Pacific Journal of Mathematics

Volume 218 No. 2 February 2005

PACIFIC JOURNAL OF MATHEMATICS

http://www.pjmath.org

Founded in 1951 by

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

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

Regular subscription rate: $400.00 a year (10 issues). Special rate: $200.00 a year to individual members of supporting institutions.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94707-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840 is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94707-0163.

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

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

CHEVALLEY COHOMOLOGY FOR KONTSEVICH'S GRAPHS

DIDIER ARNAL, ANGELA GAMMELLA AND MOHSEN MASMOUDI

We introduce the Chevalley cohomology for the graded Lie algebra of polyvector fields on $\mathbb{R}^d$. This cohomology occurs naturally in the problem of construction and classification of formalities on the space $\mathbb{R}^d$. Considering only graph formalities, that is, formalities defined with the help of graphs as in the original construction of Kontsevich, we define (as the first and third authors did earlier for the Hochschild cohomology) the Chevalley cohomology directly on spaces of graphs. More precisely, observing first a noteworthy property for Kontsevich's explicit formality on $\mathbb{R}^d$, we restrict ourselves to graph formalities with that property. With this restriction, we obtain some simple expressions for the Chevalley coboundary operator; in particular, we can write this cohomology directly on the space of purely aerial, nonoriented graphs. We also give examples and applications.

1. Introduction

In this article, we study formalities on the space $\mathbb{R}^d$, which are defined as follows. Let $T_{\text{poly}}(\mathbb{R}^d)[1]$ be the space of polyvector fields on $\mathbb{R}^d$ graded by $|\alpha| = \text{degree}(\alpha) = k-2$ if α is a k-vector field (the [1] stands for this choice of translation on degrees). Similarly, $D_{\text{poly}}(\mathbb{R}^d)[1])$ will denote the polydifferential operators on $\mathbb{R}^d$ graded by $|D| = m-2$ if D is an m-differential operator. We view both spaces as formal graded manifolds; see [Kontsevich 1997; 2003]. A *formality* is a formal nonlinear mapping $\mathcal{F}$ between $T_{\text{poly}}(\mathbb{R}^d)[1]$ and $D_{\text{poly}}(\mathbb{R}^d)[1]$, intertwining their natural vector fields Q and Q'.

The monomial functions $\alpha_1 \centerdot \alpha_2 \centerdot \dots \centerdot \alpha_n$ on $T_{\text{poly}}(\mathbb{R}^d)$ are elements of the space $S^n(T_{\text{poly}}(\mathbb{R}^d)[1])$ of symmetric n-polyvector fields on $T_{\text{poly}}(\mathbb{R}^d)[1]$ (this means that $\alpha_2 \centerdot \alpha_1 = (-1)^{|\alpha_1||\alpha_2|}\alpha_1 \centerdot \alpha_2$). The manifold $T_{\text{poly}}(\mathbb{R}^d)[1]$ is equipped with the formal bilinear vector field $Q = Q_2$, defined with the help of the Schouten bracket $[\ ,\ ]_S$:

$$Q_2(\alpha_1 \centerdot \alpha_2) = (-1)^{(|\alpha_1|-1)|\alpha_2|}[\alpha_1, \alpha_2]_S.$$

Similarly, $D_{\text{poly}}(\mathbb{R}^d)[1]$ is equipped with the formal vector field

$$Q' = Q'_1 + Q'_2,$$

defined by

$$Q_1'(D_1) = -d_H D_1, \quad Q_2'(D_1 \centerdot D_2) = (-1)^{(|D_1|-1)|D_2|}[D_1, D_2]_G.$$

Here $[\ ,\ ]_G$ is the Gerstenhaber bracket and d_H denotes the usual Hochschild coboundary operator: if D is an m-differential operator,

$$\begin{aligned}&d_H D(f_1, \dots, f_{m+1})\\ &\quad = f_1 D(f_2, \dots, f_{m+1}) - D(f_1 f_2, \dots, f_{m+1}) + \dots + (-1)^m D(f_1, \dots, f_m) f_{m+1}\end{aligned}$$

A formality $\mathcal{F}$ is then given by a sequence of mappings

$$\mathcal{F}_n : S^n\big(T_{\text{poly}}(\mathbb{R}^d)[1]\big) \to D_{\text{poly}}(\mathbb{R}^d)[1],$$

homogeneous of degree 0 and satisfying the *formality equation*

$$\begin{aligned}d_H(\mathcal{F}_n)(\alpha_1 \centerdot \dots \centerdot \alpha_n) &= \frac{1}{2} \sum_{\substack{I \sqcup J = \{1,\dots,n\}\\ |I|\neq 0,\, |J|\neq 0}} \varepsilon_\alpha(I, J) Q_2'\big(\mathcal{F}_{|I|}(\alpha_I) \centerdot \mathcal{F}_{|J|}(\alpha_J)\big)\\ &\quad - \frac{1}{2} \sum_{k\neq \ell} \varepsilon_\alpha(k\ell, 1 \dots \widehat{k\ell} \dots n) \mathcal{F}_{n-1}\big(Q_2(\alpha_k \centerdot \alpha_\ell) \centerdot \alpha_1 \centerdot \dots \centerdot \widehat{\alpha_k \alpha_\ell} \centerdot \dots \centerdot \alpha_n\big).\end{aligned}$$

Here, if $I = \{i_1 < \dots < i_\ell\}$, the notation α_I means $\alpha_{i_1} \centerdot \dots \centerdot \alpha_{i_\ell}$.

We shall impose moreover the condition that $\mathcal{F}_1$ is the canonical mapping $\mathcal{F}_1^{(0)}$ from $T_{\text{poly}}(\mathbb{R}^d)$ to $D_{\text{poly}}(\mathbb{R}^d)$ defined by

$$\mathcal{F}_1^{(0)}(\xi_1 \wedge \dots \wedge \xi_n)(f_1, \dots, f_n) = \frac{1}{n!} \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} \xi_{\sigma(i)}(f_i),$$

for any vector field ξ_k and any function f_i.

Now choose a coordinate system (x_t) on $\mathbb{R}^d$. M. Kontsevich [2003] has built explicitly a formality $\mathcal{U}$ for $\mathbb{R}^d$, using families of graphs drawn on configuration spaces. A graph Γ has aerial and terrestrial vertices. The aerial vertices are labeled $p_1, \dots, p_n$ and are elements of the Poincaré half-plane

$$\mathcal{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}.$$

The terrestrial vertices $q_1 < \dots < q_m$ are on the real line. The edges of Γ are arrows starting from an aerial vertex and ending in a terrestrial or aerial vertex; there are no arrows of the form $\overrightarrow{p_i p_i}$ and no multiple arrows. If we fix a total ordering O on the edges of Γ, we get an oriented graph (Γ, O). We say that O is compatible if, for all i, the arrows starting from p_i precede those starting from p_{i+1}. We denote by $GO_{n,m}$ the set of oriented graphs (Γ, O) with n labeled aerial vertices and m labeled terrestrial vertices, and such that O is compatible.

Consider such an oriented graph $(\Gamma, O) \in GO_{n,m}$. Suppose there are k_i edges starting from the vertex p_i $(1 \le i \le n)$. Kontsevich [2003] defines a natural operator $B_{(\Gamma,O)}$ assigning an m-differential operator $B_{(\Gamma,O)}(\alpha_1 \otimes \cdots \otimes \alpha_n)$ to an n-uple $(\alpha_1, \dots, \alpha_n)$ of polyvector fields α_i. This operator vanishes unless, for each i, α_i belongs to $T^{k_i-1}_{\text{poly}}(\mathbb{R}^d)$ (α_i is a k_i-polyvector field). We first consider all the multi-indexes $(t_1, \dots, t_{|k|})$ with $|k| = \sum k_i$ and $1 \le t_r \le d$ for all $1 \le r \le |k|$. We denote by $\text{end}(a)$ the set of edges ending at the vertex a; if these edges are $e_{i_1}, \dots, e_{i_\ell}$, we let $\partial_{\text{end}(a)}$ be the operator

$$\partial_{\text{end}(a)} = \frac{\partial^l}{\partial x_{t_{i_1}} \dots \partial x_{t_{i_\ell}}}.$$

Then, we denote by $\text{strt}(p_i)$ the ordered set $e^i_{j_1} < \cdots < e^i_{j_{k_i}}$ of edges starting from p_i and, if α_i is a k_i-vector field, by $\alpha_i^{\text{strt}(p_i)}$ the following component of α_i:

$$\alpha_i^{\text{strt}(p_i)} = \alpha_i^{t_{j_1} \dots t_{j_{k_i}}}.$$

Finally, if α_i is a k_i-vector field for each i, we set

$$B_{(\Gamma,O)}(\alpha_1 \otimes \cdots \otimes \alpha_n)(f_1, \dots, f_m) = \sum_{1 \le t_1, \dots, t_{|k|} \le d} \prod_{i=1}^{n} \partial_{\text{end}(p_i)} \alpha_i^{\text{strt}(p_i)} \prod_{j=1}^{m} \partial_{\text{end}(q_j)} f_j.$$

$B_{(\Gamma,O)}$ will be called the graph operator associated with (Γ, O).

The explicit formality $\mathcal{U}$ of Kontsevich can now be written as a sum $\mathcal{U} = \sum_n \mathcal{U}_n$ with

$$\mathcal{U}_n = \sum_{m \ge 0} \sum_{(\Gamma,O) \in GO_{n,m}} w_{(\Gamma,O)} B_{(\Gamma,O)},$$

where the coefficient $w_{(\Gamma,O)}$, the *weight* of (Γ, O), is an integral on a compactified configuration space. To be precise, for $2n + m - 2 \ge 0$, let $\text{Conf}(n, m)$ be the space of $(n+m)$-tuples consisting of n distinct points p_i in $\mathcal{H}$ and m distinct points q_j on the real line $\partial\mathcal{H}$. Consider on $\text{Conf}(n, m)$ the action of the group G of transformations $z \mapsto az + b$ ($a > 0$ and b real), and form the quotient space

$$C_{n,m} = \text{Conf}(n, m)/G.$$

Kontsevich associates with each oriented graph (Γ, O) the form

$$\omega_{(\Gamma,O)} = \frac{1}{k!} \bigwedge_{i=1}^{n} \left(d\Phi_{e^i_1} \wedge \cdots \wedge d\Phi_{e^i_{k_i}} \right)$$

on $C_{n,m}$, where $\{e^i_1 < e^i_2 < \cdots < e^i_{k_i}\}$ denotes the ordered set $\mathrm{strt}(p_i)$ formed by the k_i edges starting from p_i, $k! := k_1! \dots k_n!$ and, if $e^i_\ell = \overrightarrow{p_i a}$,

$$\Phi_{e^i_\ell} = \Phi_{\overrightarrow{p_i a}} = \frac{1}{2\pi}\mathrm{Arg}\,\frac{a-p_i}{a-\bar{p}_i}.$$

The weight $w_{(\Gamma,O)}$ is then defined as the value of the integral $\omega_{(\Gamma,O)}$ on the connected component $C^+_{n,m}$ of $C_{n,m}$ for which $q_1 < \cdots < q_m$.

In this work, we are looking for graph formalities, that is, formalities on the space $\mathbb{R}^d$ of the form $\mathcal{F} = \sum_n \mathcal{F}_n$, where the $\mathcal{F}_n$ are homogeneous mappings (of degree 0) of the form

$$\mathcal{F}_n = \sum_{m\geq 0}\ \sum_{(\Gamma,O)\in GO_{n,m}} c_{(\Gamma,O)} B_{(\Gamma,O)},$$

with real coefficients $c_{(\Gamma,O)}$. We shall use the notation $\mathcal{F}_n = B_{\gamma_n}$, where γ_n is the linear combination

$$\gamma_n = \sum_{m\geq 0}\ \sum_{(\Gamma,O)\in GO_{n,m}} c_{(\Gamma,O)}(\Gamma, O).$$

Now assume we have found $\mathcal{F}_1, \dots, \mathcal{F}_{n-1}$ (with $\mathcal{F}_1 = \mathcal{F}^{(0)}_1 = \mathcal{U}_1$) such that the formality equation holds up to order $n-1$. The next term $\mathcal{F}_n$, if it exists, must be a solution of an equation

$$d_H \circ \mathcal{F}_n = E_n,$$

that is,

$$d_H(\mathcal{F}_n(\alpha_1 \cdot \dots \cdot \alpha_n)) = E_n(\alpha_1 \cdot \dots \cdot \alpha_n) = E_n(\alpha_{\{1,\dots,n\}}),$$

where $E_n(\alpha_{\{1,\dots,n\}})$ is a Hochschild cocycle. The Hochschild cohomology is localized in $T_{\mathrm{poly}}(\mathbb{R}^d)[1]$; more precisely, the total skewsymmetrization $\mathfrak{a} \circ E_n(\alpha_{\{1,\dots,n\}})$ of $E_n(\alpha_{\{1,\dots,n\}})$ is a polydifferential operator of order $1, \dots, 1$, that is, the image under $\mathcal{F}^{(0)}_1$ of a polyvector field. Moreover, there exists an operator A_n such that

$$E_n(\alpha_{\{1,\dots,n\}}) = (\mathfrak{a} \circ E_n + d_H \circ A_n)(\alpha_{\{1,\dots,n\}}).$$

Now put

$$\varphi_n = \mathcal{F}_1^{-1} \circ \mathfrak{a} \circ E_n,$$

that is,

$$\varphi_n(\alpha_{\{1,\dots,n\}}) = \mathcal{F}_1^{-1}\big(\mathfrak{a}(E_n(\alpha_{\{1,\dots,n\}}))\big);$$

then $\varphi_n : S^n\big(T_{\mathrm{poly}}(\mathbb{R}^d)[1]\big) \to T_{\mathrm{poly}}(\mathbb{R}^d)[1]$ is homogeneous of degree $|\varphi_n| = 1$.

In Section 2, we define the Chevalley coboundary operator ∂ on $T_{\mathrm{poly}}(\mathbb{R}^d)$. We show that the mapping φ_n described above is a Chevalley cocycle, and, if it is a coboundary ($\varphi_n = \partial\phi_{n-1}$), we can add to $\mathcal{F}_{n-1}$ a Hochschild coboundary so that

$\mathfrak{a}(E_n)$ vanishes and thus find a $\mathcal{F}_n$ for which the formality equation holds up to order n.

In Section 3, we establish a noteworthy property for the Kontsevich weights. For any graph Γ (with k_i edges starting from p_i), denote by Δ the purely aerial graph obtained by removing the legs $\overrightarrow{p_i q_j}$ and the feet q_j of Γ, and by ℓ_i the number of aerial edges starting from p_i. We prove that

$$\mathfrak{a}\Bigg(\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Gamma,O)} B_{(\Gamma,O)}\Bigg) = \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w_{(\Delta,O_\Delta)} \frac{1}{m!} \sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}} \frac{\ell!}{k!}\varepsilon(\Gamma) B_{(\Gamma,O)}.$$

Here $GO^{(1)}_{n,m}$ denotes the subspace of $GO_{n,m}$ formed by the oriented graphs having exactly one leg for each foot, $GO^{(0)}_n$ is the set of purely aerial oriented graphs (Δ, O_Δ) with n aerial vertices and O_Δ compatible and $\varepsilon(\Gamma)$ is an explicit sign depending only on Γ.

This property suggests that we study what we call K-graph formalities. A *K-graph formality* up to order n is a graph formality $\mathcal{F}$ at order $n-1$ such that $\varphi_n = \mathcal{F}_1^{-1} \circ \mathfrak{a} \circ E_n$ has the form

$$\varphi_n = \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} c_{(\Delta,O_\Delta)} C_{(\Delta,O_\Delta)}$$

with real coefficients $c_{(\Delta,O_\Delta)}$ and where

$$C_{(\Delta,O_\Delta)} = \sum_{m\geq 0} \frac{1}{m!} \sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}} \frac{\ell!}{k!}\varepsilon(\Gamma) B_{(\Gamma,O)}.$$

In Section 4 we give some simple expressions of our Chevalley coboundary operator. Then we restrict ourselves to K-graph formalities and study the Chevalley cohomology related to the question of building such formalities.

In Section 5 we show that the coboundary operator ∂ can be written directly on the aerial part of the graphs.

We devote Section 6 to explicit computations and applications. In particular, we prove the triviality of the cohomology for small values of n and give the restriction of the cohomology for linear formalities.

2. Chevalley cohomology and formalities

We start by defining a graded Chevalley cohomology in a general algebraic setting — that is, for cochains $C : S^n(\mathfrak{g}[1]) \to \mathfrak{M}[1]$, where $\mathfrak{g}$ is a graded Lie algebra and $\mathfrak{M}$ a graded $\mathfrak{g}$-module. In fact two Chevalley coboundary operators are naturally associated with the formality equation for $\mathbb{R}^d$. The first, ∂', is obtained by endowing $D_{\mathrm{poly}}(\mathbb{R}^d)$ with a $T_{\mathrm{poly}}(\mathbb{R}^d)$-graded module structure; cochains are mappings $C : S^n\big(T_{\mathrm{poly}}(\mathbb{R}^d)[1]\big) \to D_{\mathrm{poly}}(\mathbb{R}^d)[1]$. The other one, ∂, is obtained by considering $T_{\mathrm{poly}}(\mathbb{R}^d)$ as a graded module over itself; cochains are mappings $C : S^n\big(T_{\mathrm{poly}}(\mathbb{R}^d)[1]\big) \to T_{\mathrm{poly}}(\mathbb{R}^d)[1]$. Using both ∂ and ∂', we show that the obstructions to formalities can be interpreted as cocycles for ∂.

2.1. ***Chevalley cohomology.*** Let $(\mathfrak{g}, [\ ,\ ])$ be a graded Lie algebra and $\mathfrak{M}$ a graded module over $\mathfrak{g}$. For reasons of homogeneity, we prefer to work with $\mathfrak{g}[1]$ and $\mathfrak{M}[1]$. Thus, we replace $[\ ,\ ]$ and the action of $\mathfrak{g}$ on $\mathfrak{M}$ respectively by $[\ ,\ ]'$ and $[\ ,\ ]_{\mathfrak{M}}$, defined for homogeneous α, β in $\mathfrak{g}[1]$ of degrees $|\alpha|$, $|\beta|$ and for m in $\mathfrak{M}[1]$ of degree $|m|$ by

$$[\alpha, \beta]' = (-1)^{(|\alpha|+1)|\beta|}[\alpha, \beta],$$
$$[\alpha, m]_{\mathfrak{M}} = (-1)^{(|\alpha|+1)|m|}\alpha.m.$$

The space $C^n(\mathfrak{g}, \mathfrak{M})$ of n-cochains consists of mappings C from $S^n(\mathfrak{g}[1])$ to $\mathfrak{M}[1]$. The Chevalley coboundary ∂C of an n-cochain C, homogeneous of degree $|C|$, is the $(n+1)$-cochain defined by

$$\begin{aligned}\partial C&(\alpha_1 \cdot \ldots \cdot \alpha_{n+1})\\ &= \sum_{i=1}^{n+1} (-1)^{|C||\alpha_i|}\varepsilon_\alpha(i, 1 \ldots \hat{\imath} \ldots n+1)\big[\alpha_i, C(\alpha_1 \cdot \ldots \cdot \hat{\alpha_i} \cdot \ldots \cdot \alpha_{n+1})\big]_{\mathfrak{M}}\\ &\quad - \frac{1}{2}\sum_{i\neq j} \varepsilon_\alpha(ij, 1 \ldots \widehat{ij} \ldots, n+1)(-1)^{|C|}C\big([\alpha_i, \alpha_j]' \cdot \alpha_1 \ldots \widehat{\alpha_i\alpha_j} \cdot \ldots \cdot \alpha_{n+1}\big).\end{aligned}$$

Here the α_i are homogeneous elements of $\mathfrak{g}$, $|\alpha_i|$ denotes the degree of α_i in $\mathfrak{g}[1]$ and for any permutation σ of $\{1, \ldots, n\}$, $\varepsilon_\alpha(\sigma)$ is the sign of σ in the graded sense. We shall denote by $C^n_{[q]}(\mathfrak{g}, \mathfrak{M})$ the subspace of $C^n(\mathfrak{g}, \mathfrak{M})$ formed by the n-cochains of degree q and by $H^n_{[q]}(\mathfrak{g}, \mathfrak{M})$ the corresponding cohomology group. Note that ∂ sends $C^n_{[q]}(\mathfrak{g}, \mathfrak{M})$ into $C^{n+1}_{[q+1]}(\mathfrak{g}, \mathfrak{M})$.

Extending usual techniques to the graded case (See [Gammella 2001] for an explicit computation), it is possible to prove:

Lemma 2.1. *The operator ∂ is a cohomology operator*, *that is*, $\partial^2 = \partial \circ \partial = 0$.

We now return to the graded Lie algebras

$$\big(T_{\mathrm{poly}}(\mathbb{R}^d), [\ ,\ ]_S\big) \quad \text{and} \quad \big(D_{\mathrm{poly}}(\mathbb{R}^d), [\ ,\ ]_G\big),$$

where $[\ ,\ ]_S$ is the Schouten bracket and $[\ ,\ ]_G$ the Gerstenhaber bracket. Let us first make our conventions for these spaces and brackets precise.

Let α be a k-vector field and $\{e_i\}$ the canonical basis of $\mathbb{R}^d$. We put

$$\begin{aligned}\alpha &= \sum_{j_1,\dots,j_k} \alpha^{j_1\dots j_k} e_{j_1}\otimes\cdots\otimes e_{j_k} = \sum_{j_1<j_2<\cdots<j_k} \alpha^{j_1\dots j_k} e_{j_1}\wedge\cdots\wedge e_{j_k}\\ &= \frac{1}{k!}\sum_{j_1\dots j_k} \alpha^{j_1\dots j_k} e_{j_1}\wedge\cdots\wedge e_{j_k}.\end{aligned}$$

For any k_1-vector field α_1 and k_2-vector field α_2 (the degree of α_i is k_i-1 in $T_{\text{poly}}(\mathbb{R}^d)$), we define first a polyvector field $\alpha_1\bullet\alpha_2$ with components

$$\begin{aligned}\alpha_1\bullet\alpha_2^{i_1\cdots i_{k_1+k_2-1}} = \frac{1}{k_1!\,k_2!}\sum_{\sigma\in S_{k_1+k_2-1}}\Bigg(\varepsilon(\sigma)\sum_{\ell=1}^{k_1}(-1)^{\ell-1}&\\ \times\sum_{s=1}^{d}\alpha_1^{i_{\sigma(1)}\dots i_{\sigma(\ell-1)} s\, i_{\sigma(\ell)}\dots i_{\sigma(k_1-1)}}\partial_s\alpha_2^{i_{\sigma(k_1)}\dots i_{\sigma(k_1+k_2-1)}}\Bigg)&.\end{aligned}$$

Now, $[\alpha_1,\alpha_2]_S$ can be written as

$$[\alpha_1,\alpha_2]_S = (-1)^{k_2(k_1-1)}\alpha_1\bullet\alpha_2 - (-1)^{k_2-1}\alpha_2\bullet\alpha_1.$$

(This choice for the Schouten bracket is denoted $[\ ,\ ]'_S$ in [Arnal et al. 2002] and [Manchon and Torossian 2003].)

On the other hand, for any m_1-differential operator D_1 and any m_2-differential operator D_2 (the degree of D_i is m_i-1 in $D_{\text{poly}}(\mathbb{R}^d)$), we may write $[D_1,D_2]_G$ in the form

$$[D_1,D_2]_G = D_1\circ D_2 - (-1)^{(m_1-1)(m_2-1)}D_2\circ D_1,$$

where

$$\begin{aligned}&D_1\circ D_2(f_1,\dots,f_{m_1+m_2-1}) =\\ &\sum_{j=1}^{m_1}(-1)^{(m_2-1)(j-1)}D_1\big(f_1,\dots,f_{j-1},D_2(f_j,\dots,f_{j+m_2-1}),f_{j+m_2},\dots,f_{m_1+m_2-1}\big).\end{aligned}$$

Recall the canonical mapping $\mathcal{F}_1^{(0)}$ from $T_{\text{poly}}(\mathbb{R}^d)$ into $D_{\text{poly}}(\mathbb{R}^d)$: each k-vector field α can be viewed as a k-differential operator $\mathcal{F}_1^{(0)}(\alpha)$ of order $1,\dots,1$:

$$\big(\mathcal{F}_1^{(0)}(\alpha)\big)(f_1,\dots,f_k) = \langle\alpha, df_1\wedge\cdots\wedge df_k\rangle = \frac{1}{k!}\alpha^{i_1\cdots i_k}\partial_{i_1}f_1\dots\partial_{i_k}f_k.$$

Now consider the action of $T_{\text{poly}}(\mathbb{R}^d)$ given by

$$\alpha.D = \mathfrak{a}\circ\big[\mathcal{F}_1^{(0)}(\alpha), D\big]_G \qquad \text{for } \alpha\in T_{\text{poly}}(\mathbb{R}^d) \text{ and } D\in D_{\text{poly}}(\mathbb{R}^d),$$

where $\mathfrak{a}$ denotes the usual skewsymmetrization of differential operators and $[\ ,\ ]_G$ is the Gerstenhaber bracket. This action defines a $T_{\mathrm{poly}}(\mathbb{R}^d)$-graded module structure on $D_{\mathrm{poly}}(\mathbb{R}^d)$. Indeed, one can prove:

Proposition 2.2. *The following equalities hold for any D_1, D_2, D in $D_{\mathrm{poly}}(\mathbb{R}^d)$, any k_1-vector field α_1 and k_2-vector field α_2 in $T_{\mathrm{poly}}(\mathbb{R}^d)$:*

(i) $\mathfrak{a}\circ[D_1, D_2]_G = \mathfrak{a}\circ[D_1, \mathfrak{a}\circ D_2]_G$;

(ii) $\mathscr{F}_1^{(0)}([\alpha_1,\alpha_2]_S) = \mathfrak{a}\circ[\mathscr{F}_1^{(0)}(\alpha_1), \mathscr{F}_1^{(0)}(\alpha_2)]_G$;

(iii) $$\mathfrak{a}\circ\big[\mathscr{F}_1^{(0)}([\alpha_1,\alpha_2]_S), D\big]_G = \mathfrak{a}\circ\big[\mathscr{F}_1^{(0)}(\alpha_1), \mathfrak{a}\circ[\mathscr{F}_1^{(0)}(\alpha_2), D]_G\big]_G \\ -(-1)^{(k_1-1)(k_2-1)}\mathfrak{a}\circ\big[\mathscr{F}_1^{(0)}(\alpha_2), \mathfrak{a}\circ[\mathscr{F}_1^{(0)}(\alpha_1), D]_G\big]_G.$$

From (iii) *it follows that*

$$[\alpha_1,\alpha_2]_S.D = \alpha_1.(\alpha_2.D) - (-1)^{(k_1-1)(k_2-1)}\alpha_2.(\alpha_1.D),$$

and thus $D_{\mathrm{poly}}(\mathbb{R}^d)$ is a $T_{\mathrm{poly}}(\mathbb{R}^d)$-module.

Now endow $D_{\mathrm{poly}}(\mathbb{R}^d)$ with the $T_{\mathrm{poly}}(\mathbb{R}^d)$-graded structure described above. If $C : \bigwedge^n\big(T_{\mathrm{poly}}(\mathbb{R}^d)\big) = S^n\big(T_{\mathrm{poly}}(\mathbb{R}^d)[1]\big) \to D_{\mathrm{poly}}(\mathbb{R}^d)[1]$ is a homogeneous mapping of degree $|C|$, we can define its Chevalley coboundary $\partial' C$. The latter can be written using the vector fields Q and Q', associated respectively with $T_{\mathrm{poly}}(\mathbb{R}^d)$ and $D_{\mathrm{poly}}(\mathbb{R}^d)$:

$$\partial' C(\alpha_1\cdot\ldots\cdot\alpha_{n+1}) \\ = \sum_{i=1}^{n+1}(-1)^{|C||\alpha_i|}\varepsilon_\alpha(i, 1\ldots\hat{\imath}\ldots n+1)\mathfrak{a}\circ Q_2'\big(\mathscr{F}_1^{(0)}(\alpha_i)\cdot C(\alpha_1\cdot\ldots\hat{\alpha}_i\ldots\cdot\alpha_{n+1})\big) \\ -\frac{1}{2}\sum_{i\neq j}\varepsilon_\alpha(ij, 1\ldots\widehat{\imath\jmath}\ldots n+1)(-1)^{|C|}C\big(Q_2(\alpha_i\cdot\alpha_j)\cdot\alpha_1\ldots\widehat{\alpha_i\alpha_j}\ldots\cdot\alpha_{n+1}\big).$$

To simplify the writing, we will sometimes write α_i instead of $\mathscr{F}_1^{(0)}(\alpha_i)$.

At the same time, considering $T_{\mathrm{poly}}(\mathbb{R}^d)$ as a graded module over itself, one can define the Chevalley cohomology for $T_{\mathrm{poly}}(\mathbb{R}^d)$. If $C : S^n\big(T_{\mathrm{poly}}(\mathbb{R}^d)[1]\big) \to T_{\mathrm{poly}}(\mathbb{R}^d)[1]$ is an n-cochain, homogeneous of degree $|C|$, its coboundary ∂C is

$$\partial C(\alpha_1\cdot\ldots\cdot\alpha_{n+1}) \\ = \sum_{i=1}^{n+1}(-1)^{|C||\alpha_i|}\varepsilon_\alpha(i, 1\ldots\hat{\imath}\ldots n+1)Q_2\big(\alpha_i\cdot C(\alpha_1\cdot\ldots\hat{\alpha}_i\ldots\cdot\alpha_{n+1})\big) \\ -\frac{1}{2}\sum_{i\neq j}\varepsilon_\alpha(ij, 1\ldots\widehat{\imath\jmath}\ldots n+1)(-1)^{|C|}C\big(Q_2(\alpha_i\cdot\alpha_j)\cdot\alpha_1\ldots\widehat{\alpha_i\alpha_j}\ldots\cdot\alpha_{n+1}\big).$$

Remark. For any $\varphi : S^n\big(T_{\text{poly}}(\mathbb{R}^d)[1]\big) \to T_{\text{poly}}(\mathbb{R}^d)[1]$, we have

$$\partial'(\mathcal{F}_1^{(0)} \circ \varphi) = \mathcal{F}_1^{(0)} \circ \partial\varphi.$$

2.2. ***Obstruction to formalities.*** The two Chevalley coboundary operators ∂ and ∂' enable us to reformulate the formality equation. Indeed, suppose we want to construct a formality $\mathcal{F}$ from $T_{\text{poly}}(\mathbb{R}^d)$ to $D_{\text{poly}}(\mathbb{R}^d)$. We thus need to solve recursively the formality equation (see [Kontsevich 1997; Arnal et al. 2002] for notations)

$$\begin{aligned} d_H(\mathcal{F}_n)(\alpha_1 \ldots\ldots \alpha_n) = \frac{1}{2} &\sum_{\substack{I \sqcup J = \{1,\ldots,n\} \\ |I|\geq 1, |J|\geq 1}} \varepsilon_\alpha(I, J)\, Q_2'\big(\mathcal{F}_{|I|}(\alpha_I) \,.\, \mathcal{F}_{|J|}(\alpha_J)\big) \\ &- \frac{1}{2}\sum_{k\neq\ell} \varepsilon_\alpha(k\ell, 1\ldots\widehat{k\ell}\ldots n)\mathcal{F}_{n-1}\big(Q_2(\alpha_k \,.\, \alpha_\ell) \,.\, \alpha_1 \ldots\ldots \widehat{\alpha_k \alpha_\ell}\partial \ldots . \alpha_n\big), \end{aligned}$$

where d_H is the Hochschild coboundary operator.

Now impose the condition that the first component $\mathcal{F}_1$ be $\mathcal{F}_1^{(0)}$. Assume there are mappings $\mathcal{F}_2, \ldots, \mathcal{F}_{n-1}$, homogeneous of degree 0, and satisfying the formality equation up to order $n-1$. Denote by E_n the right-hand side of the equation at the order n. Then E_n is a Hochschild cocycle: $d_H E_n = 0$ (see [Arnal and Masmoudi 2002] for instance). Thus

$$E_n = \mathfrak{a} \circ E_n + d_H C,$$

where $\mathfrak{a} \circ E_n$ is a differential operator of order $1, \ldots, 1$ and E_n is a coboundary if and only if $\mathfrak{a} \circ E_n = 0$. But

$$\mathfrak{a} \circ E_n(\alpha_1 \ldots\ldots \alpha_n) = \partial'\mathfrak{a}\mathcal{F}_{n-1}(\alpha_1 \ldots\ldots \alpha_n) + \mathfrak{a}R_n(\alpha_1 \ldots\ldots \alpha_n),$$

where

$$R_n(\alpha_1 \ldots\ldots \alpha_n) = \frac{1}{2}\sum_{I\sqcup J=\{1,\ldots,n\}, |I|\geq 2, |J|\geq 2} \varepsilon_\alpha(I, J)\, Q_2'\Big(\mathcal{F}_{|I|}(\alpha_I) \,.\, \mathcal{F}_{|J|}(\alpha_J)\Big).$$

It follows directly from this expression that R_n and $\mathfrak{a} \circ R_n$ both have degree 1: $|R_n| = |\mathfrak{a} \circ R_n| = 1$. Moreover,

Theorem 2.3. *The skewsymmetrization $\mathfrak{a} \circ E_n$ of E_n can be identified through the inverse mapping of $\mathcal{F}_1$ with a ∂-cocycle. If this cocycle is exact, we can find $\mathcal{F}'_{n-1}$ and $\mathcal{F}'_n$, homogeneous of degree 0, such that $\mathcal{F}_2, \ldots, \mathcal{F}_{n-2}, \mathcal{F}'_{n-1}, \mathcal{F}'_n$ satisfy the formality equation up to order n.*

Proof. The proof proceeds in three steps.

Step 1. First we check that $\mathfrak{a}\circ R_n$ is a cocycle for ∂':

$$\begin{aligned}
&\partial'\mathfrak{a}R_n(\alpha_1\,.\,\dots\,.\,\alpha_{n+1})\\
&\quad=\sum_{i=1}^{n+1}(-1)^{|\alpha_i|}\varepsilon_\alpha(i,1\dots\hat{\imath}\dots n+1)\mathfrak{a}Q_2'\big(\alpha_i\,.\,\mathfrak{a}R_n(\alpha_1\,.\,\dots\,.\,\widehat{\alpha_i}\,.\,\dots\,.\,\alpha_{n+1})\big)\\
&\qquad+\frac{1}{2}\sum_{i\neq j}\varepsilon_\alpha(ij,1\dots\widehat{\imath\jmath}\dots n+1)\mathfrak{a}R_n\big(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_1\,.\,\dots\,.\,\widehat{\alpha_i\alpha_j}\,.\,\dots\,.\,\alpha_{n+1}\big)\\
&\quad=\frac{1}{2}\sum_{i=1}^{n+1}\Bigg((-1)^{|\alpha_i|}\varepsilon_\alpha(i,1\dots\hat{\imath}\dots n+1)\\
&\qquad\qquad\times\sum_{\substack{I\sqcup J=\{1\dots\hat{\imath}\dots n+1\}\\|I|\geq2,|J|\geq2}}\varepsilon_{\alpha'}(I,J)\mathfrak{a}Q_2'\big(\alpha_i\,.\,\mathfrak{a}Q_2'(\mathcal{F}_{|I|}(\alpha_I)\,.\,\mathcal{F}_{|J|}(\alpha_J))\big)\Bigg)\\
&\qquad+\frac{1}{4}\sum_{i\neq j}\Bigg(\varepsilon_\alpha(ij,1\dots\widehat{\imath\jmath}\dots n+1)\\
&\qquad\qquad\times\sum_{\substack{I\sqcup J=\{0,1\dots\widehat{\imath\jmath}\dots n+1\}\\|I|\geq2,|J|\geq2}}\varepsilon_{\alpha''}(I,J)\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(\alpha_I)\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)\Bigg)\\
&\quad=\tfrac{1}{2}(\mathrm{I})+\tfrac{1}{4}(\mathrm{II}),
\end{aligned}$$

where we have set $\alpha_0:=Q_2(\alpha_i\,.\,\alpha_j)$, $\varepsilon_{\alpha'}:=\varepsilon_{\alpha\setminus\{\alpha_i\}}$ and $\varepsilon_{\alpha''}:=\varepsilon_{(\alpha\cup\{\alpha_0\})\setminus\{\alpha_i,\alpha_j\}}$.

The term (I) above equals

$$\sum_{i=1}^{n+1}\sum_{\substack{I\sqcup J=\{1\dots\hat{\imath}\dots n+1\}\\|I|\geq2,|J|\geq2}}(-1)^{(|\alpha_I|+|\alpha_J|)|\alpha_i|}\varepsilon_\alpha(i,I,J)\mathfrak{a}Q_2'\big(\mathfrak{a}Q_2'(\mathcal{F}_{|I|}(\alpha_I)\,.\,\mathcal{F}_{|J|}(\alpha_J))\,.\,\alpha_i\big).$$

By Proposition 2.2, $\mathfrak{a}Q_2'$ satisfies the graded Jacobi identity; thus (I) equals

$$\begin{aligned}
&-\sum_{i=1}^{n+1}\sum_{\substack{I\sqcup J=\{1\dots\hat{\imath}\dots n+1\}\\|I|\geq2,|J|\geq2}}(-1)^{|\alpha_J|(|\alpha_I|+|\alpha_i|)}\varepsilon_\alpha(i,I,J)\mathfrak{a}Q_2'\big(\mathfrak{a}Q_2'(\mathcal{F}_{|J|}(\alpha_J)\,.\,\alpha_i)\,.\,\mathcal{F}_{|I|}(\alpha_I)\big)\\
&\qquad\qquad-\sum_{\substack{I\sqcup J=\{1\dots\hat{\imath}\dots n+1\}\\|I|\geq2,|J|\geq2}}\varepsilon_\alpha(i,I,J)\mathfrak{a}Q_2'\big(\mathfrak{a}Q_2'(\alpha_i\,.\,\mathcal{F}_{|I|}(\alpha_I))\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)\\
&\qquad=-2\sum_{\substack{I\sqcup J=\{1\dots\hat{\imath}\dots n+1\}\\|I|\geq2,|J|\geq2}}\varepsilon_\alpha(i,I,J)\mathfrak{a}Q_2'\big(\mathfrak{a}Q_2'(\alpha_i\,.\,\mathcal{F}_{|I|}(\alpha_I))\,.\,\mathcal{F}_{|J|}(\alpha_J)\big).
\end{aligned}$$

Similarly, the second term, (II), is equal to

$$\sum_{i\neq j}\varepsilon_\alpha(ij,1\ldots\widehat{ij}\ldots n+1)\sum_{\substack{I\sqcup J=\{0,1\ldots\widehat{ij}\ldots n+1\}\\|I|\geq 2,|J|\geq 2}}\varepsilon_{\alpha''}(I,J)\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(\alpha_I)\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)$$

$$=\sum_{i\neq j}\sum_{\substack{I=I_1\sqcup\{0\}\\I_1\sqcup J=\{1\ldots\widehat{ij}\ldots n+1\}}}\varepsilon_\alpha(ij,I_1,J)\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_{I_1})\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)$$

$$+\sum_{i\neq j}\Bigg(\sum_{\substack{J=J_1\sqcup\{0\}\\I\sqcup J_1=\{1\ldots\widehat{ij}\ldots n+1\}}}\varepsilon_\alpha(ij,1\ldots\widehat{ij}\ldots n+1)\varepsilon_{\alpha''}(I,\{0\},J_1)\times\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(\alpha_I)\,.\,\mathcal{F}_{|J|}(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_{J_1})\big)\Bigg)$$

$$=\sum_{i\neq j}\Bigg(\sum_{\substack{I=I_1\sqcup\{0\}\\I_1\sqcup J=\{1\ldots\widehat{ij}\ldots n+1\}}}\varepsilon_\alpha(ij,I_1,J)\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_{I_1})\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)$$

$$+\sum_{\substack{J=J_1\sqcup\{0\}\\I\sqcup J_1=\{1\ldots\widehat{ij}\ldots n+1\}}}\varepsilon_\alpha(ij,I,J_1)(-1)^{(|\alpha_i|+|\alpha_j|+1)|\alpha_J|}\times\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(\alpha_I)\,.\,\mathcal{F}_{|J|}(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_{J_1})\big)\Bigg)$$

$$=2\sum_{i\neq j}\sum_{\substack{I=I_1\sqcup\{0\}\\I_1\sqcup J=\{1\ldots\widehat{ij}\ldots n+1\}}}\varepsilon_\alpha(ij,I_1,J)\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_{I_1})\,.\,\mathcal{F}_{|J|}(\alpha_J)\big).$$

Putting (I) and (II) together, we get

$$\partial'(\mathfrak{a}R_n)(\alpha_1\,.\,\ldots\,.\,\alpha_{n+1})=\tfrac{1}{2}(\mathrm{I})+\tfrac{1}{4}(\mathrm{II})$$

$$=\sum_{\substack{I'\sqcup J=\{1\ldots n+1\}\\|J|\geq 2,|I'|\geq 3}}\varepsilon_\alpha(I',J)\Bigg(\sum_{\substack{i\in I'\\(I'=I\sqcup\{i\})}}\varepsilon_{\alpha_{\{i\}\sqcup I}}(i,I)\mathfrak{a}Q_2'\big(\mathfrak{a}Q_2'(\alpha_i\,.\,\mathcal{F}_{|I|}(\alpha_I))\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)$$

$$+\frac{1}{2}\sum_{\substack{i\neq j\in I',I'=I_1\sqcup\{ij\}\\I=I_1\sqcup\{0\}}}\varepsilon_{\alpha_{\{ij\}\cup I_1}}(ij,I_1)\mathfrak{a}Q_2'\big(\mathcal{F}_{|I|}(Q_2(\alpha_i\,.\,\alpha_j)\,.\,\alpha_{I_1})\,.\,\mathcal{F}_{|J|}(\alpha_J)\big)\Bigg).$$

Now, Proposition 2.2 and the definition of ∂' yield

(∗)

$$\partial'(\mathfrak{a}R_n)(\alpha_1\,.\,\ldots\,.\,\alpha_{n+1})=-\sum_{\substack{I'\sqcup J=\{1\ldots n+1\}\\|J|\geq 2,\ |I'|\geq 3}}\varepsilon_\alpha(I,J)\mathfrak{a}Q_2'\big(\partial'\mathfrak{a}\mathcal{F}_{|I'|-1}(\alpha_{I'})\,.\,\mathcal{F}_{|J|}(\alpha_J)\big).$$

On the other hand, since the formality equation holds up to order $n-1$, we have

$$\partial'\mathfrak{a}\mathcal{F}_{p-1}+\mathfrak{a}R_p=\mathfrak{a}(E_p)=\mathfrak{a}(d_H(\mathcal{F}_p))=0\quad\text{for } p\leq n-1.$$

But $|I'| \leq n-1$ for all I' in the expression $(*)$; thus

$$-\partial' \mathfrak{a}\mathscr{F}_{|I'|-1}(\alpha_{I'}) = \mathfrak{a}R_{|I'|}(\alpha_{I'}) = \frac{1}{2} \sum_{\substack{S \sqcup T = I' \\ |S|\geq 2, |T|\geq 2}} \varepsilon_{\alpha_{S\sqcup T}}(S,T)\mathfrak{a}Q'_2(\mathscr{F}_{|S|}(\alpha_S) \,.\, \mathscr{F}_{|T|}(\alpha_T)).$$

Finally, $(*)$ becomes

$$\begin{aligned}
&\partial'(\mathfrak{a}R_n)(\alpha_1 \,.\, \dots \,.\, \alpha_{n+1}) \\
&\quad = \frac{1}{2} \sum_{\substack{S\sqcup T\sqcup J=\{1\dots n+1\} \\ |S|\geq 2, |T|\geq 2, |J|\geq 2}} \varepsilon_\alpha(S\cup T, J)\varepsilon_{\alpha_{S\sqcup T}}(S,T) \\
&\qquad\qquad\qquad\qquad \times \mathfrak{a}Q'_2\big(\mathfrak{a}Q'_2(\mathscr{F}_{|S|}(\alpha_S) \,.\, \mathscr{F}_{|T|}(\alpha_T)) \,.\, \mathscr{F}_{|J|}(\alpha_J)\big) \\
&\quad = \frac{1}{2} \sum_{\substack{S\sqcup T\sqcup J=\{1\dots n+1\} \\ |S|\geq 2, |T|\geq 2, |J|\geq 2}} \varepsilon_\alpha(S,T,J)\mathfrak{a}Q'_2\big(\mathfrak{a}Q'_2(\mathscr{F}_{|S|}(\alpha_S) \,.\, \mathscr{F}_{|T|}(\alpha_T)) \,.\, \mathscr{F}_{|J|}(\alpha_J)\big).
\end{aligned}$$

Thanks to the Jacobi identity, the quantity on the last line vanishes. Hence $\partial'(\mathfrak{a}R_n)$ and $\partial'(\mathfrak{a}E_n)$ both vanish.

Step 2. Put

$$\varphi_n = \mathscr{F}_1^{-1} \circ \mathfrak{a} \circ E_n.$$

Since

$$\partial'(\mathfrak{a} \circ E_n) = \partial'\mathscr{F}_1(\varphi_n) = \mathscr{F}_1(\partial\varphi_n) = 0,$$

φ_n is a cocycle for ∂.

Step 3. Assume that $\varphi_n = \partial\phi_{n-1}$, where $\phi_{n-1} : S^{n-1}\big(T_{\mathrm{poly}}(\mathbb{R}^d)[1]\big) \to T_{\mathrm{poly}}(\mathbb{R}^d)[1]$. Of course, $d_H\mathscr{F}_1(\phi_{n-1}) = 0$. Therefore, the mappings $\mathscr{F}'_2 = \mathscr{F}_2, \dots, \mathscr{F}'_{n-2} = \mathscr{F}_{n-2}$, $\mathscr{F}'_{n-1} = \mathscr{F}_{n-1} - \mathscr{F}_1 \circ \phi_{n-1}$ satisfy the formality equation up to order $n-1$. Moreover, the Hochschild cocycle E'_n corresponding to these $\mathscr{F}'_p$ satisfies

$$\mathfrak{a} \circ E'_n = \mathfrak{a} \circ E_n - \partial'(\mathscr{F}_1 \circ \phi_{n-1}) = \mathfrak{a} \circ E_n - \mathscr{F}_1(\partial\phi_{n-1}) = 0.$$

We are now able to find $\mathscr{F}'_n$ such that $E'_n = d_H\mathscr{F}'_n$. This ends the proof. □

3. Skewsymmetrization

The aim of this section is to prove a noteworthy property of Kontsevich's weights and the definition of K-graph formalities.

3.1. ***Skewsymmetrization and 1-graphs.*** Consider an m-differential operator D on $\mathbb{R}^d$, vanishing on constants. We can decompose D as

$$D = D^{(1)} + D^{(>1)},$$

where $D^{(1)}$ has order 1 in each of its arguments and $D^{(>1)}$ has order greater than 1 in at least one of its arguments. The skewsymmetrization $\mathfrak{a}(D)$ of D, defined by

$$\mathfrak{a}(D)(f_1, \dots, f_m) = \frac{1}{m!} \sum_{\sigma \in S_m} \varepsilon(\sigma) D\big(f_{\sigma^{-1}(1)}, \dots, f_{\sigma^{-1}(m)}\big),$$

satisfies $\mathfrak{a}(D) = \mathfrak{a}(D^{(1)}) + \mathfrak{a}(D^{(>1)})$, and therefore

$$\mathfrak{a}(D)^{(1)} = \mathfrak{a}(D^{(1)}).$$

We assume D is defined with the help of graphs:

$$D_{(\alpha_1,\dots,\alpha_n)} = \sum_{\Gamma} c_\Gamma B_\Gamma(\alpha_1 \otimes \cdots \otimes \alpha_n),$$

where the c_Γ are real. To compute $a(D)^{(1)}$, we need only consider

$$D^{(1)}_{(\alpha_1,\dots,\alpha_n)} = \sum_{\Gamma \in G^{(1)}} c_\Gamma B_\Gamma(\alpha_1 \otimes \cdots \otimes \alpha_n),$$

where $G^{(1)}$ denotes the family of 1-graphs, that is, graphs having exactly one leg for each foot.

However, as in [Kontsevich 2003], to define B_Γ we need to choose a total ordering O on the set $E(\Gamma)$ of edges of Γ. To be precise, we first choose a labeling on the aerial vertices of Γ, say $p_1, \dots, p_n$. Then we put away the arrows starting from p_1, from p_2, and so on, and finally from p_n. We get a total ordering of $E(\Gamma)$ compatible with the ordering $p_1 < p_2 < \cdots < p_n$ in the sense that the arrows starting from p_i precede those starting from p_{i+1}.

From now on, we denote by $GO_{n,m}$ the set of oriented graphs (Γ, O) with n labeled aerial vertices, m labeled terrestrial vertices and compatible ordering O, and by $GO^{(1)}_{n,m}$ the subset of $GO_{n,m}$ formed by the oriented 1-graphs. Our earlier notation $\sum c_\Gamma B_\Gamma$ actually means

$$\sum_\Gamma c_\Gamma B_\Gamma = \sum_{(\Gamma,O)\in GO_{n,m}} c_{(\Gamma,O)} B_{(\Gamma,O)} \quad \text{and} \quad \sum_{\Gamma\in G^{(1)}} c_\Gamma B_\Gamma = \sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} c_{(\Gamma,O)} B_{(\Gamma,O)}.$$

3.2. *A noteworthy property of Kontsevich weights.*

Kontsevich weights. Let (Γ, O) be an oriented graph in $GO^{(1)}_{n,m}$ with aerial vertices $p_1 < \cdots < p_n$. We denote by k_i the number of edges starting from p_i, by U_i the ordered set of legs starting from p_i, and by V_i the ordered set of aerial edges starting from the same point. Let ℓ_i be the number of elements in V_i, U_i has $m_i = k_i - \ell_i$ elements. By the definition of $GO^{(1)}_{n,m}$, the number of legs is exactly the number of terrestrial vertices; that is, $m = \sum_{i=1}^n m_i$.

Given (Γ, O), it will be helpful to consider the permutation s_O defined by

$$s_O : E(\Gamma) \mapsto V_1 \cup \cdots \cup V_n \cup U_1 \cup \cdots \cup U_n.$$

After this permutation we get a new (and no longer compatible) ordering O' on $E(\Gamma)$ such that all the legs are put at the end, and we can define a permutation τ_O of the legs of (Γ, O') by putting first the leg ending at q_1, then the leg ending at q_2, and so on, with the the leg ending at q_m last. We extend the permutation τ_O to $V_1 \cup \ldots V_n \cup U_1 \cup \cdots \cup U_n$ by setting $\tau_O(v) = v$ for all v in $\bigcup V_i$. Finally, note Δ the aerial graph obtained from Γ by cutting the legs and the feet and by O_Δ the (compatible) ordering on Δ induced by O.

Let $GO_n^{(0)}$ be the set of oriented, purely aerial graphs (Δ, O_Δ) with n vertices.

With these notations, the Kontsevich weight associated with (Γ, O) can be written as

$$w_{(\Gamma,O)} = \frac{1}{k!}\varepsilon(s_O)\varepsilon(\tau_O) \int_{C^+_{n,0}} \bigwedge_{r=1}^{|\ell|} d\Phi_{e_r^\Delta} \int_{\substack{q_1<\cdots<q_m \text{ oriented} \\ \text{by } dq_1\wedge\cdots\wedge dq_m}} \bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i_j}q_j}},$$

where $k! = k_1! \ldots k_n!$, $|\ell| := \sum \ell_i$, $V_1 \cup \cdots \cup V_n := \{e_1^\Delta < \cdots < e_{|\ell|}^\Delta\}$ and i_j stands for the unique index i such that the leg arriving on q_j is exactly $\overrightarrow{p_i q_j}$.

The Kontsevich weight of (Δ, O_Δ) is just

$$w_{(\Delta,O_\Delta)} = \frac{1}{\ell!} \int_{C^+_{n,0}} \bigwedge_{r=1}^{|\ell|} d\Phi_{e_r^\Delta},$$

$(\ell! = \ell_1! \ldots \ell_n!)$. Thus

$$w_{(\Gamma,O)} = w_{(\Delta,O_\Delta)} \frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O) \int_{\substack{q_1<\cdots<q_m \text{ oriented} \\ \text{by } dq_1\wedge\cdots\wedge dq_m}} \bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i_j}q_j}}.$$

The S_m action on $GO^{(1)}_{n,m}$. Let σ be an element in the permutation group S_m. With any graph (Γ, O) in $GO^{(1)}_{n,m}$, we associate a new graph $(\sigma(\Gamma), \sigma(O))$. We keep for $\sigma(\Gamma)$ the vertices of Γ. But, if $E(\Gamma) = \{e_1 < \cdots < e_{|k|}\}$, we put $E(\sigma(\Gamma)) = \{e'_1 < \cdots < e'_{|k|}\}$, where $e'_r := e_r$ if e_r is an aerial edge and $e'_r := \overrightarrow{p_i q_{\sigma(j)}}$ if $e_r = \overrightarrow{p_i q_j}$ is a leg (see Figure 1). In this way we get a free action of S_m on $GO^{(1)}_{n,m}$.

Lemma 3.1. *For all σ in S_m and all (Γ, O) in $GO^{(1)}_{n,m}$,*

$$B_{(\sigma(\Gamma),\sigma(O))}(\alpha)(f_1, \ldots, f_m) = B_{(\Gamma,O)}(\alpha)(f_{\sigma(1)}, \ldots, f_{\sigma(m)}) \quad f_i \in C^\infty(\mathbb{R}^d).$$

Proof. Let r_j be the label of the leg arriving on q_j in (Γ, O). In $(\sigma(\Gamma), \sigma(O))$, this leg has the same label r_j, but it ends at $q_{\sigma(j)}$. The aerial edges are kept unchanged. The result follows easily. ∎

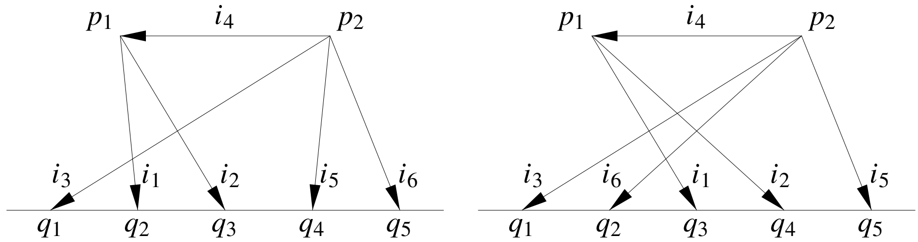


Figure 1. Left: (Γ, O). Right: $(\sigma(\Gamma), \sigma(O))$ with $\sigma = (2345)$.

Lemma 3.2. *Let σ be in S_m and (Γ, O) in $GO^{(1)}_{n,m}$. Then*

$$\varepsilon(s_{\sigma(O)}) = \varepsilon(s_O) \quad \textit{and} \quad \varepsilon(\tau_{\sigma(O)}) = \varepsilon(\sigma)\varepsilon(\tau_O).$$

Proof. When building $(\sigma(\Gamma), \sigma(O))$, we get a bijective mapping from $E(\Gamma)$ to $E(\sigma(\Gamma))$, say $\tilde{\sigma}$. In fact, $s_{\sigma(O)} = \tilde{\sigma} \circ s_O \circ \tilde{\sigma}^{-1}$. Thus $\varepsilon(s_{\sigma(O)}) = \varepsilon(s_O)$.

Now let $q_{a^i_1}, \dots, q_{a^i_{m_i}}$ be the feet of the legs starting from p_i. By definition, τ_O is the permutation

$$\overrightarrow{p_1 q_{a^1_1}}, \overrightarrow{p_1 q_{a^1_2}}, \dots, \overrightarrow{p_n q_{a^n_{m_n}}} \mapsto \overrightarrow{p_{i_1} q_1}, \dots, \overrightarrow{p_{i_m} q_m}.$$

We may write

$$\tau_O^{-1} : (1, \dots, m) \mapsto (a^1_1, \dots, a^n_{m_n}).$$

By the definition of $(\sigma(\Gamma), \sigma(O))$,

$$\tau_{\sigma(O)}^{-1} : (1, \dots, m) \mapsto \big(\sigma(a^1_1), \dots, \sigma(a^n_{m_n})\big).$$

Thus $\tau_{\sigma(O)}^{-1} \circ \tau_O = \sigma$. The result follows. □

A noteworthy property.

Proposition 3.3. *We keep our notations. In particular, for any (Γ, O) in $GO^{(1)}_{n,m}$ and any (Δ, O_Δ) in $GO^{(0)}_n$, we denote by $w_{(\Gamma,O)}$ and $w_{(\Delta,O_\Delta)}$ the corresponding weights. Then*

$$\mathfrak{a}\Bigg(\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Gamma,O)} B_{(\Gamma,O)}(\alpha)\Bigg) = \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w_{(\Delta,O_\Delta)} \sum_{m\geq 0} \frac{1}{m!} \sum_{\substack{(\Gamma,O)\supset(\Delta,O_\Delta)\\ (\Gamma,O)\in GO^{(1)}_{n,m}}} \frac{\ell!}{k!} \varepsilon(s_O)\varepsilon(\tau_O) B_{(\Gamma,O)}(\alpha).$$

Proof. Skewsymmetrizing and using Lemma 3.1, we get

$$\begin{aligned}
\mathfrak{a}\Bigg(\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Gamma,O)} B_{(\Gamma,O)}(\alpha)\Bigg)&(f_1\otimes\cdots\otimes f_m)\\
&= \frac{1}{m!}\sum_{\sigma\in S_m}\varepsilon(\sigma)\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Gamma,O)} B_{(\Gamma,O)}(\alpha)(f_{\sigma^{-1}(1)}\otimes\cdots\otimes f_{\sigma^{-1}(m)})\\
&= \frac{1}{m!}\sum_{\sigma\in S_m}\varepsilon(\sigma)\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Gamma,O)} B_{(\sigma^{-1}(\Gamma),\sigma^{-1}(O))}(\alpha)(f_1\otimes\cdots\otimes f_m)\\
&= \frac{1}{m!}\sum_{\sigma\in S_m}\varepsilon(\sigma)\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\sigma(\Gamma),\sigma(O))} B_{(\Gamma,O)}(\alpha)(f_1\otimes\cdots\otimes f_m).
\end{aligned}$$

By definition,

$$w_{(\sigma(\Gamma),\sigma(O))} = \varepsilon(s_{\sigma(O)})\varepsilon(\tau_{\sigma(O)})\frac{\ell!}{k!}\int_{C^+_{n,0}}\bigwedge_{r=1}^{|\ell|} d\Phi_{e^\Delta_r}\int_{\substack{q_1<\cdots<q_m \text{ oriented}\\ \text{by } dq_1\wedge\cdots\wedge dq_m}}\bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i'_j}q_j}},$$

where i'_j stands for the unique index i' such that the leg arriving on q_j is exactly $\overrightarrow{p'_i q_j}$. Now $\bigwedge_{j=1}^m d\Phi_{\overrightarrow{p_{i'_j}q_j}} = \varepsilon(\sigma)\bigwedge_{j=1}^m d\Phi_{\overrightarrow{p_{i_j}q_{\sigma(j)}}}$; then, by Lemma 3.2,

$$w_{(\sigma(\Gamma),\sigma(O))} = \frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)\int_{C^+_{n,0}}\bigwedge_{r=1}^{|\ell|} d\Phi_{e^\Delta_r}\int_{\substack{q_1<\cdots<q_m \text{ oriented}\\ \text{by } dq_1\wedge\cdots\wedge dq_m}}\bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i_j}q_{\sigma(j)}}}.$$

With the new variables $q'_j = q_{\sigma(j)}$, we get

$$w_{(\sigma(\Gamma),\sigma(O))} = \frac{\ell!}{k!} w_{(\Delta,O_\Delta)}\varepsilon(s_O)\varepsilon(\tau_O)\int_{D^\sigma}\bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i_j}q'_j}}$$

where D^σ is the domain $q'_{\sigma^{-1}(1)} < \cdots < q'_{\sigma^{-1}(m)}$ oriented by $dq'_1\wedge\cdots\wedge dq'_m$. Thus

$$\begin{aligned}
&\mathfrak{a}\Bigg(\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Gamma,O)} B_{(\Gamma,O)}(\alpha)\Bigg)\\
&= \frac{1}{m!}\sum_{\sigma\in S_m}\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w_{(\Delta,O_\Delta)}\frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)\int_{D^\sigma}\bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i_j}q'_j}} B_{(\Gamma,O)}(\alpha)\\
&= \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w_{(\Delta,O_\Delta)}\frac{1}{m!}\sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}}\frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)\Bigg(\sum_{\sigma\in S_m}\int_{D^\sigma}\bigwedge_{j=1}^{m} d\Phi_{\overrightarrow{p_{i_j}q'_j}}\Bigg) B_{(\Gamma,O)}(\alpha)\\
&= \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w_{(\Delta,O_\Delta)}\frac{1}{m!}\sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}}\frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O) B_{(\Gamma,O)}(\alpha). \qquad \square
\end{aligned}$$

3.3. ***K-graph formalities.*** Consider the explicit Kontsevich formality $\mathcal{U} = \sum_n \mathcal{U}_n$ on $\mathbb{R}^d$. If (Γ, O) is an oriented graph with O not compatible, we write, as in [Arnal et al. 2002],

$$B_{(\Gamma,O)} = \varepsilon(\sigma_{(O,O_0)}) B_{(\Gamma,O_0)},$$

where O_0 is any compatible orientation on Γ and $\sigma_{(O,O_0)}$ stands for the permutation of $E(\Gamma)$ obtained by changing (Γ, O) into (Γ, O_0). We also put

$$\omega'_{(\Gamma,O)} = \frac{k!}{|k|!}\omega_{(\Gamma,O)} \quad \text{and} \quad w'_{(\Gamma,O)} = \int_{C^+_{n,m}} \omega'_{(\Gamma,O)},$$

where $k! = k_1! \dots k_n!$ and $|k| = \sum k_i$ if k_i is the number of edges emanating from the vertex p_i of Γ, and $\omega_{(\Gamma,O)} = d\Phi_{e_1} \wedge \dots \wedge d\Phi_{e_{|k|}}$ if $E(\Gamma) = \{e_1 < \dots < e_{|k|}\}$.

We denote by $GO'_{n,m}$ the set of oriented graphs (Γ', O'), with O' not necessarily compatible. Then

$$\mathcal{U}_n = \sum_{m \geq 0} \sum_{(\Gamma', O') \in GO'_{n,m}} w'_{(\Gamma',O')} B_{(\Gamma',O')}.$$

We write the formality equation for $\mathcal{U}$ as

$$F_n = E_n - d_H(\mathcal{U}_n) = 0.$$

Rewriting the proof of the formality theorem by Kontsevich, one can see that F_n looks like a sum over the faces F of the boundary $\partial C^+_{n,m}$ of $C^+_{n,m}$ (see [Arnal et al. 2002] for details):

$$F_n = \sum_{m \geq 0} \sum_{F \subset \partial C^+_{n,m}} \sum_{(\Gamma', O') \in GO'_{n,m}} {w'}^F_{(\Gamma',O')} B_{(\Gamma',O')},$$

where ${w'}^F_{(\Gamma',O')}$ is the integral over F of the closed 2-form $\omega'_{(\Gamma',O')}$.

That $F_n = 0$ then follows directly from the Stokes formula. In particular, we have $\mathfrak{a}(E_n) = 0$.

Now, we saw that $\mathfrak{a}(E_n) = \mathfrak{a}({E_n}^{(1)})$. Thus, for a fixed face F of $\partial C^+_{n,m}$, the corresponding term in $\mathfrak{a}(E_n)$ is a sum over 1-graphs of the form

$$\mathfrak{a}\Bigg(\sum_{(\Gamma', O') \in {GO'}^{(1)}_{n,m}} {w'}^F_{(\Gamma',O')} B_{(\Gamma',O')} \Bigg).$$

Each term of this sum satisfies our relation

$$\mathfrak{a}\Bigg(\sum_{(\Gamma', O') \in {GO'}^{(1)}_{n,m}} {w'}^F_{(\Gamma',O')} B_{(\Gamma',O')}(\alpha) \Bigg) = \sum_{(\Delta, O_\Delta)} w^F_{(\Delta,O_\Delta)} \frac{1}{m!} \sum_{GO^{(1)}_{n,m} \ni (\Gamma,O) \supset (\Delta, O_\Delta)} \frac{\ell!}{k!} \varepsilon(s_O)\varepsilon(\tau_O) B_{(\Gamma,O)}(\alpha),$$

where $w^F_{(\Delta,O_\Delta)} = \int_F \omega_{(\Delta,O_\Delta)}$. Let's prove this:

A face is of either type 1 or type 2 (see [Kontsevich 2003] or [Arnal et al. 2002]). We consider only the faces such that ${w'}^F_{(\Gamma',O')}$ can be different from 0.

(i) If the face F has type 1: Two vertices p_i, p_j of Γ', related by exactly one edge, are collapsing and the face is $F = C_{\{p_i,p_j\}} \times C^+_{\{p,p_1,\dots,\widehat{p_i p_j},\dots,p_n\};\{q_1,\dots,q_m\}}$. We parametrize $C^+_{n,m}$ by

$$\rho = \frac{|p_j - p_i|}{\operatorname{Im} p_i}, \quad p'_j = \frac{p_j - p_i}{|p_j - p_i|} \quad p'_r = \frac{p_r - \operatorname{Re} p_i}{\operatorname{Im} p_i} \quad q'_s = \frac{q_s - \operatorname{Re} p_i}{\operatorname{Im} p_i}.$$

With the signs computed in [Arnal et al. 2002], we can write

$${w'}^F_{(\Gamma',O')} = -\int_{C_2} d\Phi_{\overrightarrow{p_i p_j}} \int_{C^+_{n-1,m}} \omega'_{(\Gamma_2,O_2)},$$

where Γ_2 is the graph obtained from Γ' by gluing together p_i and p_j at the point p and suppressing the edge $\overrightarrow{p_i p_j}$. This weight ${w'}^F_{(\Gamma',O')}$ corresponds to a limit when ρ tends to zero. In fact, if we put

$$C^+_{n,m}(\varepsilon) = C^+_{n,m} \cap \{(p,q) : \rho = \varepsilon\},$$

we get

$${w'}^F_{(\Gamma',O')} = \lim_{\varepsilon\to 0} \frac{k!}{|k|!} \int_{C^+_{n,m}(\varepsilon)} \omega_{(\Gamma',O')} := \lim_{\varepsilon\to 0} {w'}^F_{(\Gamma',O')}(\varepsilon).$$

This limit vanishes for graphs (Γ', O') whose vertices p_i and p_j are linked by two edges or no edges at all. We can thus also consider these graphs in our sum. Then

$$\mathfrak{a}\Bigg(\sum_{(\Gamma',O')\in {GO'}^{(1)}_{n,m}} {w'}^F_{(\Gamma',O')} B_{(\Gamma',O')}(\alpha) \Bigg) = \lim_{\varepsilon\to 0} \mathfrak{a}\Bigg(\sum_{(\Gamma',O')\in {GO'}^{(1)}_{n,m}} {w'}^F_{(\Gamma',O')}(\varepsilon) B_{(\Gamma',O')}(\alpha) \Bigg).$$

Passing to compatible orderings, we obtain

$$\mathfrak{a}\Bigg(\sum_{(\Gamma',O')\in {GO'}^{(1)}_{n,m}} {w'}^F_{(\Gamma',O')} B_{(\Gamma',O')}(\alpha) \Bigg) = \lim_{\varepsilon\to 0} \mathfrak{a}\Bigg(\sum_{(\Gamma,O)\in GO^{(1)}_{n,m}} w^F_{(\Gamma,O)}(\varepsilon) B_{(\Gamma,O)}(\alpha) \Bigg).$$

By Proposition 3.3, we get, as announced,

$$\mathfrak{a}\Bigg(\sum_{(\Gamma',O')\in GO'^{(1)}_{n,m}} w'^{F}_{(\Gamma',O')} B_{(\Gamma',O')}(\alpha)\Bigg)$$
$$= \lim_{\varepsilon\to 0}\frac{1}{m!}\sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w^F_{(\Delta,O_\Delta)}(\varepsilon)\sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}}\frac{l!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma,O)}(\alpha)$$
$$= \frac{1}{m!}\sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w^F_{(\Delta,O_\Delta)}\sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}}\frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma,O)}(\alpha).$$

(ii) If F has type 2: Since our graphs (Γ', O') have exactly one leg for each foot, F is isomorphic to $C^+_{n_1,m_1}\times C^+_{n_2,m_2}$ with $n_2>0$ and $n_1>0$. This case corresponds to the subcase 1 of [Arnal et al. 2002]. Suppose that $p_{i_1},\dots,p_{i_{n_1}}$ and $q_{\ell+1},\dots,q_{\ell+m}$ are collapsing on $q\in\mathbb{R}$. Denote by p_j the first aerial vertex of Γ' that is not a p_{i_s}, and impose the condition $p_j=\sqrt{-1}$. The other parameters are then fixed and we get a parametrization of our configuration space $C^+_{n,m}$ by variables a_r, b_s, q_t (see the notation of [Arnal et al. 2002]). We put $a_{i_1}=q$, $b=\operatorname{Im} p_{i_1}$, and

$$p'_{i_k}=\frac{p_{i_k}-q}{b}\quad(2\le k\le n_1),\qquad q'_{\ell+r}=\frac{q_{\ell+r}-q}{b}\quad(1\le r\le m_1).$$

That is, $p_{i_k}=bp'_{i_k}+qb$ and $q_{\ell+r}=q'_{\ell+r}+qb$, and when b tends to zero, the p_{i_k} and the $q_{\ell+r}$ tend to q. We finally set

$$C^+_{n,m}(\varepsilon)=\{(p,q)\in C^+_{n,m}: b=\varepsilon\}.$$

We get

$$w'^{F}_{(\Gamma',O')}=\lim_{\varepsilon\to 0}\frac{k!}{|k|!}\int_{C^+_{n,m}(\varepsilon)}\omega_{(\Gamma',O')}=\lim_{\varepsilon\to 0} w'^{F}_{(\Gamma',O')}(\varepsilon).$$

If Γ' has a bad edge, the weight $w'^{F}_{(\Gamma',O')}$ vanishes. We can thus consider also these graphs in our sum. Now, a computation similar to that of (i) gives the result.

From now on, for any aerial oriented graph (Δ, O_Δ) in $GO^{(0)}_n$, denote by $C_{(\Delta,O_\Delta)}$ the operator $C_{(\Delta,O_\Delta)}: T^{\otimes n}_{\text{poly}}(\mathbb{R}^d)\to D_{\text{poly}}(\mathbb{R}^d)^{(1)}\simeq T_{\text{poly}}(\mathbb{R}^d)$ defined by

$$C_{(\Delta,O_\Delta)}(\alpha_1\otimes\cdots\otimes\alpha_n)=\sum_{m\ge 0}\frac{1}{m!}\sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}}\frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma,O)}(\alpha_1\otimes\cdots\otimes\alpha_n),$$

where $\varepsilon(s_O)$ and $\varepsilon(\tau_O)$ have the same meaning as above.

Remark. The definition of $C_{(\Delta,O_\Delta)}$ can be extended naturally to the space $GO'^{(0)}_n$ of aerial graphs (Δ', O'_Δ) with O'_Δ not necessarily compatible just by putting

$$C_{(\Delta',O'_\Delta)} = \sum_{m\geq 0} \frac{1}{m!} \sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\ (\Gamma,O)\supset(\Delta,O_\Delta)}} \frac{\ell!}{k!}\varepsilon(s_{O'})\varepsilon(\tau_{O'})B_{(\Gamma',O')}.$$

We will need to use this extension in Section 5.

Summing up:

Proposition 3.4. *Consider the explicit Kontsevich formality $\mathcal{U}$ on $\mathbb{R}^d$. The formality equation can be read as*

$$F_n = E_n - d_H\mathcal{U}_n = 0,$$

and the skewsymmetrization of E_n has the form

$$\mathfrak{a}\circ E_n = \sum_{m\geq 0}\ \sum_{F \text{ face of } \partial C^+_{n,m}}\ \sum_{(\Gamma',O')\in GO'^{(1)}_{n,m}} w'^{F}_{(\Gamma',O')}B_{(\Gamma',O')},$$

where $w'^{F}_{(\Gamma',O')} = \int_{F\in\partial C^+_{n,m}} \omega'_{(\Gamma',O')}$. Then, for each face F,

$$\mathfrak{a}\Bigg(\sum_{(\Gamma',O')\in GO'^{(1)}_{n,m}} w'^{F}_{(\Gamma',O')}B_{(\Gamma',O')}(\alpha)\Bigg) = \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} w^F_{(\Delta,O_\Delta)}C_{(\Delta,O_\Delta)}(\alpha).$$

This proposition suggests that we define:

Definition 3.5. A mapping φ from $T_{\text{poly}}(\mathbb{R}^d)^{\otimes n}$ to $D_{\text{poly}}(\mathbb{R}^d)^{(1)} \simeq T_{\text{poly}}(\mathbb{R}^d)$ is called a *K-graph mapping* if it can be written

$$\varphi = \sum_{(\Delta,O_\Delta)\in GO^{(0)}_n} c_{(\Delta,O_\Delta)}C_{(\Delta,O_\Delta)}$$

with real coefficients $c_{(\Delta,O_\Delta)}$. Such a mapping is homogeneous of degree s if $c_{(\Delta,O_\Delta)} = 0$ for all Δ such that $\# E(\Delta) + s \neq 2n-2$.

Definition 3.6. A *K-graph formality* $\mathcal{F}$ at order n is a graph formality up to order $n-1$ such that $\varphi_n = \mathcal{F}_1^{-1}\circ\mathfrak{a}\circ E_n$ is a K-graph mapping.

4. Symmetrization

4.1. ***Expressions for*** ∂. If B is an n-linear mapping $B : T_{\text{poly}}(\mathbb{R}^d)^{\otimes n} \to T_{\text{poly}}(\mathbb{R}^d)$, we define SB by setting

$$SB(\alpha_1\otimes\cdots\otimes\alpha_n) = \frac{1}{n!}\sum_{\sigma\in S_n} \varepsilon_\alpha(\sigma)B(\alpha_{\sigma(1)}\otimes\cdots\otimes\alpha_{\sigma(n)}),$$

and say that B is symmetric if $SB = B$. Any symmetric mapping can be viewed as a map $\varphi : S^n(T_{\mathrm{poly}}(\mathbb{R}^d)) \to T_{\mathrm{poly}}(\mathbb{R}^d)$. With this symmetrization operator S, the expression of the Chevalley coboundary operator can be conveniently simplified:

Proposition 4.1. *Let $\varphi : S^n(T_{\mathrm{poly}}(\mathbb{R}^d)[1]) \to T_{\mathrm{poly}}(\mathbb{R}^d)[1]$ be an n-cochain for ∂, homogeneous of degree $|\varphi|$. Then we can write*

$$\partial\varphi = S(\tilde{\partial}\varphi),$$

where $\tilde{\partial}\varphi$ is given by

$$\begin{aligned}\tilde{\partial}\varphi(\alpha_1 \otimes \cdots \otimes \alpha_{n+1}) &= (n+1)\big(\varphi(\alpha_1 \otimes \cdots \otimes \alpha_n) \bullet \alpha_{n+1} \\ &+ (-1)^{|\varphi||\alpha_1|}\alpha_1 \bullet \varphi(\alpha_2 \otimes \cdots \otimes \alpha_{n+1}) + (-1)^{|\varphi|+1} n\varphi(\alpha_1 \bullet \alpha_2 \otimes \alpha_3 \otimes \cdots \otimes \alpha_{n+1})\big),\end{aligned}$$

or else by an expression imitating the Hochschild coboundary operator :

$$\begin{aligned}\tilde{\partial}\varphi(\alpha_1 \otimes \cdots \otimes \alpha_{n+1}) = (n+1)\Bigg(&\varphi(\alpha_1 \otimes \cdots \otimes \alpha_n) \bullet \alpha_{n+1} \\ &+ (-1)^{|\varphi|+1} \sum_{k=2}^{n+1} (-1)^{\sum_{s=1}^{k-2}|\alpha_s|} \varphi(\alpha_1 \otimes \cdots \otimes \alpha_{k-1} \bullet \alpha_k \otimes \cdots \otimes \alpha_{n+1}) \\ &+ (-1)^{|\varphi||\alpha_1|}\alpha_1 \bullet \varphi(\alpha_2 \otimes \cdots \otimes \alpha_{n+1})\Bigg).\end{aligned}$$

Proof. By the definition of ∂, we have

$$\partial\varphi(\alpha_1 \ldotp \ldots \ldotp \alpha_{n+1}) = (1) + (2) + (3),$$

with

$$\begin{aligned}(1) &= \sum_{i=1}^{n+1} \varepsilon_\alpha(1 \ldots \hat{\imath} \ldots n{+}1, i)\varphi(\alpha_1 \ldotp \ldots \ldotp \hat{\alpha}_i \ldotp \ldots \ldotp \alpha_{n+1}) \bullet \alpha_i, \\ (2) &= \sum_{i=1}^{n+1} (-1)^{|\varphi||\alpha_i|} \varepsilon_\alpha(i, 1 \ldots \hat{\imath} \ldots n{+}1)\alpha_i \bullet \varphi(\alpha_1 \ldotp \ldots \ldotp \hat{\alpha}_i \ldotp \ldots \ldotp \alpha_{n+1}), \\ (3) &= \sum_{i \neq j} (-1)^{|\varphi|+1} \varepsilon_\alpha(ij, 1 \ldots \widehat{ij} \ldots n{+}1)\varphi(\alpha_i \bullet \alpha_j \ldotp \alpha_1 \ldots \widehat{\alpha_i\alpha_j} \ldotp \ldots \ldotp \alpha_{n+1}).\end{aligned}$$

Now, put

$$\begin{aligned}\psi_1(\alpha_1 \otimes \cdots \otimes \alpha_{n+1}) &= (n{+}1)\varphi(\alpha_1 \otimes \cdots \otimes \alpha_n) \bullet \alpha_{n+1}, \\ \psi_2(\alpha_1 \otimes \cdots \otimes \alpha_{n+1}) &= (-1)^{|\varphi||\alpha_1|}(n{+}1)\alpha_1 \bullet \varphi(\alpha_2 \otimes \cdots \otimes \alpha_{n+1}), \\ \psi_3(\alpha_1 \otimes \cdots \otimes \alpha_{n+1}) &= (-1)^{|\varphi|+1}(n{+}1)n\varphi(\alpha_1 \bullet \alpha_2 \otimes \cdots \otimes \alpha_{n+1}), \\ \psi_3'(\alpha_1 \otimes \cdots \otimes \alpha_{n+1}) &= (-1)^{|\varphi|+1} \sum_{k=2}^{n+1} \big((-1)^{\sum_{s=1}^{k-2}|\alpha_s|} \\ &\qquad \times \varphi(\alpha_1 \otimes \cdots \otimes \alpha_{k-1} \bullet \alpha_k \otimes \cdots \otimes \alpha_{n+1})\big).\end{aligned}$$

First

$$S\psi_1(\alpha_1 \ldots\ldots \alpha_{n+1}) = \frac{(n+1)}{(n+1)!} \sum_{\sigma \in S_{n+1}} \varepsilon_\alpha(\sigma)\varphi(\alpha_{\sigma(1)} \ldots\ldots \alpha_{\sigma(n)}) \bullet \alpha_{\sigma(n+1)}$$
$$= \frac{(n+1)}{(n+1)!} \sum_{i=1}^{n+1} \sum_{\sigma:\sigma(n+1)=i} \varepsilon_\alpha(\sigma)\varphi(\alpha_{\sigma(1)} \ldots\ldots \alpha_{\sigma(n)}) \bullet \alpha_i$$
$$= \frac{(n+1)}{(n+1)!} \sum_{i=1}^{n+1} \sum_{\tau:\tau(i)=i} \varepsilon_\alpha(\tau \circ \sigma_i)\varphi(\alpha_{\tau(\sigma_i(1))} \ldots\ldots \alpha_{\tau(\sigma_i(n))}) \bullet \alpha_i.$$

Here σ_i is the permutation of S_{n+1} sending $(1, \ldots, n+1)$ to $(1, \ldots \hat{\imath} \ldots, n+1, i)$. And, denoting by $\bar{\tau}$ the restriction of τ to $\{1, \ldots \hat{\imath} \ldots, n+1\}$, we easily get

$$S\psi_1(\alpha_1 \ldots\ldots \alpha_{n+1})$$
$$= \frac{n+1}{(n+1)!} \sum_{i=1}^{n+1} \sum_{\bar{\tau} \in S_n} \varepsilon_{\alpha \setminus \{\alpha_i\}}(\bar{\tau})\varepsilon_\alpha(\sigma_i)\varphi(\alpha_{\bar{\tau}(1)} \ldots\ldots \alpha_{\bar{\tau}(n+1)}) \bullet \alpha_i$$
$$= \frac{(n+1)}{(n+1)!} n! \sum_{i=1}^{n+1} \varepsilon_\alpha(1 \ldots \hat{\imath} \ldots n+1, i)\varphi(\alpha_1 \ldots . \hat{\alpha_i} \ldots . \alpha_{n+1}) \bullet \alpha_i = (1).$$

With exactly the same argument, we obtain

$$S\psi_2(\alpha_1 \ldots\ldots \alpha_{n+1}) = (2).$$

Now,

$$S\psi_3(\alpha_1 \ldots\ldots \alpha_{n+1})$$
$$= \sum_{\sigma \in S_{n+1}} \frac{1}{(n+1)!} (-1)^{|\varphi|+1} \varepsilon_\alpha(\sigma)(n+1)n\varphi(\alpha_{\sigma(1)} \bullet \alpha_{\sigma(2)} \otimes \alpha_{\sigma(3)} \otimes \cdots \otimes \alpha_{\sigma(n+1)})$$
$$= \sum_{i \neq j} \sum_{\sigma:\sigma(1)=i, \sigma(2)=j} \Big(\varepsilon_\alpha(\sigma) \frac{1}{(n+1)!} (-1)^{|\varphi|+1}(n+1)n$$
$$\varphi(\alpha_i \bullet \alpha_j \otimes \alpha_1 \otimes \alpha_{\sigma(3)} \otimes \cdots \otimes \alpha_{\sigma(n+1)})\Big)$$
$$= \sum_{i \neq j} \sum_{\tau:\tau(i)=i, \tau(j)=j} \Big(\varepsilon_\alpha(\tau) \frac{1}{(n+1)!} \varepsilon_\alpha(\sigma_{ij})(-1)^{|\varphi|+1}(n+1)n$$
$$\varphi(\alpha_i \bullet \alpha_j \otimes \alpha_{\tau(\sigma_{ij}(3))} \ldots \cdots \otimes \alpha_{\tau(\sigma_{ij}(n+1))})\Big),$$

where σ_{ij} is the permutation of S_{n+1} sending $(1, \ldots, n+1)$ to $(ij, 1 \ldots \widehat{ij} \ldots n+1)$. Now, if $\bar{\tau}$ denotes the restriction of τ to $\{1, \ldots \widehat{ij} \ldots, n+1\}$, we get

$$S\psi_3(\alpha_1 \ldots\ldots \alpha_{n+1})$$
$$= \sum_{i \neq j} \frac{(-1)^{|\varphi|+1}}{(n-1)!} \sum_{\bar{\tau}:\tau(i)=i, \tau(j)=j} \Big(\varepsilon_\alpha(\sigma_{ij})\varepsilon_\alpha(\bar{\tau})$$
$$\varphi(\alpha_i \bullet \alpha_j \otimes \alpha_{\bar{\tau}(1)} \otimes \cdots \widehat{\alpha_i \alpha_j} \cdots \otimes \alpha_{\bar{\tau}(n+1)})\Big)$$

$$= \sum_{i\neq j}(-1)^{|\varphi|+1}\varepsilon_\alpha(ij, 1\dots\widehat{ij}\dots n+1)\varphi(\alpha_i\bullet\alpha_j\dots.\widehat{\alpha_i\alpha_j}\dots.\alpha_{n+1}) = (3).$$

Finally,

$$\begin{aligned}
&S\psi_3'(\alpha_1\dots.\alpha_{n+1})\\
&\quad= \frac{(n+1)}{(n+1)!}\sum_{\sigma\in S_{n+1}}(-1)^{|\varphi|+1}\sum_{k=2}^{n+1}(-1)^{\sum_{s=1}^{k-2}|\alpha_s|}\big(\varepsilon_\alpha(\sigma)\\
&\qquad\qquad \varphi(\alpha_{\sigma(1)}\otimes\dots\otimes\alpha_{\sigma(k-2)}\otimes\alpha_{\sigma(k-1)}\bullet\alpha_{\sigma(k)}\otimes\dots\otimes\alpha_{\sigma(n+1)})\big)\\
&\quad= \frac{(-1)^{|\varphi|+1}}{n!}\sum_{k=2}^{n+1}\sum_{i\neq j}\sum_{\sigma:\sigma(k-1)=i,\,\sigma(k)=j}(-1)^{\sum_{s=1}^{k-2}|\alpha_s|}\big(\varepsilon_\alpha(\sigma)\\
&\qquad\qquad \varphi(\alpha_{\sigma(1)}\dots\otimes\alpha_i\bullet\alpha_j\otimes\dots\otimes\alpha_{\sigma(n+1)})\big).
\end{aligned}$$

Let σ_{ij}^k be the permutation

$$\sigma_{ij}^k : (1\dots n+1)\mapsto(1,\dots,k-2,i,j,k-1,k,\dots,n+1).$$

Then

$$\begin{aligned}
&(-1)^{|\varphi|+1}S\psi_3'(\alpha_1\dots.\alpha_{n+1})\\
&\quad= \frac{1}{n!}\sum_{\substack{2\le k\le n+1\\ i\neq j}}\big((-1)^{\sum_{s=1}^{k-2}|\alpha_s|}\varepsilon_\alpha(\sigma_{ij}^k)(n-1)!\\
&\qquad\qquad \varphi(\alpha_1\otimes\dots\otimes\alpha_{(k-2)}\otimes\alpha_i\bullet\alpha_j\otimes\dots\otimes\alpha_{n+1})\big)\\
&\quad= \frac{1}{n}\sum_{\substack{2\le k\le n+1\\ i\neq j}}\big((-1)^{\sum_{s=1}^{k-2}|\alpha_s|}\varepsilon_\alpha(\sigma_{ij})\varepsilon_\alpha(\rho_{ij}^k)\varphi(\alpha_i\bullet\alpha_j\otimes\alpha_1\otimes\dots\otimes\alpha_{n+1})(-1)^{a_{ijk}}\big),
\end{aligned}$$

with $a_{ijk} = (|\alpha_i|+|\alpha_j|+1)\big(\sum_{s=1}^{k-2}|\alpha_s|\big)$. Here $\sigma_{ij} = (ij1\dots\widehat{ij}\dots n+1)$ and ρ_{ij}^k is the permutation

$$\rho_{ij}^k : (ij1\dots\widehat{ij}\dots n+1)\mapsto(1,\dots,k-2,i,j,k-1,k,\dots,n+1);$$

thus we have used the composition $\sigma_{ij}^k = \rho_{ij}^k\circ\sigma_{ij}$. Now, since

$$\varepsilon_\alpha(\rho_{ij}^k) = (-1)^{(|\alpha_i|+|\alpha_j|)(\sum_{s=1}^{k-2}|\alpha_s|)},$$

we get

$$S\psi_3'(\alpha_1\dots.\alpha_{n+1}) = \sum_{i\neq j}(-1)^{|\varphi|+1}\varepsilon_\alpha(\sigma_{ij})\varphi(\alpha_i\bullet\alpha_j\otimes\alpha_1\otimes\dots\otimes\alpha_{n+1}) = (3).$$

This ends the proof. □

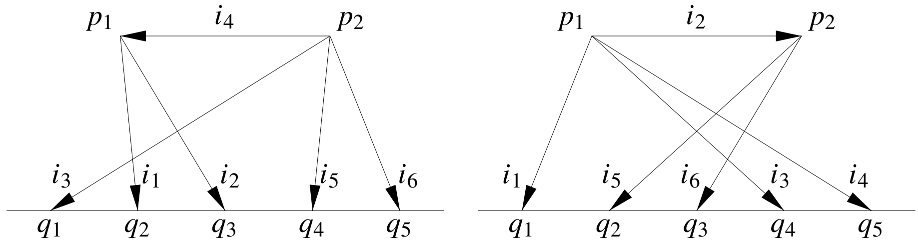


Figure 2. Left: (Γ, O). Right: $(\sigma(\Gamma), \sigma(O))$ with $\sigma = (12)$.

