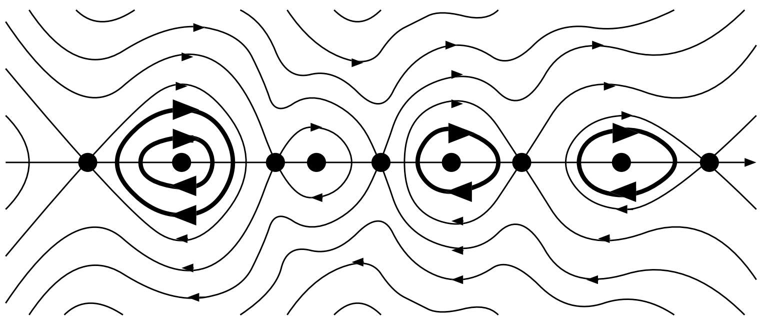
Documenta Mathematica

Gegründet 1996 durch die

Deutsche Mathematiker-Vereinigung



Phase portrait, cf. page 467

Band 9 · 2004

Documenta Mathematica veröffentlicht Forschungsarbeiten aus allen mathematischen Gebieten und wird in traditioneller Weise referiert.

Artikel können als TEX-Dateien per E-Mail bei einem der Herausgeber eingereicht werden. Hinweise für die Vorbereitung der Artikel können unter der unten angegebenen WWW-Adresse gefunden werden.

Documenta Mathematica publishes research manuscripts out of all mathematical fields and is refereed in the traditional manner.

Manuscripts should be submitted as TEX -files by e-mail to one of the editors. Hints for manuscript preparation can be found under the following WWW-address.

`http://www.math.uni-bielefeld.de/documenta`

ISSN 1431-0635 (Print), ISSN 1431-0643 (Internet)

SPARC
Leading Edge

Documenta Mathematica is a Leading Edge Partner of SPARC, the Scholarly Publishing and Academic Resource Coalition of the Association of Research Libraries (ARL), Washington DC, USA.

Address of Technical Managing Editor: Ulf Rehmann, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Typesetting in TEX.

DOCUMENTA MATHEMATICA

BAND 9, 2004

Tropical Convexity

Mike Develin and Bernd Sturmfels

Received: August 28, 2003
Revised: January 21, 2004

Communicated by Günter Ziegler

Abstract. The notions of convexity and convex polytopes are introduced in the setting of tropical geometry. Combinatorial types of tropical polytopes are shown to be in bijection with regular triangulations of products of two simplices. Applications to phylogenetic trees are discussed.

2000 Mathematics Subject Classification: 52A30; 92B10

1 Introduction

The *tropical semiring* $(\mathbb{R}, \oplus, \odot)$ is the set of real numbers with the arithmetic operations of *tropical addition*, which is taking the minimum of two numbers, and *tropical multiplication*, which is ordinary addition. Thus the two arithmetic operations are defined as follows:

$$a \oplus b \;:=\; \min(a,b) \qquad \text{and} \qquad a \odot b \;:=\; a + b.$$

The n-dimensional space $\mathbb{R}^n$ is a semimodule over the tropical semiring, with tropical addition

$$(x_1, \ldots, x_n) \oplus (y_1, \ldots, y_n) \quad = \quad (x_1 \oplus y_1, \ldots, x_n \oplus y_n),$$

and tropical scalar multiplication

$$c \odot (x_1, x_2, \ldots, x_n) \quad = \quad (c \odot x_1, c \odot x_2, \ldots, c \odot x_n).$$

The semiring $(\mathbb{R}, \oplus, \odot)$ and its semimodule $\mathbb{R}^n$ obey the usual distributive and associative laws.
The purpose of this paper is to propose a tropical theory of convex polytopes. Convexity in arbitrary idempotent semimodules was introduced by Cohen,

Gaubert and Quadrat [3] and Litvinov, Maslov and Shpiz [13]. Some of our results (such as Theorem 23 and Propositions 20 and 21) are known in a different guise in idempotent analysis. Our objective is to provide a combinatorial approach to convexity in the tropical semiring which is consistent with the recent developments in tropical algebraic geometry (see [15], [18], [20]). The connection to tropical methods in representation theory (see [12], [16]) is less clear and deserves further study.

