemma 3.3. We call this the *decomposition coaction*, because $B \times_{\delta_B, W_B} G$ is naturally isomorphic to $A \times_\delta G$ by [PR], Theorem 3.1. It can be routinely verified that δ_B is nondegenerate whenever δ is nondegenerate.

Multipliers of imprimitivity bimodules. An *imprimitivity bimodule* ${}_AX_B$ is an A-B bimodule which is simultaneously a full left Hilbert A-module and a full right Hilbert B-module in such a way that the following compatibility conditions are satisfied:

$$\langle a \cdot x, y\rangle_B = \langle x, a^* \cdot y\rangle_B; \quad {}_A\langle x \cdot b, y\rangle = {}_A\langle x, y \cdot b^*\rangle; \quad {}_A\langle x, y\rangle \cdot z = x \cdot \langle y, z\rangle_B.$$

If there is an A-B imprimitivity bimodule ${}_AX_B$, we say that the C^*-algebras A and B are *Morita equivalent*. Such algebras have essentially the same representation theory, and, more importantly for us, the lattices of ideals in both A and B are isomorphic to the lattice of closed sub-bimodules of X [Rieff 2], §3. (This is unambiguous: the norms on X induced by the two inner products coincide [Rieff 2], Proposition 3.1.)

If X and Y are Hilbert B-modules, then $\mathcal{L}_B(X, Y)$, or just $\mathcal{L}(X, Y)$, will denote the Banach space of adjointable B-linear maps $T: X \to Y$, and $\mathcal{K}(X, Y)$ the subspace of compact adjointable maps. The *multiplier bimodule* $\mathcal{M}({}_AX_B)$ of an imprimitivity bimodule ${}_AX_B$ is the collection of pairs $m_A \in \mathcal{L}_A(A, X)$, $m_B \in \mathcal{L}_B(B, X)$ such that $m_A(a) \cdot b = a \cdot m_B(b)$; one thinks of the multiplier $m = (m_A, m_B)$ as specifying the maps

$$m_A : a \mapsto a \cdot m, \quad m_B : b \mapsto m \cdot b.$$

The space $\mathcal{M}(X)$ is naturally an $\mathcal{M}(A)$-$\mathcal{M}(B)$ bimodule, and the inner products on X extend uniquely to $\mathcal{M}(A)$-, $\mathcal{M}(B)$-valued inner products on $\mathcal{M}(X)$, though $\mathcal{M}(X)$ is not in general an $\mathcal{M}(A)$-$\mathcal{M}(B)$ imprimitivity bimodule (see [ER], §1). It was proved in [ER], 1.3, that $m \mapsto m_B$ is an isomorphism of $\mathcal{M}(X)$ onto $\mathcal{L}_B(B, X)$, and $m \mapsto m_A$ an isomorphism of $\mathcal{M}(X)$ onto $\mathcal{L}_A(A, X)$; conversely, if we start with a full Hilbert B-module X, we can view it as a $\mathcal{K}(X)$-B imprimitivity bimodule, and use $\mathcal{L}_B(B, X)$ in place of the multiplier bimodule $\mathcal{M}({}_{\mathcal{K}(X)}X_B)$. We shall also need the concept of an *imprimitivity bimodule homomorphism* $\Phi : {}_AX_B \to \mathcal{M}({}_CY_D)$: a triple $\Phi = (\Phi_A, \Phi_X, \Phi_B)$ in which

$$\Phi_A : A \to \mathcal{M}(C), \quad \Phi_B : B \to \mathcal{M}(D)$$

are homomorphisms, and $\Phi_X : X \to \mathcal{M}(Y)$ satisfies

$${}_{\mathcal{M}(C)}\langle \Phi_X(x), \Phi_X(y)\rangle = \Phi_A({}_A\langle x, y\rangle), \quad \Phi_X(a \cdot x) = \Phi_A(a) \cdot \Phi_X(y),$$

and similarly on the right [ER], 1.8.

We shall denote the dual of an imprimitivity bimodule ${}_AX_B$ by ${}_B\tilde{X}_A$. The set $L := \begin{pmatrix} A & X \\ \tilde{X} & B \end{pmatrix}$, equipped with multiplication and involution given by

$$\begin{pmatrix} a_1 & x_1 \\ \tilde{y}_1 & b_1 \end{pmatrix}\begin{pmatrix} a_2 & x_2 \\ \tilde{y}_2 & b_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + {}_A\langle x_1, y_2\rangle & a_1 \cdot x_2 + x_1 \cdot b_2 \\ \tilde{y}_1 \cdot a_2 + b_1 \cdot \tilde{y}_2 & \langle y_1, x_2\rangle_B + b_1b_2 \end{pmatrix},$$

$$\begin{pmatrix} a & x \\ \tilde{y} & b \end{pmatrix}^* = \begin{pmatrix} a^* & y \\ \tilde{x} & b^* \end{pmatrix},$$

is a C^*-algebra called the *linking algebra* for X. (It is naturally isomorphic to the algebra $\mathcal{L}(X \oplus B)$ of adjointable operators on the Hilbert B-module $X_B \oplus B_B$ [BGR].) The C^*-norm on L induces the C^*-norms on the diagonal corners A, B, and the norm on X as the upper right-hand corner of L agrees with the norm on X induced by the inner product. Conversely, if there is a C^*-algebra L containing copies of A and B as full complementary corners, then the upper right-hand corner is naturally an A-B imprimitivity bimodule. It was shown in [ER], Proposition A1, that the multiplier algebra $\mathcal{M}(L)$ of the linking algebra can be conveniently identified with $\begin{pmatrix} \mathcal{M}(A) & \mathcal{M}(X) \\ \mathcal{M}(\tilde{X}) & \mathcal{M}(B) \end{pmatrix}$, with multiplication and involution defined similarly. In the same way, we can identify $\mathcal{M}(L \otimes C_r^*(G))$ with $\begin{pmatrix} \mathcal{M}(A \otimes C_r^*(G)) & \mathcal{M}(X \otimes C_r^*(G)) \\ \mathcal{M}(\tilde{X} \otimes C_r^*(G)) & \mathcal{M}(B \otimes C_r^*(G)) \end{pmatrix}$.

§ 2. Morita equivalence of twisted coactions

Definition 2.1. Suppose that δ_B is a coaction of G on a C^*-algebra B, and X is a Hilbert B-module. Following [BS], 2.2, a *coaction of G on X compatible with* δ_B is a linear map

$$\delta_X : X \to \mathcal{L}_{B \otimes C_r^*(G)}(B \otimes C_r^*(G), X \otimes C_r^*(G))$$

satisfying

(1) $\delta_X(xb) = \delta_X(x) \cdot \delta_B(b)$ and

$$\delta_B(\langle x, y \rangle_B) = \langle \delta_X(x), \delta_X(y) \rangle_{\mathcal{M}(B)} (:= \delta_X(x)^* \circ \delta_X(y));$$

(2) $(1 \otimes z) \cdot \delta_X(x), \delta_X(x) \cdot (1 \otimes z) \in X \otimes C_r^*(G)$;

(3) $\delta_X(X) \cdot (B \otimes C_r^*(G))$ is dense in $X \otimes C_r^*(G)$;

(4) $(\delta_X \otimes \mathrm{id}_G) \circ \delta_X = (\mathrm{id}_X \otimes \delta_G) \circ \delta_X$.

Following [ER], one can make sense of these statements by identifying

$$\mathcal{L}(B \otimes C_r^*(G), X \otimes C_r^*(G))$$

with the multiplier bimodule of the imprimitivity bimodule

$$_{\mathcal{K}(X) \otimes C_r^*(G)}(X \otimes C_r^*(G))_{B \otimes C_r^*(G)},$$

so that δ_X by definition maps X into $\mathcal{M}(X \otimes C_r^*(G))$. Now the general properties of multiplier bimodules allow us to interpret (1), (2) and (3). As in [ER], Example 1.10, a pair of maps (δ_X, δ_B) satisfying these three conditions induces a nondegenerate homomorphism $\delta_{\mathcal{K}(X)}: \mathcal{K}(X) \to \mathcal{L}(X \otimes C_r^*(G))$ such that $(\delta_{\mathcal{K}(X)}, \delta_X, \delta_B)$ is an imprimitivity bimodule homomorphism, which therefore extends uniquely to the multiplier bimodule $\mathcal{M}(X)$ [ER], Proposition 1.9. Using this and similar extensions one can see that the *coaction identity* (4) makes sense: both sides are linear maps of $\mathcal{M}(X)$ into $\mathcal{M}(X \otimes C_r^*(G) \otimes C_r^*(G))$.

Remarks. In [BS] it is also hypothesised that the maps $\delta_X \otimes \mathrm{id}_G$ and $\mathrm{id}_X \otimes \delta_G$ extend to the multiplier bimodules: the above discussion suggests that this is automatic. One can verify quite easily that the coaction $\delta_{\mathcal{K}(X)}$ we obtain is the same as the one defined in [BS], 2.8, [Bui], 2.8, by $k \mapsto V(k \otimes 1)V^*$.

Definition 2.2. Following [Bui], we shall say that two coactions δ_A, δ_B are *Morita equivalent* if there is an A-B imprimitivity bimodule X, and a coaction δ_X of G on X which is compatible with δ_B and satisfies $\delta_A = \delta_{\mathcal{K}(X)}$ (under the canonical identification of A with the algebra $\mathcal{K}(X)$ of compact operators on the Hilbert B-module X).

Two twisted coactions (δ_A, W_A) and (δ_B, W_B) are *Morita equivalent* if there is a Morita equivalence (X, δ_X) of δ_A and δ_B such that

(5) $(\mathrm{id}_X \otimes q)(\delta_X(x)) = W_A(x \otimes 1)W_B^*$ for all $x \in X$.

As in the comments in Definition 2.1, the equation $\delta_A = \delta_{\mathcal{K}(X)}$ implies that the triple $(\delta_A, \delta_X, \delta_B)$ is an imprimitivity bimodule homomorphism. Thus any Morita equivalence in the sense of Definition 2.2 is also a Morita equivalence in the sense of [ER], and the results of that paper apply. In fact, we show in the appendix that the two notions are usually equivalent. (We point out that Condition (1) in [ER], Definition 3.1, is redundant:

if $(\delta_A, \delta_X, \delta_B)$ is an imprimitivity bimodule homomorphism, which is the content of (2) and (3), and $x = a \cdot y$, then $(1 \otimes z) \cdot \delta_X(x) = (1 \otimes z)\,\delta_A(a) \cdot \delta_X(y)$ is in $X \otimes C_r^*(G)$ because $(1 \otimes z)\delta_A$ is in $A \otimes C_r^*(G)$.)

Example 2.3. Suppose that ε_A, ε_B are coactions of the amenable subgroup N. Then we can inflate ε_A, ε_B to twisted coactions $(\operatorname{Inf}\varepsilon_A, 1 \otimes 1)$, $(\operatorname{Inf}\varepsilon_B, 1 \otimes 1)$ of $(G, G/N)$. If (X, δ_X) is a Morita equivalence between ε_A and ε_B, and $\operatorname{Inf}\delta_X := (\mathrm{id}_X \otimes (\lambda^G|_N)) \circ \delta_X$, then it can be routinely verified that $(X, \operatorname{Inf}\delta_X)$ is a Morita equivalence between $(\operatorname{Inf}\varepsilon_A, 1 \otimes 1)$ and $(\operatorname{Inf}\varepsilon_B, 1 \otimes 1)$. We shall see in Corollary 4.3 that the converse is also true, though not so easy to prove.

Our main Morita equivalence will be a quotient of a Morita equivalence of twisted coactions corresponding to ideals in the algebras which are invariant under the coactions. The notion of δ-invariant ideal in [LPRS], § 4, amounts to saying that δ induces a coaction on the ideal; here we want a coaction on the quotient, for which another definition seems more useful.

Definition 2.4. Let δ be a coaction of G on A, I a closed ideal of A, and $\phi : A \to A/I$ the quotient map. We say that I is *G-invariant* if it is the kernel of the homomorphism

$$(\phi \otimes \mathrm{id}_G) \circ \delta : A \to \mathscr{M}(A/I \otimes C_r^*(G)).$$

If G is amenable, it follows from [LPRS], 4.3, that this is equivalent to the definition of δ-invariant in [LPRS]; we do not know if this is true in general. However, it is an easy matter to check that this one does what we want:

Lemma 2.5 (cf. [LPRS], Lemma 4.6). *Suppose that I is a G-invariant ideal of A with respect to the coaction δ, and $\phi : A \to A/I$ is the quotient map. Then there is a coaction $\delta_{A/I} : A/I \to \mathscr{M}(A/I \otimes C_r^*(G))$ such that*

$$\delta_{A/I}(\phi(a)) = (\phi \otimes \mathrm{id}_G)(\delta(a)) \quad \textit{for } a \in A\,.$$

If δ is nondegenerate, then so is $\delta_{A/I}$, and if W is a twist for δ, then $W_{A/I} := (\phi \otimes \mathrm{id}_{G/N})(W)$ is a twist for $\delta_{A/I}$.

Proposition 2.6. *Suppose that (X, δ_X) implements a Morita equivalence between twisted coactions (δ_A, W_A) and (δ_B, W_B) of $(G, G/N)$. Let J be an ideal in B, and*

$$I = \overline{\mathrm{sp}}_A\langle X \cdot J, X \cdot J\rangle$$

the ideal of A induced from J via X. Then I is a G-invariant ideal of A if and only if J is a G-invariant ideal of B. Moreover, if J is G-invariant, then $(\delta_{A/I}, W_{A/I})$ is Morita equivalent to $(\delta_{B/J}, W_{B/J})$.

For the proof we need:

Lemma 2.7. *Suppose that $\Phi = (\Phi_A, \Phi_X, \Phi_B) : {}_AX_B \to \mathscr{M}({}_CY_D)$ is an imprimitivity bimodule homomorphism, $J = \ker \Phi_B$, and I is the ideal of A induced from J via X. Then $X \cdot J = \ker \Phi_X$ and $I = \ker \Phi_A$.*

Proof. Since $\Phi_X(x \cdot j) = \Phi_X(x) \cdot \Phi_B(j) = 0$ for all $j \in J$, we have $X \cdot J \subseteqq \ker \Phi_X$. On the other hand, $V = \ker \Phi_X$ is a submodule of X such that $\langle V, V \rangle_B$ is contained in $\ker \Phi_B$. Since $Y \mapsto \overline{\mathrm{sp}}\langle Y, Y \rangle_B$ is an isomorphism of the lattice of submodules of X onto the lattice of closed ideals of B, with inverse given by $K \mapsto X \cdot K$ ([Rief2], Theorem 3.1), it follows that $V = X \cdot J$. Similarly, we have $V = \ker \Phi_A \cdot X$, and $\ker \Phi_A$ must be the ideal induced from J. □

Proof of Proposition 2.6. We equip $Y := X/X \cdot J$ with the canonical A/I- and B/J-inner products and actions to obtain an A/I-B/J imprimitivity bimodule. We claim that Y is complete with respect to the norm given by these inner products. To see this, we observe that if $L = \begin{pmatrix} A & X \\ \tilde{X} & B \end{pmatrix}$ is the linking algebra for ${}_A X_B$, and $K = \begin{pmatrix} I & X \cdot J \\ \widetilde{X \cdot J} & J \end{pmatrix}$ is the ideal of L corresponding to J, then $\begin{pmatrix} 0 & Y \\ 0 & 0 \end{pmatrix}$ is just the upper right corner of L/K, and hence complete in the quotient norm. But the quotient norm on L/K is the C^*-norm, and similarly on the diagonal corners, so

$$\begin{aligned} \| x + V \|_{X/V}^2 &= \left\| \begin{pmatrix} 0 & x+V \\ 0 & 0 \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} 0 & x+V \\ 0 & 0 \end{pmatrix}^* \begin{pmatrix} 0 & x+V \\ 0 & 0 \end{pmatrix} \right\| \\ &= \left\| \begin{pmatrix} {}_{A/I}\langle x+V, x+V \rangle & 0 \\ 0 & 0 \end{pmatrix} \right\|_{L/K} \\ &= \| {}_{A/I}\langle x+V, x+V \rangle \| = \| x + V \|_{A/I}^2 ; \end{aligned}$$

thus Y is complete, as claimed.