4.2. ***Symmetrization on graphs.*** We now want to describe the symmetrization directly on the space of graphs. Since we are mainly interested in K-graph formalities, we will restrict ourselves to linear combinations of graphs for which the associated operator is a K-graph mapping (see Section 3.3).

The S_n action on $GO_{n,m}$ and $GO_n^{(0)}$. There is a natural action of S_n on $GO_{n,m}$ and $GO_n^{(0)}$, which we now define. Let σ be a permutation in S_n. Let (Γ, O) be in $GO_{n,m}$; for the moment, denote by P_i the set $\operatorname{strt}(p_i)$, ordered by O. Let σ_Γ be the permutation of the ordered set $E(\Gamma)$ of edges of Γ sending $P_1 \cup \cdots \cup P_n$ to $P_{\sigma(1)} \cup \cdots \cup P_{\sigma(n)}$. We denote by $\varepsilon_\Gamma(\sigma_\Gamma)$ the sign of σ_Γ and by $\sigma(\Gamma, O) := (\sigma(\Gamma), \sigma(O))$ the graph with aerial vertices $p'_1 = p_{\sigma(1)}, \ldots, p'_n = p_{\sigma(n)}$ oriented by $\sigma_\Gamma(E(\Gamma))$ (see Figure 2). We apply the same definition to aerial graphs in $GO_n^{(O)}$. Clearly, σ sends $GO_{n,m}$ (and $GO_n^{(0)}$) onto itself.

This S_n action on $GO_{n,m}^{(1)}$ is entirely different from the action of S_m defined in Section 3. But there is an analog of Lemma 3.1:

Lemma 4.2. *For all σ in S_n, all (Γ, O) in $GO_{n,m}^{(1)}$ and all polyvector fields α_i,*

$$B_{(\sigma(\Gamma),\sigma(O))}(\alpha_1 \otimes \cdots \otimes \alpha_n) = B_{(\Gamma,O)}(\alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(n)}).$$

Proof. With our notations,

$$B_{(\Gamma,O)}(\alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(n)})(f_1, \ldots, f_m) = \sum_{1 \le i_{t_1} \ldots i_{t_{|k|}} \le d} \prod_{i=1}^{n} \partial_{\operatorname{end}(p_i)} \alpha_{\sigma(i)}^{P_i} \prod_{j=1}^{m} \partial_{\operatorname{end}(q_j)} f_j.$$

Since the permutation σ_Γ does not affect the order inside each P_i, we have

$$\begin{aligned} B_{(\Gamma,O)}(\alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(n)})(f_1, \ldots, f_m) &= \sum_{1 \le i_{t_1} \ldots i_{t_{|k|}} \le d} \prod_{i=1}^{n} \partial_{\operatorname{end}(p_{\sigma(i)})} \alpha_{\sigma(i)}^{P_{\sigma(i)}} \prod_{j=1}^{m} \partial_{\operatorname{end}(q_j)} f_j \\ &= \sum_{1 \le i_{t_1} \ldots i_{t_{|k|}} \le d} \prod_{i'=1}^{n} \partial_{\operatorname{end}(p'_i)} \alpha_{i'}{}^{P_{i'}} \prod_{j=1}^{m} \partial_{\operatorname{end}(q_j)} f_j \\ &= B_{\sigma(\Gamma,O)}(\alpha_1 \otimes \cdots \otimes \alpha_n)(f_1, \ldots, f_m). \qquad \square \end{aligned}$$

Symmetrization for K-graph mappings.

Definition 4.3. Let

$$(\delta, O_\delta) = \sum_{(\Delta, O_\Delta)\in GO_n^{(0)}} c_{(\Delta, O_\Delta)}(\Delta, O_\Delta)$$

be a linear combination of aerial graphs with n vertices. We say that (δ, O_δ) is symmetric if

$$c_{(\sigma(\Delta),\sigma(O_\Delta))} = \varepsilon_\Delta(\sigma_\Delta)c_{(\Delta, O_\Delta)} \quad \text{for all } (\Delta, O_\Delta) \text{ and } \sigma \in S_n.$$

Proposition 4.4. *If* $(\delta, O_\delta) = \sum_{(\Delta, O_\Delta)\in GO_n^{(0)}} c_{(\Delta, O_\Delta)}(\Delta, O_\Delta)$ *is symmetric, so is the corresponding K-graph mapping*

$$C_{(\delta, O_\delta)} = \sum_{(\Delta, O_\Delta)\in GO_n^{(0)}} c_{(\Delta, O_\Delta)} C_{(\Delta, O_\Delta)}.$$

Proof. Let σ be in S_n and let $\alpha_1, \dots, \alpha_n$ be n polyvector fields on $\mathbb{R}^d$. By Lemma 4.2 and using the fact that δ is symmetric, we have

$$\begin{aligned}
&C_{(\delta, O_\delta)}(\alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(n)}) \\
&\quad= \sum_{(\Delta, O_\Delta)} c_{(\Delta, O_\Delta)} \sum_{m\geq 0} \frac{1}{m!} \sum_{(\Gamma, O)\supset(\Delta, O_\Delta)} \frac{\ell!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\sigma(\Gamma),\sigma(O))}(\alpha_1\otimes\cdots\otimes\alpha_n) \\
&\quad= \sum_{\sigma^{-1}(\Delta, O_\Delta)} c_{(\sigma^{-1}(\Delta),\sigma^{-1}(O_\Delta))} \sum_{m\geq 0} \frac{1}{m!} \sum_{\sigma^{-1}(\Gamma, O)\supset\sigma^{-1}(\Delta, O_\Delta)} \\
&\qquad\qquad \frac{\ell!}{k!}\varepsilon(s_{\sigma^{-1}(O)})\varepsilon(\tau_{\sigma^{-1}(O)})B_{(\Gamma, O)}(\alpha_1 \otimes \cdots \otimes \alpha_n) \\
&\quad= \sum_{(\Delta, O_\Delta)} \varepsilon_\Delta(\sigma_\Delta)c_{(\Delta, O_\Delta)} \sum_{m\geq 0} \frac{1}{m!} \sum_{(\Gamma, O)\supset(\Delta, O_\Delta)} \\
&\qquad\qquad \frac{\ell!}{k!}\varepsilon(s_{\sigma^{-1}(O)})\varepsilon(\tau_{\sigma^{-1}(O)})B_{(\Gamma, O)}(\alpha_1 \otimes \cdots \otimes \alpha_n).
\end{aligned}$$

Extending σ_Δ to $E(\Gamma)$ in the obvious way, we can write

$$\tau_O \circ s_O = \sigma_\Delta \circ \tau_{\sigma^{-1}(O)} \circ s_{\sigma^{-1}(O)} \circ \sigma_\Gamma^{-1}.$$

Thus

$$\begin{aligned}
&C_\delta(\alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(n)}) \\
&= \sum_{(\Delta, O_\Delta)} c_{(\Delta, O_\Delta)} \sum_{m\geq 0} \frac{1}{m!} \sum_{(\Gamma, O)\supset(\Delta, O_\Delta)} \varepsilon_\Gamma(\sigma_\Gamma)\frac{l!}{k!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma, O)}(\alpha_1 \otimes \cdots \otimes \alpha_n).
\end{aligned}$$

Since each $\varepsilon_\Gamma(\sigma_\Gamma)$ clearly coincides with the sign $\varepsilon_\alpha(\sigma)$ of σ, we get

$$C_\delta(\alpha_{\sigma(1)} \otimes \cdots \otimes \alpha_{\sigma(n)}) = \varepsilon_\alpha(\sigma) C_\delta(\alpha_1 \otimes \cdots \otimes \alpha_n).$$

This proves the result. $\square$

5. Chevalley cohomology for graphs

We will now prove that, on K-graph mappings, the Chevalley coboundary operator can be nicely reduced to an operator acting on purely aerial graphs.

5.1. ***Purely aerial and compatible oriented graphs.*** For any (Δ, O_Δ) in $GO_n^{(0)}$ with vertices $p_1 < \cdots < p_n$, we still write $\ell_i = \#\,\mathrm{strt}^\Delta(p_i)$. We also put $|\Delta| = \sum \ell_i = |\ell|$.

Fix two indexes $i \neq j$. We say that an aerial graph $(\Delta', O_{\Delta'})$ in $GO'^{(0)}_{n+1}$ (with $O_{\Delta'}$ not necessarily compatible) with vertices $p'_1 < \cdots < p'_{n+1}$ reduces to (Δ, O_Δ) in the indexes i, j if the two following assertions hold:

(i) The vertices p'_i and p'_j of Δ' are linked by only the edge $\overrightarrow{p'_i p'_j}$.

(ii) the new graph $(\Delta'_{ij}, O_{\Delta'_{ij}})$, obtained by gluing together the vertices p'_i, p'_j of Δ', by suppressing the edge $\overrightarrow{p'_i p'_j}$ and considering the induced ordering, coincides with (O, Δ).

We say that $(\Delta', O_{\Delta'})$ reduces properly to (Δ, O_Δ) in the indexes i, j if $(\Delta', O_{\Delta'})$ reduces to (Δ, O_Δ) in the same indexes and in addition

$$\inf\big(\#\,\mathrm{strt}^{\Delta'}(p'_i) + \#\,\mathrm{end}^{\Delta'}(p'_i),\ \#\,\mathrm{strt}^{\Delta'}(p'_j) + \#\,\mathrm{end}^{\Delta'}(p'_j)\big) > 1.$$

In the situations above we write

$$(\Delta', O_{\Delta'}) \to_{i,j} (\Delta, O_\Delta) \quad \text{and} \quad (\Delta', O_{\Delta'}) \to_{i,j}^{\mathrm{prop}} (\Delta, O_\Delta),$$

respectively. We use the same notation for graphs (Γ, O) in $GO^{(1)}_{n,m}$.

Definition 5.1. If (Δ, O_Δ) is an aerial oriented graph in $GO_n^{(0)}$, we define the coboundary $\partial(\Delta, O_\Delta)$ of (Δ, O_Δ) by

$$\partial(\Delta, O_\Delta) = (-1)^{|\Delta|+1} \sum_{i \neq j} \sum_{(\Delta', O_{\Delta'}) \to_{i,j}^{\mathrm{prop}} (\Delta, O_\Delta)} \varepsilon(\Delta', O_{\Delta'}, \Delta, O_\Delta)(\Delta', O_{\Delta'}).$$

Here $\varepsilon(\Delta', O_{\Delta'}, \Delta, O_\Delta)$ is the sign of the permutation of $E(\Delta')$, that consists in putting first the edge $\overrightarrow{p'_i p'_j}$, then the other edges starting from p'_i (with the ordering induced by $O_{\Delta'}$), then the edges starting from p'_j (also with the induced ordering), and finally all the remaining edges (with the ordering given by O_Δ).

We extend ∂ by linearity to all combinations $(\delta, O_\delta) = \sum_{(\Delta,O_\Delta)} c_{(\Delta,O_\Delta)}(\Delta, O_\Delta)$. Note that the restriction of ∂ to symmetric combinations of graphs is an operator of cohomology.

More precisely:

Proposition 5.2. *With the same notations as above and for any symmetric combination of graphs* (δ, O_δ), *we have*

$$\partial(C_{(\delta,O_\delta)}) = C_{\partial(\delta,O_\delta)}.$$

Proof. First, $C_{(\Delta,O_\Delta)}$ is a linear combination of m-differential operators $B_{(\Gamma,O)}(\alpha)$, for certain k_i-vector fields α_i:

$$m-2 = \left|B_{(\Gamma,O)}(\alpha_1 \otimes \cdots \otimes \alpha_n)\right| = \sum_{i=1}^{n} |\alpha_i| + |B_{(\Gamma,O)}| = \sum_{i=1}^{n} k_i - 2n + |B_{(\Gamma,O)}|,$$

where $|\ |$ stands for the degree in $T_{\mathrm{poly}}(\mathbb{R}^d)[1]$ and $D_{\mathrm{poly}}(\mathbb{R}^d)[1]$. Now, since the graphs (Γ, O) occurring in $C_{(\Delta,O_\Delta)}$ are 1-graphs, we have $k_i = \ell_i + m_i$ for each i and $m = \sum_{i=1}^{n} m_i$. Thus

$$|B_{(\Gamma,O)}| = \sum_{i=1}^{n} \ell_i = |\Delta| \bmod 2.$$

Now, by the definition of ∂ on operators,

$$\begin{aligned}
&\partial C_{(\Delta,O_\Delta)}(\alpha_1 \ldots\ldots \alpha_{n+1})\\
&\quad= \sum_{j=1}^{n+1} \varepsilon_\alpha(1\ldots\hat{\jmath}\ldots n{+}1, j) C_{(\Delta,O_\Delta)}(\alpha_1 \ldots\ldots \hat{\alpha}_j \ldots\ldots \alpha_{n+1}) \bullet \alpha_j\\
&\qquad+ \sum_{i=1}^{n+1} (-1)^{|\Delta||\alpha_i|} \varepsilon_\alpha(i, 1\ldots\hat{\imath}\ldots n{+}1) \alpha_i \bullet C_{(\Delta,O_\Delta)}(\alpha_1 \ldots\ldots \hat{\alpha}_i \ldots\ldots \alpha_{n+1})\\
&\qquad+ \sum_{i\neq j} (-1)^{|\Delta|+1} \varepsilon_\alpha(ij, 1\ldots\widehat{\imath\jmath}\ldots n{+}1) C_{(\Delta,O_\Delta)}(\alpha_i \bullet \alpha_j \,.\, \alpha_1 \ldots \widehat{\alpha_i\alpha_j} \ldots\ldots \alpha_{n+1})\\
&\quad= \text{(i)} + \text{(ii)} + \text{(iii)}.
\end{aligned}$$

We first consider the term (iii). We have

$$\begin{aligned}
&C_{(\Delta,O_\Delta)}(\alpha_i \bullet \alpha_j \ldots\ldots \widehat{\alpha_i\alpha_j} \ldots\ldots \alpha_{n+1})\\
&\quad= \sum_{m\geq 0} \frac{1}{m!} \sum_{GO^{(1)}_{n,m} \ni (\Gamma,O) \supset (\Delta,O_\Delta)} \frac{\ell!}{k!} \varepsilon(s_O)\varepsilon(\tau_O) B_{(\Gamma,O)}(\alpha_i \bullet \alpha_j \ldots\ldots \widehat{\alpha_i\alpha_j} \ldots\ldots \alpha_{n+1}).
\end{aligned}$$

Now, we can write (see [Arnal et al. 2002] for details)

$$B_{(\Gamma,O)}(\alpha_i \bullet \alpha_j \,.\, \alpha_1 \ldots . \,\alpha_{n+1}) = \sum_{(\Gamma',O')\to_{i,j}(\Gamma,O)} (-1)^{\ell_{\Gamma'}-1} B_{(\Gamma',O')}(\alpha_1 \,. \ldots . \,\alpha_{n+1}),$$

where $\ell_{\Gamma'}$ denotes the position of the edge $\overrightarrow{p'_i p'_j}$ in Γ', and the sign $(-1)^{\ell_{\Gamma'}-1}$ comes directly from the definition of $\bullet$.

Next consider a graph (Γ', O') that reduces to (Γ, O) in the indexes i, j. We permute the edges as follows: we put first the edge $\overrightarrow{p'_i p'_j}$, then the other edges starting from p'_i, then the edges starting from p'_j, and finally we put all the legs at the end in order of their feet. This gives a sign that can be written as

$$\varepsilon_\alpha(ij, 1 \ldots \widehat{ij} \ldots n{+}1)(-1)^{\ell_{\Gamma'}-1}\varepsilon(s_O)\varepsilon(\tau_O).$$

Starting from (Γ', O'), one can also place the legs at the end in order of their feet, preceded by the aerial edges starting from p'_i and those starting from p'_j, and then by the aerial edge $\overrightarrow{p'_i p'_j}$ at the first position. If we denote by Δ' the aerial part of Γ' and by $\ell_{\Delta'}$ the position of the edge $\overrightarrow{p'_i p'_j}$ in Δ', the resulting sign is

$$\varepsilon(s_{O'})\varepsilon(\tau_{O'})\varepsilon(\Delta', \Delta)(-1)^{\ell_{\Delta'}-1}.$$

These two permutations of the edges of Γ' obviously coincide; thus

$$\varepsilon_\alpha(ij, 1 \ldots \widehat{ij} \ldots n{+}1)(-1)^{\ell_{\Gamma'}-1}\varepsilon(s_O)\varepsilon(\tau_O) = \varepsilon(s_{O'})\varepsilon(\tau_{O'})\varepsilon(\Delta', \Delta)(-1)^{\ell_{\Delta'}-1}.$$

It follows that

$$\begin{aligned}
&C_{(\Delta,O_\Delta)}(\alpha_i \bullet \alpha_j \,. \ldots . \,\widehat{\alpha_i\alpha_j} \,. \ldots . \,\alpha_{n+1})\\
&\quad = \sum_{m\ge 0}\frac{1}{m!}\sum_{\substack{(\Gamma,O)\in GO^{(1)}_{n,m}\\(\Gamma,O)\supset(\Delta,O_\Delta)}} \varepsilon_\alpha(ij1\ldots n{+}1)\Biggl(\sum_{(\Gamma',O')\underset{i,j}{\to}(\Gamma,O)} \frac{\ell!}{k!}\varepsilon(s_{O'})\varepsilon(\tau_{O'})(-1)^{\ell_{\Delta'}-1}\\
&\qquad\qquad\qquad\qquad\qquad\qquad \varepsilon(\Delta', \Delta)B_{(\Gamma',O')}(\alpha_1 \,. \ldots . \,\alpha_{n+1})\Biggr)\\
&\quad = \varepsilon_\alpha(ij, 1\ldots n{+}1)\sum_{m\ge 0}\frac{1}{m!}\Biggl(\sum_{(\Delta',O_{\Delta'})\to_{i,j}(\Delta,O_\Delta)} (-1)^{\ell_{\Delta'}-1}\varepsilon(\Delta',\Delta)\\
&\qquad\qquad\qquad \sum_{(\Gamma',O')\supset(\Delta',O_{\Delta'})} \frac{\ell!}{k!}\varepsilon(s'_O)\varepsilon(\tau'_O)B_{(\Gamma',O')}(\alpha_1 \,. \ldots . \,\alpha_{n+1})\Biggr)\\
&\quad = \varepsilon_\alpha(ij, 1\ldots n{+}1)\sum_{(\Delta',O_{\Delta'})\to_{i,j}(\Delta,O_\Delta)} (-1)^{\ell_{\Delta'}-1}\varepsilon(\Delta', \Delta)C_{(\Delta',O_{\Delta'})}(\alpha_1 \,. \ldots . \,\alpha_{n+1}).
\end{aligned}$$

Finally,

$$\begin{aligned}(\mathrm{iii}) &= (-1)^{|\Delta|+1} \sum_{(\Delta', O_{\Delta'}) \to_{i,j} (\Delta, O_\Delta)} (-1)^{\ell_{\Delta'}-1} \varepsilon(\Delta', \Delta) C_{(\Delta', O_{\Delta'})}(\alpha_1 \bullet \dots \bullet \alpha_{n+1}) \\ &= (-1)^{|\Delta|+1} \sum_{(\Delta', O_{\Delta'}) \to_{i,j} (\Delta, O_\Delta)} \varepsilon(\Delta', O_{\Delta'}, \Delta, O_\Delta) C_{(\Delta', O_{\Delta'})}(\alpha_1 \bullet \dots \bullet \alpha_{n+1}).\end{aligned}$$

Now let $(\delta, O_\delta) = \sum c_{(\Delta, O_\Delta)}(\Delta, O_\Delta)$ be a symmetric combination of graphs and put

$$C_{(\delta, O_\delta)} = (\mathrm{i})_\delta + (\mathrm{ii})_\delta + (\mathrm{iii})_\delta.$$

We have to prove that $-\big((\mathrm{i})_\delta + (\mathrm{ii})_\delta\big)$ coincides with the nonproper terms of $(\mathrm{iii})_\delta$, that is, with

$$\sum_{(\Delta, O_\Delta)} c_{(\Delta, O_\Delta)} (-1)^{|\Delta|+1} \sum_{i \neq j} \sum_{(\Delta', O_{\Delta'}) \to^{\mathrm{nonprop}}_{i,j} (\Delta, O_\Delta)} (-1)^{\ell_{\Delta'}-1} \varepsilon(\Delta', O_{\Delta'})(\Delta', O_{\Delta'}).$$

Consider first the term

$$(\mathrm{ii})_\delta = \sum_{(\Delta, O_\Delta)} c_{(\Delta, O_\Delta)} \sum_{i=1^n} (-1)^{|\Delta||\alpha_i|} \varepsilon_\alpha(i, 1 \dots n+1) \alpha_i \bullet C_{(\Delta, O_\Delta)}(\alpha_1 \dots \hat{\alpha}_i \dots \alpha_{n+1}).$$

We identify $C_{(\Delta, O_\Delta)}(\alpha)$ with a polyvector field, and put

$$C_{(\Delta, O_\Delta)}(\alpha_1 \dots \hat{\alpha}_i \dots \alpha_{n+1}) = \big(C_{(\Delta, O_\Delta)}(\alpha_1 \dots \hat{\alpha}_i \dots \alpha_{n+1})\big)^{r_1 \dots r_m} \partial_{r_1} \wedge \dots \wedge \partial_{r_m}.$$

Thus

$$\begin{aligned}&\alpha_i \bullet C_{(\Delta, O_\Delta)}(\alpha_1 \dots \hat{\alpha}_i \dots \alpha_{n+1}) \\ &\quad = \sum_{j \neq i} \sum_{l \leq k_i} (-1)^{l-1} \alpha_i^{i_1 \dots s \dots k_i - 1} \big(C_{(\Delta, O_\Delta)}(\alpha_1 \dots \partial_s(\alpha_j) \dots \hat{\alpha}_i \dots \alpha_{n+1})\big)^{r_1 \dots r_m} \\ &\qquad\qquad \partial_{i_1} \wedge \dots \partial_{i_{k_1 - 1}} \wedge \partial_{r_1} \wedge \dots \wedge \partial_{r_m}.\end{aligned}$$

Let σ be the permutation $(j 1 \dots \widehat{ij} \dots n+1)$ and $(\Delta^\sigma, O_{\Delta^\sigma})$ be the aerial graph obtained by relabeling the vertices of Δ in the ordering given by σ. Then

$$C_{(\Delta, O_\Delta)}(\alpha_1 \dots \partial_s(\alpha_j) \dots \hat{\alpha}_i \dots \alpha_{n+1}) = C_{(\Delta^\sigma, O_{\Delta^\sigma})}(\partial_s(\alpha_j) \alpha_1 \dots \widehat{\alpha_i \alpha_j} \dots \alpha_{n+1}).$$

But (δ, O_δ) is symmetric; thus

$$c_{(\Delta^\sigma, O_{\Delta^\sigma})} = c_{(\Delta, O_\Delta)} \varepsilon_\alpha(j, 1 \dots \widehat{ij} \dots n+1).$$

Hence,

$$\begin{aligned}(\mathrm{ii})_\delta = &\sum_{(\Delta, O_\Delta)} \sum_{i \neq j} (-1)^{|\Delta||\alpha_i|} \varepsilon_\alpha(i j 1 \dots n+1) c_{(\Delta, O_\Delta)} \sum_{\ell \leq k_i} (-1)^{\ell - 1} \alpha_i^{i_1 \dots s \dots i_{k_i - 1}} \\ &\big(C_{(\Delta, O_\Delta)}(\partial_s \alpha_j \alpha_1 \dots \widehat{\alpha_i \alpha_j} \dots \alpha_{n+1})\big)^{r_1 \dots r_m} \partial_{i_1} \wedge \dots \wedge \partial_{i_{k_i - 1}} \wedge \partial_{r_1} \wedge \dots \wedge \partial_{r_m}.\end{aligned}$$

It is now easy to see that $-(\mathrm{ii})_\delta$ coincides with certain nonproper terms of $(\mathrm{iii})_\delta$ — more precisely, with those corresponding to the graphs Δ' with

$$(\Delta', O_{\Delta'}) \to_{i,j} (\Delta, O_\Delta) \quad \text{and} \quad \big(\#\,\mathrm{strt}^{\Delta'}(p'_i) + \#\,\mathrm{end}^{\Delta'}(p'_i)\big) = 1.$$

(In this case, $\ell_{\Delta'} = 1$.) In the same way, one can check that $-(\mathrm{i})_\delta$ coincides with the remaining nonproper terms of $(\mathrm{iii})_\delta$, that is, with the nonproper terms corresponding to the case

$$(\Delta', O_{\Delta'}) \to_{i,j} (\Delta, O_\Delta) \quad \text{and} \quad (\#\,\mathrm{strt}^{\Delta'}(p'_j) + \#\,\mathrm{end}^{\Delta'}(p'_j)) = 1.$$

The result follows. □

5.2. ***Purely aerial and nonoriented graphs.*** We say that a graph is nonoriented if there is an ordering only on the aerial vertices but no ordering on the edges of the graph. We are now interested in translating our cohomology on nonoriented graphs. Let Δ be an aerial nonoriented graph with n vertices $p_1 < \cdots < p_n$. We still write $\ell_i = \mathrm{strt}^\Delta(p_i)$ and $\ell! = \ell_1! \ldots \ell_n!$. We order the edges of Δ lexicographically:

$$\overrightarrow{ab} \le \overrightarrow{a'b'} \quad \text{if and only if} \quad (a = a' \text{ and } a < b') \text{ or } (a < a').$$

This yields a compatible ordering on Δ, called the standard ordering. We denote by $(\Delta, O_\Delta^{\mathrm{std}})$ the resulting oriented graph.

Now put

$$\Delta = \frac{1}{\ell!} \sum_{O_\Delta : (\Delta, O_\Delta) \in GO_n^{(0)}} \varepsilon(\sigma_{(O_\Delta^{std}, O_\Delta)})(\Delta, O_\Delta).$$

By the definition of ∂ on compatible oriented graphs, we have:

$$\partial\Delta = \frac{1}{\ell!}\Bigg(\sum_{O_\Delta : (\Delta, O_\Delta) \in GO_n^{(0)}} \varepsilon(\sigma_{(O_\Delta^{std}, O_\Delta)})(-1)^{|\Delta|+1} \sum_{i \neq j} \sum_{(\Delta', O_{\Delta'}) \to_{i,j}^{\mathrm{prop}} (\Delta, O_\Delta)} \varepsilon(\Delta', O_{\Delta'}, \Delta, O_\Delta)(\Delta', O_{\Delta'}) \Bigg).$$

Note that the sign

$$\tilde{\varepsilon}(\Delta, \Delta') := \varepsilon(O_\Delta^{\mathrm{std}}, O_\Delta)\varepsilon(\Delta', O_{\Delta'}, \Delta, O_\Delta)\varepsilon(O_{\Delta'}^{\mathrm{std}}, O_{\Delta'})$$

does not depend on O_Δ or O'_Δ. This yields a very simple expression for the coboundary $\partial\Delta$ of Δ:

$$\partial\Delta = \frac{1}{\ell!} \sum_{i \neq j} \sum_{\Delta' \supset \Delta} \tilde{\varepsilon}(\Delta', \Delta)\Delta'.$$

We extend ∂ to linear combination of graphs $\delta = \sum_\Delta c_\Delta \Delta$.

Now, if Δ is a nonoriented graph with vertices $p_1 < \cdots < p_n$ and if σ is a permutation in S_n, we denote by $\sigma(\Delta)$ the nonoriented graph with vertices $p_{\sigma(1)} < \cdots < p_{\sigma(n)}$. A linear combination $\delta = \sum_\Delta c_\Delta \Delta$ of nonoriented graphs with n labeled vertices is said to be symmetric if for any σ in S_n, we have $c_\Delta = c_{\sigma(\Delta)}$. Our operator ∂ restricted to symmetric δ is clearly a cohomology operator.

More precisely, for an aerial nonoriented graph Δ, let

$$C_\Delta = \frac{1}{\ell!} \sum_{O_\Delta:(\Delta,O_\Delta)\in GO_n^{(0)}} \varepsilon(\sigma_{(O_\Delta^{\mathrm{std}}),O_\Delta)})C_{(\Delta,O_\Delta)}.$$

Extend this definition by linearity to all linear combinations. Then, by computations similar to those we did before for oriented graphs, we can prove:

Proposition 5.3. *For any symmetric combination $\delta = \sum_\Delta c_\Delta C_\Delta$ of graphs with n labeled vertices, we have*

$$\partial(C_\delta) = C_{\partial(\delta)}.$$

5.3. ***Examples.*** Let Δ_1 be the graph with only one vertex p_1. Let α_1 be a k_1-vector field. Then

$$C_{\Delta_1}(\alpha_1) = \frac{1}{(k_1!)^2} \sum_{GO_{n,m}^{(1)}\ni(\Gamma,O)\supset\Delta_1} \varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma,O)}(\alpha_1).$$

There is only one graph occurring in this sum, namely the graph Γ with one aerial vertex p_1, k_1 terrestrial vertices $q_1,\ldots,q_{k_1}$ and k_1 edges $\overrightarrow{p_1q_1},\ldots,\overrightarrow{p_1q_{k_1}}$. For any σ in S_{k_1}, denote by (Γ, O^σ) the graph Γ endowed with the ordering given by $\overrightarrow{p_1q_{\sigma(1)}}\ldots\overrightarrow{p_1q_{\sigma(k_1)}}$. Clearly,

$$C_{\Delta_1}(\alpha_1) = \frac{1}{(k_1!)^2}\sum_{\sigma\in S_{k_1}} \varepsilon(\sigma)B_{(\Gamma,O^\sigma)}(\alpha_1) = \mathscr{F}_1^{(0)}(\alpha_1) \simeq \alpha_1,$$

and C_{Δ_1} just corresponds to the identity mapping.

Now let Δ_2 be the aerial graph with two vertices $p_1 < p_2$ and one edge $\overrightarrow{p_1p_2}$. Let α_1 be a k_1-vector field and α_2 a k_2-vector field. Then

$$C_{\Delta_2}(\alpha_1\otimes\alpha_2) = \frac{1}{(k_1+k_2-1)!} \sum_{\substack{(\Gamma,O)\supset\Delta_2\\(\Gamma,O)\in GO_{n,m}^{(1)}}} \frac{1}{k_1!\,k_2!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma,O)}(\alpha_1\otimes\alpha_2).$$

There are exactly $(k_1+k_2-1)!/\big((k_1-1)!\,k_2!\big)$ graphs Γ containing Δ_2 and having exactly (k_1-1) legs starting from p_1 and k_2 legs starting from p_2. For each of them, we choose a compatible ordering. There are $k_1!\,k_2!$ possibilities to do it. Thus, there are exactly $k_1(k_1+k_2-1)!$ compatible oriented graphs (Γ, O) occurring in C_{Δ_2}. For each of these graphs, $\varepsilon(s_O)$ corresponds to the permutation of S_{k_1} that consists in putting the aerial edge of (Γ, O) at the first position and $\varepsilon(\tau_O)$ corresponds to the permutation of $S_{k_1+k_2-1}$ that consists in putting the legs in the order of the feet. There is thus $k_1(k_1+k_2-1)!$ terms in C_{Δ_2}, each of which looks like

$$\frac{1}{(k_1+k_2-1)!\,k_1!\,k_2!}\varepsilon(s_O)\varepsilon(\tau_O)B_{(\Gamma,O)} = \frac{1}{(k_1+k_2-1)!\,k_1!\,k_2!}(-1)^{\ell-1}\varepsilon(\sigma)$$
$$\alpha_1^{i_{\sigma(1)}\dots i_{\sigma(\ell-1)} s\, i_{\sigma(\ell)}\dots i_{\sigma(k_1-1)}}\partial_s\big(\alpha_2^{i_{\sigma(k_1)}\dots i_{\sigma(k_1+k_2-1)}}\big)\partial_{i_{\sigma(1)}}\otimes\cdots\otimes\partial_{i_{s(k_1+k_2-1)}}.$$

Thus

$$C_{\Delta_2}(\alpha_1\otimes\alpha_2)=\mathscr{F}_1^{(0)}(\alpha_1\bullet\alpha_2)\simeq\alpha_1\bullet\alpha_2.$$

Now consider the aerial graph Δ_2^- with two vertices $p_1<p_2$ and one edge $\overrightarrow{p_2p_1}$. In the same way as above, one can see that

$$C_{\Delta_2^-}(\alpha_1\otimes\alpha_2)=(-1)^{k_1k_2}\alpha_2\bullet\alpha_1.$$

In other words, $C_{\Delta_2+\Delta_2^-}$ coincides with Q_2.

The identity map Id and Q_2 are thus easy examples of K-graph mappings, and the fact that Q_2 is the Chevalley coboundary of Id can be checked directly on the graphs. Indeed, we have with our notations:

$$\partial\Delta_1=\tilde{\varepsilon}(\Delta_2,\Delta_1)\Delta_2+\tilde{\varepsilon}(\Delta_2{}^-,\Delta_1)\Delta_2{}^-=\Delta_2+\Delta_2^-.$$

Hence,

$$Q_2=C_{\Delta_2+\Delta_2^-}=C_{\partial\Delta_1}=\partial C_{\Delta_1}=\partial\,\mathrm{Id}\,.$$

6. Triviality of the cohomology for small n

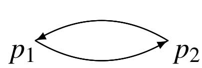
Our first example proves that the first cohomology group H^1 is trivial, since, for $n=1$, there is only one purely aerial graph, namely Δ_1.

Now suppose $n=2$. There is one graph Δ with two vertices and with degree 0 $|\Delta|=0$, the nonconnected symmetric graph denoted $\Delta_1\times\Delta_1$ without any edges. Its coboundary does not vanish; in the obvious notation, we have

$$\partial(\Delta_1\times\Delta_1)=S\big((\Delta_2^++\Delta_2^-)\times\Delta_1+\Delta_1\times(\Delta_2^2+\Delta_2^-)\big)\neq 0.$$

In degree 1 ($|\Delta|=1$), there is only one symmetrized graph, $\Delta_2^++\Delta_2^-$. Our second example shows that this graph is a coboundary.

Finally, there is no graph with degree larger than 1; indeed, the number of edges for a graph with 2 vertices is at most 2, but there is only one graph Δ with $|\Delta|=2$, the graph $\Delta_{2,2}$ given by



But the symmetrization of this graph is $\Delta_{2,2} - \Delta_{2,2} = 0$. Thus the second cohomology group H^2 vanishes.

It is possible to prove with elementary arguments that $H^3 = 0$ too. For that, we consider the different cases, $|\Delta| = 0, \ldots, 6$, then we define the order of a graph in the following way:

We define the order o_i of a vertex p_i as the pair (ℓ_i, r_i) of number ℓ_i of edges starting from p_i and the number r_i of edges ending at p_i, we shall say that $o = (\ell, r)$ is smaller than $o' = (\ell', r')$ and note $o < o'$ if and only if $\ell + r < \ell' + r'$ or $\ell + r = \ell' + r'$ and $\ell < \ell'$.

We define then the order $o(\Delta)$ of a graph Δ as $o(\Delta) = (o_1, \ldots, o_n)$ if Δ has n vertices. The order $o(\delta)$ of a linear combination $\delta = \sum c_\Delta \Delta$ of graphs is the maximum of $o(\Delta)$ for $c_\Delta \neq 0$ for the lexicographic ordering. We define the symbol of δ by

$$\text{symb } \delta = \sum_{o(\Delta) = o(\delta)} c_\Delta C_\Delta.$$

Case 1: $|\Delta| = 0$. There is only one graph, disconnected and symmetric: the graph $\Delta_1 \times \Delta_1 \times \Delta_1$. It is not a cocycle since

$$\partial(\Delta_1 \times \Delta_1 \times \Delta_1) = S\big((\Delta_2^+ + \Delta_2^-) \times \Delta_1 \times \Delta_1\big) \neq 0.$$

Case 2: $|\Delta| = 1$. There is, up to the ordering of vertices, only one symmetrized, disconnected graph: $\delta = S(\Delta_2^+ \times \Delta_1)$. This graph is a coboundary:

$$\partial(\Delta_1 \times \Delta_1) = \tfrac{1}{3} S\big((\Delta_2^+ + \Delta_2^-) \times \Delta_1\big) = \tfrac{2}{3}\delta.$$

Case 3: $|\Delta| = 2$. There is, up to the ordering of vertices, a disconnected graph $\Delta_{2,2} \times \Delta_1$ and three connected graphs, listed below. (We choose the ordering of vertices that maximizes the order, and for a given order maximizes, for the lexicographic ordering, the set $E(\Delta)$ of edges of graphs Δ.)

$$\begin{aligned}
&\Delta_{3,2,1} \quad \text{with} \quad E(\Delta_{3,2,1}) = \{\overrightarrow{p_1 p_2}, \overrightarrow{p_1 p_3}\},\\
&\Delta_{3,2,2} \quad \text{with} \quad E(\Delta_{3,2,2}) = \{\overrightarrow{p_2 p_1}, \overrightarrow{p_1 p_3}\},\\
&\Delta_{3,2,3} \quad \text{with} \quad E(\Delta_{3,2,3}) = \{\overrightarrow{p_2 p_1}, \overrightarrow{p_3 p_1}\}.
\end{aligned}$$

After symmetrization, we get $S(\Delta_{2,2} \times \Delta_1) = 0$, $S(\Delta_{3,2,1})) = S(\Delta_{3,2,3}) = 0$ and

$$\text{symb } S(\Delta_{3,2,2}) = \tfrac{1}{6}\Delta_{3,2,2}, \quad o(S(\Delta_{3,2,2})) = \big((1,1), (1,0), (0,1)\big).$$

When we compute $\partial(S(D))$, we have to consider the blow-up of each vertex of each graph in $S(\Delta)$. If the vertex p has order $o = (\ell, r)$, we get a few graphs with two vertices p' and p" at the place of p; these vertices have order $o' = (\ell', r')$,

$o" = (\ell", r")$, with conditions

$$\ell' + r' \geq 2, \quad \ell'' + r'' \geq 2, \quad \ell' + \ell'' = \ell + 1, \quad r' + r'' = r + 1.$$

Then we look for $o(\partial\Delta)$. If $r > 0$, the maximal possible order among those $(o', o")$ is $\big((\ell+1, r-1), (0, 2)\big)$; if $r = 0$, it is $\big((\ell, r), (1, 1)\big) = \big((\ell, 0), (1, 1)\big)$.

Thus $o(\partial(S(D_{3,2,2}))) \leq \big((2, 0), (0, 2), (1, 0), (0, 1)\big)$; more precisely,

$$\text{symb } \partial(S(\Delta_{3,2,2})) = \tfrac{1}{6}\Delta', \quad E(\Delta') = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_4}, \overrightarrow{p_3p_2}\},$$

and, since there is only one graph in the symbol,

$$o(\partial(S(\Delta_{3,2,2}))) = \big((2, 0), (0, 2), (1, 0), (0, 1)\big).$$

No vector in this case is a cocycle; ∂ is an one-to-one mapping.

Case 4: $|\Delta| = 3$. From now on, all our graphs are connected. Repeating the argument of the preceding case, we get the following results:

They are, up to a permutation of vertices, four graphs:

$$\begin{aligned}
&\Delta_{3,3,1} \quad \text{with} \quad E(\Delta_{3,3,1}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_1}\},\\
&\Delta_{3,3,2} \quad \text{with} \quad E(\Delta_{3,3,2}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_2p_1}, \overrightarrow{p_3p_1}\},\\
&\Delta_{3,3,3} \quad \text{with} \quad E(\Delta_{3,3,3}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_3}\},\\
&\Delta_{3,3,4} \quad \text{with} \quad E(\Delta_{3,3,4}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_2p_3}, \overrightarrow{p_3p_1}\}.
\end{aligned}$$

Their symmetrizations do not vanish:

$$\begin{aligned}
o(S(\Delta_{3,3,1})) &= \big((2, 1), (1, 1), (0, 1)\big),\\
o(\partial(S(\Delta_{3,3,1}))) &= \big((3, 0), (1, 1), (0, 2), (0, 1)\big),\\
o(S(\Delta_{3,3,2})) &= \big((1, 2), (1, 1), (1, 0)\big),\\
o(\partial(S(\Delta_{3,3,2}))) &= \big((2, 1), (1, 1), (0, 2), (0, 1)\big),\\
o(S(\Delta_{3,3,3})) &= \big((2, 0), (1, 1), (0, 2)\big),\\
o(\partial(S(\Delta_{3,3,3}))) &= \big((2, 0), (2, 0), (0, 2), (0, 2)\big),\\
o(S(\Delta_{3,3,4})) &= \big((1, 1), (1, 1), (1, 1)\big),\\
o(\partial(S(\Delta_{3,3,4}))) &= \big((2, 0), (1, 1), (1, 1), (0, 2)\big).
\end{aligned}$$

Then ∂ is still a one-to-one mapping on that space of graphs.

Case 5: $|\Delta| = 4$. They are, up to a permutation of vertices, four graphs:

$$\begin{aligned}
&\Delta_{3,4,1} \quad \text{with} \quad E(\Delta_{3,4,1}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_1}, \overrightarrow{p_3p_1}\},\\
&\Delta_{3,4,2} \quad \text{with} \quad E(\Delta_{3,4,2}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_1}, \overrightarrow{p_2p_3}\},\\
&\Delta_{3,4,3} \quad \text{with} \quad E(\Delta_{3,4,3}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_1}, \overrightarrow{p_3p_2}\},\\
&\Delta_{3,4,4} \quad \text{with} \quad E(\Delta_{3,4,4}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_2p_1}, \overrightarrow{p_3p_1}, \overrightarrow{p_3p_2}\}.
\end{aligned}$$

Their symmetrizations do not vanish:

$$\begin{aligned}
o(S(\Delta_{3,4,1})) &= \big((2,2),(1,1),(1,1)\big),\\
o(\partial(S(\Delta_{3,4,1}))) &= \big((3,1),(1,1),(1,1),(0,2)\big),\\
o(S(\Delta_{3,4,2})) &= \big((2,1),(2,1),(0,2)\big),\\
o(\partial(S(\Delta_{3,4,2}))) &= \big((3,0),(2,1),(0,2),(0,2)\big),\\
o(S(\Delta_{3,4,3})) &= \big((2,1),(1,2),(1,1)\big),\\
o(\partial(S(\Delta_{3,4,3}))) &= \big((3,0),(1,2),(1,1),(0,2)\big),\\
o(S(\Delta_{3,4,4})) &= \big((1,2),(1,2),(2,0)\big),\\
o(\partial(S(\Delta_{3,4,4}))) &= \big((2,1),(1,2),(2,0),(0,2)\big).
\end{aligned}$$

Then ∂ is still a one-to-one mapping on that space of graphs.

Case 6: $|\Delta| = 5$. Up to a permutation of vertices, this space contains only one graph:

$$\Delta_{3,5,1} \quad \text{with} \quad E(\Delta_{3,5,1}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_1}, \overrightarrow{p_2p_3}, \overrightarrow{p_3p_1}\}.$$

Its symmetrization does not vanish,

$$\begin{aligned}
o(S(\Delta_{3,5,1})) &= \big((2,2),(2,1),(1,2)\big),\\
o(\partial(S(\Delta_{3,6,1}))) &= \big((3,1),(2,1),(1,2),(0,2)\big).
\end{aligned}$$

Then ∂ is still a one-to-one mapping on that space of graphs.

Case 7: $|\Delta| = 6$. In this last case, there is only one graph:

$$\Delta_{3,6,1} \quad \text{with} \quad E(\Delta_{3,6,1}) = \{\overrightarrow{p_1p_2}, \overrightarrow{p_1p_3}, \overrightarrow{p_2p_1}, \overrightarrow{p_2p_3}, \overrightarrow{p_3p_1}, \overrightarrow{p_3p_2}\}.$$

But its symmetrization does vanish.

This proves:

Proposition 6.1. *The three first spaces H^1, H^2 and H^3 of the Chevalley cohomology for graphs vanish.*

7. Canonical cocycles for the linear case

We first recall the construction of the relevant cocycles for the cohomology of the Lie algebra of vector fields $\mathscr{X}(\mathbb{R}^d)$ associated to the Lie derivative of smooth functions; see for instance [De Wilde and Lecomte 1983] for an explicit presentation of this cohomology.

A basis of the Lie algebra $\bigwedge^{\mathrm{inv}}\big(\mathfrak{gl}(d,\mathbb{R})\big)$ of multilinear, skewsymmetric, invariant forms on $\mathfrak{gl}(d,\mathbb{R})$ is given by

$$\zeta^{(j_1)}\wedge\cdots\wedge\zeta^{(j_q)} \quad \text{with } j_k \text{ odd and } j_1 < j_2\cdots < j_q < 2d,$$

where the $\zeta^{(j)}$ are the mappings

$$\zeta^{(j)}(A_1,\ldots,A_j) = \mathfrak{a}\big(\operatorname{Tr}(A_1\ldots A_j)\big).$$

Then, for each odd n, the linear form θ defined on $\bigwedge^n \mathscr{X}(\mathbb{R}^d)$ by

$$\theta(\xi_1,\xi_2,\ldots,\xi_n) = \zeta^{(n)}\big(\operatorname{Jac}(\xi_1),\ldots,\operatorname{Jac}(\xi_n)\big)$$

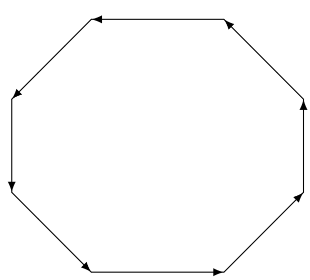
is a cocycle for the coboundary operator associated to the Lie derivative:

$$d\theta(\xi_0\ldots\ldots\xi_n) = \sum_{i=0}^{n}(-1)^i\mathscr{L}_{\xi_i}\theta(\xi_0\ldots\ldots\hat{\xi_i}\ldots\ldots\xi_n) + \frac{1}{2}\sum_{i\neq j}(-1)^{i+j}\theta\big([\xi_i,\xi_j]\,.\,\xi_0\ldots\widehat{\xi_i\xi_j}\ldots\ldots\xi_n\big).$$

This cocycle is not a coboundary; see [De Wilde and Lecomte 1983].

Let Ψ be an n-cochain on $T_{\mathrm{poly}}(\mathbb{R}^d)$ with values in the space $T_{\mathrm{poly}}(\mathbb{R}^d)^{-1}$ (that is, in $C^\infty(\mathbb{R}^d)$), and let ψ be its restriction to $\mathscr{X}(\mathbb{R}^d)$. Then the restriction of $\partial\Psi$ to $\mathscr{X}(\mathbb{R}^d)$ is exactly $d\psi$.

For instance, we consider the "wheel without an axis", the graph Δ of this form:



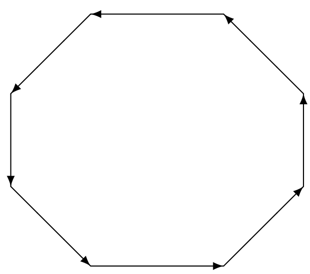
Denote by δ its symmetrization, which defines a cochain $\Psi = C_\delta$. By construction, on vector fields ξ_i, we get

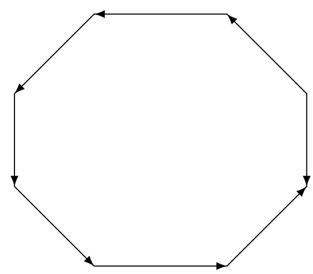
$$\begin{aligned}\psi(\xi_1 \cdot \ldots \cdot \xi_n) = \Psi(\xi_1 \cdot \ldots \cdot \xi_n) &= C_d(\xi_1 \cdot \ldots \cdot \xi_n) \\ &= \frac{1}{n!} \sum_{\sigma \in S_n} \varepsilon(\sigma) \partial_{i_n} \xi^{i_1}_{\sigma(1)} \partial_{i_1} \xi^{i_2}_{\sigma(2)} \ldots \partial_{i_{n-1}} \xi^{i_n}_{\sigma(n)} \\ &= \theta(\xi_1 \cdot \ldots \cdot \xi_n).\end{aligned}$$

Thus

$$\begin{aligned}C_{\partial\delta}(\xi_0 \cdot \ldots \cdot \xi_n) = \partial C_\delta(\xi_0 \cdot \ldots \cdot \xi_n) &= \partial\Psi(\xi_0 \cdot \ldots \cdot \xi_n) \\ &= d\theta(\xi_0 \cdot \ldots \cdot \xi_n) = 0.\end{aligned}$$

We now restrict ourselves to the space of linear polyvector fields. This is a subalgebra of $T_{\mathrm{poly}}(\mathbb{R}^d)$ equipped with the Schouten bracket; thus we can restrict our coboundary operator to cochains defined on this subalgebra. We get a new operator ∂_{lin}. Our previous computation tells us that the graphs happening in $\partial\delta$ are of the following forms:





For linear polyvector fields, only the first case appears. Then $B_{\partial_{\mathrm{lin}}(\delta)}(\alpha_0 \cdot \ldots \cdot \alpha_n)$ vanishes if one of the α_j is not a vector field. And

$$B_{\partial_{\mathrm{lin}}\delta}(\xi_0 \cdot \ldots \cdot \xi_n) = C_{\partial\delta}(\xi_0 \cdot \ldots \cdot \xi_n) = 0.$$

Since the mapping $\gamma \mapsto B_\gamma$ is one-to-one, $\partial_{\mathrm{lin}}\delta = 0$.

Now, suppose δ is a coboundary $d = \partial_{\mathrm{lin}}\beta$. Then β has $n-1$ vertices and $n-1$ edges. At each vertex there ends exactly one edge. If there is a vertex p from which no edge emanates, denote by $\overrightarrow{p'p}$ the edge ending at p. Since the graphs in β can be deduced from the graphs $\partial_{\mathrm{lin}}\beta$ only by proper reduction, there is no reduction at the vertex p, and in $\partial_{\mathrm{lin}}\beta$ there remains a unique edge $\overrightarrow{p'p}$. But there is no such graph in δ, so we can eliminate in β all the graphs with a vertex without emanating edges (we consider only "nonhanded" graphs). Now from each vertex of a graph in β, there is exactly one edge starting. As previously, the restriction of

$\partial\beta$ to the vector fields coincides with $\partial_{\mathrm{lin}}\beta$, and

$$\begin{aligned} dC_\beta(\xi_0 \cdot \ldots \cdot \xi_n) &= \partial C_\beta(\xi_0 \cdot \ldots \cdot \xi_n) = C_{\partial\beta}(\xi_0 \cdot \ldots \cdot \xi_n) \\ &= C_{\partial_{\mathrm{lin}} b}(\xi_0 \cdot \ldots \cdot \xi_n) = C_\delta(\xi_0 \cdot \ldots \cdot \xi_n) = \theta(\xi_0 \cdot \ldots \cdot \xi_n). \end{aligned}$$

This is impossible.

Thus any "wheel without an axis" Δ having an odd number of vertices gives rise to a canonical true cocycle for ∂_{lin}.

Remark. Suppose we want to build a linear formality $\mathcal{F}$ from the space of linear polyvector fields to the space of multidifferential operators. As we saw in Section 2, the obstruction to such a construction is a mapping φ, of degree 1, with $n \geq 4$ arguments. Such a mapping corresponds to purely aerial graphs with n vertices and $2n-3$ edges; in the linear case, we should have $2n-3 \leq n$, which is impossible. Every linear formality at order n can be extended to a linear formality.

References

[Arnal and Masmoudi 2002] D. Arnal and M. Masmoudi, "Cohomologie de Hochschild des graphes de Kontsevich", *Bull. Soc. Math. France* **130**:1 (2002), 49–69. MR 2003k:53122 Zbl 1028.46099

[Arnal et al. 2002] D. Arnal, D. Manchon, and M. Masmoudi, "Choix des signes pour la formalité de M. Kontsevich", *Pacific J. Math.* **203**:1 (2002), 23–66. MR 2003k:53123 Zbl 01818958

[De Wilde and Lecomte 1983] M. De Wilde and P. B. A. Lecomte, "Cohomology of the Lie algebra of smooth vector fields of a manifold, associated to the Lie derivative of smooth forms", *J. Math. Pures Appl.* (9) **62**:2 (1983), 197–214. MR 85j:17017 Zbl 0481.58032

[Gammella 2001] A. Gammella, *Déformations sur les variétés de Poisson et cohomologies appropriées*, Ph.D. thesis, Université de Metz, 2001.

[Kontsevich 1997] M. Kontsevich, "Formality conjecture", pp. 139–156 in *Deformation theory and symplectic geometry* (Ascona, 1996), edited by D. Sternheimer et al., Math. Phys. Stud. **20**, Kluwer Acad. Publ., Dordrecht, 1997. MR 98m:58044

[Kontsevich 2003] M. Kontsevich, "Deformation quantization of Poisson manifolds", *Lett. Math. Phys.* **66**:3 (2003), 157–216. MR 2062626

[Manchon and Torossian 2003] D. Manchon and C. Torossian, "Cohomologie tangente et cup-produit pour la quantification de Kontsevich", *Ann. Math. Blaise Pascal* **10**:1 (2003), 75–106. MR 2004f:53106 Zbl 1051.53072

Received December 20, 2002. Revised January 10, 2004.

DIDIER ARNAL
INSTITUT DE MATHÉMATIQUES DE BOURGOGNE
UNIVERSITÉ DE BOURGOGNE
U.F.R. SCIENCES ET TECHNIQUES B.P. 47870
F-21078 DIJON CEDEX
FRANCE

arnald@u-bourgogne.fr

ANGELA GAMMELLA
angela.gamella@wanadoo.fr

MOHSEN MASMOUDI
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ DE METZ
ILE DU SAULCY 57045
METZ CEDEX 01
FRANCE
masmoudi@poncelet.sciences.univ-metz.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

LIE ALGEBRAS AND GROWTH IN BRANCH GROUPS

LAURENT BARTHOLDI

We compute the structure of the Lie algebras associated to two examples of branch groups, and show that one has finite width while the other, the Gupta–Sidki group, has unbounded width and Lie algebra of Gelfand–Kirillov dimension $\log 3/\log(1+\sqrt{2})$.

We then draw a general result relating the growth of a branch group, of its Lie algebra, of its graded group ring, and of a natural homogeneous space we call *parabolic space*, namely the quotient of the group by the stabilizer of an infinite ray. The growth of the group is bounded from below by the growth of its graded group ring, which connects to the growth of the Lie algebra by a product-sum formula, and the growth of the parabolic space is bounded from below by the growth of the Lie algebra.

Finally we use this information to explicitly describe the normal subgroups of $\mathfrak{G}$, the *Grigorchuk group*. All normal subgroups are characteristic, and the number b_n of normal subgroups of $\mathfrak{G}$ of index 2^n is odd and satisfies $\limsup b_n/n^{\log_2 3} = 5^{\log_2 3}$, $\liminf b_n/n^{\log_2 3} = \frac{2}{9}$.

1. Introduction

The first purpose of this paper is to describe explicitly the Lie algebra associated to the Gupta–Sidki group $\ddot{\Gamma}$ [Gupta and Sidki 1983], and show in this way that this group is not of finite width (Corollary 3.9). We shall describe in Theorem 3.8 the Lie algebra as a graph, somewhat similar to a Cayley graph, in a formalism close to that introduced in [Bartholdi and Grigorchuk 2000a].

We shall then consider another group, Γ, and show in Corollary 3.14 that although many similarities exist between $\ddot{\Gamma}$ and Γ, the Lie algebra of Γ does have finite width.

These results follow from a description of group elements as branch portraits, exhibiting the relation between the group and its Lie algebra. They lead to the

The author acknowledges support from the Swiss National Fund for Scientific Research and the Hebrew University of Jerusalem.

MSC2000: 20F14, 20F40, 17B70, 16P90, 20E08.

Keywords: Lie algebra, growth of groups, lower central series.

notion of infinitely iterated wreath algebras, similar to wreath products of groups [Bartholdi $\geq$ 2005].

We shall show in Theorem 4.4 that, in the class of branch groups, the growth of the homogeneous space G/P (where P is a parabolic subgroup) is larger than the growth of the Lie algebra $\mathscr{L}(G)$. This result parallels a lower bound on the growth of G by that of its graded group ring $\overline{\Bbbk G}$ (Proposition 1.10).

Finally, we shall describe all the normal subgroups of the first Grigorchuk group, using the same formalism as that used to describe the lower central series. We confirm the description by Ceccherini et al. [2001] of the low-index normal subgroups of $\mathfrak{G}$. It turns out that all nontrivial normal subgroups are characteristic, and have finite index a power of 2. Call b_n the number of normal subgroups of index 2^n (Finite-index, not necessarily normal subgroups always have index a power of 2; this follows from G being a 2-torsion group.) Then there are 3^k+2 subgroups of index $2^{5\cdot 2^k+1}$ and $\frac{2}{9}3^k+1$ subgroups of index 2^{2^k+2}; these two values are extreme, in the sense that $b_n/n^{\log_2 3}$ has lower limit $5^{-\log_2 3}$ and upper limit $\frac{2}{9}$. Also, b_n is odd for all n (see Corollaries 5.4 and 5.5).

1.1. ***Philosophy.*** One can hardly exaggerate the importance of Lie algebras in the study of Lie groups. Lie subgroups correspond to subalgebras, normal subgroups correspond to ideals; simplicity, nilpotence and other properties match perfectly. This is due to the existence of mutually-inverse functions exp and log between a group and its algebra, and the Campbell–Hausdorff formula expressing the group operation in terms of the Lie bracket.

In the context of (discrete) p-groups and Lie algebras of characteristic p, the correspondence is not so perfect. First, in general, there is no exponential, and the best one can consider is the degree-1 truncations

$$\exp x = 1 + x + \mathbb{O}(x^2),$$
$$\log(1+x) = x + \mathbb{O}(x^2);$$

more terms would introduce denominators that in general are not invertible; and no reasonable definition of convergence can be imposed on $\mathbb{F}_p$. As a consequence, the group has to be subjected to a filtration to yield a Lie algebra. Then there is no perfect bijection between group and Lie-algebra objects.

However, the numerous results obtained in the area show that much can be gained from consideration of these imperfect algebras. To name a few, the theory of groups of finite width is closely related to the classification of finite p-groups (see [Leedham-Green 1994; Shalev and Zelmanov 1992]) and the theory of pro-p-groups is intimately Lie-algebraic; see [Shalev 1995a], [Shalev 1995b, §8] and [Klaas et al. 1997] with its bibliography. The solution to Burnside's problems by Efim Zelmanov relies also on Lie algebras. The results by Lev Kaloujnine on the

p-Sylow subgroups of $\mathfrak{S}_{p^n}$, even if in principle independent, can be restated in terms of Lie algebras in a very natural way (see Theorem 3.4).

In this paper, I argue that questions of growth, geometry and normal subgroup structure are illuminated by Lie-algebraic considerations.

1.2. ***Notation.*** We shall always write commutators as $[g, h] = g^{-1}h^{-1}gh$, conjugates as $g^h = h^{-1}gh$, and the adjoint operators $\mathsf{Ad}(g) = [g, -]$ and $\mathsf{ad}(x) = [x, -]$ on the group and Lie algebra respectively. $\mathfrak{S}_n$ is the symmetric group on n letters, and $\mathfrak{A}_n$ is the alternate subgroup of $\mathfrak{S}_n$. Polynomials and power series are all written over the formal variable $\hbar$, as is customary in the theory of quantum algebras. The Galois field with p elements is written $\mathbb{F}_p$. The cyclic group of order n is written C_n.

The lower central series of G is $\{\gamma_n(G)\}$, the lower p-central series is $\{P_n(G)\}$, the dimension series is $\{G_n\}$, the Lie dimension series is $\{L_n(G)\}$, and the derived series is $G^{(n)}$, and in particular $G' = [G, G]$ — the definitions shall be given below.

For $H \le G$, the subgroup of H generated by n-th powers of elements in H is written $\mho_n(H)$, and $H^{\times n}$ denotes the direct product of n copies of H, avoiding the ambiguous H^n. The normal closure of H in G is H^G.

Finally, $*$ stands for anything — something a speaker would abbreviate as "blah, blah, blah" in a talk. It is used to mean either that the value is irrelevant to the rest of the computation, or that it is the only unknown in an equation and therefore does not warrant a special name.

1.3. ***N-series.*** We first recall a classical construction of Magnus [1940], described for instance in [Lazard 1954] and [Huppert and Blackburn 1982, Chapter VIII].

Definition 1.1. Let G be a group. An *N-series* is a series $\{H_n\}$ of normal subgroups with $H_1 = G$, $H_{n+1} \le H_n$ and $[H_m, H_n] \le H_{m+n}$ for all $m, n \ge 1$. The associated Lie ring is

$$\mathscr{L}(G) = \bigoplus_{n=1}^{\infty} \mathscr{L}_n,$$

with $\mathscr{L}_n = H_n/H_{n+1}$ and the bracket operation $\mathscr{L}_n \otimes \mathscr{L}_m \to \mathscr{L}_{m+n}$ induced by commutation in G.

For p a prime, an *N_p-series* is an N-series $\{H_n\}$ such that $\mho_p(H_n) \le H_{pn}$, and the associated Lie ring is a restricted Lie algebra over $\mathbb{F}_p$:

$$\mathscr{L}_{\mathbb{F}_p}(G) = \bigoplus_{n=1}^{\infty} \mathscr{L}_n,$$

with the p-mapping $\mathscr{L}_n \to \mathscr{L}_{pn}$ induced by raising to the power p in H_n.

We recall that $\mathscr{L}$ is a *restricted* Lie algebra (see [Jacobson 1941] or [Strade and Farnsteiner 1988, Section 2.1]) if it is over a field $\Bbbk$ of characteristic p, and there exists a mapping $x \mapsto x^{[p]}$ such that $\operatorname{ad} x^{[p]} = \operatorname{ad}(x)^p$, $(\alpha x)^{[p]} = \alpha^p x^{[p]}$ and $(x+y)^{[p]} = x^{[p]} + y^{[p]} + \sum_{i=1}^{p-1} s_i(x,y)$, where the s_i are obtained by expanding $\operatorname{ad}(x \otimes \hbar + y \otimes 1)^{p-1}(a \otimes 1) = \sum_{i=1}^{p-1} s_i(x,y) \otimes i\hbar^{i-1}$ in $\mathscr{L} \otimes \Bbbk[\hbar]$. Equivalently:

Proposition 1.2 (Jacobson). *Let (e_i) be a basis of $\mathscr{L}$ such that, for some $y_i \in \mathscr{L}$, we have $\operatorname{ad}(e_i)^p = \operatorname{ad}(y_i)$. Then $\mathscr{L}$ is restricted; more precisely, there exists a unique p-mapping such that $e_i^{[p]} = y_i$.*

The standard examples of an N-series are the *lower central series*, $\{\gamma_n(G)\}_{n=1}^{\infty}$, given by $\gamma_1(G) = G$ and $\gamma_n(G) = [G, \gamma_{n-1}(G)]$, and the *lower exponent-p central series* or *Frattini series* given by $P_1(G) = G$ and

$$P_n(G) = [G, P_{n-1}(G)]\,\mho_p(P_{n-1}(G)).$$

The Frattini series differs from the lower central series in that its successive quotients are all elementary p-groups.

The standard example of an N_p-series is the *dimension series*, also known as the p-lower central, Zassenhaus [1940], Jennings [1941], Lazard [1954] or Brauer series, given by $G_1 = G$ and $G_n = [G, G_{n-1}]\,\mho_p(G_{\lceil n/p \rceil})$, where $\lceil n/p \rceil$ is the least integer no less than n/p. It can alternatively be described, by a result of Lazard [1954], as

$$G_n = \prod_{i \cdot p^j \geq n} \mho_{p^j}(\gamma_i(G)),$$

or as

$$G_n = \{g \in G \mid g - 1 \in \varpi^n\},$$

where ϖ is the augmentation (or fundamental) ideal of the group algebra $\mathbb{F}_p G$. Note that this last definition extends to characteristic 0, giving a graded Lie algebra $\mathscr{L}_{\mathbb{Q}}(G)$ over $\mathbb{Q}$. In that case, the subgroup G_n is the isolator of $\gamma_n(G)$:

$$G_n = \sqrt{\gamma_n(G)} = \bigl\{g \in G \mid \langle g \rangle \cap \gamma_n(G) \neq \{1\}\bigr\}.$$

A good reference for these results is [Passi 1979, Chapter VIII].

We mention finally for completeness another N_p-series, the *Lie dimension series*

$$L_n(G) = \{g \in G \mid g - 1 \in \varpi^{(n)}\},$$

where $\varpi^{(n)}$ is the n-th Lie power of $\varpi \leq \Bbbk G$, given by $\varpi^{(1)} = \varpi$ and $\varpi^{(n+1)} = [\varpi^{(n)}, \varpi] = \{xy - yx \mid x \in \varpi^{(n)}, y \in \varpi\}$. As shown in [Passi and Sehgal 1975],

$$L_n(G) = \prod_{(i-1)\cdot p^j \geq n} \mho_{p^j}(\gamma_i(G))$$

if $\Bbbk$ is of characteristic p, and

$$L_n(G) = \sqrt{\gamma_n(G)} \cap [G, G]$$

if $\Bbbk$ is of characteristic 0.

In the sequel we will only consider the N-series $\{\gamma_n(G)\}$ and $\{P_n(G)\}$ and the N_p-series $\{G_n\}$ of dimension subgroups. We reserve the symbols $\mathscr{L}$ and $\mathscr{L}_{\mathbb{F}_p}$ for their respective Lie algebras.

Definition 1.3. Let $\{H_n\}$ be an N-series for G. The *degree* of $g \in G$ is the maximal $n \in \mathbb{N} \cup \{\infty\}$ such that g belongs to H_n.

Recall that the rank of an abelian group A is the minimal number of elements that generate A. A series $\{H_n\}$ has *finite width* if there is a constant W such that $\ell_n := \operatorname{rank}[H_n : H_{n+1}] \leq W$ for all n. A group has *finite width* if its lower central series has finite width; this definition comes from [Klaas et al. 1997].

Definition 1.4. Let $a = \{a_n\}$ and $b = \{b_n\}$ be sequences of real numbers. We write $a \precsim b$ if there is an integer $C > 0$ such that $a_n < Cb_{Cn+C} + C$ for all $n \in \mathbb{N}$, and write $a \sim b$ if $a \precsim b$ and $b \precsim a$.

In the sense of this definition, a group has finite width if and only if $\{\ell_n\} \sim \{1\}$.

Question 1. If the rank of $\gamma_n(G)/\gamma_{n+1}(G)$ is bounded, does it follow that the rank of G_n/G_{n+1}, $P_n(G)/P_{n+1}(G)$ or $L_n(G)/L_{n+1}(G)$ is bounded? How about a converse?

More generally, say an N-series $\{H_n\}$ has *finite width* if $\operatorname{rank}(H_n/H_{n+1})$ is bounded over $n \in \mathbb{N}$. If G has a finite-width N-series intersecting to $\{1\}$, are all N-series of G of finite width?

I do not know the answer to these natural questions.

The following result is well-known, and shows that sometimes the Lie ring $\mathscr{L}(G)$ is actually a Lie algebra over $\mathbb{F}_p$.

Lemma 1.5. *Let G be a group generated by a set S. Let $\mathscr{L}(G)$ be the Lie ring associated to the lower central series.*

(1) *If S is finite, $\mathscr{L}_n$ is a finite-rank $\mathbb{Z}$-module for all n.*

(2) *If there is a prime p such that all generators $s \in S$ have order p, then $\mathscr{L}_n$ is a vector space over $\mathbb{F}_p$ for all n. It follows that the Frattini series (for that prime p) and the lower central series coincide.*

Proof. First, $\mathscr{L}_1$ is generated by $\bar{S}$, the image of S in G/G'. Since $\mathscr{L}$ is generated by $\mathscr{L}_1$, in particular $\mathscr{L}_n$ is generated by the finitely many $(n-1)$-fold products of elements of $\bar{S}$; this proves the first point.

Actually, far fewer generators are required for $\mathscr{L}_n$; in the extremal case when G is a free group, a basis of $\mathscr{L}_n$ is given in terms of "standard monomials" of degree n. See Section 3.2 or [Hall 1950].

For the second claim, assume more generally that $s^p \in G'$ for all $s \in S$, so that G/G' is an $\mathbb{F}_p$-vector space. We use the identity $[x, y]^p \equiv [x, y^p] \bmod \gamma_3\langle x, y\rangle$, due to Philip Hall. Let $g = [x, y]$ be a generator of $\gamma_n(G)$, with $x \in G$ and $y \in$

$\gamma_{n-1}(G)$. Then $y^p \in \gamma_n(G)$ by induction, so $g^p \in \gamma_{n+1}(G)$ and $\mathscr{L}_n$ is an $\mathbb{F}_p$-vector space. □

Anticipating, we note that the groups $\ddot{\Gamma}$ and Γ we shall consider satisfy these hypotheses for $p = 3$, and $\mathfrak{G}$ satisfies them for $p = 2$.

1.4. ***Growth of groups and vector spaces.*** Let G be a group generated by a finite set S. The *length* $|g|$ of an element $g \in G$ is the minimal number n such that g can be written as $s_1 \dots s_n$ with $s_i \in S$. The *growth series* of G is the formal power series

$$\mathsf{growth}(G) = \sum_{g \in G} \hbar^{|g|} = \sum_{n \geq 0} f_n \hbar^n,$$

where $f_n = \#\{g \in G \mid |g| = n\}$. The *growth function* of G is the $\sim$-equivalence class of the sequence $\{f_n\}$. Note that although $\mathsf{growth}(G)$ depends on S, this equivalence class is independent of the choice of S.

Let X be a transitive G-set and $x_0 \in X$ be a fixed base point. The *length* $|x|$ of an element $x \in X$ is the minimal length of a $g \in G$ moving x_0 to x. The *growth series* of X is the formal power series

$$\mathsf{growth}(X, x_0) = \sum_{x \in X} \hbar^{|x|} = \sum_{n \geq 0} f_n \hbar^n,$$

where $f_n = \{x \in X \mid \min_{gx_0 = x} |g| = n\}$. The *growth function* of X is the equivalence class under $\sim$ of the sequence $\{f_n\}$. It is again independent of the choice of x_0 and of generators of G.

Let $V = \bigoplus_{n \geq 0} V_n$ be a graded vector space. The *Hilbert–Poincaré* series of V is the formal power series

$$\mathsf{growth}(V) = \sum_{n \geq 0} v_n \hbar^n = \sum_{n \geq 0} \dim V_n \hbar^n.$$

We return to the dimension series of G. Consider the graded algebra

$$\overline{\mathbb{F}_p G} = \bigoplus_{n=0}^{\infty} \varpi^n / \varpi^{n+1}.$$

Here a fundamental result connecting $\mathscr{L}_{\mathbb{F}_p}(G)$ and $\overline{\mathbb{F}_p G}$:

Theorem 1.6 [Quillen 1968]. *$\overline{\mathbb{F}_p G}$ is the restricted enveloping algebra of the Lie algebra $\mathscr{L}_{\mathbb{F}_p}(G)$ associated to the dimension series.*

The Poincaré–Birkhoff–Witt Theorem then gives a basis of $\overline{\mathbb{F}_p G}$ consisting of monomials over a basis of $\mathscr{L}_{\mathbb{F}_p}(G)$, with exponents at most $p - 1$. Therefore:

Proposition 1.7 [Jennings 1941]. *Let G be a group, and let $\sum_{n\geq 1} \ell_n \hbar^n$ be the Hilbert–Poincaré series of $\mathscr{L}_{\mathbb{F}_p}(G)$. Then*

$$\mathsf{growth}(\overline{\mathbb{F}_p G}) = \prod_{n=1}^{\infty} \left(\frac{1-\hbar^{pn}}{1-\hbar^n} \right)^{\ell_n}.$$

Approximations from analytical number theory [Li 1996] and complex analysis then give:

Proposition 1.8 [Bartholdi and Grigorchuk 2000a, Proposition 2.2; Petrogradsky 1999, Theorem 2.1]. *Let G be a group and expand the power series*

$$\mathsf{growth}(\mathscr{L}_{\mathbb{F}_p}(G)) = \sum_{n\geq 1} \ell_n \hbar^n \quad \text{and} \quad \mathsf{growth}(\overline{\mathbb{F}_p G}) = \sum_{n\geq 0} f_n \hbar^n.$$

(1) *$\{f_n\}$ grows exponentially if and only if $\{\ell_n\}$ does, and*

$$\limsup_{n\to\infty} \frac{\ln \ell_n}{n} = \limsup_{n\to\infty} \frac{\ln f_n}{n}.$$

(2) *If $\ell_n \sim n^d$, then $f_n \sim e^{n^{(d+1)/(d+2)}}$.*

The Lie algebras we consider have polynomial growth, i.e., finite Gelfand–Kirillov dimension. This notion is more commonly studied for associative rings [Gelfand and Kirillov 1966]:

Definition 1.9. The *Gelfand–Kirillov dimension* of a graded Lie algebra $\mathscr{L} = \bigoplus \mathscr{L}_n$ is

$$\dim_{GK}(\mathscr{L}) = \limsup_{n\to\infty} \frac{\log(\dim \mathscr{L}_1 + \cdots + \dim \mathscr{L}_n)}{\log n}.$$

If $\ell_n \sim n^d$, then $\mathscr{L}$ has Gelfand–Kirillov dimension $d+1$. However, the converse is not true, since the sequence $\log(\ell_1 + \cdots + \ell_n)/\log n$ need not converge. If the group G has finite width, its algebra $\mathscr{L}(G)$ has Gelfand–Kirillov dimension 1.

Note also that if A is any algebra generated in degree 1, then $\dim_{GK}(A) = 0$ or $\dim_{GK}(A) \geq 1$. Furthermore, George Bergman [1978] has shown that if A is associative, then $\dim_{GK}(A) = 1$ or $\dim_{GK}(A) \geq 2$; see [Krause and Lenagan 1985, Theorem 2.5] for a proof. Victor Petrogradsky [1997] showed that there exist Lie algebras of any Gelfand–Kirillov dimension ≥ 1.

Finally, we recall a connection between the growth of G and that of $\overline{\mathbb{F}_p G}$. We use the notation $\sum f_n \hbar^n \geq \sum g_n \hbar^n$ to mean $f_n \geq g_n$ for all $n \in \mathbb{N}$.

Proposition 1.10 [Grigorchuk 1989, Lemma 8]. *Let G be a group generated by a finite set S. Then*

$$\frac{\mathsf{growth}(G)}{1-\hbar} \geq \mathsf{growth}(\overline{\Bbbk G}).$$

2. Branch groups

Branch groups were introduced by Rostislav Grigorchuk [2000], where he developed a general theory of groups acting on rooted trees. We shall content ourselves with a restricted definition; recall that $G \wr \mathfrak{S}_d$ is the *wreath product* $G^{\times d} \rtimes \mathfrak{S}_d$, the action of $\mathfrak{S}_d$ on the direct product induced by the permutation action of $\mathfrak{S}_d$ on $\Sigma = \{1, \dots, \mathsf{d}\}$.

Definition 2.1. A group G is a *regular branch group* if for some $d \in \mathbb{N}$ there is

(1) an embedding $\psi : G \hookrightarrow G \wr \mathfrak{S}_d$ such that the image of $\psi(G)$ in $\mathfrak{S}_d$ acts transitively on Σ. Define for $n \in \mathbb{N}$ the subgroups $\mathsf{Stab}_G(n)$ of G by $\mathsf{Stab}_G(0) = G$, and inductively

$$\mathsf{Stab}_G(n) = \psi^{-1}(\mathsf{Stab}_G(n-1)^{\times d})$$

where $\mathsf{Stab}_G(n-1)^{\times d}$ is seen as a subgroup of $G \wr \mathfrak{S}_d$. One requires then that $\bigcap_{n\in\mathbb{N}} \mathsf{Stab}_G(n) = \{1\}$;

(2) a subgroup $K \le G$ of finite index with $\psi(K) \le K^{\times d}$.

To avoid ambiguous bracket notations, we write the decomposition map

$$\psi(g) = \ll g_1, \dots, g_d \gg \pi,$$

with π expressed as a permutation in disjoint cycle notation.

We shall abbreviate "regular branch group" to "branch group", since all branch groups in this paper are actually regular branch. We shall usually omit d from the description, and say that "G branches over K".

Lemma 2.2. *If G is a branch group, then G branches over a subgroup K of G such that K is normal in G, and $K^{\times d}$ is normal in $\psi(K)$.*

Proof. Let G be branch over L of finite index, and set $K = \bigcap_{g\in G} L^g$, the *core* of L. Then obviously $L \lhd G$; and since $(L^{\times d})^{\psi g} \le \psi(K^g)$ for all $g \in G$, we have, writing $\psi(g) = \ll g_1, \dots, g_d \gg \pi$,

$$K^{\times d} \le \bigcap_{g\in G} (L^{g_{1\pi}} \times \cdots \times L^{g_{d\pi}}) = \bigcap_{g\in G} (L^{\times d})^{\psi g} \le K,$$

and $(K^{\times d})^{\psi(g)} = K^{g_{1\pi}} \times \cdots \times K^{g_{d\pi}} = K^{\times d}$, so $K^{\times d} \lhd \psi(G)$. □

Let G be a branch group, with d, Σ and K as in the definition. The *rooted tree* on Σ is the free monoid Σ^*, with root the empty sequence $\varnothing$; it is a metric space for the distance

$$\operatorname{dist}(\sigma, \tau) = |\sigma| + |\tau| - 2 \max \{n \in \mathbb{N} \mid \sigma_n = \tau_n\}.$$

The *natural action* of G is an action on Σ^* defined inductively by

$$(\sigma_1\sigma_2\dots\sigma_n)^g = (\sigma_1)^\pi(\sigma_2\dots\sigma_n)^{g_{\sigma_1}} \text{ for } \sigma_1,\dots,\sigma_n\in\Sigma, \tag{2–1}$$

where $\psi(g) = \ll g_1,\dots,g_d\gg\pi$. By the condition $\bigcap \mathsf{Stab}_G(n) = \{1\}$, this action is faithful and G is residually finite. Note that $\mathsf{Stab}_G(n)$ is the fixator of Σ^n in this action.

Note that the action (2–1) gives a geometrical meaning to the branch structure of G that closely parallels the structure of the tree Σ^*. Indeed one may consider G as a group acting on the tree Σ^*; then the choice of a vertex σ of Σ^* and of a subgroup J of K determines a subgroup L_σ of K, namely the group of tree automorphisms of Σ^* that fix $\Sigma^*\setminus\sigma\Sigma^*$ and whose action on $\sigma\Sigma^*$ is that of an element of J on Σ^*. The choice of a subgroup J_σ for all $\sigma\in\Sigma^*$ determines a subgroup M of K, namely the closure of the L_σ associated to σ and J_σ when σ ranges over Σ^*.

This geometrical vision can also give pictorial descriptions of group elements:

Definition 2.3. Suppose G branches over K; let T be a transversal of K in G, and let U be a transversal of $\psi^{-1}(K^{\times d})$ in K. The *branch portrait* of an element $g\in G$ is a labeling of Σ^*, as follows: the root vertex $\varnothing$ is labeled by an element of TU, and all other vertices are labeled by an element of U.