There are many notions of discrete convexity in the computational geometry literature, but none of them seems to be quite like tropical convexity. For instance, the notion of *directional convexity* studied by Matoušek [14] has similar features but it is different and much harder to compute with.

A subset S of $\mathbb{R}^n$ is called *tropically convex* if the set S contains the point $a \odot x \,\oplus\, b \odot y$ for all $x, y \in S$ and all $a, b \in \mathbb{R}$. The *tropical convex hull* of a given subset $V \subset \mathbb{R}^n$ is the smallest tropically convex subset of $\mathbb{R}^n$ which contains V. We shall see in Proposition 4 that the tropical convex hull of V coincides with the set of all tropical linear combinations

$$a_1 \odot v_1 \oplus a_2 \odot v_2 \oplus \cdots \oplus a_r \odot v_r \,, \text{where } \; v_1, \ldots, v_r \in V \; \text{ and } \; a_1, \ldots, a_r \in \mathbb{R}. \quad (1)$$

Any tropically convex subset S of $\mathbb{R}^n$ is closed under tropical scalar multiplication, $\mathbb{R} \odot S \subseteq S$. In other words, if $x \in S$ then $x + \lambda(1, \ldots, 1) \in S$ for all $\lambda \in \mathbb{R}$. We will therefore identify the tropically convex set S with its image in the $(n-1)$-dimensional *tropical projective space*

$$\mathbb{TP}^{n-1} \quad = \quad \mathbb{R}^n/(1, \ldots, 1)\mathbb{R}.$$

Basic properties of (tropically) convex subsets in $\mathbb{TP}^{n-1}$ will be presented in Section 2. In Section 3 we introduce tropical polytopes and study their combinatorial structure. A *tropical polytope* is the tropical convex hull of a finite subset V in $\mathbb{TP}^{n-1}$. Every tropical polytope is a finite union of convex polytopes in the usual sense: given a set $V = \{v_1, \ldots, v_n\}$, their convex hull has a natural decomposition as a polyhedral complex, which we call the *tropical complex* generated by V. The following main result will be proved in Section 4:

THEOREM 1. *The combinatorial types of tropical complexes generated by a set of r vertices in $\mathbb{TP}^{n-1}$ are in natural bijection with the regular polyhedral subdivisions of the product of two simplices $\Delta_{n-1} \times \Delta_{r-1}$.*

This implies a remarkable duality between tropical $(n-1)$-polytopes with r vertices and tropical $(r-1)$-polytopes with n vertices. Another consequence of Theorem 1 is a formula for the f-vector of a generic tropical complex. In Section 5 we discuss applications of tropical convexity to phylogenetic analysis, extending known results on injective hulls of finite metric spaces (cf. [7], [8], [9] and [20]).

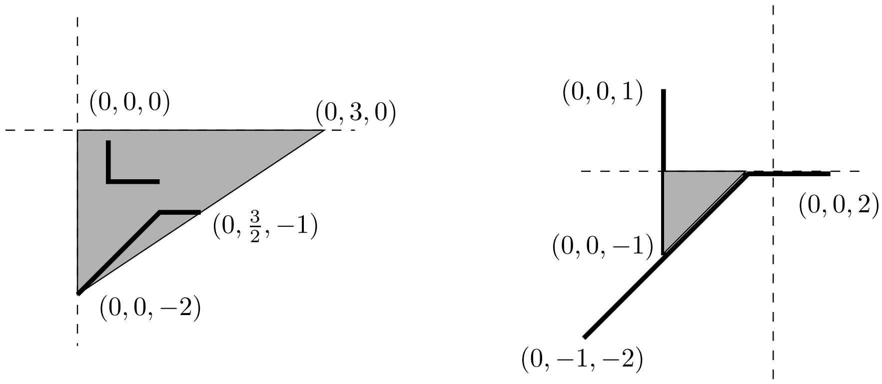


Figure 1: Tropical convex sets and tropical line segments in $\mathbb{TP}^2$.