If $\Phi_A : A \to A/I$, $\Phi_X : X \to Y = X/X \cdot J$ and $\Phi_B : B \to B/J$ are the quotient maps, the triple $\Phi = (\Phi_A, \Phi_X, \Phi_B) : {}_A X_B \to {}_{A/I} Y_{B/J}$ is an imprimitivity bimodule homomorphism, as are the triple $\Phi \otimes \mathrm{id}_G = (\Phi_A \otimes \mathrm{id}_G, \Phi_X \otimes \mathrm{id}_G, \Phi_B \otimes \mathrm{id}_G)$ and the composition

$$(\Phi \otimes \mathrm{id}_G) \circ \delta = \big((\Phi_A \otimes \mathrm{id}_G) \circ \delta_A, (\Phi_X \otimes \mathrm{id}_G) \circ \delta_X, (\Phi_B \otimes \mathrm{id}_G) \circ \delta_B \big).$$

If J is a G-invariant ideal of B, so that $J = \ker\big((\Phi_B \otimes \mathrm{id}_G) \circ \delta_B\big)$, then Lemma 2.7 implies that $I = \ker\big((\Phi_A \otimes \mathrm{id}_G) \circ \delta_A\big)$ is G-invariant. Similarly, if I is G-invariant, so is J.

Suppose now that J, and hence also I, are G-invariant. Since $X \cdot J = \ker \Phi_X$ by Lemma 2.7, there is a well defined linear map δ_Y on $Y = X/X \cdot J$ such that

$$\delta_Y\big(\Phi_X(x)\big) = (\Phi_X \otimes \mathrm{id}_G) \circ \delta_X(x),$$

and the triple

$$(\delta_{A/I}, \delta_Y, \delta_{B/J}) : {}_{A/I}(Y)_{B/J} \to \mathscr{M}\big({}_{A/I \otimes C_r^*(G)}(Y \otimes C_r^*(G))_{B/J \otimes C_r^*(G)}\big)$$

is an injective bimodule homomorphism. It is easy to check, using the corresponding property of δ_X and the equation

$$\delta_Y(Y)\big(B/J \otimes C_r^*(G)\big) = (\Phi_X \otimes \mathrm{id}_G)\big(\delta_X(X)\big(B \otimes C_r^*(G)\big)\big),$$

that δ_Y satisfies condition (3) of Definition 2.1. Thus $\delta_{\mathscr{K}(Y)}$ makes sense, and because we already know $(\delta_{A/I}, \delta_Y, \delta_{B/J})$ is a bimodule homomorphism, it follows that $\delta_{A/I} = \delta_{\mathscr{K}(Y)}$. To see that δ_Y satisfies the coaction identity (4), we compute:

$$\begin{aligned}(\delta_Y \otimes \mathrm{id}_G) \circ \delta_Y \circ \Phi_X &= (\delta_Y \otimes \mathrm{id}_G) \circ (\Phi_X \otimes \mathrm{id}_G) \circ \delta_X \\ &= (\Phi_X \otimes \mathrm{id}_G \otimes \mathrm{id}_G) \circ (\delta_X \otimes \mathrm{id}_G) \circ \delta_X \\ &= (\Phi_X \otimes \mathrm{id}_G \otimes \mathrm{id}_G) \circ (\mathrm{id}_X \otimes \delta_G) \circ \delta_X \\ &= (\mathrm{id}_Y \otimes \delta_G) \circ (\Phi_X \otimes \mathrm{id}_G) \circ \delta_X \\ &= (\mathrm{id}_Y \otimes \delta_G) \circ \delta_Y \circ \Phi_X .\end{aligned}$$

We have now shown that (Y, δ_Y) is a Morita equivalence between $\delta_{A/I}$ and $\delta_{B/J}$. To check that δ_Y respects the twists, we write $y = \Phi_X(x)$, and verify that:

$$\begin{aligned}(\mathrm{id}_Y \otimes q)(\delta_Y(y)) &= \big((\mathrm{id}_Y \otimes q) \circ (\Phi_X \otimes \mathrm{id}_G)\big)(\delta_X(x)) \\ &= (\Phi_X \otimes \mathrm{id}_{G/N})\big((\mathrm{id}_X \otimes q)(\delta_X(x))\big) \\ &= (\Phi_X \otimes \mathrm{id}_{G/N})\big(W_A \cdot (x \otimes 1) \cdot W_B^*\big) \\ &= \big(\Phi_A \otimes \mathrm{id}_{G/N}(W_A)\big) \cdot \big(\Phi_X \otimes \mathrm{id}_{G/N}(x \otimes 1)\big) \cdot \big(\Phi_B \otimes \mathrm{id}_{G/N}(W_B^*)\big) \\ &= W_{A/I} \cdot (y \otimes 1) \cdot W_{B/J}^* .\end{aligned}$$

This completes the proof. □

We finish this section with some examples of G-invariant ideals we shall need later.

Example 2.8. Let (A, G, α) be a C^*-dynamical system with G amenable. Suppose that I is an α-invariant ideal of A, and denote by α_I and $\alpha_{A/I}$ the induced actions of G on I and A/I. Then $J = I \times_{\alpha_I} G$ is an ideal in $B = A \times_\alpha G$, and $B/J \cong A/I \times_{\alpha_{A/I}} G$: the quotient map ϕ is the integrated form of the covariant representation $(j_{A/I} \circ \psi, j_G)$, where $\psi : A \to A/I$ and $j_{A/I}, j_G$ are the canonical embeddings of A/I, G in $\mathscr{M}(A/I \times_{\alpha_{A/I}} G)$ [G], Proposition 12. The dual coaction $\delta_B = \hat{\alpha}$ of G on $B = A \times_\alpha G$ is the integrated form of $(i_A \otimes 1,\ i_G \otimes \lambda^G)$. Thus for $a \in A$ and $s \in G$ we have

$$(\phi \otimes \mathrm{id}_G)\big(\delta_B(i_A(a))\big) = (\phi \otimes \mathrm{id}_G)\big(i_A(a) \otimes 1\big) = j_{A/I}\big(\psi(a)\big) \otimes 1 ,$$

and

$$(\phi \otimes \mathrm{id}_G)\big(\delta_B(i_G(s))\big) = (\phi \otimes \mathrm{id}_G)\big(i_G(s) \otimes \lambda_s^G\big) = j_G(s) \otimes \lambda_s^G .$$

It follows that

$$(\phi \otimes \mathrm{id}_G) \circ \delta_B = \widehat{\alpha_{A/I}} \circ \phi ,$$

which implies that J is a G-invariant ideal in B with $\delta_{B/J} = \widehat{\alpha_{A/I}}$. Similar computations show that the restriction of δ_B to J coincides with $\widehat{\alpha_I}$.

Example 2.9. Let (δ, W) be a twisted coaction of $(G, G/N)$ on A, and let (δ_B, W_B) be the decomposition coaction on $B = A \times_{\delta|} G/N$, as in Example 1.4. Then B/I_W is canonically isomorphic to A by [PR], Example 2.13. We claim that $I := I_W$ is G-invariant with respect to (δ_B, W_B), and that the canonical isomorphism carries $(\delta_{B/I}, W_{B/I})$ into (δ, W). We shall show that

$$(*) \qquad (\phi \otimes \mathrm{id}_G) \circ \delta_B = \delta \circ \phi\,,$$

where $\phi : B \to A$ is the quotient map. Recall from [PR], Example 2.13, that ϕ is the integrated form of the covariant representation (id_A, j), where $j: C_0(G/N) \to \mathcal{M}(A)$ is the homomorphism obtained by slicing the corepresentation W: in other words, $j(f) = S_f(W)$ for f in the Fourier algebra $A(G/N)$ [PR], Remark 2.2. Thus $\phi(\delta(a)(1 \otimes M_f)) = aj(f)$: in particular, if we view δ as a $*$-homomorphism from A into $\mathcal{M}(B)$, we have $\phi \circ \delta = \mathrm{id}_A$. Hence for $a \in A$, $f \in C_0(G/N)$ we get:

$$\begin{aligned}(\phi \otimes \mathrm{id}_G)\big(\delta_B(\delta(a)(1 \otimes M_f))\big) &= (\phi \otimes \mathrm{id}_G)\big((\delta \otimes \mathrm{id}_G)(\delta(a))(1 \otimes M_f \otimes 1)\big)\\ &= (\mathrm{id}_A \otimes \mathrm{id}_G)(\delta(a))\big(j(f) \otimes 1\big)\\ &= \delta\big(aj(f)\big)\,,\end{aligned}$$

where we used that $\delta\big(j(f)\big) = j(f) \otimes 1$ [PR], Remark 2.2. Since the elements $\delta(a)(1 \otimes M_f)$ span a dense subspace of B, this gives $(*)$, which implies that I is G-invariant and that $\delta_{B/I} = \delta$. To finish off, we have

$$W_{B/I} = \phi \otimes \mathrm{id}_{G/N}(W_B) = \big(\phi \circ (1 \otimes M)\big) \otimes \mathrm{id}_{G/N}(w_{G/N}) = j \otimes \mathrm{id}_{G/N}(w_{G/N}) = W,$$

which establishes the claim.

§ 3. The main theorem

Theorem 3.1. *Let N be a closed amenable subgroup of a locally compact group G, let (δ, W) be a nondegenerate twisted coaction of $(G, G/N)$ on a C^*-algebra A, and let α denote the dual action of N on the twisted crossed product $A \times_{\delta, W} G$. Then the twisted system $(A, G, G/N, \delta, W)$ is Morita equivalent to the untwisted system $\big((A \times_{\delta, W} G) \times_\alpha N, N, \hat{\alpha}\big)$, in the sense that (δ, W) is Morita equivalent to the inflated twisted coaction* $(\mathrm{Inf}\,\hat{\alpha}, 1 \otimes 1)$ *of $(G, G/N)$ on $(A \times_{\delta, W} G) \times_\alpha N$.*

We already know from [PR], Theorem 4.1, that A is Morita equivalent to

$$D := (A \times_{\delta, W} G) \times_\alpha N\,,$$

so we shall try to build a coaction of G on the imprimitivity bimodule ${}_D Y_A$ constructed in [PR]. Since ${}_D Y_A$ is a quotient of the $(A \times_\delta G) \times_{\hat{\delta}} N$-$A \times_{\delta|} G/N$ bimodule X in Mansfield's imprimitivity theorem, we shall construct a coaction on X, and then use Proposition 2.6 to pass to the quotient bimodule. Unfortunately, this is not as easy as it sounds, and not just because Mansfield's construction is complicated. He does not construct his imprimitivity bimodule explicitly: he builds a Hilbert module $X_{A\times G/N}$, and then proves that $(A\times_\delta G)\times_{\hat\delta} N$ is isomorphic to the algebra $\mathscr{K}(X)$ of compact operators on X, without directly writing down a formula for the $(A\times_\delta G)\times_{\hat\delta} N$-valued inner product. Thus we shall be forced to use the description of $M(X\otimes C_r^*(G))$ as $\mathscr{L}((A\times G/N)\otimes C_r^*(G), X\otimes C_r^*(G))$, and it will take quite a bit of work to check that our formulas make sense.

The coaction of G on $X_{A\times G/N}$ which we construct will be compatible with the decomposition twisted coaction of $(G, G/N)$ on $A\times_{\delta|} G/N$. To be precise:

Proposition 3.2. *Let δ be a nondegenerate coaction of G on a C^*-algebra A, let $\delta|$ be the restriction of δ to G/N, $B:=A\times_{\delta|} G/N$, and let (δ_B, W_B) be the decomposition twisted coaction of $(G, G/N)$ on B (see Example 1.4). Let β denote the restriction of the dual action $\hat\delta$ to N. Then (δ_B, W_B) is Morita equivalent to the dual coaction $\hat\beta$ of N on $C:=(A\times_\delta G)\times_\beta N$.*

As we foreshadowed above, we shall have to go into the technical details of Mansfield's construction. We shall always assume that A is represented faithfully and nondegenerately on $\mathscr{H}$, and that $A\times_\delta G$, $A\times_{\delta|} G/N$ are represented on $\mathscr{H}\otimes L^2(G)$ as

$$A\times_\delta G = \overline{\mathrm{sp}}\{\delta(a)(1\otimes M_f) : a\in A, f\in C_0(G)\},$$
$$A\times_{\delta|} G/N = \overline{\mathrm{sp}}\{\delta(a)(1\otimes M_f) : a\in A, f\in C_0(G/N)\}$$

(it is proved in [M], Proposition 7, that the latter is a faithful representation of $A\times_{\delta|} G/N$). Mansfield defined dense subalgebras $\mathscr{D}$, $\mathscr{D}_N$ of $A\times_\delta G$, $A\times_{\delta|} G/N$ as follows. Choose Haar measures on N, G and G/N such that

$$\int_G g(s)\,ds = \int_{G/N}\int_N g(sn)\,dn\,d(sN)$$

for all $g\in C_c(G)$, and let $\phi: C_c(G)\to C_c(G/N)$ be the surjective map satisfying

$$\phi(f)(sN) = \int_N f(sn)\,dn\,.$$

We write $A_c(G)$ for the compactly supported elements in the Fourier algebra $A(G)$ of G, and $C_E(G)$ for the functions in $C_c(G)$ with support in a fixed compact subset E of G. For fixed E and $u\in A_c(G)$, Mansfield says an operator $T\in B(\mathscr{H}\otimes L^2(G))$ is (u, E, N) if T is the norm-limit of a sequence in

$$\mathrm{sp}\{\delta(\delta_u(a))(1\otimes M(\phi(f))) : f\in C_E(G), a\in A\},$$

where $\delta_u: A\to A$ is the composition of δ with the extension to $\mathscr{M}(A\otimes C_r^*(G))$ of the slice map $S_u: A\otimes C_r^*(G)\to A$. If N is trivial, he says T is (u, E). Mansfield proved that the space $\mathscr{D}_N$ of operators in $B(\mathscr{H}\otimes L^2(G))$ which are (u, E, N) for some $u\in A_c(G)$ and $E\subseteq G$ compact is a dense $*$-subalgebra of $A\times_{\delta|} G/N$; if $N=\{e\}$, we write $\mathscr{D}:=\mathscr{D}_N$. We shall now define similar subspaces of $(A\times_{\delta|} G/N)\otimes S$, where S is an arbitrary C^*-algebra; in our applications, S will be $C_r^*(G)$.

Definition 3.3. Let S be a C^*-algebra represented as a nondegenerate subalgebra of $B(\mathscr{K})$ for some Hilbert space $\mathscr{K}$. Define $\psi: C_c(G, S)\to C_c(G/N, S)$ by

$$\psi(g)(sN) = \int_N g(sn)\,dn \quad \text{for } g \in C_c(G,S).$$

For $g \in C_b(G,S)$, we define an operator $M^S(g)$ on $L^2(G) \otimes \mathscr{K} \cong L^2(G,\mathscr{K})$ by

$$\big(M^S(g)\xi\big)(s) = g(s)\big(\xi(s)\big).$$

Then, for $u \in A_c(G)$ and $E \subseteqq G$ compact, an operator $T \in B\big(\mathscr{H} \otimes L^2(G) \otimes \mathscr{K}\big)$ is called $(u, E, N)^S$ if T is the norm-limit of a sequence in

$$\operatorname{sp}\big\{\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^S(\psi(g))\big) : g \in C_E(G,S),\, a \in A\big\}.$$

If N is trivial, we say T is $(u,E)^S$. The set of all elements in $B\big(\mathscr{H} \otimes L^2(G) \otimes \mathscr{K}\big)$ which are $(u,E,N)^S$ is denoted by $\mathscr{D}_N^S$, and $\mathscr{D}^S := \mathscr{D}_{\{e\}}^S$.