Given $g\in G$, write first $g = kt$ with $k\in K$ and $t\in T$, then write

$$k = \psi^{-1}(k_1,\dots,k_{\mathsf{d}})u_\varnothing,$$

and inductively $k_\sigma = \psi^{-1}(k_{\sigma 1},\dots,k_{\sigma\mathsf{d}})u_\sigma$ for all $\sigma\in\Sigma^*$. Label the root vertex by $tu_\varnothing$ and then label each vertex $\sigma\neq\varnothing$ by u_σ.

There are uncountably many branch portraits, even for a countable branch group. We therefore introduce the following notion:

Definition 2.4. Let G be a branch group. Its *completion* $\overline{G}$ is the inverse limit

$$\operatorname*{proj\,lim}_{n\to\infty} G/\mathsf{Stab}_G(n).$$

This is also the closure in $\mathsf{Aut}\,\Sigma^*$ of G seen through its natural action (2–1).

Since $\overline{G}$ is closed in $\mathsf{Aut}\,\Sigma^*$ it is a profinite group, and thus is compact, and totally disconnected. If G has the *congruence subgroup property* [Grigorchuk 2000], meaning that all finite-index subgroups of G contain $\mathsf{Stab}_G(n)$ for some n, then $\overline{G}$ is also the profinite completion of G.

Lemma 2.5. *Let G be a branch group and $\overline{G}$ its completion. Then Definition 2.3 yields a bijection between the set of branch portraits and $\overline{G}$.*

We shall often simplify notation by omitting ψ from subgroup descriptions, as in statements like $\mathsf{Stab}_G(n)\le\mathsf{Stab}_G(n-1)^{\times d}$.

2.1. ***The Grigorchuk group*** $\mathfrak{G}$**.** We shall consider more carefully three examples of branch groups in the sequel. The first example of a branch group was considered by Grigorchuk in 1980, and has appeared innumerable times in recent mathematics — the entire chapter VIII of [de la Harpe 2000] is devoted to it. It is defined as follows: it is a 4-generated group $\mathfrak{G}$ (with generators a, b, c, d), its map ψ is given by

$$\psi : \mathfrak{G} \hookrightarrow (\mathfrak{G} \times \mathfrak{G}) \rtimes \mathfrak{S}_2,$$
$$a \mapsto \langle\!\langle 1, 1\rangle\!\rangle(1,2), \quad b \mapsto \langle\!\langle a, c\rangle\!\rangle, \quad c \mapsto \langle\!\langle a, d\rangle\!\rangle, \quad \mapsto \langle\!\langle 1, b\rangle\!\rangle,$$

and its subgroup K is the normal closure of $[a, b]$, of index 16. Grigorchuk [1980; 1983] proved that $\mathfrak{G}$ is an intermediate-growth, infinite-torsion group. Its lower central series was computed in [Bartholdi and Grigorchuk 2000a], along with a description of its Lie algebra. We shall reproduce that result using a more general method.

2.2. ***The Gupta–Sidki group*** $\ddot{\Gamma}$**.** This 2-generated group was introduced by Narain Gupta and Said Sidki in [Gupta and Sidki 1983], where they proved it to be an infinite torsion group. Later Sidki obtained a complete description of its automorphism group [Sidki 1987], along with information on its subgroups. It is a branch group with generators a, t, its map ψ is given by

$$\psi : \ddot{\Gamma} \hookrightarrow (\ddot{\Gamma} \times \ddot{\Gamma} \times \ddot{\Gamma}) \rtimes \mathfrak{A}_3,$$
$$a \mapsto \langle\!\langle 1, 1, 1\rangle\!\rangle(1,2,3), \quad t \mapsto \langle\!\langle a, a^{-1}, t\rangle\!\rangle,$$

and its subgroup K is $\ddot{\Gamma}'$, of index 9.

It was recently proved in [Bartholdi 2000] that $\ddot{\Gamma}$ has intermediate growth, which increases its analogy with the Grigorchuk group mentioned above. An outstanding question was whether $\ddot{\Gamma}$ has finite width. Ana Cristina Vieira [1998; 1999] computed the first 9 terms of the lower central series and showed that there are all of rank at most 2. We shall shortly see, however, that $\ddot{\Gamma}$ has unbounded width.

The following lemma is straightforward:

Lemma 2.6. $\ddot{\Gamma}'/(\ddot{\Gamma}' \times \ddot{\Gamma}' \times \ddot{\Gamma}')$ *is isomorphic to* $C_3 \times C_3$, *generated by* $c = [a, t]$ *and* $u = [a, c]$.

Note finally that the notations in [Sidki 1987] are slightly different: his x is our a, and his y is our t. In [Vieira 1998] her $y^{[1]}$ is our u, and more generally her g_1 is our $\mathbb{0}(g)$ and her $g^{[1]}$ is our $\mathbb{2}(g)$. In [Bartholdi and Grigorchuk 2002], where a great deal of information on $\ddot{\Gamma}$ is gathered, the group is called $\overset{\equiv}{\Gamma}$.

2.3. ***The Fabrykowski–Gupta group*** Γ**.** This other group is at first sight close to $\ddot{\Gamma}$: it is also a branch group, generated by two elements a, t. Its map ψ is given by

$$\psi : \Gamma \hookrightarrow (\Gamma \times \Gamma \times \Gamma) \rtimes \mathfrak{A}_3,$$
$$a \mapsto \ll 1, 1, 1 \gg (1, 2, 3), \quad t \mapsto \ll a, 1, t \gg,$$

and its subgroup K is Γ', of index 9.

This group was first considered in [Fabrykowski and Gupta 1991], where its growth was studied. In [Bartholdi and Grigorchuk 2002] it was proved that it is a branch group, and that its subgroup $L = \langle at, ta \rangle$ has index 3 and is torsion-free. In [Bartholdi 2000] another proof of the subexponential growth of Γ is given.

3. Lie algebras

We now describe the Lie algebras associated to the groups $\mathfrak{G}$, $\ddot{\Gamma}$ and Γ defined in the previous section. We start by considering a group G, and make the following hypotheses on G, which will be satisfied by $\mathfrak{G}$, $\ddot{\Gamma}$ and Γ:

(1) G is finitely generated by a set S;

(2) there is a prime p such that all $s \in S$ have order p.

Under these conditions, it follows from Lemma 1.5 that $\gamma_n(G)/\gamma_{n+1}(G)$ is a finite-dimensional vector space over $\mathbb{F}_p$, and therefore that $\mathscr{L}(G)$ is a Lie algebra over $\mathbb{F}_p$ that is finite at each dimension. Clearly the same property holds for the restricted algebra $\mathscr{L}_{\mathbb{F}_p}(G)$.

We propose the following notation for such algebras:

Definition 3.1. Let

$$\mathscr{L} = \bigoplus_{n \geq 1} \mathscr{L}_n$$

be a graded Lie algebra over $\mathbb{F}_p$, and choose a basis B_n of $\mathscr{L}_n$ for all $n \geq 1$. For $x \in \mathscr{L}_n$ and $b \in B_n$ denote by $\langle x | b \rangle$ the b-coefficient of x in the basis B_n.

The *Lie graph* associated to these choices is an abstract graph. Its vertex set is $\bigcup_{n \geq 1} B_n$, and each vertex $x \in B_n$ has degree n. Its edges are labeled as αx, with $x \in B_1$ and $\alpha \in \mathbb{F}_p$, and may only connect a vertex of degree n to a vertex of degree $n + 1$. For all $x \in B_1$, $y \in B_n$ and $z \in B_{n+1}$, there is an edge labeled $\langle [x, y] | z \rangle x$ from y to z.

If $\mathscr{L}$ is a restricted algebra of $\mathbb{F}_p$, there are additional edges, labeled $\alpha \cdot p$ with $\alpha \in \mathbb{F}_p$, from vertices of degree n to vertices of degree pn. For all $x \in B_n$ and $y \in B_{pn}$, there is an edge labeled $\langle x^p | y \rangle \cdot p$ from x to y.

Edges labeled $0x$ are naturally omitted. Edges labeled $1x$ are simply written x.

There is some analogy between this definition and that of a Cayley graph — this topic will be developed in Section 4. The generators (in the Cayley sense) are simply chosen to be the $\mathtt{ad}(x)$ with x running through B_1, a basis of $G/[G, G]$.

A presentation for the $\mathscr{L}$ can also be read off its Lie graph. For every n, consider the set $\mathscr{W}$ of all words of length n over B_1. For a path π in the Lie graph, define its weight as the product of the labels on its edges. Each $w \in \mathscr{W}$ defines an element of $\mathscr{L}_n$, by summing the weights of all paths labeled w in the Lie graph. Let $\mathscr{R}_n$ be the set of all linear dependence relations among these words. Then $\mathscr{L}$ admits a presentation by generators and relations as

$$\mathscr{L} = \langle B_1 \mid \mathscr{R}_1, \mathscr{R}_2, \dots \rangle.$$

We give a few examples of Lie graphs. First, if G is abelian, its Lie graph has $\operatorname{rank}(G)$ vertices of weight 1 and no other vertices. If G is the quaternion group $Q_8 = \{\pm 1, \pm i, \pm j, \pm k\}$, its Lie ring is an algebra over $\mathbb{F}_2$, and the Lie graph of $\mathscr{L}(Q_8) = \mathscr{L}_{\mathbb{F}_2}(Q_8)$ is

$$\begin{tikzcd} i \arrow[dr, "j"] & \\ & -1 \\ j \arrow[ur, "i"] & \end{tikzcd}$$

3.1. ***The infinite dihedral group.*** As another example, let G be the infinite dihedral group $D_\infty = \langle a, b \mid a^2, b^2 \rangle$. Then $\gamma_n(G) = \left\langle (ab)^{2^{n-1}} \right\rangle$ for all $n \geq 2$, and its Lie ring is again a Lie algebra over $\mathbb{F}_2$, with Lie graph

$$\begin{tikzcd} a \arrow[dr, "b"] & & & \\ & (ab)^2 \arrow[r, "{a,b}"] & (ab)^4 \arrow[r, "{a,b}"] & (ab)^8 \arrow[r, dotted, "{a,b}"] & {} \\ b \arrow[ur, "a"'] & & & \end{tikzcd}$$

The lower 2-central series of G is different: $G_{2^n} = G_{2^n+1} = \cdots = G_{2^{n+1}-1} = \gamma_{n+1}(G)$, so the Lie graph of $\mathscr{L}_{\mathbb{F}_2}(G)$ is

$$\begin{tikzcd} a \arrow[dr, "b"] & & & \\ & (ab)^2 \arrow[r, "\cdot 2"] & (ab)^4 \arrow[r, "\cdot 2"] & (ab)^8 \arrow[r, dotted, "\cdot 2"] & {} \\ b \arrow[ur, "a"'] & & & \end{tikzcd}$$

3.2. ***The free group.*** Consider, as an example producing exponential growth, the free group F_r and its Lie algebra $\mathscr{L}$; this is a free Lie algebra of rank r. Using Theorem 1.6 and Möbius inversion, we get

$$\dim_{\mathbb{Q}}\big(\gamma_n(F_r)/\gamma_{n+1}(F_r)\otimes\mathbb{Q}\big)=\#\{u\in\mathcal{M}\mid\deg u=n\}=\frac{1}{n}\sum_{d\mid n}\mu_{n/d}r^d\precsim r^n,$$

where μ is the Möbius function; therefore $\mathsf{growth}(\overline{\mathbb{Q}F_r})\le 1/(1-r\hbar)$. Recall that

$$\mathsf{growth}(F_r)=\frac{1+\hbar}{1-(2r-1)\hbar},$$

so the group growth rate can be strictly larger than the algebra growth rate in Proposition 1.10.

It is an altogether different story to find explicitly a basis of $\mathscr{L}$. Pick a basis X of F_r; its image in $\mathscr{L}_1\cong\mathbb{Z}^r$ is a generating set of $\mathscr{L}$, still written X. A *Hall set* is a linearly ordered set of nonassociative words $\mathcal{M}$ with $X\subset\mathcal{M}$ and

$$[u,v]\in\mathcal{M}\text{ if and only if }u<v\in\mathcal{M}\text{ and }(u\in X\text{ or }u=[p,q],\,q\ge v);$$

furthermore one requires $[u,v]<v$. Note that an order on the nonassociative words uniquely defines a corresponding Hall set.

There are many examples of Hall sets, and for each Hall set $\mathcal{M}$ the set $\{u\in\mathcal{M}\mid |u|=r\}$ is a basis of the abelian group $\gamma_n(F_r)/\gamma_{n+1}(F_r)$. For example, the *Hall basis* [Hall 1950] is the linearly ordered set $\mathcal{M}$ having as maximal elements X in an arbitrary order, and such that $u<v$ in $\mathcal{M}$ whenever $\deg u>\deg v$. It contains then all $[x,y]$ with $x,y\in X$ and $x>y$; then all $\big[[u,v],w\big]$ whenever $[u,v]<w\le v$ and $u,v,w\in\mathcal{M}$.

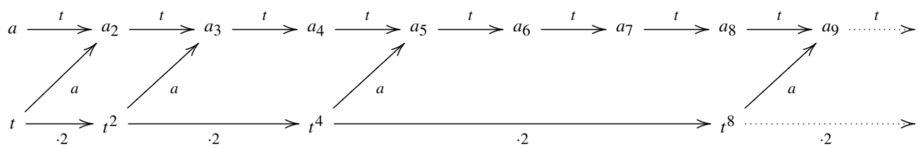
Another basis, more computationally efficient (it is a Lie algebra equivalent of Gröbner bases), is the *Lyndon–Shirshov basis* [Širšov 1962; Lothaire 1990; Reutenauer 1993]. It is defined as follows: order X arbitrarily; on the free monoid X^* put the lexicographical ordering: $u\le uv$, and $uxv<uyw$ for all $u,v,w\in X^*$ and $x<y\in X$. A nonempty word $w\in X^*$ is a *Lyndon–Shirshov word* if for any nontrivial factorization $w=uv$ we have $w<v$. If furthermore we insist that v be $<$-minimal, then u and v are again Lyndon–Shirshov words. For a Lyndon–Shirshov word w, define its *bracketing* $B(w)$ inductively as follows: if $w\in X$ then $B(w)=w$. If $w=uv$ with v minimal then $B(w)=[B(u),B(v)]$. Then $\{B(w)\}$ is a basis of $\mathscr{L}$.

From our perspective, an optimal basis B would consist only of left-ordered commutators, and be prefix-closed, i.e., be such that $[u,x]\in B$ implies $u\in B$; then indeed the Lie algebra structure of an arbitrary Lie algebra would be determined $\mathsf{ad}(u)$ for all $u\in B$, and therefore would be a tree in the case of a free Lie algebra. Kukin announced in [Kukin 1978] a construction of such bases, but his proof does not appear to be altogether complete [Blessenohl and Laue 1993], and the problem of construction of a left-ordered basis seems to be considered open.

3.3. ***The lamplighter group.*** As another example, consider the *lamplighter group* $G = C_2 \wr \mathbb{Z}$, with a generating C_2 and t generating $\mathbb{Z}$. Define the elements

$$a_n = \prod_{i=0}^{n-1} a^{(-1)^i \binom{n-1}{i} t^i} = a t^{-1} a^{-(n-1)} t^{-1} \dots a^{(-1)^{n-1}} t^{n-1}$$

of G. The Lie algebra $\mathscr{L}_{\mathbb{F}_2}(G)$ is as follows:



Note that $\mathscr{L}_{\mathbb{F}_2}(G)$ has bounded width, while G has exponential growth! This shows that in Proposition 1.10 the group growth rate can be exponential while the algebra growth rate is polynomial.

3.4. ***The Nottingham group.*** As a final example, we give the Lie graph of the Nottingham group's Lie algebra [Jennings 1954; Camina 2000]. Recall that for odd prime p the Nottingham group $J(p)$ is the group of all formal power series

$$\hbar + \sum_{i>1} a_i \hbar^i \in \mathbb{F}_p[\![\hbar]\!],$$

with composition (i.e., substitution) as binary operation. The lower central series is given by

$$J_n = \{\hbar + \sum_{i > \left\lceil \frac{np-1}{p-1} \right\rceil} a_i \hbar^i\},$$

and a basis of $\mathscr{L}$ is $\{f_i = \hbar(1+\hbar^i)\}_{i \ge 1}$, where f_i has degree $\lfloor ((p-1)i+1)/p \rfloor$. As a basis of J_1/J_2, we take $B_1 = \{x = \hbar + \hbar^2 + \hbar^3,\ y = \hbar + \hbar^3\}$. The commutations are given by

$$[f_i, x] = (i-1) f_{i+1}, \qquad [f_i, y] = \begin{cases} -2 f_{i+2} & \text{if } i \equiv 0 \bmod p \\ -f_{i+2} & \text{if } i \equiv 1 \bmod p \\ 0 & \text{otherwise,} \end{cases}.$$

This gives a Lie graph with a diamond structure [Caranti 1997]:

$$
\begin{tikzcd}
1 & 2 & 3 & \cdots & p-2 & p-1 & p & p+1 & p+2 & \\
x \arrow[dr, "-y"] & & & & & & f_{p+1} \arrow[dr, "-y"] & & & \\
& f_3 \arrow[r, "2x"] & f_4 \arrow[rr, dotted] & & f_{p-1} \arrow[r, "-2x"] & f_p \arrow[ur, "-x"] \arrow[dr, "-2y"'] & & f_{p+3} \arrow[r, "2x"] & f_{p+4} \arrow[r, dotted] & {} \\
y \arrow[ur, "x"'] & & & & & & f_{p+2} \arrow[ur, "x"'] & & &
\end{tikzcd}
$$

3.5. ***The tree automorphism group's pro-p-Sylow $\mathsf{Aut}_p(\Sigma^*)$.*** We start by considering a typical example of branch group. Let p be prime; write $p' = p - 1$ for notational simplicity. Let Σ be the p-letter alphabet $\{1, \dots, \mathsf{p}\}$, and let x_n, for $n \in \mathbb{N}$, be the p-cycle permuting the first p branches at level $n + 1$ in the tree Σ^*. Therefore x_0 acts just below the root vertex, and $x_{n+1} = \ll x_n, 1, \dots, 1\gg$ for all n.

For all $n \in \mathbb{N}$ we define $G_n = \mathsf{Aut}_p(\Sigma^*)$ as the group generated by $\{x_0, \dots, x_{n-1}\}$, and $G = \langle x_0, x_1, \dots\rangle$. Clearly $G = \operatorname{inj\,lim} G_n$, while its closure is $\overline{G} = \operatorname{proj\,lim} G_n$. Note that G_n is a p-Sylow of $\mathfrak{S}_{p^n}$, and $\overline{G}$ is a pro-p-Sylow of $\mathsf{Aut}(\Sigma^*)$.

Lemma 3.2. *$G = G \wr C_p$; therefore G is a regular branch group over itself.*

Proof. The subgroup $\langle x_1, x_2, \dots\rangle$ of G is isomorphic to G through $x_i \mapsto x_{i-1}$, and its p conjugates under powers of x_0 commute, since they act on disjoint subtrees. $\square$

Lev Kaloujnine [1948] described the lower central series of G_n, using his notion of a *tableau*. Our purpose here shall be to describe the Lie algebra of G_n (and therefore G and $\overline{G}$) using our more geometric approach. Let us just mention that in Kaloujnine's theory of tableaux his polynomials $x_1^{e_1} \dots x_n^{e_n}$ correspond to our $\mathbb{e}_1 \dots \mathbb{e}_n(x_0)$.

Lemma 3.3. *For $u, v \in G$ and $X, Y \in \{\mathbb{0}, \dots, \mathbb{p}'\}^n$ we have*

$$[X(u), Y(v)] \equiv (X_1 + Y_1 - \mathbb{p}') \dots (X_n + Y_n - \mathbb{p}')([u, v])^{\prod_{i=1}^{n} (-1)^{p'-Y_i} \binom{X_i}{p'-Y_i}},$$

modulo terms in $\big[[X(u), Y(v)], G\big]$.

Proof. The proof follows by induction, and we may suppose $n = 1$ without loss of generality. Multiplying by terms in $\big[[X(u), Y(v)], G\big]$, we may assume $Y(v)$ by some element acting only on the last Y_1 subtrees below the root vertex. Then

$$
\begin{aligned}
[X(u), Y(v)] &\equiv [\ll u, \dots, u^{(-1)^{X_1}}, 1, \dots, 1\gg, \ll 1, \dots, 1, v, \dots, v^{(-1)^{Y_1}}\gg] \\
&= \ll [u, 1], \dots, [u^{(-1)^{p'-Y_1}\binom{p}{p'-Y_1}}, v], \dots, \\
&\qquad\qquad [u^{(-1)^{X_1}}, v^{(-1)^{X_1}\binom{p}{X_1}}], \dots, [1, v]\gg \\
&\equiv (X + Y - \mathbb{p}')([u, v])^{(-1)^{p'-Y_1}\binom{p}{p'-Y_1}}.
\end{aligned}
$$
$\square$

Theorem 3.4. *Consider the following Lie graph: its vertices are the symbols X for all words $X \in \{\mathbb{0}, \dots, \mathbb{p}'\}^*$, including the empty word λ. Their degrees are given by*

$$\deg X_1 \dots X_n = 1 + \sum_{i=1}^{n} X_i p^{i-1}.$$

For all $m > n \geq 0$ and all choices of X_i, there is an arrow labeled $\mathbb{0}^n$ from

$$\mathbb{p}'^n X_{n+1} \dots X_m$$

to

$$\mathbb{0}^n (X_{n+1} + \mathbb{1}) X_{n+2} \dots X_m,$$

and an arrow labeled $\mathbb{0}^m$ from $\mathbb{p}'^n$ to $\mathbb{0}^n \mathbb{1} \mathbb{0}^{m-n-1}$.

Then the resulting graph is the Lie graph of $\mathscr{L}(G)$ and of $\mathscr{L}_{\mathbb{F}_p}(G)$.

The subgraph spanned by all words of length up to $n - 1$ is the Lie graph of $\mathscr{L}(G_n)$ and of $\mathscr{L}_{\mathbb{F}_p}(G_n)$.

Proof. We interpret X in the Lie graph as $X(x_0)$ in G. The generator x_n is then $\mathbb{0}^n(x_0)$. By Lemma 3.3, the adjoint operators $\mathsf{ad}(x_n)$ correspond to the arrows

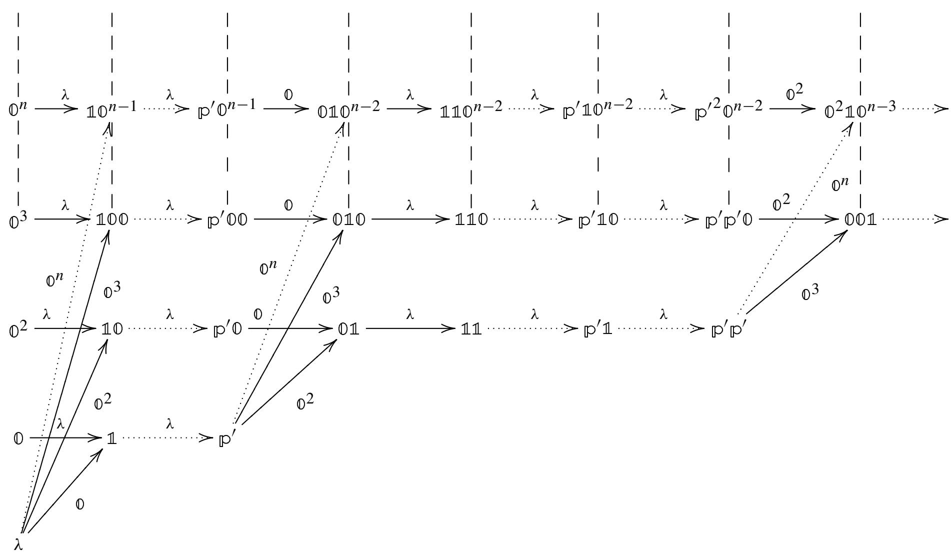


Figure 1. The beginning of the Lie graph of $\mathscr{L}(G)$ for G the p-Sylow of $\mathsf{Aut}(\Sigma^*)$.

labeled $\mathbb{0}^n$. The arrows connect elements whose degree differ by 1, so the degree of the element $X(x_0)$ is $\deg X$ as claimed.

The power maps $g \mapsto g^p$ are all trivial on the elements $X(x_0)$, so the Lie algebra and restricted Lie algebra coincide.

The elements $X(x_0)$ for $|X| \geq n$ belong to $\mathsf{Stab}_G(n)$, hence are trivial in G_n. $\square$

3.6. *The Grigorchuk group $\mathfrak{G}$.* We give an explicit description of the Lie algebra of $\mathfrak{G}$, and compute its Hilbert–Poincaré series. These results were obtained in [Bartholdi and Grigorchuk 2000a], and partly before in [Rozhkov 1996].

Set $x = [a, b]$. Then $\mathfrak{G}$ is branch over $K = \langle x \rangle^{\mathfrak{G}}$, and $K/(K \times K)$ is cyclic of order 4, generated by x.

Extend the generating set of $\mathfrak{G}$ to a formal alphabet

$$S = \left\{a, b, c, d, \left\{\begin{smallmatrix} b \\ c \end{smallmatrix}\right\}, \left\{\begin{smallmatrix} c \\ d \end{smallmatrix}\right\}, \left\{\begin{smallmatrix} d \\ b \end{smallmatrix}\right\}\right\}.$$

Define the transformation σ on words in S^* by

$$\sigma(a) = a \left\{\begin{smallmatrix} b \\ c \end{smallmatrix}\right\} a, \quad \sigma(b) = d, \quad \sigma(c) = b, \quad \sigma(d) = c,$$

extended to subsets by

$$\sigma \left\{\begin{smallmatrix} x \\ y \end{smallmatrix}\right\} = \left\{\begin{smallmatrix} \sigma x \\ \sigma y \end{smallmatrix}\right\}.$$

Note that for any fixed $g \in G$, all elements $h \in \mathsf{Stab}_{\mathfrak{G}}(1)$ such that $\psi(h) = \langle\!\langle g, * \rangle\!\rangle$ are obtained by picking a letter from each set in $\sigma(g)$. This motivates the definition of S.

Theorem 3.5. *Consider the following Lie graph: its vertices are the symbols $X(x)$ and $X(x^2)$, for words $X \in \{\mathbb{0}, \mathbb{1}\}^*$. Their degrees are given by*

$$\deg X_1 \dots X_n(x) = 1 + \sum_{i=1}^{n} X_i 2^{i-1} + 2^n,$$
$$\deg X_1 \dots X_n(x^2) = 1 + \sum_{i=1}^{n} X_i 2^{i-1} + 2^{n+1}.$$

There are four additional vertices: a, b, d of degree 1, and $[a, d]$ of degree 2.

Define the arrows as shown below, where an arrow labeled $\left\{\begin{smallmatrix} x \\ y \end{smallmatrix}\right\}$ or x, y stands for two arrows, labeled x and y, and the arrows labeled c are there to expose the *symmetry of the graph* (*indeed* $c = bd$ *is not in our chosen basis of* $G/[G,G]$):

$$\begin{array}{llll}
a \xrightarrow{b,c} x & & a \xrightarrow{c,d} [a,d] \\
b \xrightarrow{a} x & & d \xrightarrow{a} [a,d] \\
x \xrightarrow{a,b,c} x^2 & & x \xrightarrow{c,d} \mathbb{0}(x) \\
[a,d] \xrightarrow{b,c} \mathbb{0}(x) & & \mathbb{0}* \xrightarrow{a} \mathbb{1}* \\
\mathbb{1}^n(x) \xrightarrow{\sigma^n\left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\}} \mathbb{0}^{n+1}(x) & & \mathbb{1}^n(x) \xrightarrow{\sigma^n\left\{\begin{smallmatrix}b\\d\end{smallmatrix}\right\}} \mathbb{0}^n(x^2) \\
\mathbb{1}^n\mathbb{0}* \xrightarrow{\sigma^n\left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\}} \mathbb{0}^n\mathbb{1}* \quad \text{if } n \geq 1. & & &
\end{array}$$

Then the resulting graph is the Lie graph of $\mathscr{L}(\mathfrak{G})$. *A slight modification gives the Lie graph of* $\mathscr{L}_{\mathbb{F}_2}(\mathfrak{G})$: *the degree of* $X_1 \dots X_n(x^2)$ *is then* $2 \deg X_1 \dots X_n(x)$; *and the* 2*-mappings are given by*

$$X(x) \xrightarrow{\cdot 2} X(x^2),$$
$$\mathbb{1}^n(x^2) \xrightarrow{\cdot 2} \mathbb{1}^{n+1}(x^2).$$

The subgraph spanned by $a, t, X_1 \dots X_i(x)$ *for* $i \leq n-2$ *and* $X_1 \dots X_i(x^2)$ *for* $i \leq n-4$ *is the Lie graph associated to the finite quotient* $\mathfrak{G}/\operatorname{\mathsf{Stab}}_{\mathfrak{G}}(n)$.

Figure 2 describes as Lie graphs the top of the Lie algebras associated to $\mathfrak{G}$. Note the infinite path, labeled by

$$\begin{aligned}
&\left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\} a \sigma(\left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\} a) \sigma^2(\left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\} a) \dots \\
&\qquad = \left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\c\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\c\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\d\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\c\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\d\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\c\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}c\\d\end{smallmatrix}\right\} a \left\{\begin{smallmatrix}b\\c\end{smallmatrix}\right\} a \dots;
\end{aligned}$$

it is the same as the labeling of the parabolic space of $\mathfrak{G}$ — see Section 4 and [Bartholdi and Grigorchuk 2002].

The proof requires the computation, given a term N of a central series and a generator $s \in \{a, b, c, d\}$, of $[N, x]$ modulo $[N, G, G]$. We do slightly better in the following lemma — this will be useful in Section 5, where we describe all normal subgroups of G. For that purpose we introduce a symbol $\genfrac{}{}{0pt}{}{\mathbb{0}}{\mathbb{1}}(x) = \mathbb{0}(x)\mathbb{1}(x)^{-1}$. We then have

$$\mathbb{0}(x) = \ll x, 1 \gg, \quad \mathbb{1}(x) = \ll x, x^{-1} \gg, \quad \genfrac{}{}{0pt}{}{\mathbb{0}}{\mathbb{1}}(x) = \ll 1, x \gg.$$

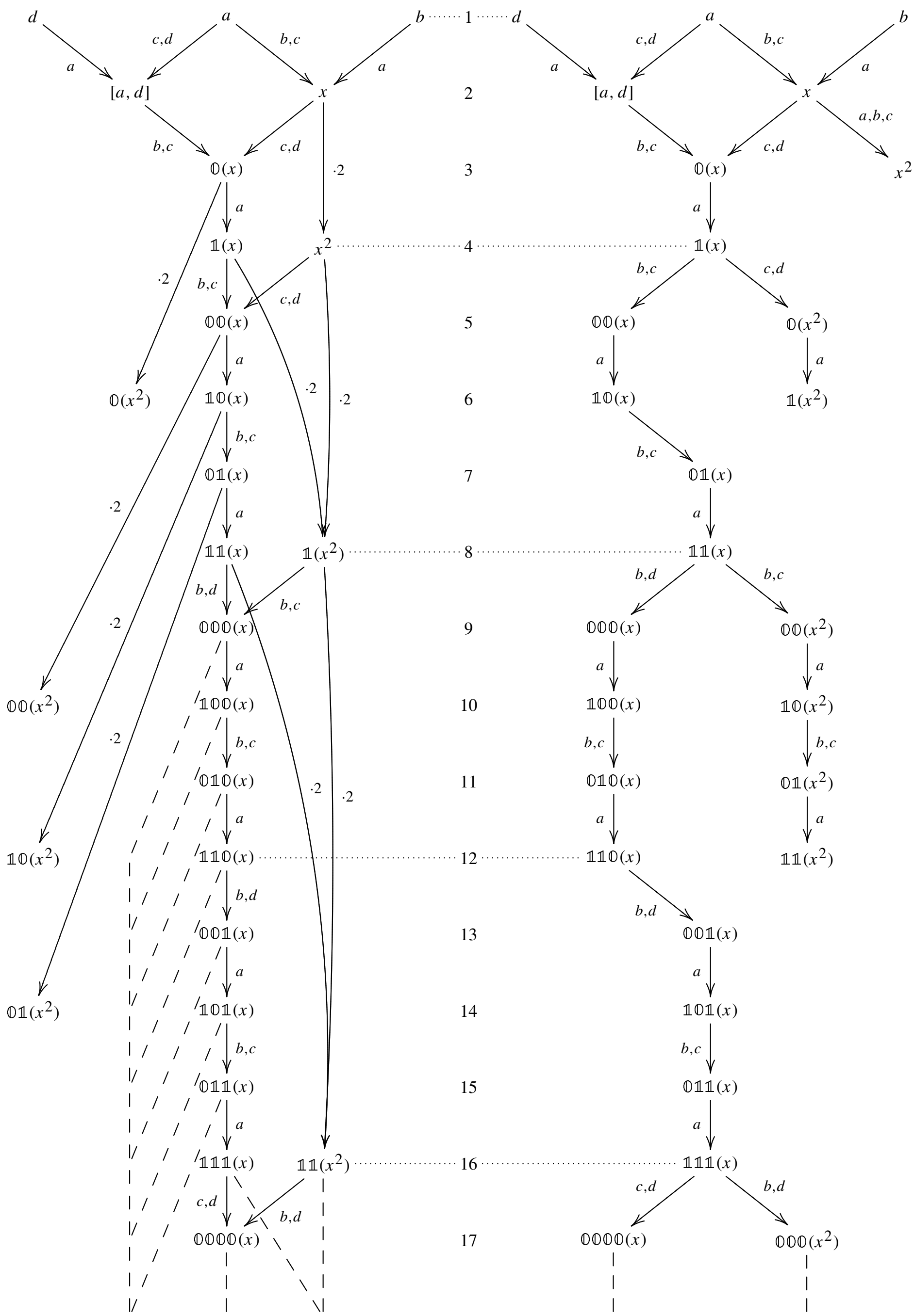


Figure 2. The beginning of the Lie graphs of $\mathscr{L}_{\mathbb{F}_2}(\mathfrak{G})$ (left) and $\mathscr{L}(\mathfrak{G})$ (right).

Lemma 3.6. *Assume N is a normal subgroup containing the left-hand operand of the commutators below. Then modulo $[N, G]'$ we have*

$$
\begin{array}{ll}
[\mathbb{0}X, a] = \mathbb{1}X, & [\mathbb{1}X, a] = \mathbb{1}X^2, \\
[\mathbb{0}X, b] = \mathbb{0}[X, a], & [\mathbb{1}X, b] = \mathbb{0}[X, a] + \genfrac{}{}{0pt}{}{\mathbb{0}}{\mathbb{1}}[X, c], \\
[\mathbb{0}X, c] = \mathbb{0}[X, a], & [\mathbb{1}X, c] = \mathbb{0}[X, a] + \genfrac{}{}{0pt}{}{\mathbb{0}}{\mathbb{1}}[X, d], \\
[\mathbb{0}X, d] = 1, & [\mathbb{1}X, d] = \genfrac{}{}{0pt}{}{\mathbb{0}}{\mathbb{1}}[X, b], \\
[x, a] = x^2, & [x^2, a] = x^4 = \mathbb{1}(x^2 + \mathbb{1}x), \\
[x, b] = x^2, & [x^2, b] = \mathbb{1}(x^2 + \mathbb{1}x), \\
[x, c] = \mathbb{0}(x) + x^2, & [x^2, c] = \mathbb{0}(x^2 + \mathbb{0}x) + \mathbb{1}(x^2 + \mathbb{1}x), \\
[x, d] = \mathbb{0}(x), & [x^2, d] = \mathbb{0}(x^2 + \mathbb{0}x).
\end{array}
$$

Proof. Direct computation, using the decompositions $\psi(b) = (a, c) = \mathbb{0}(a) \cdot \genfrac{}{}{0pt}{}{\mathbb{0}}{\mathbb{1}}(c)$ etc. and linearizing. □

Proof of Theorem 3.5. The proof proceeds by induction on length of words, or, what amounts to the same, on depth in the lower central series.

First, the assertion is checked "manually" up to degree 3. The details of the computations are the same as in [Bartholdi and Grigorchuk 2000a].

We claim that for all words X, Y with $\deg Y(x) > \deg X(x)$ we have $Y(x) \in \langle X(x)\rangle^{\mathfrak{G}}$, and similarly $Y(x^2) \in \langle X(x^2)\rangle^{\mathfrak{G}}$. The claim is verified by induction on $\deg X$.

We then claim that for any nonempty word X, either $\mathsf{ad}(a)X(*) = 0$ (if X starts with $\mathbb{1}$) or $\mathsf{ad}(v)X(*) = 0$ for $v \in \{b, c, d\}$ (if X starts with $\mathbb{0}$). Again this holds by induction.

We then prove that the arrows are as described above; this follows from Lemma 3.6. For instance,

$$
\mathsf{ad}(\sigma^n \left\{\genfrac{}{}{0pt}{}{c}{d}\right\})\mathbb{1}^n\mathbb{0}* = \begin{cases} \big(\mathsf{ad}(\sigma^n \left\{\genfrac{}{}{0pt}{}{d}{b}\right\})\mathbb{1}^{n-1}\mathbb{0}*, \mathsf{ad}(\left\{\genfrac{}{}{0pt}{}{a}{1}\right\})\mathbb{1}^{n-1}\mathbb{0}*\big) \\ \qquad = \mathbb{0}\, \mathsf{ad}(\sigma^{n-1} \left\{\genfrac{}{}{0pt}{}{c}{d}\right\})\mathbb{1}^{n-1}\mathbb{0}* = \mathbb{0}^n\mathbb{1}* & \text{if } n \geq 2, \\ \big(\mathsf{ad}(\left\{\genfrac{}{}{0pt}{}{b}{c}\right\})\mathbb{0}*, \mathsf{ad}(a)\mathbb{0}*\big) = \mathbb{0}\mathbb{1}* & \text{if } n = 1. \end{cases}
$$

Finally we check that the degrees of all basis elements are as claimed. For that purpose, we first check that the degree of an arrow's destination is always one more than the degree of its source. Then fix a word $X(*)$ and consider the largest n such that $X(*)$ belongs to $\gamma_n(\mathfrak{G})$. There is an expression of $X(*)$ as a product of n-place commutators on elements of $\mathfrak{G} \setminus [\mathfrak{G}, \mathfrak{G}]$, and therefore in the Lie graph there is a family of paths starting at some element of B_1 and following $n - 1$ arrows to reach $X(*)$. This implies that the degree of $X(*)$ is n, as required.

The modification giving the Lie graph of $\mathscr{L}_{\mathbb{F}_2}(\mathfrak{G})$ is justified by the fact that in $\mathscr{L}(\mathfrak{G})$ we always have $\deg X(x^2) \leq 2\deg X(x)$, so the element $X(x^2)$ appears always last as the image of $X(x)$ through the square map. The degrees are modified accordingly. Now $X(x^2) = X\mathbb{1}(x^2)$, and $2\deg X\mathbb{1}(x) \geq 4\deg X(x)$, with equality only when $X = \mathbb{1}^n$. This gives an additional square map from $\mathbb{1}^n(x^2)$ to $\mathbb{1}^{n+1}(x^2)$, and requires no adjustment of the degrees. □

Corollary 3.7. *Define the polynomials*

$$\begin{aligned} Q_2 &= -1 - \hbar, \\ Q_3 &= \hbar + \hbar^2 + \hbar^3, \\ Q_n(\hbar) &= (1+\hbar)Q_{n-1}(\hbar^2) + \hbar + \hbar^2 \quad \textit{for } n \geq 4. \end{aligned}$$

Then Q_n is a polynomial of degree $2^{n-1} - 1$, and the first $2^{n-3} - 1$ coefficients of Q_n and Q_{n+1} coincide. The termwise limit $Q_\infty = \lim_{n\to\infty} Q_n$ therefore exists.

The Hilbert–Poincaré series of $\mathscr{L}(\mathfrak{G}/\operatorname{Stab}_{\mathfrak{G}}(n))$ is $3\hbar + \hbar^2 + \hbar Q_n$, and the Hilbert–Poincaré series of $\mathscr{L}(\mathfrak{G})$ is $3\hbar + \hbar^2 + \hbar Q_\infty$.

The Hilbert–Poincaré series of $\mathscr{L}_{\mathbb{F}_2}(\mathfrak{G})$ is $3 + (2\hbar + \hbar^2)/(1-\hbar^2)$.

As a consequence, $\mathfrak{G}/\operatorname{Stab}_{\mathfrak{G}}(n)$ is nilpotent of class 2^{n-1}, and $\mathfrak{G}$ has finite width.

Proof. Consider the sequence of coefficients of Q_n. They are, in condensed form,

$$1, 2^{2^0}, 1^{2^0}, 2^{2^1}, 1^{2^1}, \dots, 2^{n-4}, 1^{n-4}, 1^{n-2}.$$

The i-th coefficient is 2 if there are $X(x)$ and $Y(x^2)$ of degree i in $\mathfrak{G}/\operatorname{Stab}_{\mathfrak{G}}(n)$, and is 1 if there is only $X(x)$. All conclusions follow from this remark. □

3.7. ***The Gupta–Sidki group*** $\ddot{\Gamma}$**.** We now give an explicit description of the Lie algebra of $\ddot{\Gamma}$, and compute its Hilbert–Poincaré series.

Introduce the sequence of integers

$$\alpha_1 = 1, \quad \alpha_2 = 2, \quad \alpha_n = 2\alpha_{n-1} + \alpha_{n-2} \quad \text{for } n \geq 3,$$

and set $\beta_n = \sum_{i=1}^n \alpha_i$. One has

$$\begin{aligned} \alpha_n &= \frac{1}{2\sqrt{2}}\big((1+\sqrt{2})^n - (1-\sqrt{2})^n\big), \\ \beta_n &= \frac{1}{4}\big((1+\sqrt{2})^{n+1} + (1-\sqrt{2})^{n+1} - 2\big). \end{aligned}$$

The first few values are

n	1	2	3	4	5	6	7	8
α_n	1	2	5	12	29	70	169	398
β_n	1	3	8	20	49	119	288	686

Theorem 3.8. *In $\ddot\Gamma$ write $c=[a,t]$ and $u=[a,c]=\mathbb{2}(t)$. Consider the following Lie graph: its vertices are the symbols $X_1\dots X_n(x)$ with $X_i\in\{\mathbb{0},\mathbb{1},\mathbb{2}\}$ and $x\in\{c,u\}$. Their degrees are given by*

$$\deg X_1\dots X_n(c)=1+\sum_{i=1}^n X_i\alpha_i+\alpha_{n+1},$$
$$\deg X_1\dots X_n(u)=1+\sum_{i=1}^n X_i\alpha_i+2\alpha_{n+1}.$$

There are two additional vertices, labeled a and t, of degree 1.

Define the arrows as follows:

$$\begin{array}{lll}
a \xrightarrow{-t} c & & c \xrightarrow{t} \mathbb{0}(c)\\
t \overset{a}{\dashrightarrow} c & & c \overset{a}{\dashrightarrow} u\\
& u \xrightarrow{t} \mathbb{1}(c) & \\
\mathbb{0}* \overset{a}{\dashrightarrow} \mathbb{1}* & & \mathbb{1}* \overset{a}{\dashrightarrow} \mathbb{2}*\\
\mathbb{2}* \xrightarrow{t} \mathbb{0} & \text{whenever } * \xrightarrow{t} \# & \\
\mathbb{2}(c) \xrightarrow{t} \mathbb{1}(u) & & \mathbb{1}(c) \xrightarrow{-t} \mathbb{0}(u)\\
\mathbb{1}\mathbb{0}* \xrightarrow{-t} \mathbb{0}\mathbb{1}* & & \mathbb{1}\mathbb{1}* \xrightarrow{-t} \mathbb{0}\mathbb{2}*\\
\mathbb{2}\mathbb{0}* \xrightarrow{t} \mathbb{1}\mathbb{1}* & & \mathbb{2}\mathbb{1}* \xrightarrow{t} \mathbb{1}\mathbb{2}*
\end{array}$$

(*The last three lines can be replaced by the rules $\mathbb{2}*\xrightarrow{t}\mathbb{1}\#$ and $\mathbb{1}*\xrightarrow{-t}\mathbb{0}\#$ for all arrows $*\xrightarrow{a}\#$.*)

Then the resulting graph is the Lie graph of $\mathscr{L}(\ddot\Gamma)$. It is also the Lie graph of $\mathscr{L}_{\mathbb{F}_3}(\ddot\Gamma)$, with the only nontrivial cube maps given by

$$\mathbb{2}^n(c)\xrightarrow{\cdot 3}\mathbb{2}^n\mathbb{0}\mathbb{0}(c),\qquad \mathbb{2}^n(c)\xrightarrow{\cdot 3}\mathbb{2}^n\mathbb{1}(u).$$

The subgraph spanned by a, t, the $X_1\dots X_i(c)$ for $i\le n-2$ and the $X_1\dots X_i(u)$ for $i\le n-3$ is the Lie graph associated to the finite quotient $\ddot\Gamma/\operatorname{Stab}_{\ddot\Gamma}(n)$.

Proof. We perform the computations in the completion of $\ddot\Gamma$, still written $\ddot\Gamma$. With Lemma 2.5 in mind, $\ddot\Gamma'$ is the subgroup generated by all $X(c)$ and $X(u)$, for $X\in\{\mathbb{0},\mathbb{1},\mathbb{2}\}^*$.

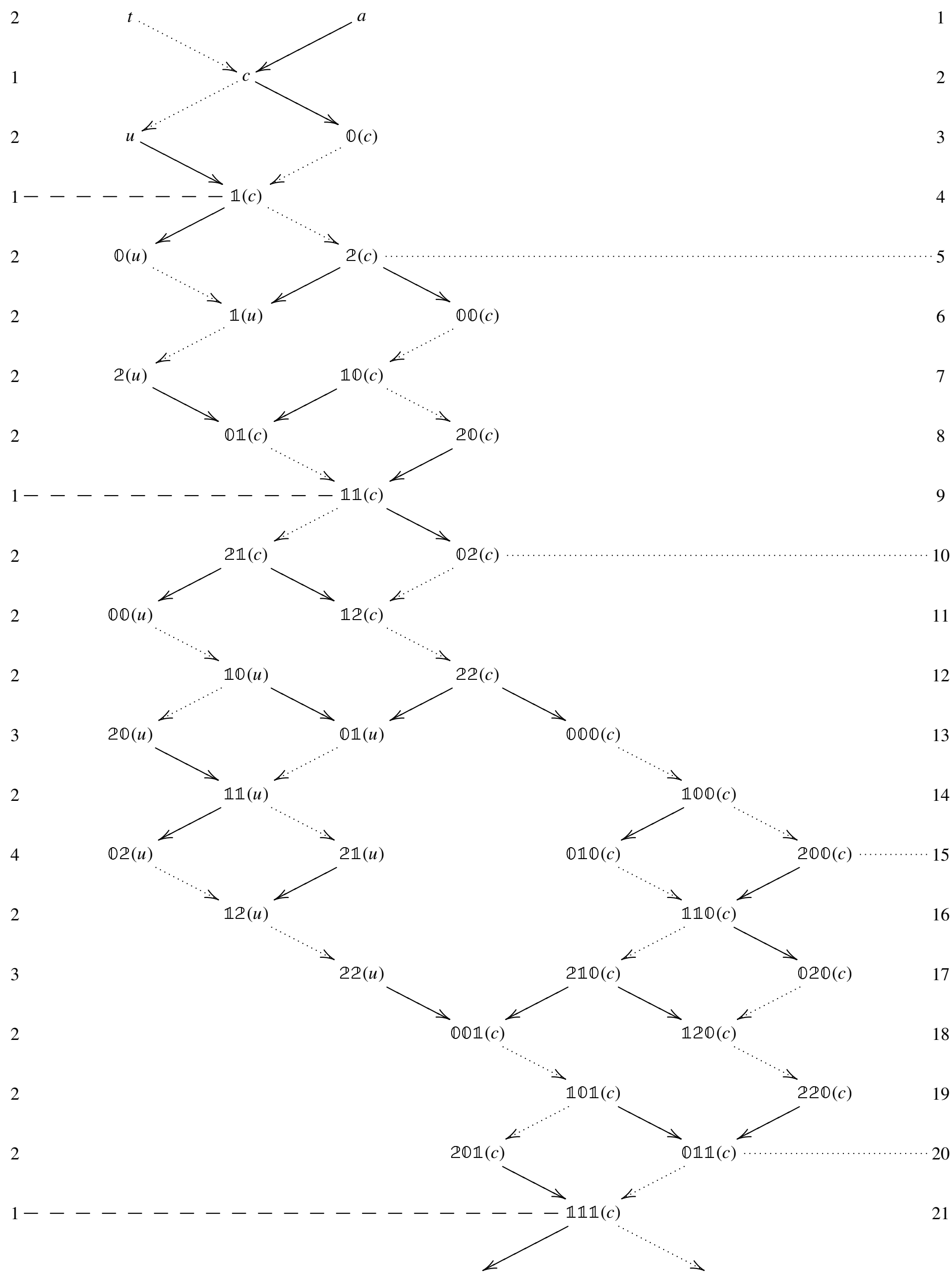


Figure 3. The beginning of the Lie graph of $\mathscr{L}(\ddot{\Gamma})$. The generator $\mathsf{ad}(t)$ is shown by plain arrows, and the generator $\mathsf{ad}(a)$ is shown by dotted arrows. The left column indicates the dimensions of $\mathscr{L}_n$.

We claim inductively that if $X_i \geq Y_i$ at all positions i, then $X(c) \in \langle Y(c)\rangle^{\ddot{\Gamma}}$, and similarly for u. Therefore some terms may be neglected in the computations of brackets.

Now we compute $\mathsf{ad}(x)y$ for $x, y \in \{a, t, c, u\}$. Here $\equiv$ means some terms of greater degree have been neglected:

$$\begin{aligned}
&[a, \mathbb{0}*] = \mathbb{1}*, \quad [a, \mathbb{1}*] = \mathbb{2}*, \quad [a, \mathbb{2}*] = 1 \text{ by definition,}\\
&[t, \mathbb{0}*] = [\ll a, a^{-1}, t\gg, \ll *, 1, 1\gg] = \ll [a, *], 1, 1\gg = \mathbb{0}[a, *]\\
&\qquad \equiv [\ll a^{-1}, t, a\gg, \ll *, 1, 1\gg] = -\mathbb{0}[a, *], \quad \text{so } [t, \mathbb{0}*] = 1,\\
&[t, \mathbb{1}*] = [\ll a, a^{-1}, t\gg, \ll *, *^{-1}, 1\gg] \equiv -\mathbb{0}[a, *],\\
&[t, \mathbb{2}*] = [\ll a, a^{-1}, t\gg, \ll *, *, *\gg] \equiv \mathbb{1}[a, *] + \mathbb{0}[t, *].
\end{aligned}$$

All asserted arrows follow from these equations.

Finally, we prove that the degrees of $X(c)$ and $X(u)$ are as claimed, by remarking that $\deg c = 3$ and $\deg u = 4$, that $\deg \mathsf{ad}(s)* \geq \deg(*)$ for $s = a, t$ and all words $*$ (so the claimed degrees smaller of equal to their actual value), and that each word of claimed degree n appears only as $\mathsf{ad}(s)*$ for words $*$ of degree at most $n-1$ (so the claimed degrees are greater or equal to their actual value).

The last point to check concerns the cube map; we skip the details. □

Corollary 3.9. *Define the polynomials*

$$\begin{aligned}
&Q_1 = 0,\\
&Q_2 = \hbar + \hbar^2,\\
&Q_3 = \hbar + \hbar^2 + 2\hbar^3 + \hbar^4 + \hbar^5,\\
&Q_n = (1 + \hbar^{\alpha_n - \alpha_{n-1}})Q_{n-1} + \hbar^{\alpha_{n-1}}(\hbar^{-\alpha_{n-2}} + 1 + \hbar^{\alpha_{n-2}})Q_{n-2} \quad \textit{for } n \geq 3.
\end{aligned}$$

Then Q_n is a polynomial of degree α_n, and the polynomials Q_n and Q_{n+1} coincide on their first $2\alpha_{n-1}$ terms. Thus the coefficientwise limit $Q_\infty = \lim_{n\to\infty} Q_n$ exists.

The largest coefficient in Q_{2n+1} is 2^n, at position $\frac{1}{2}(\alpha_{2n+1}+1)$, so the coefficients of Q_∞ are unbounded. The integers k such that $\hbar^k$ has coefficient 1 in Q_∞ are precisely the $\beta_n + 1$.

The Hilbert–Poincaré series of $\mathscr{L}(\ddot{\Gamma}/\mathsf{Stab}_{\ddot{\Gamma}}(n))$ is $\hbar + Q_n$, and the Hilbert–Poincaré series of $\mathscr{L}(\ddot{\Gamma})$ is $\hbar + Q_\infty$. The same holds for the Lie algebras

$$\mathscr{L}_{\mathbb{F}_3}(\ddot{\Gamma}/\mathsf{Stab}_{\ddot{\Gamma}}(n)) \quad \textit{and} \quad \mathscr{L}_{\mathbb{F}_3}(\ddot{\Gamma}).$$

As a consequence, $\ddot{\Gamma}/\mathsf{Stab}_{\ddot{\Gamma}}(n)$ is nilpotent of class α_n, and $\ddot{\Gamma}$ does not have finite width.

Proof. Define polynomials

$$R_n = \sum_{w \in \{\mathbb{0},\mathbb{1},\mathbb{2}\}^n} \hbar^{\deg w(c)} + \sum_{w \in \{\mathbb{0},\mathbb{1},\mathbb{2}\}^{n-1}} \hbar^{\deg w(u)} + \hbar.$$

One checks directly that the polynomials R_n satisfy the same initial values and recurrence relation as Q_n, hence are equal. All convergence properties also follow from the definition of R_n.

The words of degree $\frac{1}{2}(\alpha_{2n+1}+1)$ are $(\mathbb{0}\mathbb{1})^{n-1}\mathbb{0}(c)$, $(\mathbb{0}\mathbb{1})^{n-2}\mathbb{0}\mathbb{2}(u)$, and all the words that can be obtained from these by iterating the substitutions $\mathbb{0}\mathbb{0}\mathbb{1} \mapsto \mathbb{1}\mathbb{2}\mathbb{0}$, $\mathbb{1}\mathbb{0}\mathbb{1} \mapsto \mathbb{2}\mathbb{2}\mathbb{0}$, $\mathbb{0}\mathbb{0}\mathbb{2} \mapsto \mathbb{1}\mathbb{2}\mathbb{1}$, $\mathbb{1}\mathbb{0}\mathbb{2} \mapsto \mathbb{2}\mathbb{2}\mathbb{1}$ along with $\mathbb{0}\mathbb{1} \mapsto \mathbb{2}\mathbb{0}$ and $\mathbb{0}\mathbb{2} \mapsto \mathbb{2}\mathbb{1}$ at the beginning of the word. This gives 2^n words in total, half of the form $X(c)$ and half $X(u)$.

There is a unique word of degree $\beta_n + 1$, and that is $\mathbb{1}^n(c)$.

Note that these last two claims have a simple interpretation: there are 2^{n-1} ways of writing $\frac{1}{2}(\alpha_{2n+1}) - 1 - \alpha_{n+1}$ in base α using only the digits $0, 1, 2$; there is a unique way of writing β_n in base α using these digits. $\square$

We note as an immediate consequence that

$$\big[\ddot{\Gamma} : \gamma_{\beta_n+1}(\ddot{\Gamma})\big] = 3^{(3^n+1)/2},$$

so that the asymptotic growth of $\ell_n = \dim \gamma_n(\ddot{\Gamma})/\gamma_{n+1}(\ddot{\Gamma})$ is polynomial of degree $d = \log 3/\log(1+\sqrt{2}) - 1$:

Corollary 3.10. *The Gelfand–Kirillov dimension of* $\mathscr{L}(\ddot{\Gamma})$ *is* $\log 3/\log(1+\sqrt{2}) - 1$.

We then deduce:

Corollary 3.11. *The growth of* $\ddot{\Gamma}$ *is at least* $e^{n^{\frac{\log 3}{\log(1+\sqrt{2})+\log 3}}} \cong e^{n^{0.554}}$.

Proof. Apply Proposition 1.10 to the series $\sum n^d \hbar^n$, which is comparable to the Hilbert–Poincaré series of $\mathscr{L}(\ddot{\Gamma})$. $\square$

Turning to the derived series, we may also improve the general result $\ddot{\Gamma}^{(k)} \le \gamma_{2^k}(\ddot{\Gamma})$ to the following:

Theorem 3.12. *For all* $k \in \mathbb{N}$ *we have*

$$\ddot{\Gamma}^{(k)} \le \gamma_{\alpha_{k+1}}(\ddot{\Gamma}).$$

Proof. Clearly true for $k = 0, 1$; then a direct consequence of $\ddot{\Gamma}^{(k)} = \gamma_5(\ddot{\Gamma})^{\times 3^{k-2}}$ (obtained in [Vieira 1998]) and $\gamma_{\alpha_j}(\ddot{\Gamma})^{\times 3} \le \gamma_{\alpha_{j+1}}(\ddot{\Gamma})$ for $j = 3, \dots, k$. $\square$

3.8. *The Fabrykowski–Gupta group* Γ. We now give an explicit description of the Lie algebra of Γ, and compute its Hilbert–Poincaré series.

Theorem 3.13. *In Γ write $c = [a, t]$ and $u = [a, c] \equiv \mathbb{2}(at)$. For words $X = X_1 \dots X_n$ with $X_i \in \{\mathbb{0}, \mathbb{1}, \mathbb{2}\}$ define symbols $\overline{X_1 \dots X_n}(c)$ (representing elements of Γ) by*

$$\begin{aligned}
\overline{\mathbb{i}\mathbb{0}}(c) &= \mathbb{i}\mathbb{0}(c)/\mathbb{i}(u),\\
\overline{\mathbb{i}\mathbb{2}^{m+1}\mathbb{1}^n}(c) &= \mathbb{i}\big(\overline{\mathbb{2}^{m+1}\mathbb{1}^n}(c)\cdot \mathbb{0}\mathbb{1}^m\mathbb{0}^n(u)^{(-1)^n}\big),\\
\overline{\mathbb{i}X}(c) &= \mathbb{i}\overline{X}(c) \textit{ for all other } X.
\end{aligned}$$

Consider the following Lie graph: its vertices are the symbols $\overline{X}(c)$ and $X(u)$. Their degrees are given by

$$\deg \overline{X_1 \dots X_n}(c) = 1 + \sum_{i=1}^{n} X_i 3^{i-1} + \frac{1}{2}(3^n + 1),$$

$$\deg X_1 \dots X_n(u) = 1 + \sum_{i=1}^{n} X_i 3^{i-1} + (3^n + 1).$$

There are two additional vertices, labeled a and t, of degree 1.

Define the arrows as follows, for all $n \geq 1$:

$$\begin{array}{ll}
a \xrightarrow{-t} c & t \overset{a}{\dashrightarrow} c\\
c \xrightarrow{-t} \mathbb{0}(c) & c \xrightarrow{a} u\\
u \xrightarrow{-t} \mathbb{1}(c) & \overline{\mathbb{2}^n}(c) \xrightarrow{-t} \overline{\mathbb{0}^{n+1}}(c)\\
\mathbb{0}* \overset{a}{\dashrightarrow} \mathbb{1}* & \mathbb{1}* \overset{a}{\dashrightarrow} \mathbb{2}*\\
\mathbb{2}^n\mathbb{0}* \xrightarrow{t} \mathbb{0}^n\mathbb{1}* & \mathbb{2}^n\mathbb{1}* \xrightarrow{t} \mathbb{0}^n\mathbb{2}*
\end{array}$$

$$\overline{X_1 \dots X_n}(c) \xrightarrow{-(-1)^{\sum X_i} t} (X_1 - 1)\dots(X_n - 1)(u)$$

Then the resulting graph is the Lie graph of $\mathscr{L}(\Gamma)$.

The subgraph spanned by a, t, the $\overline{X_1 \dots X_i}(c)$ for $i \leq n-2$ and the $X_1 \dots X_i(u)$ for $i \leq n-3$ is the Lie graph associated to the finite quotient $\Gamma/\operatorname{Stab}_\Gamma(n)$.

Proof. The proof is similar to that of Theorems 3.5 and 3.8, but a bit more tricky. Again we perform the computations in the completion of Γ, still written Γ. Again Γ' is the subgroup generated by all $\overline{X}(c)$ and $X(u)$, for $X \in \{\mathbb{0}, \mathbb{1}, \mathbb{2}\}^*$.

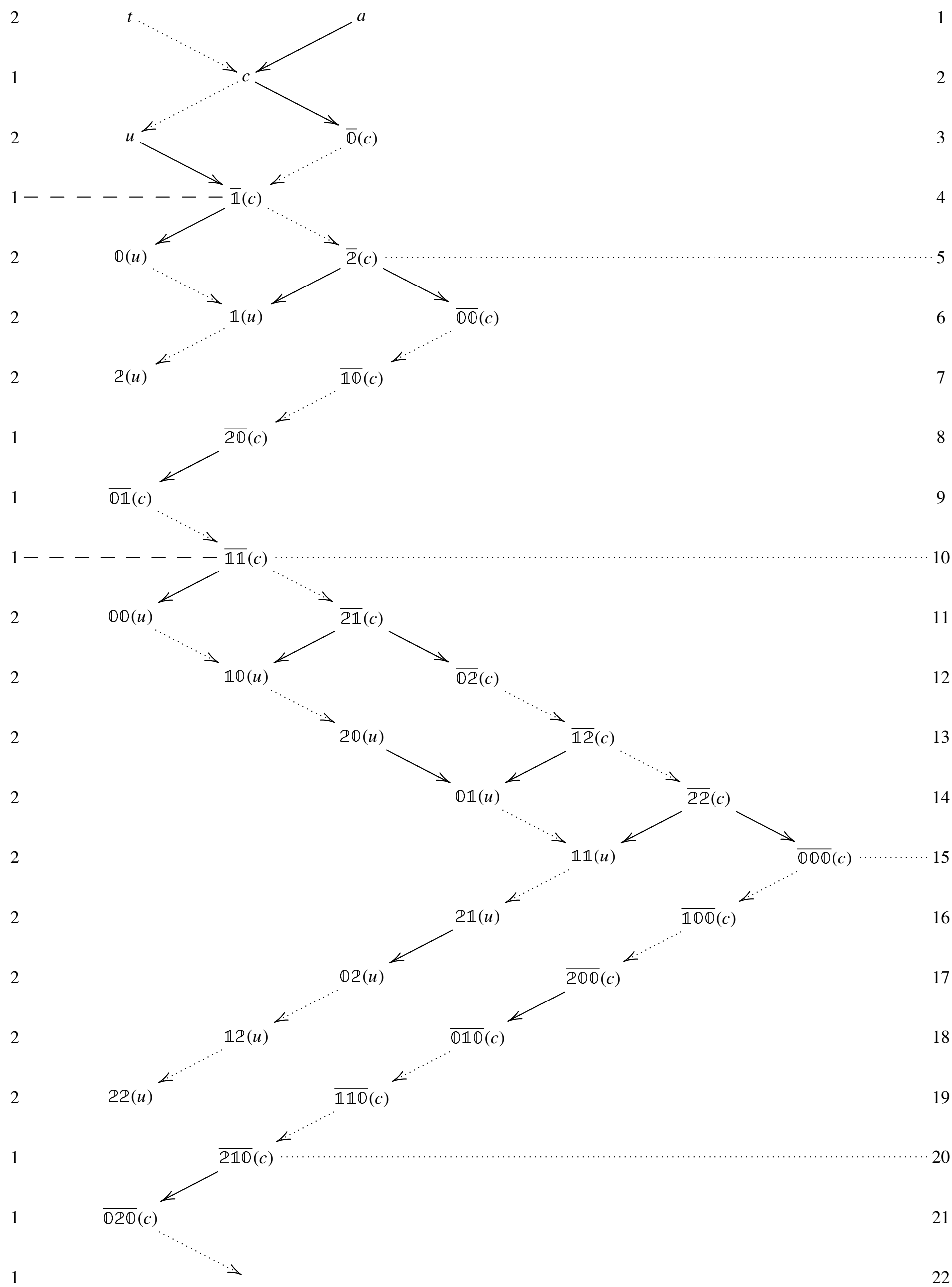


Figure 4. The beginning of the Lie graph of $\mathscr{L}(\Gamma)$. The generator $\mathsf{ad}(t)$ is shown by plain arrows, and the generator $\mathsf{ad}(a)$ is shown by dotted arrows. The left row indicates the dimensions of $\mathscr{L}_n$.

We claim inductively that if $X_i \geq Y_i$ at all positions i, then $X(c) \in \langle Y(c)\rangle^\Gamma$, and similarly for u. Therefore some terms may be neglected in the computations of brackets.

Now we compute $\mathsf{ad}(x)y$ for $x, y \in \{a, t, c, u\}$. Here $\equiv$ means some terms of greater degree have been neglected:

$$\begin{aligned}
&[a, \mathbb{0}*] = \mathbb{1}*, \quad [a, \mathbb{1}*] = \mathbb{2}*, \quad [a, \mathbb{2}*] = 1 \text{ by definition,}\\
&[t, \mathbb{0}*] \equiv [\ll 1, t, a\gg, \ll *, 1, 1\gg] = 1,\\
&[t, \mathbb{1}*] = [\ll a, 1, t\gg, \ll *, *^{-1}, 1\gg] = \mathbb{0}[a, *]\\
&\qquad\qquad \equiv [\ll 1, t, a\gg, \ll *, *^{-1}, 1\gg] \equiv -\mathbb{0}[t, *],\\
&[t, \mathbb{2}*] = [\ll a, 1, t\gg, \ll *, *, *\gg] \equiv \mathbb{0}[a, *] + \mathbb{0}[t, *] + \mathbb{1}[t, *].
\end{aligned}$$

Note that in the last line the "negligible" term $\mathbb{1}[t, *]$ has been kept; this is necessary since sometimes the $\mathbb{0}[t, *]$ term cancels out.

Now we check each of the asserted arrows against the relations described above. First the a arrows are clearly as described, and so are the t arrows on $X(u)$; for instance,

$$\begin{aligned}
\mathsf{ad}(t)\mathbb{2}^n\mathbb{1}*(u) &= \mathbb{0}\,\mathsf{ad}(a)\mathbb{2}^{n-1}\mathbb{1}*(u) + \mathbb{0}\,\mathsf{ad}(t)\mathbb{2}^{n-1}\mathbb{1}*(u) + \mathbb{1}\,\mathsf{ad}(t)\mathbb{2}^{n-1}\mathbb{1}*(u)\\
&\equiv \mathbb{0}^n\big(\mathsf{ad}(a)\mathbb{1}*(u) + \mathsf{ad}(t)\mathbb{1}*(u)\big) \equiv \mathbb{0}^n\mathbb{2}*(u),
\end{aligned}$$

which holds by induction on the length of $*$. Next, the t arrows on $\bar{X}(c)$ agree; for instance,

$$\begin{aligned}
\mathsf{ad}(t)\overline{\mathbb{2}\mathbb{1}^n}(c) &= \mathbb{0}\,\mathsf{ad}(a)\mathbb{1}^n(c) + \mathbb{0}\,\mathsf{ad}(t)\mathbb{1}^n(c) + \mathbb{1}\,\mathsf{ad}(t)\mathbb{1}^n(c)\\
&= \mathbb{0}\mathbb{2}\mathbb{1}^{n-1}(c) + (-1)^n\cdot\mathbb{0}^{n+1}(u) + (-1)^n\cdot\mathbb{1}\mathbb{0}^n(u)\big)\\
&= \overline{\mathbb{0}\mathbb{2}\mathbb{1}^{n-1}}(c) + (-1)^n\cdot\mathbb{1}\mathbb{0}^n(u) \text{ by induction on } n,\\
\mathsf{ad}(t)\overline{\mathbb{2}^n}(c) &= \mathsf{ad}(t)\mathbb{2}\big(\overline{\mathbb{2}^{n-1}}(c)\cdot\mathbb{0}\mathbb{1}^{n-2}(u)\big)\\
&\equiv \mathbb{0}\mathbb{1}^{n-1}(u) + \mathbb{0}\big(-\overline{\mathbb{0}^n}(c) - \mathbb{1}^{n-1}(u)\big) + \mathbb{1}\big(-\overline{\mathbb{0}^n}(c) - \mathbb{1}^{n-1}(u)\big)\\
&\equiv -\overline{\mathbb{0}^{n+1}}(c) - \mathbb{1}^n(u).
\end{aligned}$$

All other cases are similar. Note how the calculation for $\overline{\mathbb{2}\mathbb{1}^n}(c)$ explains the definition of $\bar{X}(c)$: both $\mathbb{0}\mathbb{2}\mathbb{1}^{n-1}(c)$ and $\mathbb{0}^{n+1}(u)$ have degree smaller than $d = \deg\overline{\mathbb{2}\mathbb{1}^n}(c)$ in $\mathscr{L}(\Gamma)$, but they are linearly dependent in $\gamma_{d-1}(\Gamma)/\gamma_d(\Gamma)$.

Finally, we prove that the degrees of $X(c)$ and $X(u)$ are as claimed, by remarking that $\deg c = 3$ and $\deg u = 4$, that $\deg\mathsf{ad}(s)* \geq \deg(*)$ for $s = a, t$ and all words $*$ (so the claimed degrees smaller of equal to their actual value), and that each word of claimed degree n appears only as $\mathsf{ad}(s)*$ for words $*$ of degree at most $n-1$ (so the claimed degrees are greater or equal to their actual value). $\square$

Corollary 3.14. *Define the integers* $\alpha_n = \frac{1}{2}(5 \cdot 3^{n-2} + 1)$ *and the polynomials*

$$Q_2 = 1,$$
$$Q_3 = 1 + 2\hbar + \hbar^2 + \hbar^3 + \hbar^4 + \hbar^5 + \hbar^6,$$
$$Q_n(\hbar) = (1 + \hbar + \hbar^2) Q_{n-1}(\hbar^3) + \hbar + \hbar^{\alpha_n - 2} \quad \text{for } n \geq 4.$$

Then Q_n *is a polynomial of degree* $\alpha_n - 2$, *and the first* $3^{n-2}+1$ *coefficients of* Q_n *and* Q_{n+1} *coincide. The termwise limit* $Q_\infty = \lim_{n\to\infty} Q_n$ *therefore exists.*

The Hilbert–Poincaré series of $\mathscr{L}(\Gamma/\operatorname{\mathsf{Stab}}_\Gamma(n))$ *is* $2\hbar + \hbar^2 Q_n$, *and the Hilbert–Poincaré series of* $\mathscr{L}(\Gamma)$ *is* $2\hbar + \hbar^2 Q_\infty$.

As a consequence, $\Gamma/\operatorname{\mathsf{Stab}}_\Gamma(n)$ *is nilpotent of class* α_n, *and* Γ *has finite width.*

Proof. Consider the sequence of coefficients of $2\hbar + \hbar^2 Q_n$. They are, in condensed form,

$$2, 1, 2^{3^0}, 1^{3^0}, 2^{3^1}, 1^{3^1}, \dots, 2^{3^{n-3}}, 1^{3^{n-3}}, 1^{(3^{n-1}+1)/2}.$$

The i-th coefficient is 2 if there are $\overline{X}(c)$ and $Y(u)$ of degree i in $\Gamma/\operatorname{\mathsf{Stab}}_\Gamma(n)$, and is 1 if there is only $\overline{X}(c)$. All conclusions follow from this remark. $\square$

In quite the same way as for $\ddot{\Gamma}$, we may improve the general result $\Gamma^{(k)} \leq \gamma_{2^k}(\Gamma)$:

Theorem 3.15. *The derived series of* Γ *satisfies* $\Gamma' = \gamma_2(\Gamma)$ *and* $\Gamma^{(k)} = \gamma_5(\Gamma)^{\times 3^{k-2}}$ *for* $k \geq 2$. *We have*

$$\Gamma^{(k)} \leq \gamma_{2+3^{k-1}}(\Gamma) \quad \text{for all } k \in \mathbb{N}.$$

Proof. It is a general fact for a 2-generated group Γ that $\Gamma'' \leq \gamma_5(\Gamma)$. Since $[c, \mathbb{0}(c)] \equiv \mathbb{0}(u)^{-1}$ and $[c, u] \equiv \mathbb{2}(c)^{-1}$ (modulo $\gamma_6(\Gamma)$), we have $[\gamma_2(\Gamma), \gamma_3(\Gamma)] = \gamma_5(\Gamma)$ and therefore $\Gamma'' = \gamma_5(\Gamma)$.

Next, $\gamma_5(\Gamma) = \gamma_3(\Gamma)^{\times 3} \cdot \mathbb{2}(c)$, so $\Gamma^{(3)} = [\gamma_3(\Gamma), c]^{\times 3} = \gamma_5^{\times 3}$, and the claimed formula holds for all $\Gamma^{(k)}$ by induction. Finally $\gamma_{2+3^{j-2}}(\Gamma)^{\times 3} \leq \gamma_{2+3^{j-1}}(\Gamma)$ for all $j = 3, \dots, k$. $\square$

We omit altogether the proofs of the next two results, since they are completely analogous to that of Theorem 3.13.

Theorem 3.16. *Keep the notations of Theorem 3.13. Define furthermore symbols* $\overline{X_1 \dots X_n}(u)$ (*representing elements of* Γ) *by*

$$\overline{\mathbb{2}^n}(u) = \mathbb{2}^n(u) \cdot \mathbb{2}^{n-1}\mathbb{0}(c) \cdot \mathbb{2}^{n-2}\mathbb{0}\mathbb{1}(c) \cdots \mathbb{2}\mathbb{0}\mathbb{1}^{n-2}(c),$$
$$\overline{X}(u) = X(u) \quad \text{for all other } X.$$

Consider the following Lie graph: its vertices are the symbols $\bar{X}(c)$ *and* $\bar{X}(u)$. *Their degrees are given by*

$$\begin{aligned}
\deg \overline{X_1 \dots X_n}(c) &= 1 + \sum_{i=1}^{n} X_i 3^{i-1} + \frac{1}{2}(3^n + 1),\\
\deg \overline{\mathbb{2}^n}(u) &= 3^{n+1},\\
\deg X_1 \dots X_n(u) &= \max\left\{1 + \sum_{i=1}^{n} X_i 3^{i-1} + (3^n+1),\ \frac{1}{2}(9-3^n) + 3\sum_{i=1}^{n} X_i 3^{i-1}\right\}.
\end{aligned}$$

There are two additional vertices, labeled a and t, of degree 1.

Define the arrows as follows, for all $n \geq 1$:

$$\begin{array}{ll}
a \xrightarrow{-t} c & t \overset{a}{\dashrightarrow} c\\
c \xrightarrow{-t} \mathbb{0}(c) & c \xrightarrow{a} u\\
u \xrightarrow{-t} \mathbb{1}(c) & \overline{\mathbb{2}^n}(c) \xrightarrow{-t} \overline{\mathbb{0}^{n+1}}(c)\\
\mathbb{0}* \overset{a}{\dashrightarrow} \mathbb{1}* & \mathbb{1}* \overset{a}{\dashrightarrow} \mathbb{2}*\\
\mathbb{2}^n\mathbb{0}* \xrightarrow{t} \mathbb{0}^n\mathbb{1}* & \mathbb{2}^n\mathbb{1}* \xrightarrow{t} \mathbb{0}^n\mathbb{2}*
\end{array}$$

$$\overline{X_1 \dots X_n}(c) \xrightarrow{-(-1)^{\sum X_i} t} (X_1-1)\dots(X_n-1)(u)$$

$$\begin{array}{ll}
c \xrightarrow{\cdot 3} \overline{\mathbb{0}\mathbb{0}}(c) & \overline{\mathbb{2}^n}(u) \xrightarrow{\cdot 3} \overline{\mathbb{2}^{n+1}}(u)
\end{array}$$

$$*\mathbb{0}(c) \xrightarrow{\cdot 3} *\mathbb{2}(u) \quad \textit{if } 3\deg *\mathbb{0}(c) = \deg *\mathbb{2}(u)$$

Then the resulting graph is the Lie graph of $\mathscr{L}_{\mathbb{F}_3}(\Gamma)$.

The subgraph spanned by a, t, *the* $\overline{X_1 \dots X_i}(c)$ *for* $i \leq n-2$ *and the* $X_1 \dots X_i(u)$ *for* $i \leq n-3$ *is the Lie graph of the Lie algebra* $\mathscr{L}_{\mathbb{F}_3}(\Gamma/\operatorname{Stab}_\Gamma(n))$.

As a consequence, the dimension series of $\Gamma/\operatorname{Stab}_\Gamma(n)$ *has length* 3^{n-1} (*the degree of* $\overline{\mathbb{2}^n}(u)$), *and* Γ *has finite width.*

Proposition 1.8 then implies:

Corollary 3.17. *The growth of* Γ *is at least* $e^{\sqrt{n}}$.

4. Parabolic space

In the natural action of a branch group G on the tree Σ^*, consider a "parabolic subgroup" P, the stabilizer of an infinite ray in Σ^*. (The terminology comes from geometry, where a parabolic subgroup is the stabilizer of a point on the boundary of an appropriate G-space.) Such a parabolic subgroup may be defined directly as follows: let $\omega = \omega_1\omega_2 \cdots \in \Sigma^\infty$ be an infinite sequence. Set $P_0^\omega = G$ and inductively set

$$P_n^\omega = \psi^{-1}(G \times \cdots \times P_{n-1}^\omega \times \cdots \times G),$$

with the P_{n-1}^ω in position ω_n. Set $P^\omega = \bigcap_{n\geq 0} P_n^\omega$.

In the natural tree action (2–1) of G on Σ^* or on Σ^∞ its boundary, P_n^ω is the stabilizer of the point $\omega_1 \dots \omega_n$, and P^ω is the stabilizer of the infinite sequence ω.

The following facts easily follow from the definitions:

Lemma 4.1. $\bigcap_{\omega\in\Sigma^\infty} P^\omega = 1$. *The index of P_n^ω in G is d^n, and that of P^ω is infinite.*

Definition 4.2. Let G be a branch group. A *parabolic space* for G is a homogeneous space G/P, where P is a parabolic subgroup.

Suppose now that G is finitely generated by a set S.

Proposition 4.3 [Bartholdi and Grigorchuk 2000b]. *Suppose that the length $|\cdot|$ on the branch group G is such that, for certain constants λ, μ and for all $g \in \mathsf{Stab}_G(1)$, one has $|g_i| < \lambda|g| + \mu$, where we have written $\psi(g) = (g_1, \dots, g_d)$. Then all parabolic spaces of G have polynomial growth of degree at most* $\log_{1/\lambda}(d)$.

Theorem 4.4. *Let G be a finitely generated branch group. There exists a constant C such that, for any $x_0 \in G$,*

$$\frac{C\,\mathsf{growth}(G/P, x_0P)}{1-\hbar} \geq \frac{\mathsf{growth}\,\mathscr{L}(G)}{1-\hbar}.$$

Proof. Assume G acts on a d-regular tree, and write as before $d' = d - 1$. The proof relies on an identification of the Lie action on group elements and the natural action on tree levels. We first claim that for any $u \in K$ and $W \in \{\mathbb{0}, \mathbb{1}, \dots, \mathbb{d}'\}^*$

$$\deg W(u) \geq \deg(\mathbb{0}^{|W|}(u)) + d_{G/P}(\mathbb{0}^{|W|}, W),$$

where $d(W, X)$ is the length of a minimal word moving W to X in the tree Σ^*.

Therefore the growth of $\mathscr{L}(G)$ and G/P may be compared just by considering the degrees of elements of the form $\mathbb{0}^n(u)$ for some fixed $u \in K$; indeed the other $W(u)$ will contribute a smaller growth to the Lie growth series than the corresponding vertices to the parabolic growth series, and the N finitely many values u may take in a branch portrait description will be taken care of by the constant C.

Now there is a constant $\ell \in \mathbb{N}$ such that $\mathbb{0}^{\ell+m}(u)$ has greater degree than $(\mathbb{d}')^m(u)$ for all $m \in \mathbb{N}$. Indeed there exists $k \in K$ and $\ell \in \mathbb{N}$ such that $[k, u] = \mathbb{0}^\ell(u)$, and then $[\mathbb{0}^m k, \mathbb{d}'^m(u)] = \mathbb{0}^{\ell+m}(u)$, proving the claim.

We may now take $C = \ell N$. The Lie growth series is the sum over all $n \in \mathbb{N}$ and coset representatives $u \in T$ of the power series counting the growth of $W(u)$ over words W of length n. There are N choices for u, and for given u at most ℓ of these power series overlap. $\square$

Note that this result is valid even if the action on the rooted tree is not cyclic, i.e., even if in the decomposition map $G \to G \wr A$ the finite group A is not cyclic. If A is not nilpotent, the Lie algebra $\mathscr{L}$ is no longer isomorphic to G, so the best we can hope for is an inequality bounding the growth of $\mathscr{L}$ by that of G/P.

5. Normal subgroups

Using the notion of a branch portrait, it is not too difficult to determine the exact structure of normal subgroups in a branch group. Consider a p-group G and its p-Lie algebra $\mathscr{L}$ over $\mathbb{F}_p$. Normal subgroups of G correspond to ideals of $\mathscr{L}$, just as subgroups of G correspond to subalgebras of $\mathscr{L}$; and the index of $H \leq G$ is $p^{\dim \mathscr{L}/\mathscr{M}}$, where the subgroup H corresponds to the subalgebra $\mathscr{M}$. This correspondence is not exact, and we shall neither use it nor make it explicit; however it serves as a motivation for relating subgroup growth and the study of Lie algebras. In all cases, sufficient knowledge of $\mathscr{L}$, as well as its finiteness of width, allow an explicit description of the normal subgroup lattice of G.

We focus on the first and most important example, $\mathfrak{G}$, for which we obtain an explicit answer. The computations presented here clearly extend, *mutatis mutandis*, to any regular branch group.

Set $\mathscr{W} = \{\mathbb{0}, \mathbb{1}\}^*$, and order words $X \in \mathscr{W}$ by reverse shortlex: the rank of $X_1 \dots X_n$ is

$$\#X_1 \dots X_n = 1 + \sum_{i=1}^{n} X_i 2^{i-1} + 2^n.$$

(Note that $\#X = \deg X(x)$ according to the definition in Section 3.6.) We write $<$ the order induced by rank.

Theorem 5.1. *The nontrivial normal subgroups of* $\mathfrak{G}$ *are as follows*:

- *there are respectively* 1, 7, 7, 1 *subgroups of index* 1, 2, 4, 8 *corresponding to the lifts to* $\mathfrak{G}$ *of subgroups of* $\mathfrak{G}/[\mathfrak{G}, \mathfrak{G}] = C_2^{\times 3}$;
- *there are* 12 *other subgroups of* $\mathfrak{G}$ *not contained in* K: *six of index* 8, *namely* $\langle [a,c], d^a b\rangle^{\mathfrak{G}}$, $\langle c\rangle^{\mathfrak{G}}$, $\langle x, c^a d\rangle^{\mathfrak{G}}$, $\langle b\rangle^{\mathfrak{G}}$, $\langle [a,d], b^a c\rangle^{\mathfrak{G}}$, *and* $\langle d, x^2\rangle^{\mathfrak{G}}$; *four of index* 16, *namely* $\langle [a,c]\rangle^{\mathfrak{G}}$, $\langle [a,d], x^2\rangle^{\mathfrak{G}}$, $\langle d\rangle^{\mathfrak{G}}$, *and* $\langle [a,d], x^2 d\rangle^{\mathfrak{G}}$; *and two of index* 32, *namely* $\langle [a,d]x^2\rangle^{\mathfrak{G}}$ *and* $\langle [a,d]\rangle^{\mathfrak{G}}$;
- *all normal subgroups* $N \triangleleft \mathfrak{G}$ *contained in* K *are of the form*

$$(*) \qquad W(A; B_1, \dots, B_m; C) := \left\langle A(x) B_1(x^2) \dots B_m(x^2), C(x^2) \right\rangle^{\mathfrak{G}},$$

for words $A, B_i, C \in \mathscr{W}$. *There are functions* $M(A, \{B_i\}, C)$ *and* $S(A, \{B_i\}, C)$ (*defined in the proof*), *with values in* $\mathscr{W}$, *such that there is a unique description* *of N in the form $(*)$ satisfying*

$$B_1 < B_2 < \cdots < B_m \le S(A, \{B_i\}, C) < C \le M(A, \{B_i\}).$$

The index of N is $2^{\#A+\#S(A,\{B_i\},C)}$. The groups can furthermore be subdivided into three types:

I: *$C \le \mathbb{0}^{|A|}$ and $A \le \mathbb{0}^{|C|}\mathbb{1}\mathbb{0}$. Then all B_i are optional, i.e., there are 2^m groups with these A and C, obtained by choosing any subset of the B_i's*;

II: *$C > \mathbb{0}^{|A|}$ and $C \le \mathbb{0}^{|A|+1}$. Then $A = B_1\mathbb{1}$ and all other B_i's are optional*;

III: *$A = \mathbb{0}^n$ and some $B_i = \mathbb{0}^{n-1}$. Then in fact an alternate description exists, obtained by suppressing A and B_i from the description.*

Note that we have only described finite-index subgroups of $\mathfrak{G}$. Since $\mathfrak{G}$ is just-infinite, all its nontrivial normal subgroups have finite index.