2 Tropically convex sets

We begin with two pictures of tropical convex sets in the tropical plane $\mathbb{TP}^2$. A point $(x_1, x_2, x_3) \in \mathbb{TP}^2$ is represented by drawing the point with coordinates $(x_2 - x_1, x_3 - x_1)$ in the plane of the paper. The triangle on the left hand side in Figure 1 is tropically convex, but it is not a tropical polytope because it is not the tropical convex hull of finitely many points. The thick edges indicate two tropical line segments. The picture on the right hand side is a *tropical triangle*, namely, it is the tropical convex hull of the three points $(0, 0, 1)$, $(0, 2, 0)$ and $(0, -1, -2)$ in the tropical plane $\mathbb{TP}^2$. The thick edges represent the tropical segments connecting any two of these three points.

We next show that tropical convex sets enjoy many of the features of ordinary convex sets.

Theorem 2. *The intersection of two tropically convex sets in $\mathbb{R}^n$ or in $\mathbb{TP}^{n-1}$ is tropically convex. The projection of a tropically convex set onto a coordinate hyperplane is tropically convex. The ordinary hyperplane $\{x_i - x_j = l\}$ is tropically convex, and the projection map from this hyperplane to $\mathbb{R}^{n-1}$ given by eliminating x_i is an isomorphism of tropical semimodules. Tropically convex sets are contractible spaces. The Cartesian product of two tropically convex sets is tropically convex.*

Proof. We prove the statements in the order given. If S and T are tropically convex, then for any two points $x, y \in S \cap T$, both S and T contain the tropical line segment between x and y, and consequently so does $S \cap T$. Therefore $S \cap T$ is tropically convex by definition.

Suppose S is a tropically convex set in $\mathbb{R}^n$. We wish to show that the image of S under the coordinate projection $\phi : \mathbb{R}^n \to \mathbb{R}^{n-1}$, $(x_1, x_2, \ldots, x_n) \mapsto (x_2, \ldots, x_n)$ is a tropically convex subset of $\mathbb{R}^{n-1}$. If $x, y \in S$ then we have the obvious identity

$$\phi\big(c \odot x \,\oplus\, d \odot y\big) \quad = \quad c \odot \phi(x) \,\oplus\, d \odot \phi(y).$$

This means that ϕ is a homomorphism of tropical semimodules. Therefore, if S contains the tropical line segment between x and y, then $\phi(S)$ contains the tropical line segment between $\phi(x)$ and $\phi(y)$ and hence is tropically convex. The same holds for the induced map $\phi : \mathbb{TP}^{n-1} \to \mathbb{TP}^{n-2}$.

Most ordinary hyperplanes in $\mathbb{R}^n$ are not tropically convex, but we are claiming that hyperplanes of the special form $x_i - x_j = k$ are tropically convex. If x and y lie in that hyperplane then $x_i - y_i = x_j - y_j$. This last equation implies the following identity for any real numbers $c, d \in \mathbb{R}$:

$$(c \odot x \oplus d \odot y)_i - (c \odot x \oplus d \odot y)_j = \min{(x_i{+}c, y_i{+}d)} - \min{(x_j{+}c, y_j{+}d)} = k.$$

Hence the tropical line segment between x and y also lies in the hyperplane $\{x_i - x_j = k\}$.

Consider the map from $\{x_i - x_j = k\}$ to $\mathbb{R}^{n-1}$ given by deleting the i-th coordinate. This map is injective: if two points differ in the x_i coordinate they must also differ in the x_j coordinate. It is clearly surjective because we can recover an i-th coordinate by setting $x_i = x_j + k$. Hence this map is an isomorphism of $\mathbb{R}$-vector spaces and it is also an isomorphism of $(\mathbb{R}, \oplus, \odot)$-semimodules.

Let S be a tropically convex set in $\mathbb{R}^n$ or $\mathbb{TP}^{n-1}$. Consider the family of hyperplanes $H_l = \{x_1 - x_2 = l\}$ for $l \in \mathbb{R}$. We know that the intersection $S \cap H_l$ is tropically convex, and isomorphic to its (convex) image under the map deleting the first coordinate. This image is contractible by induction on the dimension n of the ambient space. Therefore, $S \cap H_l$ is contractible. The result then follows from the topological result that if S is connected, which all tropically convex sets obviously are, and if $S \cap H_l$ is contractible for each l, then S itself is also contractible.