Mansfield did most of his calculations on the spaces $\mathscr{D}$ and $\mathscr{D}_N$, and we shall do most of ours on the spaces $\mathscr{D}^S$ and $\mathscr{D}_N^S$. The next lemma shows that the spaces $\mathscr{D}^S$, $\mathscr{D}_N^S$ are closely related to the algebraic tensor products $\mathscr{D} \odot S$, $\mathscr{D}_N \odot S$; this will allow us to use his results in our calculations.

Lemma 3.4. *Fix $u \in A_c(G)$ and a compact subset E of G, and suppose that $x \in \mathscr{D}_N^S$ is $(u,E,N)^S$. Suppose further that F is a compact subset of G such that E is contained in the interior of F. Then x is the norm-limit of a sequence of finite sums of the form*

$$\sum_{i=1}^{n} \delta(\delta_u(a_i))\big(1 \otimes M(\phi(f_i))\big) \otimes z_i,$$

in which $a_i \in A$, $f_i \in C_F(G)$ and $z_i \in S$. In particular, x is the norm-limit of a sequence of elements $\sum_i d_i \otimes z_i \in \mathscr{D}_N \odot S$ such that every d_i is (u,F,N).

Proof. It is enough to show that each element $\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^S(\psi(g))\big)$ for $g \in C_E(G,S)$ can be thus approximated. By a standard compactness argument, for each $\varepsilon > 0$ there exist $f_i \in C_F(G)$ and $z_i \in S$ such that $\|g(s) - \sum_i f_i(s) z_i\| < \varepsilon$ for all $s \in G$. Fix a constant $c_F > 0$ such that $\|\psi(g)\|_\infty \leqq c_F \|g\|_\infty$ for all $g \in C_F(G)$. Then

$$\begin{aligned}
\Big\|\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^S(\psi(g))\big) - \sum_i \delta(\delta_u(a))\big(1 \otimes M(\phi(f_i))\big) \otimes z_i\Big\| \\
&= \Big\|\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^S\big(\psi\big(g - \textstyle\sum f_i \otimes z_i\big)\big)\big)\Big\| \\
&\leqq \|a\|\,\big\|\psi\big(g - \textstyle\sum f_i \otimes z_i\big)\big\|_\infty \\
&\leqq \|a\|\, c_F \big\|g - \textstyle\sum f_i \otimes z_i\big\|_\infty \\
&\leqq c_F \|a\| \varepsilon,
\end{aligned}$$

as required. □

Proposition 3.5. *$\mathscr{D}^S$ is a dense $*$-subalgebra of $(A \times_\delta G) \otimes S$ containing $\mathscr{D} \odot S$, and $\mathscr{D}_N^S$ is a dense $*$-subalgebra of $(A \times_{\delta|} G/N) \otimes S$ containing $\mathscr{D}_N \odot S$. Further, if $x \in \mathscr{D}^S$ and $y \in \mathscr{D}_N^S$, then xy and yx are in $\mathscr{D}^S$, and $\mathscr{D}^S \mathscr{D}_N^S$ is dense in $\mathscr{D}^S$.*

Proof. The arguments used in the proofs of [M], Lemma 10 (i) and Lemma 11 (i), show that $\mathscr{D}_N^S$ is closed under addition. Since the elementary tensors $d \otimes z$ are trivially contained in $\mathscr{D}_N^S$, it follows that $\mathscr{D}_N \odot S \subset \mathscr{D}_N^S$, and since $\mathscr{D}_N$ is dense in $A \times_{\delta|} G/N$, we deduce that $\mathscr{D}_N^S$ is dense in $(A \times_{\delta|} G/N) \otimes S$. The other algebraic assertions follow from Lemma 3.4 and [M], Lemma 11, while the density of $\mathscr{D}^S \mathscr{D}_N^S$ in $\mathscr{D}^S$ follows from Lemma 3.4 and [M], Lemma 17. □

Recall from [M], Proposition 16, that there is a well-defined map $\Psi : \mathscr{D} \to \mathscr{D}_N$ such that: if x is the norm-limit of a sequence $x_j = \sum_{i=1}^{n_j} \delta(\delta_u(a_{ij}))(1 \otimes M(f_{ij}))$, and all f_{ij} have support in a fixed compact subset of G, then $\Psi(x)$ is the norm-limit of the sequence $y_j = \sum_{i=1}^{n_j} \delta(\delta_u(a_{ij}))(1 \otimes M(\phi(f_{ij})))$. We want to define a similar map $\Psi^S : \mathscr{D}^S \to \mathscr{D}_N^S$ such that: if $x \in \mathscr{D}^S$ is the norm-limit of

$$\sum_{i=1}^{n_j} (\delta(\delta_u(a_{ij})) \otimes 1)(1 \otimes M^S(g_{ij})),$$

with $u \in A_c(G)$ and $g_{ij} \in C_E(G, S)$, then $\Psi^S(x)$ will be the norm-limit of

$$\Psi^S(x_j) := \sum_{i=1}^{n_j} (\delta(\delta_u(a_{ij})) \otimes 1)(1 \otimes M^S(\psi(g_{ij}))).$$

Lemma 3.6. *The map $\Psi^S : \mathscr{D}^S \to \mathscr{D}_N^S$ as given above is well-defined. Further, for each $u \in A_c(G)$ and compact subset $E \subseteq G$, there is a constant $c_{u,E}$ such that*

$$\|\Psi^S(x)\| \leqq c_{u,E} \|x\|$$

for all $x \in \mathscr{D}^S$ which are $(u, E)^S$.

Proof. We can replace $L^2(G)$, $L^2(G/N)$ by $L^2(G, \mathscr{K})$, $L^2(G/N, \mathscr{K})$ in the proofs of [M], Lemma 14 and Lemma 15, and deduce that, for each pair (u, E), there is a constant $c_{u,E}$ such that

$$\left\| \sum_{i=1}^{n} (\delta(\delta_u(a_i)) \otimes 1)(1 \otimes M^S(\psi(g_i))) \right\| \leqq c_{u,E} \left\| \sum_{i=1}^{n} (\delta(\delta_u(a_i)) \otimes 1)(1 \otimes M^S(g_i)) \right\|,$$

provided all the g_i have support in E. (It helps to realise that the norm of an element $x \in A \times_{\delta|} G/N$ is the operator norm of its image in $B(\mathscr{H} \otimes L^2(G))$.) The arguments in the proof of [M], Proposition 16 show that Ψ^S is well-defined, and taking limits gives the desired inequality.

The dual action $\hat{\delta}$ of G on $A \times_\delta G$ is given by $s \mapsto \operatorname{Ad}(1 \otimes \varrho_s)$, where ϱ denotes the right regular representation of G. It was shown in [M], Lemma 11, that $\mathscr{D}$ is invariant under $\hat{\delta}|_N$. Hence, again using Lemma 3.4, we see that there is a well-defined action γ of N on $\mathscr{D}^S$ such that

$$\gamma_n = \hat{\delta}(n) \otimes \mathrm{id}_S = \mathrm{Ad}(1 \otimes \varrho_n \otimes 1).$$

Lemma 3.7. *For $x, y \in \mathscr{D}^S$, the maps $n \mapsto \gamma_n(x)y$ and $n \mapsto y\gamma_n(x)$ are continuous with compact support, and the action of $\Psi^S(x)$ as a multiplier on $(A \times_\delta G) \otimes S$ is given by the formulas*

$$\Psi^S(x)y = \int_N \gamma_n(x)y\,dn \quad \text{and} \quad y\Psi^S(x) = \int_N y\gamma_n(x)\,dn.$$

Proof. By Lemma 3.4 there exist $u, v \in A_c(G)$ and compact sets $E, F \subseteq G$, such that x is a limit of elements $x_j = \sum_{i=1}^{n_j} d_{ij} \otimes z_{ij}$, where all the $d_{ij} \in \mathscr{D}$ are (u, E), and y is the limit of elements $y_j = \sum_{i=1}^{m_j} c_{ij} \otimes w_{ij}$, where the c_{ij} are (v, F). The proof of [M], Lemma 18 shows that there is a compact set K, depending only on E and F, such that the maps

$$n \mapsto \hat{\delta}(n)(d_{ij})\,c_{kj}$$

have support in K. This shows that $n \mapsto \gamma_n(x_j)y_j$ has support in K, and hence also that $n \mapsto \gamma_n(x)y$ has support in K. The continuity of this map follows from the continuity of $\hat{\delta}$ on $A \times_\delta G$. It follows also from [M], Lemma 18, that $\Psi^S(x_j)y_j = \int_N \gamma_n(x_j)y_j\,dn$ for all j; taking limits on both sides, and repeating the argument for $y\Psi^S(x)$, gives the result. □

The $\mathscr{D}_N$-valued inner product on $\mathscr{D}$ was defined in [M] by $\langle d, e\rangle_{\mathscr{D}_N} = \Psi(d^*e)$, and we shall define a $\mathscr{D}_N^S$-valued inner product on $\mathscr{D}^S$ by

$$\langle x, y\rangle_{\mathscr{D}_N^S} = \Psi^S(x^*y).$$

In fact, since $\Psi^S(d \otimes z) = \Psi(d) \otimes z$ for all elementary tensors $d \otimes z \in \mathscr{D} \odot S$, this inner product coincides on $\mathscr{D} \odot S$ with the $\mathscr{D}_N \odot S$-valued inner product satisfying

$$\langle d \otimes z, c \otimes w\rangle = \langle d, c\rangle \otimes z^*w.$$

Let X be the completion of $\mathscr{D}$ with respect to the $\mathscr{D}_N$-valued inner product on $\mathscr{D}$, and recall that we have already denoted the completion $A \times_{\delta|} G/N$ of $\mathscr{D}_N$ by B. Then $X \otimes S$ denotes the completion of the algebraic tensor product $X \odot S$ with respect to the $(B \otimes S)$-valued inner product, which has $\mathscr{D} \odot S$ as a dense subspace.

Lemma 3.8. *$\mathscr{D}^S$ embeds naturally as a dense subspace of $X \otimes S$, so that the $\mathscr{D}_N^S$-valued inner product on $\mathscr{D}^S$ as defined above is the restriction of the $(B \otimes S)$-valued inner product on $X \otimes S$.*

Proof. Let $x, y \in \mathscr{D}^S$. Then by Lemma 3.4 we can write $x = \lim x_j$, $y = \lim y_j$, and assume there exist $u \in A_c(G)$ and $E \subseteq G$ compact such that $x_j = \sum_{i=1}^{n_j} d_{ij} \otimes z_{ij}$ and

$y_j = \sum_{i=1}^{m_j} c_{ij} \otimes w_{ij}$, where all d_{ij}, c_{ij} are (u, E). It follows then from [M], Lemma 11 (ii) and (iii) (by observing that the set D' given in (ii) and the w given in (iii) only depend in our case on u and E) that there exist $w \in A_c(G)$ and $F \subseteq G$ compact such that $x_j^* y_j$ is $(w, F)^S$ for all j, and the same is true for all $(x_i - x_j)^*(x_i - x_j)$ and $(y_i - y_j)^*(y_i - y_j)$. It follows from this and the inequality in Lemma 3.6 that (x_j) and (y_j) are Cauchy sequences in $X \otimes S$. Denote by $\tilde{x}$ and $\tilde{y}$ the limits of these sequences. Then, again using the inequality in Lemma 3.6 and the continuity of the inner product on $X \otimes S$, we see that $\langle \tilde{x}, \tilde{y} \rangle = \Psi^S(x^* y)$. Hence the map $x \mapsto \tilde{x}$ gives the desired embedding. □

Lemma 3.9. *Let $\pi : S \to B(\mathscr{H}_\pi)$ be a nondegenerate $*$-representation of S and let $\pi(S) = D$. Then the homomorphism* $\mathrm{id}_X \otimes \pi : X \otimes S \to X \otimes D$ *satisfies*

$$\mathrm{id}_X \otimes \pi\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^S(f)\big) = \big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^D(\pi \circ f)\big) \in \mathscr{D}^D,$$

for any $a \in A$, $f \in C_c(G, S)$.

Proof. The equation is clear when $f = g \otimes z \in C_c(G) \odot S$, since in that case $\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^S(f)\big) = \big(\delta(\delta_u(a))(1 \otimes M_g)\big) \otimes z$. Now let $f \in C_c(G, S)$ be arbitrary. By Lemma 3.4 we can approximate $y := \big(\delta(\delta_u(a)) \otimes\big)\big(1 \otimes M^S(f)\big)$ by finite sums of elements of the form $\big(\delta(\delta_u(a))(1 \otimes M_g)\big) \otimes z$ such that all the g have support in a fixed compact set $E \subseteq G$. Since there exist $v \in A_c(G)$ and $F \subseteq G$ compact such that $x^* x$ is $(v, F)^S$ whenever x is $(u, E)^S$, the inequality in Lemma 3.6 implies that y can be approximated by similar finite sums in the $X \otimes S$-norm on $\mathscr{D}^S$. The same arguments show that the images of the sums under $\mathrm{id}_X \otimes \pi$ converge to $\big(\delta(\delta_u(a)) \otimes 1\big)\big(1 \otimes M^D(\pi \circ f)\big)$, and the lemma follows from the continuity of $\mathrm{id}_X \otimes \pi$. □

We shall now return to the setting of Proposition 3.2. Thus S will now be $C_r^*(G)$, acting faithfully on $L^2(G)$ via the left regular representation λ^G of G. The spaces $\mathscr{D}^{C_r^*(G)}$ and $\mathscr{D}_N^{C_r^*(G)}$ will be denoted $\mathscr{D}^G$ and $\mathscr{D}_N^G$, and we write M^G for $M^{C_r^*(G)} = M \otimes \mathrm{id}_G$.

Recall that the decomposition coaction of G on $B = A \times_{\delta|} G/N$ is given by $\delta_B := \pi \times \mu$, where $\pi = (\delta \otimes \mathrm{id}_G) \circ \delta$ and $\mu = (1 \otimes M|_{C_0(G/N)} \otimes 1)$ (since $A \times_{\delta|} G/N$ is represented faithfully on $\mathscr{H} \otimes L^2(G)$ via the covariant representation $(\delta, 1 \otimes M|_{C_0(G/N)})$ we have $j_A = \delta$ and $j_{C(G/N)} = 1 \otimes M|_{C_0(G/N)}$). We shall define a coaction δ_X of G on X which is compatible with δ_B (see Definition 2.1). For this we recall that $W_G = M^G(w_G)$ is given by

$$(W_G \xi)(s, t) = \xi(s, s^{-1} t) \quad \text{for } \xi \in L^2(G \times G), s, t \in G,$$

and define $W \in B(\mathscr{H} \otimes L^2(G) \otimes L^2(G))$ by $W = 1 \otimes W_G$. We want to define δ_X on $\mathscr{D} \subset X$ by $\delta_X(d) = (d \otimes 1) W^*$, but first we have to prove that this defines an element of

$$\mathscr{L}_{B \otimes C_r^*(G)}\big(B \otimes C_r^*(G), X \otimes C_r^*(G)\big),$$

and for this we need some lemmas.