We depict the top of the lattice in Figure 5, which shows all normal subgroups of index at most 2^{13} (there are never more than 7 subgroups of a given lesser index).

The first few subgroups of K are described in Table 1, sorted by their index in $\mathfrak{G}$, and identified by their type in $\{(\mathrm{I}), (\mathrm{II}), (\mathrm{III})\}$. We write λ for the empty sequence. An argument $[B_i]$ means that term is optional, and therefore stands for two groups, one with that term and one without.

Among the remarkable subgroups are: $K^{\times 2^n} = \langle \mathbb{0}^n(x)\rangle^{\mathfrak{G}}$, written K_n in [Bartholdi and Grigorchuk 2002]; the subgroup $K^{\times 2^n}\mho_2(K)^{\times 2^{n-1}} = \langle \mathbb{0}^n(x), \mathbb{0}^{n-1}(x^2)\rangle$, written N_n in the same reference; and $\mathsf{Stab}_G(n) = \big\langle \mathbb{0}^{n-3}(\mathbb{1}(x)x^2), \mathbb{0}^{n-2}(x^2)\big\rangle$.

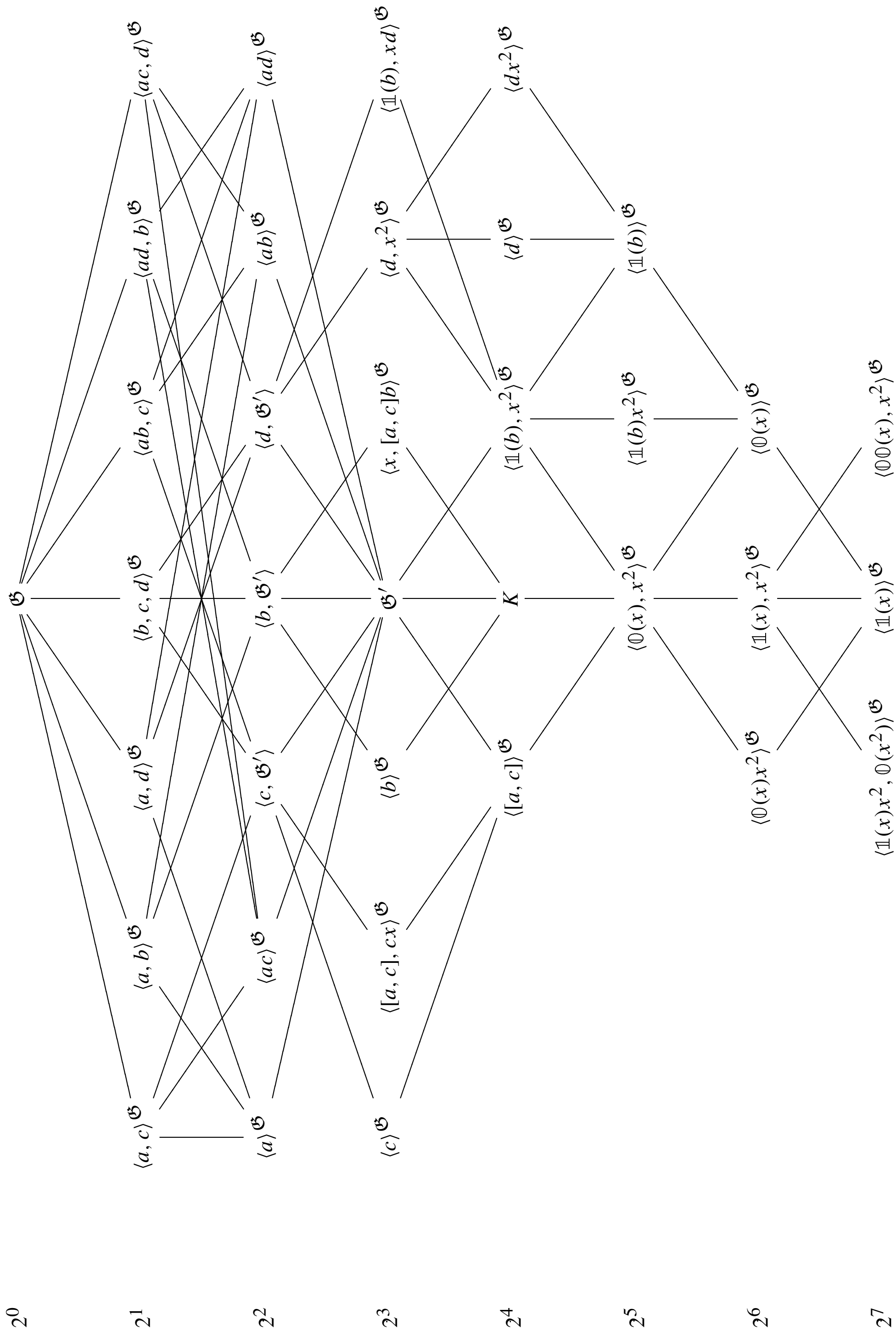
The lattice of normal subgroups of $\mathfrak{G}$ is described in Figure 5. Even though I do not understand completely the lattice's structure, some remarks can be made: the lattice has a fractal appearance; all its nodes have 1 or 3 descendants, and 1 or 3 ascendants. Large portions of it have a grid-like structure. This can be explained by the construction $N \rightsquigarrow N \times N$ of normal subgroups, lending the lattice some self-similarity.

Proof of Theorem 5.1. The first two assertions are checked directly as follows. Let $\mathcal{F}$ be the set of finite-index subgroups of $\mathfrak{G}$ not in K. Consider the finite quotient $Q = \mathfrak{G}/\mathsf{Stab}_6(\mathfrak{G})$, and the preimage P of $\mathfrak{G}$ defined as

$$P = \big\langle a, b, c, d \;\big|\; a^2, b^2, c^2, d^2, bcd, \sigma^i(ad)^4, \sigma^i(adacac)^4 \quad (i = 0 \dots 5)\big\rangle.$$

Clearly the image of $\mathcal{F}$ in Q is at most as large as $\mathcal{F}$, and the preimage of $\mathcal{F}$ in P is at least as large as $\mathcal{F}$. Now we use the algorithms in GAP [GAP 2002] computing the top of the lattice of normal subgroups for finite groups (Q) and finitely presented groups (P). The number of subgroups not contained in K agree in P and Q, so give the structure of the lattice not below K in $\mathfrak{G}$.

Let now N be a normal subgroup of $\mathfrak{G}$, contained in K. If N is nontrivial, then it has finite index [Grigorchuk 2000, Corollary to Proposition 9]. It is easy to


$\mathfrak{G}$
$\langle ac, d\rangle^{\mathfrak{G}}$
$\langle ad, b\rangle^{\mathfrak{G}}$
$\langle ab, c\rangle^{\mathfrak{G}}$
$\langle b, c, d\rangle^{\mathfrak{G}}$
$\langle a, d\rangle^{\mathfrak{G}}$
$\langle a, b\rangle^{\mathfrak{G}}$
$\langle a, c\rangle^{\mathfrak{G}}$
$\langle ad\rangle^{\mathfrak{G}}$
$\langle ab\rangle^{\mathfrak{G}}$
$\langle d, \mathfrak{G}'\rangle$
$\langle b, \mathfrak{G}'\rangle$
$\langle c, \mathfrak{G}'\rangle$
$\langle ac\rangle^{\mathfrak{G}}$
$\langle a\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(b), xd\rangle^{\mathfrak{G}}$
$\langle d, x^2\rangle^{\mathfrak{G}}$
$\langle x, [a,c]b\rangle^{\mathfrak{G}}$
$\mathfrak{G}'$
$\langle b\rangle^{\mathfrak{G}}$
$\langle [a,c], cx\rangle^{\mathfrak{G}}$
$\langle c\rangle^{\mathfrak{G}}$
$\langle dx^2\rangle^{\mathfrak{G}}$
$\langle d\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(b), x^2\rangle^{\mathfrak{G}}$
K
$\langle [a,c]\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(b)\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(b)x^2\rangle^{\mathfrak{G}}$
$\langle \mathbb{0}(x), x^2\rangle^{\mathfrak{G}}$
$\langle \mathbb{0}(x)\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(x), x^2\rangle^{\mathfrak{G}}$
$\langle \mathbb{0}(x)x^2\rangle^{\mathfrak{G}}$
$\langle \mathbb{0}\mathbb{0}(x), x^2\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(x)\rangle^{\mathfrak{G}}$
$\langle \mathbb{1}(x)x^2, \mathbb{0}(x^2)\rangle^{\mathfrak{G}}$
2^0
2^1
2^2
2^3
2^4
2^5
2^6
2^7

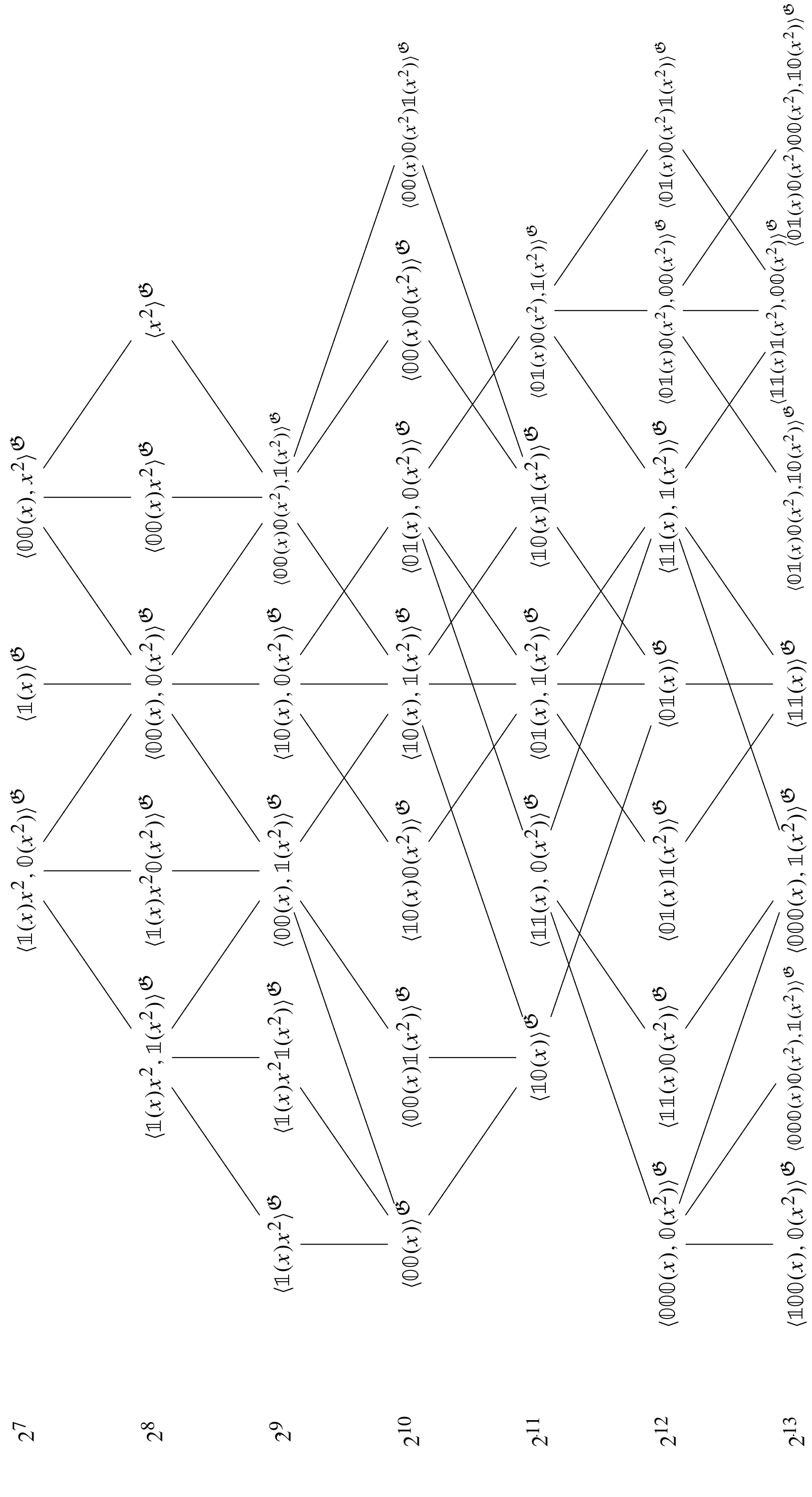


Figure 5. The top of the lattice of normal subgroups of $\mathfrak{G}$, of index at most 2^{13}.

Index	Count	Description		
2^4	1	$W(\lambda;\,;\lambda)_{\mathrm{I}}=K$		
2^5	1	$W(\mathbb{0};\,;\lambda)_{\mathrm{I}}$		
2^6	3	$W(\mathbb{1};\,;\lambda)_{\mathrm{I}}$	$W(\mathbb{0};[\lambda];\mathbb{0})_{\mathrm{I}}$	
2^7	3	$W(\mathbb{00};\,;\lambda)_{\mathrm{I}}$	$W(\mathbb{1};[\lambda];\mathbb{0})_{\mathrm{I}}$	
2^8	5	$W(\mathbb{10};\,;\lambda)_{\mathrm{I}}$	$W(\mathbb{00};\,;\mathbb{0})_{\mathrm{I}}$	
			$W(\mathbb{1};\lambda,[\mathbb{0}];\mathbb{1})_{\mathrm{II}}$	$W(\infty;\lambda,\mathbb{0};\mathbb{1})_{\mathrm{III}}$
2^9	5	$W(\mathbb{10};\,;\mathbb{0})_{\mathrm{I}}$	$W(\mathbb{00};[\mathbb{0}];\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{1};\lambda,[\mathbb{1}];\mathbb{00})_{\mathrm{II}}$
2^{10}	7	$W(\mathbb{01};\,;\mathbb{0})_{\mathrm{I}}$	$W(\mathbb{10};[\mathbb{0}];\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{00};[\mathbb{0}],[\mathbb{1}];\mathbb{00})_{\mathrm{I}}$
2^{11}	5	$W(\mathbb{11};\,;\mathbb{0})_{\mathrm{I}}$	$W(\mathbb{01};[\mathbb{0}];\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{10};[\mathbb{1}];\mathbb{00})_{\mathrm{I}}$
2^{12}	7	$W(\mathbb{000};\,;\mathbb{0})_{\mathrm{I}}$	$W(\mathbb{11};[\mathbb{0}];\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{01};[\mathbb{0}],[\mathbb{1}];\mathbb{00})_{\mathrm{I}}$
2^{13}	7	$W(\mathbb{100};\,;\mathbb{0})_{\mathrm{I}}$	$W(\mathbb{000};[\mathbb{0}];\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{11};[\mathbb{1}];\mathbb{00})_{\mathrm{I}}$
				$W(\mathbb{01};\mathbb{0},[\mathbb{00}];\mathbb{10})_{\mathrm{II}}$
2^{14}	13	$W(\mathbb{010};\,;\mathbb{0})_{\mathrm{I}}$	$W(\mathbb{100};[\mathbb{0}];\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{000};\,;\mathbb{00})_{\mathrm{I}}$
			$W(\mathbb{11};\mathbb{1},[\mathbb{00}];\mathbb{10})_{\mathrm{II}}$	$W(\mathbb{01};\mathbb{0},[\mathbb{00}],[\mathbb{10}];\mathbb{01})_{\mathrm{II}}$
			$W(\infty;\mathbb{1},\mathbb{00};\mathbb{10})_{\mathrm{III}}$	$W(\infty;\mathbb{0},[\mathbb{1}],\mathbb{00};\mathbb{01})_{\mathrm{III}}$
2^{15}	9	$W(\mathbb{010};\,;\mathbb{1})_{\mathrm{I}}$	$W(\mathbb{100};\,;\mathbb{00})_{\mathrm{I}}$	$W(\mathbb{000};[\mathbb{00}];\mathbb{10})_{\mathrm{I}}$
		$W(\mathbb{11};\mathbb{1},[\mathbb{10}];\mathbb{01})_{\mathrm{II}}$	$W(\mathbb{01};\mathbb{0},[\mathbb{01}];\mathbb{11})_{\mathrm{II}}$	$W(\infty;\mathbb{1},\mathbb{10};\mathbb{01})_{\mathrm{III}}$
2^{16}	13	$W(\mathbb{010};\,;\mathbb{00})_{\mathrm{I}}$	$W(\mathbb{100};[\mathbb{00}];\mathbb{10})_{\mathrm{I}}$	$W(\mathbb{000};[\mathbb{00}],[\mathbb{10}];\mathbb{01})_{\mathrm{I}}$
			$W(\mathbb{11};\mathbb{1},[\mathbb{01}];\mathbb{11})_{\mathrm{II}}$	$W(\mathbb{01};\mathbb{0},[\mathbb{01}],[\mathbb{11}];\mathbb{000})_{\mathrm{II}}$
2^{17}	11	$W(\mathbb{110};\,;\mathbb{00})_{\mathrm{I}}$	$W(\mathbb{010};[\mathbb{00}];\mathbb{10})_{\mathrm{I}}$	$W(\mathbb{100};[\mathbb{10}];\mathbb{01})_{\mathrm{I}}$
		$W(\mathbb{000};[\mathbb{00}],[\mathbb{01}];\mathbb{11})_{\mathrm{I}}$		$W(\mathbb{11};\mathbb{1},[\mathbb{11}];\mathbb{000})_{\mathrm{II}}$
2^{18}	19	$W(\mathbb{001};\,;\mathbb{00})_{\mathrm{I}}$	$W(\mathbb{110};[\mathbb{00}];\mathbb{10})_{\mathrm{I}}$	$W(\mathbb{010};[\mathbb{00}],[\mathbb{10}];\mathbb{01})_{\mathrm{I}}$
		$W(\mathbb{100};[\mathbb{10}],[\mathbb{01}];\mathbb{11})_{\mathrm{I}}$	$W(\mathbb{000};[\mathbb{00}],[\mathbb{01}],[\mathbb{11}];\mathbb{000})_{\mathrm{I}}$	

Table 1. Normal subgroups of index up to 2^{18} in $\mathfrak{G}$, contained in K.

see that N contains $C(x^2)$ and $D(x)$ for some words C, D, using for instance the congruence property [Grigorchuk 2000, Proposition 10]; therefore the generators of N may be chosen as

$$\big\{A_1(x)\cdots A_n(x)B_1(x^2)\cdots B_m(x^2),\, A'_1(x)\cdots A'_{n'}(x)B'_1(x^2)\cdots B'_{m'}(x^2),\,\dots,\,C(x^2),\,D(x)\big\},$$

with $A_i^{(j)} < D$ and $B_i^{(j)} < C$ for all i, j.

Taking the commutators of these generators with the appropriately chosen generator among $\{a, b, c, d\}$, we shift the ranks of the A-terms up by 1, and multiplying a generator by another we may get rid of all generators except $C(x^2)$ and the one with A_1 of smallest rank.

We therefore consider all subgroups $W(A; B_1, \dots, B_m; C)$, and seek conditions on A, $\{B_i\}$ and C so that to each normal subgroup in K there corresponds a unique expression of the form $W(A; B_1, \dots, B_m; C)$.

Let first C be minimal such that $C(x^2) \in N$; then take A minimal such that for some $B_1 < \cdots < B_m < C$ we have $A(x)B_1(x^2)\cdots B_m(x^2) \in N$. Take also B'_1 minimal such that $B'_1(x^2)\cdots B'_{m'}(x^2) \in N$ for some B'_i.

Define the functions $M, S : \mathcal{W} \times 2^{\mathcal{W}} \times \mathcal{W} \to \mathcal{W}$ as follows (M stands for "monomial" and S stands for "squares"): Consider $A(x)B_1(x^2)\dots B_m(x^2)$ as an element of $\mathcal{L}_{\mathbb{F}_2}(\mathfrak{G})$, truncated at degree C. Successive commutations with generators $s \in \{a, b, c, d\}$, *according the the rules of Lemma 3.6*, give rise to other elements of $\mathcal{L}_{\mathbb{F}_2}(\mathfrak{G})$. We stress that we use the complete computations of commutators, and not just those in the filtered Lie algebra. Define $M(A, \{B_i\})$ as the minimal word D such that $D(x^2)$ that arises in this process; if no such word occurs, $M(A, \{B_i\}, C) = C$. Define $S(A, \{B_i\})$ as the minimal $B'_{m'}$ such that $B'_1(x^2)\cdots B'_{m'}(x^2)$ occurs in this process; if no such product occurs, $S(A, \{B_i\}, C) = C - 1$.

Now, since $M(A, \{B_i\}, C)(x^2) \in N$, we necessarily have $C \le M(A, \{B_i\})$. Also, all B_i of degree at least $B'_{m'}$ can be replaced by terms of lower degree $B'_1, \dots, B'_{m-1}$. This proves the claimed inequalities. Conversely, if there existed another description $A(x)\tilde{B}_1(x^2)\dots\tilde{B}_m(x^2) \in N$ for another choice of $\tilde{B}$'s, then by dividing we would obtain a product of $B_i(x^2)$ in N, contradicting $B_m < S(A, \{B_i\}, C)$. The data $(A; B_1, \dots, B_m; C)$ subjected to the theorem's constraints therefore correspond bijectively to N's.

The index of N can be computed in $\mathcal{L}_{\mathbb{F}_2}(\mathfrak{G})$. Seeing elements of N as inside $\mathcal{L}$, a vector-space complement of N is spanned by all $\tilde{A}(x)$ of rank less than A, and all $\tilde{B}(x^2)$ of rank less than $S(A, \{B_i\}, C)$.

We consider finally three cases: first assume $C \le \mathbb{0}^{|A|}$ and $|B_1| \ge |A| - 1$. Then $C(x^2)$ gives $\mathbb{0}^{|C|+1}(x^2)\mathbb{0}^{|C|+2}(x)$ by commutation with $\sigma^{|A|}(d)$, which itself gives $\mathbb{0}^{|C|}\mathbb{1}\mathbb{0}(x)$ by commutation with a, so we may suppose $A \le \mathbb{0}^{|C|}\mathbb{1}\mathbb{0}$. Various B_i's can be added, giving the description (I).

Now assume $C > \mathbb{0}^{|A|}$. Then since $A(x)$ would produce $\mathbb{0}^{|A|}(x^2)$ by commutation with an appropriate conjugate of $\sigma^{|A|}(b)$, we must have $A = B_1\mathbb{1}$ so that the same commutation vanishes, giving the description (II).

Finally assume $C \le \mathbb{0}^{|A|}$ and $|B_1| < |A| - 1$. Then necessarily $A = \mathbb{0}^n$; taking appropriate commutations we see that the normal subgroup in question contains $\mathbb{0}^n(x)\mathbb{0}^{n-1}(x^2)$. We may then replace the generator $A(x)B_1(x^2)\dots B_m(x^2)$ by $\mathbb{0}^{n-1}(x^2)B_1(x^2)\dots B_m(x^2)$, and obtain the description (III). $\square$

Corollary 5.2. *Let N be a normal subgroup of $\mathfrak{G}$. Then $N/[N, \mathfrak{G}]$ is an elementary 2-group of rank 1 or 2, unless it is $N = \mathfrak{G}$ (of rank 3).*

Corollary 5.3. *Every normal subgroup of G is characteristic.*

Proof. The automorphism group of $\mathfrak{G}$ is determined in [Bartholdi and Sidki 2003]: it also acts on the binary tree, and is

$$\operatorname{Aut}\mathfrak{G} = \langle G, \mathbb{1}^j\mathbb{0}[a,d] \text{ for all } j \in \mathbb{N}\rangle.$$

It then follows that $[K, \operatorname{Aut}\mathfrak{G}] = \langle \mathbb{0}(x), x^2\rangle^{\mathfrak{G}}$ is a strict subgroup of K; and hence $[N, \operatorname{Aut}\mathfrak{G}] < N$ for any normal subgroup that is generated by expressions in $W(x)$ and $W(x^2)$ for words $W \in \{\mathbb{0}, \mathbb{1}\}^*$. The theorem asserts that all normal subgroups of $\mathfrak{G}$ below K have this form; it then suffices to check, for instance using the algorithms in GAP, that the finitely many normal subgroups of $\mathfrak{G}$ not in K are characteristic. □

Corollary 5.4. *The number b_n of normal subgroups of $\mathfrak{G}$ of index 2^n starts as follows, and is asymptotically $n^{\log_2 3}$. More precisely, we have* $\liminf b_n/n^{\log_2(3)} = 5^{-\log_2 3} \approx 0.078$ *and* $\limsup b_n/n^{\log_2 3} = \frac{2}{9} \approx 0.222$.

index 2^n	2^0	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}
$\lvert\{N \lhd \mathfrak{G}\}\rvert$	1	7	7	7	5	3	3	3	5	5	7	5
	2^{12}	2^{13}	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}	2^{19}	2^{20}	2^{21}	2^{22}	2^{23}
	7	7	13	9	13	11	19	11	13	11	19	15
	2^{24}	2^{25}	2^{26}	2^{27}	2^{28}	2^{29}	2^{30}	2^{31}	2^{32}	2^{33}	2^{34}	
	25	21	37	23	31	23	37	25	37	31	55	

Proof. The number of subgroups of index 2^n behaves in a somewhat erratic way, but is greater when n is of the form $2^k + 2$, so that there is a maximal number of choices for A and C, and is smaller when n is of the form $5 \cdot 2^k + 1$. We compute the numbers F_k and f_k of normal subgroups of $\mathfrak{G}$ contained in K of index 2^n, with respectively $n = 2^k + 2$ and $n = 5 \cdot 2^k + 1$, yielding the upper and lower bounds. The computations are simplified by the fact that for these two values of n there are only subgroups of type I.

We start with the upper bound, when $n = 2^k + 2$. First, for $k = 2$, the subgroups of index 2^n are $W(\mathbb{0};;\mathbb{0})$, $W(\mathbb{0};\lambda;\mathbb{0})$ and $W(\mathbb{1};;\lambda)$, giving $F_2 = 3$. Then, for $k > 2$, the subgroups can of index 2^n can be described as follows:

(1) $W(A\mathbb{1}\mathbb{0}; \mathcal{B}\mathbb{0}; C\mathbb{0})$ for all $W(A\mathbb{0}; \mathcal{B}; C)$ counted in F_{k-1}, except when $C = \mathbb{0}^{k-3}$, when no subgroup appears in F_k, and when $C = \mathbb{0}^{k-2}$, when $C\mathbb{0}$ should be replaced by $\mathbb{0}^{k-3}\mathbb{1}$;

(2) $W(A\mathbb{0}; \mathcal{B}\mathbb{1}; C\mathbb{1})$ for all $W(A; \mathcal{B}; C)$ counted in F_{k-1}, except when $C = \mathbb{0}^{k-3}$, when no subgroup appears in F_k, and when $C = \mathbb{0}^{k-2}$, when $C\mathbb{1}$ should be replaced by $\mathbb{0}^{k-1}$;

(3) $W(A\mathbb{0}; \{A\} \cup \mathcal{B}\mathbb{1}; C\mathbb{1})$, with the same qualifications as above;

(4) $W(\mathbb{0}^{k-2}\mathbb{1};; \mathbb{0}^{k-2})$.

It follows that $F_k = 3(F_{k-1} - 1) + 1$, so $F_k = \frac{2}{9}3^k + 1$ for all $k \geq 2$.

For the lower bound, we have $f_0 = F_2 = 3$; and for $k > 0$, when $n = 5 \cdot 2^k + 1$, the subgroups can of index 2^n can be described as follows:

(1) $W(A\mathbb{1}\mathbb{1}; \mathcal{B}\mathbb{0}; C\mathbb{0})$ for all $W(A\mathbb{1}; \mathcal{B}; C)$ counted in f_{k-1};

(2) $W(A\mathbb{0}\mathbb{1}; \mathcal{B}\mathbb{1}; C\mathbb{1})$ for all $W(A\mathbb{1}; \mathcal{B}; C)$ counted in f_{k-1};

(3) $W(A\mathbb{0}\mathbb{1}; \{A\mathbb{0}\} \cup \mathcal{B}\mathbb{1}; C\mathbb{1})$, with the same qualifications as above;

(4) $W(\mathbb{1}^k\mathbb{0}; ; \mathbb{0}^{k+1})$ and $W(\mathbb{1}^k\mathbb{0}; \mathbb{1}^k; \mathbb{0}^{k+1})$.

It follows that $f_k = 3(f_{k-1} - 2) + 2$, so $F_k = 3^k + 2$ for all $k \geq 0$.

In summary, the number of normal subgroups of index 2^n oscillates between $3^{\log_2((n-1)/5)} + 2$ and $\frac{2}{9}3^{\log_2(n-2)} + 1$ for $n \geq 6$ (when all normal subgroups of $\mathfrak{G}$ are contained in K). These bounds give respectively

$$5^{-\log_2 3}(n-1)^{\log_2 3} \quad \text{and} \quad \frac{2}{9}(n-2)^{\log_2 3}. \qquad \square$$

Note also the following curiosity:

Corollary 5.5. *The number of normal subgroups of index r of G is odd for all r's a power of* 2, *and even* (*in fact*, 0) *for all other r*.

(The same congruence phenomenon holds for the group $C_2 * C_3$, as observed by Thomas Müller [1996].)

Proof. The proof follows from the description of Theorem 5.1. Assume $r = 2^k$. To determine the parity of the number of subgroups of index r, it suffices to consider which $W(A; \mathcal{B}; C)$ expressions have no choices for $\mathcal{B}$. These are precisely the $W(A; ; \mathbb{0}^n)_{\mathrm{I}}$ with $2^{n+1} < \#A \leq 5 \cdot 2^n$, the $W(\mathbb{0}^n\mathbb{1}\mathbb{0}; ; C)_{\mathrm{I}}$ with $2^n < \#C \leq 2^{n+1}$ and the $W(\infty; \mathbb{1}^n, C-1; C)_{\mathrm{III}}$ with $2^{n+1} + 1 < \#C \leq 3 \cdot 2^n + 1$.

Now these last two families yield a subgroup for precisely the same values of k, namely those satisfying $6 \cdot 2^j + 2 \leq k \leq 7 \cdot 2^j + 1$, and therefore contribute nothing modulo 2. The first family contributes a subgroup for all k. $\square$

5.1. ***Normal subgroups in*** $\ddot{\Gamma}$**.** The normal subgroup growth of $\ddot{\Gamma}$ is much larger. As a crude lower bound, consider the quotient $A = \gamma_k(\ddot{\Gamma})/\gamma_{k+1}(\ddot{\Gamma})$, where we take $k = \frac{1}{2}(\alpha_{2n+1} + 1)$. It is abelian of rank 2^n; indeed, the index of $\gamma_k(\ddot{\Gamma})$ is $3^{3^{2n-1} - 2^{n-1} + 1}$, and that of $\gamma_{k+1}(\ddot{\Gamma})$ is $3^{3^{2n-1} + 2^{n-1} + 1}$.

In the vector space $\mathbb{F}_3{}^j$, there are roughly $3^{\binom{j}{2}}$ subspaces; so A has about 3^{4^n} subgroups $S = N/\gamma_{k+1}(\ddot{\Gamma})$, each of them giving rise to a subgroup N of index roughly 3^{9^n}.

It then follows that the number of normal subgroups of $\ddot{\Gamma}$ of index 3^n is at least $3^{n^{\log_3 2}}$, a function intermediate between polynomial and exponential growth. More precise estimations of the normal subgroup growth of $\ddot{\Gamma}$ will be the topic of a future paper.

Acknowledgments

I am immensely grateful to the referee, who helped me clarify many parts of the present and forthcoming papers.

References

[Bartholdi 2000] L. Bartholdi, *Croissance des groupes agissant sur des arbres*, Ph.D. Thesis, Université de Genève, 2000.

[Bartholdi $\geq$ 2005] L. Bartholdi, "Self-similar Lie algebras and groups". in preparation.

[Bartholdi and Grigorchuk 2000a] L. Bartholdi and R. I. Grigorchuk, "Lie methods in growth of groups and groups of finite width", pp. 1–27 in *Computational and geometric aspects of modern algebra* (Edinburgh, 1998), edited by M. Atkinson et al., London Math. Soc. Lecture Note Ser. **275**, Cambridge Univ. Press, Cambridge, 2000. MR 2001h:20046 Zbl 1032.20026

[Bartholdi and Grigorchuk 2000b] L. Bartholdi and R. I. Grigorchuk, "On the spectrum of Hecke type operators related to some fractal groups", pp. 5–45 in *Din. Sist., Avtom. i Beskon. Gruppy*, Tr. Mat. Inst. Steklova **231**, 2000. MR 2002d:37017 Zbl 01729296

[Bartholdi and Grigorchuk 2002] L. Bartholdi and R. I. Grigorchuk, "On parabolic subgroups and Hecke algebras of some fractal groups", *Serdica Math. J.* **28**:1 (2002), 47–90. MR 2003c:20027 Zbl 1011.20028

[Bartholdi and Sidki 2003] L. Bartholdi and S. N. Sidki, "The automorphism tower of groups acting on rooted trees", preprint, 2003. To appear in *Trans. Amer. Math. Soc.* math.GR/0308127

[Bergman 1978] G. M. Bergman, "A note on growth functions of algebras and semigroups", 1978.

[Blessenohl and Laue 1993] D. Blessenohl and H. Laue, "A basis construction for free Lie algebras", *Exposition. Math.* **11**:2 (1993), 145–152. MR 94d:17006 Zbl 0782.17005

[Camina 2000] R. Camina, "The Nottingham group", pp. 205–221 in *New horizons in pro-p groups*, Progr. Math. **184**, Birkhäuser, Boston, 2000. MR 2001f:20054 Zbl 0977.20020

[Caranti 1997] A. Caranti, "Presenting the graded Lie algebra associated to the Nottingham group", *J. Algebra* **198**:1 (1997), 266–289. MR 99b:17019 Zbl 0977.17016

[Ceccherini-Silberstein et al. 2001] T. Ceccherini-Silberstein, F. Scarabotti, and F. Tolli, "The top of the lattice of normal subgroups of the Grigorchuk group", *J. Algebra* **246**:1 (2001), 292–310. MR 2002k:20048 Zbl 0998.20036

[Fabrykowski and Gupta 1991] J. Fabrykowski and N. Gupta, "On groups with sub-exponential growth functions, II", *J. Indian Math. Society* (*N.S.*) **56**:1-4 (1991), 217–228. MR 93g:20053 Zbl 0868.20029

[GAP 2002] The GAP Group, "GAP: Groups, Algorithms, and Programming", 2002, Available at http://www.gap-system.org/. Version 4.3.

[Gelfand and Kirillov 1966] I. M. Gel'fand and A. A. Kirillov, "On fields connected with the enveloping algebras of Lie algebras", *Dokl. Akad. Nauk SSSR* **167** (1966), 503–505. In Russian; translated in *Sov. Math. Dokl.* **7** (1966), 407–409. MR 33 #4108 Zbl 0149.02903

[Grigorchuk 1980] R. I. Grigorčuk, "On Burnside's problem on periodic groups", *Funktsional. Anal. i Prilozhen.* **14**:1 (1980), 53–54. In Russian; translated in *Functional Anal. Appl.* **14**:1 (1980), 41–43. MR 81m:20045 Zbl 0595.20029

[Grigorchuk 1983] R. I. Grigorčuk, "On the Milnor problem of group growth", *Dokl. Akad. Nauk SSSR* **271**:1 (1983), 30–33. In Russian; translated in *Soviet Math. Doklady* **28**:1 (1983), 23–26. MR 85g:20042 Zbl 0547.20025

[Grigorchuk 1989] R. I. Grigorčuk, "On the Hilbert–Poincaré series of graded algebras that are associated with groups", *Mat. Sb.* **180**:2 (1989), 207–225, 304. In Russian; translated in *Math. USSR Sb.* **66**:1 (1990), 211–229. MR 90j:20063 Zbl 0695.16009

[Grigorchuk 2000] R. I. Grigorchuk, "Just infinite branch groups", pp. 121–179 in *New horizons in pro-p groups*, edited by D. S. Marcus du Sautoy and A. Shalev, Progr. Math. **184**, Birkhäuser, Boston, 2000. MR 2002f:20044 Zbl 0982.20024

[Gupta and Sidki 1983] N. Gupta and S. Sidki, "On the Burnside problem for periodic groups", *Math. Z.* **182**:3 (1983), 385–388. MR 84g:20075 Zbl 0513.20024

[Hall 1950] M. Hall, Jr., "A basis for free Lie rings and higher commutators in free groups", *Proc. Amer. Math. Soc.* **1** (1950), 575–581. MR 12,388a Zbl 0039.26302

[de la Harpe 2000] P. de la Harpe, *Topics in geometric group theory*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 2000. MR 2001i:20081 Zbl 0965.20025

[Huppert and Blackburn 1982] B. Huppert and N. Blackburn, *Finite groups. II*, vol. 242, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1982. AMD, 44. MR 84i:20001a Zbl 0477.20001

[Jacobson 1941] N. Jacobson, "Restricted Lie algebras of characteristic p", *Trans. Amer. Math. Soc.* **50** (1941), 15–25. MR 3,103g JFM 0025.30301

[Jennings 1941] S. A. Jennings, "The structure of the group ring of a p-group over a modular field", *Trans. Amer. Math. Soc.* **50** (1941), 175–185. MR 3,34f JFM 0025.24401

[Jennings 1954] S. A. Jennings, "Substitution groups of formal power series", *Canadian J. Math.* **6** (1954), 325–340. MR 15,853f Zbl 0058.02201

[Kaloujnine 1948] L. Kaloujnine, "La structure des p-groupes de Sylow des groupes symétriques finis", *Ann. Sci. École Norm. Sup.* (3) **65** (1948), 239–276. MR 10,505b Zbl 0034.30501

[Klaas et al. 1997] G. Klaas, C. R. Leedham-Green, and W. Plesken, *Linear pro-p-groups of finite width*, vol. 1674, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1997. MR 98m:20028 Zbl 0901.20013

[Krause and Lenagan 1985] G. R. Krause and T. H. Lenagan, *Growth of algebras and Gel'fand–Kirillov dimension*, Research Notes in Mathematics **116**, Pitman, Boston, 1985. Revised edition, Grad. Stud. Math. **22**, Amer. Math. Soc., Providence, 2000. MR 86g:16001 Zbl 0564.16001

[Kukin 1978] G. P. Kukin, "Bases of a free Lie algebra", *Mat. Zametki* **24**:3 (1978), 375–382, 447. MR 80b:17003 Zbl 0404.17013

[Lazard 1954] M. Lazard, "Sur les groupes nilpotents et les anneaux de Lie", *Ann. Sci. Ecole Norm. Sup.* (3) **71** (1954), 101–190. MR 19,529b Zbl 0055.25103

[Leedham-Green 1994] C. R. Leedham-Green, "Pro-p-groups of finite coclass", *J. London Math. Soc.* (2) **50**:1 (1994), 43–48. MR 95j:20021 Zbl 0817.20031

[Li 1996] W. C. W. Li, *Number theory with applications*, vol. 7, Series on University Mathematics, World Scientific Publishing Co. Inc., River Edge, NJ, 1996. MR 98b:11001 Zbl 0849.11006

[Lothaire 1990] M. Lothaire, *Mots*, Langue, Raisonnement, Calcul. [Language, Reasoning, Computation], Hermès, Paris, 1990. MR 94g:05004 Zbl 0862.05001

[Magnus 1940] W. Magnus, "Über Gruppen und zugeordnete Liesche Ringe", *J. Reine Angew. Math.* **182** (1940), 142–149. MR 2,214d Zbl 0025.24201 JFM 66.1208.03

[Müller 1996] T. Müller, "Subgroup growth of free products", *Invent. Math.* **126**:1 (1996), 111–131. MR 97f:20031 Zbl 0862.20019

[Passi 1979] I. B. S. Passi, *Group rings and their augmentation ideals*, vol. 715, Lecture Notes in Mathematics, Springer, Berlin, 1979. MR 80k:20009 Zbl 0405.20007

[Passi and Sehgal 1975] I. B. S. Passi and S. K. Sehgal, "Lie dimension subgroups", *Comm. Algebra* **3** (1975), 59–73. MR 51 #5725 Zbl 0319.20055

[Petrogradsky 1997] V. M. Petrogradsky, "On Lie algebras with nonintegral q-dimensions", *Proc. Amer. Math. Soc.* **125**:3 (1997), 649–656. MR 97e:16055 Zbl 0873.17014

[Petrogradsky 1999] V. M. Petrogradsky, "Growth of finitely generated polynilpotent Lie algebras and groups, generalized partitions, and functions analytic in the unit circle", *Internat. J. Algebra Comput.* **9**:2 (1999), 179–212. MR 2000i:17007 Zbl 0931.17007

[Quillen 1968] D. G. Quillen, "On the associated graded ring of a group ring", *J. Algebra* **10** (1968), 411–418. MR 38 #245 Zbl 0192.35803

[Reutenauer 1993] C. Reutenauer, *Free Lie algebras*, vol. 7, London Mathematical Society Monographs. New Series, The Clarendon Press Oxford University Press, New York, 1993. Oxford Science Publications. MR 94j:17002 Zbl 0798.17001

[Rozhkov 1996] A. V. Rozhkov, "Lower central series of a group of tree automorphisms", *Mat. Zametki* **60**:2 (1996), 225–237, 319. MR 98b:20048 Zbl 0899.20014

[Shalev 1995a] A. Shalev, "Some problems and results in the theory of pro-p groups", pp. 528–542 in *Groups '93 Galway/St. Andrews* (Galway, 1993), vol. 2, edited by C. M. Campbell et al., London Math. Soc. Lecture Note Ser. **212**, Cambridge Univ. Press, Cambridge, 1995. MR 96k:20040 Zbl 0866.20020

[Shalev 1995b] A. Shalev, "Finite p-groups", pp. 401–450 in *Finite and locally finite groups* (Istanbul, 1994), edited by A. V. B. B. Hartley, G. M. Seitz and R. M. Bryant, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **471**, Kluwer Acad. Publ., Dordrecht, 1995. MR 97h:20023 Zbl 0839.20028

[Shalev and Zelmanov 1992] A. Shalev and E. I. Zel'manov, "Pro-p groups of finite coclass", *Math. Proc. Cambridge Philos. Soc.* **111**:3 (1992), 417–421. MR 93e:20030 Zbl 0813.20030

[Sidki 1987] S. Sidki, "On a 2-generated infinite 3-group: subgroups and automorphisms", *J. Algebra* **110**:1 (1987), 24–55. MR 89b:20081b Zbl 0623.20025

[Širšov 1962] A. I. Širšov, "On the bases of a free Lie algebra", *Algebra i Logika Sem.* **1**:1 (1962), 14–19. MR 27 #183

[Strade and Farnsteiner 1988] H. Strade and R. Farnsteiner, *Modular Lie algebras and their representations*, Pure Appl. Math. **116**, Dekker, New York, 1988. MR 89h:17021 Zbl 0648.17003

[Vieira 1998] A. C. Vieira, "On the lower central series and the derived series of the Gupta–Sidki 3-group", *Comm. Algebra* **26**:4 (1998), 1319–1333. MR 99i:20041 Zbl 0899.20015

[Vieira 1999] A. C. Vieira, "On some series of normal subgroups of the Gupta–Sidki 3-group", pp. 728–737 in *Groups St. Andrews 1997* (Bath, 1997), vol. 2, edited by C. M. Campbell et al., London Math. Soc. Lect. Note Ser. **261**, Cambridge Univ. Press, 1999. MR 2000h:20076 Zbl 1004.20502

[Zassenhaus 1940] H. Zassenhaus, "Ein Verfahren, jeder endlichen p-Gruppe einen Lie-Ring mit der Charakteristik p zuzuordnen", *Abhandlungen Math. Sem. Univ. Hamburg* **13** (1940), 200–207. Zbl 0021.20001 JFM 65.0091.01

Received April 23, 2002. Revised October 10, 2003.

LAURENT BARTHOLDI
ÉCOLE POLYTECHNIQUE FÉDÉRALE
SB/IGAT/MAD, BÂTIMENT BCH
CH-1015 LAUSANNE
SWITZERLAND
laurent.bartholdi@epfl.ch
http://mad.epfl.ch/laurent/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

PEAK-INTERPOLATING CURVES FOR $A(\Omega)$ FOR FINITE-TYPE DOMAINS IN $\mathbb{C}^2$

GAUTAM BHARALI

Let Ω be a bounded, weakly pseudoconvex domain in $\mathbb{C}^2$, having smooth boundary. $A(\Omega)$ is the algebra of all functions holomorphic in Ω and continuous up to the boundary. A smooth curve $C \subset \partial\Omega$ is said to be complex-tangential if $T_p(C)$ lies in the maximal complex subspace of $T_p(\partial\Omega)$ for each $p \in C$. We show that if C is complex-tangential and $\partial\Omega$ is of constant type along C, then every compact subset of C is a peak-interpolation set for $A(\Omega)$. Furthermore, we show that if $\partial\Omega$ is real-analytic and C is an arbitrary real-analytic, complex-tangential curve in $\partial\Omega$, compact subsets of C are peak-interpolation sets for $A(\Omega)$.

1. Statement of the main result

Let Ω be a bounded domain in $\mathbb{C}^n$, and let $A(\Omega)$ be the algebra of functions continuous on $\overline{\Omega}$ and holomorphic in Ω. Recall that a compact subset $K \subset \partial\Omega$ is called a *peak-interpolation set for* $A(\Omega)$ if given any $f \in \mathscr{C}(K)$, $f \not\equiv 0$, there exists a function $F \in A(\Omega)$ such that $F|_K = f$ and $|F(\zeta)| < \sup_K |f|$ for every $\zeta \in \overline{\Omega} \setminus K$.

We are interested in determining when a smooth submanifold $M \subset \partial\Omega$ is a peak-interpolation set for $A(\Omega)$. When Ω is a strictly pseudoconvex domain having $\mathscr{C}^2$ boundary and M is of class $\mathscr{C}^2$, the situation is very well understood; see [Henkin and Tumanov 1976; Nagel 1976; Rudin 1978]. In the strictly pseudoconvex setting, M is a peak-interpolation set for $A(\Omega)$ if and only if M is *complex-tangential*, meaning that $T_p(M) \subset H_p(\partial\Omega)$ for all $p \in M$. (Here and in what follows, for any submanifold $M \subseteq \partial\Omega$, $T_p(M)$ will denote the real tangent space to M at the point $p \in M$, while $H_p(\partial\Omega)$ will denote the maximal complex subspace of $T_p(\partial\Omega)$.)

Very little is known, however, when Ω is a weakly pseudoconvex of finite type. (There are several notions of type for domains in $\mathbb{C}^n$, $n \geq 2$, but they all coincide for pseudoconvex domains in $\mathbb{C}^2$. See Section 2 below.) In view of a result by Henkin and Tumanov [1976] or a similar result by Nagel and Rudin [1978], it is still necessary for M to be complex-tangential. It was recently shown [Bharali 2004] that for bounded (weakly) convex domains $\Omega \subset \mathbb{C}^n$ with real-analytic boundaries,

MSC2000: primary 32A38, 32T25; secondary 32C25, 32D99.
Keywords: complex-tangential, finite type domain, interpolation set, pseudoconvex domain.

complex-tangential submanifolds $M \subset \partial\Omega$ are peak-interpolation sets. However, showing even that any smooth compact complex-tangential arc in $\partial\Omega$ is a peak-interpolation set for $A(\Omega)$, for a general smoothly bounded weakly pseudoconvex domain of finite type, is a difficult problem. This is because doing so would necessarily imply that every point in $\partial\Omega$ is a peak point for $A(\Omega)$. Whether or not this is true for general pseudoconvex domains of finite type is an extremely difficult open question in the theory of functions in several complex variables, but this fact is certainly known for smoothly bounded finite type domains in $\mathbb{C}^2$ [Bedford and Fornæss 1978; Fornæss and McNeal 1994; Fornæss and Sibony 1989], and we will use it in one of our results below. In this paper we show, among other things, that when Ω is a bounded domain in $\mathbb{C}^2$, $\partial\Omega$ is real-analytic and $C \subset \partial\Omega$ is a real-analytic curve, it suffices that C be complex-tangential for every compact subset of C to be a peak-interpolation set for $A(\Omega)$.

More precisely, our main result is:

Theorem 1.1. *Let Ω be a bounded pseudoconvex domain in $\mathbb{C}^2$ having smooth boundary, and let $C \subset \partial\Omega$ be a smooth curve.*

(i) *Let $\partial\Omega$ be of class $\mathscr{C}^\infty$ and Ω be of finite type. If C is complex-tangential, and if $\partial\Omega$ is of constant type along C, then each compact subset of C is a peak-interpolation set for $A(\Omega)$.*

(ii) *Let Ω have real-analytic boundary and let $C \subset \partial\Omega$ be a real-analytic complex-tangential curve. Then each compact subset of C is a peak interpolation set for $A(\Omega)$.*

In (ii) above, we do *not* assume that $\partial\Omega$ is of constant type along C.

2. Some notation and introductory remarks

We begin by defining the notion of type.

Definition 2.1. Let $\Omega \subset \mathbb{C}^2$ be a bounded domain having a smooth boundary. Let $p \in \partial\Omega$. The *type* of p, denoted by $\tau(p)$, is the maximum order of contact that the germ of a 1-dimensional complex variety through p can have with $\partial\Omega$ at p. The point p is said to be of finite type if $\tau(p) < \infty$. The domain Ω is said to be of *finite type* if there is an $N \in \mathbb{N}$ such that $\tau(p) \leq N$ for each $p \in \partial\Omega$.

Remark 2.2. Let $\Omega \subset \mathbb{C}^2$ be a smoothly bounded pseudoconvex domain. Suppose $p \in \partial\Omega$ has type $\tau(p) = N$ and there are local holomorphic coordinates $(U; \zeta_1, \zeta_2)$, near p, relative to which $p = 0$ and relative to which $U \cap \partial\Omega$ is defined by

$$\rho(\zeta) = A(\zeta_1) + O\big(v_2^2, |\zeta_1||v_2|\big) - u_2, \tag{2–1}$$

where $\zeta_k := u_k + i v_k$, and $A(\zeta_1) = O(|\zeta_1|^2)$. Then:

(1) N is the leading order in ζ_1 of A.

(2) N is an even number, because Ω is pseudoconvex.

These are consequences of a computation on smoothly bounded pseudoconvex domains in $\mathbb{C}^2$ of finite type at $p \in \partial\Omega$, given in [Fornæss and Stensønes 1987, Lecture 28]. Examining this calculation, we can infer that:

(3) Suppose $\Phi = (\phi_1, \phi_2) : (U, p) \to (\mathbb{C}^2, 0)$ is a smooth change of coordinate such that $\bar{\partial}\phi_1$ and $\bar{\partial}\phi_2$ vanish to infinite order at p, and such that $(U \cap \partial\Omega)$ (with respect to these new coordinates) has a defining function of the form (2–1), where we have written $\zeta_j = \phi_j(z_1, z_2)$ for $j = 1, 2$. Then conclusions (1) and (2) above continue to hold.

We now present some notation. For a $\mathscr{C}^2$ function ϕ defined in some open set in $\mathbb{C}^n$, we set

$$\partial_j\phi = \frac{\partial\phi}{\partial z_j}, \quad \partial_{\bar{j}}\phi = \frac{\partial\phi}{\partial\bar{z}_j},$$
$$\partial^2_{jk}\phi = \frac{\partial^2\phi}{\partial z_j\partial z_k}, \quad \partial^2_{j\bar{k}}\phi = \frac{\partial^2\phi}{\partial z_j\partial\bar{z}_k}, \quad \partial^2_{\bar{j}\bar{k}}\phi = \frac{\partial^2\phi}{\partial\bar{z}_j\partial\bar{z}_k}.$$

If F is a smooth function defined in a neighborhood of $0 \in \mathbb{R}^N$, we define (borrowing our notation from [Bloom 1978a])

$$\begin{aligned} \mathrm{In}(F) &:= \text{the leading homogeneous polynomial} \\ &\qquad \text{in the Taylor expansion of } F \text{ around } 0, \\ \mathrm{ord}(F) &:= \text{the degree of } \mathrm{In}(F). \end{aligned}$$

In what follows, $B(p; r)$ will denote the open Euclidean ball in $\mathbb{C}^2$ centered at $p \in \mathbb{C}^2$ and having radius r, while $D(a; r)$ will denote the open disc in $\mathbb{C}$ centered at $a \in \mathbb{C}$ and having radius r. Several parameters occur in our analysis and the independence of the quantitative estimates in the results below from these parameters will be of some concern. We will express such estimates via the notation $X \lesssim Y$ — meaning that there is a constant $C > 0$, independent of all parameters, such that $X \le CY$.

A standard approach [Henkin and Tumanov 1976; Rudin 1978] to proving that $C \subset \partial\Omega$, with C, $\partial\Omega$ smooth, is a peak-interpolation set makes use of Bishop's theorem:

Theorem [Bishop 1962]. *Let Ω be a bounded domain in $\mathbb{C}^n$. A compact subset $K \subset \partial\Omega$ is a peak-interpolation set for $A(\Omega)$ if and only if $|\mu|(K) = 0$ for every annihilating measure $\mu \perp A(\Omega)$.*

In this theorem, an *annihilating measure* is a regular, complex Borel measure on $\overline{\Omega}$ which, viewed as a bounded linear functional on $\mathscr{C}(\overline{\Omega})$, annihilates $A(\Omega)$. A variation of the aforementioned approach — needed in the proof of our main theorem — involves showing that if for any $p \in C$ there is a small neighborhood $V_p \ni p$ such that for each bump function $\chi \in \mathscr{C}_c^\infty(V_p; [0, 1])$ with $\operatorname{int}\big(\chi^{-1}\{1\}\big) \cap C$ being an open arc in C, there is a sequence of functions $\{h_k\}_{k\in\mathbb{N}}$ such that

(i) $\{h_k\}_{k\in\mathbb{N}} \subset A(\Omega)$ and is uniformly bounded on $\overline{\Omega}$;

(ii) $\lim_{k\to\infty} h_k(z) = 0 \quad$ for all $z \in \overline{\Omega} \setminus (C \cap V_p)$;

(iii) $\lim_{k\to\infty} h_k(z) = \chi(z) \quad$ for all $z \in C \cap V_p$.

We explain in the next section why Theorem 1.1(i) follows from the existence of such a $\{h_k\}_{k\in\mathbb{N}}$.

The key step in our proof is to show that if C is as described in Theorem 1.1(i), then for each $p \in C$ we can find a small neighborhood $V_p \ni p$ such that for any $U \Subset V_p$ for which $C \cap U$ is an arc, there is a smooth function G in V_p that is almost holomorphic with respect to $C \cap V_p$ and peaks on $C \cap \overline{U}$. Further, one requires that this almost holomorphic peak function must approach the value 1 at a controlled rate. We show that

$$|G(z)| \le 1 - C\,\mathrm{dist}[z, C \cap V_p]^{2M} \quad \text{for all } z \in \overline{\Omega} \cap V_p. \tag{2–2}$$

Here $2M$ represents the type of $\partial\Omega$ along C. The above result is strongly reminiscent of [Noell 1985, Lemma 2.1]. In that lemma, if C — where C is *not necessarily* complex-tangential, but $\partial\Omega$ is of type $2M$ along C — has the property that at each $p \in C$ there is a holomorphic function, smooth up to $\partial\Omega$, that peaks on a small closed sub-arc of C passing through p, then we can find a holomorphic peak function, smooth up to $\partial\Omega$, that satisfies the estimate (2–2). In our situation we do not, of course, have holomorphic functions that peak locally along C. However, we can use some of Noell's ideas (which in turn rely on an estimate from [Bloom 1978a]) and exploit the complex-tangency of C to construct an *almost-holomorphic* local peak function that satisfies good estimates. This construction is presented in Section 4.

We complete the proof of Theorem 1.1 in Section 5. Part (i) of the theorem will follow from the construction of the family $\{h_k\}_{k\in\mathbb{N}}$ described above. Each h_k is, near C, a holomorphic correction of the k-th power of G (G as introduced above). This correction is achieved by solving an appropriate $\bar\partial$-equation in $\overline{\Omega}$, and the estimate (2–2) is used to show that h_k satisfies the three properties listed above. Theorem 1.1(ii) will follow from the fact that in the real-analytic setting $\partial\Omega$ is of constant type along C *except* for a discrete set of points in C. Using part (i) of the theorem and the fact that each point in this discrete set is a peak point for $A(\Omega)$, we deduce part (ii).

3. A technical lemma

In this section, we present an abstract lemma that is instrumental to the proof of our main theorem.

Definition 3.1. Given an open set $V \subset \mathbb{R}^N$, a *bump function* f in V is a function belonging to $\mathscr{C}_c^\infty(V;[0,1])$ such that $\mathrm{int}\big(f^{-1}\{1\}\big) \neq \varnothing$.

Lemma 3.2. *Let Ω be a bounded domain in $\mathbb{C}^2$ having smooth boundary and let C be a smooth curve in $\partial\Omega$. Assume that for each $p \in C$, there exists a small neighborhood V_p of p such that for each bump function $\chi \in \mathscr{C}_c^\infty(V_p;[0,1])$ for which $\mathrm{int}\big(\chi^{-1}\{1\}\big) \cap C$ is an arc, we can find a sequence of functions $\{h_k\}_{k\in\mathbb{N}} \subset A(\Omega)$ (depending on χ) satisfying*

(i) $\{h_k\}_{k\in\mathbb{N}} \subset A(\Omega)$ *is uniformly bounded on* $\overline{\Omega}$;

(ii) $\lim_{k\to\infty} h_k(z) = 0$ *for all* $z \in \overline{\Omega} \setminus (C \cap V_p)$;

(iii) $\lim_{k\to\infty} h_k(z) = \chi(z)$ *for all* $z \in C \cap V_p$.

Then C is a countable union of peak-interpolation sets for $A(\Omega)$.

Remark 3.3. A form of this lemma is true if Ω is a bounded domain in $\mathbb{C}^n$ and C is replaced by $M \subset \partial\Omega$, where M is a smooth submanifold of $\partial\Omega \cap U$, U being an open subset of $\mathbb{C}^n$. However, to be able to derive the conclusion of the lemma in this new setting with $\dim_{\mathbb{R}}(M) > 1$, one would have to produce, for *every* bump function $\chi \in \mathscr{C}_c^\infty(V_p;[0,1])$ (not *merely* those for which $\mathrm{int}\big(\chi^{-1}\{1\}\big) \cap M$ is nice), an $h \in A(\Omega)$ such that $\{h^k : k \in \mathbb{N}\}$ would satisfy conditions (i)–(iii) above. Being able to find such an h could be rather difficult if $\dim_{\mathbb{R}}(M) > 1$, because $\mathrm{int}\big(\chi^{-1}\{1\}\big) \cap M$ could be structurally quite complicated in this situation. We add that if $\partial\Omega$ is *strictly* pseudoconvex, a less exacting form of the above lemma — see, for instance, [Henkin and Tumanov 1976, Lemma 6] — suffices to infer peak-interpolation in higher dimensions.

Proof of Lemma 3.2. Fix $p \in C$. We may assume that $C \cap V_p$ is an arc in C. Let K be any compact subset of $C \cap V_p$ and let μ be any annihilating measure. Then

$$K = (C \cap V_p) \setminus \textstyle\coprod_{k\in\mathbb{N}} \mathscr{A}_k,$$

where each $\mathscr{A}_k$ is an *open* sub-arc of $C \cap V_p$. If we could show that $\mu(\mathscr{A}_k) = 0$ for each k and that $\mu(C \cap V_p) = 0$, we could conclude by the additivity of μ that $\mu(K) = 0$.

Let $\mathscr{C} \subset C \cap V_p$ be any *closed* sub-arc of C. Let $\{\mathfrak{D}_\nu\}_{\nu\in\mathbb{N}}$ be a shrinking family of compact subsets of $\mathbb{C}^2$ such that

(a) $\mathfrak{D}_{\nu+1} \subset \operatorname{int}(\mathfrak{D}_\nu)$,

(b) $\bigcap_{\nu\in\mathbb{N}} \mathfrak{D}_\nu = \mathscr{C}$,

(c) $\mathfrak{D}_\nu \subseteq V_p$,

(d) $C \cap \mathfrak{D}_\nu$ is an arc.

Let $\chi_\nu \in \mathscr{C}_c^\infty(V_p;\ [0, 1])$ be a bump function with

$$\chi_\nu|_{\mathfrak{D}_{\nu+1}} \equiv 1 \quad \text{and} \quad \operatorname{supp} \chi_\nu \subseteq \mathfrak{D}_\nu.$$

Finally, define $\{h_{k,\nu}\}_{k\in\mathbb{N}}$ to be the sequence of functions corresponding to χ_ν given by the hypothesis of this lemma.

Choose any $\mu \perp A(\Omega)$. By the bounded convergence theorem,

$$0 = \lim_{k\to\infty} \int_{\overline{\Omega}} h_{k,\nu}\, d\mu = \int_{C\cap V_p} \chi_\nu\, d\mu.$$

Another passage to the limit yields $\mu(\mathscr{C}) = 0$, and this is true for any $\mu \perp A(\Omega)$. As μ is a regular measure, this shows that $\mu(\mathscr{A}) = 0$ for any *open* sub-arc $\mathscr{A} \subset C \cap V_p$; in particular $\mu(C \cap V_p) = 0$. Let $\mathscr{V}_p$ be any neighborhood of p such that $\mathscr{V}_p \Subset V_p$. In view of our remarks in the first paragraph of this proof we have just shown that $|\mu|(C \cap \overline{\mathscr{V}_p}) = 0$ for any $\mu \perp A(\Omega)$. By Bishop's theorem, $C \cap \overline{\mathscr{V}_p}$ is a peak-interpolation set for $A(\Omega)$. Letting p vary over a countable dense subset of C, we have the desired result. □

4. Constructing an almost holomorphic function that peaks locally on C

Let $p \in \partial\Omega$. In this section, we will study $\partial\Omega$ near p with respect to a convenient system of local coordinates that are almost holomorphic with respect to C (near p), where Ω and C are as in Theorem 1.1(i). The following lemma asserts the existence of local coordinates having the desired properties:

Lemma 4.1. *Let Ω be a bounded domain in $\mathbb{C}^2$ having smooth boundary and let $C \subset \partial\Omega$ be a complex-tangential curve. Let $p \in C$. There is a neighborhood $\omega \ni p$ and a $\mathscr{C}^\infty$-diffeomorphism $\Phi : (\omega, p) \to (\mathbb{C}^2, 0)$ which is almost holomorphic with respect to $(C \cap \omega)$ and such that, writing $(\zeta_1, \zeta_2) := \Phi(z_1, z_2)$, we have:*

(1) *$\Phi(C \cap \omega) \subset \{(\zeta_1, \zeta_2) : \operatorname{im}(\zeta_1) = \zeta_2 = 0\}$.*

(2) *$\Phi(\partial\Omega \cap \omega)$ is defined by a defining function of the form*

$$\rho(\zeta) = A(\zeta_1) + B(\zeta_1)v_2 + R(\zeta_1, v_2) - u_2,$$

where $\zeta_k = u_k + iv_k$ for $k = 1, 2$, $A(\zeta_1) = O(|\zeta_1|^2)$, $R(\zeta_1, v_2) = O(|v_2|^2)$, and

$$A(u_1) = B(u_1) = 0 \quad \textit{and} \quad \nabla A(u_1) = \nabla B(u_1) = 0 \quad \textit{for all } u_1 \textit{ near } 0.$$

Proof. Without loss of generality, we may let p be the origin, and assume that, near p, $\partial\Omega$ is defined by

$$r(z_1, z_2) = h(z_1, \operatorname{im} z_2) - \operatorname{Re} z_2,$$

where $h(0) = 0$ and $\nabla h(0) = 0$.

Let ω be a neighborhood of $p = 0$ and let $\mathcal{M} \subset \partial\Omega$ be the smooth 2-manifold of ω formed by the integral curves to the vector-field $-\mathbb{J}(\nabla r)$ passing through $(C \cap \omega)$. $\mathcal{M}$ is totally real. Let

$$\gamma = (\gamma_1, \gamma_2) : (B(0; \varepsilon), (u_1, u_2) = 0) \to ((\mathcal{M} \cap \omega), p = 0)$$

parametrize $\mathcal{M}$ near $p = 0$ in such a way that, for each c, $\operatorname{Image}(\gamma|_{\{u_2 = c\}})$ is an integral curve to the unit section of $T(\mathcal{M}) \cap H(\partial\Omega)|_{\mathcal{M}}$, with

$$\operatorname{Image}(\gamma|_{\{u_2=0\}}) = C \cap \omega \quad \text{and} \quad \frac{\partial \gamma(0,0)}{\partial u_2} = -\mathbb{J}(\nabla r)(0,0).$$

Shrinking ω if necessary, we construct a diffeomorphism $\Phi : (\omega, p = 0) \to (\mathbb{C}^2, 0)$ of class $\mathscr{C}^\infty$ that is almost holomorphic with respect to $(\mathcal{M} \cap \omega)$, by defining

$$\Phi^{-1}(\zeta_1, \zeta_2) = \big(\Gamma_1(\zeta_1, -i\zeta_2),\ \Gamma_2(\zeta_1, -i\zeta_2)\big) := \eta(\zeta_1, \zeta_2),$$

where $\zeta_k := u_k + i v_k$ for $k = 1, 2$ and Γ_k is an almost holomorphic extension of γ_k for $k = 1, 2$. By construction,

$$\begin{aligned} &\Phi(\mathcal{M} \cap \omega) \subset \{(\zeta_1, \zeta_2) : v_1 = u_2 = 0\}, \\ &\Phi(C \cap \omega) \subset \{(\zeta_1, \zeta_2) : v_1 = \zeta_2 = 0\}. \end{aligned} \tag{4–1}$$

Now, $\Phi(\partial\Omega \cap \omega)$ is defined by

$$\rho(\zeta_1, \zeta_2) = r \circ \Phi^{-1}(\zeta_1, \zeta_2).$$

We expand ρ around the origin in a Taylor series. We make use of the fact that Γ_k are almost holomorphic with respect to $\{(\zeta_1, \zeta_2) |\ v_1 = v_2 = 0\}$ to get

$$\rho(\zeta) = 2 \operatorname{Re}\left(\sum_{j=1}^{2} \frac{\partial r}{\partial z_j}(\eta(0,0))\Big(\frac{\partial \Gamma_j}{\partial \zeta_1}(0,0)\zeta_1 + (-i)\frac{\partial \Gamma_j}{\partial \zeta_2}(0,0)\zeta_2\Big)\right) + O(|\zeta|^2).$$

Using the fact that

$$\frac{\partial \Gamma_j}{\partial \zeta_k}(0,0) = \frac{\partial \gamma_j}{\partial u_k}(0,0) \quad \text{for } j, k = 1, 2,$$

we get

$$\begin{aligned}\rho(\zeta) = 2\,\mathrm{Re}\biggl(\sum_{j=1}^{2}\frac{\partial r}{\partial z_j}(\gamma(0,0))\frac{\partial\gamma_j}{\partial u_1}(0,0)\zeta_1 \\ + (-i)\sum_{j=1}^{2}\frac{\partial r}{\partial z_j}(\gamma(0,0))\frac{\partial\gamma_j}{\partial u_2}(0,0)\zeta_2\biggr) + O(|\zeta|^2) \\ = 2\,\mathrm{Re}\biggl((-i)\sum_{j=1}^{2}\frac{\partial r}{\partial z_j}(\gamma(0,0))\frac{\partial\gamma_j}{\partial u_2}(0,0)\zeta_2\biggr) + O(|\zeta|^2) \\ = -u_2 + O(|\zeta|^2).\end{aligned}$$

The second equality follows from the complex-tangency of $\mathrm{Image}\bigl(\gamma(\,\cdot\,,0)\bigr)$, which implies

$$\sum_{j=1}^{2}\frac{\partial r}{\partial z_j}(\gamma(u_1,0))\frac{\partial\gamma_j}{\partial u_1}(u_1,0) = 0 \quad \text{for all } u_1 \in (-\varepsilon,\varepsilon),$$

and the last equality follows from the normalization condition on $\partial\gamma(0,0)/\partial u_2$. We see that the only term in the expansion above that has first order in either ζ_1 or ζ_2 is $-u_2$. Hence, the hypersurface $\Phi(\partial\Omega\cap\omega)$ is tangent at 0 to the hyperplane $\boldsymbol{H} := \bigl\{(\zeta_1,\zeta_2)\in\mathbb{C}^2 \mid u_2 = 0\bigr\}$. Thus, we can find, near $0\in\mathbb{C}^2$, a defining function — and for convenience of notation, we will continue to call it ρ — having the form

$$\rho(\zeta) = A(\zeta_1) + B(\zeta_1)v_2 + R(\zeta_1,v_2) - u_2, \tag{4–2}$$

where $A(\zeta_1) = O(|\zeta_1|^2)$ and $R(\zeta_1,v_2) = O(|v_2|^2)$. Then, since $\Phi(C\cap\omega)$ is contained in $\Phi(\partial\Omega\cap\omega)$, setting $v_1 = \zeta_2 = 0$ in (4–2), we get

$$A(u_1) = 0 \quad \text{for all } (u_1,0)\in\Phi(C\cap\omega). \tag{4–3}$$

And since $\Phi(\mathcal{M}\cap\omega)\subset\Phi(\partial\Omega\cap\omega)$, setting $v_1 = u_2 = 0$ in (4–2), we see that $B(u_1)v_2 + O(|v_2|^2) = 0$ for all (u_1,v_2) belonging to a small neighborhood 0. Thus

$$B(u_1) = 0 \quad \text{for all } (u_1,0)\in\Phi(C\cap\omega). \tag{4–4}$$

By construction, $(\nabla\rho)(u_1,v_2)$ is a normal vector to $\Phi(\mathcal{M}\cap\omega)$ for all $(u_1,v_2)\in\Phi(\mathcal{M}\cap\omega)$. This implies that $T_{(u_1,v_2)}\bigl(\Phi(\partial\Omega\cap\omega)\bigr) = \boldsymbol{H}$ for all $(u_1,v_2)\in\Phi(\mathcal{M}\cap\omega)$. Computing $(\nabla\rho)(u_1,v_2)$, we see that $\nabla A(u_1) + \nabla B(u_1)v_2 = 0$ for all (u_1,v_2) in a neighborhood of 0. Thus

$$\nabla A(u_1) = \nabla B(u_1) = 0 \quad \text{for all } (u_1,0)\in\Phi(C\cap\omega). \tag{4–5}$$

By (4–3), (4–4) and (4–5), we have the desired result. □

We now state the key lemma of this paper. It concerns the construction of an almost holomorphic peak function of the type discussed in Section 2.

Proposition 4.2. *Let Ω be a bounded pseudoconvex domain in $\mathbb{C}^2$ of finite type, and let $\partial\Omega$ be of class $\mathscr{C}^\infty$. Let $C \subset \partial\Omega$ be a complex-tangential curve of class $\mathscr{C}^\infty$, and let $\partial\Omega$ be of constant type $2M$ along C. Let $p \in C$. There exists a neighborhood $V \equiv V(p)$ of p and a uniform constant $C > 0$, and for any open set $U \Subset V$ such that $C \cap U$ is an arc, there is a neighborhood $V_1 \equiv V(p, U)$ of p satisfying $C \cap V_1 = C \cap V$ and a function $G \in \mathscr{C}^\infty(V_1)$ — G depending on p and U — that satisfies*

(1) $G^{-1}\{1\} = C \cap \overline{U}$;

(2) $\bar{\partial}G$ *vanishes to infinite order on* $V \cap C$;

(3) $|G(z)| \leq 1 - C\,\mathrm{dist}[z, C \cap V]^{2M}$ *for each* $z \in \overline{\Omega} \cap V_1$.

Proof. Let $\omega \ni p$ and $\Phi : (\omega, p) \to (\mathbb{C}^2, 0)$ be the change of coordinate described in Lemma 4.1. Let $\Phi(\partial\Omega \cap \omega)$ be defined by

$$\rho(\zeta_1, \zeta_2) = A(\zeta_1) + B(\zeta_1)v_2 + R(\zeta_1, v_2) - u_2. \tag{4–6}$$

Consider a point $(x_0, 0) \in \Phi(C \cap \omega)$ and let

$$\varrho_{x_0}(\zeta_1^*, \zeta_2) = \mathscr{A}_{x_0}(\zeta_1^*) + \mathscr{B}_{x_0}(\zeta_1^*)v_2 + \mathscr{R}_{x_0}(\zeta_1^*, v_2) - u_2 \tag{4–7}$$

represent the expansion of ρ in (4–6) around $(x_0, 0)$, where $\zeta_1^* := \zeta_1 - x_0$.

Claim 1. Shrinking ω if necessary, there is a $c > 0$ such that

$$A(u_1 + iv_1) \geq cv_1^{2M}, \quad \text{for all } \zeta_1 \text{ such that } \zeta \in \Phi(\omega). \tag{4–8}$$

As $A(x_0) = B(x_0) = 0$ and $\nabla A(x_0) = \nabla B(x_0) = 0$ for each $(x_0, 0) \in \Phi(\partial\Omega \cap \omega)$, the right-hand side of (4–7) represents a defining function of the form (2–1). By Remark 2.2(3), the function $\mathscr{A}_{x_0}$ in (4–7) must vanish to order $2M$ at 0, whereby the function A in (4–6) must vanish precisely to order $2M$ at each $(u_1, 0) \in \Phi(\partial\Omega \cap \omega)$. Now write

$$A(u_1 + iv_1) = a_J(u_1)v_1^J + O(|v_1|^{J+1}), \tag{4–9}$$

where J is the least positive integer k such that $a_k \not\equiv 0$ near $u_1 = 0$. By our remarks above, it is clear that $J \leq 2M$. But, if $J < 2M$, then if $\tilde{u}_1$ is such that $a_J(\tilde{u}_1) \neq 0$, then A vanishes to order $< 2M$ at $u_1 + iv_1 = \tilde{u}_1$, which contradicts our remarks above. Thus, $J = 2M$ in (4–9) and

$$A(u_1 + iv_1) = a_{2M}(u_1)v_1^{2M} + O(|v_1|^{2M+1}).$$

and $a_{2M}(0) \neq 0$. Now recall that Φ is almost-holomorphic with respect to $(\mathscr{M} \cap \omega)$. If, in fact, $(u_1 + iv_1, u_2 + iv_2)$ were *holomorphic* coordinates, the pseudoconvexity of Ω would have implied that

$$\alpha : (u_1, v_1) \mapsto a_{2M}(u_1) v_1^{2M} \text{ is subharmonic,}$$
$$\Delta\alpha(u_1, v_1) > 0 \text{ off } \{v_1 = 0\}, \text{ and } (u_1, v_1) \text{ close to } 0.$$

This would have implied that $a_{2M}(u_1) > 0$ for u_1 close to 0 (the second statement above follows from an obvious calculation). In our present situation, the coordinates $(u_1 + iv_1, u_2 + iv_2)$ differ from holomorphic ones by terms vanishing to arbitrarily high order along $(C \cap \omega)$. From the last two facts, we can conclude, after shrinking ω if necessary, that

$$a_{2M}(u_1) > 0 \quad \text{for all } (u_1, 0) \in \Phi(\partial\Omega \cap \omega).$$

From this final fact, we deduce (4–8). Hence the claim.

Claim 2. We can find $\omega_1 \Subset \omega$ and a uniform constant $T > 0$ such that

$$B(\zeta_1)^2 \leq T A(\zeta_1) \quad \text{for all } \zeta \in \Phi(\overline{\Omega} \cap \omega_1). \tag{4–10}$$

To see this, we use a procedure originating in [Bloom 1978a, Section 3]. Write $q = (x_0, 0) \in \Phi(C \cap \omega)$. The positivity of the Levi form for $\partial\Omega$ on the complex tangent vectors implies that, *were* $(u_1 + iv_1, u_2 + iv_2)$ *holomorphic coordinates*, there would be a $\delta > 0$ such that the function $\mathfrak{L}$ induced by the Levi form

$$\mathfrak{L} : D(x_0; \delta) \times (-\delta, \delta) \to \mathbb{R}$$

defined by

$$\mathfrak{L} = |\partial_{\bar{2}}\rho|^2 \; \partial^2_{1\bar{1}}\rho + |\partial_{\bar{1}}\rho|^2 \; \partial^2_{2\bar{2}}\rho - 2\mathrm{Re}\left(\partial_1\rho \; \partial_{\bar{2}}\rho \; \partial^2_{\bar{1}2}\rho\right)$$

would be nonnegative (notice that $\mathfrak{L}$ is independent of u_2). In our present situation, however, $\mathfrak{L}(u_1, v_2) \geq 0$ for all $(u_1, v_2) \in \Phi(\mathcal{M} \cap \omega)$.

Write

$$\mathfrak{L}(\zeta_1, v_2) = \mathfrak{L}^{(0)}(\zeta_1) + v_2\mathfrak{L}^{(1)}(\zeta_1) + v_2^2\mathfrak{L}^{(2)}(\zeta_1) + O(|v_2|^3).$$

It has been shown in [Bloom 1978a] that if $\operatorname{ord} B < \operatorname{ord} A$, then

$$\begin{aligned} \mathrm{In}(\mathfrak{L}^{(0)}) &= \tfrac{1}{4}\mathrm{In}(\partial^2_{1\bar{1}}A), \qquad & \operatorname{ord}\mathfrak{L}^{(0)} &= \operatorname{ord} A - 2, \\ \mathrm{In}(\mathfrak{L}^{(1)}) &= \tfrac{1}{4}\mathrm{In}(\partial^2_{1\bar{1}}B), \qquad & \operatorname{ord}\mathfrak{L}^{(1)} &= \operatorname{ord} B - 2. \end{aligned} \tag{4–11}$$

If already $2 \operatorname{ord} B \geq \operatorname{ord} A$, then (4–10) would follow trivially. Thus, *assume* that $2 \operatorname{ord} B < \operatorname{ord} A$. Write $r = \operatorname{ord} B$. We have

$$\frac{1}{\lambda^{2r-2}}\mathfrak{L}(\lambda(u_1 + i0), \lambda^r v_2) \geq 0$$

for all $(u_1, v_2) \in (x_0 - \delta, x_0 + \delta) \times (-\delta, \delta)$ and $\lambda \in \mathbb{R}_+$. But from (4–11) and our assumption, we get

$$\lim_{\lambda \to 0_+} \frac{1}{\lambda^{2r-2}} \mathfrak{L}(\lambda u_1, \lambda^r v_2) = \frac{v_2}{4} \operatorname{In}(\partial^2_{1\bar{1}} B)(\zeta_1). \tag{4–12}$$

Write

$$B(u_1 + i v_1) = b_J(u_1) v_1^J + O(|v_1|^{J+1}),$$

where J is the least positive integer k such that $b_k \not\equiv 0$ near $u_1 = 0$. By Lemma 4.1(2), $J \geq 2$, whence $\operatorname{In}(B)$ is nonharmonic near 0. So, as v_2 occurs linearly in the right-hand side of (4–12), it is impossible that

$$\tfrac{1}{4} v_2 \operatorname{In}(\partial^2_{1\bar{1}} B)(u_1) \geq 0, \quad \text{for all } (u_1, v_2) \in (x_0 - \delta, x_0 + \delta) \times (-\delta, \delta).$$

This results in a contradiction. So $2 \operatorname{ord} B \geq \operatorname{ord} A$, which, in conjunction with the positivity of A, namely (4–8), yields (4–10).

Finally, define $H : \Phi(\overline{\Omega} \cap \omega_1) \to \mathbb{C}$ by

$$H(\zeta) = \zeta_2 - \alpha \zeta_2^2,$$

for $\alpha > 0$ chosen appropriately large. We choose α as follows: Observe that

$$\begin{aligned} &\tfrac{1}{2} A(\zeta_1) + B(\zeta_1) v_2 + R(\zeta_1, v_2) + \tfrac{1}{6} \alpha v_2^2 \\ &\quad = \left(\frac{T}{\sqrt{2}} v_2 + \frac{B(\zeta_1)}{\sqrt{2} T} \right)^2 + \frac{1}{2T} \big(T A(\zeta_1) - B(\zeta_1)^2 \big) - \frac{T^2}{2} v_2^2 + R(\zeta_1, v_2) + \frac{\alpha}{6} v_2^2. \end{aligned}$$

The first two terms of the right-hand side are positive, in view of (4–10). So we shrink ω_1 appropriately and choose $\alpha > 0$ so large that

$$\tfrac{1}{2} A(\zeta_1) + B(\zeta_1) v_2 + R(\zeta_1, v_2) + \tfrac{1}{6} \alpha v_2^2 \geq 0, \quad \text{for all } \zeta \in \Phi(\overline{\Omega} \cap \omega_1). \tag{4–13}$$

Now consider:

Case (i): $u_2 \geq 0$. Let $\varepsilon_1 > 0$ be so small that $B(p; \varepsilon_1) \subset \omega_1$ and

$$(u_2 - \alpha u_2^2) \geq \tfrac{1}{2} u_2 \quad \text{for } \zeta \in \Phi\big(\overline{\Omega} \cap B(p; \varepsilon_1)\big).$$

Then, for all such ζ, we have

(4–14)

$$\begin{aligned} \operatorname{Re} H(\zeta) &= (u_2 - \alpha u_2^2) + \alpha v_2^2 \\ &\geq \tfrac{1}{2} u_2 + \alpha v_2^2 = \tfrac{1}{4} u_2 + \tfrac{1}{2} \alpha v_2^2 + \tfrac{1}{4} (u_2 + 2\alpha v_2^2) \\ &\geq \tfrac{1}{4} u_2 + \tfrac{1}{2} \alpha v_2^2 + \tfrac{1}{4} \big((A(\zeta_1) + B(\zeta_1) v_2 + R(\zeta_1, v_2)) + 2\alpha v_2^2 \big) \\ &= \tfrac{1}{4} u_2 + \tfrac{1}{8} A(\zeta_1) + \tfrac{1}{2} \alpha v_2^2 + \tfrac{1}{4} \big(\tfrac{1}{2} A(\zeta_1) + B(\zeta_1) v_2 + R(\zeta_1, v_2) + 2\alpha v_2^2 \big) \\ &\gtrsim u_2^2 + v_2^2 + A(\zeta_1), \quad \text{using (4–13).} \end{aligned}$$

Case (ii): $u_2 < 0$. Let $\varepsilon_2 > 0$ be so small that $B(p; \varepsilon_2) \subset \omega_1$ and that $(u_2 - \alpha u_2^2) \geq 2u_2$ for $\zeta \in \Phi(\overline{\Omega} \cap B(p; \varepsilon_2))$. Then, for all such ζ, we have (arguing exactly as before)

(4–15)
$$\begin{aligned}\operatorname{Re} H(\zeta) &\geq -u_2 + \tfrac{1}{2}\alpha v_2^2 + 3\big(u_2 + \tfrac{1}{6}\alpha v_2^2\big)\\ &\geq -u_2 + \tfrac{3}{2}A(\zeta_1) + \tfrac{1}{2}\alpha v_2^2 + 3\big(\tfrac{1}{2}A(\zeta_1) + B(\zeta_1)v_2 + R(\zeta_1, v_2) + \tfrac{1}{6}\alpha v_2^2\big)\\ &\gtrsim u_2^2 + v_2^2 + A(\zeta_1), \quad \text{using (4–13)}.\end{aligned}$$

Now let $\varepsilon_0 = \min(\varepsilon_1, \varepsilon_2)$. From (4–8), (4–14) and (4–15) we see that there is a uniform constant $\kappa > 0$ such that

$$\begin{aligned}\text{(4–16)}\quad \operatorname{Re} H(\zeta) &\geq \kappa(u_2^2 + v_2^2 + v_1^{2M})\\ &\geq \kappa \operatorname{dist}\big[\zeta, \Phi\big(C \cap B(p; \varepsilon_0)\big)\big]^{2M} \quad \text{for } \zeta \in \Phi\big(\overline{\Omega} \cap B(p; \varepsilon_0)\big).\end{aligned}$$

Write $\Phi(C \cap U) = (a, b)$, and without loss of generality, assume that $a < 0 < b$. Define the function ϕ by

$$\phi(u_1) = \begin{cases} \exp\big(1/(u_1 - a)\big) & \text{if } u_1 < a,\\ 0 & \text{if } a \leq u_1 \leq b,\\ \exp\big(-1/(u_1 - b)\big) & \text{if } u_1 > b.\end{cases}$$

Let $r > 0$ such that $B(0; r) \supset \Phi\big(B(p; \varepsilon_0)\big)$, and let $R(\sigma)$ be the rectangle

$$R(\sigma) = \big\{(u_1 + iv_1) \in \mathbb{C} \;\big|\; |u_1| < r,\ |v_1| < \sigma\big\}.$$

By an argument given in [Noell 1985, Lemma 2.1], there exists a smooth almost holomorphic extension $\tilde{\phi}$ of ϕ and a $\sigma > 0$ small enough that

$$\text{(4–17)}\qquad \operatorname{Re}\big(\tilde{\phi}(u_1 + iv_1)\big) \geq -\tfrac{1}{2}\kappa v_1^{2M}, \quad u_1 + iv_1 \in R(\sigma).$$

We set

$$V_1(p, U) = B(p; \varepsilon_0) \cap \Phi^{-1}\big(\operatorname{Image}\Phi \,\cap\, (R(\sigma) \times \mathbb{C})\big).$$

From (4–16) and (4–17), we infer that the function $G(z) = (1 - \tilde{\phi}) \circ \Phi(z) - H \circ \Phi(z)$ satisfies (1)–(3). $\square$

5. The proof of Theorem 1.1

Statement (i). Let C be as in Theorem 1.1(i), and fix $p \in C$. Let $V(p)$ be a neighborhood of p as given by Proposition 4.2. We will use Lemma 3.2 to provide a proof. Take V_p, in the notation of that lemma, to be $V(p)$. In the notation of Lemma 3.2, let $\chi \in \mathscr{C}_c^\infty(V_p; [0, 1])$ be a bump function such that $\operatorname{int}\big(\chi^{-1}\{1\}\big) \cap C$

is an arc. Write $U = \operatorname{int}\bigl(\chi^{-1}\{1\}\bigr)$. Now set $V_1 = V_1(p, U)$ and let $G \in \mathscr{C}^\infty(V_1)$ be as given by Proposition 4.2.