Suppose that $S \subset \mathbb{R}^n$ and $T \subset \mathbb{R}^m$ are tropically convex. Our last assertion states that $S \times T$ is a tropically convex subset of $\mathbb{R}^{n+m}$. Take any (x, y) and (x', y') in $S \times T$ and $c, d \in \mathbb{R}$. Then

$$c \odot (x, y) \,\oplus\, d \odot (x', y') \quad = \quad \big(\, c \odot x \oplus d \odot x' \,,\, c \odot y \oplus d \odot y' \,\big)$$

lies in $S \times T$ since S and T are tropically convex. $\square$

We next give a more precise description of what tropical line segments look like.

PROPOSITION 3. *The tropical line segment between two points x and y in $\mathbb{TP}^{n-1}$ is the concatenation of at most $n-1$ ordinary line segments. The slope of each line segment is a zero-one vector.*

Proof. After relabeling the coordinates of $x = (x_1, , \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, we may assume

$$y_1 - x_1 \leq y_2 - x_2 \leq \cdots \leq y_n - x_n. \tag{2}$$

The following points lie in the given order on the tropical segment between x and y:

$$\begin{aligned}
x = (y_1 - x_1) \odot x \oplus y &= \big(y_1,\, y_1{-}x_1{+}x_2, \ldots, y_1{-}x_1{+}x_{n-1}, y_1{-}x_1{+}x_n\big)\\
(y_2 - x_2) \odot x \oplus y &= \big(y_1,\, y_2,\, y_2{-}x_2{+}x_3, \ldots, y_2{-}x_2{+}x_{n-1}, y_2{-}x_2{+}x_n\big)\\
(y_3 - x_3) \odot x \oplus y &= \big(y_1,\, y_2,\, y_3,\, \ldots, y_3{-}x_3{+}x_{n-1}, y_3{-}x_3{+}x_n\big)\\
\vdots \qquad & \qquad \vdots \qquad\quad \vdots\\
(y_{n-1}{-}x_{n-1}) \odot x \oplus y &= \big(y_1,\, y_2,\, y_3,\, \ldots, y_{n-1}, y_{n-1}{-}x_{n-1}{+}x_n\big)\\
y = (y_n - x_n) \odot x \oplus y &= \big(y_1,\, y_2,\, y_3,\, \ldots,\, y_{n-1},\, y_n\big).
\end{aligned}$$

Between any two consecutive points, the tropical line segment agrees with the ordinary line segment, which has slope $(0, 0, \ldots, 0, 1, 1, \ldots, 1)$. Hence the tropical line segment between x and y is the concatenation of at most $n - 1$ ordinary line segments, one for each strict inequality in (2). □

This description of tropical segments shows an important feature of tropical polytopes: their edges use a limited set of directions. The following result characterizes the *tropical convex hull*.

Proposition 4. *The smallest tropically convex subset of $\mathbb{TP}^{n-1}$ which contains a given set V coincides with the set of all tropical linear combinations (1). We denote this set by* $\mathrm{tconv}(V)$.

Proof. Let $x = \bigoplus_{i=1}^{r} a_i \odot v_i$ be the point in (1). If $r \leq 2$ then x is clearly in the tropical convex hull of V. If $r > 2$ then we write $x = a_1 \odot v_1 \oplus (\bigoplus_{i=2}^{r} a_i \odot v_i)$. The parenthesized vector lies the tropical convex hull, by induction on r, and hence so does x. For the converse, consider any two tropical linear combinations $x = \bigoplus_{i=1}^{r} c_i \odot v_i$ and $y = \bigoplus_{j=1}^{r} d_i \odot v_i$. By the distributive law, $a \odot x \oplus b \odot y$ is also a tropical linear combination of $v_1, \ldots, v_r \in V$. Hence the set of all tropical linear combinations of V is tropically convex, so it contains the tropical convex hull of V. □

If V is a finite subset of $\mathbb{TP}^{n-1}$ then $\mathrm{tconv}(V)$ is a *tropical polytope*. In Figure 2 we see three small examples of tropical polytopes. The first and second are tropical convex hulls of three points in $\mathbb{TP}^2$. The third tropical polytope lies in $\mathbb{TP}^3$ and is the union of three squares.