Lemma 3.10. *Let $d \in \mathscr{D}$, $x \in \mathscr{D}^G$ and $b \in \mathscr{D}_N^G$. Then xW, Wx and $(d \otimes 1) W^* b$ are all elements of $\mathscr{D}^G$, where the products are taken in $B(\mathscr{H} \otimes L^2(G) \otimes L^2(G))$.*

Proof. Since $W_G^* = M^G(w_G^*)$, we have

$$\big(\delta(\delta_u(a))\otimes 1\big)\big(1\otimes M^G(g)\big)W^* = \big(\delta(\delta_u(a))\otimes 1\big)\big(1\otimes M^G(gw_G^*)\big),$$

which is $(u, E)^G$ if and only if $\operatorname{supp} g \subseteq E$. Because $\mathscr{D}^G$ is invariant under taking adjoints, this shows that both xW^* and Wx are in $\mathscr{D}^G$.

To prove that $(d\otimes 1)W^*b \in \mathscr{D}^G$ for all $d\in\mathscr{D}$ and $b\in\mathscr{D}_N^G$, we note first that $\mathscr{D}_N\otimes 1$ lies naturally in the (algebraic) multiplier algebra of $\mathscr{D}^G$. For if $y\in\mathscr{D}^G$ is the limit of elements of the form $\sum_{i=1}^{n} d_i\otimes z_i$ where the d_i are all (u, E), then, for fixed $c\in\mathscr{D}_N$, there exist $w\in A_c(G)$ and $F\subseteq G$ compact such that $d_i c$ is (w, F) for all i. Then $y(c\otimes 1)$ is the limit of the $\sum_{i=1}^{n} d_i c\otimes z_i$, and hence is (w, F), and belongs to $\mathscr{D}^G$; similarly, $(c\otimes 1)y\in\mathscr{D}^G$. Suppose now that $b = \sum_j c_j\otimes z_j$ for some $c_j\in\mathscr{D}_N$ and $z_j\in C_r^*(G)$. Then, if d has the form $\sum_i \delta(\delta_u(a_i))\big(1\otimes M(f_i)\big)$, we can compute

$$\begin{aligned}(d\otimes 1)W^*b &= \sum_{i,j}\Big(\big(\delta(\delta_u(a_i))\big(1\otimes M(f_i)\big)\big)\otimes 1\Big)W^*(c_j\otimes z_j)\\ &= \sum_{i,j}\big(\delta(\delta_u(a_i))\otimes 1\big)\big(1\otimes M^G(f_i w_G^*)\big)(1\otimes z_j)(c_j\otimes 1)\\ &= \sum_{i,j}\big(\delta(\delta_u(a_i))\otimes 1\big)\big(1\otimes M^G(f_{ij})\big)(c_j\otimes 1),\end{aligned}$$

which belongs to $\mathscr{D}^G$ because $f_{ij} = f_i w_G^*(1\otimes z_j)$ is in $C_c\big(G, C_r^*(G)\big)$ with the same support as f_i. Taking limits, and doing a little bit of bookkeeping concerning the supports, gives the desired result. □

Lemma 3.11. *For $n\in N$, let $\gamma_n = \operatorname{Ad}(1\otimes\varrho_n\otimes 1)$ acting on $B\big(\mathscr{H}\otimes L^2(G)\otimes L^2(G)\big)$. Then $\gamma_n(W) = W(1\otimes 1\otimes\lambda_n^G)$.*

Proof. For $\xi\in L^2(G\times G)$, we have

$$\big((\varrho_n\otimes 1)W_G(\varrho_n^*\otimes 1)\xi\big)(s,t) = \big(W_G(\varrho_n^*\otimes 1)\xi\big)(sn,t) = \big((\varrho_n^*\otimes 1)\xi\big)(sn, n^{-1}s^{-1}t)$$

$$= \xi(s, n^{-1}s^{-1}t) = \big(W_G(1\otimes\lambda_n^G)\xi\big)(s,t).$$

Since $W = 1\otimes W_G$, this gives the result. □

Proposition 3.12. *For each $d\in\mathscr{D}$ we define $\delta_X(d):\mathscr{D}_N^G\to\mathscr{D}^G$ by*

$$\delta_X(d)(b) = (d\otimes 1)W^*b.$$

Then $\delta_X(d)$ extends to a bounded operator $\delta_X(d): B\otimes C_r^*(G)\to X\otimes C_r^*(G)$, with adjoint given by $\delta_X(d)^*(y) = \Psi^G\big(W(d^*\otimes 1)y\big)$ for $y\in\mathscr{D}^G$. The linear map $d\mapsto\delta_X(d)$ extends to a bounded linear map $\delta_X: X\to\mathscr{L}_{B\otimes C_r^*(G)}\big(B\otimes C_r^*(G), X\otimes C_r^*(G)\big)$ satisfying

$$(1 \otimes z) \cdot \delta_X(x), \quad \delta_X(x) \cdot (1 \otimes z) \in X \otimes C_r^*(G) \quad \textit{for } x \in X,\ z \in C_r^*(G).$$

Proof. It follows from Lemma 3.10 that $\delta_X(d)$ is a map from $\mathscr{D}_N^G$ into $\mathscr{D}^G$. To see that it can be extended to all of X, we start by fixing $y \in \mathscr{D}^G$, $b \in \mathscr{D}_N^G$ and computing

$$\begin{aligned} \|y\langle \delta_X(d)(b), \delta_X(d)(b)\rangle_{\mathscr{D}_N^G}\| &= \|y\Psi^G\big(b^* W(d^*d \otimes 1) W^* b\big)\| \\ &= \|\int_N y\gamma_n\big(b^* W(d^*d \otimes 1) W^* b\big)\, dn\| \\ &= \|\int_N yb^*\gamma_n\big(W(d^*d \otimes 1) W^*\big) b\, dn\|, \end{aligned}$$

because $\gamma_n(c) = c$ for all $c \in \mathscr{D}_N^G$. (This last observation holds because $\hat{\delta}(n)$ fixes $\mathscr{D}_N$.) From Lemma 3.11 we know that $\gamma_n(W^*) = (1 \otimes 1 \otimes \lambda_{n^{-1}}^G) W^*$, and thus

$$\begin{aligned} \|y\langle \delta_X(d)(b), \delta_X(d)(b)\rangle_{\mathscr{D}_N^G}\| &= \|\int_N yb^* W\gamma_n\big((d^*d \otimes 1)\big) W^* b\, dn\| \\ &= \|\big(\int_N yb^* W\big(\hat{\delta}(n)(d^*d) \otimes 1\big) dn\big) W^* b\| \\ &= \|yb^* W\big(\Psi(d^*d) \otimes 1\big) W^* b\|, \end{aligned}$$

where the last equation follows from [M], Lemma 18, since we can approximate $yb^* W \in \mathscr{D}^G$ in norm by appropriate elements in $\mathscr{D} \odot C_r^*(G)$. Thus

$$\begin{aligned} \|y\langle \delta_X(d)(b), \delta_X(d)(b)\rangle_{\mathscr{D}_N^G}\| &\leqq \|y\|\,\|b^* W(\langle d, d\rangle_{\mathscr{D}_N} \otimes 1) W^* b\| \\ &= \|y\|\,\|b\|^2\,\|\langle d, d\rangle_{\mathscr{D}_N}\|. \end{aligned}$$

By letting y run through an approximate identity in $\mathscr{D} \odot C_r^*(G) \subseteqq \mathscr{D}^G$ (see [M], Lemma 17), we conclude that

$$\|\delta_X(d)(b)\|^2 = \|\langle \delta_X(d)(b), \delta_X(d)(b)\rangle_{\mathscr{D}_N^G}\| \leqq \|\langle d, d\rangle_{\mathscr{D}_N}\|\,\|b\|^2 = \|d\|^2\,\|b\|^2.$$

Thus we can extend $\delta_X(d)$ to an operator from $B \otimes C_r^*(G)$ to $X \otimes C_r^*(G)$ with

$$\|\delta_X(d)\| \leqq \|d\|_X.$$

Easy calculations show that $\delta_X(d)^*$, defined on $\mathscr{D}^G$ as in the proposition, is an algebraic adjoint on the dense subset $\mathscr{D}^G$; it follows from the Cauchy-Schwartz inequality [Rieff 1], Proposition 2.9, that it extends to an adjoint of $\delta_X(d)$ on $X \otimes C_r^*(G)$ with

$$\|\delta_X(d)^*\| \leqq \|\delta_X(d)\|.$$

Finally, if $d \in \mathscr{D}$ and $z \in C_r^*(G)$, then $(1 \otimes z)\,\delta_X(d) b = (d \otimes z) W^* b$; since

$$(d \otimes z) W^* \in \mathscr{D}^G \subseteqq X \otimes C_r^*(G),$$

this implies that $(1 \otimes z)\,\delta_X(d) \in X \otimes C_r^*(G)$ for $d \in \mathscr{D}$. A computation like that in the previous paragraph shows that

$$\|(1 \otimes z)\delta_X(d)\|_{X \otimes C_r^*(C)} \leqq \|d\|_X \|z\|,$$

and hence it follows by continuity that $(1 \otimes z)\delta_X(x) \in X \otimes C_r^*(G)$ for all $x \in X$. Almost the same arguments show that $\delta_X(x)(1 \otimes z) \in X \otimes C_r^*(G)$ for all $x \in X$ and $z \in C_r^*(G)$. □

Lemma 3.13. *$W(b \otimes 1)W^* = \delta_B(b)$ for all $b \in B$.*

Proof. Since $(\delta, 1 \otimes M)$ is a covariant representation of (A, G, δ) and

$$W = (1 \otimes M \otimes \mathrm{id}_G)(w_G),$$

we have

$$W(\delta(a) \otimes 1)W^* = (\delta \otimes \mathrm{id}_G) \circ \delta(a) = \pi(a)$$

for all $a \in A$. The operator W commutes with $\mu(f) = 1 \otimes M(f) \otimes 1$ for all $f \in C_b(G)$, and especially for $f \in C_0(G/N)$. Thus for $a \in A$ and $f \in C_0(G/N)$ we have

$$\begin{aligned} W\big((\delta(a)(1 \otimes M_f)) \otimes 1\big)W^* &= W(\delta(a) \otimes 1)(1 \otimes M_f \otimes 1)W^* \\ &= \pi(a)\mu(f) = \delta_B\big(\delta(a)(1 \otimes M_f)\big). \end{aligned}$$

This equality extends to all of B by continuity. □

We now come to the key technical result.

Proposition 3.14. *$\delta_X : X \to \mathscr{L}_{B \otimes C_r^*(G)}\big(B \otimes C_r^*(G), X \otimes C_r^*(G)\big)$ is a coaction of G on X which is compatible with δ_B.*

Proof. By Proposition 3.12 it remains to show that δ_X satisfies Conditions (1), (3) and (4) of Definition 2.1. For $x \in \mathscr{D}$ and $b \in \mathscr{D}_N$, Lemma 3.13 gives

$$\delta_X(xb) = (xb \otimes 1)W^* = (x \otimes 1)(b \otimes 1)W^* = (x \otimes 1)W^*\delta_B(b) = \delta_X(x)\delta_B(b),$$

and this extends by continuity to $x \in X$, $b \in B$. Next let $x, y \in \mathscr{D}$. Then for $z \in \mathscr{D}^G$ and $b \in \mathscr{D}_N^D$, we have

$$\begin{aligned} z\delta_B(\langle x, y\rangle_B)b &= zW(\langle x, y\rangle_B \otimes 1)W^*b \\ &= \int_N zW\big(\hat{\delta}(n)(x^*y) \otimes 1\big)W^*b\,dn = \int_N z\gamma_n\big(W(x^*y \otimes 1)W^*b\big)\,dn \\ &= z\Psi^G\big(W(x^*y \otimes 1)W^*b\big) = z\delta_X(x)^*\big((y \otimes 1)W^*b\big) \\ &= z\big(\delta_X(x)^*\delta_X(y)(b)\big), \end{aligned}$$

which implies that $\delta_B(\langle x, y\rangle_B) = \delta_X(x)^*\delta_X(y)$ for all $x, y \in X$, and completes the verification of condition (1).

Next we show (3). Using Lemma 3.6 to reduce to norm convergence, and a partition of unity argument to approximate $f \in C_c(G, C_r^*(G))$ by a sum of functions of the form $g w_G^*(1 \otimes z)$ for $g \in C_c(G)$ and $z \in C_r^*(G)$, we can prove that $(\mathscr{D} \otimes 1) W^*(1 \otimes C_r^*(G))$ is dense in $X \otimes C_r^*(G)$. Since $\delta_X(X)(B \otimes C_r^*(G))$ contains $(\mathscr{D} \otimes 1) W^*(1 \otimes C_r^*(G))(\mathscr{D}_N \otimes 1)$, (3) now follows from the density of $X\mathscr{D}_N$ in X [M], Theorem 19.

Finally, we come to the hard bit: we have to check the coaction identity

$$(\delta_X \otimes \mathrm{id}_G) \circ \delta_X = (\mathrm{id}_X \otimes \delta_G) \circ \delta_X . \tag{3.1}$$

Recall that for $y \in X$ the expression $(\delta_X \otimes \mathrm{id}_G)(\delta_X(y))$ is an adjointable map in

$$\mathscr{L}(B \otimes C_r^*(G) \otimes C_r^*(G), X \otimes C_r^*(G) \otimes C_r^*(G)),$$

which is given on an element of the form $(\delta_B \otimes \mathrm{id}_G)(c)d$, with $c \in B \otimes C_r^*(G)$ and

$$d \in B \otimes C_r^*(G) \otimes C_r^*(G),$$

by the equation

$$(\delta_X \otimes \mathrm{id}_G)(\delta_X(y))\big((\delta_B \otimes \mathrm{id}_G)(c)d\big) = (\delta_X \otimes \mathrm{id}_G)(\delta_X(y)(c))(d). \tag{3.2}$$

On the other hand, $(\mathrm{id}_X \otimes \delta_G)(\delta_X(y)) \in \mathscr{L}(B \otimes C_r^*(G) \otimes C_r^*(G), X \otimes C_r^*(G) \otimes C_r^*(G))$ is given on elements of the form $(\mathrm{id}_B \otimes \delta_G)(c)d$ by

$$(\mathrm{id}_X \otimes \delta_G)(\delta_X(y))\big((\mathrm{id}_B \otimes \delta_G)(c)d\big) = (\mathrm{id}_X \otimes \delta_G)(\delta_X(y)(c))(d). \tag{3.3}$$

The difficulty in proving (3.1) is that the two sides are naturally defined, in (3.2) and (3.3), on elements of a different form. We shall start with (3.1) for $c = \delta_B(b)(1 \otimes z)$ (recall that δ_B is nondegenerate), and then approximate $(1 \otimes 1 \otimes z)d$ by a linear combination $\sum_i (1 \otimes \delta_G(z_i))\, d_i$, to give us an argument of the form required in (3.3).

To simplify the formulas in the calculation, which concerns operators on

$$(\mathscr{H} \otimes L^2(G)) \otimes L^2(G) \otimes L^2(G),$$

we shall adopt subscript notation: for example, T_{23} means that $T \in B(L^2(G) \otimes L^2(G))$ acts on the second and third spaces, so that $T_{23} = 1 \otimes T \otimes 1$, and

$$T_{24} = (1 \otimes 1 \otimes \Sigma) T_{23} (1 \otimes 1 \otimes \Sigma),$$

where $\Sigma(h \otimes k) := k \otimes h$. With this convention, the operator W_{123} is by definition $(W_G)_{23}$. Since W_G satisfies the pentagonal identity

$$(W_G)_{23}(W_G)_{24} = \mathrm{id} \otimes \delta_G(W_G)_{234} = (W_G)_{34}(W_G)_{23}(W_G^*)_{34},$$

we have $W_{123}W_{124} = (W_G)_{34}W_{123}(W_G^*)_{34}$. Recall also from Lemma 3.13 that W implements δ_B.