Define

$$G_k(z) = \begin{cases} G(z)^k\,\chi(z) & \text{if } z \in \overline{\Omega} \cap V_1, \\ 0 & \text{if } z \in \overline{\Omega} \setminus V_1. \end{cases}$$

Also define

$$f_k(z) = \bar\partial G_k(z) = k\,G(z)^{k-1}\,\bar\partial G(z)\,\chi(z) + G(z)^k\,\bar\partial\chi(z). \tag{5–1}$$

For a (0, 1) form $\phi(z) = \phi_1(z)d\bar z_1 + \phi_2(z)d\bar z_2$ defined on $\overline{\Omega}$, define

$$\|\phi\|_{\overline{\Omega}} := \max\bigl\{\sup_{\overline{\Omega}}|\phi_1(z)|,\ \sup_{\overline{\Omega}}|\phi_2(z)|\bigr\}.$$

By construction,

$$\|G^k\,\bar\partial\chi\|_{\overline{\Omega}} \to 0 \quad \text{as } k \to \infty. \tag{5–2}$$

Notice that $\bar\partial G$ vanishes to infinite order wherever $G(z) = 1$. Thus, for $j = 1, 2$,

$$\bigl|k\,G(z)^{k-1}\,\partial_{\bar j}G(z)\,\chi(z)\bigr| \lesssim k\bigl(1 - C\,\mathrm{dist}[z, C\cap V_p]^{2M}\bigr)^{k-1}\,\bigl|\partial_{\bar j}G(z)\bigr| \to 0 \quad \text{uniformly as } k \to \infty. \tag{5–3}$$

From (5–2) and (5–3),

$$\|f_k\|_{\overline{\Omega}} \to 0 \quad \text{as } k \to \infty. \tag{5–4}$$

Now consider on Ω the $\bar\partial$-equations

$$\bar\partial u_k = f_k.$$

We need Lipschitz estimates for the solution of the $\bar\partial$-equation on pseudoconvex domains in $\mathbb{C}^2$ of finite type. Such estimates may be found in several places in the literature; for instance, in the results of Chang, Nagel and Stein [Chang et al. 1992], which imply that

$$\|u_k\|_{\overline{\Omega}} \le \|u_k\|_{\Lambda^{1/N}(\overline{\Omega})} \le C^*\|f_k\|_{\overline{\Omega}}, \tag{5–5}$$

where N is a positive integer such that $\tau(p) \le N$ for each $p \in \partial\Omega$, $\Lambda^{1/N}(\overline{\Omega})$ is the class of complex-valued Lipschitz functions on $\overline{\Omega}$ of order $1/N$, and $C^* > 0$ is a constant depending only on Ω. From (5–4) and (5–5) we see that $\|u_k\|_{\overline{\Omega}} \to 0$, whence, defining

$$h_k(z) = G_k(z) - u_k(z) \quad \text{for all } z \in \overline{\Omega},$$

we have a sequence of $A(\Omega)$ functions with

$$\lim_{k\to\infty} h_k(z) = \lim_{k\to\infty} G_k(z) = \begin{cases} \chi(z) & \text{if } z \in C \cap V_p, \\ 0 & \text{if } z \in \overline{\Omega} \setminus (C \cap V_p). \end{cases}$$

Notice that, by construction, the sequence $\{h_k\}_{k\in\mathbb{N}}$ is uniformly bounded. The sequence $\{h_k\}_{k\in\mathbb{N}} \subset A(\Omega)$ satisfies hypotheses (i)–(iii) in Lemma 3.2 for the bump function $\chi \in \mathscr{C}_c^\infty(V_p; [0, 1])$ such that $\operatorname{int}\big(\chi^{-1}\{1\}\big) \cap C$ is an arc. Thus we conclude, using Lemma 3.2, that any compact subset of C is a peak-interpolation set for $A(\Omega)$.

Statement (ii). In the present situation, Ω is a bounded domain having a real-analytic boundary and C is a real-analytic complex-tangential curve. Let B be an open ball in $\mathbb{C}^2$ and let $\gamma : (-2\varepsilon, 2\varepsilon) \to C$ be an injective real-analytic parametrization of C locally such that $\operatorname{Image}(\gamma|_{[-\varepsilon,\varepsilon]}) = (C \cap \bar{B})$. Let $p \in (C \cap \bar{B})$ be such that

$$\tau(p) = \min_{q\in C\cap \bar{B}} \tau(q).$$

Write $\tau(p) = 2M$.

Recall that

$$H_p \otimes \mathbb{C}(\partial\Omega) = H_p^{1,0}(\partial\Omega) \oplus H_p^{0,1}(\partial\Omega),$$

where $H \otimes \mathbb{C}(\partial\Omega)$ is the complexification of $H(\partial\Omega)$, and that $H_p^{1,0}(\partial\Omega)$ and $H_p^{0,1}(\partial\Omega)$ are the eigenspaces of the complex-structure map $\mathbb{J}$ corresponding to $+i$ and $-i$ respectively. Without loss of generality, we may assume that there is an open set $U \supset \bar{B}$ and a real-analytic section L of $H^{1,0}(\partial\Omega)|_U$ such that $L(q)$ spans $H_q^{1,0}(\partial\Omega)$ and

$$L(q) \in \big\{v \in H_q^{1,0}(\partial\Omega) : \|v\| = 1\big\}$$

for each $q \in (\partial\Omega \cap U)$. Now consider the real-analytic function $\mathfrak{L} : S^1 \times I \to \mathbb{R}$ defined by

$$\mathfrak{L}(\zeta, t) = \sum_{\substack{j+k=2M\\ 1\le j<2M}} L^{j-1}\bar{L}^{k-1}\big\langle [L, \bar{L}], \partial\rho\big\rangle(\gamma(t))\, \zeta^j \bar{\zeta}^k,$$

where I is an open interval around $[-\varepsilon, \varepsilon]$, S^1 is the unit circle in $\mathbb{C}$ and ρ is a defining function of $\partial\Omega$. Let t_0 be such that $\gamma(t_0) = p$. By [Bloom 1978b, Theorem 3.3], $\tau(p) = 2M$ implies that there exists a $\zeta_0 \in S^1$ such that $\mathfrak{L}(\zeta_0, t_0) \neq 0$. Then, by the real-analyticity of $\mathfrak{L}$, we conclude that

$$\{t \in [-\varepsilon, \varepsilon] : \mathfrak{L}(\zeta_0, t) = 0\} \text{ is a finite set } \mathfrak{S} \subset [-\varepsilon, \varepsilon].$$

Write $\mathfrak{S} = \{t_1, \dots, t_N\}$. Again by [Bloom 1978b, Theorem 3.3], $\partial\Omega$ is of constant type $2M$ in each connected component of $(C \cap \bar{B}) \setminus \{\gamma_1(t_1), \dots, \gamma(t_N)\}$. Therefore, by Theorem 1.1(i),

$$(C \cap \bar{B}) \setminus \{\gamma(t_1), \dots, \gamma(t_N)\} \text{ is a countable union of peak-interpolation sets.}$$

Recall that Ω is a bounded domain with real-analytic boundary. By [Bedford and Fornæss 1978], therefore, every point of $\partial\Omega$ is a peak point for $A(\Omega)$. So, each $\gamma(t_j)$, for $j = 1, \dots, N$, is a peak point for $A(\Omega)$. This, together with the fact that $(C \cap \bar{B}) \setminus \{\gamma(t_1), \dots, \gamma(t_N)\}$ is a countable union of peak-interpolation sets, implies that C is a countable union of peak-interpolation sets for $A(\Omega)$, and that each compact subset of C is a peak-interpolation set for $A(\Omega)$.

Acknowledgement

The author wishes to thank Alexander Nagel for several useful discussions during the course of this work.

References

[Bedford and Fornæss 1978] E. Bedford and J. E. Fornaess, "A construction of peak functions on weakly pseudoconvex domains", *Ann. of Math.* (2) **107**:3 (1978), 555–568. MR 58 #11520 Zbl 0392.32004

[Bharali 2004] G. Bharali, "On peak-interpolation manifolds for $A(\Omega)$ for convex domains in $\mathbb{C}^n$", *Trans. Amer. Math. Soc.* **356**:12 (2004), 4811–4827. MR 2084399 Zbl 02097164

[Bishop 1962] E. Bishop, "A general Rudin–Carleson theorem", *Proc. Amer. Math. Soc.* **13** (1962), 140–143. MR 24 #A3293 Zbl 0101.08807

[Bloom 1978a] T. Bloom, "$\mathscr{C}^\infty$ peak functions for pseudoconvex domains of strict type", *Duke Math. J.* **45**:1 (1978), 133–147. MR 58 #6353 Zbl 0376.32014

[Bloom 1978b] T. Bloom, "Remarks on type conditions for real hypersurfaces in $\mathbf{C}^n$", pp. 14–24 in *Several complex variables* (Cortona, 1976/1977), Scuola Norm. Sup. Pisa, 1978. MR 84f:32022 Zbl 0421.32022

[Chang et al. 1992] D.-C. Chang, A. Nagel, and E. M. Stein, "Estimates for the $\bar\partial$-Neumann problem in pseudoconvex domains of finite type in $\mathbb{C}^2$", *Acta Math.* **169**:3 (1992), 153–228. MR 93k:32025 Zbl 0821.32011

[Fornæss and McNeal 1994] J. E. Fornæss and J. D. McNeal, "A construction of peak functions on some finite type domains", *Amer. J. Math.* **116**:3 (1994), 737–755. MR 95j:32023 Zbl 0809.32005

[Fornæss and Sibony 1989] J. E. Fornæss and N. Sibony, "Construction of P.S.H. functions on weakly pseudoconvex domains", *Duke Math. J.* **58**:3 (1989), 633–655. MR 90m:32034 ZBL 0679. 32017

[Fornæss and Stensønes 1987] J. E. Fornæss and B. Stensønes, *Lectures on counterexamples in several complex variables*, Mathematical Notes **33**, Princeton University Press, Princeton, NJ, 1987. MR 88f:32001 Zbl 0626.32001

[Henkin and Tumanov 1976] A. E. Tumanov and G. M. Henkin, "Interpolation submanifolds of pseudoconvex manifolds", pp. 74–86 in Математическое программирование и смежные вопросы: Выпуклое программирование (Drogobych, 1974), edited by B. S. Mitjagin and B. P. Titarenko, Central. Èkonom.-Mat. Inst. Akad. Nauk SSSR, Moscow, 1976. In Russian; translated in *Translations Amer. Math. Soc.*, **115** (1980), 59–69. MR 58 #22665 Zbl 0455.32009

[Nagel 1976] A. Nagel, "Smooth zero sets and interpolation sets for some algebras of holomorphic functions on strictly pseudoconvex domains", *Duke Math. J.* **43**:2 (1976), 323–348. MR 56 #670 Zbl 0343.32016

[Nagel and Rudin 1978] A. Nagel and W. Rudin, "Local boundary behavior of bounded holomorphic functions", *Canad. J. Math.* **30**:3 (1978), 583–592. MR 58 #6315 Zbl 0427.32006

[Noell 1985] A. V. Noell, "Interpolation in weakly pseudoconvex domains in $\mathbb{C}^2$", *Math. Ann.* **270**:3 (1985), 339–348. MR 86j:32043 Zbl 0534.32003

[Rudin 1978] W. Rudin, "Peak-interpolation sets of class C^1", *Pacific J. Math.* **75**:1 (1978), 267–279. MR 58 #6346 Zbl 0383.32007

Received December 2, 2002. Revised April 20, 2004.

GAUTAM BHARALI
MATHEMATICS DEPARTMENT
THE UNIVERSITY OF MICHIGAN
525 EAST UNIVERSITY AVENUE
ANN ARBOR, MI 48109

bharali@umich.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

BOUNDING THE BENDING OF A HYPERBOLIC 3-MANIFOLD

MARTIN BRIDGEMAN AND RICHARD D. CANARY

We obtain bounds on the total bending of the boundary of the convex core of a hyperbolic 3-manifold. These bounds will depend on the geometry of the boundary of the convex hull of the limit set.

1. Introduction

The boundary of the convex core of a hyperbolic 3-manifold is a hyperbolic surface in its intrinsic metric. This surface is totally geodesic except along a lamination, called the bending lamination. The bending lamination inherits a transverse measure that keeps track of how much the surface is bent along the lamination. The length (or mass) of the bending lamination, regarded as a measured lamination, records the total bending of the boundary of the convex core. For example, if the boundary of the convex core is bent by an angle of θ along a single simple closed geodesic of length L, then the length of the bending lamination is $L\theta$.

Our main result is an upper bound on the mass of the bending lamination, which depends on a lower bound for the injectivity radius of the boundary of the convex hull of the limit set. An upper bound on the mass of the bending lamination is also implicit in the techniques developed by Bonahon and Otal [2001, Lemma 12].

If $N = \mathbb{H}^3/\Gamma$ is an orientable hyperbolic 3-manifold and Γ is a nonabelian group of orientation-preserving isometries of $\mathbb{H}^3$, the limit set L_Γ of Γ is the smallest closed nonempty Γ-invariant subset of $\partial_\infty\mathbb{H}^3 = \hat{\mathbb{C}}$. The convex core $C(N)$ of N is $CH(L_\Gamma)/\Gamma$, where $CH(L_\Gamma)$ is the convex hull of L_Γ in $\mathbb{H}^3$. Notice that ρ_0 is a lower bound for the injectivity radius of the boundary $\partial CH(L_\Gamma)$ of the convex hull of the limit set if and only if $2\rho_0$ is a lower bound for the length of a compressible curve on the boundary of the convex core (i.e., a closed curve in $\partial C(N)$ that is null-homotopic in $C(N)$ but not in $\partial C(N)$).

Theorem 1. *There exist constants S and T such that if N is an orientable hyperbolic 3-manifold with finitely generated, nonabelian fundamental group and*

Research supported in part by grants from the National Science Foundation.
MSC2000: 57M50.
Keywords: hyperbolic 3-manifold, convex core, bending lamination.

bending lamination β_N and $\rho_0 \in (0, 1]$ is a lower bound for the injectivity radius of the boundary $\partial CH(L_\Gamma)$ of the convex hull of the limit set, then

$$l_{\partial C(N)}(\beta_N) \leq \left|\chi(\partial C(N))\right| \left(S \log \frac{1}{\rho_0} + T\right),$$

where $l_{\partial C(N)}(\beta_N)$ is the length of β_N and $\chi(\partial C(N))$ is the Euler characteristic of the boundary of the convex core.

We also obtain a lower bound for the mass of the bending lamination in the case that $\partial C(N)$ has a short compressible curve. This lower bound makes clear that the dependence on the geometry of the convex hull of the limit set in our first result cannot be removed and that the form of the estimate cannot be substantially improved. Also, notice that if one passes to a degree-d cover of N, both the length of the bending lamination and the Euler characteristic of the boundary of the convex core get multiplied by d, while the convex hull of the limit set is the same, so any upper bound must depend linearly on $\left|\chi(\partial C(N))\right|$.

Theorem 2. *Let $N = \mathbb{H}^3/\Gamma$ be an orientable hyperbolic 3-manifold with finitely generated, nonabelian fundamental group. If $\partial CH(L_\Gamma)$ contains a closed geodesic of length $\rho \leq 2\sinh^{-1} 1$, then*

$$l_{\partial C(N)}(\beta_N) \geq 4\pi \log \frac{4\sinh^{-1} 1}{\rho}.$$

If the boundary of the convex core is incompressible, Proposition 4.2 gives the following stronger result:

Theorem 3. *If N is an orientable hyperbolic 3-manifold with finitely generated, nonabelian fundamental group and $\partial C(N)$ is incompressible in N, then*

$$l_{\partial C(N)}(\beta_N) \leq \frac{\pi^3}{\sinh^{-1} 1} \left|\chi(\partial C(N))\right|.$$

In related work, Epstein, Marden and Markovic (see, for example, [Epstein et al. 2004, Theorem 4.2] have studied the possible bending laminations of embedded convex hyperbolic planes in $\mathbb{H}^3$.

Thurston [1979] (see also [Kourouniotis 1985; Johnson and Millson 1987; Epstein and Marden 1987]) studied the operation of obtaining a quasifuchsian group by bending a Fuchsian group along a simple closed geodesic, or more generally along a measured lamination. Theorem 3 may be used to quantify the observation that if this geodesic is "long," one may only bend by a "small" angle.

This paper is based on earlier work [Bridgeman 1998; Bridgeman and Canary 2003; Canary 2001], which explored the relationship between the boundary of the convex core and the conformal boundary. In particular, we make central use of a result (repeated here as Lemma 3.1) that ensures the existence of a lower bound, depending only on the injectivity radius of its basepoint, for the length of a geodesic arc in $\partial CH(L_\Gamma)$ whose intersection with the bending lamination is at least 2π. We will combine this estimate with a Crofton-like formula (Lemma 4.1) for the length of the bending lamination to prove Theorem 1.

In Section 7, we will apply the results of [Bridgeman and Canary 2003] and [Canary 2001] to obtain analogues of Theorems 1 and 2, which depend on the geometry of the domain of discontinuity $\Omega(\Gamma)$ for the action of Γ on $\hat{\mathbb{C}}$.

2. Background

Let $N = \mathbb{H}^3/\Gamma$ be an orientable hyperbolic 3-manifold with nonabelian fundamental group. Then Γ acts properly discontinuously on the domain of discontinuity $\Omega(\Gamma) = \hat{\mathbb{C}} - L_\Gamma$. The domain of discontinuity admits a canonical conformally invariant hyperbolic metric $p(z)|dz|$, called the *Poincaré metric*. The quotient surface $\partial_c N = \Omega(\Gamma)/\Gamma$, called the conformal boundary of N, is then naturally a hyperbolic surface. The hyperbolic 3-manifold N is said to be *analytically finite* if $\partial_c N$ has finite area in this metric. Ahlfors' Finiteness Theorem [Ahlfors 1964] asserts that N is analytically finite if Γ is finitely generated. All of our results hold for analytically finite hyperbolic 3-manifolds.

If N is analytically finite then there is always a positive lower bound for the injectivity radius on $\Omega(\Gamma)$. By Lemma 8.1 of [Bridgeman and Canary 2003], a lower bound on the injectivity radius of $\Omega(\Gamma)$ implies a lower bound on the injectivity radius of $\partial CH(L_\Gamma)$. In particular, if N is analytically finite then there is a positive lower bound on the injectivity radius of $\partial CH(L_\Gamma)$. The boundary of the convex hull of the limit set is a hyperbolic surface in its intrinsic metric and is totally geodesic in the complement of a closed union β_Γ of disjoint geodesics, called the bending lamination of $CH(L_\Gamma)$. The bending lamination β_N of the convex core $C(N)$ is simply the projection of β_Γ to $\partial C(N)$.

A *measured lamination* on a hyperbolic surface S consists of a closed subset λ of S that is the disjoint union of simple geodesics, together with countably additive invariant (with respect to projection along λ) measures on arcs transverse to λ. The bending laminations β_Γ and β_N come equipped with bending measures on arcs transverse to the lamination, which record the total bending along the arc. These bending measures give β_Γ and β_N the structure of measured laminations. Real multiples of simple closed geodesics are dense in the space $ML(S)$ of all measured laminations on a finite-area hyperbolic surface S. Moreover, the length of a simple closed geodesic and the intersection number of two simple closed geodesics extend naturally to continuous functions on $ML(S)$ and $ML(S) \times ML(S)$ respectively. See [Thurston 1979] or [Bonahon 2001] for fuller discussions of measured lamination spaces and [Thurston 1979] or [Epstein and Marden 1987] for a fuller discussion of convex cores and bending laminations.

3. Local intersection number estimates

In [Bridgeman and Canary 2003] we obtained bounds on the intersection of a transverse geodesic arc with the bending lamination. There we defined a function

$$F(x) = \frac{x}{2} + \sinh^{-1}\left(\frac{\sinh(x/2)}{\sqrt{1-\sinh^2(x/2)}}\right)$$

and its inverse $G(x) = F^{-1}(x)$. The function F is monotonically increasing and has domain $(0, 2\sinh^{-1} 1)$. The function $G(x)$ has domain $(0, \infty)$, has asymptotic behavior $G(x) \asymp x$ as x tends to 0, and $G(x)$ approaches $2\sinh^{-1} 1$ as x tends to ∞. We define $G_\infty = 2\sinh^{-1} 1 \approx 1.76275$.

Lemma 3.1 [Bridgeman and Canary 2003, Lemma 4.3]. *Let $N = \mathbb{H}^3/\Gamma$ be an analytically finite hyperbolic 3-manifold such that L_Γ is not contained in a round circle. Let $\alpha : [0, 1) \to \partial CH(L_\Gamma)$ be a geodesic path (in the intrinsic metric on $\partial CH(L_\Gamma)$) with length $l(\alpha)$. If either*

(1) $l(\alpha) \le G(\mathrm{inj}_{\partial CH(L_\Gamma)}(\alpha(0)))$, *or*

(2) $\alpha([0, 1))$ *is contained in a simply connected component of* $\partial CH(L_\Gamma)$ *and* $l(\alpha) \le G_\infty$,

then

$$i(\alpha, \beta_\Gamma) \le 2\pi.$$

A geodesic arc α is either transverse to β_Γ or contained within β_Γ, in which case we define $i(\alpha, \beta_\Gamma) = 0$.

If $\alpha : [0, 1) \to \partial C(N)$ is a geodesic in the boundary of the convex core, consider its lift $\tilde{\alpha} : [0, 1) \to \partial CH(L_\Gamma)$. If we subdivide this lift into pieces to which Lemma 3.1 applies, as in the proof of [Bridgeman and Canary 2003, Proposition 5.1], we obtain:

Corollary 3.2. *Let N be an analytically finite hyperbolic 3-manifold. Let $\alpha : [0, 1) \to \partial C(N)$ be a geodesic path with length $l(\alpha)$. If α is contained in an incompressible component of $\partial C(N)$, let $G = G_\infty$. Otherwise, let ρ_α be a lower bound on the injectivity radius of $\partial CH(L_\Gamma)$ at every point in $\tilde{\alpha}([0, 1))$ and let $G = G(\rho_\alpha)$. Then*

$$i(\alpha, \beta_N) \le 2\pi \left\lceil \frac{l(\alpha)}{G} \right\rceil.$$

Here $\lceil x \rceil$, as usual, denotes the least integer greater than or equal to x.

We have so far avoided, for simplicity of exposition, discussing the case that the limit set is contained in a round circle. In this case, the convex core is a totally geodesic surface with geodesic boundary. It is natural to consider the boundary of the convex core to be the double of the convex core (where one considers the two sheets of the convex core to have opposite normal vectors.) With this convention, the boundary of the convex core is still a finite-area hyperbolic surface with boundary if our manifold is analytically finite. One can easily see, just as in the proof of [Bridgeman and Canary 2003, Proposition 5.1], that Corollary 3.2 remains valid in this situation.

4. A length formula

In order to prove Theorem 1 we first represent the length of the bending lamination as the integral of the intersection number over all geodesics of a fixed length. Our formula is similar to the Crofton formula for the area of a region in the plane. See also [Bonahon 1988, Proposition 14].

Let S be a hyperbolic surface. If $v \in T^1(S)$ is a unit tangent vector, let $\bar{\alpha}(v) : (0, \infty) \to S$ be the unit-speed geodesic ray originating at the basepoint of v and in the direction of v. Let $\alpha^L(v) = \bar{\alpha}|_{(0,L)}$ be the open geodesic segment of length L emanating from the basepoint of v in the direction v.

Lemma 4.1. *Let β be a measured lamination on a hyperbolic surface S of finite area. Then*

$$l_S(\beta) = \frac{1}{4L} \int_{T^1(S)} i(\alpha^L(v), \beta)\, d\Omega(v),$$

where $d\Omega$ is the volume form on $T_1(S)$.

Proof. We define a function F_L on the space $ML(S)$ of measured laminations by setting

$$F_L(\beta) = \frac{1}{4L} \int_{T_1(S)} i(\alpha^L(v), \beta)\, d\Omega(v).$$

As F_L and l_S are both continuous on $ML(S)$ and real multiples of closed geodesics are dense in $ML(S)$, it suffices to prove that $F_L(\beta) = l_S(\beta)$ for real multiples of closed geodesics. Since $F_L(k\beta) = kF_L(\beta)$ and $l_S(k\beta) = kl_S(\beta)$ for all $\beta \in ML(S)$ and all $k > 0$, we may assume that β is a single closed geodesic with unit transverse measure.

Let C be the hyperbolic cylinder covering S corresponding to β and let $\tilde{\beta}$ be the lift of β to C. If $v \in T^1(S)$, then $i(\alpha^L(v), \beta)$ is precisely the number of lifts of $\alpha^L(v)$ to C that intersect $\tilde{\beta}$. Let

$$U = \left\{ v \in T^1(C) \mid \alpha^L(v) \text{ intersects } \tilde{\beta} \right\}.$$

Lifting the integral to C we see that

$$\int_{T^1(S)} i(\alpha^L(v), \beta)\, d\Omega(v) = \int_U d\Omega(v).$$

The metric on C is given by

$$ds^2 = dx^2 + \cosh^2 x\, dl^2,$$

where x is the perpendicular distance to the core geodesic and l is a length coordinate along the core geodesic; see [Buser 1992, Example 1.3.2]. The hyperbolic area element is $dA = \cosh x\, dx\, dl$.

Let

$$N = \{c \in C \mid 0 < d(\widetilde{\beta}, c) < L\}.$$

If $v \in U$, the basepoint p of v is in N. If $p \in N$, let U_p denote the cone of tangent vectors in $U \cap T^1_p(C)$. Let w_p denote the unit vector tangent to the geodesic ray through p perpendicular to $\widetilde{\beta}$. Then U_p consists of all vectors in $T^1_p(C)$ making an angle of at most $\theta(p)$ with w_p, where

$$\theta(p) = \cos^{-1} \frac{\tanh x}{\tanh L}.$$

Therefore,

$$\int_U d\Omega(v) = \int_N 2\cos^{-1} \frac{\tanh x}{\tanh L}\, dA.$$

Integrating over the core of the annulus we obtain

$$\begin{aligned}\int_N 2\cos^{-1} \frac{\tanh x}{\tanh L}\, dA &= 2l_S(\beta) \int_{-L}^{L} \cosh x \cos^{-1} \frac{\tanh x}{\tanh L}\, dx \\ &= 4l_S(\beta) \int_0^L \cosh x \cos^{-1} \frac{\tanh x}{\tanh L}\, dx.\end{aligned}$$

Therefore,

$$F_L(\beta) = \frac{l_S(\beta)}{L} \int_0^L \cosh x \cos^{-1} \frac{\tanh x}{\tanh L}\, dx.$$

Substituting $u = \dfrac{\tanh x}{\tanh L}$ we obtain

$$F_L(\beta) = \frac{l_S(\beta)\tanh L}{L} \int_0^1 \frac{\cos^{-1} u}{(1 - u^2 \tanh^2 L)^{3/2}}\, du.$$

We may then integrate by parts and evaluate the result to check that $F_L(\beta)$ has the claimed form. □

We now prove a version of Theorem 1 that is a direct application of Corollary 3.2 and Lemma 4.1. Recall that $G(x) \asymp x$ as x tends to 0. If $\rho_0 \geq 1$, this estimate is better than the one provided by Theorem 1, but it is much weaker as ρ_0 approaches 0, since the upper bound provided by Proposition 4.2 is $O\big(|\chi(\partial C(N))|/\rho_0\big)$, while the estimate provided by Theorem 1 is $O\big(|\chi(\partial C(N))| \log(\rho_0^{-1})\big)$. Notice that Theorem 3 is case (2) of Proposition 4.2.

Proposition 4.2. *Let $N = \mathbb{H}^3/\Gamma$ be an analytically finite hyperbolic 3-manifold with bending lamination β_N.*

(1) *If $\rho_0 > 0$ is a lower bound for the injectivity radius of $\partial CH(L_\Gamma)$, then*

$$l_{\partial C(N)}(\beta_N) \leq \frac{2\pi^3}{G(\rho_0)} \left|\chi(\partial C(N))\right|.$$

(2) *If $\partial C(N)$ is incompressible in N, then*

$$l_{\partial C(N)}(\beta_N) \leq \frac{\pi^3}{\sinh^{-1} 1} \left|\chi(\partial C(N))\right|.$$

Proof. If $\partial C(N)$ is incompressible, we let $G = G_\infty = 2\sinh^{-1} 1$. If not, we let $G = G(\rho_0)$. Corollary 3.2 implies that, for all $v \in T^1(\partial C(N))$,

$$i(\alpha^L(v), \beta_N) \leq 2\pi \left\lceil \frac{L}{G} \right\rceil \leq 2\pi \Big(\frac{L}{G} + 1\Big).$$

Therefore, by Lemma 4.1,

$$l_{\partial C(N)}(\beta_N) \leq \frac{\pi}{2L} \int_{T^1(\partial C(N))} \Big(\frac{L}{G} + 1\Big)\, d\Omega \leq \operatorname{vol} T^1(\partial C(N)) \left(\frac{\pi}{2G} + \frac{\pi}{2L}\right).$$

The volume of the unit tangent bundle $T^1(\partial C(N))$ is $4\pi^2 \left|\chi(\partial C(N))\right|$. Thus, by letting L tend to infinity, we see that

$$l_{\partial C(N)}(\beta_N) \leq 4\pi^2 \left|\chi(\partial C(N))\right| \left(\frac{\pi}{2G}\right) = \frac{2\pi^3}{G} \left|\chi(\partial C(N))\right|.$$

□

5. Proof of Theorem 1

To obtain the sharper bound on the length of the bending lamination given by Theorem 1, we must decompose $\partial C(N)$ using the Collar Lemma. We will use the following explicit version of the Collar Lemma, which combines [Buser 1992, Theorem 4.4.6] and [Yamada 1982, Lemma 7] (which guarantees that curves of length at most $2\sinh^{-1} 1$ are simple).

Collar Lemma. *Let S be a finite-area hyperbolic surface of genus g with n punctures. Let $\{\nu_1, \dots, \nu_k\}$ be the collection of all primitive closed geodesics on S of length at most* $2\sinh^{-1} 1$. *Then*:

(1) $k \le 3g - 3 + n$.

(2) $\{\nu_1, \dots, \nu_k\}$ *is a disjoint collection of simple closed geodesics.*

(3) *There exists a disjoint collection $\{B_1, \dots, B_k\}$ of metric collar neighborhoods of $\{\nu_1, \dots, \nu_k\}$ such that each B_i is isometric to the quotient of*

$$[-w(\nu_i), w(\nu_i)] \times [0, l_S(\nu_i)]$$

by the identification map $(t, 0) \mapsto (t, l_S(\nu_i))$, where $l_S(\nu_i)$ is the length of ν_i,

$$w(\nu_i) = \sinh^{-1} \frac{1}{\sinh\left(\frac{1}{2} l_S(\nu_i)\right)},$$

and the product has the metric

$$ds^2 = dx^2 + \cosh^2 x \, dl^2.$$

(4) *If $x \in B_i$, then* $\sinh \operatorname{inj}_S(x) = \sinh\left(\frac{1}{2} l_S(\nu_i)\right) \cosh d(x, \nu_i)$.

(5) *If there is a curve through $x \in S$ homotopic to ν_i of length at most* $2\sinh^{-1} 1$, *then $x \in B_i$.*

We now restate Theorem 1 for analytically finite hyperbolic 3-manifolds.

Theorem 1. *There exist constants S and T such that if $N = \mathbb{H}^3/\Gamma$ is an analytically finite hyperbolic* 3*-manifold with bending lamination β_N and $\rho_0 \in (0, 1]$ is a lower bound for the injectivity radius of the boundary $\partial CH(L_\Gamma)$ of the convex hull of the limit set, then*

$$l_{\partial C(N)}(\beta_N) \le \left|\chi(\partial C(N))\right| \left(S \log \frac{1}{\rho_0} + T\right),$$

where $l_{\partial C(N)}(\beta_N)$ is the length of β_N and $\chi(\partial C(N))$ is the Euler characteristic of the boundary of the convex core.

Proof. As the proof is rather technical, we begin with a brief outline. We first decompose $\partial C(N)$ into the set X of collars of short compressible geodesics and its complement Y. We choose $\epsilon = \sinh^{-1} 1$ and $L = G(\epsilon)$. By Lemma 4.1

$$\begin{aligned} l_{\partial C(N)}(\beta_N) &= \frac{1}{4L} \int_{T^1(S)} i(\alpha^L(v), \beta_N) \, d\Omega(v) \\ &= \frac{1}{4L} \left(\int_{T^1(X)} i(\alpha^L(v), \beta_N) \, d\Omega(v) + \int_{T^1(Y)} i(\alpha^L(v), \beta_N) \, d\Omega(v) \right). \end{aligned}$$

Lemma 3.1 implies that $i(\alpha^L(v), \beta_N) \leq 2\pi$ for $v \in T^1(Y)$, so, just as in the proof of Proposition 4.2,

$$\int_{T^1(Y)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega \leq 2\pi \operatorname{vol} T^1(Y).$$

To handle the integral over $T^1(X)$, we use Corollary 3.2, which implies that

$$i(\alpha^L(v), \beta_N) \leq 2\pi \left\lceil \frac{L}{G(r(v))} \right\rceil,$$

where $r(v)$ is a lower bound on the injectivity radius of $\partial C(N)$ at any point on $\alpha^L(v)$. If B is a component of X with core geodesic ν and $v \in T^1(B)$, we observe that

$$r(v) \geq \sinh^{-1}\left(\frac{1}{e^{G(\epsilon)}} \sinh \frac{l_S(\nu)}{2} \cosh d(v)\right)$$

where $d(v)$ is the distance from the basepoint of v to ν. Combining the resulting bounds and integrating, we obtain an upper bound on the integral of $i(\alpha^L(v), \beta_N)$ over $T^1(B)$ in terms of the length of ν. Summing the resulting bounds over $T^1(Y)$ and all components of $T^1(X)$ gives our result.

Let $\{\nu_1, \ldots, \nu_k\}$ be the primitive closed geodesics of length at most $2\sinh^{-1} 1$ on $\partial C(N)$. Let $\{B_1, \ldots, B_k\}$ be the collar neighborhoods of $\{\nu_1, \ldots, \nu_k\}$ provided by the Collar Lemma.

Let $\pi : \partial CH(L_\Gamma) \to \partial C(N)$ be the covering map from the boundary of the convex hull to the boundary of the convex core. Set $\epsilon = \sinh^{-1} 1$ and

$$\widetilde{V} = \left\{x \in \partial CH(L_\Gamma) \;\middle|\; \operatorname{inj}_{\partial CH(L_\Gamma)}(x) \leq \epsilon\right\}.$$

If $x \in \widetilde{V}$, then x lies on a homotopically nontrivial curve n_x of length at most 2ϵ. Since there is a lower bound on the injectivity radius of $\partial CH(L_\Gamma)$, n_x is homotopic to a closed geodesic $\tilde{\nu}_x$ of length at most 2ϵ. Then $\tilde{\nu}_x$ projects to (a multiple of) one of the curves $\{\nu_1, \ldots, \nu_k\}$, so $\pi(x)$ lies in some collar neighborhood B_i and x lies in a lift of B_i to $\partial CH(L_\Gamma)$. Let X denote the union of all collar neighborhoods B_i containing some component of $\pi(\widetilde{V})$. Let $Y = \partial C(N) - X$. We may renumber $\{B_1, \ldots, B_k\}$ so that $X = \bigcup_{i=1}^m B_i$ for some $m \leq k$. Notice that $\operatorname{inj}_{\partial CH(L_\Gamma)}(y) > \epsilon$ for $y \in \pi^{-1}(Y)$.

We choose $L = G(\epsilon)$ in the formula for $l_{\partial C(N)}(\beta_N)$ in Lemma 4.1. We split the integral into two integrals using the decomposition, so that

$$l_{\partial C(N)}(\beta_N) = \frac{1}{4G(\epsilon)}\left(\int_{T^1(X)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega + \int_{T^1(Y)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega\right).$$

We first estimate the portion of the integral with domain $T^1(Y)$. If v has basepoint in Y and $\tilde{\alpha}^{G(\epsilon)}(v)$ is a lift of $\alpha^{G(\epsilon)}(v)$ to $\partial CH(L_\Gamma)$, then $\tilde{\alpha}^{G(\epsilon)}(v)$ originates at a point $\tilde{y}$ such that $\mathrm{inj}_{\partial CH(L_\Gamma)}(\tilde{y}) > \epsilon$ and has length $G(\epsilon) < G(\mathrm{inj}_{\partial CH(L_\Gamma)}(\tilde{y}))$. Therefore, Lemma 3.1 implies that $i(\widetilde{\alpha}^{G(\epsilon)}(v), \beta_\Gamma) \le 2\pi$ and hence that

$$i(\alpha^{G(\epsilon)}(v), \beta_N) \le 2\pi.$$

Therefore

$$\int_{T^1(Y)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega \le \int_{T^1(Y)} 2\pi\, d\Omega \le 2\pi \operatorname{vol} T^1(Y). \tag{1}$$

We now estimate the portion of the integral with domain $T^1(X)$. If X is empty, we are done. Otherwise, let B_i be a component of X. Let $v \in T_1(B_i)$ and $d_i(v)$ be the distance from ν_i to the basepoint b_v of v.

We now derive a lower bound for the injectivity radius along the geodesic $\alpha^{G(\epsilon)}(v)$ as a function of $d_i(v)$. One can readily check that if S is a hyperbolic surface, $w, z \in S$ and $\delta = d_S(z, w)$, then $\sinh \mathrm{inj}_S(w) \ge e^{-\delta} \sinh \mathrm{inj}_S(z)$. (This follows, for example, from [Beardon 1983, Theorem 7.35.1].) Since, by the Collar Lemma,

$$\sinh \mathrm{inj}_S(b_v) = \sinh \frac{l_S(\nu_i)}{2} \cosh d_i(v),$$

we see that if x is any point on $\alpha^{G(\epsilon)}(v)$, then

$$\sinh \mathrm{inj}_{\partial C(N)}(x) \ge \frac{1}{e^{G(\epsilon)}} \sinh \frac{l_S(\nu_i)}{2} \cosh d_i(v).$$

We define $R_i : [0, w(\nu_i)] \to \mathbb{R}$ by

$$R_i(t) = \sinh^{-1}\left(\frac{1}{e^{G(\epsilon)}} \sinh \frac{l_S(\nu_i)}{2} \cosh t\right).$$

The injectivity radius at any point of $\alpha^{G(\epsilon)}(v)$ is bounded from below by $R_i(d_i(v))$, so if $\tilde{\alpha}^{G(\epsilon)}(v)$ is a lift of $\alpha^{G(\epsilon)}(v)$ to $\partial CH(L_\Gamma)$, the injectivity radius of $\partial CH(L_\Gamma)$ at every point of $\tilde{\alpha}^{G(\epsilon)}(v)$ is also bounded from below by $R_i(d_i(v))$. Thus, by Corollary 3.2,

$$i(\alpha^{G(\epsilon)}(v), \beta_N) \le 2\pi \left\lceil \frac{G(\epsilon)}{G(R_i(d_i(v)))} \right\rceil.$$

So

$$\begin{aligned} \int_{T^1(B_i)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega &\le \int_{T^1(B_i)} 2\pi \left\lceil \frac{G(\epsilon)}{G(R_i(d_i(v)))} \right\rceil d\Omega \\ &\le 2\pi G(\epsilon) \int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega + 2\pi \operatorname{vol} T^1(B_i). \end{aligned} \tag{2}$$

Since the integral depends only on $d_i(v)$,

$$\int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega \le 2\pi \int_0^{l_S(\nu_i)} \int_{-\omega(\nu_i)}^{\omega(\nu_i)} \frac{1}{G(R_i(|x|))} \cosh x\, dx\, dl,$$

where x and l are the coordinates on B_i provided by the Collar Lemma.

As $R_i(|x|) < \epsilon$ on B_i, we need only consider G on the domain $[0, \epsilon]$. Since $t/G(t)$ tends to 1 as t tends to 0 and is continuous on $(0, \epsilon]$, there exists a constant $K_1 > 0$ such that $t/G(t) \le K_1$ for all $t \in (0, \epsilon]$. Therefore

$$\int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega \le 2\pi \int_0^{l_S(\nu_i)} \int_{-\omega(\nu_i)}^{\omega(\nu_i)} \frac{K_1 \cosh x}{R_i(|x|)}\, dx\, dl.$$

Integrating over the core curve and making use of the symmetry about the core geodesic, we see that

$$\int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega \le 4\pi K_1 l_S(\nu_i) \int_0^{w(\nu_i)} \frac{\cosh x}{R_i(x)} dx. \tag{3}$$

Since $\sinh x/x$ is increasing on $(0, \infty)$, $\sinh x/x \le K_2 = \sinh \epsilon/\epsilon$ for all $x \in (0, \epsilon]$. Thus, for all $x \in (0, w(\nu_i))$,

$$\frac{1}{R_i(x)} \le \frac{K_2}{\sinh R_i(x)}.$$

Therefore,

$$\int_0^{w(\nu_i)} \frac{\cosh x}{R_i(x)}\, dx \le \int_0^{w(\nu_i)} \frac{K_2 e^{G(\epsilon)}}{\sinh(l(\nu_i)/2)} dx \le \frac{w(\nu_i) K_2 e^{G(\epsilon)}}{\sinh(l(\nu_i)/2)}. \tag{4}$$

Combining inequalities (3) and (4) we see that

$$\int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega \le \frac{4\pi K_1 l(\nu_i) w(\nu_i) K_2 e^{G(\epsilon)}}{\sinh(l(\nu_i)/2)}.$$

Since $\sinh x \ge x$, we get

$$\int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega \le 8\pi K_1 K_2 e^{G(\epsilon)} w(\nu_i).$$

Applying the equality $\sinh^{-1} x = \log(x + \sqrt{x^2+1})$, we see that

$$w(\nu_i) = \sinh^{-1} \frac{1}{\sinh(l(\nu_i)/2)} = \log \frac{1 + \cosh(l(\nu_i)/2)}{\sinh(l(\nu_i)/2)}.$$

Thus

$$\begin{aligned} w(\nu_i) &\le \log\left(1 + \cosh \frac{l(\nu_i)}{2}\right) + \log \frac{1}{\sinh(l(\nu_i)/2)} \\ &\le \log(1 + \cosh \epsilon) + \log \frac{2}{l(\nu_i)}. \end{aligned}$$

This yields

$$\int_{T^1(B_i)} \frac{1}{G(R_i(d_i(v)))}\, d\Omega \le S_0 \log \frac{2}{l(\nu_i)} + T_0, \tag{5}$$

where $S_0 = 8\pi K_1 K_2 e^{G(\epsilon)}$ and $T_0 = S_0 \log(1+\cosh\epsilon)$.

Since $X = \bigcup_{i=1}^m B_i$, we can combine inequalities (2) and (5) to obtain

$$\begin{aligned}\int_{T^1(X)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega &= \sum_{i=1}^m \int_{T^1(B_i)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega \\ &\le \sum_{i=1}^m 2\pi G(\epsilon) \Big(S_0 \log \frac{2}{l(\nu_i)} + T_0 \Big) + 2\pi \operatorname{vol} T^1(B_i).\end{aligned}$$

Since m is bounded above by the number of disjoint geodesics in $\partial C(N)$,

$$m \le \tfrac{3}{2} \big| \chi(\partial C(N)) \big|.$$

Moreover, as ρ_0 is a lower bound for the injectivity radius of $\partial CH(L_\Gamma)$, we have $\rho_0 \le l(\nu_i)/2$ for all i. Therefore,

$$\begin{aligned}&\int_{T^1(X)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega \\ &\qquad \le 3\pi G(\epsilon) \big|\chi(\partial C(N))\big| \Big(S_0 \log \frac{1}{\rho_0} + T_0 \Big) + 2\pi \operatorname{vol} T^1(X).\end{aligned} \tag{6}$$

Combining estimates (1) and (6) for the integral over $T^1(X)$ and $T^1(Y)$, we get

$$\begin{aligned}&\int_{T^1(\partial C(N))} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega \\ &\quad = \int_{T^1(X)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega + \int_{T^1(Y)} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega \\ &\quad \le 3\pi G(\epsilon) \big|\chi(\partial C(N))\big| \Big(S_0 \log \frac{1}{\rho_0} + T_0 \Big) + 2\pi \operatorname{vol} T^1(X) + 2\pi \operatorname{vol} T^1(Y) \\ &\quad \le 3\pi G(\epsilon) \big|\chi(\partial C(N))\big| \Big(S_0 \log \frac{1}{\rho_0} + T_0 \Big) + 2\pi \operatorname{vol} T^1(\partial C(N)).\end{aligned}$$

Recalling that

$$l_{\partial C(N)}(\beta_N) = \frac{1}{4G(\epsilon)} \int_{T^1(\partial C(N))} i(\alpha^{G(\epsilon)}(v), \beta_N)\, d\Omega$$

and that $\operatorname{vol} T^1(\partial C(N)) = 4\pi^2 \big|\chi(\partial C(N))\big|$, we see that this implies that

$$l_{\partial C(N)}(\beta_N) \le \big|\chi(\partial C(N))\big| \Big(S \log \frac{1}{\rho_0} + T \Big),$$

where

$$S = \frac{3\pi S_0}{4} \qquad \text{and} \qquad T = \frac{3\pi T_0}{4} + \frac{2\pi^3}{G(\epsilon)}. \qquad \square$$

Remark. One can evaluate the constants used in the proof to check that $\epsilon = \sinh^{-1} 1 \approx 0.8814$, $G(\epsilon) = F^{-1}(\epsilon) \approx 0.8387$, $K_1 = \epsilon/G(\epsilon) \approx 1.0509$ (since $t/G(t)$ is increasing), and $K_2 = \sinh\epsilon/\epsilon \approx 1.1346$. Therefore, $S \le 164$ and $T \le 218$.

6. A lower bound on the length of the bending lamination

If the boundary of the convex core contains a short compressible curve we obtain a lower bound on the length of the bending lamination, having the same asymptotic form as the upper bound obtained in Theorem 1. Notice that if N is Fuchsian, the bending lamination has length zero, so no general lower bound is possible.

Theorem 2. *Let $N = \mathbb{H}^3/\Gamma$ be an analytically finite hyperbolic 3-manifold. If $\partial CH(L_\Gamma)$ contains a closed geodesic of length $\rho \le 2\sinh^{-1} 1$, then*

$$l_{\partial C(N)}(\beta_N) \ge 4\pi \log \frac{4\sinh^{-1} 1}{\rho}.$$

Proof. Let $\tilde\alpha$ be the closed geodesic of length ρ on $\partial CH(L_\Gamma)$ and let $\epsilon = \sinh^{-1} 1$. Let α be the projection of $\tilde\alpha$ to $\partial C(N)$. It follows from the Collar Lemma that α is a multiple of a simple closed geodesic ν. Let B be the collar of ν provided by the Collar Lemma. The collar B has width $w \ge \sinh^{-1}(1/\sinh(\rho/2))$. Since $\sinh^{-1} x = \log(x + \sqrt{x^2+1})$,

$$w \ge \log \frac{1+\cosh(\rho/2)}{\sinh(\rho/2)} \ge \log \frac{2}{\sinh(\rho/2)}.$$

Since $\sinh x/x$ is an increasing function on $(0, \infty)$,

$$\sinh\frac{\rho}{2} \le \frac{\sinh\epsilon}{\epsilon}\,\frac{\rho}{2} = \frac{\rho}{2\epsilon},$$

so

$$w \ge \log \frac{4\epsilon}{\rho}.$$

Any leaf of $\beta_N \cap B$ that intersects α intersects it exactly once and runs from one boundary component of B to the other and has length at least $2w$. By [Lecuire 2002, Proposition 4] we have $i(\alpha, \beta_N) > 2\pi$ (see also [Bonahon and Otal 2001, Proposition 7] for the case when β_N is finite-leaved). Thus, the total (measured) length of $\beta_N \cap B$ is at least $2\pi(2w) = 4\pi w$. Therefore

$$l_{\partial C(N)}(\beta_N) \ge 4\pi \log \frac{4\epsilon}{\rho},$$

as claimed. $\square$

7. Bounds depending on the geometry of $\Omega(\Gamma)$

We observed in [Bridgeman and Canary 2003] that a lower bound on the injectivity radius of the boundary of the convex hull implies a lower bound on the injectivity radius of the domain of discontinuity, while in [Canary 2001] we saw that a short geodesic in the domain of discontinuity implies the existence of an even shorter geodesic in the boundary of the convex hull. Therefore, we can give versions of Theorems 1 and 2 where the constants depend on the geometry of the domain of discontinuity.

If $N = \mathbb{H}^3/\Gamma$ is an analytically finite hyperbolic 3-manifold, then [Bridgeman and Canary 2003, Lemma 8.1] implies that

$$\tfrac{1}{2}e^{-m}e^{-\pi^2/(2r_0)}$$

is a lower bound for the injectivity radius of $\partial CH(L_\Gamma)$, where $m = \cosh^{-1} e^2$ and r_0 is a lower bound for the injectivity radius of the domain of discontinuity $\Omega(\Gamma)$ of Γ. Therefore, we obtain the following version of Theorem 1, where $S' = \frac{1}{2}\pi^2 S$ and $T' = S\log 2 + Sm + T$.

Theorem 1′. *There exist constants S' and T' such that if N is an analytically finite hyperbolic 3-manifold with bending lamination β_N and r_0 is a lower bound for the injectivity radius of the domain of discontinuity $\Omega(\Gamma)$, then*

$$l_{\partial C(N)}(\beta_N) \le \left|\chi(\partial C(N))\right|\left(\frac{S'}{r_0} + T'\right),$$

where $l_{\partial C(N)}(\beta_N)$ is the length of β_N and $\chi(\partial C(N))$ is the Euler characteristic of the boundary of the convex core.

Theorem 5.1 of [Canary 2001] implies that if $\Omega(\Gamma)$ contains a closed geodesic of length $r \le 1$, then $\partial CH(L_\Gamma)$ contains a closed geodesic of length at most

$$\frac{4\pi e^{0.502\,\pi}}{e^{\pi^2/(\sqrt{e}r)}} \le .153\, r.$$

Thus, we obtain the following version of Theorem 2, where $P = 4\pi^3/\sqrt{e}$ and $Q = 4\pi \log\left(4\pi e^{0.502\,\pi}/\sinh^{-1} 1\right)$.

Theorem 2′. *There exist positive constants P and Q such that if $N = \mathbb{H}^3/\Gamma$ is an analytically finite hyperbolic 3-manifold, $\Omega(\Gamma)$ contains a closed geodesic of length $r \le 1$, then*

$$l_{\partial C(N)}(\beta_N) \ge \frac{P}{r} - Q.$$

References

[Ahlfors 1964] L. V. Ahlfors, "Finitely generated Kleinian groups", *Amer. J. Math.* **86** (1964), 413–429. MR 29 #4890 Zbl 0133.04201

[Beardon 1983] A. F. Beardon, *The geometry of discrete groups*, Graduate Texts in Mathematics **91**, Springer, New York, 1983. MR 85d:22026 Zbl 0528.30001

[Bonahon 1988] F. Bonahon, "The geometry of Teichmüller space via geodesic currents", *Invent. Math.* **92**:1 (1988), 139–162. MR 90a:32025 Zbl 0653.32022

[Bonahon 2001] F. Bonahon, "Geodesic laminations on surfaces", pp. 1–37 in *Laminations and foliations in dynamics, geometry and topology* (Stony Brook, NY, 1998), edited by M. Lyubich et al., Contemp. Math. **269**, Amer. Math. Soc., Providence, RI, 2001. MR 2001m:57023 Zbl 0996.53029

[Bonahon and Otal 2001] F. Bonahon and J.-P. Otal, "Laminations mesurées de plissage des variétés hyperboliques de dimension 3", preprint, Université d'Orléans (MAPMO), 2001, Available at http://www.univ-orleans.fr/SCIENCES/MAPMO/publications/mapmo2001/mapmo0115.pdf.

[Bridgeman 1998] M. Bridgeman, "Average bending of convex pleated planes in hyperbolic three-space", *Invent. Math.* **132**:2 (1998), 381–391. MR 99f:57015 Zbl 0912.30028

[Bridgeman and Canary 2003] M. Bridgeman and R. D. Canary, "From the boundary of the convex core to the conformal boundary", *Geom. Dedicata* **96** (2003), 211–240. MR 2003j:57026 Zbl 01901496

[Buser 1992] P. Buser, *Geometry and spectra of compact Riemann surfaces*, Progress in Mathematics **106**, Birkhäuser, Boston, 1992. MR 93g:58149 Zbl 0770.53001

[Canary 2001] R. D. Canary, "The conformal boundary and the boundary of the convex core", *Duke Math. J.* **106**:1 (2001), 193–207. MR 2001m:57024 Zbl 1012.57021

[Epstein and Marden 1987] D. B. A. Epstein and A. Marden, "Convex hulls in hyperbolic space, a theorem of Sullivan, and measured pleated surfaces", pp. 113–253 in *Analytical and geometric aspects of hyperbolic space* (Coventry/Durham, 1984), edited by D. B. A. Epstein, London Math. Soc. Lecture Note Ser. **111**, Cambridge Univ. Press, Cambridge, 1987. MR 89c:52014 Zbl 0612.57010

[Epstein et al. 2004] D. B. A. Epstein, A. Marden, and V. Markovic, "Quasiconformal homeomorphisms and the convex hull boundary", *Ann. of Math.* (2) **159**:1 (2004), 305–336. MR 2052356

[Johnson and Millson 1987] D. Johnson and J. J. Millson, "Deformation spaces associated to compact hyperbolic manifolds", pp. 48–106 in *Discrete groups in geometry and analysis* (New Haven, 1984), edited by R. Howe, Progr. Math. **67**, Birkhäuser, Boston, 1987. MR 88j:22010 Zbl 0664.53023

[Kourouniotis 1985] C. Kourouniotis, "Deformations of hyperbolic structures", *Math. Proc. Cambridge Philos. Soc.* **98**:2 (1985), 247–261. MR 87g:32022 Zbl 0577.53041

[Lecuire 2002] C. Lecuire, "Plissage des variétés hyperboliques de dimension 3", preprint 0301, Université de Lyon, 2002.

[Thurston 1979] W. P. Thurston, "The geometry and topology of 3-manifolds", lecture notes, Princeton University, 1979, Available at http://www.msri.org/publications/books/gt3m/.

[Yamada 1982] A. Yamada, "On Marden's universal constant of Fuchsian groups, II", *J. Analyse Math.* **41** (1982), 234–248. MR 84h:30073 Zbl 0513.30037

Received January 3, 2003. Revised June 7, 2003.

MARTIN BRIDGEMAN
DEPARTMENT OF MATHEMATICS
BOSTON COLLEGE
CHESTNUT HILL, MA 02167
bridgem@bc.edu

RICHARD D. CANARY
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MI 48109-1109
canary@umich.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

RELATIVE FAMILY GROMOV–WITTEN INVARIANTS AND SYMPLECTOMORPHISMS

OLGUȚA BUȘE

We study the symplectomorphism groups $G_\lambda = \mathrm{Symp}_0(M, \omega_\lambda)$ of a closed manifold M equipped with a one-parameter family of symplectic forms ω_λ with variable cohomology class. We show that the existence of nontrivial elements in $\pi_*(\mathcal{A}, \mathcal{A}')$, where $(\mathcal{A}, \mathcal{A}')$ is a suitable pair of spaces of almost complex structures, implies the existence of nontrivial elements in $\pi_{*-i}(G_\lambda)$, for $i = 1$ or 2. Suitable parametric Gromov–Witten invariants detect nontrivial elements in $\pi_*(\mathcal{A}, \mathcal{A}')$. By looking at certain resolutions of quotient singularities we investigate the situation $(M, \omega_\lambda) = (S^2 \times S^2 \times X, \sigma_F \oplus \lambda\sigma_B \oplus \omega_{\mathrm{arb}})$, with $(X, \omega_{\mathrm{arb}})$ an arbitrary symplectic manifold. We find nontrivial elements in higher homotopy groups of G^X_λ, for various values of λ. In particular we show that the fragile elements w_ℓ found by Abreu and McDuff in $\pi_{4\ell}(G^{\mathrm{pt}}_{\ell+1})$ do not disappear when we consider them in $S^2 \times S^2 \times X$.

1. Introduction

Let (M^{2n}, ω) be a $2n$-dimensional compact symplectic manifold. The group of symplectomorphisms $\mathrm{Symp}(M, \omega)$ of M is a basic invariant that distinguishes among different symplectic structures on M. It is an infinite-dimensional group endowed with a natural C^∞ topology.

Two natural questions arise in relation with $\mathrm{Symp}(M, \omega)$:

(1) What can be said about the topological type of $\mathrm{Symp}(M, \omega)$?

(2) How does the topological type change as ω varies?

Research has been done in this direction by various authors [Abreu 1998; Lê and Ono 2001; McDuff 2000; Seidel 1999; 1997] by using information on J-holomorphic curves. We investigate these questions by defining relative parametric GW invariants, which are sensitive to the topology of appropriate spaces of almost

MSC2000: primary 57R17; secondary 53D35, 53D45, 57S05.
Keywords: symplectomorphism group, Gromov–Witten invariant, almost complex structure.

complex structures. The connection between the spaces of almost complex structures and the symplectomorphism groups is achieved by means of the following fibration, introduced in [Kronheimer 1998] and used in [McDuff 2001]:

$$\operatorname{Symp}_0(M,\omega) \longrightarrow \operatorname{Diff}_0 M \xrightarrow{\psi\to(\psi^{-1})^*\omega} S_{[\omega]}, \tag{1}$$

where $S_{[\omega]}$ is the space of symplectic forms that can be joined to ω through a path of cohomologous symplectic forms, $\operatorname{Diff}_0 M$ is the connected component of the identity inside the group of diffeomorphism and $\operatorname{Symp}_0(M,\omega) = \operatorname{Symp}(M,\omega) \cap \operatorname{Diff}_0 M$. Now consider the space $\mathcal{A}_{[\omega]}$ of almost complex structures that are tamed by some symplectic form in $\mathcal{S}_{[\omega]}$. By [McDuff 2001], $\mathcal{A}_{[\omega]}$ is homotopy equivalent to $\mathcal{S}_{[\omega]}$. This yields the homotopy fibration

$$\operatorname{Symp}_0(M,\omega) \longrightarrow \operatorname{Diff}_0 M \longrightarrow \mathcal{A}_{[\omega]}. \tag{2}$$

Our strategy will be to define suitable pairs $(\mathcal{A},\mathcal{A}')$ of spaces of almost complex structures, such that information on nontrivial homotopy groups in $(\mathcal{A},\mathcal{A}')$ extends to information on $\operatorname{Symp}_0(M,\omega)$. We develop relative family GW invariants that detect such nontrivial elements in $\pi_*(\mathcal{A},\mathcal{A}')$.

Outline of the methods. In Section 2 we define the invariants as follows: Consider a smooth family of symplectic forms $(\omega_\lambda)_{\lambda\in I}$, where the parameter λ varies in the interval I in $\mathbb{R}$ in such a manner that the cohomology classes $[\omega_\lambda]$ may also vary along a line L inside $H^2(M,\mathbb{R})$. For convenience we set $\mathcal{A}_\lambda := \mathcal{A}_{[\omega_\lambda]}$. Consider $D \in H_2(M,\mathbb{Z})$ and let $\mathcal{A}^c_{\lambda,D} \subset \mathcal{A}_\lambda$ be the subspace of those almost complex structures J which do not admit J-holomorphic stable maps in the class D. Further define $\mathcal{A}_I = \bigcup_{\lambda\in I}\mathcal{A}_\lambda$, and similarly let $\mathcal{A}^c_{I,D}$ be its subset consisting of $\bigcup_{\lambda\in I}\mathcal{A}^c_{\lambda,D}$. By a similar argument as in [McDuff 2001], $\mathcal{A}_I$ is homotopy equivalent with $\bigcup_{\lambda\in I}\mathcal{S}_{[\omega_\lambda]}$ and hence is connected. We will assume that there is a special almost complex structure $* = J_{\text{basepoint}}$ that belongs to all the spaces $\mathcal{A}^c_{\lambda,D}$. Consider a family of almost complex structures $(J_B, \partial J_B, *)$ that represent an element in $\pi_*(\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$. We will define a homomorphism

$$\mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k} : \bigoplus_{i=1}^{k} H^{a_i}(M,\mathbb{Q})^k \to \mathbb{Q} \tag{3}$$

by counting J_b-holomorphic stable maps in class D, for all $b \in B$. This is well defined because the class D is never represented as a J_b-holomorphic stable map if $b \in \partial B$.

Theorem 1.1. (i) *The* $\mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k}$ *are symplectic deformation invariants and depend only on the relative homotopy class of the triple* $(J_B, \partial J_B, *)$.

(ii) *For a fixed choice of* k, D *and* α_i *the map* $\Theta_{0,k,\alpha_1,\ldots,\alpha_k} : \pi_*(\mathcal{A}_I, \mathcal{A}^c_{I,D}, *) \to \mathbb{Q}$ *given by*

$$\Theta_{k,\alpha_1,\ldots,\alpha_k}\big([(J_B, \partial J_B)]\big) = \mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k}(\alpha_1, \ldots, \alpha_k).$$

is a homomorphism.

The reason (i) holds is that the class D is never represented for a J_b with $b \in \partial B$.

In Section 3 we will exhibit some examples of nontrivial PGW. There we consider the case where $M = S^2 \times S^2 \times X$ for X an arbitrary symplectic manifold and where $\omega = \omega_\lambda \oplus \omega_{\mathrm{arb}}$ with ω_{arb} an arbitrary symplectic form on X and $\omega_\lambda = \sigma_F \oplus \lambda\sigma_B$, for σ_F, σ_B forms (of total area 1) on the fiber and base, and $\lambda \geq 1$. The families $(J_B, \partial J_B)$ of almost complex structures are provided for $S^2 \times S^2$ in [Kronheimer 1998] and then further investigated in [Abreu and McDuff 2000]. One has to look at a quotient singularity, $\mathbb{C}^2/C_{2\ell}$, where $C_{2\ell}$ is the cyclic group of order 2ℓ acting diagonally by scalars on $\mathbb{C}^2$. The deformation space for the canonical resolution of this singularity provides a $4\ell-2$ family $(J_{B_\ell}, \partial J_{B_\ell}) \in (\mathcal{A}_{[\ell,\ell+\epsilon]}, \mathcal{A}_\ell)$ for which suitable PGWs are nontrivial.

The link between these examples and the corresponding groups of symplectomorphisms will be explained in Section 4. It explain there the extent to which the known homotopy properties (see [Abreu and McDuff 2000]) of $\mathrm{Symp}_0(S^2 \times S^2, \omega_\lambda)$ are reflected in the higher homotopy groups of

$$G^X_\lambda := \mathrm{Symp}_0(S^2 \times S^2 \times X, \omega_\lambda \oplus \omega_{\mathrm{arb}}).$$

For every (M, ω_λ) a general symplectic manifold, we set $G_\lambda := \mathrm{Symp}_0(M, \omega_\lambda)$.

To be able to address the two questions posed at the beginning, one has to establish first a more precise language in which they make sense. One of the difficulties is that in general there is no direct map $G_\lambda \to G_{\lambda+\epsilon}$. In the particular situation $M = S^2 \times S^2 \times \mathrm{pt}$, Abreu and McDuff [2000; McDuff 2001] find natural maps $G^{\mathrm{pt}}_\lambda \to G^{\mathrm{pt}}_{\lambda+\epsilon}$, well defined up to homotopy, and prove:

Theorem 1.2 (Abreu and McDuff). (i) *The homotopy type of* G^{pt}_λ *is constant on all the intervals* $(\ell-1, \ell]$, *with* $\ell \geq 2$ *a natural number. Moreover*, *as* λ *passes an integer* $\ell \geq 2$, *the groups* $\pi_i(G^{\mathrm{pt}}_\lambda)$, *for* $i \leq 4\ell - 5$, *do not change.*

(ii) *There is an element* $w_\ell \in \pi_{4\ell-4}(G^{\mathrm{pt}}_\lambda) \times \mathbb{Q}$ *when* $\ell - 1 < \lambda \leq \ell$ *that vanishes for* $\lambda > \ell$.

To get around the fact that there is no map $G_\lambda \to G_{\lambda+\epsilon}$ when dealing with a general manifold M, we show that for any compact $K \subset G_\lambda$ the inclusion $0 \times K \subset$

G_λ extends to a map h that fits into the commutative diagram

$$
\begin{tikzcd}
{[-\epsilon,\epsilon]\times K} \arrow[r, "h"] \arrow[d, "\mathrm{pr}_1"'] & \mathcal{G} := \bigcup (G_\lambda \times \lambda) \subset \mathrm{Diff}_0 M \times \mathbb{R} \arrow[d, "\mathrm{pr}_2"] \\
{[-\epsilon,\epsilon]} \arrow[r, "\mathrm{incl}"] & \mathbb{R}
\end{tikzcd}
\tag{4}
$$

Moreover, for any two such maps h and h' coinciding on $0\times K$, there is, for ϵ' small enough, a homotopy $H : [0,1]\times[-\epsilon',\epsilon']\times K \to \mathcal{G}$ between h and h' preserving the fibers of the natural projections. Therefore, for any cycle ρ in G_λ, there are *extensions* ρ_ϵ in $G_{\lambda+\epsilon}$ that, for ϵ sufficiently small, are unique up to homotopy. Hence they yield well defined elements in $\pi_*(G_{\lambda+\epsilon})$.

It will therefore make sense to ask what will become of an element $\rho\in\pi_*(G_\lambda)$ inside $\pi_*(G_{\lambda+\epsilon})$, for small ϵ. In this language an element $\theta_\ell\in\pi_*(G_\ell)$ is called *fragile* if any extension $\theta_{\ell+\epsilon}$ is null-homotopic in $\pi_*(G_{\ell+\epsilon})$ for $\epsilon>0$. Also, we say that a family $\eta_{\ell+\epsilon}\in\pi_*(G_{\ell+\epsilon})$, $0<\epsilon$ is *new* if there is no $\eta_\ell\in\pi_*(G_\ell)$ whose extension is $\eta_{\ell+\epsilon}$. We consider the space $\mathcal{A}_{\ell^+}$ roughly given by $\mathcal{A}_{\ell^+} := \big(\bigcap_{0<\epsilon<\epsilon_0}\mathcal{A}_{\ell+\epsilon}\big)\cup\mathcal{A}_\ell$; for the precise definition see (7). We say that an element $\alpha\in\pi_*(\mathcal{A}_{\ell^+},\mathcal{A}_\ell)$ is *persistent* if it has nonzero image under the map $\pi_*(\mathcal{A}_{\ell^+},\mathcal{A}_\ell)\to\pi_*(\mathcal{A}_{[\ell,\ell+\epsilon]},\mathcal{A}_\ell)$.

Our main theorem is the following:

Theorem 1.3. *Assume that we have a persistent element* $0\neq\beta_\ell\in\pi_k(\mathcal{A}_{\ell^+},\mathcal{A}_\ell,*)$. *Then we can construct an element* $\theta_\ell\in\pi_{k-2}(G_\ell)$ *such that either*

(A) $\theta_\ell\in\pi_{k-2}(G_\ell)$ *is a nonzero* ***fragile*** *element, or*

(B) $\theta_\ell=0$ *and there is* $\epsilon_\ell>0$ *such that we can construct a family of* ***new*** *elements* $0\neq\eta_{\ell+\epsilon}\in\pi_{k-1}(G_{\ell+\epsilon})$, *where* $0<\epsilon<\epsilon_\ell$.

Any fragile element is null-homotopic when viewed inside $\mathrm{Diff}_0 M$. Our methods do not allow us to decide in general whether or not the image of $\eta_{\ell+\epsilon}$ in $\pi_{k-1}\mathrm{Diff}_0 M$ is zero.

We show that the hypothesis of the theorem is satisfied when $M=S^2\times S^2\times X$. We consider $D=A-\ell F$. Since $\big(\sigma_F\oplus\ell\sigma_B\oplus\omega_{\mathrm{arb}}\big)(A-\ell F)=0$ we get $\mathcal{A}_\ell\subset\mathcal{A}^c_{[\ell,\ell+\epsilon],\,D}$. In this situation the $(4\ell-2)$-dimensional elements $(B_\ell,\partial B_\ell)$ obtained in Section 3 are detected as nontrivial in $\pi_{4\ell-2}(\mathcal{A}_{\ell^+},\mathcal{A}_\ell)$ and are persistent. In fact, in general PGW invariants detect persistent elements. By varying the value of the integer ℓ we obtain infinitely many values of λ for which higher-order homotopy groups of G^X_λ are nontrivial and we discuss in more detail the stability of the elements w_ℓ provided by Theorem 1.2 inside G^X_λ. We obtain:

Corollary 1.4. *For any natural number* $\ell\geq1$, *exactly one of the statements below holds.*

(A) *We can construct a nonzero* ***fragile*** *element* $w_\ell^X \in \pi_{4\ell-4}(G_\ell^X)$, *which can be identified with* $w_\ell \times \mathrm{id}$.

(B) *There exists an* $\epsilon_\ell > 0$ *for which we can construct a family of* ***new*** *elements* $0 \neq \eta_{\ell+\epsilon}^X \in \pi_{4\ell-3}(G_{\ell+\epsilon}^X)$, $0 < \epsilon < \epsilon_\ell$.

In particular this shows that the fragile elements obtained by Abreu and McDuff for $\ell > 1$ do not disappear when we consider them inside $S^2 \times S^2 \times X$. Either $0 \neq w_\ell \times \mathrm{id} \in \pi_{4\ell-4}(G_\ell^X)$ as in (A) or, if $w_\ell \times \mathrm{id} = 0$ then it yields the associated new $4\ell - 3$ dimensional elements $0 \neq \eta_{\ell+\epsilon}^X$ in $\pi_{4\ell-3}(G_{\ell+\epsilon}^X)$ for small $\epsilon > 0$ — this is case (B). For general X and for $\ell = 1$ it is known by work of Lê and Ono that (B) takes place, and moreover that $0 \neq i_*(\eta_{\ell+\epsilon}) \in \pi_1(\mathrm{Diff}_0(S^2 \times S^2 \times X))$, where i is the inclusion of a symplectomorphism group into the diffeomorphism group. Also, for $X = \mathrm{pt}$ and $\ell > 1$ we know by work of Abreu and McDuff that (A) takes place.

We don't know of any examples where case (B) takes place and $i_*(\eta_{\ell+\epsilon}) = 0 \in \pi_*(\mathrm{Diff}_0 M)$.

Our method has been inspired by the work of P. Kronheimer, who uses parametric Seiberg–Witten invariants in dimension 4, as well as by [McDuff 2000]. Similar work has been done in this direction in [Lê and Ono 2001]; by looking at related but slightly different parametric GW invariants, these authors get results about $\pi_k\big(\mathrm{Symp}_0(S^2 \times S^2 \times X, \omega_1 \oplus \omega_{\mathrm{arb}})\big)$ when $k = 1, 3$. In Section 3 we could consider $\mathbb{C}^2/C_{2\ell+1}$ instead and, by carrying out similar arguments, get the same type of results for $\mathbb{C}P^2 \# \overline{\mathbb{C}P^2} \times X$.

2. Relative parametric GW invariants

General setting. Consider a compact manifold B with boundary and a smooth map $i : (B, \partial B) \to (\mathcal{A}_I, \mathcal{A}_{I,D}^c)$. Although the invariants can be defined in this generality, for the applications we have in mind we will consider B to be an n-ball such that i represents a relative homotopy class in $\pi_*(\mathcal{A}_I, \mathcal{A}_{I,D}^c, *)$. We will often write $J_b := i(b)$ and $J_B = \mathrm{im}\, i$, and refer to $\mathrm{im}\, B$ in $\mathcal{A}_I$ as J_B. Consider also a smooth family of symplectic forms $\omega_B := (\omega_b)_{b \in B}$ where ω_b tames J_b. The ω_b need not be cohomologous, since the taming condition is an open condition. Our goal here is to show how we can define parametric GW invariants relative to the boundary ∂J_B of J_B, invariants that count J_b-holomorphic maps for some $b \in B$. These will not depend either on deformations of the family ω_B or on the representative $(J_B, \partial J_B)$ of a relative homotopy class in $(\mathcal{A}_I, \mathcal{A}_{I,D}^c)$.

Consider the space $\widetilde{\mathcal{M}}_{0,k}^*(M, D, (J_B, \partial J_B))$ of tuples $(b, f, x_1, \ldots, x_k)$, where $f : S^2 \to M$ is a simple[1] J_b-holomorphic map in class D, for some $b \in B$, and the

[1] We say that $f : \Sigma \to M$ is simple if it is not the composite of a holomorphic branched covering map $(\Sigma, j) \to (\Sigma', j')$ of degree greater than 1 with a J-holomorphic map $\Sigma' \to M$.

x_i are pairwise distinct points on S^2. We will consider

$$\mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B)) = \widetilde{\mathcal{M}}^*_{0,k}(M, D, (J_B, \partial J_B))/G,$$

where $G = \mathrm{PSL}(2, \mathbb{C})$ acts on the moduli space by reparametrizations of the domain. Denote the elements of $\mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B))$ by $[b, f, x_1, \dots, x_k]$.

In the best scenario, for a good choice of $(J_B, \partial J_B)$,

(P1) $\widetilde{\mathcal{M}}^*_{0,k}(M, D, (J_B, \partial J_B))$ is a manifold of dimension $2n+2c_1(D)+2k+\dim B$, and

(P2) $\mathcal{M}^*_{0,k} := \mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B))$ is compact.

Then the image of the map

$$ev : \mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B)) \to M^k \tag{5}$$

with $ev([b, f, x_1, \dots, x_k]) := (f(x_1), \dots, f(x_k))$ will provide a cycle $ev_*(\mathcal{M}^*_{0,k})$ in M^k which, by intersection with homology classes of complementary dimension in M^k, gives the parametric Gromov–Witten invariants.

Definition and properties of PGW. As the regularity discussion below will make clear, condition (P1) can always be achieved by the Sard–Smale theorem. However, even in situations when (P1) holds, (P2) is seldom true; the compactification $\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))$ of $\mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B))$ contains both *stable J-holomorphic maps*[2] and *nonsimple curves*, which we sometimes call multiple cover curves. These nonsimple curves could potentially produce strata of high dimension in the compactification $\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))$, and hence this space would not necessarily carry a fundamental class.

In the situation $B = \mathrm{pt}$, there are various procedures [Li and Tian 1998; Ruan 1999; Fukaya and Ono 1999] to build up a theory that would provide *a virtual moduli cycle*, that is, an object carrying a fundamental class required for the definition of the invariants.

Roughly speaking, locally one needs to consider here all the stable holomorphic maps as well as small perturbations of them. There are then various procedures to pass to a global object with the required properties. These go through without essential changes if one considers parameter spaces with no boundary; see [Bryan and Leung 2000; Ruan 1999].

In our situation we need to make sure that the boundary causes no problem. In what follows denote by $[f, \Sigma, x_1, \dots, x_k]$ the equivalence class of a stable map $(f, \Sigma, x_1, \dots, x_k)$, where two maps are equivalent if they differ by an automorphism of the domain. The elements of $\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))$ consist of such

[2] These are rational maps $f : (\Sigma, x_1, \dots, x_k) \to M$ with the most normal crossing singularities and no infinitesimal automorphisms; see [Li and Tian 1998; Bryan and Leung 2000] for details.

equivalence classes. The next result states that if we consider an appropriately small open neighborhood of $\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))$ consisting of almost holomorphic stable maps, its projection onto J_B stays away from ∂J_B.

Lemma 2.1. *For any compact set $J_B \in \mathcal{A}_I$ such that $\partial J_B \subset \mathcal{A}^c_{I,D}$, there exist $\delta > 0$ and $\epsilon(\delta) > 0$ for which there is no stable map $(f, \Sigma, x_1, \dots, x_k)$ such that $\bar\partial_J f = \nu$, when $d(J, \partial J_B) < \delta$ and $\nu \in L^p(\Lambda^{0,1} \otimes_J f^*TM)$ with $|\nu| \le \epsilon(\delta)$.*

Proof. We will prove this by assuming the opposite. Assume we have sequences J_i, ν_i and f_i such that $d(J_i, \partial J_B) \to 0$, $|\nu_i| = \epsilon_i \to 0$ and each f_i is a stable map in class D with the property that $\bar\partial_{J_i} f_i = \nu_i$. Since J_B is compact we find a convergent subsequence J_i whose limit J_∞ is in ∂J_B. But by the Gromov compactness theorem there is a subsequence of f_i converging to a J_∞ stable holomorphic map in class D. This contradicts the fact that $J_\infty \in \partial J_B \subset \mathcal{A}^c_{I,D}$. □

With this lemma one shows, as in [Li and Tian 1998], that every moduli space $\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))$ carries a virtual fundamental cycle

$$[\mathcal{M}]^{\mathrm{vir}} := [\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))]^{\mathrm{vir}}$$

of degree $r = 2c_1(D) + 2k + 2n - 6 + \dim B$.

Moreover, if we take two homotopic maps $i : (B, \partial B, *) \to (\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$ and $i' : (B', \partial B', *) \to (\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$ representing the same element in $\pi_*(\mathcal{A}, \mathcal{A}^c_D, *)$, then the corresponding fundamental cycles given by $[\overline{\mathcal{M}}_{0,k}(M, D, (J_B, \partial J_B))]^{\mathrm{vir}}$ and $[\overline{\mathcal{M}}_{0,k}(M, D, (J_{B'}, \partial J_{B'}))]^{\mathrm{vir}}$ are oriented cobordant and hence the virtual fundamental class $[\mathcal{M}]^{\mathrm{vir}}$ is independent of the choice of $(J_B, *)$ within the same class in $\pi_*(\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$. Note that $[\mathcal{M}]^{\mathrm{vir}}$ is also invariant under symplectic deformation of the family of taming symplectic forms $(\omega_b)_{b\in B}$. We denote by $\overline{\mathcal{F}}_D(M, 0, k)$ the space of all equivalences classes of stable maps $[f, \Sigma, x_1, \dots, x_k]$ with total homology D. To define relative parametric Gromov–Witten invariants we consider $ev_i : B \times \overline{\mathcal{F}}_D(M, 0, k) \to M$ given by

$$ev_i(b, [f, \Sigma, x_1, \dots, x_k]) = f(x_i).$$

We then can define

$$\mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k} : \bigoplus_{i=1}^{k} H^{a_i}(M, \mathbb{Q})^k \to \mathbb{Q}$$

by

$$\mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k}(\alpha_1, \dots, \alpha_k) = ev_1^*(\alpha_1) \wedge \cdots \wedge ev_k^*(\alpha_k)[\mathcal{M}]^{\mathrm{vir}}$$

which are zero unless

$$\sum_{i=1}^{k} a_i = 2c_1(D) + 2k + 2n - 6 + \dim B. \tag{6}$$

Changing the orientation of B just changes the sign of the invariant.

Theorem 2.2. (i) *The invariants* $\mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k}$ *are symplectic deformation invariants and depend only on the relative homotopy class of* $(J_B, \partial J_B)$.

(ii) *For a fixed choice of* k, D *and* α_i *the map* $\Theta_{0,k,\alpha_1,\ldots,\alpha_k} : \pi_*(\mathcal{A}_I, \mathcal{A}^c_{I,D}, *) \to \mathbb{Q}$ *given by*

$$\Theta_{k,\alpha_1,\ldots,\alpha_k}([(J_B, \partial J_B)]) = \mathrm{PGW}^{M,(J_B,\partial J_B)}_{D,0,k}(\alpha_1, \ldots, \alpha_k)$$

is a group homomorphism.[3]

Proof. Point (i) and the well definedness of Θ follow from the properties of PGW listed above. To show that Θ is a homeomorphism, choose $(B_1, \partial B_1, *)$, and $(B_2, \partial B_2, *)$ representing two maps from the standard n-ball with boundary to $(\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$, giving two elements β_1 and β_2 inside $\pi_*(\mathcal{A}, \mathcal{A}^c_D *)$. We choose them in such a way that by their concatenation we represent the element $\beta_1 + \beta_2$ by a map $j : (B, \partial B, *) \to (\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$ with $j(B \setminus \partial B) = (B_1 \setminus \partial B_1) \cup (B_2 \setminus \partial B_2)$, so that $j^{-1}(\mathcal{A}_I \setminus \mathcal{A}^c_{I,D})$ is included in the disjoint union of two open subdiscs in B. Then the new virtual cycle corresponding to the classes $\beta_1 + \beta_2$ is a disjoint union of the virtual neighborhoods corresponding to β_1 and β_2. But this implies that the parametric invariants corresponding to the new class $\beta_1 + \beta_2$ are the sum of the PGW corresponding to β_1 and β_2. Therefore Θ is a homomorpism. □

More on the relation between PGW and almost complex structures. We will now see that PGW detects only certain kinds of relative homotopy classes of almost complex structures. As before, we write $\mathcal{A}_\lambda = \mathcal{A}_{\omega_\lambda}$. Set

$$\mathcal{A}_{\ell^+} = \{J \mid \text{there is } \epsilon_J > 0 \text{ such that } J \in \mathcal{A}_{\ell+\epsilon} \text{ for all } 0 < \epsilon < \epsilon_J\} \tag{7}$$

Then $\mathcal{A}_\ell \subset \mathcal{A}_{\ell^+}$ by Lemma 4.1 below. Note that $\mathcal{A}_{\ell^+}$ may not be connected, but $\mathcal{A}_\ell$ is and we will consider our basepoint $* = J_{\text{basepoint}} \in \mathcal{A}_\ell$.