One of the basic results in the usual theory of convex polytopes is Carathéodory's theorem. This theorem holds in the tropical setting.

Proposition 5 (Tropical Carathéodory's Theorem). *If x is in the tropical convex hull of a set of r points v_i in $\mathbb{TP}^{n-1}$, then x is in the tropical convex hull of at most n of them.*

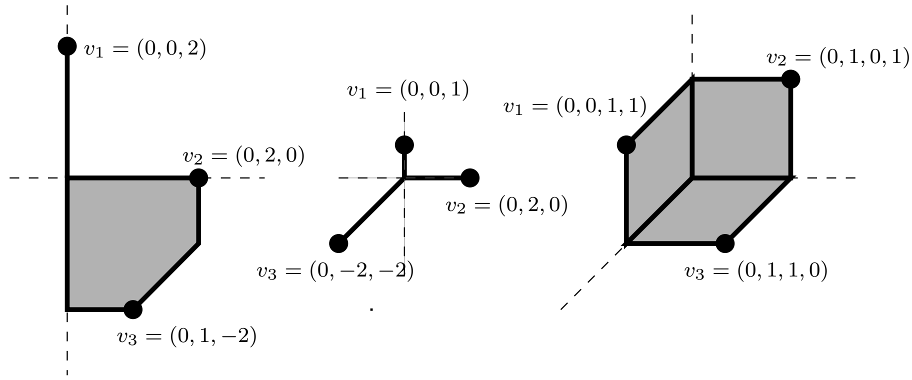


Figure 2: Three tropical polytopes. The first two live in $\mathbb{TP}^2$, the last in $\mathbb{TP}^3$.

Proof. Let $x = \bigoplus_{i=1}^{r} a_i \odot v_i$ and suppose $r > n$. For each coordinate $j \in \{1,\ldots,n\}$, there exists an index $i \in \{1,\ldots,r\}$ such that $x_j = c_i + v_{ij}$. Take a subset I of $\{1,\ldots,r\}$ composed of one such i for each j. Then we also have $x = \bigoplus_{i\in I} a_i \odot v_i$, where $\#(I) \leq n$. $\square$

The basic theory of tropical linear subspaces in $\mathbb{TP}^{n-1}$ was developed in [18] and [20]. Recall that the *tropical hyperplane* defined by a tropical linear form $a_1 \odot x_1 \oplus a_2 \odot x_2 \oplus \cdots \oplus a_n \odot x_n$ consists of all points $x = (x_1, x_2, \ldots, x_n)$ in $\mathbb{TP}^{n-1}$ such that the following holds (in ordinary arithmetic):

$$a_i + x_i \;=\; a_j + x_j \;=\; \min\{a_k + x_k \,:\, k = 1,\ldots,n\} \qquad \text{for some indices } i \neq j. \tag{3}$$

Just like in ordinary geometry, hyperplanes are convex sets:

PROPOSITION 6. *Tropical hyperplanes in $\mathbb{TP}^{n-1}$ are tropically convex.*

Proof. Let H be the hyperplane defined by (3). Suppose that x and y lie in H and consider any tropical linear combination $z = c \odot x \oplus d \odot y$. Let i be an index which minimizes $a_i + z_i$. We need to show that this minimum is attained at least twice. By definition, z_i is equal to either $c + x_i$ or $d + y_i$, and, after permuting x and y, we may assume $z_i = c + x_i \leq d + y_i$. Since, for all k, $a_i + z_i \leq a_k + z_k$ and $z_k \leq c + x_k$, it follows that $a_i + x_i \leq a_k + x_k$ for all k, so that $a_i + x_i$ achieves the minimum of $\{a_1 + x_1, \ldots, a_n + x_n\}$. Since x is in H, there exists some index $j \neq i$ for which $a_i + x_i = a_j + x_j$. But now $a_j + z_j \leq a_j + c + x_j = c + a_i + x_i = a_i + z_i$. Since $a_i + z_i$ is the minimum of all $a_j + z_j$, the two must be equal, and this minimum is obtained at least twice as desired. $\square$