For the calculation itself, we start with $d \in \mathscr{D} \odot C_r^*(G) \odot C_r^*(G)$ and $c = \delta_B(b)(1 \otimes z)$ for $b \in \mathscr{D}_N$, $z \in C_r^*(G)$, and choose $z_i \in C_r^*(G)$, $d_i \in \mathscr{D} \odot C_r^*(G) \odot C_r^*(G)$ such that

$$z_4 d := (1 \otimes 1 \otimes 1 \otimes z)\, d \sim \sum_i \delta_G(z_i)_{34} d_i$$

(in the operator norm; equivalently, $(1 \otimes 1 \otimes z)\, d \sim \sum (1 \otimes \delta_G(z_i))\, d_i$ in

$$B \otimes C_r^*(G) \otimes C_r^r(G))\,.$$

Then

$$\begin{aligned}
&(\delta_X \otimes \mathrm{id}_G)(\delta_X(y))\big((\delta_B \otimes \mathrm{id}_G)(c)d\big)\\
&\quad = (\delta_X \otimes \mathrm{id}_G)\big(\delta_X(y)(c)\big)(d)\\
&\quad = (\delta_X \otimes \mathrm{id}_G)\big((y \otimes 1)W^*\delta_B(b)(1 \otimes z)\big)(d)\\
&\quad = (\delta_X \otimes \mathrm{id}_G)\big((y \otimes 1)(b \otimes 1)W^*(1 \otimes z)\big)(d)\\
&\quad = (yb)_{12}W_{124}^* z_4 W_{123}^* d\\
&\quad = (yb)_{12}W_{124}^* W_{123}^* z_4 d\\
&\quad = (yb)_{12}(W_G)_{34}W_{123}^*(W_G^*)_{34} z_4 d\\
&\quad \sim (yb)_{12}(W_G)_{34}W_{123}^*(W_G^*)_{34}\sum_i \delta_G(z_i)_{34} d_i\\
&\quad = \sum_i (W_G)_{34}(yb)_{12}W_{123}^*(z_i)_3(W_G^*)_{34} d_i\\
&\quad = \sum_i (W_G)_{34}y_{12}W_{123}^*(W_{123}b_{12}W_{123}^*)(z_i)_3(W_G^*)_{34} d_i\\
&\quad = \sum_i (W_G)_{34}\big(\delta_X(y)(\delta_B(b)(1 \otimes z_i))\big)_{123}(W_G^*)_{34} d_i\\
&\quad = \sum_i \mathrm{id}_X \otimes \delta_G\big(\delta_X(y)(\delta_B(b)(1 \otimes z_i))\big)(d_i)\\
&\quad = \sum_i \mathrm{id}_X \otimes \delta_G\big(\delta_X(y)\big)\big(\mathrm{id}_B \otimes \delta_G(\delta_B(b)(1 \otimes z_i))\, d_i\big) \quad \text{using (3.3)}\\
&\quad = \mathrm{id}_X \otimes \delta_G\big(\delta_X(y)\big)\big(\mathrm{id}_B \otimes \delta_G(\delta_B(b))\big(\sum_i (1 \otimes \delta_G(z_i))\, d_i\big)\big)\\
&\quad \sim \mathrm{id}_X \otimes \delta_G\big(\delta_X(y)\big)\big(\delta_B \otimes \mathrm{id}_G(\delta_B(b))(1 \otimes 1 \otimes z)d\big)\\
&\quad = \mathrm{id}_X \otimes \delta_G\big(\delta_X(y)\big)\big(\delta_B \otimes \mathrm{id}_G(\delta_B(b)(1 \otimes z))d\big)\\
&\quad = \mathrm{id}_X \otimes \delta_G\big(\delta_X(y)\big)\big(\delta_B \otimes \mathrm{id}_G(c)d\big)\,.
\end{aligned}$$

We have now proved that both sides of (3.1) agree on elements y of the dense subspace $\mathscr{D}$ of X; since δ_X and δ_G are isometric, it follows that (3.1) holds on all of X. □

Proof of Proposition 3.2. After Proposition 3.14, it only remains to identify $\operatorname{Inf}\hat{\beta}$ with the coaction $\delta_{\mathscr{K}(X)}$ on $C = \mathscr{K}(X_B)$ induced by the δ_B-compatible coaction δ_X, and check the equation for the twists. For the first, it is enough to check that they agree on the image $i_{A\times_\delta G}(\mathscr{D})$ of the dense subalgebra $\mathscr{D}$ of A, and on $i_N(N)$. Recall that $\hat{\beta}$ is by definition the integrated form of the covariant representation $(i_{A\times_\delta G}\otimes 1, i_N \otimes \lambda^N)$ of the system $(A\times_\delta G, N, \beta := \hat{\delta}|_N)$. Since $\lambda^G|_N(\lambda^N_n) = \lambda^G_n$ for $n \in N$, the coaction $\operatorname{Inf}\hat{\beta}$ of G on C is just the integrated form of the covariant representation $(i_{A\times_\delta G}\otimes 1, i_N\otimes \lambda^G|_N)$. From the description of the action of $C_c(N, \mathscr{D})$ on X in [M], Proposition 22, we see that the action of $i_N(n)$ on $d \in \mathscr{D} \subseteq X$ is given by $i_N(n)\,d = \Delta(n)^{1/2}\beta_n(d)$, where Δ is the modular function of N. Hence, if $d\otimes z \in \mathscr{D} \odot C^*_r(G)$,

$$\operatorname{Inf}\hat{\beta}(i_N(n))(d\otimes z) = \Delta(n)^{1/2}\beta_n(d)\otimes \lambda^G_n z = \Delta(n)^{1/2}(1\otimes 1\otimes \lambda^G_n)\gamma_n(d\otimes z),$$

and $\operatorname{Inf}\hat{\beta}(i_{A\times_\delta G}(x))(d\otimes z) = xd\otimes z$. On the other hand, for any $c\in C$ the action of $\delta_{\mathscr{K}(X)}(c)$ on $y \in X\otimes C^*_r(G)$ is given by $\delta_{\mathscr{K}(X)}(c)(y) = V(c\otimes 1)V^*y$, where V is the unitary in $\mathscr{L}_{B\otimes C^*_r(G)}(X\otimes_{\delta_B}(B\otimes C^*_r(G)), X\otimes C^*_r(G))$ given on elementary tensors $y\otimes b \in \mathscr{D}\odot \mathscr{D}^G_N$ by

$$V(y\otimes b) = \delta_X(y)(b) = (y\otimes 1)W^*b.$$

Because $\mathscr{D}\odot\mathscr{D}^G_N$ has dense image in $X\otimes_{\delta_B}(B\otimes C^*_r(G))$, it is enough to show that

$$\operatorname{Inf}\hat{\beta}(i_N(n))V(y\otimes b) = V(\Delta(n)^{1/2}\beta_n(y)\otimes b),$$

and

$$\operatorname{Inf}\hat{\beta}(i_{A\times_\delta G}(x))V(y\otimes b) = V(xy\otimes b)$$

for all $x, y\in\mathscr{D}$, $n\in N$ and $b\in\mathscr{D}^G_N$. But

$$\begin{aligned}\operatorname{Inf}\hat{\beta}(i_N(n))V(y\otimes b) &= \operatorname{Inf}\hat{\beta}(i_N(n))(y\otimes 1)W^*b\\ &= \Delta(n)^{1/2}(1\otimes 1\otimes\lambda^G_n)\gamma_n((y\otimes 1)W^*b)\\ &= \Delta(n)^{1/2}\gamma_n(y\otimes 1)(1\otimes 1\otimes \lambda^G_n)\gamma_n(W^*)b\\ &= \Delta(n)^{1/2}(\beta_n(y)\otimes 1)W^*b,\end{aligned}$$

where we have used that $1\otimes 1\otimes\lambda^G_n$ and b are fixed points for γ, and Lemma 3.11. For $x\in\mathscr{D}$ we have

$$\begin{aligned}\operatorname{Inf}\hat{\beta}(i_{A\times_\delta G}(x))V(y\otimes b) &= \operatorname{Inf}\hat{\beta}(i_{A\times_\delta G}(x))(y\otimes 1)W^*b\\ &= (xy\otimes 1)W^*b = V(xy\otimes b),\end{aligned}$$

and we have proved that $\operatorname{Inf}\hat{\beta} = \delta_{\mathscr{K}(X)}$.

Finally, we have to show that $(\mathrm{id}_X\otimes q)(\delta_X(x)) = (1\otimes 1)(x\otimes 1)W_B^*$, which amounts to showing that

$$(\mathrm{id}_X\otimes q)(\delta_X(x)(b)) = (x\otimes 1)W_B^*(\mathrm{id}_B\otimes q)(b)$$

for $x \in X$ and $b \in B \otimes C_r^*(G)$. It is enough to show this for $x = \delta(\delta_u(a))(1 \otimes M_g)$, where $u \in A_c(G)$, $a \in A$ and $g \in C_c(G)$, and $b = d \otimes z \in \mathscr{D} \odot C_r^*(G)$. Since $W^* = (1 \otimes M^G)(w_G^*)$, Lemma 3.10 gives

$$\begin{aligned}
&(\mathrm{id}_X \otimes q)(\delta_X(x)(b)) \\
&\quad = (\mathrm{id}_X \otimes q)\big((\delta(\delta_u(a))(1 \otimes M_g) \otimes 1)\,W^*(d \otimes z)\big) \\
&\quad = (\mathrm{id}_X \otimes q)\big((\delta(\delta_u(a)) \otimes 1)(1 \otimes M^G(g w_G^*))(d \otimes z)\big) \\
&\quad = \big(\delta(\delta_u(a)) \otimes 1\big)\big((1 \otimes 1 \otimes q)((1 \otimes M \otimes \mathrm{id}_G)(g w_G^*))(d \otimes q(z))\big) \\
&\quad = \big(\delta(\delta_u(a))(1 \otimes M_g) \otimes 1\big)\big((1 \otimes M \otimes \mathrm{id}_{G/N})(w_{G/N}^*)\big)(d \otimes q(z)) \\
&\quad = (x \otimes 1)\,W_B^*\big((\mathrm{id}_B \otimes q)(d \otimes z)\big),
\end{aligned}$$

which finishes the proof. □

With Proposition 3.2 available, it is easy to prove Theorem 3.1.

Proof of Theorem 3.1. As above, we write $B = A \times_{\delta|} G/N$, β for the dual action of N on $A \times_\delta G$, $C = (A \times_\delta G) \times_\beta N$, $D = (A \times_{\delta, W} G) \times_\alpha N$, and (δ_B, W_B) for the decomposition coaction of G on B. Thus Proposition 3.2 gives us a Morita equivalence (X, δ_X) between $(C, \mathrm{Inf}\,\hat{\beta}, 1 \otimes 1)$ and (B, δ_B, W_B). If I_W is the twisting ideal for W, then $A \times_{\delta|, W} G/N$ is by definition B/I_W. If J_W is the corresponding ideal in $A \times_\delta G$, then it is shown in [PR], p. 337–8, that J_W is a β-invariant ideal; that the ideal $J_W \times_\beta N$ in C coincides with the one induced from I_W via the bimodule X; and hence that $C/(J_W \times_\beta N)$ is Morita equivalent to $A \times_{\delta|, W} G/N$. Since the quotient map of $A \times_\delta G$ onto $A \times_\delta G/J_W \cong A \times_{\delta, W} G$ intertwines β and α, it induces an isomorphism of $C/(J_W \times_\beta N)$ onto D. On the other hand, the quotient $B/I_W = A \times_{\delta|, W} G/N$ is naturally isomorphic to A.

We have now constructed a Morita equivalence of D and A. We showed in Example 2.9 that the ideal $I := I_W$ is G-invariant for the coaction (δ_B, W_B), and that the isomorphism of B/I with A takes the quotient coaction $(\delta_{B/I}, W_{B/I})$ into (δ, W). Thus it follows from Proposition 2.6 that the corresponding ideal $J_W \times_\beta N$ in C is G-invariant, and that $(X/(X \cdot I), \delta_{X/(X \cdot I)})$ implements a Morita equivalence between

$$\big(C/J_W \times_\beta N, (\mathrm{Inf}\,\hat{\beta})_{C/J_W \times_\beta N}, 1 \otimes 1\big) \quad \text{and} \quad (B/I, \delta_{B/I}, W_B) \cong (A, \delta, W).$$

To finish the proof, we recall from Example 2.8 that, for any β-invariant ideal J in any C^*-algebra E,

$$\big((E \times_\beta N)/(J \times_\beta N), N, \hat{\beta}\big) \cong \big((E/J) \times_\alpha N, N, \hat{\alpha}\big),$$

where α is the action of N on E/J induced by β. But the dual action α of N on

$$(A \times_\delta G)/J_W = A \times_{\delta, W} G$$

is by definition that induced by the dual action β of $N \subset G$ on $A \times_\delta G$. Thus we obtain the desired Morita equivalence of $(D, \mathrm{Inf}\,\hat{\alpha}, 1 \otimes 1)$ and (A, δ, W). □

§ 4. Applications

Morita equivalence and duality. Recall that strongly continuous actions $\alpha : G \to \operatorname{Aut} A$ and $\beta : G \to \operatorname{Aut} B$ are *Morita equivalent* if there is an A-B imprimitivity bimodule X and a strongly continuous action u of G on X such that:

$$u_s(x \cdot \langle y, z \rangle_B) = u_s(x) \cdot \langle u_s(y), u_s(z) \rangle_B,$$

$$\alpha_s({}_A\langle x, y \rangle) = {}_A\langle u_s(x), u_s(y) \rangle, \quad \text{and}$$

$$\beta_s(\langle x, y \rangle_B) = \langle u_s(x), u_s(y) \rangle_B .$$

If (X, u) implements such a Morita equivalence, then the crossed products $A \times_\alpha G$ and $B \times_\beta G$ are Morita equivalent via an imprimitivity bimodule $X \times_u G$ which is the completion of a $C_c(G, A)$-$C_c(G, B)$ pre-imprimitivity bimodule obtained by putting appropriate actions and inner products on $C_c(G, X)$ [Com].

Theorem 4.1. *Suppose that (δ_A, W_A) and (δ_B, W_B) are nondegenerate Morita equivalent twisted coactions of $(G, G/N)$ on A and B, respectively, and let α and β be the dual actions of N on $A \times_{\delta_A, W_A} G$ and $B \times_{\delta_B, W_B} G$, respectively. Then (δ_A, W_A) is Morita equivalent to (δ_B, W_B) if and only if α is Morita equivalent to β.*

The theorem will follow from Theorem 3.1 and the following lemma, which at least in part is already known (e.g., see [BS], § 6). Recall from [ER], Corollary 2.3, that every A-B imprimitivity bimodule X can be represented faithfully on a pair of Hilbert spaces $(\mathscr{H}, \mathscr{K})$ so that A and B are closed nondegenerate subalgebras of $B(\mathscr{H})$ and $B(\mathscr{K})$, X is a linear subspace of $B(\mathscr{K}, \mathscr{H})$, and the actions and inner products are given by composition of operators: e.g., ${}_A\langle x, y \rangle := x \circ y^*$.