Definition 2.3. Consider a nontrivial element $\beta_\ell \in \pi_*(\mathcal{A}_{\ell^+}, \mathcal{A}_\ell)$. We say that β_ℓ is *persistent* if its image under the natural morphism

$$i_* : \pi_*(\mathcal{A}_{\ell^+}, \mathcal{A}_\ell, *) \to \pi_*(\mathcal{A}_{[\ell,\ell+\epsilon]}, \mathcal{A}_\ell, *)$$

is nonzero for any arbitrary small ϵ.

Proposition 2.4. *Assume there is an* ℓ *such that no* J *in* $\mathcal{A}_\ell$ *admits* J*-holomorphic stable maps in class* D. *Consider an element* $0 \neq \beta_\ell \in \pi_*(\mathcal{A}_{\ell^+}, \mathcal{A}_\ell, *)$ *obtained by counting nontrivial parametric Gromov–Witten invariants in class* D. *Then* β_ℓ *is a persistent element.*

Proof. The proof follows directly from Theorem 2.2. □

[3] Except in the case of $\pi_1(\mathcal{A}_I, \mathcal{A}^c_{I,D}, *)$, which is not a group.

Computability of PGW. We will now get back to the two conditions we posed at the beginning of the section, sufficient to imply that the image of the map (5) is a cycle. Below we will provide sufficient hypotheses on the parameter space $(J_B, \partial J_B, *)$ and on the class D such that (P1) and (P2) are satisfied, as well as a criterion for how to check one of the hypothesis. It will follow for such a family $(J_B, \partial J_B, *)$ the invariants PGW defined above are integer-valued and can be obtained by intersecting the image of the cycle $ev_*\big(\overline{\mathcal{M}}^*_{0,k}(M, D, (J_B, \partial J_B))\big)$ with the classes $(PD(\alpha_1), \dots, PD(\alpha_k))$ in $H_*(M)^k$. Moreover, they can be obtained by counting the number of J_b-holomorphic maps in class D with k marked points which intersect generic cycles representing $(PD(\alpha_1), \dots, PD(\alpha_k))$ in $f(z_i)$.

Parametric regularity. We now show that *D-parametric regular* families $(J_B, \partial J_B)$ are ones for which (P1) is satisfied. We begin by explaining what D-parametric regularity is and contrasting it with the usual D-regularity for J (see [McDuff and Salamon 1994]). For this we need the following facts.

Let $\mathscr{X} = \mathrm{Map}(\Sigma, M; D)$ be the space of *somewhere injective*[4] smooth maps $f : \Sigma \to M$ representing class D. This is an infinite-dimensional manifold with $T_f\mathscr{X} = C^\infty(f^*TM)$. We will next consider the following generalized vector bundle $\mathscr{E} \to B \times \mathscr{X}$, whose fiber at (b, f) is the space $\mathscr{E}_{b,f} = \Omega^{0,1}_{J_b}(f^*TM)$ of smooth J_b antilinear forms with values in f^*TM. In this vector bundle we consider a section $\Phi : B \times \mathscr{X} \to \mathscr{E}$, given by

$$\Phi(b, f) = \frac{1}{2}(df + J_b \circ df \circ j). \tag{8}$$

The zeros of Φ are precisely the J_b-holomorphic maps and thus the moduli space

$$\widetilde{\mathcal{M}}^*_{0,0}(M, D, (J_B, \partial J_B)) = \Phi^{-1}(0),$$

is the intersection of $\operatorname{im} \Phi$ with the zero section of the bundle. Since we would like $\widetilde{\mathcal{M}}^*_{0,k}(M, D, (J_B, \partial J_B))$ to be a manifold, we require that Φ be transversal to the zero section. This means that the image of $d\Phi(b, f)$ is complementary to the tangent space $T_bB \oplus T_f\mathscr{X}$ of the zero section. But for any f which is J_b-holomorphic, $d\Phi$ is given by

$$d\Phi(b, f) : T_bB \oplus C^\infty(f^*TM) \longrightarrow T_bB \oplus T_f\mathscr{X} \oplus \mathscr{E}_{b,f}.$$

If we now consider the projection onto the vertical space of the bundle,

$$proj_2 : T_bB \oplus T_f\mathscr{X} \oplus \mathscr{E}_{b,f} \longrightarrow \mathscr{E}_{b,f},$$

[4]We say that a map $f : \sigma \to M$ is *somewhere injective* if $df(z) \neq 0$ and $f^{-1}(f(z)) = z$ for some $z \in \Sigma$. A simple J-holomorphic map is somewhere injective; see [McDuff and Salamon 1994].

the transversality mentioned above translates into the fact that

(9) $$\mathrm{proj}_2 \circ d\Phi(b, f) : T_bB \oplus C^\infty(f^*TM) \longrightarrow \Omega^{0,1}_{J_b}(\Sigma, f^*TM)$$

is onto. We introduce the notation $D\Phi(b, f) = \mathrm{proj}_2 \circ d\Phi(b, f)$.

Definition 2.5. We say that a J_b-holomorphic map f is J_B-parametric regular if $D\Phi(b, f)$ is onto.

Observation. The linearized operator is well defined if there is no pair (b, f) with f a J_b-holomorphic and $b \in \partial B$. This is precisely the condition we imposed on $(J_B, \partial J_B)$ to give a relative cycle in $(\mathcal{A}_I, \mathcal{A}^c_{I,D})$.

Definition 2.6. Consider (J_B, ω_B) as above. We say that $(J_B, \partial J_B)$ is a D-*parametric regular family of almost complex structures* if any J_b-holomorphic map in class D is parametric regular. We denote by $J_{\mathrm{preg}}(D)$ the set of all D-parametric regular families $(J_B, \partial J_B) \subset (\mathcal{A}_I, \mathcal{A}^c_{I,D})$.

To apply the implicit function theorem and the Sard–Smale theorem, we must work on Banach manifolds and hence complete all spaces under suitable Sobolev norms. For example, one should work on spaces consisting of almost complex structures of class C^l, on $\mathcal{X}^{k,p}$, with $kp > 2$, the space of maps whose k-th derivatives are of class L^p. Also, we should work on

$$\mathcal{E}^p_f = L^p(\Lambda^{0,1} \otimes_J f^*TM)$$

rather that with $\Omega^{0,1}_J(\Sigma, f^*TM)$. There are standard arguments [McDuff and Salamon 1994] to show that one can transfer the following arguments from spaces of C^l objects (which are Banach manifolds) to spaces of C^∞ objects (which are Fréchet manifolds). For simplicity we will drop the superscripts l, k, p unless specifying them is relevant.

Theorem 2.7. *If $J_B \in J_{\mathrm{preg}}(D)$, the moduli space $\widetilde{\mathcal{M}}^*_{0,0}(M, D, (J_B, \partial J_B))$ is a smooth open manifold of dimension $2n + 2c_1(D) + \dim B$, with a natural orientation.*

Moreover, if one considers $\widetilde{\mathcal{M}}^*_{0,0}(M, D, (J_B, \partial J_B)) \times (S^2)^k$ and takes away all the diagonals of the type $\widetilde{\mathcal{M}}^*_{0,0}(M, D, (J_B, \partial J_B)) \times \mathrm{diag}_{i,j}$, one obtains precisely $\widetilde{\mathcal{M}}^*_{0,k}(M, D, (J_B, \partial J_B))$. This will therefore be a manifold of dimension $2n + 2c_1(D) + \dim B + 2k$.

Let $\widetilde{\mathcal{M}}^*_{0,0}(M, D, \mathcal{A}_I)$ be the *universal* moduli space consisting of pairs (f, J), where $J \in \mathcal{A}_I$ and f is J-holomorphic. It will be relevant for the results we have in mind to point out the following characterization of parametric regularity.

Proposition 2.8. *Consider the diagram*

$$\begin{tikzcd} & \widetilde{\mathcal{M}}^*_{0,0}(M, D, \mathcal{A}_I) \arrow[d, "\Pi"] \\ (B, \partial B) \arrow[r, "i"] & (\mathcal{A}_I, \mathcal{A}^c_{I,D}) \end{tikzcd} \tag{10}$$

Then $J_B \in J_{\text{preg}}(A)$ *if and only if* $i \pitchfork \Pi$.

Proof. For simplicity we will write $D_{f,b} = D\Phi(b, f)_{|C^\infty(f^*(TM))}$. By (9), the surjectivity of $D\Phi(b, f)$ is then equivalent to the surjectivity of the linear operator

$$D\phi_{|T_bB} : T_bB \to \operatorname{coker} D_{b,f}.$$

We will set $i(b) = J$. The tangent space $T_J\mathcal{A}_I$ to $\mathcal{A}_I$ consists of all sections Y of the bundle $\operatorname{End}(TM, J)$ whose fiber at $p \in M$ is the space of linear maps $Y : T_pM \to T_pM$ such that $YJ + JY = 0$; we will consider the map

$$R : T_J\mathcal{A}_I \to \Omega^{0,1}_J \big(\Sigma, f^*TM\big)$$

given by $R(Y) = \frac{1}{2}Y \circ df \circ j$. The map

$$d\Pi : T_{f,J}\widetilde{\mathcal{M}}^*_{0,0}(M, D, \mathcal{A}_I) \to T_J\mathcal{A}_I$$

is given by $d\Pi(\xi, Y) = Y$, where the pair (ξ, Y) is in $T_{f,J}\widetilde{\mathcal{M}}^*_{0,0}(M, D, \mathcal{A}_I)$ if and only if

$$D_{f,b}(\xi) + R(Y) = 0. \tag{11}$$

From this one can see that $\operatorname{im} D_{f,b} = R(\operatorname{im} d\Pi)$. Since $D_{b,f}$ is elliptic and $\ker R \subset \operatorname{im} d\Pi$, it follows that $\operatorname{coker} d\Pi$ has finite dimension. If we consider the map $\mathcal{F} : \mathcal{X} \times \mathcal{A}_I \to \mathcal{E}$, given by $\mathcal{F}(f, J) = \bar{\partial}_J(f)$ then (see [McDuff and Salamon 1994]) the linearization at a zero (f, J) with f simple is onto. That is

$$D\mathcal{F}(f, J)(\xi, Y) = D_f\xi + R(Y)$$

is onto. This implies that $\operatorname{coker} D_f$ is covered by R. We can show that there is an induced map

$$\tilde{R} : \operatorname{coker} d\Pi \to \operatorname{coker} D_{b,f}$$

which is isomorphism. We have $D\Phi_{|T_bB}(Y) = R \circ di$, so

$$i \pitchfork \Pi \iff di \to \operatorname{coker} d\Pi \ \text{ onto} \iff \tilde{R} \circ di \to \operatorname{coker} D_{b,f} \ \text{ onto}.$$

The proposition follows. □

We call attention to a few key points. Parametric regularity is a generalization of the usual regularity. Indeed, if we consider $J_b = J$ to be constant for b in a neighborhood around b_0, the regularity of an almost complex structure J simply says, following the diagram above, that $d\Pi$ is surjective. If we now regard J within an arbitrary family J_B, this no longer needs to be the case. It will then suffice that the cokernel of $d\Pi$ is covered by the variation of J in the direction of B.

In fact, when we count rational maps, the criterion of parametric regularity described below reduces the problem to the usual regularity in some suitable ambient space.

More precisely, note that the regularity of a holomorphic map is a local statement within B and it only concerns the almost complex structure data. Therefore, for each $b \in \mathring{B}$, we can restrict our attention to a neighborhood of b, and without loss of generality the following discussion can be made for smoothly trivial fibrations. We say that a family (J_B, ω_B) *descends from a fibration* $M \to \widetilde{M} \to B$ if there is a diagram

$$\begin{tikzcd} M \arrow[r, "i"] & \widetilde{M} \arrow[d, "\pi"] \\ & B \end{tikzcd} \tag{12}$$

such that the almost complex structure $\tilde{J}$ on $\widetilde{M}$ yields, by restriction to each fiber $M \times b$, the almost complex structure J_b on M, and such that the closed two-form $\widetilde{\omega}$ on $\widetilde{M}$ also gives, by restriction to each fiber, the symplectic form ω_b, which tames J_b. Here we have chosen a trivialization of the fibration such that $\widetilde{M} = B \times M$ smoothly and π is just the projection on the first factor. In the following theorem we consider the family of parameters B to be a subset of $\mathbb{C}^m$ and we denote by z the parameter.

Theorem 2.9. *Let $(J_z, \omega_z)_{z \in B \subset \mathbb{C}^m}$ be a family on M descending from the symplectic fibration $(\widetilde{M}, \tilde{J}, \widetilde{\omega})$. Suppose that $f : \Sigma \to M$ is a J_0-holomorphic map and consider the composite map*

$$\tilde{f} = i \circ f, \quad \tilde{f} : \Sigma \to M \times 0 \subset \widetilde{M},$$

which is $\tilde{J}$-holomorphic. If $\tilde{f}$ is regular, f is (J_z)-parametric regular. If $\Sigma = S^2$, the reverse statement also holds.

For the proof of the Theorem see the Appendix.

There exists a large subset of parametric regular families of almost complex structures inside $(\mathscr{A}_I, \mathscr{A}^c_{I,D})$. This is because one can employ the Sard–Smale theorem [Smale 1965] and show that any map $i : (B, \partial B) \to (\mathscr{A}_I, \mathscr{A}^c_{I,D})$ in Proposition 2.8 can be perturbed to an i' such that $i' \pitchfork \Pi$.

Definition 2.10. We will say that $(J_B, \partial J_B)$ *satisfies hypothesis* H_1 if it is a *D-parametric regular* family of almost complex structures.

Compactness. Even in those situations when (P1) is easily achieved using Sard–Smale, (P2) is seldom true. However, (P2) is true when k is either 0 or 1, and the class D is J_b *indecomposable* for any $b \in B$. This means that no J_b-holomorphic map in class D can decompose into a connected union of J_b-holomorphic spheres $C = C^1 \cup C^2 \cup \cdots \cup C^N$ such that each C^i represents the class D_i and $D = D_1 + \cdots + D_N$. Then as a consequence of Gromov's compactness theorem it follows that $\mathcal{M}^*_{0,k} := \mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B))$ is compact and hence in this situation the image of $ev : \mathcal{M}^*_{0,k}(M, D, (J_B, \partial J_B)) \to M^k$ is a cycle.

Definition 2.11. We will say that the *hypothesis* H_2 *is satisfied* by $(J_B, \partial J_B)$ and D if the class D is J_b indecomposable for every $b \in B$.

Note that if D is J_b-indecomposable and $k \geq 2$ then in order to compactify the image of the evaluation map one only needs to add the limits of sequences of J-holomorphic maps for which two distinct marked points converge to each other. Hence $ev(\mathcal{M}^*_{0,k})$ will have boundary of codimension 2 or more and hence it will carry a fundamental class.

3. Resolutions of singularities and relative PGW

Quotient singularities. We now give an overview of work of Kronheimer [1998] and Abreu and McDuff [2000] on how to construct special families of almost complex structures arising from the study of the total spaces of deformations for some quotient singularities. At the end of the section we will explain how these families serve our purpose of counting nontrivial PGWs. The local picture is as follows [Kronheimer 1998]:

We consider the particular type of Hirzebruch–Jung singularity $Y_0 = \mathbb{C}^2/C_{2\ell}$, given by the diagonal action by scalars of $C_{2\ell}$ on $\mathbb{C}^2$, where $C_{2\ell}$ is the cyclic group of order 2ℓ. This admits a resolution $\sigma_0 : \widetilde{Y}_0 \to Y_0$, where $\widetilde{Y}_0$ is the total space of the line bundle of degree -2ℓ over $\mathbb{C}P^1$. The exceptional curve of the resolution, we will call it E, is a curve of self-intersection -2ℓ and is the zero section of $\widetilde{Y}_0$. This resolution admits a $(2\ell - 1)$-complex-dimensional parameter family of deformations $\tilde{Y}_t$, $t \in \mathbb{C}^{2\ell-1}$. With the exception of the case $\ell = 2$ the total space $\widetilde{Y} = \bigcup \widetilde{Y}_t$ of the family of deformations is the total space of the vector bundle $\mathbb{O}(-1)^{2\ell}$. More precisely, we consider the exact sequence of bundles

$$\mathbb{O}(-2\ell) \longrightarrow \mathbb{O}(-1)^{2\ell} \xrightarrow{r} \mathbb{O}^{2\ell-1}, \tag{13}$$

where r is given by evaluating at $2\ell - 1$ generic sections of the dual, $\widetilde{Y}^* = \mathbb{O}(1)^{2\ell}$ of $\widetilde{Y}$. Since holomorphically $\mathbb{O}^{2\ell-1}$ is trivial, we can project it to its fiber $\mathbb{C}^{2\ell-1}$

and hence obtain a submersion $\tilde{q} : \mathbb{O}(-1)^{2\ell} \to \mathbb{C}^{2\ell-1}$ with $\widetilde{Y}_t = \tilde{q}^{-1}(t)$. Also it can be seen that $\widetilde{Y}$ is diffeomorphic with $\widetilde{Y}_0 \times \mathbb{C}^{2\ell-1}$ and a choice of trivialization provides a fiberwise diffeomorphism

$$\theta : \widetilde{Y} \overset{C^\infty}{=\!=\!=} \widetilde{Y}_0 \times \mathbb{C}^{2\ell-1}, \tag{14}$$

where $\widetilde{Y}_0$ is the total space of the bundle $\mathbb{O}(-2\ell)$. Now consider a 4ℓ-dimensional basis of sections in the dual $\widetilde{Y}^*$. Here the space of holomorphic sections is given by $\bigoplus_{i=1}^{2\ell} H^0(CP^1, \mathbb{O}(1)) \tilde{=} (\mathbb{C}^2)^{2\ell}$. Denote by Y the subspace of $(\mathbb{C}^2)^{2\ell}$ consisting of 2ℓ-tuples of vectors in $\mathbb{C}^2$ spanning either zero or a line. By evaluating all the 4ℓ section we obtain a map

$$\sigma : \widetilde{Y} \longrightarrow Y \subset \mathbb{C}^{4\ell}$$

that contracts E to a point $y_0 = \sigma(E)$. Moreover, y_0 is the only singular point of Y and the morphism is one-to-one outside E. Define a map $q : Y \to \mathbb{C}^{2\ell-1}$ by evaluating at the original $2\ell - 1$ generic sections. The diagram

$$\begin{tikzcd} \widetilde{Y} \arrow[r, "\sigma"] \arrow[d, "\tilde{q}"'] & Y \arrow[d, "q"] \\ \mathbb{C}^{2\ell-1} \arrow[r, "\mathrm{id}"] & \mathbb{C}^{2\ell-1} \end{tikzcd} \tag{15}$$

commutes. We can obtain a two-form τ on Y by pulling back a Kähler form from $\mathbb{C}^{4\ell}$. Via σ^* this can be seen as a two-form on $\widetilde{Y}$ that restricts to a Kähler form τ_t on each fiber $\widetilde{Y}_t$ if $t \neq 0$ but degenerates along E when $t = 0$. If we further push forward through θ, these forms can be seen as a family of forms on $\widetilde{Y}_0$.

As in [Abreu and McDuff 2000], we can choose an appropriate compactification of the local picture as follows:

Let $B^{4\ell-2}$ be the unit ball in $\mathbb{C}^{2\ell-1}$. We have a family $(\overline{Y}_t, J_t^\ell, \tau_t)_{t \in B^{4\ell-2}}$, where each $(\overline{Y}_t, J_t^\ell, \tau_t^\ell)$, $t \neq 0$ is a Kähler manifold diffeomorphic with $S^2 \times S^2$, and, $(\overline{Y}_0, J_0^\ell)$ is a complex manifold, also diffeomorphic with $S^2 \times S^2$ and τ_0 degenerates along E which represents the homology class $A - \ell F$. We take $A = [S^2_{\text{base}}]$ and $F = [S^2_{\text{fiber}}]$. The total space of the family has the following properties:

- The space $\overline{Y} = \cup_{t \in B^{4\ell-2}} \overline{Y}_t$ is smoothly diffeomorphic to $S^2 \times S^2 \times B^{4\ell-2}$. Moreover $\overline{Y}$ is a complex manifold with a complex structure $\tilde{J}^\ell$ which restricts to each fiber $\overline{Y}_t$ to the complex structure J_t^ℓ. Also, $\overline{Y}$ has a closed $(1, 1)$ form τ which is satisfies all the properties of a Kähler form outside the zero fiber and restricts at each fiber to the forms τ_t .
- The restriction of τ to $\overline{Y}_0$ degenerates along the curve E representing the class $A - \ell F$.

Since the forms τ_t are obtained by restricting the closed form τ to fibers, it is clear that they are all in the same cohomology class. From $(\tau_0)_{|E} = 0$ we have $[\tau_0](A - \ell F) = 0$, and hence $[\tau_t^\ell] = [\omega_\ell]$ for all $t \in B^{4\ell-2}$, where, as in the introduction, $\omega_\ell = \sigma_F \oplus \ell\sigma_B$ is a symplectic form on $S^2 \times S^2$.

From (a) we see that there is a holomorphic projection $\pi : \overline{Y} \to S^2 \times B^{4\ell-2}$. This is because every $\overline{Y}_t$ is a ruled surface therefore it fibers over S^2. If we denote by α the area form on S^2 we can construct a two-form

$$\tau^\lambda = \tau + (\lambda - \ell)\pi^*(\alpha)$$

For $\lambda > \ell$ these forms are Kähler forms, and restricted to each $\overline{Y}_t$ they yield symplectic forms in the class $[\omega_\lambda]$. This proves that any J_t^ℓ (including J_0^ℓ) is tamed by a form isotopic to ω_λ, as long as $\lambda > \ell$. We now follow a similar procedure to construct a family of symplectic forms ω_t, for $t \in B^{4\ell-2}$, such that each ω_t tames J_t^ℓ. We next change the forms τ_t by perturbing with a a positive factor of $\pi^*(\alpha)$ only around $t = 0$ and smooth with a cutoff function. With this procedure we obtain symplectic forms ω_t with variable cohomology classes.

In conclusion, we have pairs

$$(S^2 \times S^2, J_t^\ell, \omega_t)_{t\in B^{4\ell-2}},$$

where ω_t is a *symplectic* structure on $S^2 \times S^2$ that tames J_t^ℓ. Moreover $[\omega_t]_{t\in S^{4\ell-3}} = [\omega_\ell]$. This gives a family of almost complex structures (which we denote B_ℓ, by abuse of notation) such that $(B_\ell, \partial B_\ell) \in (\mathcal{A}_{[\ell,\ell+\epsilon]}, A_\ell)$ for any $\epsilon > 0$. More importantly, for cohomological reasons, only J_0^ℓ admits almost holomorphic stable curves in the class $A - \ell F$.

We then obtain a family of almost complex structures on $(S^2 \times S^2 \times X)$ by taking $(J_t^\ell \times J_{\mathrm{arb}})$, and by abuse of notation, we call this family also B_ℓ. Thus we have just produced on $(S^2 \times S^2 \times X)$ pairs $(B_\ell, \partial B_\ell) \subset (\mathcal{A}_{[\ell,\ell+\epsilon]}, A_\ell)$, with $\epsilon > 0$, representing an element β_ℓ in $\pi_*(\mathcal{A}_{[\ell,\ell+\epsilon]}, A_\ell)$. Moreover each B_ℓ is contained in $\mathcal{A}_{\ell+\epsilon}$ for any small $\epsilon > 0$.

From the choice of the J's we know that the only almost complex structure admitting $A - \ell F$ almost complex stable curves is $J_0 \times J_{\mathrm{arb}}$.

The computation of PGW. Here we prove that (H_1) and (H_2) are satisfied for the family $(B_\ell, \partial B_\ell)$, and therefore the invariant is integer-valued and can be obtained by counting holomorphic maps intersecting generic cycles of appropriate dimension.

Claim 1. *The family $(B_\ell, \partial B_\ell)$ satisfies H_1.*

Proof. From the sequence (13) we see that the exceptional curve E, which is $\tilde{J}^\ell$-holomorphic, has normal bundle $\mathcal{O}(-1)^{2\ell}$; therefore we can apply [McDuff and Salamon 1994, Lemma 3.5.1, p. 38] for the integrable almost complex structure $\tilde{J}$.

If follows that E is $\tilde{J}^\ell$-regular inside $\overline{Y}$. If we now consider $\overline{Y} \times X$ and $\tilde{J}^\ell \times J_{\mathrm{arb}}$, the curve E lies entirely inside $\overline{Y}$ and therefore the normal bundle inside $\overline{Y} \times X$ is $\mathbb{O}(-1)^{2\ell} \times$ trivial, and therefore the curve is $(\tilde{J}^\ell \times J_{\mathrm{arb}})$-regular. This splitting and therefore regularity use the fact that the map E is of genus zero. Theorem 2.9 implies parametric regularity and therefore (H_1) holds. ☐

Claim 2. *The family* $(B_\ell, \partial B_\ell)$ *satisfies* H_2.

Proof. This is proved by inspection. Only $J_0^\ell \times J_{\mathrm{arb}}$ admits $(A - \ell F)$-stable maps, and the only maps in this class are copies of the embedded map E in any fiber $S^2 \times S^2 \times \mathrm{pt}$. Hence there are no decomposable J_b-holomorphic maps. ☐

Remark. For other almost complex structures J on $S^2 \times S^2 \times X$ one could have decomposable J-holomorphic maps in the class $A - \ell F$. For an example, consider $M = S^2 \times S^2 \times \mathbb{C}P^n$ and $\omega = \omega_{1+\epsilon} \oplus \omega_{\mathrm{arb}}$. If H denotes the hyperplane class in $\mathbb{C}P^n$, we can take ω_{arb} such that $\omega(A - F - H) > 0$ and get a symplectic embedding of S^2 into M, in the class $A - F - H$. We can choose an ω-tamed almost complex structure $\tilde{J}$ on M that fibers over the base $S^2 \times S^2$ and such that the class H has a $\tilde{J}$-holomorphic representative. Then the class $A - F$ is $\tilde{J}$-decomposable, where the decomposition is given by a C with $C = C_1 \cup C_2$ with $[C_1] = A - F - H$ and $[C_2] = H$.

We conclude that the invariants

$$\mathrm{PGW}_{A-\ell F,0,k}^{S^2\times S^2\times X,(B_\ell,\partial B_\ell)} : \bigoplus_{i=1}^{k} H^{a_i}(S^2 \times S^2 \times X, \mathbb{Q})^k \to \mathbb{Z}$$

are integer-valued. We have two situations. First, if $X = \mathrm{pt}$, the moduli space of unparametrized curves has dimension 0, so we would count isolated curves. This follows immediately from the equality $c_1(A - \ell F) = -4\ell + 2$ (adjunction formula), so that

$$\begin{aligned}\dim \mathcal{M}_{0,0}^*(S^2 \times S^2, A - \ell F, (B_\ell, \partial B_\ell)) &= 2 \times 2 + 2c_1(A - \ell F) + \dim B^\ell - 6 \\ &= 4 - 4\ell + 4 + 4\ell - 2 - 6 = 0.\end{aligned}$$

Moreover, the invariant $\mathrm{PGW}_{A-\ell F,0,0}^{S^2\times S^2\times X,(B_\ell,\partial B_\ell)}([\mathrm{pt}])$ equals 1 because it counts E, the only J_b-map (where $b \in B^\ell$) in the class $A - \ell F$.

In the situation $\dim X = 2n > 0$, we will count maps with one marked point. Then $c_1(A - \ell F)$ is the same, since the holomorphic maps in class $A - \ell F$ will be copies of the curve E and hence will have the image entirely in the fibers $S^2 \times S^2 \times \mathrm{pt} \subset S^2 \times S^2 \times X$. We therefore have

$$\begin{aligned}\dim \mathcal{M}_{0,1}^*\big(S^2 \times S^2 \times X, A - \ell F, (B_\ell, \partial B_\ell)\big) &= 2 \times (2+n) + 2c_1(A - \ell F) + \dim B^\ell - 6 + 2 \\ &= 2n + 2.\end{aligned}$$

We consider a cycle in the homology class F lying in a fiber $S^2 \times S^2 \times \mathrm{pt}$ inside $S^2 \times S^2 \times X$. It easily follows that the only J_{b_ℓ}-holomorphic map with one marked point that intersects this cycle transversely is a copy of the map E inside the fiber $S^2 \times S^2 \times \mathrm{pt}$. We obtain

$$\mathrm{PGW}_{A-\ell F,0,1}^{S^2\times S^2\times X,(B_\ell,\partial B_\ell)}(PD([F])) = \pm 1,$$

where the sign depends on the orientation of the parameter space B_ℓ. Applying Theorem 2.2 we conclude that the morphism Θ in both situations is nontrivial and therefore there is a nonzero element

$$\beta_\ell \in \pi_{4\ell-2}((\mathcal{A}_{[\ell,\ell+\epsilon]}, A_\ell)) \quad \text{for all } \epsilon > 0 \tag{16}$$

represented by the cycle $(B_\ell, \partial B_\ell) \subset (\mathcal{A}_{\ell+\epsilon}, \mathcal{A}^c_{\ell+\epsilon,D})$.

4. Almost complex structures and symplectomorphism groups

Almost complex structures and symplectomorphisms; deformations along compact subsets. We now give a quick overview of what can be said about the behavior of spaces of almost complex structures and about the symplectomorphism groups as the symplectic form varies along the line L.

If L happens to be a ray $\lambda\omega$, $\lambda > 0$, then G_λ is independent of λ. Thus we may as well assume L is not a ray.

If $M = S^2 \times S^2$, much is known about the structure of $\mathcal{A}_\lambda$; see [McDuff 2000]. For example, one can establish that there is a direct inclusion $\mathcal{A}_\lambda \subset \mathcal{A}_{\lambda'}$, for $\lambda < \lambda'$. Moreover, the homotopy type of the spaces $\mathcal{A}_\lambda$ changes only as λ strictly passes an integer ℓ.

None of this is known to hold when M is an arbitrary symplectic manifold. Nevertheless, as a consequence of the fact that taming is an open condition, we are able to establish the following lemma, which we use in the proof of Theorem 2.9.

Lemma 4.1. (a) *Let K' to be an arbitrary compact subset of $\mathcal{A}_\lambda$. There is an $\epsilon_{K'} > 0$ such that K' is contained in $\mathcal{A}_{\lambda+\epsilon}$, for $|\epsilon| < \epsilon_{K'}$.*

(b) *Consider K an arbitrary compact set in G_λ. For $\mathcal{G}$ as in (4), there is an $\epsilon_K > 0$ and a map $h : [-\epsilon_K, \epsilon_K] \times K \to \mathcal{G}_{|L}$ such that the diagram*

$$\begin{tikzcd} {[-\epsilon_K, \epsilon_K] \times K} \arrow[r, "h"] \arrow[d, "\mathrm{pr}_1"'] & \mathcal{G}_{|L} \arrow[d, "\mathrm{pr}_2"] \\ {[-\epsilon_K, \epsilon_K]} \arrow[r, "\mathrm{incl}"] & \mathbb{R} \end{tikzcd} \tag{17}$$

commutes.

For any two such maps h and h' coinciding on $0 \times K$, there exists, for ϵ' small enough, a homotopy $H : [0,1] \times [-\epsilon', \epsilon'] \times K \to \mathcal{G}_{|L}$ between them that satisfies

$$
\begin{tikzcd}
{[0,1] \times [-\epsilon', \epsilon'] \times K} \arrow[r, "H"] \arrow[d, "\mathrm{pr}_1"'] & \mathcal{G}_{|L} \arrow[d, "\mathrm{pr}_2"] \\
{[-\epsilon', \epsilon']} \arrow[r, "\mathrm{incl}"] & \mathbb{R}
\end{tikzcd}
\tag{18}
$$

Proof. Part (i) is an immediate consequence of the openness of the taming condition.

For the proof of (ii), let's first notice that, since the symplectic condition is an open condition, there is a convex open neighborhood U of ω_λ inside the space of 2-forms such that any closed ω' in U is still symplectic.

Moreover since K is compact there is an $\epsilon(K) > 0$ such that, for any $g_k \in K$,

$$g_k^* \omega_{\lambda+\epsilon} \in U \quad \text{for all } \ 0 \le \epsilon < \epsilon(K).$$

This is true because we can produce such an ϵ for an open set around each element $g \in K$ and hence find a global $\epsilon(K)$ by following a standard compactness argument.

We will construct the elements $h(\epsilon, k)$ as follows. For $t \in [0,1]$ the forms

$$\omega^t_{k,\lambda+\epsilon} := t g_k^* \omega_{\lambda+\epsilon} + (1-t)\omega_{\lambda+\epsilon}$$

are symplectic, since both $g_k^* \omega_{\lambda+\epsilon}$ and $\omega_{\lambda+\epsilon}$ are inside the convex set U. Moreover, since $K \subset G_\lambda \subset \mathrm{Diff}_0 M$, any g_k is smoothly isotopic to the identity and hence $[g_k^* \omega_{\lambda+\epsilon}] = [\omega_{\lambda+\epsilon}]$. Therefore the forms $\omega^t_{k,\lambda+\epsilon}$ are cohomologous as we vary t. We now apply Moser's argument for the one-parameter family of symplectic forms $\omega^t_{k,\lambda+\epsilon}$ and obtain a family of diffeomorphisms $\xi_{k,\lambda+\epsilon,t}$ having the property that $\xi^*_{k,\lambda+\epsilon,t} \omega^t_{k,\lambda+\epsilon} = \omega_{\lambda+\epsilon}$. We next define $h(\epsilon, k) := g_k \circ \xi_{k,\lambda+\epsilon,1}$. Then h has the required properties.

For an arbitrary $h : [-\epsilon, \epsilon] \times K$ satisfying (17) we take the homotopy

$$F : [0,1] \times [-\epsilon, \epsilon] \times K \to \mathbb{R} \times \mathrm{Diff}_0 M$$

given by $F(t, \epsilon, k) := (\epsilon, h(t\epsilon, k))$.

This gives a homotopy between h and $h_0 : [-\epsilon, \epsilon] \times K \to \mathbb{R} \times \mathrm{Diff}_0 M$, where $h_0(\epsilon', k) = h(0, k)$. We similarly obtain a homotopy F' between h' and h_0, where h' also satisfies (17). By concatenating one homotopy with the opposite of the other we obtain a homotopy between h and h', which we call $G : [0,1] \times [-\epsilon_1, \epsilon_1] \times K \to \mathbb{R} \times \mathrm{Diff}_0 M$. We set $g_{s,\epsilon,k} := G(s, \epsilon, k)$ and follow the same procedure as before: we restrict to a short interval $[-\epsilon', \epsilon']$ such that, if we define

$$\omega^t_{s,k,\lambda+\epsilon} := t g^*_{s,\epsilon,k} \omega_{\lambda+\epsilon} + (1-t)\omega_{\lambda+\epsilon},$$

these maps are symplectic for all $0 \leq |\epsilon| < \epsilon'$ and $t, s \in [0, 1]$. This is possible because $\omega^t_{s,k,\lambda} = \omega_\lambda$. Again, the diffeomorphisms $g_{s,\epsilon,k}$ are smoothly isotopic to the identity and, as above, we can apply Moser's argument to the isotopic forms $\omega^t_{s,k,\lambda+\epsilon}$, to obtain diffeomorphisms $\xi_{s,k,\lambda+\epsilon,t}$ such that $\xi^*_{s,k,\lambda+\epsilon,t}\omega^t_{s,k,\lambda+\epsilon,} = \omega_{\lambda+\epsilon}$. If we define $H(s, \epsilon, k) := g_{s,\epsilon,k} \circ \xi_{s,k,\lambda+\epsilon,1}$, the map H has the required properties. $\square$

Definition 4.2. Let $\rho : B \to G_\lambda$ be a cycle in G_λ. An *extension* ρ^ϵ of ρ is a smooth family of cycles $\rho^\epsilon : B \to G_{\lambda+\epsilon}$ defined for $|\epsilon| \leq \epsilon_0$ such that $\rho^0 = \rho$ and satisfying (18). Using Lemma 4.1(i) we see that every cycle ρ has an extension.

Observation. Consider two extensions ρ^ϵ_1, where $0 \leq |\epsilon| < \epsilon_1$, and ρ^ϵ_2, where $0 \leq |\epsilon| < \epsilon_2$. By (18) there is an $\epsilon' > 0$ and a homotopy between ρ^ϵ_1 and ρ^ϵ_2 defined for all $0 \leq \epsilon \leq \epsilon'$. Hence any extension provides well defined elements in $\pi_*(G_{\lambda+\epsilon})$ for small values of ϵ. Therefore each $[\rho] \in \pi_*(G_\lambda)$ has an extension $[\rho^\epsilon] \in \pi_*(G_{\lambda+\epsilon})$ whose *germ* at $\epsilon = 0$ is independent of the choices of ρ.

Definition 4.3. We say that a smooth family of elements $[\rho^\epsilon] \in \pi_*(G_{\lambda+\epsilon})$, with $0 < \epsilon < \epsilon_\rho$, is *new* if it is not the extension for $\epsilon > 0$ of any element $[\rho] \in \pi_*(G_\lambda)$.

In the next section we will use the same letter ρ to refer both to cycles as well as to the homotopy class they represent.

Relation between almost complex structures and symplectomorphism groups; proof of Theorem 1.3. We consider the long exact sequence of relative homotopy groups of the pair $(\mathscr{A}_{\ell^+}, \mathscr{A}_\ell)$:

$$\cdots \longrightarrow \pi_k(\mathscr{A}_{\ell^+}) \longrightarrow \pi_k(\mathscr{A}_{\ell^+}, \mathscr{A}_\ell) \longrightarrow \pi_{k-1}(\mathscr{A}_\ell) \longrightarrow \pi_{k-1}(\mathscr{A}_{\ell^+}) \longrightarrow \cdots$$

Since by construction $\beta_\ell \in \pi_k(\mathscr{A}_{\ell^+}, \mathscr{A}_\ell)$ is nontrivial, one of the two following cases can happen:

1. $\beta_\ell \mapsto \gamma_\ell \neq 0 \in \pi_{k-1}(\mathscr{A}_\ell)$.
2. $\beta_\ell \mapsto 0 \in \pi_{k-1}(\mathscr{A}_\ell)$. In this situation, there is a nonzero element $\alpha_\ell \in \pi_k(\mathscr{A}_{\ell^+})$ that maps to β_ℓ.

We analyze each case in turn:

Case 1. Consider the fibration (2), which yields $G_\ell \longrightarrow \mathrm{Diff}_0\, M \longrightarrow \mathscr{A}_\ell$, and then the long exact sequence in homotopy,

$$\cdots \longrightarrow \pi_{k-1}(G_\ell) \longrightarrow \pi_{k-1}(\mathrm{Diff}_0\, M) \longrightarrow$$
$$\longrightarrow \pi_{k-1}(\mathscr{A}_\ell) \longrightarrow \pi_{k-2}(G_\ell) \longrightarrow \pi_{k-2}(\mathrm{Diff}_0\, M) \longrightarrow \cdots$$

Again, there are two possibilities:

(i) $\gamma_\ell \to \theta_\ell \neq 0 \in \pi_{k-2}(G_\ell)$. In this situation, we have a nontrivial element $\theta_\ell \in \pi_{k-2}(G_\ell)$, such that $\theta_\ell \mapsto 0 \in \pi_{k-2}(\mathrm{Diff}_0 M)$. Then we are in case (A) of Theorem 1.3.

This element is *fragile*. For assume it isn't; then θ_ℓ can be extended by $\theta_{\ell+\epsilon}$, which yields nontrivial classes in $\pi_{k-2}(G_{\ell+\epsilon})$. Then $\theta_{\ell+\epsilon} \mapsto 0 \in \pi_{k-2}(\mathrm{Diff}_0 M)$ as well. Therefore $\theta_{\ell+\epsilon}$ appears as a boundary of an element $\gamma_{\ell+\epsilon} \in \pi_{k-1}(\mathcal{A}_{\ell+\epsilon})$, which is homotopic to γ_ℓ. But by construction and Lemma 4.1, we know that γ_ℓ is a contractible cycle inside $\mathcal{A}_{\ell+\epsilon}$. This contradicts the existence of $\gamma_{\ell+\epsilon}$.

(ii) $\gamma_\ell \mapsto 0 \in \pi_{k-2}(G_\ell)$. Then γ_ℓ is in the image of the morphism $\pi_{k-1}(\mathrm{Diff}_0 M) \to \pi_{k-1}(\mathcal{A}_\ell)$, so there is an element $\gamma'_\ell \in \pi_{k-1}(\mathrm{Diff}_0 M)$ such that $0 \neq \gamma'_\ell \mapsto \gamma_\ell$.

In this situation, we can choose a cycle $S \subset \mathcal{A}_\ell$ representing $\gamma_\ell \in \pi_{k-1}(\mathcal{A}_\ell)$, and, using Lemma 4.1, there is an $\epsilon_S > 0$ such that $S \subset \mathcal{A}_{\ell+\epsilon}$ for any $\epsilon \in (0, \epsilon_S)$. Now we claim that

$$0 = [S] \in \pi_{k-1}(\mathcal{A}_{\ell+\epsilon}).$$

For, by hypothesis, S is the boundary of a cycle B_ℓ such that $B_\ell \subset \mathcal{A}_{\ell+\epsilon}$ for all small $\epsilon > 0$. Therefore we have a k-dimensional ball inside $\mathcal{A}_{\ell+\epsilon}$ whose boundary is S, which proves the claim. We therefore have $\pi_{k-1}(\mathrm{Diff}_0 M) \ni \gamma'_\ell \mapsto [S] = 0 \in \pi_{k-1}(\mathcal{A}_{\ell+\epsilon})$ on the top row of the commutative diagram

$$\begin{tikzcd}
\cdots \arrow[r] & \pi_{k-1}(G_{\ell+\epsilon}) \arrow[r] & \pi_{k-1}(\mathrm{Diff}_0 M) \arrow[r] & \pi_{k-1}(\mathcal{A}_{\ell+\epsilon}) \arrow[r] & \pi_{k-2}(G_{\ell+\epsilon}) \arrow[r] & {} \\
\cdots \arrow[r] & \pi_{k-1}(G_\ell) \arrow[r] & \pi_{k-1}(\mathrm{Diff}_0 M) \arrow[r] \arrow[u, equal] & \pi_{k-1}(\mathcal{A}_\ell) \arrow[r] \arrow[u, dotted, "i_{|k}"'] & \pi_{k-2}(G_\ell) \arrow[r] & {}
\end{tikzcd}$$

while on the bottom row the same γ'_ℓ maps to $\gamma_\ell \in \pi_{k-1}(\mathcal{A}_{\ell+\epsilon})$. By the exactness of the first row, γ'_ℓ is in the image of the map $\pi_{k-1}(G_{\ell+\epsilon}) \to \pi_{k-1}(\mathrm{Diff}_0 M)$, and therefore we are able to produce an element $0 \neq \eta_{\ell+\epsilon} \in \pi_{k-1}(G_{\ell+\epsilon})$ such that $\eta_{\ell+\epsilon}$ persists in the topology of the group of diffeomorphisms. Thus we are in case (B).

The elements we obtain here are *new*. This follows easily by assuming the opposite. That is, if we consider that there is an element $0 \neq \eta_\ell \in \pi_{k-1}(G_\ell)$ whose germ is given by $\eta_{\ell+\epsilon}$, then the image of η_ℓ in $\mathrm{Diff}_0 M$ has to be γ'_ℓ. But this contradicts the fact that $\gamma'_\ell \mapsto \gamma_\ell \neq 0$.

Case 2. In this situation we have a nontrivial element $\alpha_\ell \in \pi_k(\mathcal{A}_{\ell^+})$. Then we shall see that there is an ϵ such that for $0 < \delta < \epsilon$, α_ℓ has a representative C inside $\mathcal{A}_{\ell+\delta}$, with $0 \neq [C] \in \pi_k(\mathcal{A}_{\ell+\delta})$. The proof of this follows from the construction of α_ℓ. Namely, since $\beta_\ell \mapsto 0 \in \pi_{k-1}(\mathcal{A}_\ell)$, there exists a k-dimensional disk D inside $\mathcal{A}_\ell$ whose boundary is ∂B_ℓ; by Lemma 4.1(i), this can be viewed inside $\mathcal{A}_{\ell+\delta}$ for small δ. We can now glue B_ℓ and D along their boundary ∂B_ℓ. In this manner we get a cycle $C \subset \mathcal{A}_{\ell+\delta}$ representing the class α_ℓ. We can therefore consider again the sequence

$$\cdots \longrightarrow \pi_k(G_{\ell+\delta}) \longrightarrow \pi_k(\mathrm{Diff}_0 M) \longrightarrow \\ \longrightarrow \pi_k(\mathcal{A}_{\ell+\delta}) \longrightarrow \pi_{k-1}(G_{\ell+\delta}) \longrightarrow \pi_{k-1}(\mathrm{Diff}_0 M) \longrightarrow \cdots$$

Next we claim that $[C]$ *doesn't lift to a nontrivial element in* $\pi_k(\mathrm{Diff}_0 M)$. Indeed, there is a map

$$\pi_k(\mathrm{Diff}_0 M) \longrightarrow \pi_k(\mathcal{A}_\lambda) \tag{19}$$

for any λ, and as λ varies these maps vary homotopically in $\mathcal{A}_I$. If C did lift, the map $\pi_k(\mathrm{Diff}_0 M) \to \pi_k(\mathcal{A}_\ell)$ would produce a cycle $[B] \in \mathcal{A}_\ell$, which by means of Lemma 4.1 could be viewed inside all $\mathcal{A}_{\ell+\epsilon}$ for small ϵ and which moreover would be homotopic to C inside $\mathcal{A}_{[\ell,\ell+\epsilon]}$. Therefore $[C]$ would map to $0 \in \pi_k(\mathcal{A}_{\ell^+}, \mathcal{A}_\ell)$, contradicting its definition.

Since $[C]$ cannot be in the image of the map $\pi_k(\mathrm{Diff}_0 M) \to \pi_k(\mathcal{A}_{\ell+\delta})$, we know that $[C]$ must have nonzero image $[C] \mapsto \eta_{\ell+\delta} \neq 0$ in $\pi_{k-1}(G_{\ell+\delta})$. Moreover, from the obvious properties of exact sequences again, $\eta_{\ell+\delta} \to 0$ through the natural inclusion map $\pi_{k-1}(G_{\ell+\delta}) \to \pi_{k-1}(\mathrm{Diff}_0 M)$. That these elements are *new* follows again by assuming the opposite. If they formed the *germ* of an element η_ℓ in $\pi_{k-1}(G_\ell)$, then η_ℓ would also be null-homotopic inside $\mathrm{Diff}_0 M$, so it would come from a class $[C']$ in $\pi_k(\mathcal{A}_\ell)$. Moreover, C' would be homotopic with C inside $\mathcal{A}_{[\ell,\ell+\delta]}$, therefore also in $(\mathcal{A}_{[\ell,\ell+\delta]}, \mathcal{A}_\ell)$, which is false given that C has to yield a nontrivial element in $\pi_k(\mathcal{A}_{[\ell,\ell+\delta]}, \mathcal{A}_\ell)$. Thus we are in case (B) of the theorem.

With this, we have exhausted all the possible cases given by the nontrivial PGW, and the proof of Theorem 1.3 is complete. □

Now to prove Corollary 1.4, consider the manifold $(S^2 \times S^2 \times X, \omega_\lambda \oplus \omega_{\mathrm{arb}})$. As seen in (16), the cycles $(B^\ell, \partial B^\ell)$ satisfy the definition (7), so by Proposition 2.4 they give *persistent* elements in $\pi_{4\ell-2}(\mathcal{A}_{\ell^+}, \mathcal{A}_\ell)$. Therefore Theorem 1.3 applies and the corollary holds.

Appendix: A proof of the criterion of parametric regularity, Theorem 2.9

Let $T_{|_{\pi^{-1}(0)}}\widetilde{M}$ be the tangent space along the preimage of $0 \in \mathbb{C}^m$. Denote by H the subbundle of $T_{|_{\pi^{-1}(0)}}\widetilde{M}$ which is $\widetilde{\omega}$-orthogonal to the fiber $\{0\} \times M$. We would like H to coincide with the horizontal space of $T\widetilde{M}$ with respect to the trivialization π and to be $\tilde{J}$-invariant. This can be arranged by deforming the form $\widetilde{\omega}$ so that near the zero fiber $\{0\} \times M$ it is given by

$$\widetilde{\omega} = \omega_0 + \pi^*(\sigma_{\mathrm{base}}),$$

where σ_{base} is a standard symplectic two-form on the holomorphic base B. Throughout this deformation process $\tilde{J}$ is still $\widetilde{\omega}$-tamed.

Let g_0 be a metric on M_0 and ∇ the Levi-Civita connection on M associated with it. Also let ∇^{st} be the standard Levi-Civita connection on $\mathbb{C}^m$, and set $\widetilde{\nabla} = \nabla \times \nabla^{st}$, the product connection on $\widetilde{M} \simeq \mathbb{C}^m \times M$. The regularity of $\tilde{f} : \Sigma \to \widetilde{M}$ is by definition equivalent to the surjectivity of $D_{\tilde{f}}$, the linearization of $\bar{\partial}$:

$$D_{\tilde{f}} : C^\infty(\tilde{f}^* T\widetilde{M}) \twoheadrightarrow \Omega^{0,1}_{\tilde{J}}(\Sigma, \tilde{f}^* T\widetilde{M}).$$

Using the connection $\widetilde{\nabla}$ we will derive formulas for $D_{\tilde{f}}$ and express them in terms of the linearization $D\Phi$.

Since $\widetilde{M} \simeq \mathbb{C}^m \times M$ and $\operatorname{im} \tilde{f} \subset \{0\} \times M$, we have the relations

$$\tilde{f}^*(T\widetilde{M}) = \tilde{f}^*\big(T\widetilde{M}_{\pi^{-1}(0)}\big) = \tilde{f}^*(H \oplus TM) = \text{triv} \oplus f^*(TM),$$

where by triv we denote the trivial m-dimensional complex bundle over Σ. This gives

$$C^\infty(\tilde{f}^* T\widetilde{M}) \simeq C^\infty(\text{triv}) \oplus C^\infty(f^* TM) \tag{20}$$

Since each fiber is $\tilde{J}$-invariant and H is $\tilde{J}$-invariant along $\pi^{-1}(0)$, we obtain

$$\Omega^{0,1}_{\tilde{J}}(\Sigma, \tilde{f}^* T\widetilde{M}) \simeq \Omega^{0,1}_{J}\big(\Sigma, f^* TM\big) \oplus \Omega^{0,1}_{\tilde{J}}(\Sigma, H). \tag{21}$$

From (20) and (21) we obtain

$$D_{\tilde{f}} : C^\infty(\text{triv}) \oplus C^\infty(f^* TM) \twoheadrightarrow \Omega^{0,1}_{J}\big(\Sigma, f^* TM\big) \oplus \Omega^{0,1}_{\tilde{J}}(\Sigma, H),$$

and by considering the appropriate restrictions we obtain the operators

$$D_{1,\text{vert}} : C^\infty(\text{triv}) \to \Omega^{0,1}_{J}(\Sigma, f^* TM), \quad D_{2,\text{vert}} : C^\infty(f^* TM) \to \Omega^{0,1}_{J}(\Sigma, f^* TM),$$
$$D_{1,\text{hor}} : C^\infty(\text{triv}) \to \Omega^{0,1}_{\tilde{J}}(\Sigma, H), \qquad D_{2,\text{hor}} : C^\infty(f^* TM) \to \Omega^{0,1}_{\tilde{J}}(\Sigma, H).$$

We sometimes write $D_k = (D_{k,\text{vert}}, D_{k,\text{hor}})$, for $k = 1, 2$.

To compute the formulas for these operators we use a general method found in [Aebischer et al. 1994]: Consider $\xi \in C^\infty(\Sigma, \tilde{f}^* T\widetilde{M})$ and $\widetilde{F}_\xi : [0, 1] \times \Sigma \to \widetilde{M}$ given by $\widetilde{F}_\xi(t, x) = \exp^{\widetilde{\nabla}}_{\tilde{f}(x)}(t\xi(x))$, for ξ sufficiently small. Let $s : \Sigma \to T\Sigma$ be a section and let $\tilde{s}$ be its lift to $T([0, 1] \times \Sigma)$. Denote by $\partial/\partial t$ the vector field in $T([0, 1] \times \Sigma)$ corresponding to the parameter in $[0, 1]$. Define $\tilde{f}_t(x) := \widetilde{F}_\xi(t, x)$. For any $x \in \Sigma$, define the path $\tilde{\gamma}^\xi_x : [0, 1] \to \widetilde{M}$ given by $\tilde{\gamma}^\xi_x(t) = \widetilde{F}_\xi(t, x)$, the image under $\widetilde{F}_\xi$ of $[0, 1] \times x$ in $\widetilde{M}$. By the definition of $\widetilde{F}_\xi$, $\tilde{\gamma}^\xi_x$ is a geodesic path in $\widetilde{M}$ relative to the connection $\widetilde{\nabla}$. Denote by $\tau^\xi_{t,x} : T_{\gamma_x(t)}\widetilde{M} \to T_{\gamma_x(0)}\widetilde{M}$ the parallel transport in $\widetilde{M}$ along the curve $\gamma_x := \tilde{\gamma}^\xi_x$. To compute $D_{\tilde{f}}(\xi)(s)$ in general, one needs to consider the expression $\frac{1}{2}\tau^{\xi}_{t,x}\big(d\tilde{f}_t(s)+\tilde{J}d\tilde{f}_t(js)\big)$ and take its derivative with respect to t at $t=0$:

$$D_{\tilde{f}}(\xi)(s)=\frac{1}{2}\frac{\partial}{\partial t}\big(\tau^{\xi}_{t,x}(d\tilde{f}_t(s)+\tilde{J}d\tilde{f}_t(js))\big)_{|t=0} \tag{22}$$

We define Const to be the subspace of $C^\infty(\mathrm{triv})$ made out of constant sections. For the proof of the theorem, we are particularly interested in computing $D_{1,\mathrm{hor}}$ and the restriction of $D_{1,\mathrm{vert}}$ to Const.

To simplify the notation, we denote by x the coordinate on Σ and write the points in $\mathbb{C}^m\times M$ as $(z_1,\dots,z_m,y)$, where $z_1=w_1+iv_1$ and so on. For simplicity we denote coordinate vector fields in Const by $\partial_{w_k}:=\partial/\partial w_k$ and so on. Since we are going to work with an arbitrary choice of w_k and v_k we will refer to them simply as ∂_w, unless we need to be more specific.

Lemma 4.4. *Let the notation be as above.*

(i) $D_{2,\mathrm{hor}}=0$.

(ii) $D_{2,\mathrm{vert}}=D_f$.

(iii) $D_{1,\mathrm{hor}}(\xi)=\bar{\partial}_{\mathbb{C}^m}(\xi)$ *for all* $\xi\in C^\infty(\mathrm{triv})$, *where* $\bar{\partial}_{\mathbb{C}^m}$ *is the delbar operator in* $\mathbb{C}^m$.

(iv) $(D_{1,\mathrm{vert}})(\partial_z)(s)=\frac{1}{2}(\partial/\partial z)(J(z))_{|z=0}(df(js))$ *for* ∂_z *a coordinate vector field in* Const $\subset C^\infty(\mathrm{triv})$.

Proof. Since $\tilde{f}=f\circ i\subset\{0\}\times M$ we can naturally view any $\xi\in C^\infty(f^*TM)$ as an element in $C^\infty(\tilde{f}^*T\widetilde{M})$ with values in the vertical direction tangent to $\{0\}\times M$. We have

$$\widetilde{F}_\xi(t,x)=\exp^{\widetilde{\nabla}}_{\tilde{f}(x)}(t\xi)=\exp^{\nabla}_{f(x)}(t\xi),$$

with $\operatorname{im}\widetilde{F}\subset\{0\}\times M$. This implies that the $d\tilde{f}_t(s)$ are also vertical vector fields supported in $\{0\}\times M$ and, since $\tilde{J}$ keeps $T(\{0\}\times M)$ invariant, we have as well that the $\tilde{J}d\tilde{f}_t(js)$ are vertical vector fields in $\{0\}\times M$. Similarly, $\widetilde{F}^*_\xi(\partial/\partial t)$ is a vertical section in $T\widetilde{M}$ supported in $\{0\}\times M$ and parallel transport along $\tilde{f}(x)$ with respect to $\widetilde{\nabla}$ is the same as parallel transport with respect to ∇.

A direct application of (22) is that

$$(D_{\tilde{f}}\xi)(s)=\frac{1}{2}\frac{\partial}{\partial t}\big(\tau^{\xi}_{t,x}d\tilde{f}_t(s)+\tau^{\xi}_{t,x}\tilde{J}df_t(js)\big)_{|t=0}=(D_f\xi)(s),$$

which proves (i). Relation (ii) follows immediately from the formula above, taking into account that $D_{\tilde{f}}\xi=D_{2,\mathrm{vert}}(\xi)$ and that $\operatorname{im}D_{\tilde{f}}|_{C^\infty f^*TM}\subset\Omega^{0,1}_J(\Sigma,f^*TM)$.

For the proofs of (iii) and (iv) we now consider $\xi\in C^\infty(\mathrm{triv})$. We can assume

$\xi = \phi(x)\,\partial_w$, where $\phi : \Sigma \to \mathbb{C}^m$. In this situation, $\widetilde{F}_\xi(t,x) = \exp^{\tilde{\widetilde{\nabla}}}_{f(x)}(t\partial_w) = (\phi(x)t, 0, \dots, 0, f(x))$. Thus the paths γ_x are straight lines in $\mathbb{C}^n \times f(x) \subset \widetilde{M}$ and parallel transport $\tau_{t,x} : T_{(t,f(x))}\widetilde{M} \to T_{0,f(x))}\widetilde{M}$ along γ_x is the identity. We are also going to consider the coordinates $x \in \Sigma$ of the type $x = x_1 + ix_2$, and do our computations for $s = \partial_{x_1}$.

If $\tilde{J}(t)$ is the almost complex structure at $\tilde{\gamma}^\xi_x(t)$, then $\tilde{J}(t)$ has the form

$$\begin{pmatrix} A_t & 0 \\ B_t & J_t \end{pmatrix}$$

with respect to the product structure $\mathbb{C}^m \times M$. Moreover along $\pi^{-1}(0)$ we have

$$\tilde{J}(0) = \begin{pmatrix} J_{\mathbb{C}^m} & 0 \\ 0 & J_t \end{pmatrix}.$$

Therefore $(\partial/\partial t)\tilde{J}(t)$ preserves the fibers, as does $\tilde{J}(t)$. Moreover, along $\{0\} \times M$, $\tilde{J}(0)$ preserves the splitting into TM and H. As we have seen, parallel transport along $\tilde{\gamma}^\xi_x(t)$ is the identity.

Considering local coordinates $x = x_1 + ix_2$ on Σ and taking $s = \partial_{x_1}$, we have

$$\begin{aligned}
D_{1,\mathrm{hor}}(\phi\partial_w)(\partial_{x_1}) &= \tfrac{1}{2}\operatorname{proj}_H \frac{\partial}{\partial t}\bigl(\tau^\xi_{t,x} d\tilde{f}_t(\partial_{x_1}) + \tfrac{1}{2}\tau^\xi_{t,x}\tilde{J} d\tilde{f}_t(j\partial_{x_1})\bigr)_{|t=0} \\
&= \tfrac{1}{2}\operatorname{proj}_H \frac{\partial}{\partial t}\bigl(d\tilde{f}_t(\partial_{x_1}) + \tfrac{1}{2}\tilde{J} d\tilde{f}_t(\partial_{x_2})\bigr)_{|t=0} \\
&= \frac{1}{2}\frac{\partial}{\partial t}\bigl(\partial_{x_1}(\phi(x))t, 0, \dots, 0\bigr)_{|t=0} + \tfrac{1}{2}\operatorname{proj}_H \frac{\partial}{\partial t}(\tilde{J}_t)_{|t=0} df(\partial_{x_2}) \\
&\qquad + \tfrac{1}{2}\operatorname{proj}_H \tilde{J}_0 \frac{\partial}{\partial t}\bigl(\partial_{x_2}(\phi(x))t, 0, \dots, 0, df(x)\bigr)_{|t=0},
\end{aligned}$$

where, as mentioned before, $\phi : \Sigma \to \mathbb{C}^m$. But the middle term on the right-hand side vanishes because $df(\partial_{x_2})$ is a vertical vector and $\partial/\partial t\tilde{J}$ preserves fibers, so $(\partial/\partial t)(\tilde{J}_t)_{|t=0} df(\partial_{x_2})$ is also a vertical vector. Then

$$D_{1,\mathrm{hor}}(\phi\partial_w)(\partial_{x_1}) = \tfrac{1}{2}\partial_{x_1}\phi(x) + \tfrac{1}{2}J_{\mathbb{C}^m}(\partial_{x_2})\phi(x) \tag{23}$$

For the last expression we have to use that along $\pi^{-1}(0)$, $\tilde{J}_0$ preserves the horizontal space H, so $\operatorname{proj}_H \circ \tilde{J}_0 = \tilde{J}_{\mathbb{C}^m} \circ \operatorname{proj}_H$. Therefore, the conclusion follows that $D_{1,\mathrm{hor}} = \bar{\partial}_{\mathbb{C}^m}$.

To prove point (iv) of the theorem we now need to consider $\xi = \partial_w \in \mathrm{Const}$. Under this assumption we have $\tau^{\partial_w}_{t,x} d\tilde{f}_t = df_0$. Thus

$$\frac{\partial}{\partial t}\tau^{\partial_w}_{t,x} d\tilde{f}_t(s) = 0.$$

As before, s is a just a section in $T\Sigma$. Then

$$\begin{aligned}
D_{1,\mathrm{vert}}(\partial_w)(s) &= \tfrac{1}{2}\operatorname{proj}_V \frac{\partial}{\partial t}\big(\tau_{t,x}^{\partial w} d\tilde f_t(s) + \tfrac{1}{2}\tau_{t,x}^{\partial w}\tilde J d\tilde f_t(js)\big)_{|t=0} \\
&= \tfrac{1}{2}\operatorname{proj}_V \frac{\partial}{\partial t}\big(\tau_{t,x}^{\partial w} d\tilde f_t(s)\big)_{|t=0} + \tfrac{1}{2}\operatorname{proj}_V \frac{\partial}{\partial t}\big(\tau_{t,x}^{\partial w}\tilde J(\tau_{t,x}^{\partial w})^{-1}\big)_{|t=0}\cdot df(js) \\
&\qquad\qquad + \tfrac{1}{2}\operatorname{proj}_V \tilde J_0\Big(\frac{\partial}{\partial t}\tau_{t,x}^{\partial w} d\tilde f_t(js)\Big)_{|t=0} \\
&= \tfrac{1}{2}\operatorname{proj}_V(\widetilde\nabla_{\partial_w}\tilde J)df(js),
\end{aligned}$$

where we denote by proj_V the projection onto the fibers. Recall that $(\partial/\partial t)\tilde J$ takes vertical vector fields into vertical vector fields. Therefore

$$\tfrac{1}{2}\operatorname{proj}_V \widetilde\nabla_{\partial_w}\tilde J df(js) = \frac{1}{2}\frac{\partial J(z)}{\partial w}(df(js)),$$

precisely because $df(js)$ is a vertical vector field and the covariant derivative along horizontal vector fields was chosen to be the standard connection in $\mathbb{C}^m$. Applying the same reasoning to $i\partial_v$, we see that

$$(D_{1,\mathrm{vert}})(\partial_z)(s) = \frac{1}{2}\frac{\partial}{\partial z}(J(z))_{|z=0}(df(js)).$$

It is worth pointing out that $(\partial/\partial z)(J(z))_{|z=0} = d\psi_0^*(\partial/\partial z)$. $\square$

Proof of Theorem 2.9. Direct implication: Using Lemma 4.4(v) we get the commutativity of the diagram

$$\begin{tikzcd}
T_0\mathbb{C}^m \arrow[r, "d\psi"] \arrow[d, "i"'] & T_J\mathcal{A}_I \arrow[d, "R"] \\
\mathrm{Const} \arrow[r, "D_{1,\mathrm{vert}}"] & \Omega_J^{0,1}(\Sigma, f^*TM),
\end{tikzcd} \tag{24}$$

where $i : T_0\mathbb{C}^n \to \mathrm{Const} \subset C^\infty(\mathrm{triv})$ is the natural identification map and ψ is the morphism from the parameter space to the space of almost complex structures. R is, as mentioned before, given by $R(Y) = \frac{1}{2}Y \circ df \circ j$.

Since $D_{\tilde f}$ is surjective by hypothesis, this means that $D_1 \oplus D_2$ is surjective. We therefore conclude, by Lemma 4.4(i,ii), that

$$D_1 = (D_{1,\mathrm{vert}}, D_{1,\mathrm{hor}}) : C^\infty(\mathrm{triv}) \longrightarrow \operatorname{coker} D_f \oplus \Omega_{\tilde J}^{0,1}(\Sigma, H) \tag{25}$$

is surjective. Since the kernel of the $\bar\partial_{\mathbb{C}^m}$ operator on $\mathbb{C}^m$ consists precisely of constant sections, Lemma 4.4(iii) implies that $D_{1,\mathrm{hor}}^{-1}(0) =$Const. Therefore the operator $(D_{1,\mathrm{vert}})_{|\mathrm{Const}} : \mathrm{Const} \to \operatorname{coker} D_f$ is surjective. But this will imply that

$$(D_{1,\mathrm{vert}})_{|\mathrm{Const}} \circ i : T_0\mathbb{C}^m \to \operatorname{coker} D_f$$

is surjective.

As we saw in the proof of Proposition 2.8, R induces an isomorphism

$$\tilde{R} : \widetilde{\operatorname{coker} d\Pi} \longrightarrow \operatorname{coker} D_2,$$

and moreover the diagram (24) will be still commutative if we restrict $d\psi$ and $D_{1,\mathrm{vert}}$ to $\operatorname{coker} d\Pi$ and $\operatorname{coker} D_2$ respectively. Therefore $d\psi : T_0\mathbb{C}^n \to \operatorname{coker} d\Pi$ is surjective. By Proposition 2.8, this yields parametric regularity.

For the inverse implication, notice that $D_{1,\mathrm{hor}}$ will cover the space $\Omega^{0,1}_{\tilde{J}}(\Sigma, H)$ when $\Sigma = S^2$, because $D_{1,\mathrm{hor}} = \bar{\partial}_{\mathbb{C}^m}$ in this case. By hypothesis, $d\psi : T_0\mathbb{C}^n \to \operatorname{coker} d\Pi$ is surjective and the preceding observation implies that

$$D_1 = (D_{1,\mathrm{vert}}, D_{1,\mathrm{hor}}) : C^\infty(\mathrm{triv}) \longrightarrow \operatorname{coker} D_f \oplus \Omega^{0,1}_{\tilde{J}}(\Sigma, H)$$

is also surjective. Therefore $D_{\tilde{f}}$ is a surjective operator. □

Acknowledgments

This work is part of the author's doctoral research at SUNY Stony Brook. The author thanks her advisor, Dusa McDuff, for her suggestions, advice and comments on earlier drafts.

References

[Abreu 1998] M. Abreu, "Topology of symplectomorphism groups of $S^2 \times S^2$", *Invent. Math.* **131**:1 (1998), 1–23. MR 99k:57065 Zbl 0902.53025

[Abreu and McDuff 2000] M. Abreu and D. McDuff, "Topology of symplectomorphism groups of rational ruled surfaces", *J. Amer. Math. Soc.* **13**:4 (2000), 971–1009. MR 2001k:57035 ZBL 0965.57031

[Aebischer et al. 1994] B. Aebischer, M. Borer, M. Kälin, C. Leuenberger, and H. M. Reimann, *Symplectic geometry*, Progress in Mathematics **124**, Birkhäuser, Basel, 1994. MR 96a:58082 Zbl 0932.53002

[Bryan and Leung 2000] J. Bryan and N. C. Leung, "The enumerative geometry of $K3$ surfaces and modular forms", *J. Amer. Math. Soc.* **13**:2 (2000), 371–410. MR 2001i:14071 Zbl 0963.14031

[Fukaya and Ono 1999] K. Fukaya and K. Ono, "Arnold conjecture and Gromov–Witten invariant", *Topology* **38**:5 (1999), 933–1048. MR 2000j:53116 Zbl 0946.53047

[Kronheimer 1998] P. Kronheimer, "Some non-trivial families of symplectic structures", preprint, Harvard University, 1998, Available at http://www.math.harvard.edu/~kronheim/diffsymp.pdf.

[Lê and Ono 2001] H. V. Lê and K. Ono, "Parameterized Gromov-Witten invariant and topology of symplectomorphism groups", preprint 28/2001, Max Planck Institute for Mathematics in the Sciences, 2001, Available at http://www.mis.mpg.de/preprints/2001/prepr2801-abstr.html.

[Li and Tian 1998] J. Li and G. Tian, "Virtual moduli cycles and Gromov–Witten invariants of general symplectic manifolds", pp. 47–83 in *Topics in symplectic 4-manifolds* (Irvine, CA, 1996), edited by R. J. Stern, Internat. Press, Cambridge, MA, 1998. MR 2000d:53137 Zbl 0978.53136

[McDuff 2000] D. McDuff, "Almost complex structures on $S^2 \times S^2$", *Duke Math. J.* **101**:1 (2000), 135–177. MR 2001h:53126 Zbl 0974.53020

[McDuff 2001] D. McDuff, "Symplectomorphism groups and almost complex structures", pp. 527–556 in *Essays on geometry and related topics*, vol. 2, Monogr. Enseign. Math. **38**, Enseignement Math., Geneva, 2001. MR 2003i:57042 Zbl 1010.53064

[McDuff and Salamon 1994] D. McDuff and D. Salamon, *J-holomorphic curves and quantum cohomology*, vol. 6, University Lecture Series, American Mathematical Society, Providence, RI, 1994. MR 95g:58026 Zbl 0809.53002

[Ruan 1999] Y. Ruan, "Virtual neighborhoods and pseudo-holomorphic curves", *Turkish J. Math.* **23**:1 (1999), 161–231. MR 2002b:53138 Zbl 0967.53055

[Seidel 1997] P. Seidel, *Floer homology and the symplectic isotopy problem*, Ph.D. thesis, Oxford University, 1997.

[Seidel 1999] P. Seidel, "On the group of symplectic automorphisms of $\mathbf{CP^m} \times \mathbf{CP^n}$", pp. 237–250 in *Northern California Symplectic Geometry Seminar*, edited by Y. Eliashberg et al., Amer. Math. Soc. Transl. Ser. 2 **196**, Amer. Math. Soc., Providence, RI, 1999. MR 2000m:53124 Zbl 0954.58009

[Smale 1965] S. Smale, "An infinite dimensional version of Sard's theorem", *Amer. J. Math.* **87** (1965), 861–866. MR 32 #3067 Zbl 0143.35301

Received January 29, 2003.

OLGUŢA BUŞE
DEPARTMENT OF MATHEMATICS
A-320 WELLS HALL
MICHIGAN STATE UNIVERSITY
EAST LANSING, MI 48824
UNITED STATES
buse@math.msu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

A FAMILY OF ISOCHRONOUS FOCI WITH DARBOUX FIRST INTEGRAL

JAUME GINÉ AND JAUME LLIBRE

We consider the class of polynomial differential equations $\dot{x} = \lambda x - y + P_n(x, y) + P_{2n-1}(x, y)$, $\dot{y} = x + \lambda y + Q_n(x, y) + Q_{2n-1}(x, y)$ with $n \geq 2$, where P_i and Q_i are homogeneous polynomials of degree i. These systems have a focus at the origin if $\lambda \neq 0$, and have either a center or a focus if $\lambda = 0$. Inside this class we identify a new subclass of Darboux integrable systems having either a focus or a center at the origin. Under generic conditions such Darboux integrable systems can have at most two limit cycles, and when they exist are algebraic. For the case $n = 2$ and $n = 3$ we present new classes of Darboux integrable systems having a focus.

1. Introduction and statement of the results

Three of the main problems in the qualitative theory of real planar differential systems are the determination of centers, limit cycles and first integrals. This paper deals mainly with the determination of first integrals and limit cycles.

As usual a *center* is a singular point having a neighborhood filled of periodic orbits, and a *focus* is a singular point having a neighborhood where all the orbits spiral in forward or in backward time to it.

Here we study real planar polynomial differential systems of the form

$$\begin{aligned} \dot{x} &= \lambda x - y + P_n(x, y) + P_{2n-1}(x, y), \\ \dot{y} &= x + \lambda y + Q_n(x, y) + Q_{2n-1}(x, y), \end{aligned} \tag{1}$$

where P_i and Q_i are homogeneous polynomials of degree i. Inside this class we will characterize a new subclass of Darboux integrable systems having either a focus or a center at the origin.

Giné is partially supported by MCYT grant BFM 2002-04236-C02-01, by a University of Lleida Project P01 and by DURSI of Government of Catalonia's Acció Integrada ACI2001-26. Llibre is partially supported by DGICYT grant BFM 2002-04236-C02-02 and by CICYT grant 2001SGR00173.
MSC2000: 34C35, 34D30.
Keywords: integrability, algebraic limit cycle, focus, center.

We establish some notation and preliminary results. In polar coordinates (r, θ), defined by

$$x = r\cos\theta, \qquad y = r\sin\theta, \tag{2}$$

system (1) becomes

$$\begin{aligned} \dot r &= \lambda r + f_{n+1}(\theta)r^n + f_{2n}(\theta)r^{2n-1},\\ \dot\theta &= 1 + g_{n+1}(\theta)r^{n-1} + g_{2n}(\theta)r^{2n-2}, \end{aligned} \tag{3}$$

where

$$\begin{aligned} f_i(\theta) &= \cos\theta\, P_{i-1}(\cos\theta, \sin\theta) + \sin\theta\, Q_{i-1}(\cos\theta, \sin\theta),\\ g_i(\theta) &= \cos\theta\, Q_{i-1}(\cos\theta, \sin\theta) - \sin\theta\, P_{i-1}(\cos\theta, \sin\theta) \end{aligned}$$

are certain homogeneous trigonometric polynomials in the variables $\cos\theta$ and $\sin\theta$ having degree in the set $\{i, i-2, i-4, \dots\}$. Indeed, $f_i(\theta)$ can be of the form $(\cos^2\theta + \sin^2\theta)^s f_{i-2s}$ with f_{i-2s} a trigonometric polynomial of degree $i - 2s \geq 0$, and a similar situation occurs for $g_i(\theta)$.

If we impose $g_{n+1}(\theta) = g_{2n}(\theta) = 0$ and make the change $R = r^{n-1}$, system (3) becomes the differential equation

$$\frac{dR}{d\theta} = (n-1)\big(\lambda R + f_{n+1}(\theta)R^2 + f_{2n}(\theta)R^3\big). \tag{4}$$

Differential equations of this kind appeared in Abel's studies on the theory of elliptic functions. For more details on Abel differential equations, see [Kamke 1943; Cheb-Terrab and Roche 2003; Gasull and Llibre 1990].

We say that all polynomial differential systems (1) with $g_{n+1}(\theta) = g_{2n}(\theta) = 0$ define *class* $\mathcal{F}$ if $f_{2n}(\theta)$ and $f_{n+1}(\theta)$ satisfy

$$f_{2n}'(\theta)f_{n+1}(\theta) - f_{2n}(\theta)f_{n+1}'(\theta) = (n-1)\big(af_{n+1}(\theta)^3 - \lambda f_{n+1}(\theta)f_{2n}(\theta)\big) \tag{5}$$

for some $a \in \mathbb{R}$. Clearly, the class of our polynomial differential systems (1) has dimension $6n+4$ in the space of all coefficients, and the subclass $\mathcal{F}$ is an algebraic subvariety of it.

We shall prove that all polynomial differential systems (1) in class $\mathcal{F}$ have a Darboux first integral. We have found the subclass $\mathcal{F}$ thanks to the Abel differential equations studied in [Kamke 1943, pp. 24–25, cases (a–d)]. Using these same techniques new Darboux integrable systems are found in [Giné and Llibre 2004] for polynomial systems formed by a linear part plus homogeneous nonlinearities.

A function of the form $f_1^{\lambda_1} \dots f_p^{\lambda_p} \exp(h/g)$, where f_i, g and h are polynomials in $\mathbb{C}[x, y]$ and the λ_i's are complex numbers, is called a *Darboux function*. System (1) is called *Darboux integrable* if the system has a first integral or an integrating factor which is a Darboux function (for a definition of a first integral and of an integrating factor, see [Chavarriga et al. 1999; Christopher and Llibre 2000], for instance). Our main result is the following:

Theorem 1. *For polynomial differential systems* (1) *in the class* $\mathcal{F}$ *the following statements hold.*

(a) *If* $\lambda \neq 0$ *and* $f_{2n}(\theta) f_{n+1}(\theta) \neq 0$, *then the origin is a focus and the system has the Darboux first integral* $\tilde{H}(x, y)$ *obtained from*

$$H(R,\theta)=\begin{cases} \dfrac{R\exp((n-1)\lambda\theta)\exp\left(-\frac{1}{\sqrt{4a-1}}\arctan\frac{(1+2Rf_{2n}(\theta)/f_{n+1}(\theta))}{\sqrt{4a-1}}\right)}{\sqrt{R^2 f_{2n}^2(\theta)/f_{n+1}^2(\theta)+Rf_{2n}(\theta)/f_{n+1}(\theta)+a}} & \text{if } a>\frac{1}{4},\\[2ex] \dfrac{R\exp((n-1)\lambda\theta)\exp\frac{1}{1+2Rf_{2n}(\theta)/f_{n+1}(\theta)}}{1+2Rf_{2n}(\theta)/f_{n+1}(\theta)} & \text{if } a=\frac{1}{4},\\[2ex] \dfrac{R\exp((n-1)\lambda\theta)\left(\sqrt{1-4a}+1+\frac{2Rf_{2n}(\theta)}{f_{n+1}(\theta)}\right)^{(-1+1/\sqrt{1-4a})/2}}{\left(\sqrt{1-4a}-1-\frac{2Rf_{2n}(\theta)}{f_{n+1}(\theta)}\right)^{(1+1/\sqrt{1-4a})/2}} & \text{if } a<\frac{1}{4},\ a\neq 0,\\[2ex] \dfrac{\exp((n-1)\lambda\theta)f_{2n}(\theta)}{f_{n+1}(\theta)} & \text{if } a=0,\end{cases}$$

through the changes of variables (2) *and with* $R = r^{n-1}$.

(b) *If* $\lambda \neq 0$ *and* $a = f_{2n}(\theta) f_{n+1}(\theta) = 0$, *then the origin is a focus and the system has the Darboux first integral* $\tilde{H}(x, y)$ *obtained from*

$$H(R,\theta)=\begin{cases} \frac{\exp((n-1)\lambda\theta)}{R}+(n-1)\int\exp((n-1)\lambda\theta)f_{n+1}(\theta)\,d\theta & \text{if } f_{2n}(\theta)=0,\\[1ex] \frac{\exp(2(n-1)\lambda\theta)}{R^2}+2(n-1)\int\exp(2(n-1)\lambda\theta)f_{2n}(\theta)\,d\theta & \text{if } f_{n+1}(\theta)=0,\end{cases}$$

through the changes of variables (2) *and with* $R = r^{n-1}$.

(c) *If* $\lambda = 0$, *the origin is a center, and the system has an analytic first integral* $\tilde{H}(x, y)$ *obtained by taking* $\lambda = 0$ *in the expressions for* $\tilde{H}(x, y)$ *in* (a) *and* (b).

(d) *If* $\lambda = 0$, *the origin is a center, and the following systems have a rational first integral*:

- (d1) *Systems with* $f_{n+1}(\theta) = f_{2n}(\theta) = 0$.
- (d2) *Systems with* $f_{2n}(\theta) = 0$ *and* $\int_0^{2\pi} f_{n+1}(\theta)\,d\theta = 0$.
- (d3) *Systems with* $f_{n+1}(\theta) = 0$ *and* $\int_0^{2\pi} f_{2n}(\theta)\,d\theta = 0$.
- (d4) *Systems whose* a (*defined in* (5)) *satisfies* $a < \frac{1}{4}$, $a \neq 0$, *and* $\sqrt{1-4a}$ *is rational*.

Theorem 1 will be proved in Section 2. Part (c) follows easily from (a) and (b).

A *limit cycle* of system (1) is a periodic orbit isolated in the set of periodic orbits of system (1). We say that a limit cycle γ is *algebraic* if it is contained in an algebraic curve.

Theorem 2. *If a system* (1) *in class* $\mathcal{F}$ *with* $\lambda \neq 0$ *and* $f_{2n}(\theta) f_{n+1}(\theta) \neq 0$ *has a limit cycle*, *it is algebraic*. *Moreover*, *such a system can have at most two limit cycles*. *There are systems with* 0, 1 *or* 2 *limit cycles*.

In the course of the proof, given in Section 3, we provide an explicit expression for algebraic limit cycles.

Systems (1) for $n = 2$ are cubic differential systems. The problem of determining when a cubic differential system (1) has a center at a singular point is open. Trying to distinguish whether a weak focus of a general cubic system is a center or a focus has produced disappointing results, due to the huge expressions obtained for its Poincaré–Liapunov constants; see [Schlomiuk 1993]. But several authors have studied particular subclasses of cubic polynomial differential systems; see for instance [Pearson et al. 1996] and the references therein. The center problem for cubic polynomial differential system (1) satisfying $x Q_3(x, y) - y P_3(x, y) = 0$ has been totally solved in [Chavarriga and Giné 1998; Lloyd et al. 1997].

Other polynomial differential systems recently investigated are those of the form

$$\dot{x} = y + x F(x, y), \quad \dot{y} = -x + y F(x, y), \tag{6}$$

where $F(x, y) = \sum_{i=1}^{4} F_i(x, y)$ for homogeneous polynomials $F_i(x, y)$ of degree i. Such systems satisfy $x Q_i(x, y) - y P_i(x, y) = 0$ for $i = 1, \dots, 4$; therefore, they have constant angular speed $\dot{\theta} = 1$. When a system (6) has a center at the origin, this center is called a *uniformly isochronous center* [Conti 1994]. If $F(x, y) = 0$, the origin is a linear center. The conditions for a system (6) to have a center have been studied in [Collins 1997] when $F_3 = F_4 = 0$. Systems of the form (6) have been studied in [Giné and Santallusia 2001] in the case that $F(x, y)$ is of degree 3 with $F_2 = 0$, and in [Chavarriga et al. 2001] in the case that $F(x, y)$ is of degree 3. The case where $F(x, y)$ is of degree 4 is totally solved in [Volokitin 2002] when $F_1 = F_3 = 0$.

It is easy to check that systems (1) with $n = 2$ satisfying $g_3(\theta) = g_4(\theta) = 0$ can be written into the form

$$\begin{aligned} \dot{x} &= \lambda x - y + x(\alpha x + \beta y + A x^2 + B x y + C y^2), \\ \dot{y} &= x + \lambda y + y(\alpha x + \beta y + A x^2 + B x y + C y^2), \end{aligned} \tag{7}$$

where α, β, A, B and C are arbitrary constants. In [Collins 1997] it has been proved that the origin of system (7) is a center if and only if

$$\lambda = 0, \quad A + C = 0 \quad \text{and} \quad A\alpha^2 + B\alpha\beta + C\beta^2 = 0.$$

In Corollary 5 we compute the class of cubic polynomial differential systems satisfying (a) and (b) in Theorem 1, thereby exhibiting new classes of Darboux integrable cubic systems having a focus.