Proposition 6 implies that if V is a subset of $\mathbb{TP}^{n-1}$ which happens to lie in a tropical hyperplane H, then its tropical convex hull $\operatorname{tconv}(V)$ will lie in H

as well. The same holds for tropical planes of higher codimension. Recall that every *tropical plane* is an intersection of tropical hyperplanes [20]. But the converse does not hold: not every intersection of tropical hyperplanes qualifies as a tropical plane (see [18, §5]). Proposition 6 and the first statement in Theorem 2 imply:

Corollary 7. *Tropical planes in* $\mathbb{TP}^{n-1}$ *are tropically convex.*

A theorem in classical geometry states that every point outside a closed convex set can be separated from the convex set by a hyperplane. The same statement holds in tropical geometry. This follows from the results in [3]. Some caution is needed, however, since the definition of hyperplane in [3] differs from our definition of hyperplane, as explained in [18]. In our definition, a tropical hyperplane is a fan which divides $\mathbb{TP}^{n-1}$ into n convex cones, each of which is also tropically convex. Rather than stating the most general separation theorem, we will now focus our attention on tropical polytopes, in which case the separation theorem is the Farkas Lemma stated in the next section.

3 Tropical polytopes and cell complexes

Throughout this section we fix a finite subset $V = \{v_1, v_2, \ldots, v_r\}$ of tropical projective space $\mathbb{TP}^{n-1}$. Here $v_i = (v_{i1}, v_{i2}, \ldots, v_{in})$. Our goal is to study the tropical polytope $P = \mathrm{tconv}(V)$. We begin by describing the natural cell decomposition of $\mathbb{TP}^{n-1}$ induced by the fixed finite subset V.
Let x be any point in $\mathbb{TP}^{n-1}$. The *type* of x relative to V is the ordered n-tuple $(S_1, \ldots, S_n)$ of subsets $S_j \subseteq \{1, 2, \ldots, r\}$ which is defined as follows: An index i is in S_j if

$$v_{ij} - x_j \quad = \quad \min(v_{i1} - x_1, v_{i2} - x_2, \ldots, v_{in} - x_n).$$

Equivalently, if we set $\lambda_i = \min\{\, \lambda \in \mathbb{R} : \lambda \odot v_i \oplus x = x \,\}$ then S_j is the set of all indices i such that $\lambda_i \odot v_i$ and x have the same j-th coordinate. We say that an n-tuple of indices $S = (S_1, \ldots, S_n)$ is a *type* if it arises in this manner. Note that every i must be in some S_j.

Example 8. Let $r = n = 3$, $v_1 = (0,0,2)$, $v_2 = (0,2,0)$ and $v_3 = (0,1,-2)$. There are 30 possible types as x ranges over the plane $\mathbb{TP}^2$. The corresponding cell decomposition has six convex regions (one bounded, five unbounded), 15 edges (6 bounded, 9 unbounded) and 6 vertices. For instance, the point $x = (0,1,-1)$ has $\mathrm{type}(x) = \big(\{2\},\{1\},\{3\}\big)$ and its cell is a bounded pentagon. The point $x' = (0,0,0)$ has $\mathrm{type}(x') = \big(\{1,2\},\{1\},\{2,3\}\big)$ and its cell is one of the six vertices. The point $x'' = (0,0,-3)$ has $\mathrm{type}(x'') = \big\{\{1,2,3\},\{1\},\emptyset\big)$ and its cell is an unbounded edge.

Our first application of types is the following separation theorem.

PROPOSITION 9 (TROPICAL FARKAS LEMMA). *For all $x \in \mathbb{TP}^{n-1}$, exactly one of the following is true.*
(i) the point x is in the tropical polytope $P = \mathrm{tconv}(V)$, or
(ii) there exists a tropical hyperplane which separates x from P.

The separation statement in part (ii) means the following: if the hyperplane is given by (3) and $a_k + x_k = \min(a_1 + x_1, \ldots, a_n + x_n)$ then $a_k + y_k > \min(a_1 + y_1, \ldots, a_n + y_n)$ for all $y \in P$.