Lemma 4.2. *Suppose that (δ_A, W_A) and (δ_B, W_B) are Morita equivalent twisted coactions of $(G, G/N)$ on A and B, respectively. Then the dual actions of N on $A \times_{\delta_A, W_A} G$ and $B \times_{\delta_B, W_B} G$ are also Morita equivalent.*

Dually, if α and β are Morita equivalent actions of G on the C^*-algebras A and B, then the dual coactions $\hat{\alpha}$ and $\hat{\beta}$ on the reduced crossed products $A \times_{\alpha, r} G$ and $B \times_{\beta, r} G$ are Morita equivalent.

Proof. If we represent ${}_AX_B$ faithfully on $(\mathscr{H}, \mathscr{K})$, and represent $A \times_{\delta_A} G$ and $B \times_{\delta_B} G$ on $\mathscr{H} \otimes L^2(G)$ and $\mathscr{K} \otimes L^2(G)$ in the usual way, then

$$X \times_{\delta_X} G := \overline{\operatorname{sp}}\{\delta_X(x)(1 \otimes M_f) : x \in X, f \in C_0(G)\}$$

is an $A \times_{\delta_A} G$-$B \times_{\delta_B} G$ imprimitivity bimodule [ER], Theorem 3.2. Conjugation by $1 \otimes \varrho$ defines an action $\widehat{\delta_X}$ of G on $X \times_{\delta_X} G$ such that

$$(\widehat{\delta_X})_s(\delta_X(x)(1 \otimes M_f)) = \delta_X(x)(1 \otimes M(\sigma_s(f))),$$

where $\sigma_s(f)(t) = f(ts)$. Since $\widehat{\delta_A}$ and $\widehat{\delta_B}$ are also given by $\operatorname{Ad}(1 \otimes \varrho_s)$ for $s \in G$, it follows that $\widehat{\delta_X}$ implements a Morita equivalence between $\widehat{\delta_A}$ and $\widehat{\delta_B}$. By [ER], Corollary 3.3, the twisting ideal I_{W_A} is the ideal of $A \times_{\delta_A} G$ induced from I_{W_B} via $X \times_{\delta_X} G$. Since these ideals are invariant under $\widehat{\delta_A}|_N$ and $\widehat{\delta_B}|_N$, the actions of N on the resulting quotients are Morita equivalent, too. But these actions are just the dual actions α and β on the twisted crossed products.

Assume now that $({}_A X_B, u)$ is a Morita equivalence for actions α and β of G, and suppose (π_A, π_X, π_B) is a faithful representation of ${}_A X_B$ on $(\mathscr{H}, \mathscr{K})$. Let

$$\operatorname{ind} \pi_A := \tilde{\pi}_A \times (1 \otimes \lambda^G)$$

be the regular representation of $A \times_\alpha G$, and analogously define

$$\operatorname{ind} \pi_X : C_c(G, X) \to B(L^2(G, \mathscr{K}), L^2(G, \mathscr{H}))$$

by

$$(\operatorname{ind} \pi_X(f)\xi)(t) = \int_G \pi_X(u_{t^{-1}}(f(s)))(\xi(s^{-1}t))\, ds .$$

Routine calculations show that $(\operatorname{ind}\pi_A, \operatorname{ind}\pi_X, \operatorname{ind}\pi_B)$ is a bimodule representation, and thus the image $X \times_{u,r} G := \operatorname{ind}_X(X \times_u G)$ is an imprimitivity bimodule for the reduced crossed products. In fact, since the elementary tensors $f \otimes x$, $f \in C_c(G)$, $x \in X$ generate a dense subset of $C_c(G, X)$, $(\operatorname{ind} \pi_X)(X \times_u G)$ is the closure of $\tilde{\pi}_X(X)(1 \otimes C_r^*(G))$ in

$$B(\mathscr{K} \otimes L^2(G), \mathscr{H} \otimes L^2(G)),$$

where $\tilde{\pi}_X(x)\xi(t) := \pi_X(u_t^{-1}(x))(\xi(t))$.

We now define

$$\hat{u} : X \times_{u,r} G \to B(\mathscr{K} \otimes L^2(G) \otimes L^2(G), \mathscr{H} \otimes L^2(G) \otimes L^2(G))$$

by $\hat{u}(y) = (1 \otimes W_G)(y \otimes 1)(1 \otimes W_G^*)$, and claim that $\hat{u}$ is a coaction of G on $X \times_{u,r} G$. For $x \in X$ and $z \in C_r^*(G)$, a direct computation shows that

$$\hat{u}(\tilde{\pi}_X(x)(1 \otimes z)) = (\tilde{\pi}_X(x) \otimes 1)(1 \otimes \delta_G(z)) .$$

Since $\tilde{\pi}_X(x) \otimes 1 \in \mathscr{M}((X \times_{u,r} G) \otimes C_r^*(G))$ and $1 \otimes \delta_G(z) \in \mathscr{M}((B \times_{\beta,r} G) \otimes C_r^*(G))$, it follows that $\hat{u}$ maps $X \times_{u,r} G$ into $\mathscr{M}((X \times_{u,r} G) \otimes C_r^*(G))$. Moreover, since $\hat{\alpha}$ and $\hat{\beta}$ are similarly defined using $\operatorname{Ad}(1 \otimes W_G)$, it follows that $(\hat{\alpha}, \hat{u}, \hat{\beta})$ is an imprimitivity bimodule homomorphism. To see that $(\hat{u} \otimes \mathrm{id}_G) \circ \hat{u} = (\mathrm{id}_{X \times_{u,r} G} \otimes \delta_G) \circ \hat{u}$, we just need to observe that, on elements of the form $\tilde{\pi}_X(x)$, both sides deliver the element $\tilde{\pi}_X(x) \otimes 1 \otimes 1$ of $\mathscr{M}(X \times_{u,r} G \otimes C_r^*(G) \otimes C_r^*(G))$, and that, on $1 \otimes z \in 1 \otimes C_r^*(G)$, the desired equation reduces to the comultiplication identity $(\delta_G \otimes \mathrm{id}_G) \circ \delta_G = (\mathrm{id}_G \otimes \delta_G) \circ \delta_G$. The rest is straightforward. □

Proof of Theorem 4.1. If (δ_A, W_A) and (δ_B, W_B) are Morita equivalent, then it follows from Lemma 3.2 that their dual actions are Morita equivalent. If the dual actions α and β are Morita equivalent, then Lemma 4.2 implies that the coactions $\hat{\alpha}$ and $\hat{\beta}$ are Morita equivalent, and Lemma 2.7 that the twisted coactions $(\operatorname{Inf}\hat{\alpha}, 1 \otimes 1)$ and $(\operatorname{Inf}\hat{\beta}, 1 \otimes 1)$ are

Morita equivalent. Since Morita equivalence is an equivalence relation on twisted coactions, it follows from Theorem 3.1 that (δ_A, W_A) is indeed Morita equivalent to (δ_B, W_B). □

From Theorem 4.1 we obtain the desired converse to Example 2.3.

Corollary 4.3. *Suppose that ε_A and ε_B are coactions of N on A and B, respectively. Then the inflated twisted coactions* $(\operatorname{Inf}\varepsilon_A, 1\otimes 1)$ *and* $(\operatorname{Inf}\varepsilon_B, 1\otimes 1)$ *of* $(G, G/N)$ *are Morita equivalent if and only if ε_A is Morita equivalent to ε_B.*

Proof. It was shown in [PR], Example 2.14, that $A\times_{\operatorname{Inf}\varepsilon_A, 1\otimes 1} G$ is isomorphic to $A\times_{\varepsilon_A} N$, and the corollary will follow from Theorem 4.1 if we can show that this isomorphism is equivariant with respect to the dual actions of N. But the isomorphism in [PR], Example 2.14 takes $\operatorname{Inf}\varepsilon_A(a)(1\otimes M_f)\in A\times_{\operatorname{Inf}\varepsilon_A} G$ to $\varepsilon_A(a)(1\otimes M(f|_N))\in A\times_{\varepsilon_A} N$. Since the dual action of $n\in N$ on both algebras is given by right translation of f, this map is N-equivariant. □

Morita equivalence and induced representations. Suppose that $({}_AX_B, \delta_X)$ is a Morita equivalence between coactions δ_A and δ_B of G. Let M be a closed amenable normal subgroup of G, and let $q^M: C_r^*(G)\to C_r^*(G/M)$ denote the quotient map. Then with $\delta_X| = (\mathrm{id}_X\otimes q^M)\circ\delta_X$, $(X, \delta_X|)$ is a Morita equivalence between the restrictions $\delta_A|$ and $\delta_B|$ of δ_A and δ_B to G/M. Thus the Morita equivalence (X, δ_X) gives several ways of inducing representations of crossed products: using the bimodule $X\times_{\delta_X} G$, we can induce representations from $B\times_{\delta_B} G$ to $A\times_{\delta_A} G$, and similarly, we can use $X\times_{\delta_X|} G/M$ to induce from $B\times_{\delta_B|} G/M$ to $A\times_{\delta_A|} G/M$. We claim that these are compatible with Mansfield's induction process $\operatorname{ind}_{G/M}^G$. More precisely, we have four maps defined on the various spaces of unitary equivalence classes of representations

$$\begin{array}{ccc}
\operatorname{Rep}(B\times_{\delta_B|} G/M) & \xrightarrow{(X\times_{\delta_X|} G/M)\text{-ind}} & \operatorname{Rep}(A\times_{\delta_A|} G/M) \\
\operatorname{ind}_{G/M}^G \downarrow & & \downarrow \operatorname{ind}_{G/M}^G \\
\operatorname{Rep}(B\times_{\delta_B} G) & \xrightarrow[(X\times_{\delta_X} G)\text{-ind}]{} & \operatorname{Rep}(A\times_{\delta_A} G)\,,
\end{array}$$

and we claim that:

Theorem 4.4. *The above diagram is commutative.*

Recall that Mansfield's bimodule X^A is actually an $(A\times_{\delta_A} G)\times_{\widehat{\delta_A}} M$-$A\times_{\delta_A|} G/M$ imprimitivity bimodule. The key ingredient in the proof of Theorem 4.4 will be the following proposition, which may be of some independent interest.

Proposition 4.5. *Suppose that (X, δ_X) is a Morita equivalence between δ_A and δ_B. Then*

$$X^A\otimes_{A\times_{\delta_A|} G/M}(X\times_{\delta_X|} G/M)\cong \big((X\times_{\delta_X} G)\times_{\widehat{\delta_X}} M\big)\otimes_{(B\times_{\delta_B} G)\times_{\widehat{\delta_B}} M} X^B$$

as $(A\times_{\delta_A} G)\times_{\widehat{\delta_A}} M$-$B\times_{\delta_B|} G/M$ imprimitivity bimodules.

Proof. We shall exploit the linking algebras for $X\times_{\delta_X} G$ and $X\times_{\delta_X|} G/M$. Let $L = \begin{pmatrix} A & X \\ \tilde X & B\end{pmatrix}$ be the linking algebra for X; then $\delta_L = \begin{pmatrix} \delta_A & \delta_X \\ \delta_{\tilde X} & \delta_B\end{pmatrix}$ is a coaction of G on L

[ER], Appendix. Represent ${}_AX_B$ faithfully on $(\mathscr{H}, \mathscr{K})$, so that L acts faithfully on $\mathscr{H} \oplus \mathscr{K}$, and let $p = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $q = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. By representing $L \times_{\delta_L} G$ faithfully on

$$(\mathscr{H} \otimes L^2(G)) \oplus (\mathscr{K} \otimes L^2(G)) \cong (\mathscr{H} \oplus \mathscr{K}) \otimes L^2(G)$$

as $\overline{\mathrm{sp}}\{\delta_L(l)(1 \otimes M_f)\}$, we can see from [ER], Theorem 3.1, that $L \times_{\delta_L} G$ is canonically isomorphic to

$$\begin{pmatrix} A \times_{\delta_A} G & X \times_{\delta_X} G \\ \tilde{X} \times_{\delta_{\tilde{X}}} G & B \times_{\delta_B} G \end{pmatrix},$$

where $A \times_{\delta_A} G$ and $B \times_{\delta_B} G$ are represented on $\mathscr{H} \otimes L^2(G)$ and $\mathscr{K} \otimes L^2(G)$ in the usual way. Thus $L \times_{\delta_L} G$ is the linking algebra for $X \times_{\delta_X} G$ as an $A \times_{\delta_A} G$-$B \times_{\delta_B} G$ imprimitivity bimodule. Further, if $\widehat{\delta_L}$ is the dual action of G on $L \times_{\delta_L} G$, then the restrictions of $\widehat{\delta_L}$ to the corners are just the dual actions $\widehat{\delta_A}$, $\widehat{\delta_B}$ and $\widehat{\delta_X}$, where $\widehat{\delta_X}$ is defined as in the proof of Lemma 4.2.

On the other hand, if γ is a strongly continuous action of H on a linking algebra L, which restricts to actions α, u and β on the corners A, X and B, respectively (i.e., if (X, u) is a Morita equivalence for α and β), then the construction of the crossed product $L \times_\gamma H$ as the completion of $C_c(H, L) \cong \begin{pmatrix} C_c(H, A) & C_c(H, X) \\ C_c(H, \tilde{X}) & C_c(H, B) \end{pmatrix}$ shows that

$$L \times_\gamma H = \begin{pmatrix} A \times_\alpha H & X \times_u H \\ \tilde{X} \times_{\tilde{u}} H & B \times_\beta H \end{pmatrix}.$$

Applying this to the linking algebra $L \times_{\delta_L} G$ gives

$$(L \times_{\delta_L} G) \times_{\widehat{\delta_L}} M = \begin{pmatrix} (A \times_{\delta_A} G) \times_{\widehat{\delta_A}} M & (X \times_{\delta_X} G) \times_{\widehat{\delta_X}} M \\ (\tilde{X} \times_{\delta_{\tilde{X}}} G) \times_{\widehat{\delta_{\tilde{X}}}} M & (B \times_{\delta_B} G) \times_{\widehat{\delta_B}} M \end{pmatrix}.$$

On the other hand, representing $L \times_{\delta_L|} G/M$ on $(\mathscr{H} \otimes L^2(G)) \oplus (\mathscr{K} \otimes L^2(G))$ via the covariant representation $(\delta_L, 1 \otimes (M \circ \psi))$, where ψ denotes the inclusion of $C_0(G/M)$ into $C_b(G)$, shows that

$$L \times_{\delta_L|} G/M = \begin{pmatrix} A \times_{\delta_A|} G/M & X \times_{\delta_X|} G/M \\ \tilde{X} \times_{\delta_{\tilde{X}}|} G/M & B \times_{\delta_B|} G/M \end{pmatrix}.$$

(This representation of $L \times_{\delta_L|} G/M$ is faithful by [M], Proposition 7.)

We have now shown that Mansfield's $(L \times_{\delta_L} G) \times_{\widehat{\delta_L}|_M} M$-$L \times_{\delta_L|} G/M$ imprimitivity bimodule X^L is actually an imprimitivity bimodule for the linking algebras of the imprimitivity bimodules $(X \times_{\delta_X} G) \times_{\widehat{\delta_X}|_M} M$ and $X \times_{\widehat{\delta_X}|} G/M$. To proceed further, we need a lemma.

Lemma 4.6. *Let A, B, C, D be C^*-algebras, and suppose that ${}_AX_B$ and ${}_CY_D$ are imprimitivity bimodules. Let $L_{A,B}$ and $L_{C,D}$ be the linking algebras for X and Y, and view*

$p = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $q = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ *as elements of both* $\mathcal{M}(L_{C,D})$ *and* $\mathcal{M}(L_{A,B})$. *Suppose further that* E *is an* $L_{C,D}$-$L_{A,B}$ *imprimitivity bimodule. Then* pEq *is isomorphic to the balanced tensor products* $pEp \otimes_A X$ *and* $Y \otimes_D qEq$ *as* C-B *imprimitivity bimodules. In particular,* $pEp \otimes_A X$ *is isomorphic to* $Y \otimes_D qEq$.