Polynomial systems (1) with $n=3$ satisfying $g_4(\theta)=g_6(\theta)=0$ can be written to the form

$$\dot{x} = \lambda x - y + x(Ax^2+Bxy+Cy^2+Dx^4+Ex^3y+Fx^2y^2+Gxy^3+Hy^4),$$
$$\dot{y} = x+\lambda y + y(Ax^2+Bxy+Cy^2+Dx^4+Ex^3y+Fx^2y^2+Gxy^3+Hy^4), \tag{8}$$

where A, B, C, D, E, F, G and H are arbitrary constants. Volokitin [2002] has proved that the origin of system (8) is a center if and only if one of the following sets of conditions are satisfied:

(i) $\lambda = 0$, $A=B=C=0$, and $F=-3(D+H)$.

(ii) $\lambda = 0$ and $A=C=D=F=H=0$.

(iii) $\lambda=0$, $A\neq 0$, $C=-A$, $F=3B(AE-BD)/2A^2$,

$$H = \frac{-2A^2D+B(BD-AE)}{2A^2}, \quad G = \frac{2A^2BD+(2A^2-B^2)(BD-AE)}{2A^3}.$$

In Corollary 6 we will provide new classes of Darboux integrable systems (8) having either a focus or a center at the origin.

2. Proof of Theorem 1

Proof of Theorem 1(a). Following [Kamke 1943, p. 25, case (d)], we make the change of variables $(R,\theta)\to(\eta,\xi)$ defined by $R=u(\theta)\eta(\xi)$, where $u(\theta)=\exp((n-1)\lambda\theta)$ and $\xi=\int \exp((n-1)\lambda\theta)\,(n-1)f_{n+1}(\theta)\,d\theta$. This transformation writes the Abel differential equation (4) into the form

$$\eta'(\xi) = g(\xi)\,\eta(\xi)^3+\eta(\xi)^2, \tag{9}$$

where $g(\xi)=\exp((n-1)\lambda\theta)\,f_{2n}(\theta)/f_{n+1}(\theta)$ and $'=d/d\xi$. Making the change $\xi\to t$ in the independent variable defined by $\xi'=-1/(t\eta(\xi))$, where now $'=d/dt$, equation (9) becomes

$$t^2\xi''(t)+g(\xi(t))=0. \tag{10}$$

Note that $g(\xi)=a\xi$ means

$$\exp((n-1)\lambda\theta)f_{2n}(\theta)/f_{n+1}(\theta) = a\int \exp((n-1)\lambda\theta)(n-1)f_{n+1}(\theta)\,d\theta;$$

equivalently, by differentiating with respect to θ, we get

$$\frac{d}{d\theta}\frac{f_{2n}(\theta)}{f_{n+1}(\theta)} = a(n-1)f_{n+1}(\theta) - \frac{(n-1)\lambda f_{2n}(\theta)}{f_{n+1}(\theta)}, \tag{11}$$

which is equivalent to condition (5). Thus $g(\xi)=a\xi$, and (10) is an Euler differential equation. Applying the change $t=\exp(\tau)$ to the independent variable, equation (10) then becomes a linear ordinary differential equation with constant coefficients:

$$\xi''(\tau) - \xi'(\tau) + a\xi(\tau) = 0, \tag{12}$$

where $' = d/d\tau$. Equation (12) has the characteristic equation $k^2 - k + a = 0$, so its general solution is $\xi(\tau) = C_1 \exp(\tau/2) + C_2 \tau \exp(\tau/2)$ if $a = \frac{1}{4}$, and $\xi(\tau) = C_1 \exp(k_1 \tau) + C_2 \exp(k_2 \tau)$ if $a \neq \frac{1}{4}$, where k_1 and k_2 are the roots of the characteristic equation. Going back to the independent variable $t = \exp(\tau)$, the solution of the Euler differential equation is $\xi(t) = C_1 \sqrt{t} + C_2 \sqrt{t} \ln t$ if $a = \frac{1}{4}$ and $\xi(t) = C_1 t^{k_1} + C_2 t^{k_2}$ if $a \neq \frac{1}{4}$.

Finally, going back through the change of variables to the variables (R, θ) and taking into account whether the roots k_1 and k_2 are real or complex, we obtain the first integrals shown in statement (a), according to the value of a.

We now prove that the systems in (a) are Darboux integrable, by showing that all terms that appear in the first integral of those systems are of the form $f_1^{\lambda_1} \ldots f_p^{\lambda_p}$, with f_i a polynomial and λ_i a complex number. First, the term $\exp((1-n)\lambda\theta)$ takes the form

$$\begin{aligned}\exp((1-n)\lambda\theta) &= \exp((1-n)\lambda \arctan(y/x)) \\ &= (x+iy)^{i(n-1)\lambda/2}(x-iy)^{-i(n-1)\lambda/2}.\end{aligned}$$

Recall that if $f = 0$, with $f \in \mathbb{C}[x, y]$, is an invariant algebraic curve of a real polynomial differential system, the complex conjugate $\bar{f} = 0$ is also an invariant algebraic curve; see [Christopher and Llibre 2000], for instance. Therefore, if among the invariant algebraic curves of system (1) there occurs a complex conjugate pair $f = 0$ and $\bar{f} = 0$, the first integral has a factor of the form $f^{\mu} \bar{f}^{\bar{\mu}}$, which is the (multivalued) real function

$$\left((\operatorname{Re} f)^2 + (\operatorname{Im} f)^2\right)^{\operatorname{Re}\mu} \exp\left(-2 \operatorname{Im}\mu \arctan \frac{\operatorname{Im} f}{\operatorname{Re} f}\right).$$

On the other hand, writing $R = r^{n-1}$, it follows that

$$F = \frac{f_{n+1}(\theta) + 2R f_{2n}(\theta)}{f_{n+1}(\theta)} = \frac{r^{n+1}(f_{n+1}(\theta) + 2r^{n-1} f_{2n}(\theta))}{r^{n+1} f_{n+1}(\theta)},$$

is a rational function in cartesian coordinates because $f_{2n}(\theta)$ and $f_{n+1}(\theta)$ are homogeneous trigonometric polynomials of degree $2n$ and $n+1$, respectively. Taking into account these relations, the first integral for $a > \frac{1}{4}$ is the Darboux function

$$H(\rho, \theta) = \rho \exp((1-n)\lambda\theta) f^{\mu} \bar{f}^{\bar{\mu}},$$

where $\operatorname{Re} f = F$, $\operatorname{Im} f = \sqrt{4a-1}$, $\operatorname{Re} \mu = -\frac{1}{2}$, and $\operatorname{Im} \mu = 1/(2\sqrt{4a-1})$. The first integral for $a = \frac{1}{4}$ is the Darboux function

$$H(\rho, \theta) = \rho \exp((1-n)\lambda\theta) \exp(1/F)/F.$$

The first integral for $a < \frac{1}{4}$ and $a \neq 0$ is the Darboux function

$$H(\rho, \theta) = \rho \exp((1-n)\lambda\theta) \left|\sqrt{1-4a} + F\right|^{\mu_1} \left|\sqrt{1-4a} - F\right|^{\mu_2},$$

where $\mu_1 = \frac{1}{2}(-1 + 1/\sqrt{1-4a})$ and $\mu_2 = \frac{1}{2}(1 + 1/\sqrt{1-4a})$. Finally, the first integral for $a = 0$ is the Darboux function

$$H(\rho, \theta) = \frac{\exp((1-n)\lambda\theta) f_{2n}(\theta) r^{2n}}{r^{n-1} f_{n+1}(\theta) r^{n+1}},$$

and this completes the proof of statement (a). □

Proof of Theorem 1(b). In the cases $f_{2n}(\theta) = 0$ and $f_{n+1}(\theta) = 0$, the Abel differential equation (4) is the Bernoulli differential equation $dR/d\theta = (n-1)(f_{n+1}(\theta)R^2 + \lambda R)$ and $dR/d\theta = (n-1)(f_{2n}(\theta)R^3 + \lambda R)$, respectively. Solving these Bernoulli equations we obtain the first integrals of statement (b).

Systems of statement (b) are Darboux integrable because their first integrals are obtained by integrating elementary functions; see [Singer 1992] for more details. The integrals appearing in the first integrals in question can be computed using recurrence formulas; see for instance [Petit Bois 1961, p. 149]. □

Proof of Theorem 1(c). The proof follows easily taking $\lambda = 0$ in statements (a) and (b). □

Proof of Theorem 1(d). If $f_{n+1}(\theta) = f_{2n}(\theta) = 0$, system (3) with $\lambda = 0$ satisfies $\dot{r} = 0$ and therefore it has a polynomial first integral $H = x^2 + y^2$. Statement (d1) follows.

If $f_{2n}(\theta) = 0$, from the Abel differential equation (4) it is easy to derive that $H(R, \theta) = 1/R + (n-1)\int f_{n+1}(\theta)\, d\theta$ is a first integral. Taking into account that $\int_0^{2\pi} f_{n+1}(\theta)\, d\theta$ vanishes and going back to cartesian variables, we obtain a rational first integral and (d2) follows.

If $f_{n+1}(\theta) = 0$, again from the Abel differential equation (4) it is easy to derive that $H(R, \theta) = 1/R^2 + 2(n-1)\int f_{2n}(\theta)\, d\theta$ is a first integral. Taking into account that $\int_0^{2\pi} f_{2n}(\theta)\, d\theta$ vanishes and going back to cartesian variables, we obtain a rational first integral, and (d3) follows.

Finally, from the expression of the first integral $H(R, \theta)$ for $a < \frac{1}{4}$ and $a \neq 0$ with $\sqrt{1-4a}$ rational, we have $H^2(R, \theta) = R^2 \left|\sqrt{1-4a} + F\right|^{2\mu_1} \left|\sqrt{1-4a} - F\right|^{2\mu_2}$, where μ_1 and μ_2 are defined at the end of the proof of part (a). Therefore, a convenient power of $H^2(R, \theta)$ gives a rational first integral. There follows (d4).

□

Now we investigate whether it is possible to find other integrable classes from the well known integrable cases of the Abel differential equation. Following [Kamke 1943, p. 24, case (4)], first we perform the change of variables $(R, \theta) \to (\eta, \xi)$ defined by $R = w(\theta)\eta(\xi) - f_{n+1}(\theta)/(3 f_{2n}(\theta))$, where

$$w(\theta) = \exp\big(\textstyle\int (n-1)(\lambda - f_{n+1}^2(\theta)/(3 f_{2n}(\theta)))\, d\theta\big)$$

and $\xi = \int (n-1) f_{2n}(\theta) w^2(\theta)\, d\theta$. This puts the Abel equation (4) into the normal form

$$\eta'(\xi) = \eta(\xi)^3 + I(\theta), \tag{13}$$

where

$$I(\theta) = \frac{1}{(n-1) f_{2n}(\theta) w^3(\theta)} \left(\frac{d}{d\theta} \frac{f_{n+1}(\theta)}{3 f_{2n}(\theta)} - \frac{(n-1)\lambda f_{n+1}(\theta)}{3 f_{2n}(\theta)} + \frac{2(n-1) f_{n+1}^3(\theta)}{27 f_{2n}^2(\theta)} \right).$$

From the definition of $w(\theta)$ we have

$$\begin{aligned} \ln|w(\theta)| &= (n-1) \int \left(\lambda - \frac{f_{n+1}^2(\theta)}{3 f_{2n}(\theta)} \right) d\theta \\ &= (n-1) \int \frac{f_{n+1}(\theta)}{f_{2n}(\theta)} \left(\frac{\lambda f_{2n}(\theta)}{f_{n+1}(\theta)} - \frac{f_{n+1}(\theta)}{3} \right) d\theta. \end{aligned} \tag{14}$$

In the case $a \neq 0$, the right-hand side of (14) becomes, upon use of (5) (or, equivalently, of (11)),

$$\begin{aligned} -\frac{1}{3a} \int \frac{\frac{d}{d\theta}\big(f_{2n}(\theta)/f_{n+1}(\theta)\big)}{f_{2n}(\theta)/f_{n+1}(\theta)}\, d\theta + \Big(1 - \frac{1}{3a}\Big)(n-1) \int \lambda\, d\theta \\ = -\frac{1}{3a} \ln\left| \frac{f_{2n}(\theta)}{f_{n+1}(\theta)} \right| + \Big(1 - \frac{1}{3a}\Big)(n-1)\lambda\theta. \end{aligned}$$

This leads to $w(\theta) = \big| f_{2n}(\theta)/f_{n+1}(\theta) \big|^{-1/3a} \exp((n-1)(1 - 1/(3a))\lambda\theta)$, so $I(\theta)$ becomes

$$I(\theta) = \Big(\frac{2 - 9a}{27} \Big) \left(\frac{f_{2n}(\theta)}{f_{n+1}(\theta)} \right)^{(1-3a)/a} \exp((n-1)(1-3a)\lambda\theta/a). \tag{15}$$

It is easy to see that $I(\theta) = 0$ for $a = \frac{2}{9}$ and $I(\theta) = -\frac{1}{27}$ for $a = \frac{1}{3}$. In these two cases, we can separate variables in the differential equation (13) and obtain the associated first integrals. But $I(\theta) = 0$ and $I(\theta) = -\frac{1}{27}$ imply that (5) holds with $a = \frac{2}{9}$ and $a = \frac{1}{3}$, respectively. So we obtain cases already studied. New cases of integrability would only appear for $I(\theta) \neq 0, -\frac{1}{27}$.

Cases (b) and (c) of the Abel differential equation of [Kamke 1943, p. 25] again lead to the case already studied, with $a = \frac{2}{9}$.

3. Algebraic limit cycles with Darboux first integral

The next proposition presents what is probably the easiest example of a polynomial differential system that has a Darboux first integral and an algebraic limit cycle. Other examples of this kind were given in [Dolov 1976; Kooij and Christopher 1993; Christopher 1994]. In fact, it has now been proved that any finite configuration of limit cycles is realizable by algebraic limit cycles of a Darboux integrable polynomial differential systems [Llibre and Rodríguez 2004].

Proposition 3 [Chavarriga et al. 1999]. *The differential system*

$$x' = x - y - x(x^2 + y^2), \qquad y' = x + y - y(x^2 + y^2) \tag{16}$$

has the algebraic solution $x^2 + y^2 - 1 = 0$ *as a limit cycle. In polar coordinates* (2) *the function* $H(r, \theta) = (r^2 - 1)\exp(2\theta)/r^2 = C$ *is a Darboux first integral defined on* $\mathbb{R}^2 \setminus \Sigma$, *where* $\Sigma = \{(0, 0)\} \cup \{(x, y) : x^2 + y^2 - 1 = 0\}$.

To study the existence or nonexistence of limit cycles in system (1) we shall use the following result.

Theorem 4 [Giacomini et al. 1996, Theorem 9]. *Let* (P, Q) *be a* C^1 *vector field defined in an open subset* U *of* $\mathbb{R}^2$. *Let* $V = V(x, y)$ *be a* C^1 *solution of the linear partial differential equation*

$$P\frac{\partial V}{\partial x} + Q\frac{\partial V}{\partial y} = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right)V,$$

defined in U. *If* γ *is a limit cycle of* (P, Q), *then* γ *is contained in* $\{(x, y) \in U : V(x, y) = 0\}$.

Under the assumptions of Theorem 4, the function $1/V$ is an integrating factor in $U \setminus \{V(x, y) = 0\}$ (see [Chavarriga et al. 1999; Christopher and Llibre 2000] for details). So the function V is called an *inverse integrating factor*.

Proof of Theorem 2. For systems (1) in class $\mathcal{F}$ with $\lambda \neq 0$ and $f_{2n}(\theta) f_{n+1}(\theta) \neq 0$, it is easy to check that

$$V(\rho, \theta) = R(R^2 f_{2n}^2(\theta)/f_{n+1}^2(\theta) + R f_{2n}(\theta)/f_{n+1}(\theta) + a)$$

is an inverse integrating factor of the associated Abel differential equation (4). Notice that V is defined for all (R, θ) such that $f_{n+1}(\theta) \neq 0$. Again by Theorem 4, if system (1) and consequently its associated Abel equation (4) have limit cycles, those of the Abel equation must be contained in the set $\{V(R, \theta) = 0\}$.

From the expression of the inverse integrating factor, the unique possible limit cycles must be given by

$$R(\theta) = \begin{cases} \frac{1}{2}(-1 \pm \sqrt{1-4a}\,) f_{n+1}(\theta)/f_{2n}(\theta) & \text{if } a < \frac{1}{4}, \\ -\frac{1}{2} f_{n+1}(\theta)/f_{2n}(\theta) & \text{if } a = \frac{1}{4}. \end{cases}$$

For these expressions to define limit cycles, $R(\theta)$ must be defined and positive for all θ. Since $f_{2n}(\theta)$ and $f_{n+1}(\theta)$ are homogeneous trigonometric polynomials of degree $2n$ and $n+1$ respectively, we conclude that n must be odd and $a \le \frac{1}{4}$.

Clearly a system in our class (1) has no limit cycles if n is even, or if $a > \frac{1}{4}$ and $f_{2n}(\theta) f_{n+1}(\theta) \neq 0$. It can have one limit cycle if $a = \frac{1}{4}$: for instance, setting $A = 2$, $B = E = 0$ in Corollary 6(c) yields the system $\dot{x} = -y - x(x^2+y^2-1)^2$, $\dot{y} = x - y(x^2+y^2-1)^2$, which has exactly one limit cycle, the circle $x^2+y^2-1=0$. And the system can have two limit cycles if n is odd and $a < \frac{1}{4}$: setting $A = 4$, $B = E = 0$ in Corollary 6(c) yields the system $\dot{x} = x - y - x(x^2+y^2-2)^2$, $\dot{y} = x + y - y(x^2+y^2-2)^2$, which has exactly two limit cycles given, $x^2+y^2-3=0$ and $x^2+y^2-1=0$. This completes the proof. □

4. Some corollaries

System (1) with $n = 2$ and $g_3(\theta) = g_4(\theta) = 0$ — i.e., the cubic system (7) — has a focus or a center at the origin. The following corollary characterizes cubic polynomial systems (7) belonging to class $\mathcal{F}$.

Corollary 5. *A cubic system* (7) *with* $\lambda \neq 0$ *belongs to class* $\mathcal{F}$ *if and only if one of the following statements holds.*

(a) $\alpha = \beta = 0$. *Then* (7) *has the Darboux first integral*

$$H(x,y) = \frac{(x^2+y^2)\exp\bigl(-2\lambda \arctan \frac{y}{x}\bigr)}{\mathcal{P}_2(x,y)},$$

where $\mathcal{P}_2(x,y) = 2\lambda^3 + (A+C)(x^2+y^2) + 2\lambda^2(Ax^2 + y(Bx+Cy)) + \lambda(2 - 2Cxy + 2Axy + B(y^2-x^2))$.

(b) $A = a\alpha(\alpha\lambda - \beta)/(1+\lambda^2)$, $B = a(\alpha^2 - \beta^2 + 2\alpha\beta\lambda)/(1+\lambda^2)$, *and* $C = a\beta(\alpha+\beta\lambda)/(1+\lambda^2)$. *Then* (7) *has, if* $a > \frac{1}{4}$, *the Darboux first integral* $H(x,y) =$

$$\frac{(x^2+y^2)\exp\Bigl(-2\lambda \arctan \frac{y}{x} - \frac{2}{\sqrt{1-4a}} \arctan \frac{1+\lambda^2 - 2a(\beta-\alpha\lambda)x + 2a(\alpha+\beta\lambda)y}{(1+\lambda^2)\sqrt{1-4a}}\Bigr)}{\mathcal{P}_2(x,y)},$$

where $\mathcal{P}_2(x,y) = 1 + k^4 + a\beta^2x^2 + \alpha y + a\alpha^2y^2 + k^3(\alpha x + \beta y) + k(\alpha x + \beta y)(1 - 2a\beta x + 2a\alpha y) - \beta(x + 2a\alpha xy) + k^2(2 + a\alpha^2x^2 + \alpha y + a\beta^2y^2 + \beta x(-1 + 2a\alpha y))$;

if $a < \frac{1}{4}$ and $a \neq 0$ it has the Darboux first integral

$$H(x, y) = \frac{(x^2 + y^2)\exp\bigl(-2\lambda \arctan \frac{y}{x}\bigr)\mathcal{R}_1(x, y)}{\mathcal{R}_2(x, y)},$$

where

$$\mathcal{R}_1(x, y) = \bigl((-1+\sqrt{1-4a})(1+\lambda^2)-2a(-\beta x+\alpha\lambda x+\alpha y+\beta\lambda y)\bigr)^{-1-1/\sqrt{1-4a}},$$
$$\mathcal{R}_2(x, y) = \bigl((-1-\sqrt{1-4a})(1+\lambda^2)+2a(-\beta x+\alpha\lambda x+\alpha y+\beta\lambda y)\bigr)^{1-1/\sqrt{1-4a}};$$

in the case $a = \frac{1}{4}$ the Darboux first integral is

$$H(x, y) = \frac{(x^2 + y^2)\exp\bigl(-2\lambda \arctan \frac{y}{x} - \frac{2(1+\lambda^2)}{1+\lambda^2-2a\beta x+2a\alpha\lambda x+2a(\alpha+\beta\lambda)y}\bigr)}{\bigl(1+\lambda^2-2a\beta x+2a\alpha\lambda x+2a(\alpha+\beta\lambda)y\bigr)^2};$$

and in the case $a = 0$ the Darboux first integral is

$$H(x, y) = \frac{(x^2 + y^2)\exp\bigl(-2\lambda \arctan \frac{y}{x}\bigr)}{(1+\lambda^2-\beta x+\alpha\lambda x+\alpha y+\beta\lambda y)^2}.$$

Consequently, for $\lambda \neq 0$ these cubic systems have a focus at the origin and are Darboux integrable.

Proof. This follows from parts (a) and parts (b) of Theorem 1 for $n = 2$, after tedious computations. □

System (1) with $n = 3$ and $g_4(\theta) = g_6(\theta) = 0$ — i.e., the quintic system (8) — has a focus or a center at the origin. The following corollary characterizes quintic polynomial systems (8) belonging to class $\mathcal{F}$.

Corollary 6. *A system (8) with $\lambda \neq 0$ belongs to class $\mathcal{F}$ if and only if one of the following statements holds.*

(a) $A = B = C = 0$. *Then (8) has the Darboux first integral given by Theorem 1(b) with $n = 3$ and $f_4(\theta) = 0$.*

(b) $A = B = D = E = 0$, $F = a\,C^2/(2\lambda(1+\lambda^2))$, $G = -a\,C^2/(1+\lambda^2)$ *and* $H = a\,C^2(1+2\lambda^2)/(2\lambda(1+\lambda^2))$. *Then (8) has the Darboux first integral given by Theorem 1(a) with $n = 3$.*

(c) $B = 2\lambda A$, $C = 2\lambda D(1+\lambda^2) - aA^2/(aA)$, $E = 2aA^2$, $F = 2\lambda(2a^2A^4 + D^2 - a\lambda A^2 D + \lambda^2 D^2)/(aA^2)$, $G = 2(4a\lambda A^2 D - a^2A^4 - 2\lambda^2 D^2 + 4a\lambda^3 A^2 D - 2\lambda^4 D^2)/(aA^2)$ *and* $H = D(1+2\lambda^2)(2\lambda D(1+\lambda^2) - aA^2)/(aA^2)$. *Then (8) has the Darboux first integral given by Theorem 1(a) with $n = 3$.*

(d) $C = (-aAB + a\lambda B^2 - 2a\lambda A^2 + 2\lambda E - 4a\lambda^2 AB + 2\lambda^3 E)/(a(B - 2\lambda A))$, $D = A(2aA^2 - E)/(2\lambda A - B)$, $F = (3a^2AB^3 + 8\lambda a^2A^4 - 12a^2\lambda A^2B^2 - 8a\lambda A^2E - a\lambda B^2E + 2\lambda E^2 + 16a^2\lambda^2A^3B + 2a\lambda^2 ABE - 8a\lambda^3A^2E + 2\lambda^3E^2)/$

$(a(B-2\lambda A)^2)$, $G=(a^2B^4-4a^2A^2B^2+aB^2E-4a^2\lambda AB^3+4a\lambda ABE+4a\lambda^2A^2E-4\lambda^2E^2+8a\lambda^3ABE-4\lambda^4E^2)/(a(B-2\lambda A)^2)$, $H=(-2aA^2+aB^2+E-4a\lambda AB+2\lambda^2E)(-aAB-2a\lambda A^2+a\lambda B^2+2\lambda E-4a\lambda^2AB+2\lambda^3E)/(a(B-2\lambda A)^2)$. *Then* (8) *has the Darboux first integral given by Theorem 1*(*a*) *with* $n=3$.

(e) $D=E=F=G=H=0$. *Then* (8) *has the Darboux first integral given by Theorem 1*(*b*) *with* $n=3$ *and* $f_6(\theta)=0$.

Consequently, *for* $\lambda\neq 0$ *these quintic systems have a focus at the origin and are Darboux integrable*.

Proof. This follows from parts (a) and parts (b) of Theorem 1 for $n=3$, after tedious computations using a computer-algebra program. □

References

[Chavarriga and Giné 1998] J. Chavarriga and J. Giné, "Integrability of cubic systems with degenerate infinity", *Differential Equations Dynam. Systems* **6**:4 (1998), 425–438. MR 2001f:34052 Zbl 0998.34001

[Chavarriga et al. 1999] J. Chavarriga, H. Giacomini, J. Giné, and J. Llibre, "On the integrability of two-dimensional flows", *J. Differential Equations* **157**:1 (1999), 163–182. MR 2000h:37058 Zbl 0940.37005

[Chavarriga et al. 2001] J. Chavarriga, I. A. García, and J. Giné, "On integrability of differential equations defined by the sum of homogeneous vector fields with degenerate infinity", *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **11**:3 (2001), 711–722. MR 2002g:34062

[Cheb-Terrab and Roche 2003] E. S. Cheb-Terrab and A. D. Roche, "An Abel ordinary differential equation class generalizing known integrable classes", *European J. Appl. Math.* **14**:2 (2003), 217–229. MR 2004c:34003 Zbl 1046.34004

[Christopher 1994] C. J. Christopher, "Invariant algebraic curves and conditions for a centre", *Proc. Roy. Soc. Edinburgh Sect. A* **124**:6 (1994), 1209–1229. MR 95m:34052 Zbl 0821.34023

[Christopher and Llibre 2000] C. Christopher and J. Llibre, "Integrability via invariant algebraic curves for planar polynomial differential systems", *Ann. Differential Equations* **16**:1 (2000), 5–19. MR 2001g:34001 Zbl 0974.34005

[Collins 1997] C. B. Collins, "Conditions for a centre in a simple class of cubic systems", *Differential Integral Equations* **10**:2 (1997), 333–356. MR 98c:34043 Zbl 0894.34022

[Conti 1994] R. Conti, "Uniformly isochronous centers of polynomial systems in $\mathbb{R}^2$", pp. 21–31 in *Differential equations, dynamical systems, and control science*, edited by K. D. Elworthy et al., Lecture Notes in Pure and Appl. Math. **152**, Dekker, New York, 1994. MR 94i:34061 Zbl 0795.34021

[Dolov 1976] M. V. Dolov, "The canonical integral in the neighborhood of a focus", *Differencial'nye Uravnenija* **12**:11 (1976), 1946–1953. English translation: *Differential Equations* **12** (1977), 1362–1368. MR 56 #738 Zbl 0376.34004

[Gasull and Llibre 1990] A. Gasull and J. Llibre, "Limit cycles for a class of Abel equations", *SIAM J. Math. Anal.* **21**:5 (1990), 1235–1244. MR 91e:34036 Zbl 0732.34025

[Giacomini et al. 1996] H. Giacomini, J. Llibre, and M. Viano, "On the nonexistence, existence and uniqueness of limit cycles", *Nonlinearity* **9**:2 (1996), 501–516. MR 97a:34073 Zbl 0886.58087

[Giné and Llibre 2004] J. Giné and J. Llibre, "Integrability and algebraic limit cycles for polynomial differential systems with homogeneous nonlinearities", *J. Differential Equations* **197**:1 (2004), 147–161. MR 2004j:34077

[Giné and Santallusia 2001] J. Giné and X. Santallusia, "On the Poincaré-Lyapunov constants and the Poincaré series", *Appl. Math.* (*Warsaw*) **28**:1 (2001), 17–30. MR 2002c:34044 Zbl 1022.34028

[Kamke 1943] E. Kamke, *Differentialgleichungen: Lösungsmethoden und Lösungen*, Mathematik und ihre Anwendungen **18**, Becker & Erler, Leipzig, 1943. Zbl 0028.22702

[Kooij and Christopher 1993] R. E. Kooij and C. J. Christopher, "Algebraic invariant curves and the integrability of polynomial systems", *Applied Math. Letters* **6**:4 (1993), 51–53. MR 1348232 Zbl 0784.34008

[Llibre and Rodríguez 2004] J. Llibre and G. Rodríguez, "Configurations of limit cycles and planar polynomial vector fields", *J. Differential Equations* **198**:2 (2004), 374–380. MR 2004j:34100

[Lloyd et al. 1997] N. G. Lloyd, C. J. Christopher, J. Devlin, J. M. Pearson, and N. Yasmin, "Quadratic-like cubic systems", *Differential Eq. Dynam. Systems* **5**:3-4 (1997), 329–345. MR 99k:34061 Zbl 0898.34026

[Pearson et al. 1996] J. M. Pearson, N. G. Lloyd, and C. J. Christopher, "Algorithmic derivation of centre conditions", *SIAM Rev.* **38**:4 (1996), 619–636. MR 98c:34050 Zbl 0876.34033

[Petit Bois 1961] G. Petit Bois, *Tables of indefinite integrals*, Dover, New York, 1961. MR 23 #A256 Zbl 0098.26408

[Schlomiuk 1993] D. Schlomiuk, "Algebraic and geometric aspects of the theory of polynomial vector fields", pp. 429–467 in *Bifurcations and periodic orbits of vector fields* (Montreal, PQ, 1992), edited by D. Schlomiuk, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **408**, Kluwer Acad. Publ., Dordrecht, 1993. MR 95a:34042 Zbl 0790.34031

[Singer 1992] M. F. Singer, "Liouvillian first integrals of differential equations", *Trans. Amer. Math. Soc.* **333**:2 (1992), 673–688. MR 92m:12014 Zbl 0756.12006

[Volokitin 2002] E. P. Volokitin, "Center conditions for a simple class of quintic systems", *Int. J. Math. Math. Sci.* **29**:11 (2002), 625–632. MR 2002m:34042 Zbl 1008.34023

Received April 11, 2003.

JAUME GINÉ
DEPARTAMENT DE MATEMÀTICA
UNIVERSITAT DE LLEIDA
AV. JAUME II, 69
25001 LLEIDA
SPAIN

gine@eup.udl.es

JAUME LLIBRE
DEPARTAMENT DE MATEMÀTIQUES
UNIVERSITAT AUTÒNOMA DE BARCELONA
08193 – BELLATERRA, BARCELONA
SPAIN

jllibre@mat.uab.es

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

A NONCOMMUTATIVE BGG CORRESPONDENCE

PETER JØRGENSEN

We set up a noncommutative version of the Bernšteĭn–Gel'fand–Gel'fand (BGG) correspondence and apply it to periodic injective resolutions.

Introduction

The Bernšteĭn–Gel'fand–Gel'fand (BGG) correspondence is surprising. Originally established in [Bernšteĭn et al. 1978, Theorem 2], it gives an equivalence of categories

$$\overline{\mathsf{gr}}(E) \simeq \mathsf{D}^{\mathrm{b}}(\mathsf{coh}\,\mathbb{P}^e).$$

Let me explain this formula: On the left, E is the exterior algebra $\bigwedge(Y_1, \ldots, Y_{e+1})$ and $\overline{\mathsf{gr}}(E)$ is the category of finitely generated graded E-left-modules modulo morphisms which factor through injectives. On the right, $\mathbb{P}^e$ is e-dimensional projective space, $\mathsf{coh}\,\mathbb{P}^e$ is the category of coherent sheaves on $\mathbb{P}^e$, and $\mathsf{D}^{\mathrm{b}}(\mathsf{coh}\,\mathbb{P}^e)$ is the derived category of bounded complexes of such sheaves.

The surprising thing about the correspondence is that the geometric object on the right-hand side is equivalent to the purely algebraic object on the left-hand side. Put differently, if one did not know about the BGG correspondence, it would really not be obvious that it is possible to recover $\mathsf{D}^{\mathrm{b}}(\mathsf{coh}\,\mathbb{P}^e)$ purely algebraically!

In this paper, I will generalize the BGG correspondence to noncommutative projective geometry. Noncommutative projective geometry is well established; one of the seminal papers is [Artin and Zhang 1994] but many have been published since, showing how a range of projective geometry can be generalized in a noncommutative way. This turns out also to be true of the BGG correspondence, which is generalized in Theorem 3.1 below and now takes the form

$$\overline{\mathsf{Gr}}(A^!) \simeq \mathsf{D}(\mathsf{QGr}\,A).$$

Here A is a suitable noncommutative graded algebra with Koszul dual algebra $A^!$, and the category $\mathsf{QGr}(A)$ is a noncommutative analogue of the category $\mathsf{QCoh}(\mathbb{P}^e)$ of quasi-coherent sheaves on $\mathbb{P}^e$.

MSC2000: 14A22, 16E05, 16W50.

Keywords: noncommutative projective space, Koszul duality, Bernšteĭn–Gel'fand–Gel'fand correspondence, graded Frobenius algebra, periodic injective resolution.

After proving this, I consider an application to periodic injective resolutions. The background is a result by Eisenbud [2002, Theorem 2.2]: Let M be a finitely generated graded module without injective direct summands over the exterior algebra E, for which the Bass numbers

$$\mu^i(M) = \dim_k \operatorname{Ext}^i_E(k, M)$$

are bounded for $i \geq 0$. Then the minimal injective resolution I of M is periodic with period one: All the modules I^i and all the differentials ∂^i_I are the same, up to isomorphism and degree shift. (In fact, Eisenbud worked with minimal free resolutions, but using the Matlis duality functor $\operatorname{Hom}_k(-, k)$ on his result gives the above.)

I will show that this phenomenon can be understood geometrically in a very simple way: Using the BGG correspondence, the module M can be translated to a geometric object on $\mathbb{P}^e$. Since the Bass numbers of M are bounded, this object turns out to have zero-dimensional support, so is stable under twisting, that is, tensoring by $\mathcal{O}_{\mathbb{P}^e}(1)$. Translating back, this means that M is its own first syzygy, and periodicity of the minimal injective resolution follows.

Next, I consider the noncommutative case where a similar procedure yields remarkably different results: Let A be a noncommutative graded algebra, and let M be a finitely generated graded module over the Koszul dual $A^!$, for which the Bass numbers $\mu^i(M)$ are bounded for $i \geq 0$. Then, choosing A and M suitably, it is possible to make the minimal injective resolution of M periodic with any finite period, or to make it aperiodic.

The reason is that when translating M through the noncommutative BGG correspondence, one still obtains a geometric object with zero-dimensional support. However, due to the noncommutative (hence nonlocal) nature of the situation, it is no longer true that such an object is invariant under twisting. Rather, the object can have an orbit of any finite length, or an infinite orbit. Translating back gives the above results on periodicity of the minimal injective resolution.

The concrete example I will give of this behaviour is already known from [Smith 1996]. But the present geometric view through the BGG correspondence is new.

Here is a synopsis of the paper. Section 1 exhibits $\mathsf{D}(\mathsf{QGr}\, A)$ as a full subcategory of $\mathsf{D}(\mathsf{Gr}\, A)$. Section 2 considers a version of Koszul duality. Section 3 combines these results into the noncommutative BGG correspondence, and shows that under the correspondence, the simple module k over $A^!$ corresponds to the "structure sheaf" $\mathcal{O}$ in $\mathsf{D}(\mathsf{QGr}\, A)$.

Section 4 does a few computations that are put to use in Section 5, where the BGG correspondence is applied to periodicity of minimal injective resolutions.

To avoid a lengthy section on nomenclature, hints on notation are given along the way. The reader should rest assured that no new, let alone revolutionary, notation is introduced. The paper remains firmly on classical ground, and differs notationally only in minor details from such papers as [Artin and Zhang 1994], [Jørgensen 1999], and [Smith 1996]. However, I do need the following blanket items, which apply throughout.

Setup 0.1. k is a field, and $A = k \oplus A_1 \oplus A_2 \oplus \cdots$ is a connected $\mathbb{N}$-graded noetherian k-algebra which is AS regular and Koszul (see [Jørgensen 1999, p. 206] and [Beilinson et al. 1996, def. 1.2.1], or Remark 0.2). I assume $\operatorname{gldim} A = d \geq 2$.

Remark 0.2.

(i) For A to be AS regular means that $\operatorname{gldim} A = d$ is finite, and that the graded A-bi-module $k = A/A_{\geq 1}$ satisfies

$$\operatorname{Ext}^i_A(k, A) \cong \operatorname{Ext}^i_{A^{\mathrm{op}}}(k, A) \cong \begin{cases} 0 & \text{for } i \neq d, \\ k(\ell) & \text{for } i = d \end{cases}$$

for some ℓ. As usual, $(-)(\ell)$ denotes ℓ-th degree shift of graded modules, so $M(\ell)_i = M_{i+\ell}$.

(ii) For A to be Koszul means that the minimal free resolution L of the graded A-left-module $k = A/A_{\geq 1}$ is linear. That is, the i-th module L_i has all its generators in graded degree i, so has the form $\coprod A(-i)$.

(iii) It is easy to see that since A is Koszul, the constant ℓ in (i) must be d.

(iv) By [Beilinson et al. 1996, Cor. 2.3.3], the algebra A is quadratic, that is, it has the form

$$A \cong \mathrm{T}(V)/(R)$$

where V is a finite-dimensional vector space, $\mathrm{T}(V)$ the tensor algebra, and (R) the two sided ideal generated by a space of relations R in $V \otimes_k V$. Let $(-)'$ denote $\operatorname{Hom}_k(-, k)$ and define $R^\perp$ by the exact sequence

$$0 \to R^\perp \longrightarrow V' \otimes_k V' \longrightarrow R' \to 0.$$

Then the Koszul dual algebra of A is

$$A^! = \mathrm{T}(V')/(R^\perp);$$

see [Beilinson et al. 1996, Def. 2.8.1].

(v) By [Beilinson et al. 1996, Theorem 2.10.1] there is an isomorphism $(A^!)^{\mathrm{op}} \cong \operatorname{Ext}_A(k, k)$. Combining this with $\operatorname{gldim} A = d$ gives that $A^!$ is concentrated in graded degrees $0, \ldots, d$.

(vi) The algebra $A^!$ is graded Frobenius by [Smith 1996, Proposition 5.10]. This means that $\dim_k A^!$ is finite, and that there is an isomorphism of graded $A^!$-left-modules $(A^!)' \cong A^!(m)$, where $(A^!)' = \operatorname{Hom}_k(A^!, k)$ is the Matlis dual module of $A^!$.

(vii) Since $A^!$ is concentrated in graded degrees $0, \ldots, d$, the constant m in (vi) must be d. So there is an isomorphism of graded $A^!$-left-modules $(A^!)' \cong A^!(d)$.

1. The categories $\mathsf{Gr}(A)$ and $\mathsf{QGr}(A)$

Remark 1.1. Let me first recapitulate a few items from [Artin and Zhang 1994], to which I refer for further details and proofs.

The category $\mathsf{Gr}(A)$ has as objects all $\mathbb{Z}$-graded A-left-modules and as morphisms all homomorphisms of A-left-modules which preserve graded degree.

A module M in $\mathsf{Gr}(A)$ is called torsion if each m in M is annihilated by $A_{\geq n}$ for some n. The torsion modules form a dense subcategory $\mathsf{Tors}(A)$ of $\mathsf{Gr}(A)$, and the quotient category is

$$\mathsf{QGr}(A) = \mathsf{Gr}(A)/\mathsf{Tors}(A).$$

This category behaves like the category of quasi-coherent sheaves on the space $\operatorname{Proj}(A)$, although $\operatorname{Proj}(A)$ itself may not make sense. For instance, if A is commutative, $\mathsf{QGr}(A)$ is in fact equivalent to the category of quasi-coherent sheaves on $\operatorname{Proj}(A)$ by Serre's theorem, as given in [Artin and Zhang 1994, Theorem, p. 229].

The degree shifting functor $(-)(1)$ on $\mathsf{Gr}(A)$ induces a functor on $\mathsf{QGr}(A)$ which I will also denote $(-)(1)$.

The category $\mathsf{Gr}(A)$ has the full subcategory $\mathsf{gr}(A)$ consisting of finitely generated modules. Induced by this, $\mathsf{QGr}(A)$ has the full subcategory $\mathsf{qgr}(A)$ which behaves like the category of coherent sheaves on $\operatorname{Proj}(A)$.

The projection functor $\mathsf{Gr}(A) \xrightarrow{\pi} \mathsf{QGr}(A)$ has a right-adjoint functor

$$\mathsf{QGr}(A) \xrightarrow{\omega} \mathsf{Gr}(A)$$

by [Artin and Zhang 1994, p. 234], so there is an adjoint pair

$$\mathsf{Gr}(A) \underset{\omega}{\overset{\pi}{\rightleftarrows}} \mathsf{QGr}(A).$$

As follows from [Artin and Zhang 1994, Proposition 7.1], these functors send injective objects to injective objects, and restrict to a pair of quasi-inverse equivalences

$$\mathsf{Inj}_{\,\mathrm{tf}}(A) \underset{\omega}{\overset{\pi}{\rightleftarrows}} \mathsf{QInj}(A) \tag{1}$$

between the subcategory of torsion-free injective objects of $\mathsf{Gr}(A)$ and the subcategory of all injective objects of $\mathsf{QGr}(A)$.

Let me next turn to derived categories. The projection functor π is exact and so extends to a triangulated functor

$$\mathsf{D}(\mathsf{Gr}\,A) \xrightarrow{\pi} \mathsf{D}(\mathsf{QGr}\,A)$$

between derived categories. Moreover, since A has finite global dimension, each object of the category $\mathsf{Gr}(A)$ has a bounded resolution by injective objects. The same therefore holds for $\mathsf{QGr}(A)$, as one sees using ω and π. So right-derived functors can be defined on the unbounded derived categories $\mathsf{D}(\mathsf{Gr}\,A)$ and $\mathsf{D}(\mathsf{QGr}\,A)$ by [Weibel 1994, Section 10.5].

In particular, $\mathsf{D}(\mathsf{QGr}\,A) \xrightarrow{\mathrm{R}\omega} \mathsf{D}(\mathsf{Gr}\,A)$ exists, and it is not hard to see that

$$\mathsf{D}(\mathsf{Gr}\,A) \underset{\mathrm{R}\omega}{\overset{\pi}{\rightleftarrows}} \mathsf{D}(\mathsf{QGr}\,A) \tag{2}$$

is an adjoint pair of functors.

Definition 1.2. Let

$$k^{\perp} = \{N \in \mathsf{D}(\mathsf{Gr}\,A) \mid \mathrm{RHom}_A(k, N) = 0\}.$$

Proposition 1.3. *The functors in equation* (2) *restrict to a pair of quasi-inverse equivalences of triangulated categories*

$$k^{\perp} \underset{\mathrm{R}\omega}{\overset{\pi}{\rightleftarrows}} \mathsf{D}(\mathsf{QGr}\,A). \tag{3}$$

Proof. First observe that diagram (1) extends to a pair of quasi-inverse equivalences

$$\mathsf{K}(\mathsf{Inj}_{\,\mathrm{tf}}\,A) \underset{\omega}{\overset{\pi}{\rightleftarrows}} \mathsf{K}(\mathsf{QInj}\,A) \tag{4}$$

between the homotopy category of complexes of torsion free injective objects of $\mathsf{Gr}(A)$, and the homotopy category of complexes of injective objects of $\mathsf{QGr}(A)$.

Next, the finite global dimension of A implies that $\mathsf{K}(\mathsf{Inj}\,A)$, the homotopy category of complexes of injective objects of $\mathsf{Gr}(A)$, is equivalent to $\mathsf{D}(\mathsf{Gr}\,A)$. Under the equivalence, the restriction of a functor F to $\mathsf{K}(\mathsf{Inj}\,A)$ corresponds to the right derived functor $\mathrm{R}F$ on $\mathsf{D}(\mathsf{Gr}\,A)$. See [Weibel 1994, Section 10.5], for example. A similar remark applies to $\mathsf{K}(\mathsf{QInj}\,A)$ and $\mathsf{D}(\mathsf{QGr}\,A)$. So forming

$$\mathsf{K}(\mathsf{Inj}\,A) \underset{\omega}{\overset{\pi}{\rightleftarrows}} \mathsf{K}(\mathsf{QInj}\,A)$$

gives a diagram which, up to equivalence, is just diagram (2).

This shows that diagram (4) gives an equivalence between some subcategory of $\mathsf{D}(\mathsf{Gr}\,A)$ and the whole category $\mathsf{D}(\mathsf{QGr}\,A)$. To finish the proof, I must show that the subcategory in question is $k^{\perp}$. That is, I must show that the subcategory $\mathsf{K}(\mathsf{Inj}_{\mathrm{tf}}\,A)$ of $\mathsf{K}(\mathsf{Inj}\,A)$ corresponds to the subcategory $k^{\perp}$ of $\mathsf{D}(\mathsf{Gr}\,A)$. For this, note that by the above, the functor $\mathrm{Hom}_A(k,-)$ on $\mathsf{K}(\mathsf{Inj}\,A)$ corresponds to the derived functor $\mathrm{RHom}_A(k,-)$ on $\mathsf{D}(\mathsf{Gr}\,A)$, so I must show that $\mathsf{K}(\mathsf{Inj}_{\mathrm{tf}}\,A)$ is the subcategory of $\mathsf{K}(\mathsf{Inj}\,A)$ annihilated by $\mathrm{Hom}_A(k,-)$.

In fact, this is not quite true, but it is true and easy to see that the subcategory of $\mathsf{K}(\mathsf{Inj}\,A)$ annihilated by $\mathrm{Hom}_A(k,-)$ consists exactly of the complexes isomorphic to complexes in $\mathsf{K}(\mathsf{Inj}_{\mathrm{tf}}\,A)$, and this is enough. $\square$

2. Koszul duality

Remark 2.1. Let me recapitulate the version of Koszul duality set up by Beilinson, Ginzburg and Soergel, and elaborated on by Fløystad.

In [Beilinson et al. 1996, proof of Theorem 2.12.1] and [Fløystad 2003, Section 3.2] we find a construction for an adjoint pair of functors between categories of complexes of graded modules,

$$\mathsf{Ch}(\mathsf{Gr}\,A^{!}) \underset{\mathrm{G}}{\overset{\mathrm{F}}{\rightleftarrows}} \mathsf{Ch}(\mathsf{Gr}\,A). \tag{5}$$

These functors are defined as follows: Given M in $\mathsf{Ch}(\mathsf{Gr}\,A^{!})$, one constructs a double complex

$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
\cdots \arrow[r] & A(1)\otimes M_1^{-1} \arrow[r] \arrow[u] & A(1)\otimes M_1^{0} \arrow[r] \arrow[u] & A(1)\otimes M_1^{1} \arrow[r] \arrow[u] & \cdots \\
\cdots \arrow[r] & A\otimes M_0^{-1} \arrow[r] \arrow[u] & A\otimes M_0^{0} \arrow[r] \arrow[u] & A\otimes M_0^{1} \arrow[r] \arrow[u] & \cdots \\
\cdots \arrow[r] & A(-1)\otimes M_{-1}^{-1} \arrow[r] \arrow[u] & A(-1)\otimes M_{-1}^{0} \arrow[r] \arrow[u] & A(-1)\otimes M_{-1}^{1} \arrow[r] \arrow[u] & \cdots \\
 & \vdots \arrow[u] & \vdots \arrow[u] & \vdots \arrow[u] &
\end{tikzcd}$$

with certain differentials, and the total complex $\mathrm{Tot}^{\coprod}$, defined using coproducts, is $\mathrm{F}(M)$. In the diagram, superscripts indicate cohomological degree and subscripts indicate graded degree. Also, $\otimes$ denotes tensor product over k.

And given N in $\mathsf{Ch}(\mathsf{Gr}\,A)$, one constructs a double complex

$$\begin{tikzcd}
 & \vdots & \vdots & \vdots & \\
\cdots \arrow[r] & \mathrm{Hom}(A^!(-1), N_1^{-1}) \arrow[r] \arrow[u] & \mathrm{Hom}(A^!(-1), N_1^{0}) \arrow[r] \arrow[u] & \mathrm{Hom}(A^!(-1), N_1^{1}) \arrow[r] \arrow[u] & \cdots \\
\cdots \arrow[r] & \mathrm{Hom}(A^!, N_0^{-1}) \arrow[r] \arrow[u] & \mathrm{Hom}(A^!, N_0^{0}) \arrow[r] \arrow[u] & \mathrm{Hom}(A^!, N_0^{1}) \arrow[r] \arrow[u] & \cdots \\
\cdots \arrow[r] & \mathrm{Hom}(A^!(1), N_{-1}^{-1}) \arrow[r] \arrow[u] & \mathrm{Hom}(A^!(1), N_{-1}^{0}) \arrow[r] \arrow[u] & \mathrm{Hom}(A^!(1), N_{-1}^{1}) \arrow[r] \arrow[u] & \cdots \\
 & \vdots \arrow[u] & \vdots \arrow[u] & \vdots \arrow[u] &
\end{tikzcd}$$

with certain differentials, and the total complex $\mathrm{Tot}^{\prod}$, defined using products, is $\mathrm{G}(N)$. In the diagram, Hom denotes homomorphisms over k.

Now consider $\mathsf{CoFree}(A^!)$, the full subcategory of $\mathsf{Gr}(A^!)$ consisting of modules which have the form $\prod_j (A^!)'(m_j)$, and $\mathsf{Free}(A)$, the full subcategory of $\mathsf{Gr}(A)$ consisting of modules which have the form $\coprod_i A(n_i)$. On the corresponding homotopy categories of complexes, the functors F and G induce functors which, abusively, I will denote by the same letters,

$$\mathsf{K}(\mathsf{CoFree}\,A^!) \underset{\mathrm{G}}{\overset{\mathrm{F}}{\rightleftarrows}} \mathsf{K}(\mathsf{Free}\,A). \tag{6}$$

According to [Fløystad 2003, Proposition 5.11], this is a pair of quasi-inverse equivalences of triangulated categories.

Finite global dimension of A implies that $\mathsf{D}(\mathsf{Gr}\,A)$ is equivalent to $\mathsf{K}(\mathsf{Free}\,A)$ (see [Weibel 1994, Section 10.5]), so the equivalences (6) can also be read as

$$\mathsf{K}(\mathsf{CoFree}\,A^!) \underset{\mathrm{G}}{\overset{\mathrm{F}}{\rightleftarrows}} \mathsf{D}(\mathsf{Gr}\,A). \tag{7}$$

Remark 2.2. The name Koszul duality is potentially confusing: "duality" might lead one to think of contravariant functors, while F and G are in fact covariant.

For the following lemma, note that I use $\Sigma^i(-)$ for the i-th suspension, so if M is a complex then $(\Sigma^i M)^\ell = M^{i+\ell}$.

Lemma 2.3. *The functors* F *and* G *in equation* (7) *satisfy the following.*

(i) $\mathrm{F}(M(i)) \cong \Sigma^i(\mathrm{F}M)(-i)$.

(ii) $\mathrm{G}(N(j)) \cong \Sigma^j(\mathrm{G}N)(-j)$.

(iii) $\mathrm{F}((A^!)')$ *is isomorphic to the A-left-module k.*

Proof. (i) and (ii) can be seen by playing with the double complexes which define F and G. (iii) follows from [Beilinson et al. 1996, Theorem 2.12.5(iii)]. □

Remark 2.4. The injective stable category over a ring is defined as the module category modulo the ideal of morphisms which factor through an injective module.

The present paper uses the graded version of this, so the injective stable category $\overline{\mathsf{Gr}}(A^!)$ is defined as $\mathsf{Gr}(A^!)$ modulo the ideal of morphisms which factor through an injective object of $\mathsf{Gr}(A^!)$.

Since $A^!$ is graded Frobenius by Remark 0.2(vi), the category $\mathsf{Gr}(A^!)$ is Frobenius by the graded version of [Happel 1987, Section 9.2], and so the category $\overline{\mathsf{Gr}}(A^!)$ is triangulated by [Happel 1987, Section 9.4]. For M in $\overline{\mathsf{Gr}}(A^!)$, the suspension ΣM is the first syzygy in an injective resolution of M. So ΣM is the cokernel of an injective pre-envelope, that is, an injective homomorphism $M \longrightarrow I$ in $\mathsf{Gr}(A^!)$, where I is an injective object of $\mathsf{Gr}(A^!)$. Any injective pre-envelope can be used; changing the injective pre-envelope does not change the isomorphism class of ΣM in $\overline{\mathsf{Gr}}(A^!)$.

The degree shifting functor $(-)(1)$ on $\mathsf{Gr}(A^!)$ induces a functor on $\overline{\mathsf{Gr}}(A^!)$ which I will also denote $(-)(1)$.

Since $\mathsf{Gr}(A^!)$ is Frobenius, the methods of [Keller 1994, Section 4.3] show that the category $\overline{\mathsf{Gr}}(A^!)$ is equivalent to the full subcategory of exact complexes in $\mathsf{K}(\mathsf{CoFree}\,A^!)$. Under the equivalence, a module M corresponds to a complete cofree resolution C of M, that is, a complex C in $\mathsf{K}(\mathsf{CoFree}\,A^!)$ which is exact and has its zeroth cycle module $\mathrm{Z}^0(C)$ isomorphic to M.

Under the equivalence between $\overline{\mathsf{Gr}}(A^!)$ and the full subcategory of exact complexes in $\mathsf{K}(\mathsf{CoFree}\,A^!)$, the suspension Σ on $\overline{\mathsf{Gr}}(A^!)$ corresponds to the ordinary suspension Σ on $\mathsf{K}(\mathsf{CoFree}\,A^!)$, given by moving complexes one step to the left and switching signs of differentials. Also, the functor $(-)(1)$ on $\overline{\mathsf{Gr}}(A^!)$ corresponds to the functor $(-)(1)$ on $\mathsf{K}(\mathsf{CoFree}\,A^!)$ induced by degree shifting of $A^!$-left-modules.

Proposition 2.5. *The functors in* (7) *induce a pair of quasi-inverse equivalences of triangulated categories*

$$\overline{\mathsf{Gr}}(A^!) \rightleftarrows k^{\perp}.$$

Proof. Remark 2.4 identifies $\overline{\mathsf{Gr}}(A^!)$ with the full subcategory of exact complexes in $\mathsf{K}(\mathsf{CoFree}\,A^!)$, and Definition 1.2 defines $k^{\perp}$ as a full subcategory of $\mathsf{D}(\mathsf{Gr}\,A)$. To prove the proposition, I must show that these subcategories are mapped to each other by the functors F and G of equation (7).

However, let N be in $\mathsf{D}(\mathsf{Gr}\,A)$. Then the j-th graded component of the i-th cohomology module of the complex $\mathrm{G}N$ is

$$\begin{aligned}
\mathrm{h}^i(\mathrm{G}N)_j &\overset{\text{(a)}}{\cong} \mathrm{Hom}_{\mathsf{K}(\mathsf{Gr}\,A^!)}(A^!, \Sigma^i(\mathrm{G}N)(j)) \\
&\overset{\text{(b)}}{\cong} \mathrm{Hom}_{\mathsf{K}(\mathsf{CoFree}\,A^!)}((A^!)'(-d), \Sigma^i(\mathrm{G}N)(j)) \\
&\cong \mathrm{Hom}_{\mathsf{K}(\mathsf{CoFree}\,A^!)}(\Sigma^{-i}(A^!)'(-d-j), \mathrm{G}N) \\
&= (*),
\end{aligned}$$

where (a) is classical and (b) holds because of $A^! \cong (A^!)'(-d)$; see Remark 0.2(vii). Adjointness between F and G gives (c) in

$$\begin{aligned}
(*) &\overset{\text{(c)}}{\cong} \mathrm{Hom}_{\mathsf{D}(\mathsf{Gr}\,A)}(\mathrm{F}(\Sigma^{-i}(A^!)'(-d-j)), N) \\
&\overset{\text{(d)}}{\cong} \mathrm{Hom}_{\mathsf{D}(\mathsf{Gr}\,A)}(\Sigma^{-i-d-j}\,\mathrm{F}((A^!)')(d+j), N) \\
&\cong \mathrm{Hom}_{\mathsf{D}(\mathsf{Gr}\,A)}(\mathrm{F}((A^!)'), \Sigma^{i+d+j}N(-d-j)) \\
&\overset{\text{(e)}}{\cong} \mathrm{Hom}_{\mathsf{D}(\mathsf{Gr}\,A)}(k, \Sigma^{i+j+d}N(-j-d)) \\
&\cong \mathrm{h}^{i+j+d}\,\mathrm{RHom}_A(k, N)_{-j-d},
\end{aligned}$$

and (d) and (e) are by Lemma 2.3, parts (i) and (iii).

But now it is clear that $\mathrm{G}N$ is exact if and only if N is in $k^{\perp}$, as desired. □

3. The BGG correspondence

Composing the equivalences of categories from Propositions 1.3 and 2.5 gives the following main theorem of the paper.

Theorem 3.1 (The BGG correspondence). *There are quasi-inverse equivalences of triangulated categories*

$$\overline{\mathsf{Gr}}(A^!) \underset{\gamma}{\overset{\varphi}{\rightleftarrows}} \mathsf{D}(\mathsf{QGr}\,A).$$

Example 3.2. If A is the polynomial algebra $k[X_1, \ldots, X_d]$ then it is classical that A satisfies the conditions of Setup 0.1, and the definition of $A^!$ in Remark 0.2(iv) makes it easy to see that $A^!$ is the exterior algebra $E = \bigwedge(Y_1, \ldots, Y_d)$. Also, $\mathsf{QGr}(A)$ is equivalent to the category $\mathsf{QCoh}(\mathbb{P}^{d-1})$ of quasi-coherent sheaves on $(d-1)$-dimensional projective space by Serre's theorem, [Artin and Zhang 1994, Theorem, p. 229]. So Theorem 3.1 gives an equivalence of categories

$$\overline{\mathsf{Gr}}(E) \simeq \mathsf{D}(\mathsf{QCoh}\,\mathbb{P}^{d-1}).$$

This is the classical BGG correspondence, originally established in [Bernšteĭn et al. 1978, Theorem 2], with the slight improvement of dealing with the stable category of all modules and the unbounded derived category of quasi-coherent sheaves rather than the finite subcategories in [Bernšteĭn et al. 1978, Theorem 2].

Remark 3.3. The quasi-inverse equivalences φ and γ from Theorem 3.1 are constructed by composing some other functors. Untangling the construction gives the following concrete descriptions.

To get $\varphi(M)$, take a complete cofree resolution C of M. Then $\varphi(M) = \pi\mathrm{F}(C)$, where F is one of the functors from equation (7) and π is one of the functors from equation (2).

To get $\gamma(\mathcal{M})$, consider $\mathrm{G}(\mathrm{R}\omega(\mathcal{M}))$, where $\mathrm{R}\omega$ is one of the functors from equation (2) and G is one of the functors from equation (7). This is an object of $\mathsf{K}(\mathsf{CoFree}\,A^!)$ and in fact, it is even in the full subcategory of exact complexes of $\mathsf{K}(\mathsf{CoFree}\,A^!)$. Now $\gamma(\mathcal{M}) = \mathrm{Z}^0\mathrm{G}(\mathrm{R}\omega(\mathcal{M}))$, where Z^0 takes the zeroth cycle module.

The next lemma follows immediately from Lemma 2.3, parts (i) and (ii).

Lemma 3.4. *The functors φ and γ satisfy the following.*

(i) $\varphi(M(i)) \cong \Sigma^i(\varphi M)(-i)$.

(ii) $\gamma(\mathcal{M}(j)) \cong \Sigma^j(\gamma\mathcal{M})(-j)$.

For the following lemma, let L be the minimal free resolution of the graded $A^!$-left-module k. Each L^i is finitely generated free and hence cofree because Remark 0.2(vii) implies $A^! \cong (A^!)'(-d)$. So L is a complex in $\mathsf{K}(\mathsf{CoFree}\,A^!)$, and I can apply the functor F from equation (7) and get a complex $\mathrm{F}(L)$ in $\mathsf{D}(\mathsf{Gr}\,A)$.

Lemma 3.5. *The cohomology of* $\mathrm{F}(L)$ *is torsion.*

Proof. The version of F from equation (5) respects small colimits because it is constructed using tensor products and small coproducts. In the category of complexes $\mathsf{Ch}(\mathsf{Gr}\,A^!)$, the object L is the colimit of the objects

$$L\langle j\rangle = \cdots \longrightarrow 0 \longrightarrow L^{-j} \longrightarrow \cdots \longrightarrow L^0 \longrightarrow 0 \longrightarrow \cdots,$$

so

$$\mathrm{F}(L) \cong \mathrm{F}(\operatorname{colim} L\langle j\rangle) \cong \operatorname{colim}\mathrm{F}(L\langle j\rangle). \tag{8}$$

Now, $A^!$ is Koszul by [Beilinson et al. 1996, Proposition 2.9.1], and L is the minimal free resolution of k over $A^!$, and so

$$L^{-i} \cong \coprod A^!(-i). \tag{9}$$

This implies that L^{-i} is concentrated in graded degrees $i, \ldots, d+i$ because $A^!$ is concentrated in graded degrees $0, \ldots, d$ by Remark 0.2(v). So the construction in Remark 2.1 says that $\mathrm{F}(L\langle j\rangle)$ is $\mathrm{Tot}^{\coprod}$ of a double complex whose nonzero part can be sketched as

$$\begin{tikzcd}
A(d+j)\otimes L^{-j}_{d+j} & & & \\
A(d+j-1)\otimes L^{-j}_{d+j-1} \arrow[u] \arrow[r] & \cdots & & \\
\vdots \arrow[u] \arrow[r] & \cdots \arrow[r] & \cdots \arrow[r] & A(d)\otimes L^{0}_{d} \\
A(j)\otimes L^{-j}_{j} \arrow[u] \arrow[r] & \cdots \arrow[r] & \cdots \arrow[r] & \vdots \arrow[u] \\
 & & \cdots \arrow[r] & A(1)\otimes L^{0}_{1} \arrow[u] \\
 & & & A\otimes L^{0}_{0}. \arrow[u]
\end{tikzcd} \tag{10}$$

Also, combining equation (9) with $A^{!} \cong (A^{!})'(-d)$, which holds by Remark 0.2(vii), gives $L^{-i} \cong \coprod (A^{!})'(-d-i)$. So, up to degree shift and suspension, the $(-i)$-th column of (10) is just a coproduct of copies of the column obtained from $(A^{!})'$. This column has nonzero part

$$\begin{tikzcd}
A\otimes (A^{!})'_{0} \\
\vdots \arrow[u] \\
A(-d)\otimes (A^{!})'_{-d}, \arrow[u]
\end{tikzcd}$$

and is a free resolution of the A-left-module k, as follows from [Beilinson et al. 1996, Theorem 2.12.5(iii)]. So the columns of (10) have cohomology only at the top ends, and the cohomology in the $(-i)$-th column is $\coprod k(d+i)$.

Now consider the first spectral sequence of the double complex (10) (see [Weibel 1994, Section 5.6]). The previous part of the proof shows that the E_2-term of the spectral sequence is nonzero only at the top ends of the columns of (10), where

$$E_2^{0d} \cong \coprod k(d), \quad \ldots, \quad E_2^{-j,d+j} \cong \coprod k(d+j).$$

Since the double complex is bounded in all directions, the spectral sequence converges towards the cohomology of $\mathrm{Tot}^{\coprod}$. Consequently, $\mathrm{Tot}^{\coprod}$ of the double complex has cohomology only in cohomological degree d, and this cohomology sits in graded degrees $-d, \dots, -d-j$.

But this $\mathrm{Tot}^{\coprod}$ is $\mathrm{F}(L\langle j \rangle)$. So (8) now shows that $\mathrm{F}(L)$ has cohomology only in cohomological degree d, and that this cohomology can be nonzero only in graded degrees $-d, -d-1, \dots$. In particular, the cohomology of $\mathrm{F}(L)$ is torsion. $\square$

Now consider the graded $A^!$-left-module k viewed as an object of $\overline{\mathsf{Gr}}(A^!)$, and consider $\mathcal{O}$, the "structure sheaf" in $\mathsf{QGr}(A)$ defined by $\mathcal{O} = \pi(A)$. Then $\mathcal{O}$ can also be viewed as a complex in $\mathsf{D}(\mathsf{QGr}\,A)$ concentrated in cohomological degree zero, and the following result holds.

Theorem 3.6. *The functor φ satisfies $\varphi(k) \cong \mathcal{O}$.*

Proof. To get $\varphi(k)$, I must take $\pi\mathrm{F}(C)$, where C is a complete cofree resolution of the $A^!$-left-module k, while F and π are the functors from equations (7) and (2); see Remark 3.3.

For this, consider first the functors F and G from (7). Let X in $\mathsf{K}(\mathsf{CoFree}\,A^!)$ be a cofree resolution of k. From [Beilinson et al. 1996, Theorem 2.12.5(iii)] there follows $\mathrm{F}(X) \cong A$. Hence $\mathrm{GF}(X) \cong \mathrm{G}(A)$, and as F and G are quasi-inverse equivalences of categories, this implies $X \cong \mathrm{G}(A)$. But k is quasi-isomorphic to X, so this shows that k is quasi-isomorphic to $\mathrm{G}(A)$. However, it is clear from the construction of G in Remark 2.1 that $\mathrm{G}(A)$ is a complex of cofree modules placed in nonnegative cohomological degrees. All in all, $\mathrm{G}(A)$ must be a cofree resolution of k, so there is a canonical morphism $k \longrightarrow \mathrm{G}(A)$.

Now let L be a minimal free resolution of k as in Lemma 3.5, so there is a canonical morphism $L \longrightarrow k$. Composing the morphisms $L \longrightarrow k$ and $k \longrightarrow \mathrm{G}(A)$ gives a morphism $L \longrightarrow \mathrm{G}(A)$ whose mapping cone C is easily seen to be a complete cofree resolution of k.

The distinguished triangle $L \longrightarrow \mathrm{G}(A) \longrightarrow C \longrightarrow$ in $\mathsf{K}(\mathsf{CoFree}\,A^!)$ gives a distinguished triangle

$$\pi\mathrm{F}(L) \longrightarrow \pi\mathrm{FG}(A) \longrightarrow \pi\mathrm{F}(C) \longrightarrow$$

in $\mathsf{D}(\mathsf{QGr}\,A)$. Let me compute the three complexes here: The cohomology of $\mathrm{F}(L)$ is torsion by Lemma 3.5, so $\pi\mathrm{F}(L) \cong 0$. And F and G are quasi-inverse equivalences, so $\mathrm{FG}(A)$ is isomorphic to A, so $\pi\mathrm{FG}(A) \cong \pi(A) = \mathcal{O}$.

Finally, $\pi\mathrm{F}(C)$ is $\varphi(k)$ as mentioned above. So the distinguished triangle reads

$$0 \longrightarrow \mathcal{O} \longrightarrow \varphi(k) \longrightarrow \ ,$$

proving $\varphi(k) \cong \mathcal{O}$. $\square$

4. Computations

This section contains computations, some involving the BGG correspondence, which will be used on periodic injective resolutions in Section 5.

The following lemma is just a graded version of [Benson 1998, Cor. 2.5.4(ii)].

Lemma 4.1. *Let M be in $\overline{\mathsf{Gr}}(A^!)$. There are canonical isomorphisms*

$$\mathrm{Hom}_{\overline{\mathsf{Gr}}(A^!)}(k, \Sigma^i M) \longrightarrow \mathrm{Ext}^i_{\mathsf{Gr}(A^!)}(k, M)$$

for $i \geq 1$.

Lemma 4.2. *Let M be in $\overline{\mathsf{Gr}}(A^!)$ and consider $\mathcal{M} = \varphi(M)$ in $\mathsf{D}(\mathsf{QGr}\,A)$. Then*

$$\mathrm{Ext}^i_{\mathsf{Gr}(A^!)}(k, M(-i+j)) \cong \mathrm{Ext}^j_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}(i-j))$$

for $i \geq 1$ and each j.

Proof. This is a simple computation,

$$\begin{aligned}
\mathrm{Ext}^i_{\mathsf{Gr}(A^!)}(k, M(-i+j)) &\overset{\text{(a)}}{\cong} \mathrm{Hom}_{\overline{\mathsf{Gr}}(A^!)}(k, \Sigma^i M(-i+j)) \\
&\overset{\text{(b)}}{\cong} \mathrm{Hom}_{\mathsf{D}(\mathsf{QGr}\,A)}(\varphi k, \varphi(\Sigma^i M(-i+j))) \\
&\overset{\text{(c)}}{\cong} \mathrm{Hom}_{\mathsf{D}(\mathsf{QGr}\,A)}(\mathcal{O}, \Sigma^j \mathcal{M}(i-j)) \\
&= \mathrm{Ext}^j_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}(i-j)),
\end{aligned}$$

where (a) is by Lemma 4.1 and (b) is by the BGG correspondence, Theorem 3.1, while (c) is by Theorem 3.6 and Lemma 3.4(i). □

For the following lemma, observe that the finitely generated graded modules form a full subcategory $\overline{\mathsf{gr}}(A^!)$ of $\overline{\mathsf{Gr}}(A^!)$, and that the complexes which have bounded cohomology consisting of objects from the category $\mathsf{qgr}(A)$ form a full subcategory $\mathsf{D}^{\mathsf{f}}(\mathsf{QGr}\,A)$ of $\mathsf{D}(\mathsf{QGr}\,A)$.

Lemma 4.3. *The subcategories $\overline{\mathsf{gr}}(A^!)$ and $\mathsf{D}^{\mathsf{f}}(\mathsf{QGr}\,A)$ map to each other under the BGG correspondence*

$$\overline{\mathsf{Gr}}(A^!) \underset{\gamma}{\overset{\varphi}{\rightleftarrows}} \mathsf{D}(\mathsf{QGr}\,A).$$

Proof. It is not hard to check that $\overline{\mathsf{gr}}(A^!)$ consists of the objects of $\overline{\mathsf{Gr}}(A^!)$ which are finitely built from objects of the form $k(i)$.

Similarly, $\mathsf{D}^{\mathsf{f}}(\mathsf{QGr}\,A)$ consists of the objects of $\mathsf{D}(\mathsf{QGr}\,A)$ which are finitely built from objects of the form $\mathcal{O}(j)$.

But under the BGG correspondence, $k(i)$ corresponds to $\Sigma^i \mathcal{O}(-i)$ by Theorem 3.6 and Lemma 3.4(i), so the present lemma follows. □

Lemma 4.4. *Let $\mathcal{M}$ be in $\mathsf{D}^{\mathrm{f}}(\mathsf{QGr}\, A)$. Then for $i \gg 0$ I have*

$$\mathrm{Ext}^j_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}(i-j)) \cong \mathrm{Hom}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathrm{h}^j(\mathcal{M})(i-j))$$

for each j, where $\mathrm{h}^j(\mathcal{M})$ is the j-th cohomology of $\mathcal{M}$.

Proof. The algebra A has global dimension d by assumption, so $\mathsf{qgr}(A)$ has cohomological dimension at most $d-1$ by [Artin and Zhang 1994, Proposition 7.10(3)], so $\mathrm{Ext}^{\geq d}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{N}) = 0$ holds for each $\mathcal{N}$ in $\mathsf{qgr}(A)$.

Moreover, A is even AS regular by assumption, so $\mathsf{qgr}(A)$ satisfies Serre vanishing by [Artin and Zhang 1994, Theorems 8.1(1) and 7.4]. That is, given $\mathcal{N}$ in $\mathsf{qgr}(A)$ and given p with $1 \leq p \leq d-1$, I have $\mathrm{Ext}^p_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{N}(r)) = 0$ for $r \gg 0$.

So given $\mathcal{N}$, I can kill all the $\mathrm{Ext}^p_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{N}(r))$ with $p \geq 1$ by choosing r large enough. That is, given $\mathcal{N}$ in $\mathsf{qgr}(A)$, I have

$$r \gg 0 \;\Rightarrow\; \mathrm{Ext}^p_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{N}(r)) = 0 \text{ for } p \geq 1. \tag{11}$$

There is a convergent spectral sequence

$$E_2^{pq} = \mathrm{Ext}^p_{\mathsf{QGr}(A)}(\mathcal{O}, \mathrm{h}^q(\mathcal{M})(i-j)) \Rightarrow \mathrm{Ext}^{p+q}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}(i-j))$$

by [Weibel 1994, 5.7.9] (convergence because the cohomology $\mathrm{h}(\mathcal{M})$ is bounded). By assumption on $\mathcal{M}$, the finitely many nonzero $\mathrm{h}^q(\mathcal{M})$'s are in $\mathsf{qgr}(A)$. So equation (11) implies that for $i-j \gg 0$, the term E_2^{pq} is concentrated on the line $p=0$. So the spectral sequence collapses and gives

$$\mathrm{Hom}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathrm{h}^q(\mathcal{M})(i-j)) \cong \mathrm{Ext}^q_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}(i-j)) \tag{12}$$

for $i-j \gg 0$ and each q.

Now observe that the isomorphism (12) also holds for $q \gg 0$, simply because both sides are then zero. For the left-hand side, this is true because $\mathrm{h}(\mathcal{M})$ is bounded. For the right-hand side, use that $\mathrm{h}(\mathcal{M})$ is bounded and that $\mathsf{qgr}(A)$ has cohomological dimension at most $d-1$.

So setting q equal to j, the isomorphism (12) holds for $j \gg 0$, and for other values of j I can force $i-j \gg 0$ by picking $i \gg 0$, and then the isomorphism also holds. That is,

$$\mathrm{Hom}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathrm{h}^j(\mathcal{M})(i-j)) \cong \mathrm{Ext}^j_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}(i-j))$$

for $i \gg 0$ and each j, proving the lemma. □

5. Periodic injective resolutions

This section shows how the BGG correspondence can be used to understand the periodicity of certain injective resolutions over exterior algebras as a geometric phenomenon.

I also show an analogous noncommutative example with much more complicated behaviour, due to the more intricate nature of noncommutative geometry.

The commutative case. Denote by E the exterior algebra $\bigwedge(Y_1, \dots, Y_d)$ over k, and recall that $\mathsf{gr}(E)$ is the category of finitely generated graded E-left-modules. The following result appears in [Eisenbud 2002, Theorem 2.2].

Theorem 5.1 (Eisenbud). *Let M in $\mathsf{gr}(E)$ be without injective direct summands, and suppose that the Bass numbers*

$$\mu^i(M) = \dim_k \operatorname{Ext}^i_E(k, M)$$

are bounded for $i \geq 0$.

Then the minimal injective resolution I of M is periodic with period one in the following sense: Up to isomorphism, I^i is $I^0(i)$ and ∂^i_I is $\partial^0_I(i)$.

In other words, up to isomorphism and degree shift, all the I^i and all the ∂^i_I are the same. (In fact, Eisenbud worked with minimal free resolutions, but using Matlis duality on his result gives Theorem 5.1.)

This phenomenon can be understood geometrically in a very simple way, using the BGG correspondence: The module M can be translated to a geometric object on $\mathbb{P}^{d-1}$, and since the Bass numbers of M are bounded, this object turns out to have zero-dimensional support. Therefore the object is stable under twisting, that is, tensoring by $\mathcal{O}_{\mathbb{P}^{d-1}}(1)$, and translating back, this gives that M is its own first syzygy, and periodicity of the minimal injective resolution follows.

In more detail, let A be the polynomial algebra $k[X_1, \dots, X_d]$ so I am in the situation of Example 3.2. In particular, $A^!$ is the exterior algebra $E = \bigwedge(Y_1, \dots, Y_d)$, and $\mathsf{QGr}(A)$ is equivalent to $\mathsf{QCoh}(\mathbb{P}^{d-1})$, with the subcategory $\mathsf{qgr}(A)$ corresponding to the subcategory $\mathsf{coh}(\mathbb{P}^{d-1})$ of coherent sheaves. Let M be in $\mathsf{gr}(E)$, and suppose that the Bass numbers

$$\mu^i(M) = \dim_k \operatorname{Ext}^i_E(k, M)$$

are bounded for $i \geq 0$.

The BGG correspondence associates to M the object

$$\mathcal{M} = \varphi(M) \in \mathsf{D}(\mathsf{QCoh}\,\mathbb{P}^{d-1}).$$

In fact, Lemma 4.3 even says that only finitely many of the cohomologies $\mathrm{h}^\ell(\mathcal{M})$ are nonzero, and that each $\mathrm{h}^\ell(\mathcal{M})$ is coherent.

For $i \geq 1$ I have

$$\begin{aligned} \mu^i(M) &= \dim_k \operatorname{Ext}^i_E(k, M) \\ &\overset{\text{(a)}}{=} \sum_j \dim_k \operatorname{Ext}^i_{\mathsf{Gr}(E)}(k, M(-i+j)) \\ &\overset{\text{(b)}}{=} \sum_j \dim_k \operatorname{Ext}^j_{\mathsf{QCoh}(\mathbb{P}^{d-1})}(\mathcal{O}_{\mathbb{P}^{d-1}}, \mathcal{M}(i-j)) \\ &= (*), \end{aligned} \tag{13}$$

where in (a), I am being clever by using the degree shift $-i+j$ instead of simply j, and where (b) is by Lemma 4.2. And for $i \gg 0$ I have

$$(*) = \sum_j \dim_k \operatorname{Hom}_{\mathsf{QCoh}(\mathbb{P}^{d-1})}(\mathcal{O}_{\mathbb{P}^{d-1}}, \mathrm{h}^j(\mathcal{M})(i-j)) \tag{14}$$

by Lemma 4.4.

It follows that if $\mu^i(M)$ is bounded for $i \geq 0$, then for each j,

$$\dim_k \operatorname{Hom}_{\mathsf{QCoh}(\mathbb{P}^{d-1})}(\mathcal{O}_{\mathbb{P}^{d-1}}, \mathrm{h}^j(\mathcal{M})(i-j))$$

is bounded for $i \geq 0$. Hence for each j,

$$\dim_k \operatorname{Hom}_{\mathsf{QCoh}(\mathbb{P}^{d-1})}(\mathcal{O}_{\mathbb{P}^{d-1}}, \mathrm{h}^j(\mathcal{M})(\ell)) \tag{15}$$

is bounded for $\ell \gg 0$. However, this is now a geometric statement: For $\ell \gg 0$, the polynomial growth rate of the numbers in equation (15) equals the dimension of the support of $\mathrm{h}^j(\mathcal{M})$ on $\mathbb{P}^{d-1}$, as follows from [Hartshorne 1977, Theorem I.7.5]. So it follows that each of the finitely many nonzero $\mathrm{h}^j(\mathcal{M})$ has zero-dimensional support; in other words, the support is a finite collection of points.

Now suppose that the ground field k is infinite. Then it is possible to pick a hyperplane H in $\mathbb{P}^{d-1}$ which is disjoint from the support of each $\mathrm{h}^j(\mathcal{M})$. To H corresponds an injection $\mathcal{O}_{\mathbb{P}^{d-1}}(1) \hookrightarrow \mathcal{O}_{\mathbb{P}^{d-1}}$ which is an isomorphism away from H. Tensoring over $\mathcal{O}_{\mathbb{P}^{d-1}}$ with $\mathcal{M}$ gives a morphism

$$\mathcal{M} \otimes \mathcal{O}_{\mathbb{P}^{d-1}}(1) \xrightarrow{\mu} \mathcal{M} \otimes \mathcal{O}_{\mathbb{P}^{d-1}},$$

and $\mathrm{h}^j(\mu)$ is $\mathrm{h}^j(\mathcal{M}) \otimes \mathcal{O}_{\mathbb{P}^{d-1}}(1) \longrightarrow \mathrm{h}^j(\mathcal{M}) \otimes \mathcal{O}_{\mathbb{P}^{d-1}}$. However, this is an isomorphism for each j because $\mathcal{O}_{\mathbb{P}^{d-1}}(1) \hookrightarrow \mathcal{O}_{\mathbb{P}^{d-1}}$ is an isomorphism away from H and hence an isomorphism on the support of each $\mathrm{h}^j(\mathcal{M})$. So μ is an isomorphism in $\mathsf{D}(\mathsf{QCoh}\,\mathbb{P}^{d-1})$, proving

$$\mathcal{M}(1) \cong \mathcal{M}.$$

Under the BGG correspondence this gives $\gamma(\mathcal{M}(1)) \cong \gamma(\mathcal{M})$, and using $\gamma(\mathcal{M}) = \gamma\varphi(M) \cong M$ and Lemma 3.4(ii) this can be rearranged as

$$\Sigma M \cong M(1) \tag{16}$$

in $\overline{\mathsf{Gr}}(E)$.

In $\overline{\mathsf{Gr}}(E)$, the suspension ΣM is computed as the first syzygy of M in an injective resolution; see Remark 2.4. So equation (16) shows that in $\overline{\mathsf{Gr}}(E)$, this first syzygy is just M itself, with a degree shift of one. It is possible to improve this with a few remarks: First, if M is without injective direct summands, then it is not hard to show that the isomorphism (16) lifts to hold in $\mathsf{Gr}(E)$, if ΣM is obtained as the first syzygy in a *minimal* injective resolution of M. Secondly, the assumption that k is infinite can be dropped using [Grothendieck 1965, Proposition 2.5.8].

Iterating equation (16) now shows that in the minimal injective resolution I of M, the syzygy $\Sigma^i M$ is simply $M(i)$. Hence the module I^i must be $I^0(i)$, and the differential ∂_I^i must be $\partial_I^0(i)$. So I have recovered Theorem 5.1.

The noncommutative case. In the above argument, the minimal injective resolution is periodic with period one because points in $\mathbb{P}^{d-1}$ are invariant under twisting. It is known that this invariance breaks down when one passes to noncommutative analogues of $\mathbb{P}^{d-1}$.

Here the twist can move points, and it is possible to have orbits of length n, for any finite n, and orbits of infinite length. So it is natural to expect that suitable noncommutative analogues of the above argument might give examples of algebras $A^!$, analogous to E, and modules M where $\mu^i(M)$ is bounded for $i \geq 0$, and yet where the minimal injective resolution of M is periodic with period n, or aperiodic. Indeed, this turns out to hold.

Note that the following example of this behaviour is already known from [Smith 1996]. But the present geometric view through the BGG correspondence is new.

Setup 5.2. Assume that the ground field k is algebraically closed. Let C be an elliptic curve over k with a line bundle $\mathcal{L}$ of degree d, and an automorphism τ given by translation by a point of C. Let A be the Sklyanin algebra associated to these data in [Smith 1996, Sec. 8].

Remark 5.3. Note that A satisfies the standing assumptions from Setup 0.1. In fact, A is a noncommutative analogue of the polynomial algebra on d variables $k[X_1, \ldots, X_d]$, and hence the Koszul dual $A^!$ is a noncommutative analogue of the exterior algebra $\bigwedge(Y_1, \ldots, Y_d)$.

Remark 5.4. The construction of A in [Smith 1996, Sec. 8] is so that the curve C sits inside $\mathbb{P}(A_1')$. So each point p on C is also a point in $\mathbb{P}(A_1')$, that is, a one dimensional subspace of A_1'. This subspace has an annihilator $p^\perp$ in A_1, and the graded A-left-module $P\langle p\rangle = A/Ap^{\perp}$ is a so-called point module. That is, it is cyclic, and each graded piece in nonnegative degrees is one dimensional.

Let me now use the functor π from Remark 1.1 to write

$$\mathcal{M}\langle p\rangle = \pi(P\langle p\rangle).$$

This is an object of $\mathsf{qgr}(A)$, and I view it as a complex concentrated in cohomological degree zero. This complex is an object of $\mathsf{D}(\mathsf{QGr}\,A)$, so finally the BGG correspondence gives the object

$$M\langle p\rangle = \gamma(\mathcal{M}\langle p\rangle)$$

in $\overline{\mathsf{Gr}}(A^!)$. In fact, $\mathcal{M}\langle p\rangle$ viewed as an object of $\mathsf{D}(\mathsf{QGr}\,A)$ is in the subcategory $\mathsf{D}^{\mathrm{f}}(\mathsf{QGr}\,A)$, so Lemma 4.3 says that $M\langle p\rangle$ is even in $\overline{\mathsf{gr}}(A^!)$.

Observe that $M\langle p\rangle$ is only well-defined up to isomorphism in $\overline{\mathsf{Gr}}(A^!)$, so when looking at $M\langle p\rangle$ as a graded $A^!$-left-module, I can drop any injective direct summands, and so assume that $M\langle p\rangle$ is without injective direct summands.

Let me start by pointing out the following property of the modules $M\langle p\rangle$.

Proposition 5.5. *The Bass numbers $\mu^i(M\langle p\rangle)$ are bounded for $i \geq 0$.*

Proof. By a computation like the one in equations (13) and (14), it follows that for $i \gg 0$ I have

$$\mu^i(M\langle p\rangle) = \sum_j \dim_k \mathrm{Hom}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathrm{h}^j(\mathcal{M}\langle p\rangle)(i-j)) = (*).$$

However, the complex $\mathcal{M}\langle p\rangle$ is just the object $\mathcal{M}\langle p\rangle$ placed in cohomological degree zero, so

$$(*) = \dim_k \mathrm{Hom}_{\mathsf{QGr}(A)}(\mathcal{O}, \mathcal{M}\langle p\rangle(i)) = (**),$$

and since $\mathcal{M}\langle p\rangle$ is $\pi(P\langle p\rangle)$ and i is large, this is

$$(**) = \dim_k P\langle p\rangle_i = 1$$

by [Artin and Zhang 1994, Theorem 8.1(1) and Proposition 3.13(2)], because the algebra A is AS regular. $\square$

Now some computations with the $M\langle p\rangle$'s.

Lemma 5.6. *The module $M\langle p\rangle$ determines p.*

Proof. It is certainly true that $M\langle p\rangle$ determines $\mathcal{M}\langle p\rangle \cong \varphi(M\langle p\rangle)$. In turn, $\mathcal{M}\langle p\rangle$ determines the tail $P\langle p\rangle_{\geq n}$ for $n \gg 0$, because when viewing $\mathcal{M}\langle p\rangle$ as an object of $\mathsf{qgr}(A)$, I have

$$P\langle p\rangle_{\geq n} \cong \omega\pi(P\langle p\rangle)_{\geq n} = \omega(\mathcal{M}\langle p\rangle)_{\geq n} \tag{17}$$

for $n \gg 0$ by [Artin and Zhang 1994, Theorem 8.1(1) and Proposition 3.13(2)]. But $P\langle p\rangle_{\geq n}$ determines p by [Smith 1996, Sec. 8]. $\square$

Lemma 5.7. *Recall d and τ from Setup 5.2. The modules $M\langle p\rangle$ satisfy*

$$\Sigma(M\langle p\rangle) \cong M\langle \tau^{2-d} p\rangle(1)$$

in $\overline{\mathsf{Gr}}(A^!)$.

Proof. In [Smith 1996, Example 9.5] is proved

$$P\langle p\rangle_{\geq 1}(1) \cong P\langle \tau^{2-d} p\rangle,$$

and applying π shows

$$\mathcal{M}\langle p\rangle(1) \cong \mathcal{M}\langle \tau^{2-d} p\rangle$$

because π only sees the tail of a module. Applying γ and Lemma 3.4(ii), this can be rearranged to the lemma's isomorphism

$$\Sigma(M\langle p\rangle) \cong M\langle \tau^{2-d} p\rangle(1).$$

$\square$

Lemma 5.8. *If $\Sigma^i(M\langle p\rangle) \cong M\langle q\rangle(j)$ holds in $\overline{\mathsf{Gr}}(A^!)$ for some points p and q on C, then $i = j$.*

Proof. The lemma's isomorphism implies $\varphi(\Sigma^i(M\langle p\rangle)) \cong \varphi(M\langle q\rangle(j))$, and using Lemma 3.4(i) and $\varphi(M\langle p\rangle) = \mathcal{M}\langle p\rangle$, this becomes $\Sigma^i(\mathcal{M}\langle p\rangle) \cong \Sigma^j(\mathcal{M}\langle q\rangle)(-j)$. Since the cohomologies of $\mathcal{M}\langle p\rangle$ and $\mathcal{M}\langle q\rangle$ are concentrated in cohomological degree zero, this is only possible with $i = j$. $\square$

Finally, these lemmas can be used as follows. If there is to be periodicity in the sense

$$\Sigma^i(M\langle p\rangle) \cong M\langle p\rangle(j) \tag{18}$$

in $\overline{\mathsf{Gr}}(A^!)$ for some i and j, then $i = j$ by Lemma 5.8. Moreover, $\Sigma^i(M\langle p\rangle) \cong M\langle \tau^{(2-d)i} p\rangle(i)$ holds by Lemma 5.7. Substituting into equation (18) gives

$$M\langle \tau^{(2-d)i} p\rangle(i) \cong M\langle p\rangle(i),$$

hence $M\langle \tau^{(2-d)i} p\rangle \cong M\langle p\rangle$, and as $M\langle p\rangle$ determines p by Lemma 5.6, this implies

$$\tau^{(2-d)i}(p) = p.$$

Conversely, $\tau^{(2-d)i}(p) = p$ gives

$$\Sigma^i(M\langle p\rangle) \cong M\langle \tau^{(2-d)i} p\rangle(i) \cong M\langle p\rangle(i)$$

in $\overline{\mathsf{Gr}}(A^!)$ by Lemma 5.7.

Summing up, if d, τ and p are so that $\tau^{(2-d)i}(p) \neq p$ for $i = 1, \ldots, n-1$ but $\tau^{(2-d)n}(p) = p$, then in $\overline{\mathsf{Gr}}(A^!)$ the suspension $\Sigma^i(M\langle p\rangle)$ is not a degree shift of $M\langle p\rangle$ for $i = 1, \ldots, n-1$, but $\Sigma^n(M\langle p\rangle)$ is $M\langle p\rangle(n)$.

And if d, τ and p are so that $\tau^{(2-d)i}(p) \neq p$ for $i \geq 1$, then in $\overline{\mathsf{Gr}}(A^!)$ the suspension $\Sigma^i(M\langle p\rangle)$ is not a degree shift of $M\langle p\rangle$ for $i \geq 1$.

Using that $M\langle p\rangle$ contains no injective direct summands, this easily lifts to give the same result in $\mathsf{Gr}(A^!)$ for syzygies in *minimal* injective resolutions. So I get the following example which shows the promised contrast to Theorem 5.1 with respect to periodicity of minimal injective resolutions.

Example 5.9. (1) Let d, τ and p be so that $\tau^{(2-d)i}(p) \neq p$ for $i = 1, \ldots, n-1$ but $\tau^{(2-d)n}(p) = p$.

Then the minimal injective resolution I of $M\langle p\rangle$ is periodic with period n, in the sense that in the resolution, the i-th syzygy $\Sigma^i(M\langle p\rangle)$ is not isomorphic to a degree shift of $M\langle p\rangle$ for $i = 1, \ldots, n-1$, but the n-th syzygy $\Sigma^n(M\langle p\rangle)$ is isomorphic to $M\langle p\rangle(n)$.

Hence up to isomorphism, I^n is $I^0(n)$ and ∂_I^n is $\partial_I^0(n)$, while the same is not true with any smaller value of n.

(2) Let d, τ and p be so that $\tau^{(2-d)i}(p) \neq p$ for $i \geq 1$.

Then the minimal injective resolution I of $M\langle p\rangle$ is aperiodic, in the sense that in the resolution, no syzygy $\Sigma^i(M\langle p\rangle)$ is a degree shift of $M\langle p\rangle$ for $i \geq 1$.

Acknowledgement

The diagrams were typeset with Paul Taylor's `diagrams.tex`.

References

[Artin and Zhang 1994] M. Artin and J. J. Zhang, "Noncommutative projective schemes", *Adv. Math.* **109**:2 (1994), 228–287. MR 96a:14004 Zbl 0833.14002

[Beilinson et al. 1996] A. Beilinson, V. Ginzburg, and W. Soergel, "Koszul duality patterns in representation theory", *J. Amer. Math. Soc.* **9**:2 (1996), 473–527. MR 96k:17010 Zbl 0864.17006

[Benson 1998] D. J. Benson, *Representations and cohomology, I: Basic representation theory of finite groups and associative algebras*, Second ed., Cambridge Stud. Adv. Math. **30**, Cambridge University Press, Cambridge, 1998. MR 99f:20001a Zbl 0908.20001

[Bernšteĭn et al. 1978] I. N. Bernšteĭn, , I. M. Gel'fand, and S. I. Gel'fand, "Algebraic vector bundles on $\mathbb{P}^n$ and problems of linear algebra", *Funktsional. Anal. i Prilozhen.* **12**:3 (1978), 66–67. MR 80c:14010a Zbl 0424.14002

[Eisenbud 2002] D. Eisenbud, "Periodic resolutions over exterior algebras", *J. Algebra* **258**:1 (2002), 348–361. MR 2004a:16009 Zbl 01868092

[Fløystad 2003] G. Fløystad, "Koszul duality and equivalences of categories", revised version, 2003. math.RA/0012264

[Grothendieck 1965] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR 33 #7330 Zbl 0135.39701

[Happel 1987] D. Happel, "On the derived category of a finite-dimensional algebra", *Comment. Math. Helv.* **62**:3 (1987), 339–389. MR 89c:16029 Zbl 0626.16008

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Grad. Texts in Math. **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Jørgensen 1999] P. Jørgensen, "Non-commutative Castelnuovo–Mumford regularity", *Math. Proc. Cambridge Philos. Soc.* **125**:2 (1999), 203–221. MR 2000h:13010 Zbl 0920.14002

[Keller 1994] B. Keller, "Deriving DG categories", *Ann. Sci. École Norm. Sup.* (4) **27**:1 (1994), 63–102. MR 95e:18010 Zbl 0799.18007

[Smith 1996] S. P. Smith, "Some finite-dimensional algebras related to elliptic curves", pp. 315–348 in *Representation theory of algebras and related topics* (Mexico City, 1994), edited by R. Bautista et al., CMS Conf. Proc. **19**, Amer. Math. Soc., Providence, RI, 1996. MR 97e:16053 Zbl 0856.16009

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Stud. Adv. Math. **38**, Cambridge University Press, Cambridge, 1994. MR 95f:18001 Zbl 0834.18001

Received November 6, 2003.

PETER JØRGENSEN
DEPARTMENT OF PURE MATHEMATICS
UNIVERSITY OF LEEDS
LEEDS LS2 9JT
UNITED KINGDOM

popjoerg@maths.leeds.ac.uk
www.maths.leeds.ac.uk/~popjoerg

PACIFIC JOURNAL OF MATHEMATICS
Vol. 218, No. 2, 2005

MODULAR DIOPHANTINE INEQUALITIES AND NUMERICAL SEMIGROUPS

J. C. ROSALES, P. A. GARCÍA-SÁNCHEZ AND J. M. URBANO-BLANCO

We study the set of integer solutions to the modular diophantine inequality $ax \bmod b \leq x$.

Introduction

Given $x \in \mathbb{Q}$, we set $\lceil x \rceil = \min\{z \in \mathbb{Z} \mid z \geq x\}$ and $\lfloor x \rfloor = \max\{z \in \mathbb{Z} \mid z \leq x\}$, as usual. Given integers m, n with $n > 0$, we set $m \bmod n = m - n\lfloor m/n \rfloor$ and $m \bmod (-n) = m \bmod n$. A *modular diophantine inequality* is an expression of the form $ax \bmod b \leq x$ with a, b integers such that $b \neq 0$. Since $ax \bmod b \geq 0$, the set S of solutions to such an inequality is contained in the set $\mathbb{N}$ of nonnegative integers. S is a *numerical semigroup*, that is, S is closed under addition, $0 \in S$ and $\mathbb{N} \setminus S$ is finite. Not every numerical semigroup arises from a modular diophantine inequality, and Section 2 presents a procedure for testing numerical semigroups for this property. Theorem 12 is crucial for obtaining this algorithm, and thus Section 1 is devoted to it. One of the main consequences of this theorem is that if the inequalities $ax \bmod b \leq x$ and $cx \bmod d \leq x$ have the same solutions, then

$$b - (a, b) - (a-1, b) = d - (c, d) - (c-1, d),$$

where (x, y) denotes the greatest common divisor of the integers x and y.

A numerical semigroup S is said to be *modular with modulus b and factor a* if $S = \{x \in \mathbb{N} \mid ax \bmod b \leq x\}$. The preceding remark ensures that $b-(a, b)-(a-1, b)$ is an invariant of S, which we call the *weight* of S and denote by $\mathrm{w}(S)$.

If S is a numerical semigroup, the largest integer not in S is called the *Frobenius number* of S and is denoted by $\mathrm{g}(S)$. This integer has been widely studied; see for instance [Brauer 1942; Brauer and Shockley 1962; Johnson 1960; Selmer 1977; Sylvester 1884; Curtis 1990; Davison 1994; Djawadi and Hofmeister 1996]. In this direction it is worth highlighting [Ramírez Alfonsín 2000; ≥ 2005], where a review of this problem is given, with many references. In the literature one can also find a large number of publications devoted to the study of one-dimensional analytically

Supported by the project BFM2000-1469.
MSC2000: 20M05, 20M14, 20M30, 13H10.

irreducible local domains via their value semigroups, which are numerical semigroups; see, for instance, [Apéry 1946; Barucci et al. 1997; Bertin and Carbonne 1977; Delorme 1976; Fröberg et al. 1987; Kunz 1970; Teissier 1973; Watanabe 1973]. As a consequence of this study, some interesting kinds of numerical semigroups arise, such as symmetric and pseudo-symmetric numerical semigroups. In Section 1 we prove that a modular numerical semigroup S is symmetric if and only if $\mathrm{w}(S) = \mathrm{g}(S)$, and pseudo-symmetric if and only if $\mathrm{g}(S) = \mathrm{w}(S) + 1$. Sections 3 and 4 are devoted to modular numerical semigroups with modulus equal to their weight plus two and three, respectively. We show that those of weight plus two are obtained from a symmetric numerical semigroup by adjoining its Frobenius number to it, and that those with weight plus three arise from a pseudo-symmetric numerical semigroup by adding to it its Frobenius number and this number divided by two.

In Section 5 we study those modular numerical semigroups S such that the factor of S divides the modulus. For these numerical semigroups we can explicitly give formulas for the multiplicity, the minimal generator set, the Apéry set and the Frobenius number, so the case $a \mid b$ is now well understood.

Section 6 addresses the problem of computing the Frobenius number in the complementary case $a \nmid b$, solving it when $(a-1)(a - (b \bmod a)) < b$. We have not been able to solve the general case.

1. Modular numerical semigroups

Let a and b be integers such that $b \neq 0$. Since $ax \bmod b = (a \bmod b)x \bmod b$ and $ax \bmod b = ax \bmod (-b)$, in order to study the solutions of $ax \bmod b \leq x$, we can assume that b is a positive integer and that $0 \leq a < b$.

Proposition 1. *The set of integer solutions of a modular diophantine inequality is a numerical semigroup.*

Proof. Let a and b be two integers such that $0 \leq a < b$ and let $S = \{x \in \mathbb{N} \mid ax \bmod b \leq x\}$. Clearly $0 \in S$, and if x is an integer greater than or equal to b, then $x \in S$. Hence $\mathbb{N} \setminus S$ is finite. For $x, y \in S$, we have $a(x+y) \bmod b \leq ax \bmod b + ay \bmod b \leq x + y$, whence $x + y \in S$, so S is closed under addition. □

A numerical semigroup S arising as in the proposition is said to be *modular*. The modular semigroup with modulus b factor a will be denoted by $\mathrm{S}(a, b)$; thus $\mathrm{S}(a, b) = \{x \in \mathbb{N} \mid ax \bmod b \leq x\}$. When we write $\mathrm{S}(a, b)$ we will generally assume tacitly that a and b are integers with $0 \leq a < b$.

Example 2. $\mathrm{S}(2, 3) = \mathrm{S}(2, 4) = \{0, 2, 3, \rightarrow\}$, where $\rightarrow$ means that all the elements beyond 3 are in the set. Thus a and b don't have to be unique.

The goal of this section is to prove Theorem 12, which counts the natural numbers *absent* from $\mathrm{S}(a,b)$. We prepare the ground with some simple results.

Lemma 3. *Let a and b be integers such that $0 \le a < b$. Then $ax \bmod b \le x$ if and only if $(b+1-a)x \bmod b \le x$.*

Proof. If $ax \bmod b \le x$, there exist $q, r \in \mathbb{N}$ such that $ax = qb + r$ with $0 \le r \le x$. Hence $(b+1-a)x = (b+1)x - ax = bx - qb + x - r$ and $(b+1-a)x \bmod b \le x - r \le x$. The converse follows by interchanging a with $b+1-a$. □

Lemma 4. *Let S be a modular numerical semigroup with modulus $b \ge 2$. Then there exists a positive integer a such that $a \le \frac{1}{2}(b+1)$ and $S = \mathrm{S}(a,b)$.*

Proof. Write $S = \mathrm{S}(a,b)$ with $0 \le a < b$. By Lemma 3, $S = \mathrm{S}(b+1-a, b)$, so if $a > \frac{1}{2}(b+1)$ we can replace a by $b+1-a \le \frac{1}{2}(b+1)$. Also if $a = 0$ we can replace it by $a = 1$, since $S = \mathbb{N}$ for both these values of a. □

Lemma 5. *Let a and b be integers such that $0 \le a < b$ and let $x \in \mathbb{N}$. Then*

$$a(b-x) \bmod b = \begin{cases} 0 & \textit{if } ax \bmod b = 0, \\ b - (ax \bmod b) & \textit{if } ax \bmod b \ne 0, \end{cases}$$

and $ax \bmod b > x$ implies that $a(b-x) \bmod b < b - x$. □

Corollary 6. *If $S = \mathrm{S}(a,b)$ and $x \in \mathbb{N} \setminus S$, then $b - x \in S$.* □

Given a subset A of $\mathbb{N}$, we denote by $\mathrm{H}(A)$ the complement $\mathbb{N} \setminus A$, and by $\langle A \rangle$ the submonoid of $\mathbb{N}$ generated by A (the set of finite sums of elements of A).

Remark 7. If $S = \mathrm{S}(a,b) \ne \mathbb{N}$ for positive a and b, then $b-1 \notin \mathrm{H}(S)$, since otherwise $b - (b-1) = 1$ would be an element of S. Moreover $x \in S$ for all integers $x \ge b$. Therefore the Frobenius number $\mathrm{g}(S)$ is at most $b-2$.

We now characterize the case $\mathrm{g}(S) = b-2$. If $\mathrm{g}(S) = b-2$, Corollary 6 implies that $b - (b-2) = 2 \in S$. Hence b is odd and $S = \langle 2, b \rangle$. In addition, since $2 \in S$, $2a \bmod b \le 2$ and this leads to $2a > b$, whence $a > \frac{1}{2}b$. But Lemma 4 says we can take $a \le \frac{1}{2}(b+1)$, which then means $a = \frac{1}{2}(b+1)$. Hence $S = \mathrm{S}(\frac{1}{2}(b+1), b)$. Conversely, if $S = \mathrm{S}(\frac{1}{2}(b+1), b)$ with b odd, it is easy to check that $S = \langle 2, b \rangle$ and thus $\mathrm{g}(S) = b-2$.

Example 8. Suppose $b \ge 2$ and $S = \mathrm{S}(2,b)$. Then $S = \{0, \lfloor \frac{1}{2}(b+1) \rfloor, \rightarrow\}$. For clearly $\{b, \rightarrow\} \subseteq S$. Now take $0 < x < b$. Then $x \in S$ if and only if $2x \bmod b \le x$. However, $2x \bmod b = 2x$ if and only if $2x < b$, and thus in this case $x \notin S$. If $2x \ge b$, then $2x \bmod b = 2x - b \le x$, whence $x \in S$.

Lemma 9. *Let $S = \mathrm{S}(a,b)$ and let x be an integer such that $0 \le x \le b$. Then x and $b - x$ are both in S if and only if $ax \bmod b \in \{0, x\}$.*

Proof. If $ax \bmod b \notin \{0, x\}$, Lemma 5 gives $a(b-x) \bmod b = b - (ax \bmod b)$. If $x \in S$, the right-hand side exceeds $b - x$ (since $ax \bmod b < x$). Thus $b - x \notin S$.

Conversely, if $ax \bmod b = 0$, clearly $x \in S$ and also $b - x \in S$ by Lemma 5; whereas if $ax \bmod b = x \neq 0$, again $x \in S$, and Lemma 5 gives $a(b-x) \bmod b = b - (ax \bmod b) = b - x$, so $b - x \in S$. □

Lemma 10. *Let a and b be positive integers and x an integer such that $0 \le x < b$.*

(1) $ax \bmod b = 0$ *if and only if x is a multiple of $b/(a, b)$.*

(2) $ax \bmod b = x$ *if and only if x is a multiple of $b/(b, a-1)$.* □

Lemma 11. *Let $S = S(a, b)$ with $0 < a < b$. Let $\alpha = (b, a-1)$ and $\beta = (b, a)$, and let x be an integer such that $0 \le x \le b$. Then*

$$\{x, b-x\} \subset S \Longleftrightarrow x \in \left\{0, \frac{b}{\alpha}, 2\frac{b}{\alpha}, \ldots (\alpha-1)\frac{b}{\alpha}, \frac{b}{\beta}, 2\frac{b}{\beta}, \ldots, (\beta-1)\frac{b}{\beta}, b\right\} =: X.$$

The cardinality of X is $\alpha + \beta$.

Proof. The equivalence is just Lemmas 9 and 10 put together. To show there is no duplication in the elements of X as written, note that $(\alpha, \beta) = 1$. If $sb/\alpha = tb/\beta$ for some $s, t \in \mathbb{N}$, then $s\beta = t\alpha = k\alpha\beta$ for some $k \in \mathbb{N}$. Hence $s = k\alpha$ and $t = k\beta$. □

Theorem 12. *Let $S = \mathrm{S}(a, b)$ for some integers $0 \le a < b$. Then*

$$\#\,\mathrm{H}(S) = \frac{b + 1 - (a, b) - (a-1, b)}{2}.$$

Here as usual # denotes cardinality.

Proof. Let α, β and X be as in Lemma 11. By Corollary 6 and Lemma 11, for $0 \le x \le b$, at most one of x, $b-x$ lies in $\mathrm{H}(S)$, and it's exactly one unless $x \in X$. Thus $\#\,\mathrm{H}(S) = \frac{1}{2}(b + 1 - \#\,X) = \frac{1}{2}(b + 1 - \alpha - \beta)$. □

Example 13. If p is an odd prime, $\#\,\mathrm{H}(\mathrm{S}(a, p)) = \frac{1}{2}(p-1)$ for all a with $1 < a < p$.

As an immediate consequence of Theorem 12 we obtain:

Corollary 14. *Suppose* $\mathrm{S}(a, b) = \mathrm{S}(c, d)$. *Then*

$$b - (a, b) - (a-1, b) = d - (c, d) - (c-1, d).$$

Example 15. The converse of Corollary 14 is false. For instance, $\langle 4, 5, 6\rangle = \mathrm{S}(3, 12) \neq \mathrm{S}(2, 10) = \langle 5, 6, 7, 8, 9\rangle$.

Recall that we have defined the *weight* of $S = \mathrm{S}(a, b)$ as $\mathrm{w}(S) := b - (a, b) - (a-1, b)$; by Theorem 12, this number equals $2\,\#\,\mathrm{H}(S)-1$, and so is an invariant of S. Note that $\mathrm{w}(\mathbb{N}) = -1$. If $S \neq \mathbb{N}$, we can choose a, b with $2 \le a < b$; hence $(a, b)+(a-1, b) \le \frac{1}{2}b + \frac{1}{3}b < b$, so $\mathrm{w}(S) \ge 1$. Thus, like the Frobenius number, the weight of a modular numerical semigroup is at least 1, except for the case $S = \mathbb{N}$, where $\mathrm{w}(S) = \mathrm{g}(S) = -1$.

Theorem 12 and the inequality $\# H(S) \geq \frac{1}{2}(\mathrm{g}(S) + 1)$, valid for *any* numerical semigroup S (see [Fröberg et al. 1987], for instance), yield:

Corollary 16. *If S is a modular numerical semigroup, then* $\mathrm{w}(S)$ *is odd and greater than or equal to* $\mathrm{g}(S)$. □

In view of this, modular numerical semigroups S with $\mathrm{w}(S) = \mathrm{g}(S)$ and $\mathrm{g}(S)$ odd, or with $\mathrm{w}(S) = \mathrm{g}(S) + 1$ and $\mathrm{g}(S)$ even, have minimal possible weight with respect to their Frobenius numbers. The next result characterizes this kind of numerical semigroup, but before proving it we need to recall some concepts.

A numerical semigroup S is *symmetric* if $x \in \mathbb{N} \setminus S$ implies $\mathrm{g}(S) - x \in S$. It is straightforward to prove that a symmetric numerical semigroup has odd Frobenius number. A numerical semigroup is *pseudo-symmetric* if $\mathrm{g}(S)$ is even and $x \in \mathbb{N} \setminus S$ implies that either $x = \mathrm{g}(S)/2$ or $\mathrm{g}(S) - x \in S$. A numerical semigroup S is symmetric if and only if $\#\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + 1)$, and pseudo-symmetric if and only if $\#\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + 2))$; see [Fröberg et al. 1987], for instance.

A numerical semigroup is *irreducible* if it cannot be expressed as the intersection of two numerical semigroups containing it properly. In [Rosales and Branco 2003] it is shown that S is irreducible if and only if S is symmetric or pseudo-symmetric (depending on the parity of $\mathrm{g}(S)$).

Corollary 17. *Let S be a modular numerical semigroup.*

(1) *S is symmetric if and only if* $\mathrm{w}(S) = \mathrm{g}(S)$.

(2) *S is pseudo-symmetric if and only if* $\mathrm{w}(S) = \mathrm{g}(S) + 1$.

Proof. S is symmetric if and only if $\#\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + 1)$. By Theorem 12, $\#\mathrm{H}(S) = \frac{1}{2}(\mathrm{w}(S) + 1)$, whence S is symmetric if and only if $\mathrm{g}(S) = \mathrm{w}(S)$. The proof of (2) is analogous. □

Example 18. If b is an odd integer, there exists a modular numerical semigroup S with $\mathrm{w}(S) = b$. It suffices to take $S = \mathrm{S}(2, b+2)$, since $\mathrm{w}(\mathrm{S}(2, b+2)) = b + 2 - (2, b+2) - (1, b+2) = b + 2 - 1 - 1 = b$.

2. Determining whether a numerical semigroup is modular

In this section we give a procedure for deciding whether a given numerical semigroup is a modular numerical semigroup, and if so to express it in the form $\mathrm{S}(a, b)$.

Lemma 19. *Let S be a modular numerical semigroup with modulus b and $S \neq \mathbb{N}$. Then $b \leq 12\#\mathrm{H}(S) - 6$.*

Proof. As we saw right after Example 15, if $a \geq 2$ we have $(a, b) + (a-1, b) \leq \frac{5}{6}b$. By Theorem 12, $\#\mathrm{H}(S) \geq \frac{1}{2}(b + 1 - \frac{5}{6}b)$ and thus $b \leq 12\#\mathrm{H}(S) - 6$. □

For a numerical semigroup S, the *multiplicity* of S, denoted by $\mathrm{m}(S)$, is the least positive integer in S. Here is an immediate consequence of Lemma 11:

Lemma 20. *For $S = \mathrm{S}(a, b)$,*

$$b - \mathrm{m}(S) \in S \iff \mathrm{m}(S) = \min\left\{\frac{b}{(a, b)}, \frac{b}{(a-1, b)}\right\}.$$

Lemma 21. *Let S be a modular numerical semigroup with modulus b. Then*

$$b \geq \mathrm{g}(S) + \mathrm{m}(S).$$

Proof. Since $1, 2, \dots, \mathrm{m}(S)-1$ are not in S, Corollary 6 ensures that $b-\mathrm{m}(S)+1, \dots, b-1$ are. But $\{b, \mathrm{m}(S)\} \subset S$, so $\{b - \mathrm{m}(S) + 1, \rightarrow\} \subseteq S$. This implies that $\mathrm{g}(S) \leq b - \mathrm{m}(S)$. □

Lemma 22. *For $S = \mathrm{S}(a, b)$,*

$$b = \mathrm{g}(S) + \mathrm{m}(S) \iff \mathrm{m}(S) \neq \min\left\{\frac{b}{(a, b)}, \frac{b}{(a-1, b)}\right\}.$$

Proof. Follows from Lemmas 20 and 21. □

Now we have all the ingredients to give the algorithm announced at the start of this section, to decide whether a numerical semigroup is of the form $S(a, b)$, and if so, produce such a pair (a, b) (or all such pairs with $a \leq \frac{1}{2}(b+1)$, if the algorithm is not stopped after the first pair is found).

Algorithm 23. Given a numerical semigroup S distinct from $\mathbb{N}$:

(1) Compute $\#\mathrm{H}(S)$, $\mathrm{g}(S)$ and $\mathrm{m}(S)$.

(2) Set $b = \mathrm{g}(S) + \mathrm{m}(S)$.

(3) For every $a \in A := \left\{ a \in \mathbb{N} \;\middle|\; \begin{array}{l} 2 \leq a \leq \frac{1}{2}(b+1), \\ b = 2\#\mathrm{H}(S) + (a, b) + (a-1, b) - 1, \\ \mathrm{m}(S) < \min\{b/(a, b), b/(a-1, b)\} \end{array} \right\}$

compute $\mathrm{S}(a, b)$; if $S = \mathrm{S}(a, b)$, return this answer and stop.

(4) Compute $B = \big\{ b \in \{k \cdot \mathrm{m}(S) \mid k \in \mathbb{N}\} \;\big|\; 2\#\mathrm{H}(S) + 1 \leq b \leq 12\#\mathrm{H}(S) - 6 \big\}$.

(5) For every $b \in B$

for every $a \in A_b := \left\{ a \in \mathbb{N} \;\middle|\; \begin{array}{l} 2 \leq a \leq \frac{1}{2}(b+1), \\ b = 2\#\mathrm{H}(S) + (a, b) + (a-1, b) - 1, \\ \mathrm{m}(S) = \min\{b/(a, b), b/(a-1, b)\} \end{array} \right\}$

compute $\mathrm{S}(a, b)$; if $S = \mathrm{S}(a, b)$, return this answer and stop.

(6) Return "S is not modular".

We briefly justify the correctness of Algorithm 23. In Steps (2) and (3) we check whether S is a modular numerical semigroup with modulus $\mathrm{g}(S)+\mathrm{m}(S)$, and the correct working of these steps relies on Lemmas 4 and 22 and Theorem 12. If S is not a modular numerical semigroup with modulus $\mathrm{g}(S)+\mathrm{m}(S)$, Lemma 22 gives $\mathrm{m}(S)=\min\{b/(a,b), b/(a-1,b)\}$. This implies that $\mathrm{m}(S)$ divides b. Theorem 12 states that $b=2\,\#\,\mathrm{H}(S)+(a,b)+(a-1,b)-1$, so $b\geq 2\,\#\,\mathrm{H}(S)+1$; at the same time $b\leq 12\,\#\,\mathrm{H}(S)-6$ by Lemma 19. Therefore Steps (4) and (5) cover the case $b\neq \mathrm{g}(S)+\mathrm{m}(S)$.

Example 24. Let $S=\langle 3,5\rangle$. Then $\#\,\mathrm{H}(S)=4$, $\mathrm{g}(S)=7$ and $\mathrm{m}(S)=3$. In Step (2) we get $b=10$. Step (3) yields $A=\{2,3,4\}$, then $\mathrm{S}(2,10)=\langle 5,6,7,8,9\rangle$, $\mathrm{S}(3,10)=\langle 4,5,7\rangle$, and $\mathrm{S}(4,10)=\langle 3,5\rangle=S$, so the algorithm returns $S=S(4,10)$.

Example 25. Let $S=\langle 3,8,10\rangle$. In this case $\#\,\mathrm{H}(S)=5$, $\mathrm{g}(S)=7$ and $\mathrm{m}(S)=3$. In Step (2) we obtain $b=10$ and in Step (3), $A=\varnothing$. The only nonempty set A_b with $b\in B$ is $A_{15}=\{5\}$. Since $S\neq \mathrm{S}(5,15)=\langle 3,7,11\rangle$, the algorithm returns No.

Example 26. Let $S=\langle 10,11,12\rangle$. Then $\#\,\mathrm{H}(S)=25$, $\mathrm{g}(S)=49$ and $\mathrm{m}(S)=10$. In Step (2) we obtain $b=59$ and A is empty. Computing B, we obtain

$$B=\{60,70,80,90,100,110,120,130,140,150,160,170,180,\\ 190,200,210,220,230,240,250,260,270,280,290\}.$$

The only nonempty set A_b with $b\in B$ is $A_{60}=\{6\}$. It turns out that $S=\mathrm{S}(6,60)$.

Remark 27. If the input to Algorithm 23 is known to be symmetric, the procedure can be improved, because if $S=\mathrm{S}(a,b)$ is symmetric then b must be equal to $\mathrm{g}(S)+(a,b)+(a-1,b)$ (note that $\mathrm{w}(S)=\mathrm{g}(S)$ by Corollary 17). A similar argument applies to the pseudo-symmetric case.

Remark 28. The intersection $\bigcap_{i=1}^{n}\mathrm{S}(a_i,b_i)$ of $n\geq 1$ modular numerical semigroups is a numerical semigroup; it need not be modular, as can be seen from Example 25, since we can write $\langle 3,8,10\rangle=\langle 3,4\rangle\cap\langle 3,5\rangle=\mathrm{S}(3,8)\cap\mathrm{S}(4,10)$.

Nor can every numerical semigroup be written as such an intersection: for instance, $\langle 7,8,10,13\rangle$ is a symmetric, hence irreducible, numerical semigroup; thus it cannot be an intersection of modular numerical semigroups other than by being itself a modular numerical semigroup. Algorithm 23 says that it is not.

3. Modular numerical semigroups whose modulus is its weight plus two

We now study modular numerical semigroups $S=\mathrm{S}(a,b)$ whose modulus b equals $\mathrm{w}(S)+2$. Since $b=\mathrm{w}(S)+(a,b)+(a-1,b)\geq \mathrm{w}(S)+2$, the condition $b=\mathrm{w}(S)+2$ is equivalent to $(a,b)=(a-1,b)=1$ (so b is odd), and it characterizes modular numerical semigroups whose moduli are minimal with respect to their weights.

Every numerical semigroup S is finitely generated (as an additive monoid). This is easy to see — for instance, start with two relatively prime $r, s \in S$ and then adjoin all elements of $S \cap \{0, 1, \dots, rs-1\}$ as yet unaccounted for. Among all generating sets one can of course choose one that is *minimal*, say $\mathcal{M}(S)$. A minute's thought shows that $\mathcal{M}(S)$ is characterized by containing exactly those nonzero elements of S that cannot be expressed as a sum of two nonzero elements of S:

$$\mathcal{M}(S) = (S \setminus \{0\}) \setminus \big((S \setminus \{0\}) + (S \setminus \{0\})\big).$$

In particular, $\mathcal{M}(S)$ is unique. We set $\mathrm{e}(S) = \#\mathcal{M}(S)$ and call this number the *embedding dimension* of S; the elements of $\mathcal{M}(S)$ are called *minimal generators*.

Proposition 29. *Let* $S = \mathrm{S}(a, b)$ *with* $2 \le a < b$ *and* $(a, b) = (a-1, b) = 1$. *Then*

(1) $b = \mathrm{g}(S) + \mathrm{m}(S)$,

(2) $\#\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + \mathrm{m}(S) - 1)$,

(3) *b is the largest minimal generator of S.*

Proof. (1) We already know that $b-1 \in S$ when $2 \le a < b$. Hence $\mathrm{m}(S) \ne b$. Using Lemma 22, we get $b = \mathrm{g}(S) + \mathrm{m}(S)$.

(2) Immediate from Theorem 12.

(3) First we prove that b is a minimal generator of S. Assume to the contrary that $b = x + y$ with $x, y \in S \setminus \{0\}$. Then $ax \bmod b \le x$ and $ay \bmod b \le y$, and thus $(ax \bmod b) + (ay \bmod b) \le x + y = b$. Since $a(x + y) \bmod b = ab \bmod b = 0$, we deduce that $(ax \bmod b) + (ay \bmod b) \in \{0, b\}$. Thus either $ax \bmod b = x$ and $ay \bmod b = y$, or $ax \bmod b = 0$ and $ay \bmod b = 0$. These two cases contradict the two halves of Lemma 10.

To see that b is the largest minimal generator, take $x \in S$ with $x > b$. By applying (1) we obtain $x > \mathrm{g}(S) + \mathrm{m}(S)$, which implies that $x - \mathrm{m}(S) > \mathrm{g}(S)$; this forces $x - \mathrm{m}(S) \in S$. Thus $x = \mathrm{m}(S) + (x - \mathrm{m}(S))$ cannot be a minimal generator of S. □

Proposition 29 allows us to relate the modular numerical semigroups in question with *unitary extensions of symmetric numerical semigroups* or *UESY-semigroups* in short. A numerical semigroup S is a UESY-semigroup if there exists a symmetric numerical semigroup S' such that $S' \subset S$ and $\#(S \setminus S') = 1$. In [Rosales ≥ 2005b] this condition is shown to be equivalent to the existence of a symmetric numerical semigroup S' such that $S = S' \cup \{\mathrm{g}(S')\}$. The following result also appears there.

Proposition 30. *Let S be a numerical semigroup*, $S \ne \mathbb{N}$. *The following conditions are equivalent*:

(1) *S is a UESY-semigroup.*

(2) $\#\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + \mathrm{m}(S) - 1)$ *and* $\mathrm{g}(S) + \mathrm{m}(S)$ *is a minimal generator of S.* □

A *pseudo-Frobenius number* [Rosales and Branco 2002] of a numerical semigroup S is an integer $x \notin S$ such that $x+s \in S$ for all $s \in S \setminus \{0\}$. The set of pseudo-Frobenius numbers of S is denoted by $\mathrm{Pg}(S)$, and its cardinality, called the *type* of S, is denoted by $\mathrm{t}(S)$. Clearly $\mathrm{g}(S) \in \mathrm{Pg}(S)$. Moreover S is symmetric if and only if $\mathrm{Pg}(S) = \{\mathrm{g}(S)\}$, and S is pseudo-symmetric if and only if $\mathrm{Pg}(S) = \{\mathrm{g}(S), \frac{1}{2}\mathrm{g}(S)\}$; see [Barucci et al. 1997; Fröberg et al. 1987], for instance.

In [Rosales $\geq$ 2005b] it is proved that if S is a UESY-semigroup distinct from $\mathbb{N}$, then $\mathrm{t}(S) = \mathrm{e}(S)-1$. This, plus Propositions 29 and 30, gives:

Corollary 31. *Let $S = \mathrm{S}(a, b)$ be such that $2 \leq a < b$ and $(a, b) = (a-1, b) = 1$. Then $\mathrm{t}(S) = \mathrm{e}(S)-1$ and there exists a symmetric numerical semigroup S' such that $S = S' \cup \{\mathrm{g}(S')\}$.* □

Theorem 32. *Let $S = \mathrm{S}(a, b)$. Then $b = \mathrm{w}(S) + 2$ if and only if S is a UESY-semigroup and b is a minimal generator of S.*

Proof. If $b = \mathrm{w}(S)+2 = b-(a, b)-(a-1, b)+2$, we deduce $(a, b) = (a-1, b) = 1$. Corollary 31 then says that S is a UESY-semigroup, and Proposition 29 that b is a minimal generator of S.

Conversely, if b is a minimal generator of S it equals $\mathrm{g}(S) + \mathrm{m}(S)$, by Lemma 21 and the fact, shown in the proof of Proposition 29, that a minimal generator of S cannot exceed $\mathrm{g}(S)+\mathrm{m}(S)$. If S is a UESY, then, $\#\,\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S)+m(S)-1)$ by Proposition 30 and $\#\,\mathrm{H}(S) = \frac{1}{2}(\mathrm{w}(S)+1)$ by Theorem 12. Thus $b = \mathrm{w}(S)+2$. □

Corollary 33. *Let S be a modular numerical semigroup with modulus b. Then $b = \mathrm{w}(S) + 2$ if and only if $S \setminus \{b\}$ is a symmetric numerical semigroup. Therefore, if b is a prime integer, $S \setminus \{b\}$ is a symmetric numerical semigroup.*

Proof. If $b = \mathrm{w}(S)+2$, Theorem 32 says b is a minimal generator of S, so $S' = S \setminus \{b\}$ is a numerical semigroup with $\mathrm{g}(S') = b$. By Corollary 6, S' is symmetric.

Conversely, if $S \setminus \{b\}$ is a symmetric numerical semigroup, then S is a UESY-semigroup with b as a minimal generator. Now Theorem 32 gives $b = \mathrm{w}(S)+2$.

Finally, b prime implies $(a, b) = (a-1, b) = 1$, so $\mathrm{w}(S) = b-2$. □

Corollary 34. *Let $b \geq 3$ be an integer. Then b is prime if and only if b is the largest minimal generator of $\mathrm{S}(a, b)$ for every a such that $2 \leq a \leq \sqrt{b}$.*

Proof. If b is prime Proposition 29 applies; this proves one direction. Conversely, suppose b is not a prime — say $b = ac$ with integers $a, c \geq 2$ and $a \leq \sqrt{b}$. For $S = \mathrm{S}(a, b)$, we have $ac \bmod b = 0$ and thus $c \in S$. But then $b = ac$ cannot be a minimal generator of S. □

4. Modular numerical semigroups whose modulus is its weight plus three

We now study modular numerical semigroups $S = \mathrm{S}(a, b)$ such that $b = \mathrm{w}(S)+3$; this condition is equivalent to $(a, b) + (a-1, b) = 3$. There are two cases:

- $(a, b) = 1$ and $(a-1, b) = 2$.
- $(a, b) = 2$ and $(a-1, b) = 1$.

In both situations b must be even and by Corollary 6 we deduce that $\frac{1}{2}b \in S$.

Let S be a numerical semigroup with minimal generating set $\{n_1, \dots, n_p\}$. We say that $x \in S$ has a *unique expression* if the equality $x = a_1n_1 + \cdots + a_pn_p$, with $a_1, \dots, a_p \in \mathbb{N}$, determines $a_1, \dots, a_p$ uniquely.

Proposition 35. *Let* $S = \mathrm{S}(a, b)$ *be such that* $2 \le a < b$ *and* $(a, b) + (a-1, b) = 3$.

(1) $\mathrm{m}(S) \neq \frac{1}{2}b \Leftrightarrow S \neq \{0, \frac{1}{2}b, \to\} \Leftrightarrow b = \mathrm{g}(S)+\mathrm{m}(S) \Leftrightarrow \#\,\mathrm{H}(S) = \dfrac{\mathrm{g}(S)+\mathrm{m}(S)-2}{2}$.

(2) $\frac{1}{2}b$ *is a minimal generator of* S.

(3) b *has a unique expression in* S.

Proof. (1) Follows easily from Corollary 6, Lemma 22 and Theorem 12.

(2) Suppose $x + y = \frac{1}{2}b$ with $x, y \in S$. Then $ax \bmod b \le x$ and $ay \bmod b \le y$, whence $ax \bmod b + ay \bmod b \le x + y = \frac{1}{2}b$. Thus $\frac{1}{2}ab \bmod b = a(x+y) \bmod b = ax \bmod b + ay \bmod b$. We must show that $x = 0$ or $y = 0$. We distinguish two cases. If $(a, b) = 2$, then $\frac{1}{2}ab \bmod b = 0$, so $ax \bmod b = 0$ and $ay \bmod b = 0$; then Lemma 10 shows that both x and y are multiples of $\frac{1}{2}b$, which leads to the desired conclusion. Similarly, if $(a-1, b) = 2$, then $\frac{1}{2}ab \bmod b = \frac{1}{2}b$, so $ax \bmod b = x$ and $ay \bmod b = y$; Lemma 10 again shows that x and y are multiples of $\frac{1}{2}b$.

(3) We prove that if $x, y \in S\backslash\{0\}$ are such that $x+y = b$, then $x = y = \frac{1}{2}b$. Arguing as in the proof of Proposition 29(3), we see that either $(ax \bmod b, ay \bmod y) = (x, y)$ or $ax \bmod b = ay \bmod y = 0$. Lemma 10 implies that x and y are both multiples of $\frac{1}{2}b$, and since $x \neq 0 \neq y$, we conclude that $x = y = \frac{1}{2}b$. ☐

Paralleling what we did in Section 3 for the case $b = \mathrm{w}(S) + 2$, we can use Proposition 35 to relate modular numerical semigroups such that $b = \mathrm{w}(S) + 3$ with a previous studied class of numerical semigroups. A numerical semigroup S is called a *PESPY-semigroup* if there exists a pseudo-symmetric numerical semigroup S' such that $S = S' \cup \left\{\frac{1}{2}\mathrm{g}(S'), \mathrm{g}(S')\right\}$ (the two additional elements are the pseudo-Frobenius numbers of S'; see [Barucci et al. 1997; Fröberg et al. 1987]).

Numerical semigroups of the form $\{0, x, \to\}$ with x a positive integer are called *intervals*. The following result appears in [Rosales $\ge$ 2005a].

Proposition 36. *Let* S *be a numerical semigroup that is not an interval. The following conditions are equivalent*:

(1) S *is a PEPSY-semigroup.*

(2) $\#\,\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + \mathrm{m}(S) - 2)$, $\frac{1}{2}(\mathrm{g}(S) + \mathrm{m}(S))$ *is a minimal generator of* S *and* $\mathrm{g}(S) + \mathrm{m}(S)$ *is an element of unique expression of* S. ☐

The next result is an immediate consequence of Propositions 35 and 36.

Corollary 37. *Let* $S = \mathrm{S}(a, b)$ *be such that* $2 \le a < b$, $(a, b) + (a-1, b) = 3$ *and* S *is not an interval. Then* S *is a PEPSY-semigroup.*

In [Rosales ≥ 2005a] it is proved that if S is a PEPSY-semigroup that is not an interval, then $\mathrm{t}(S) = \mathrm{e}(S)-1$. Thus:

Corollary 38. *Let* $S = \mathrm{S}(a, b)$ *be such that* $2 \le a < b$, $(a, b) + (a-1, b) = 3$ *and* S *is not an interval. Then* $\mathrm{t}(S) = \mathrm{e}(S)-1$.

Remark 39. Among numerical semigroups, interval semigroups have maximal embedding dimension relative to multiplicity: $\mathrm{e}(S) = \mathrm{m}(S)$. For any numerical semigroup with maximal embedding dimension, $\mathrm{t}(S) = \mathrm{m}(S)-1 = \mathrm{e}(S)-1$ (see [Barucci et al. 1997], for instance). Hence the assumption "S is not an interval" can be dropped from Corollary 38.

Theorem 40. *Assume that* $S = \mathrm{S}(a, b)$ *is not an interval. Then* $b = \mathrm{w}(S) + 3$ *if and only if* S *is a PEPSY-semigroup,* $\frac{1}{2}b$ *is a minimal generator of* S *and* b *has a unique expression in* S.

Proof. Necessity follows from Corollary 37 and Proposition 35. Sufficiency: Lemma 21 says that $b \ge \mathrm{g}(S)+\mathrm{m}(S)$. If $b > \mathrm{g}(S)+\mathrm{m}(S)$, then $\mathrm{m}(S)+(b-\mathrm{m}(S))$ and $\frac{1}{2}b + \frac{1}{2}b$ are distinct expressions for b in S ($\mathrm{m}(S) \ne \frac{1}{2}b$ since otherwise S is an interval, by Corollary 6). Therefore $b = \mathrm{g}(S) + \mathrm{m}(S)$. By Proposition 36, we know that $\#\,\mathrm{H}(S) = \frac{1}{2}(\mathrm{g}(S) + \mathrm{m}(S) - 2)$ and Theorem 12 ensures that $\#\,\mathrm{H}(S) = \frac{1}{2}(\mathrm{w}(S) + 1)$, whence $b = \mathrm{g}(S) + \mathrm{m}(S) = \mathrm{w}(S) + 3$. □

Corollary 41. *Let* S *be a modular numerical semigroup with modulus* b. *Then* $b = \mathrm{w}(S)+3$ *if and only if* $S \setminus \left\{\frac{1}{2}b, b\right\}$ *is a pseudo-symmetric numerical semigroup. Therefore, if* $b = 2p$ *and* $a < p$ *for some positive prime* p, *then* $S \setminus \left\{\frac{1}{2}b, b\right\}$ *is a pseudo-symmetric numerical semigroup.*

Proof. Suppose $b = \mathrm{w}(S) + 3$. By Theorem 40, $\frac{1}{2}b$ is a minimal generator of S and b has a unique expression in S. This implies that $S' = S \setminus \left\{\frac{1}{2}b, b\right\}$ is a numerical semigroup, and clearly $\mathrm{g}(S') = b$. Using Corollary 6 we can easily deduce that S' is pseudo-symmetric.

Conversely, if $S \setminus \left\{\frac{1}{2}b, b\right\}$ is a pseudo-symmetric numerical semigroup, then S is a PEPSY-semigroup by definition, $\frac{1}{2}b$ is a minimal generator of S and $b = \frac{1}{2}b + \frac{1}{2}b$ is the unique expression of b in S. Thus $b = \mathrm{w}(S) + 3$ by Theorem 40. □

5. When the factor divides the modulus

We next focus on numerical semigroups of the form $S = \mathrm{S}(a, ab)$, where we may as well assume $a, b > 1$. First a general definition: given a numerical semigroup

S and $n \in S \setminus \{0\}$, the *Apéry set* of n in S [Apéry 1946] is

$$\mathrm{Ap}(S, n) = \{s \in S \mid s - n \notin S\}.$$

This set always has n elements $w(0) = 0$, $w(1)$, ..., $w(n-1)$, where $w(i)$ is the least element congruent to i modulo n. Note also that $x \in \mathbb{Z}$ is an element of S if and only if $x \geq w(x \bmod n)$. Consequently

$$(*) \qquad \mathrm{g}(S) = \max(\mathrm{Ap}(S, n)) - n.$$

The following result is a consequence of [Rosales 1996, Lemma 3.3] and gives a characterization of Apéry sets which will be useful later.

Lemma 42. *Let $m > 0$ be an integer and let $X = \{0 = w(0), w(1), \dots, w(m-1)\}$ be a subset of $\mathbb{N}$ such that $i < w(i) \equiv i \bmod m$ for all $i \in \{1, \dots, m-1\}$. Let S be the submonoid of $\mathbb{N}$ generated by $X \cup \{m\}$. Then S is a numerical semigroup with multiplicity m. Moreover,* $\mathrm{Ap}(S, m) = X$ *if and only if for all $i, j \in \{1, \dots, m-1\}$ there exist $k \in \{0, \dots, m-1\}$ and $t \in \mathbb{N}$ such that $w(i) + w(j) = w(k) + tm$.* □

Getting back to $S = \mathrm{S}(a, ab)$, with $a, b > 1$, we will give a description of the particular Apéry set $\mathrm{Ap}(S, \mathrm{m}(S))$ in terms of a, b, and this will lead to an explicit formula for the Frobenius number of S. We also show how the minimal generating set for such numerical semigroups can be computed from a and b as well as the corresponding sets of pseudo-Frobenius numbers.

Lemma 43. $\mathrm{m}(\mathrm{S}(a, ab)) = b$.

Proof. Let $S = \mathrm{S}(a, ab)$ and let $x \in \{1, \dots, b-1\}$. Then $ax < ab$ and thus $ax \bmod ab = ax > x$, whence $x \notin S$. Clearly $b \in S$ and consequently $\mathrm{m}(S) = b$. □

Theorem 44. $\mathrm{Ap}(\mathrm{S}(a, ab), b) = \{0, k_1 b + 1, k_2 b + 2, \dots, k_{b-1} b + b-1\}$, *where* $k_i = \lceil (a-1)i/b \rceil$ *for all* $i \in \{1, \dots, b-1\}$.

Proof. Let S' be the semigroup generated by $\{b, k_1 b+1, \dots, k_{b-1} b + b-1\}$. Since $k_i \geq 1$ for all $i \in \{1, \dots, b-1\}$ we have $\mathrm{m}(S') = b$. Clearly $k_1 \leq \cdots \leq k_{b-1}$ and $k_i + k_j \geq k_{i+j}$ for all $i, j \in \{1, \dots, b-1\}$ with $2 \leq i + j \leq b-1$. Using Lemma 42, we deduce that $\mathrm{Ap}(S', b) = \{0, k_1 b + 1, \dots, k_{b-1} b + b-1\}$. Recall that $x \in \mathbb{Z}$ belongs to S' if and only if $x \geq k_{x \bmod b} b + x \bmod b$, since this latter number is the element in $\mathrm{Ap}(S', b)$ that is congruent to x modulo b. So, for x an integer we have $x \in S' \iff \lfloor x/b \rfloor \geq k_{x \bmod b} \iff \lfloor x/b \rfloor \geq \lceil (a-1)(x \bmod b)/b \rceil \iff \lfloor x/b \rfloor \geq (a-1)(x \bmod b)/b \iff \lfloor x/b \rfloor b \geq (a-1)(x \bmod b) \iff x - (x \bmod b) \geq (a-1)(x \bmod b) \iff a(x \bmod b) \leq x \iff ax \bmod ab \leq x$. Thus $S' = \mathrm{S}(a, ab)$. □

Using this result and equality $(*)$ with $n = \mathrm{m}(S)$, we obtain:

Corollary 45. $\mathrm{g}(\mathrm{S}(a, ab)) = \left\lceil (b-1)(a-1)/b \right\rceil b - 1$.

Particularizing the formula given in Theorem 12 for the case at hand, we get

$$\#\,\mathrm{H}(\mathrm{S}(a, ab)) = \frac{a(b-1) - (a-1, b) + 1}{2}.$$

Minimal generators. We next turn our attention to the minimal generating set $\{n_0 < n_1 < \cdots < n_p\}$ of $\mathrm{S}(a, ab)$. We know that $n_0 = b$, by Lemma 43; our goal is to describe the remaining minimal generators.

Lemma 46. *Let x and y be positive integers. Then $\lceil x/b \rceil + \lceil y/b \rceil = \lceil (x+y)/b \rceil$ if and only if $x \equiv 0 \bmod b$ or $y \equiv 0 \bmod b$ or $(x \bmod b) + (y \bmod b) > b$.* □

Remark 47. If S is any numerical semigroup and $m \in S \setminus \{0\}$, then S is generated by $X = (\mathrm{Ap}(S, m) \setminus \{0\}) \cup \{m\} = \{m, w(1), \dots, w(m-1)\}$, and the minimal generating set of S is $X \setminus (X + X)$. Now, in the case of $S = \mathrm{S}(a, ab)$, Theorem 44 says that $\mathrm{Ap}(S, b) = \{0, k_1 b + 1, \dots, a k_{b-1} b + b-1\}$, with $k_i = \lceil (a-1)i/b \rceil$ for all $i \in \{1, \dots, b-1\}$. Thus $k_t b + t$ is a minimal generator of S if and only if $k_t \neq k_i + k_{t-i}$ for all $i \in \{1, \dots, t-1\}$.

Lemma 48. *Let $S = \mathrm{S}(a, ab)$ with $a, b > 1$, set $k_i = \lceil (a-1)i/b \rceil$ for all $i \in \{1, \dots, b-1\}$ and take $t \in \{1, \dots, b-1\}$.*

(i) *If $t < b/(a-1, b)$, then $k_t b + t$ is a minimal generator of S if and only if $(a-1)i \bmod b < (a-1)t \bmod b$ for all $i \in \{1, \dots, t-1\}$.*

(ii) *If $t > b/(a-1, b)$, then $k_t b + t$ is not a minimal generator of S.*

(iii) *If $t = b/(a-1, b)$, then $k_t b + t$ is a minimal generator of S.*

Proof. Using Lemma 46 and Remark 47, we see that $k_t b + t$ is a minimal generator of S if and only if $(a-1)i \not\equiv 0 \bmod b$ and $(a-1)i \bmod b + (a-1)(t-i) \bmod b \le b$ for all $i \in \{1, \dots, t-1\}$. Observe that

$$(\dagger) \qquad \frac{b}{(a-1, b)} = \frac{\mathrm{lcm}(a-1, b)}{a-1} = \min\{i \mid (a-1)i \bmod b = 0\}.$$

(i) From the foregoing we deduce that if $t < b/(a-1, b)$, then $k_t b + t$ is a minimal generator of S if and only if $(a-1)i \bmod b + (a-1)(t-i) \bmod b \le b$ for all $i \in \{1, \dots, t-1\}$. If $(a-1)i \bmod b + (a-1)(t-i) \bmod b = b$, then $(a-1)t \bmod b = 0$, which is impossible in view of (†), since $t < b/(a-1, b)$. Hence $k_t b + t$ is a minimal generator of S if and only if for all $i \in \{1, \dots, t-1\}$ one has $(a-1)i \bmod b + (a-1)(t-i) \bmod b < b$, which is equivalent to $(a-1)i \bmod b + (a-1)(t-i) \bmod b = (a-1)t \bmod b$. Since $(a-1)(t-i) \bmod b \neq 0$, we conclude that $k_t b + t$ is a minimal generator of S if and only if $(a-1)i \bmod b < (a-1)t \bmod b$ for all $i \in \{1, \dots, t-1\}$.

(ii) Let $i = b/(a-1, b)$. Then $(a-1)i \equiv 0 \bmod b$ and in view of Lemma 46 we get $k_i + k_{t-i} = k_t$, which implies that $k_t b + b$ is not a minimal generator of S.

(iii) In this setting $(a-1)t \bmod b = 0$ and $(a-1)i \bmod b \neq 0$ for all $i \in \{1, \ldots, t-1\}$. Hence for every $i \in \{1, \ldots, t-1\}$ one gets $(a-1)i \bmod b + (a-1)(t-i) \bmod b = b$, and by Lemma 46 we deduce that $k_t \neq k_i + k_{t-i}$ for any $i \in \{1, \ldots, t-1\}$. Therefore $k_t b + t$ is a minimal generator of S. □

Lemma 48 yields an explicit description of the minimal generating set of S:

Theorem 49. *Let $S = \mathrm{S}(a, ab)$ with $a, b > 1$, and set $k_i = \lceil (a-1)i/b \rceil$ for $i \in \{1, \ldots, b-1\}$.*

(1) *If $(b, a-1) = 1$, the minimal generating set of S is $\{b, k_{t_1}b+t_1, \ldots, k_{t_r}b+t_r\}$, where $\{t_1, \ldots, t_r\} = \big\{t \in \{1, \ldots, b-1\} \bigm| (a-1)i \bmod b < (a-1)t \bmod b$ for all $i \in \{1, \ldots, t-1\}\big\}$.*

(2) *If $(b, a-1) \neq 1$, let $t_{r+1} = b/(b, a-1)$. Then the minimal generating set of S is $\{b, k_{t_1}b + t_1, \ldots, k_{t_r}b + t_r, k_{t_{r+1}}b + t_{r+1}\}$, where $\big\{t_1, \ldots, t_r\} = \{t \in \{1, \ldots, t_{r+1}-1\} \bigm| (a-1)i \bmod b < (a-1)t \bmod b$ for all $i \in \{1, \ldots, t-1\}\big\}$.* □

Example 50. Let $S = \mathrm{S}(5, 35)$. Applying Theorem 49(1) with $a = 5$ and $b = 7$, we see that $\{t_1, \ldots, t_r\} = \{1, 3, 5\}$ (observe that 1 is always in $\{t_1, \ldots, t_r\}$), and that S is minimally generated by $\{7, 8, 17, 26\}$.

Example 51. Let $S = \mathrm{S}(5, 30)$. Applying Theorem 49(2) with $a = 5$ and $b = 6$, we see that $t_{r+1} = 3$, $\{t_1, \ldots, t_r\} = \{1\}$, and S is minimally generated by $\{6, 7, 15\}$.

Corollary 52. *Let $S = \mathrm{S}(a, ab)$ with $a, b > 1$. Set $k_i = \lceil (a-1)i/b \rceil$ for $i \in \{1, \ldots, b-1\}$, and*

$$t = \begin{cases} \min\{x \in \mathbb{N} \mid (a-1)x \equiv b-1 \bmod b\} & \text{if } (b, a-1) = 1, \\ b/(b, a-1) & \text{if } (b, a-1) \neq 1. \end{cases}$$

Then $k_t b + t$ is the greatest minimal generator of S. □

Corollary 53. *Let $a \geq 3$ and let b be a positive integer. Then $\mathrm{e}(\mathrm{S}(a, ab)) \geq \lfloor b/(a-1) \rfloor + 1$.*

Proof. The integer b is always a minimal generator of $\mathrm{S}(a, ab)$. Also, if $(a-1)t \leq b$, then by Lemma 48, $k_t b + t$ is a minimal generator of S. □

Pseudo-Frobenius numbers. For any numerical semigroup S, we define an order $\leq_S$ on S as follows: $a \leq_S b$ if $b - a \in S$. Given a subset A of S, denote by $\mathrm{Max}_{\leq_S} A$ the set of maximal elements of A with respect to $\leq_S$. The following result appears in [Rosales and Branco 2002].

Lemma 54. *Let S be any numerical semigroup with multiplicity m. If*

$$\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, m)) = \{w_{i_1}, \ldots, w_{i_t}\},$$

the pseudo-Frobenius numbers of S (page 387) *are precisely $w_{i_1} - m, \ldots, w_{i_t} - m$.*

Note that if $w, w' \in \mathrm{Ap}(S, m)$ and $w - w' \in S$, this forces $w - w'$ to be in $\mathrm{Ap}(S, m)$ as well. Hence

$$\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, m)) = \{w \in \mathrm{Ap}(S, m) \mid w + w' \notin \mathrm{Ap}(S, m) \text{ for all } 0 \neq w' \in \mathrm{Ap}(S, m)\}.$$

Let $S = \mathrm{S}(a, ab)$ with $a, b > 1$. Our aim is to compute the set $\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b))$ and thus, in view of Lemma 54, the pseudo-Frobenius set $\mathrm{Pg}(S)$.

Remark 55. By Theorem 44, $k_i b + i \notin \mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b))$ if and only if there exists $j \in \{1, \ldots, b-1\}$ such that $i + j \leq b-1$ and $k_i + k_j = k_{i+j}$. Minimal generators are $\leq_S$-minimal elements of $\mathrm{Ap}(S, b)$, which is why the condition just stated is similar (dual) to the one presented on the previous page for minimal generators.

Theorem 56. *Let a and b be two integers greater than one, and let $S = \mathrm{S}(a, ab)$. Let $k_i = \lceil (a-1)i/b \rceil$ for $i \in \{1, \ldots, b-1\}$. Then $k_i b + i \in \mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b))$ if and only if one of the following conditions hold*:

(i) $(a-1)i \equiv 0 \bmod b$ *and* $i = b-1$,

(ii) $(a-1)i \not\equiv 0 \bmod b$ *and for all* $t \in \{i + 1, \ldots, b-1\}$, *either* $(a-1)i \bmod b < (a-1)t \bmod b$ *or* $(a-1)t \bmod b = 0$.

Proof. Assume that $(a-1)i \equiv 0 \bmod b$ and $i < b-1$. Then by Lemma 46, we deduce that $k_i + k_1 = k_{i+1}$ and thus $k_i b + i \notin \mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b))$. If $(a-1)i \not\equiv 0 \bmod b$, then by Lemma 46 we have $k_i b_i + i \in \mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b))$ if and only if for all $t \in \{i + 1, \ldots, b-1\}$ we have $(a-1)(t - i) \not\equiv 0 \bmod b$ and $(a-1)i \bmod b + (a-1)(t - i) \bmod b \leq b$. If $(a-1)i \bmod b + (a-1)(t - i) \bmod b < b$, then $(a-1)i \bmod b + (a-1)(t - i) \bmod b = (a-1)t \bmod b$ and thus $(a-1)i \bmod b < (a-1)t \bmod b$. If $(a-1)i \bmod b + (a-1)(t-i) \bmod b = b$, then $(a-1)t \bmod b = 0$.

To prove the converse, assume $k_i b + i \notin \mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b))$. Then there exists $t \in \{1 + i, \ldots, b-1\}$ such that $k_i + k_{t-i} = k_t$. By using Lemma 46, we deduce that $(a-1)i \equiv 0 \bmod b$ or $(a-1)(t-i) \equiv 0 \bmod b$ or $(a-1)i \bmod b + (a-1)(t-i) \bmod b > b$. If $(a-1)i \equiv 0 \bmod b$, then i must be equal to $b-1$, but this is impossible since $t \in \{i + 1, \ldots, b-1\}$. If $(a-1)(t - i) \equiv 0 \bmod b$, then $(a-1)i \bmod b = (a-1)t \bmod b$, which is also impossible by hypothesis. Finally if $(a-1)i \bmod b + (a-1)(t-i) \bmod b > b$, then $(a-1)t \bmod b = (a-1)i \bmod b + (a-1)(t-i) \bmod b - b < (a-1)i \bmod b$, leading again to a contradiction. □

Example 57. Let $S = \mathrm{S}(5, 30)$. Applying Theorem 56 we get $\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, 6)) = \{29\}$, which by Lemma 54 means that $\mathrm{Pg}(S) = \{23\}$. Thus $\mathrm{S}(5, 30)$ is symmetric.

Proposition 58. *Let $S = \mathrm{S}(a, ab)$ with $a, b > 1$.*

(1) *S is symmetric if and only if* $(a-1, b) + (a-1) \bmod b = b$.

(2) *S is pseudo-symmetric if and only if* $(a-1, b) + (a-1) \bmod b = b + 1$.

Proof. (1) Combining Corollaries 45 and 17(1), we see that S is symmetric if and only if $\lceil (b-1)(a-1)/b \rceil b - 1 = ab - a - (a-1, b)$. The left-hand side can be written as $\big(a-1-\lfloor (a-1)/b \rfloor\big) b - 1 = (a-1)b - \lfloor (a-1)/b \rfloor b - 1 = ab - b - (a-1-(a-1) \bmod b) - 1$. Thus S is symmetric if and only if $(a-1) \bmod b + (a-1, b) = b$.

(2) As above, but this time using Corollary 17(2). □

Corollary 59. *Let k be a positive integer and let b be a multiple of k. Then* $\mathrm{S}(b-k+1+bn, (b-k+1+bn)b)$ *is symmetric for all* $n \in \mathbb{N}$. □

The pseudo-symmetric case is completely different:

Corollary 60. $\mathrm{S}(a, ab)$ *is not pseudo-symmetric for any choice of* $a, b > 1$.

Proof. Set $q = \lfloor (a-1)/b \rfloor$ and choose $u, v \in \mathbb{Z}$ such that $(a-1, b) = u(a-1) + vb$. If $\mathrm{S}(a, ab)$ is pseudo-symmetric, we have $(a-1, b) + (a-1) \bmod b = b+1$, hence $u(a-1) + vb + (a-1) - qb = b + 1$, or yet $(u+1)(a-1) + (v-q-1)b = 1$. But this implies $(a-1, b) = 1$ and hence $1 + (a-1) \bmod b = b+1$, an impossibility. □

Some families. We now present some families of numerical semigroups of the form $\mathrm{S}(a, ab)$ with $a, b > 1$ such that $(a-1, b) = 1$. For these families we can compute the minimal generating set and pseudo-Frobenius numbers explicitly. As a consequence of Theorems 49 and 56 one gets:

Proposition 61. *Let* $S = \mathrm{S}(a, ab)$ *with* $a, b > 1$ *and* $(a-1, b) = 1$. *Set* $k_i = \lceil (a-1)i/b \rceil$ *for* $i \in \{1, \dots, b-1\}$ *and take* $t \in \{1, \dots, b-1\}$.

(1) $k_t b + t$ *is a minimal generator of S if and only if* $(a-1)i \bmod b < (a-1)t \bmod b$ *for all* $i \in \{1, \dots, t-1\}$.

(2) $k_t b + t \in \mathrm{Max}_{\le_S}(\mathrm{Ap}(S, b))$ *if and only if* $(a-1)t \bmod b < (a-1)i \bmod b$ *for all* $i \in \{t+1, \dots, b-1\}$.

Let S_n be the symmetric group in n elements $\{1, \dots, n\}$, and for k relatively prime to $n+1$, define the permutation $\sigma_{k,n+1} \in \mathrm{S}_n$ by $\sigma(i) = ki \bmod (n+1)$ for $i = 1, \dots, n$. Such a permutation is called *modular*. Next, given any permutation $\sigma \in \mathrm{S}_n$, set

$$\mathrm{E}(\sigma) = \{t \in \{1, \dots, n\} \mid \sigma(i) < \sigma(t) \text{ for all } i \in \{1, \dots, t-1\}\},$$

$$\mathrm{T}(\sigma) = \{t \in \{1, \dots, n\} \mid \sigma(t) < \sigma(i) \text{ for all } i \in \{t+1, \dots, n\}\}.$$

With this notation we can rewrite Proposition 61 as follows.

Corollary 62. *Let* $S = \mathrm{S}(a, ab)$ *with* $a, b > 1$ *and* $(a-1, b) = 1$. *Then*

$$\mathrm{e}(S) = \#\mathrm{E}(\sigma_{a-1,b}) + 1 \quad \text{and} \quad \mathrm{t}(S) = \#\mathrm{T}(\sigma_{a-1,b}).$$

The minimal generating set of S is $\{b\} \cup \big\{ \lceil (a-1)i/b \rceil b + i \;\big|\; i \in \mathrm{E}(\sigma_{a-1,b}) \big\}$, *and*

$$\mathrm{Max}_{\le_S}(\mathrm{Ap}(S, b)) = \big\{ \lceil (a-1)i/b \rceil b + i \;\big|\; i \in \mathrm{T}(\sigma_{a-1,b}) \big\}.$$

Example 63. Let $S = \mathrm{S}(6, 42)$. Apply Corollary 62 with $a = 6$ and $b = 7$. Clearly $\sigma_{5,7} = (154623)$, $\mathrm{E}(\sigma_{5,7}) = \{1, 4\}$ and $\mathrm{T}(\sigma_{5,7}) = \{3, 6\}$. Hence $\mathrm{e}(S) = 3$ and $\mathrm{t}(S) = 2$. The set $\{7, \lceil (5\times 1)/7\rceil 7+1, \lceil (5\times 4)/7\rceil 7+4\} = \{7, 8, 25\}$ is a minimal generating set of S and $\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, 7)) = \{\lceil (5\times 3)/7\rceil 7+3, \lceil (5\times 6)/7\rceil 7+6\} = \{24, 41\}$.

Corollary 64. *Let $S = \mathrm{S}((b-1) + bn, ((b-1) + bn)b)$ with $n \in \mathbb{N}$ and $b \geq 5$ odd. Then S is minimally generated by $\left\{b, (n+1)b+1, \left(\frac{b-1}{2} + n\frac{b+1}{2}\right) b + \frac{b+1}{2}\right\}$, and*

$$\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b)) = \left\{\left(\tfrac{b-1}{2} + n\tfrac{b-1}{2}\right) b + \tfrac{b-1}{2}, ((b-2) + n(b-1))b + b-1\right\}.$$

Proof. Since $(b-2+bn, b) = (b-2, b) = 1$, we can apply Corollary 62. By inspection we see that $\mathrm{E}(\sigma_{b-2,b}) = \{1, (b+1)/2\}$ and $\mathrm{T}(\sigma_{b-2,b}) = \{(b-1)/2, b-1\}$. We can conclude the proof using Corollary 62, taking into account that

$$\left\lceil \frac{((b-2)+bn)1}{b} \right\rceil = n+1, \quad \left\lceil \frac{((b-2)+bn)(b\pm 1)/2}{b} \right\rceil = \frac{b-1}{2} + n\frac{b\pm 1}{2}, \quad \text{and}$$

$$\left\lceil \frac{((b-2)+bn)(b-1)}{b} \right\rceil = (b-2) + n(b-1). \qquad \square$$

Corollary 65. *Let b be an integer greater than or equal to two and let $n \in \mathbb{N}$. Then $S = \mathrm{S}((n+1)b, (n+1)b^2)$ is minimally generated by $\{b, (n+1)b+1\}$ and* $\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b)) = \{(n+1)(b-1)b + b-1\}$.

Proof. Use Corollary 62 and the fact that $\sigma_{(n+1)b-1,b} = \sigma_{b-1,b}$ swaps i and $b-i$. $\square$

Corollary 66. *Let $S = \mathrm{S}(2 + nb, (2 + nb)b)$ with $n \in \mathbb{N}$ and $b \geq 2$. Then S is minimally generated by*

$$X = \{b, (n+1)b+1, (2n+1)b+2, \dots, ((b-1)n+1)b + b-1\}$$

and $\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b)) = X \setminus \{b\}$.

Proof. Use Corollary 62 and the fact that $\sigma_{1+nb,b} = \sigma_{1,b}$ is the identity. $\square$

Corollary 67. *Let $S = \mathrm{S}(3 + nb, (3 + nb)b)$ with and $n \in \mathbb{N}$ $b \geq 3$ odd. Then S is minimally generated by $\left\{b, (n+1)b+1, (2n+1)b+2, \dots, \left(\frac{b-1}{2}n+1\right) + \frac{b-1}{2}\right\}$ and*

$$\mathrm{Max}_{\leq_S}(\mathrm{Ap}(S, b)) = \left\{\left(\tfrac{b+1}{2}n+2\right) b + \tfrac{b+1}{2}, \dots, ((b-1)n+2)b + b-1\right\}.$$

Proof. By considering $\sigma_{2+bn,b} = \sigma_{2,b}$ we see that $\mathrm{E}(\sigma_{2,b}) = \left\{1, \dots, \frac{1}{2}(b-1)\right\}$ and $\mathrm{T}(\sigma_{2,b}) = \left\{\frac{1}{2}(b+1), \dots, b-1\right\}$. Using Corollary 62, the proof follows easily from

$$\left\lceil \frac{(2+bn)i}{b} \right\rceil b = \begin{cases} (ni+1)b+i & \text{if } i \leq \frac{1}{2}(b-1), \\ (ni+2)b+i & \text{if } i \geq \frac{1}{2}(b+1). \end{cases} \qquad \square$$

6. The Frobenius number in other special cases

In Section 5 we studied $\mathrm{S}(a, b)$ with $a \mid b$. We now give some partial results for the Frobenius number in the complementary case, $a \nmid b$. We are able to find the number when $(a-1)(a-(b \bmod a)) < b$. We use without further comment the fact that, for q a rational number and x a positive integer, $x < \lceil q \rceil$ implies $x < q$.

Lemma 68. *Let $S = \mathrm{S}(a, b)$ with $0 < a < b$ and $b \bmod a \neq 0$. Then*

$$\mathrm{g}(\mathrm{S}(a, b)) \leq b - \lceil b/a \rceil.$$

Proof. Let x be a positive integer. If $x < \lceil b/a \rceil$, then $x < b/a$ and thus $ax \bmod b = ax > x$. Hence $x \notin S$ and in view of Corollary 6, this leads to $b - x \in S$. As $y \in S$ for all $y \geq b$, we conclude that $\mathrm{g}(S) \leq b - \lceil b/a \rceil$. □

Lemma 69. *Let a and b be positive integers such that $a < b$ and $b \bmod a \neq 0$. Then $a\lceil b/a \rceil \bmod b = a - (b \bmod a)$.* □

Proposition 70. *Let a and b be positive integers such that $a < b$ and $b \bmod a \neq 0$. Then $\mathrm{g}(\mathrm{S}(a, b)) = b - \lceil b/a \rceil$ if and only if $(a-1)(a-(b \bmod a)) < b$.*

Proof. Let $S = \mathrm{S}(a, b)$. From Lemma 68 we deduce that $\mathrm{g}(S) = b - \lceil b/a \rceil$ if and only if $b - \lceil b/a \rceil \notin S$, or in other words, $a(b - \lceil b/a \rceil) \bmod b > b - \lceil b/a \rceil$. This by Lemma 69 is equivalent to $((b \bmod a) - a) \bmod b > b - \lceil b/a \rceil$, and this condition holds if and only if $b + (b \bmod a) - a > b - \lfloor b/a \rfloor - 1$. Hence $\mathrm{g}(S) = b - \lceil b/a \rceil$ if and only if $\lfloor b/a \rfloor + 1 + (b \bmod a) > a$, or equivalently $(b - (b \bmod a))/a + 1 + (b \bmod a) > a$, and this holds if and only if $b > (a-1)(a-(b \bmod a))$. □

Corollary 71. *Let a and b be positive integers such that $a < b$,$b \bmod a \neq 0$ and $(a-1)(a-(b \bmod a)) < b$. Then $\mathrm{m}(\mathrm{S}(a, b)) = \lceil b/a \rceil$.*

Proof. Let $S = \mathrm{S}(a, b)$. By Proposition 70, we know that $\mathrm{g}(S) = b - \lceil b/a \rceil$. Thus $b - \lceil b/a \rceil \notin S$ and thus by Corollary 6, $\lceil b/a \rceil = b - (b - \lceil b/a \rceil) \in S$. Besides, if x is a positive integer such that $x < \lceil b/a \rceil$, then $x < b/a$, whence $ax \bmod b = ax > x$ and thus $x \notin S$. Therefore $\mathrm{m}(S) = \lceil b/a \rceil$. □

Though we have given an explicit formula for $\mathrm{g}(\mathrm{S}(a, b))$ for several cases, we have not been able to find such a formula for arbitrary positive integers a and b. We propose this as an open question.

Problem 1. Find a formula for $\mathrm{g}(\mathrm{S}(a, b))$ with a and b positive integers.

Acknowledgments

We thank J. I. García-García and the referee for their comments and suggestions.

References

[Apéry 1946] R. Apéry, "Sur les branches superlinéaires des courbes algébriques", *C. R. Acad. Sci. Paris* **222** (1946), 1198–1200. MR 8,221a Zbl 0061.35404

[Barucci et al. 1997] V. Barucci, D. E. Dobbs, and M. Fontana, *Maximality properties in numerical semigroups and applications to one-dimensional analytically irreducible local domains*, Mem. Amer. Math. Soc. **598**, Amer. Math. Soc., Providence, 1997. MR 97g:13039 Zbl 0868.13003

[Bertin and Carbonne 1977] J. Bertin and P. Carbonne, "Semi-groupes d'entiers et application aux branches", *J. Algebra* **49**:1 (1977), 81–95. MR 58 #27957 Zbl 0498.14016

[Brauer 1942] A. Brauer, "On a problem of partitions", *American J. Math.* **64** (1942), 299–312. MR 3,270d Zbl 0061.06801

[Brauer and Shockley 1962] A. Brauer and J. E. Shockley, "On a problem of Frobenius", *J. Reine Angew. Math.* **211** (1962), 215–220. MR 26 #6113 Zbl 0108.04604

[Curtis 1990] F. Curtis, "On formulas for the Frobenius number of a numerical semigroup", *Math. Scand.* **67**:2 (1990), 190–192. MR 92e:11019 Zbl 0734.11009

[Davison 1994] J. L. Davison, "On the linear Diophantine problem of Frobenius", *J. Number Theory* **48**:3 (1994), 353–363. MR 95j:11033 Zbl 0805.11025

[Delorme 1976] C. Delorme, "Sous-monoïdes d'intersection complète de N.", *Ann. Sci. École Norm. Sup.* (4) **9**:1 (1976), 145–154. MR 53 #10821 Zbl 0325.20065

[Djawadi and Hofmeister 1996] M. Djawadi and G. Hofmeister, "Linear Diophantine problems", *Arch. Math. (Basel)* **66**:1 (1996), 19–29. MR 96j:11029 Zbl 0854.11016

[Fröberg et al. 1987] R. Fröberg, C. Gottlieb, and R. Häggkvist, "On numerical semigroups", *Semigroup Forum* **35**:1 (1987), 63–83. MR 88d:20092 Zbl 0614.10046

[Johnson 1960] S. M. Johnson, "A linear diophantine problem", *Canad. J. Math.* **12** (1960), 390–398. MR 22 #12074 Zbl 0096.02803

[Kunz 1970] E. Kunz, "The value-semigroup of a one-dimensional Gorenstein ring", *Proc. Amer. Math. Soc.* **25** (1970), 748–751. MR 42 #263 Zbl 0197.31401

[Ramírez Alfonsín 2000] J. L. Ramírez Alfonsín, "The Diophantine Frobenius problem", preprint 00893, Forschungsinstitut für Diskrete Mathematik, Bonn, 2000.

[Ramírez Alfonsín $\geq$ 2005] J. L. Ramírez Alfonsín, "The Diophantine Frobenius problem". In preparation.

[Rosales 1996] J. C. Rosales, "On numerical semigroups", *Semigroup Forum* **52**:3 (1996), 307–318. MR 96m:20092 Zbl 0853.20041

[Rosales $\geq$ 2005a] J. C. Rosales, "Adding or removing an element from a pseudo-symmetric numerical semigroup", preprint.

[Rosales $\geq$ 2005b] J. C. Rosales, "Numerical semigroups that differ from a symmetric numerical semigroup in one element", to appear in *Algebra Colloquium*.

[Rosales and Branco 2002] J. C. Rosales and M. B. Branco, "Numerical semigroups that can be expressed as an intersection of symmetric numerical semigroups", *J. Pure Appl. Algebra* **171**:2-3 (2002), 303–314. MR 2003b:20089 Zbl 1006.20043

[Rosales and Branco 2003] J. C. Rosales and M. B. Branco, "Irreducible numerical semigroups", *Pacific J. Math.* **209**:1 (2003), 131–143. MR 2004b:20091 Zbl 02101065

[Selmer 1977] E. S. Selmer, "On the linear Diophantine problem of Frobenius", *J. Reine Angew. Math.* **293/294** (1977), 1–17. MR 56 #246 Zbl 0349.10009

[Sylvester 1884] J. Sylvester, "Mathematical questions with their solutions", *Educational Times* **41** (1884), 21.

[Teissier 1973] B. Teissier, "Appendice" in *Le problème des modules pour les branches planes*, École Polytechnique, Paris, 1973. MR 54 #2662 Zbl 0317.14004

[Watanabe 1973] K. Watanabe, "Some examples of one dimensional Gorenstein domains", *Nagoya Math. J.* **49** (1973), 101–109. MR 47 #6689 Zbl 0257.13024

Received March 20, 2003. Revised March 16, 2004.

J. C. Rosales
Departamento de Álgebra
Universidad de Granada
E-18071 Granada
Spain

jrosales@ugr.es

P. A. García-Sánchez
Departamento de Álgebra
Universidad de Granada
E-18071 Granada
Spain

pedro@ugr.es

J. M. Urbano-Blanco
Departamento de Álgebra
Universidad de Granada
E-18071 Granada
Spain

jurbano@ugr.es

CONTENTS

Volume 218, no. 1 and no. 2

Guidelines for Authors

Authors may submit manuscripts for publication to any of the editors. Submission of a manuscript acknowledges that the paper is original and has not been submitted elsewhere. Information regarding the preparation of manuscripts is provided below. For further information, write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Electronic submissions are strongly encouraged; the editors' email addresses are listed inside the front cover. Papers submitted in hard copy should be sent in triplicate and authors should keep a copy.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 218 No. 2 February 2005

0030-8730(200502)218:2;1-K