Proof. Let us first mention that the corners pEp, qEq and pEq are naturally C-A, D-B and C-B imprimitivity bimodules, respectively: just restrict the inner products on E and the actions of $L_{C,D}$ and $L_{A,B}$ on E to the appropriate corners. For example, the A-valued inner product on pEp is given by

$$\langle pep, pfp\rangle_A = \langle pep, pfp\rangle_{L_{A,B}} = p\langle pep, pfp\rangle_{L_{A,B}} p \in pL_{A,B}p \cong A\,.$$

We want to define $\Phi: pEp \otimes_A X \to pEq$ by $\Phi(z \otimes x) = z \cdot x$, where the action of

$$x \in X = pL_{A,B}q$$

on z is the right action of $L_{A,B}$ on E. Since $A = pL_{A,B}p$ is a full corner in $L_{A,B}$, we know that $L_{A,B}pL_{A,B}$ is dense in $L_{A,B}$, which implies that $(pEp)\cdot(pL_{A,B}q) = pEL_{A,B}pL_{A,B}q$ is dense in pEq. To show that Φ extends to a well-defined map on $pEp \otimes_A X$ we have to show that Φ preserves the C- and B-valued inner products. For this let w, $z \in Z$, $x, y \in X$, and compute:

$$\begin{aligned}
\langle z\cdot x, w\cdot y\rangle_B &= \langle z\cdot x, w\cdot y\rangle_{L_{A,B}}\\
&= x^*\langle z, w\rangle_{L_{A,B}}\, y\\
&= x^*\langle z, w\rangle_A\, y\\
&= (\langle w, z\rangle_A\, x)^*\, y\\
&= \langle\langle w, z\rangle_A\cdot x, y\rangle_B\\
&= \langle z\otimes x, w\otimes y\rangle_B\,.
\end{aligned}$$

A similar computation shows that Φ also preserves the C-valued inner products. Since Φ clearly preserves the left and right actions of C and B, we deduce that Φ is well-defined, and also that it induces an isomorphism of $pEp \otimes_A X$ onto pEq. The rest of the lemma follows by symmetry. □

Proof of Proposition 4.5 (*continued*). We intend to apply Lemma 4.6 to $E = X^L$. If we can show that $X^A \cong pX^Lp$ and $X^B \cong qX^Lq$, then we shall have

$$\begin{aligned}
X^A \otimes_{A\times G/M}(X\times G/M) &\cong pX^Lp \otimes_{A\times G/M} p(L\times G/M)\,q\\
&\cong p\big((L\times G)\times M\big)q \otimes_{(B\times G)\times M} qX^Lq\\
&\cong (X\times G)\times M \otimes_{(B\times G)\times M} X^B\,,
\end{aligned}$$

as required. By symmetry, it will be enough to show that $X^A \cong pX^Lp$.

If we represent both $L \times_{\delta|} G/M$ and $L \times_{\delta_L} G$ on $\mathscr{L} = (\mathscr{H} \otimes L^2(G)) \oplus (\mathscr{K} \otimes L^2(G))$, then the bimodule X^L is also represented as operators on $\mathscr{L}$, and p, q become the projections onto the two factors. Hence pX^Lp acts naturally on $\mathscr{H} \otimes L^2(G)$. When we realise $\mathscr{M}(L \otimes C_r^*(G))$ as matrices of operators on $\mathscr{L}$, the slice maps S_u corresponding to $u \in A(G)$ can be performed elementwise, and hence

$$\delta_L(S_u(\delta_L\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}))(1 \otimes M_f) = \delta_L(S\begin{pmatrix} \delta_A(a) & 0 \\ 0 & 0 \end{pmatrix})(1 \otimes M_f)$$
$$= \begin{pmatrix} \delta_A(S_u(\delta_A(a)))(1 \otimes M_f) & 0 \\ 0 & 0 \end{pmatrix}.$$

It follows that the map

$$\theta : \delta_A(S_u(\delta_A(a)))(1 \otimes M_f) \mapsto \delta_L(S_u(\delta_L\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}))(1 \otimes M_f)$$

is isometric for the operator norm, and hence extends to a map of the dense subspace $\mathscr{D}^A$ of X^A into the corresponding subspace $p\mathscr{D}^Lp$ of pX^Lp – indeed, this is a bijection because

$$p\delta_L(\begin{pmatrix} a & x \\ \tilde{y} & b \end{pmatrix})p = \begin{pmatrix} \delta_A(a) & 0 \\ 0 & 0 \end{pmatrix} \quad \text{for any} \begin{pmatrix} a & x \\ \tilde{y} & b \end{pmatrix} \in L.$$

Since the identification of $\mathscr{D}^A$ with the corner in $\mathscr{D}^L$ preserves the dual action $\hat{\delta}$, it follows from [M], Lemma 18, that θ preserves the $\mathscr{D}_M^A$-valued inner product, modulo the identification of $A \times G/M$ with the corresponding corner in $L \times G/M$. Thus it extends to an isomorphism of the Hilbert $(A \times G/M)$-module X^A onto the completion pX^Lp of $p\mathscr{D}^Lp$. This isomorphism converts the natural left action of $\mathscr{D}^A \subset A \times_\delta G$ into the left action of $p\mathscr{D}^Lp$. The action of $(A \times_{\delta_A} G) \times M$ on the left of X^A is the integrated form of the natural left action of $\mathscr{D}^A$ on itself and the representation of M given by

$$m \cdot x := \Delta(m)^{1/2} \operatorname{Ad}(1 \otimes \varrho_m)(x);$$

since θ preserves this action of M, it intertwines the left action of $(A \times_{\delta_A} G) \times M$ and that of the corresponding (isomorphic) corner $p((L \times_{\delta_L} G) \times M)p$. □

Proof of Theorem 4.4. Proposition 4.5 says that the diagram

$$\begin{array}{ccc} \operatorname{Rep}(B \times_{\delta_B|} G/M) & \xrightarrow{(X \times_{\delta_X|} G/M)\text{-ind}} & \operatorname{Rep}(A \times_{\delta_A|} G/M) \\ \downarrow \operatorname{ind}_{G/M}^G & & \downarrow \operatorname{ind}_{G/M}^G \\ \operatorname{Rep}((B \times_{\delta_B} G) \times_{\hat{\delta}_B} M) & \xrightarrow[((X \times_{\delta_X} G) \times M)\text{-ind}]{} & \operatorname{Rep}((A \times_{\delta_A} G) \times_{\hat{\delta}_A} M) \end{array} \tag{4.1}$$

commutes. However, it is shown in the discussion preceding Proposition 1 of [Ech] that, for any Morita equivalence (X, u) between ordinary systems (C, M, β) and (D, M, γ), and any covariant representation (π, U) of (D, M, γ), the induced representation $(X \times_u G)$-ind $\pi \times U$ is unitarily equivalent to a representation of the form $(X\text{-ind}\,\pi, V)$; in particular, the diagram

$$
\begin{array}{ccc}
\operatorname{Rep}((B \times_{\delta_B|} G) \times_{\hat{\delta}_B} M) & \xrightarrow{((X \times_{\delta_X} G) \times M)\text{-ind}} & \operatorname{Rep}((A \times_{\delta_A} G) \times_{\hat{\delta}_A} M) \\
\text{res} \downarrow & & \downarrow \text{res} \\
\operatorname{Rep}(B \times_{\delta_B} G) & \xrightarrow[(X \times_{\delta_X} G)\text{-ind}]{} & \operatorname{Rep}(A \times_{\delta_A} G)
\end{array} \tag{4.2}
$$

commutes. Putting together the diagrams (4.1) and (4.2) gives the theorem. □

Induced representations of crossed products by twisted coactions and the stabilisation trick. We now suppose that (δ, W) is a twisted coaction of $(G, G/N)$ on A, and that M is a closed normal subgroup of G contained in N. Since the composition $q^{N/M} \circ q^M$ of the quotient maps $q^M : C_r^*(G) \to C_r^*(G/M)$ and $q^{N/M} : C_r^*(G/M) \to C_r^*(G/N)$ is just $q^N : C_r^*(G) \to C_r^*(G/N)$, it follows that $(\delta|_{G/M})|_{(G/M)/(G/N)} = \delta|_{G/N}$, which then implies that $(\delta|_{G/M}, W)$ is a twisted coaction of $(G/M, G/N)$ on A. Unless confusion seems likely, we shall denote all these restricted coactions by $\delta|$.

Suppose now that $\varrho \times \mu$ is a representation of $A \times_{\delta|, W} G/M$. We want to define the induced representation $\operatorname{ind}_{G/M}^G(\varrho \times \mu)$ of $A \times_{\delta, W} G$ by viewing $\varrho \times \mu$ as a representation of $A \times_{\delta|} G/M$ and inducing this representation to $A \times_\delta G$ via Mansfield's bimodule X. For this to give a representation of $A \times_{\delta, W} G$ we need to know that Mansfield's induced representation $\operatorname{ind}_{G/M}^G(\varrho \times \mu)$ of $A \times_\delta G$ preserves W if $\varrho \times \mu$ does. But this follows from the proof of [PR], Lemma 4.2, which is the special case $M = N$, and we therefore have a well-defined inducing procedure $\operatorname{ind}_{G/M}^G$ on the twisted crossed products.

Note that if (X, δ_X) is a Morita equivalence for two twisted actions (δ_A, W_A) and (δ_B, W_B), then the equation $(\delta_X|_{G/M})|_{G/N} = \delta_X|_{G/N}$ implies that $(X, \delta_X|)$ is a Morita equivalence for $(\delta_A|_{G/M}, W_A)$ and $(\delta_B|_{G/M}, W_B)$. Moreover, the Morita equivalences between $A \times_{\delta_A, W_A} G$ and $B \times_{\delta_B, W_B} G$ and between $A \times_{\delta_A|, W_A} G/M$ and $B \times_{\delta_B|, W_B} G/M$ are given by quotients of $X \times_{\delta_X} G$ and $X \times_{\delta_X|} G/M$, respectively: we denote these quotients by $X \times_{\delta_X, W} G$ and $X \times_{\delta_X|, W} G/M$. Theorem 4.4 implies that we have a commutative diagram

$$
\begin{array}{ccc}
\operatorname{Rep}(B \times_{\delta_B|, W_B} G/M) & \xrightarrow{(X \times_{\delta_X|, W} G/M)\text{-ind}} & \operatorname{Rep}(A \times_{\delta_A|, W_A} G/M) \\
\operatorname{ind}_{G/M}^G \downarrow & & \downarrow \operatorname{ind}_{G/M}^G \\
\operatorname{Rep}(B \times_{\delta_B, W_B} G) & \xrightarrow[(X \times_{\delta_X, W} G)\text{-ind}]{} & \operatorname{Rep}(A \times_{\delta_A, W_B} G)\,.
\end{array}
$$

We want a similar diagram to relate the representation theory of a twisted coaction (δ, W) of (G, N) on B to that of the Morita equivalent coaction ε of N on A. By definition, this Morita equivalence is really an equivalence of (δ, W) and the inflated twisted coaction $(\operatorname{Inf}\varepsilon, 1 \otimes 1)$. Since $q^M \circ \lambda^G|_N = \lambda^{G/M}|_{N/M} \circ p^M$, where $p^M : C^*(N) \to C^*(N/M)$ denotes the quotient map, it follows that $(\operatorname{Inf}\varepsilon)|_{G/M} = \operatorname{Inf}(\varepsilon|_{N/M})$. Thus we have natural isomorphisms $A \times_{\operatorname{Inf}\varepsilon, 1 \otimes 1} G \cong A \times_\varepsilon N$, and $A \times_{(\operatorname{Inf}\varepsilon)|_{G/M}, 1 \otimes 1} G/M \cong A \times_{\varepsilon|_{N/M}} N/M$, which allows us to view $X \times_{\delta_X, W} G$ and $X \times_{\delta_X|, W} G/M$ as $A \times_\varepsilon N$-$B \times_{\delta, W} G$ and

$$A \times_{\varepsilon|} N/M\text{-}B \times_{\delta|, W} G/M$$

imprimitivity bimodules, respectively.

Theorem 4.7. *Let (δ, W) be a twisted coaction of $(G, G/N)$ on B and let $M \subset N$ be a closed normal subgroup of G. If ε is a coaction of N on A, and (X, δ_X) is a Morita equivalence between* $(\mathrm{Inf}\,\varepsilon, 1 \otimes 1)$ *and (δ, W), then we have a commutative diagram*

$$\begin{array}{ccc} \mathrm{Rep}(B \times_{\delta|, W} G/M) & \xrightarrow{(X \times_{\delta_X|, W} G/M)\text{-ind}} & \mathrm{Rep}(A \times_{\varepsilon|} N/M) \\ \mathrm{ind}_{G/M}^{G} \downarrow & & \downarrow \mathrm{ind}_{N/M}^{N} \\ \mathrm{Rep}(B \times_{\delta, W} G) & \xrightarrow[(X \times_{\delta_X, W} G)\text{-ind}]{} & \mathrm{Rep}(A \times_{\varepsilon} N)\,. \end{array}$$

Theorem 4.7 will follow from Theorem 4.4 and the proposition below. Before we state it, recall that the natural isomorphism of $A \times_{\mathrm{Inf}\,\varepsilon, 1 \otimes 1} G$ onto $A \times_{\varepsilon} N$ is given on $A \times_{\mathrm{Inf}\,\varepsilon} G$ by

$$\Phi^G(\mathrm{Inf}\,\varepsilon(a)(1 \otimes M_f)) = \varepsilon(a)(1 \otimes M_{f|_N})\,.$$

Note that, according to the definitions preceding the theorem, the representation

$$(X \times_{\delta_X, W} G)\text{-ind}\,(\varrho \times \mu)$$

of $A \times_{\varepsilon} N$ is by definition the representation $\sigma \times v$ such that

$$(X \times_{\delta_X} G)\text{-ind}\,(\varrho \times \mu) = (\sigma \times v) \circ \Phi^G\,.$$

Further, let $\Phi^{G/M}$ denote the corresponding homomorphism of $A \times_{\mathrm{Inf}\,\varepsilon|} G/M$ onto $A \times_{\varepsilon|} N/M$.

Proposition 4.8. *Let Φ^G and $\Phi^{G/M}$ be as above and let $\varrho \times \mu$ be a representation of $A \times_{\varepsilon|} N/M$. Then the representations $(\mathrm{ind}_{M/N}^{N}(\varrho \times \mu)) \circ \Phi^G$ and $\mathrm{ind}_{G/M}^{G}((\varrho \times \mu) \circ \Phi^{G/M})$ of $A \times_{\mathrm{Inf}\,\varepsilon} G$ are unitarily equivalent.*

Proof. Let $\mathscr{D}$ and $\mathscr{D}_M$ denote Mansfield's dense subalgebras of $A \times_{\mathrm{Inf}\,\varepsilon} G$ and $A \times_{\mathrm{Inf}\,\varepsilon|} G/M$, respectively, and $\mathscr{E}$ and $\mathscr{E}_M$ the corresponding dense subalgebras of $A \times_{\varepsilon} N$ and $A \times_{\varepsilon|} N/M$. We claim that Φ^G maps $\mathscr{D}$ onto a dense subalgebra of $\mathscr{E}$ and that $\Phi^{G/M}$ maps $\mathscr{D}_M$ onto a dense subalgebra of $\mathscr{E}_M$. To see this, observe that for $u \in A_c(G)$ and $z \in C_r^*(N)$ the pairings of $C_r^*(G)$ with $A(G)$ and $C_r^*(N)$ with $A(N)$ are related by

$$\langle \lambda^G|_N(z), u \rangle = \langle z, u|_N \rangle\,.$$

Thus for $u \in A_c(G)$ we have $S_u(\mathrm{Inf}\,\varepsilon(a)) = S_u((\mathrm{id}_A \otimes \lambda^G|_N) \circ \varepsilon(a)) = S_{u|_N}(\varepsilon(a))$, and

$$\Phi^G(\mathrm{Inf}\,\varepsilon(S_u(\mathrm{Inf}\,\varepsilon(a)))(1 \otimes M_f)) = \varepsilon(S_{u|_N}(\varepsilon(a)))(1 \otimes M_{f|_N})\,.$$

Since Φ^G extends to a homomorphism on $A \times_{\mathrm{Inf}\,\varepsilon} G$, it is norm-decreasing; thus if $x \in \mathscr{D}$ is (u, E) for some $u \in A_c(G)$ and $E \subseteq G$ compact, then $\Phi^G(x)$ is $(u|_N, E \cap N)$, hence in $\mathscr{E}$. The same argument shows that $\Phi^{G/M}$ maps $\mathscr{D}_M$ into $\mathscr{E}_M$. The density of $\mathscr{D}$ in $\mathscr{E}$ now follows from $A_c(G)|_N = A_c(N)$, which in turn holds because $A(G)|_N = A(N)$ [H] and because for each compact set $C \subseteq G$ there exists $w \in A_c(G)$ such that $w = 1$ on C. Similarly, the image of $\mathscr{D}_M$ is dense in $\mathscr{E}_M$.

Now view $\mathscr{D}$ and $\mathscr{E}$ as dense subspaces of the bimodules X and Y used by Mansfield to induce from G/M to G and N/M to N, respectively. Then from the definition of the $\mathscr{D}_M$- and $\mathscr{E}_M$-valued inner products, [M], Theorem 19, and [M], Proposition 17, we deduce that $\langle \Phi^G(x), \Phi^G(y) \rangle_{\mathscr{E}_M} = \Phi^{G/M}(\langle x, y \rangle_{\mathscr{D}_M})$ for all $x, y \in \mathscr{D}$ (it may help to note that $\phi(f)|_N = \phi(f|_N)$). Thus Φ^G extends to a bounded linear map Φ^X from X onto a dense subspace of Y, satisfying

$$\langle \Phi^X(x), \Phi^X(y) \rangle_{A \times_{\varepsilon|} N/M} = \Phi^{G/M}(\langle x, y \rangle_{A \times_{\operatorname{Inf} \varepsilon|} G/M}) \, .$$

Since $\Phi^G(xb) = \Phi^G(x)\Phi^{G/M}(b)$ for all $x \in \mathscr{D}$ and $b \in \mathscr{D}_M$, and Φ^G is a homomorphism, we have $\Phi^X(a \cdot x \cdot b) = \Phi^G(a) \cdot \Phi^X(x) \cdot \Phi^{G/M}(b)$ for $a \in A \times_{\operatorname{Inf} \varepsilon} G$, $x \in X$ and $b \in A \times_{\operatorname{Inf} \varepsilon|} G/M$.

If now $\varrho \times \mu$ is a representation of $A \times_{\varepsilon|} N/M$ on $\mathscr{H}$, then it follows from these properties of Φ^G that the map $U: X \otimes_{A \times_{\operatorname{Inf} \varepsilon|} G/M} \mathscr{H} \to Y \otimes_{A \times_{\varepsilon|} N/M} \mathscr{H}$ defined on elementary tensors by

$$U(x \otimes v) = \Phi^X(x) \otimes v$$

is a unitary which intertwines $\operatorname{ind}_{G/M}^G((\varrho \times \mu) \circ \Phi^{G/M})$ and $(\operatorname{ind}_{M/N}^G(\varrho \times \mu)) \circ \Phi^G$. □

Remark 4.9. Since Φ^G is equivariant with respect to the dual actions of N on

$$A \times_{\operatorname{Inf} \varepsilon} G \quad \text{and} \quad A \times_{\varepsilon} N \, ,$$

Φ^G extends to a homomorphism Φ^M of $(A \times_{\operatorname{Inf} \varepsilon} G) \times_{\widehat{\operatorname{Inf} \varepsilon}} M$ onto $(A \times_{\varepsilon} N) \times_{\hat{\varepsilon}} M$, and $(\Phi^M, \Phi^X, \Phi^{G/M})$ is in fact an imprimitivity bimodule homomorphism from X onto Y.

Proof of Theorem 4.7. Let $\varrho \times \mu$ be a representation of $B \times_{\delta|} G/M$ which preserves W. Then it follows from the proof of either [Bui], Theorem 3.3, or [ER], Corollary 3.3, that the representation $(X \times_{\delta|} G/M)$-ind$(\varrho \times \mu)$ of $A \times_{\operatorname{Inf} \varepsilon|} G/M$ preserves the trivial twist $1 \otimes 1$. Thus $(X \times_{\delta_X|} G/M)$-ind$(\varrho \times \mu)$ has the form $(\sigma \times v) \circ \Phi^{G/M}$ for some representation $\sigma \times \mu$ of $A \times_{\varepsilon|} N/M$, and then $\sigma \times \mu$ is by definition the representation of $A \times_{\varepsilon|} N/M$ induced from $\varrho \times \mu$ via $X \times_{\delta_X|} G/M$.

It follows from Proposition 4.8 that $\operatorname{ind}_{G/M}^G((X \times_{\delta_X|} G/M)\text{-ind}(\varrho \times \mu))$ is equivalent to $(\operatorname{ind}_{N/M}^N(\sigma \times v)) \circ \Phi^G$, and hence it too preserves $1 \otimes 1$. On the other hand, we know from Theorem 4.4 that $\operatorname{ind}_{G/M}^G((X \times_{\delta|} G/M)\text{-ind}(\varrho \times \mu))$ is equivalent to

$$(X \times_{\delta_X} G)\text{-ind}(\operatorname{ind}_{G/M}^G(\varrho \times \mu)) \, .$$

It follows, again using [Bui], Theorem 3.3, or [ER], Corollary 3.3, that $\operatorname{ind}_{G/M}^G(\varrho \times \mu)$ preserves the twist W. Moreover, since

$$(X \times_{\delta_X} G)\text{-ind}(\operatorname{ind}_{G/M}^G(\varrho \times \mu))$$

is equivalent to $(\operatorname{ind}_{N/M}^N(\sigma \times v)) \circ \Phi^G$, it follows that $\operatorname{ind}_{N/M}^N(\sigma \times v)$ is the representation of $A \times_{\varepsilon} N$ induced from $\operatorname{ind}_{G/M}^G(\varrho \times \mu)$ via $X \times_{\delta_X} G$. This completes the proof. □

Appendix

Suppose (X, δ_X) is a Morita equivalence between coactions δ_A and δ_B of G on C^*-algebras A and B, in the sense of [ER]; that is, δ_X satisfies conditions (1), (2) and (4) of Definition 2.1, and the natural identification of A with $\mathcal{K}(X_B)$ carries δ_A into $\delta_{\mathcal{K}(X)}$. We shall show that, provided δ_A and δ_B are nondegenerate, (X, δ_X) is also a Morita equivalence in the sense of [Bui]; since most coactions are known to be nondegenerate (see [Q] for a recent discussion), this says the two theories of Morita equivalence are effectively the same. We have to show that δ_X satisfies condition (3) of Definition 2.1, and indeed we shall do better:

Lemma A. *If δ_A and δ_B are nondegenerate, then $\delta_X(X)\cdot\big(1\otimes C_r^*(G)\big)$ is dense in $X\otimes C_r^*(G)$. As a consequence, $\delta_X(X)\cdot\big(B\otimes C_r^*(G)\big)$ is dense in $X\otimes C_r^*(G)$.*

Proof. Let $L=\begin{pmatrix} A & X\\ \tilde{X} & B\end{pmatrix}$ denote the linking algebra for ${}_AX_B$ and let $\delta_L=\begin{pmatrix}\delta_A & \delta_X\\ \delta_{\tilde{X}} & \delta_B\end{pmatrix}$ denote the corresponding coaction of G on L. We show that δ_L is nondegenerate. Then $\delta_L(L)\big(1\otimes C_r^*(G)\big)$ is dense in $L\otimes C_r^*(G)$, which implies that the upper right corner $\delta_X(X)\big(1\otimes C_r^*(G)\big)$ of $\delta_L(L)\big(1\otimes C_r^*(G)\big)$ is dense in the upper right corner $X\otimes C_r^*(G)$ of $L\otimes C_r^*(G)$.

We know from [K], Theorem 5, that δ_L is nondegenerate if and only if the closure $I(L)$ of $\{S_f\big(\delta_L(l)\big): f\in A(G),\ l\in L\}$ coincides with L. Factorising $A(G)$ via the action of $C_r^*(G)$ on the left and right, this will follow if the closed linear span of

$$\{(1\otimes z)\,\delta_L(l)(1\otimes w): z, w\in C_r^*(G),\ l\in L\}$$

equals $L\otimes C_r^*(G)$.

To see this let Y denote the closure of $\big(1\otimes C_r^*(G)\big)\,\delta_X(X)\big(1\otimes C_r^*(G)\big)$ in $X\otimes C_r^*(G)$. Since

$$\big(A\otimes C_r^*(G)\big)\cdot\big(1\otimes C_r^*(G)\big)\,\delta_X(X)\big(1\otimes C_r^*(G)\big)=\big(1\otimes C_r^*(G)\big)\,\delta_X(A\cdot X)\big(1\otimes C_r^*(G)\big),$$

it follows that Y is a left $A\otimes C_r^*(G)$-submodule of $X\otimes C_r^*(G)$, and a similar argument shows that Y is also a right $B\otimes C_r^*(G)$-submodule of $X\otimes C_r^*(G)$. Since Y contains the closure of $\delta_X(X)\big(1\otimes C_r^*(G)\big)$ (by the existence of an approximate identity in $C_r^*(G)$) we see that

$$\begin{aligned}\langle Y, Y\rangle_{\mathcal{M}(B\otimes C_r^*(G))} &\supseteq \langle \delta_X(X)\cdot\big(1\otimes C_r^*(G)\big), \delta_X(X)\cdot\big(1\otimes C_r^*(G)\big)\rangle_{\mathcal{M}(B\otimes C_r^*(G))}\\ &=\big(1\otimes C_r^*(G)\big)\,\delta_B(\langle X, X\rangle_B)\big(1\otimes C_r^*(G)\big)\end{aligned}$$

is dense in $B\otimes C_r^*(G)$, since δ_B is nondegenerate. It follows from this and the Rieffel correspondence between the closed submodules of $X\otimes C_r^*(G)$ and the closed ideals of $B\otimes C_r^*(G)$ [Rieff2], Theorem 3.1, that $Y=X\otimes C_r^*(G)$. Since

$$\big(1\otimes C_r^*(G)\big)\,\delta_A(A)\big(1\otimes C_r^*(G)\big)$$

and $(1 \otimes C_r^*(G))\,\delta_B(B)(1 \otimes C_r^*(G))$ are dense in $A \otimes C_r^*(G)$ and $B \otimes C_r^*(G)$, respectively, this implies that

$$\begin{aligned}&(1 \otimes C_r^*(G))\,\delta_L(L)(1 \otimes C_r^*(G))\\ &= \begin{pmatrix} (1 \otimes C_r^*(G))\,\delta_A(A)(1 \otimes C_r^*(G)) & (1 \otimes C_r^*(G))\,\delta_X(X)(1 \otimes C_r^*(G)) \\ (1 \otimes C_r^*(G))\,\delta_{\tilde{X}}(\tilde{X})(1 \otimes C_r^*(G)) & (1 \otimes C_r^*(G))\,\delta_B(B)(1 \otimes C_r^*(G)) \end{pmatrix}\end{aligned}$$

is dense in $L \otimes C_r^*(G)$. Thus δ_L is nondegenerate.

To see the last statement, use an approximate identity in B. □

References

[BS] *S. Baaj* and *G. Skandalis*, C^*-algèbres de Hopf et théorie de Kasparov équivariante, K-Theory **2** (1989), 683–721.

[BGR] *L.G. Brown*, *P. Green* and *M.A. Rieffel*, Stable isomorphism and strong Morita equivalence of C^*-algebras, Pacific J. Math. **71** (1977), 349–363.

[Bui] *H.H. Bui*, Morita equivalence of twisted crossed products by coactions, J. Funct. Anal. **123** (1994), 59–98.

[Com] *F. Combes*, Crossed products and Morita equivalence, Proc. London Math. Soc. **49** (1984), 289–306.

[CMW] *R.E. Curto*, *P.S. Muhly* and *D.P. Williams*, Cross products of strongly Morita equivalent C^*-algebras, Proc. Amer. Math. Soc. **90** (1984), 528–530.

[Ech] *S. Echterhoff*, Morita equivalent twisted actions and a new version of the Packer-Raeburn stabilization trick, J. London Math. Soc. (2) **50** (1994), 170–186.

[ER] *S. Echterhoff* and *I. Raeburn*, Multipliers of imprimitivity bimodules and Morita equivalence of crossed products, Math. Scand., to appear.

[G] *P. Green*, The local structure of twisted covariance algebras, Acta Math. **140** (1978), 191–250.

[H] *C. Herz*, Le rapport entre l'algèbre A_p d'un groupe et d'un sous-groupe, C.R. Acad. Sci. Paris A **271** (1970), 244–246.

[IT] *S. Imai* and *H. Takai*, On a duality for crossed products by a locally compact group, J. Math. Soc. Japan **30** (1978), 495–504.

[Kal] *S. Kaliszewski*, A note on Morita equivalence of twisted C^*-dynamical systems, Proc. Amer. Math. Soc., to appear.

[K] *Y. Katayama*, Takesaki's duality for a nondegenerate coaction, Math. Scand. **55** (1985), 141–151.

[L] *M.B. Landstad*, Duality for covariant systems, Trans. Amer. Math. Soc. **248** (1979), 223–267.

[LPRS] *M.B. Landstad*, *J. Phillips*, *I. Raeburn* and *C.E. Sutherland*, Representations of crossed products by coactions and principal bundles, Trans. Amer. Math. Soc. **299** (1987), 747–784.

[M] *K. Mansfield*, Induced representations of crossed products by coactions, J. Funct. Anal. **97** (1991), 112–161.

[PaR] *J.A. Packer* and *I. Raeburn*, Twisted crossed products of C^*-algebras, Math. Proc. Camb. Phil. Soc. **106** (1989), 293–311.

[PR] *J. Phillips* and *I. Raeburn*, Twisted crossed products by coactions, J. Austral. Math. Soc. (A) **56** (1994), 320–344.

[Q] *J.C. Quigg*, Full and reduced C^*-coactions, Math. Proc. Camb. Phil. Soc., to appear.

[R] *I. Raeburn*, On crossed products by coactions and their representation theory, Proc. London Math. Soc. **64** (1992), 625–652.

[Rieff1] *M.A. Rieffel*, Induced representations of C^*-algebras, Adv. Math. **13** (1974), 176–257.

[Rieff2] *M.A. Rieffel*, Unitary representations of group extensions; an algebraic approach to the theory of Mackey and Blattner, Studies in Analysis, Adv. Math. Supp. Stud. **4**, Academic Press, New York (1979), 43–82.

Fachbereich Mathematik-Informatik, Universität-Gesamthochschule Paderborn, D-33095 Paderborn, Germany
e-mail: echter@uni-paderborn.de

Department of Mathematics, University of Newcastle, NSW 2308, Australia
e-mail: iain@math.newcastle.edu.au

Eingegangen 26. Oktober 1994, in revidierter Fassung 10. Juli 1995