Proof. Consider any point $x \in \mathbb{TP}^{n-1}$, with $\mathrm{type}(x) = (S_1, \ldots, S_n)$, and let $\lambda_i = \min\{\, \lambda \in \mathbb{R} : \lambda \odot v_i \oplus x = x \,\}$ as before. We define

$$\pi_V(x) \quad = \quad \lambda_1 \odot v_1 \,\oplus\, \lambda_2 \odot v_2 \,\oplus\, \cdots \,\oplus\, \lambda_r \odot v_r. \tag{4}$$

There are two cases: either $\pi_V(x) = x$ or $\pi_V(x) \neq x$. The first case implies (i). Since (i) and (ii) clearly cannot occur at the same time, it suffices to prove that the second case implies (ii).
Suppose that $\pi_V(x) \neq x$. Then S_k is empty for some index $k \in \{1, \ldots, n\}$. This means that $v_{ik} + \lambda_i - x_k > 0$ for $i = 1, 2, \ldots, r$. Let $\varepsilon > 0$ be smaller than any of these r positive reals. We now choose our separating tropical hyperplane (3) as follows:

$$a_k \;:=\; -x_k - \varepsilon \qquad \text{and} \qquad a_j \;:=\; -x_j \;\text{ for } j \in \{1, \ldots, n\} \backslash \{k\}. \tag{5}$$

This certainly satisfies $a_k + x_k = \min(a_1 + x_1 \ldots, a_n + x_n)$. Now, consider any point $y = \bigoplus_{i=1}^{r} c_i \odot v_i$ in $\mathrm{tconv}(V)$. Pick any m such that $y_k = c_m + v_{mk}$. By definition of the λ_i, we have $x_k \leq \lambda_m + v_{mk}$ for all k, and there exists some j with $x_j = \lambda_m + v_{mj}$. These equations and inequalities imply

$$\begin{aligned} a_k + y_k \quad &= \quad a_k + c_m + v_{mk} \;=\; c_m + v_{mk} - x_k - \varepsilon \;>\; c_m - \lambda_m \\ = \; c_m + v_{mj} - x_j \;&\geq\; y_j - x_j \;=\; a_j + y_j \quad \geq \quad \min(a_1 + y_1, \ldots, a_n + y_n). \end{aligned}$$

Therefore, the hyperplane defined by (5) separates x from P as desired. ☐

The construction in (4) defines a map $\pi_V : \mathbb{TP}^{n-1} \to P$ whose restriction to P is the identity. This map is the tropical version of the *nearest point map* onto a closed convex set in ordinary geometry. Such maps were studied in [3] for convex subsets in arbitrary idempotent semimodules.
If $S = (S_1, \ldots, S_n)$ and $T = (T_1, \ldots, T_n)$ are n-tuples of subsets of $\{1, 2, \ldots, r\}$, then we write $S \subseteq T$ if $S_j \subseteq T_j$ for $j = 1, \ldots, n$. We also consider the set of all points whose type contains S:

$$X_S \quad := \quad \big\{\, x \in \mathbb{TP}^{n-1} \,:\, S \,\subseteq\, \mathrm{type}(x) \big\}.$$

LEMMA 10. *The set X_S is a closed convex polyhedron (in the usual sense). More precisely,*

$$X_S \,=\, \big\{\, x \in \mathbb{TP}^{n-1} \,:\, x_k - x_j \leq v_{ik} - v_{ij} \;\; \textit{for all } j, k \in \{1, \ldots, n\} \textit{ with } i \in S_j \,\big\}. \tag{6}$$

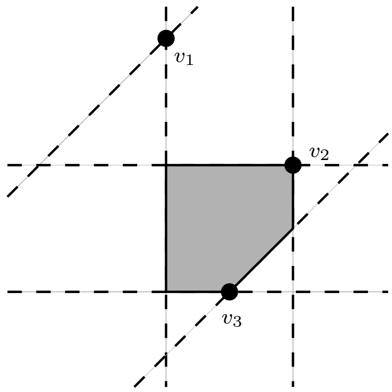


Figure 3: The region $X_{(2,1,3)}$ in the tropical convex hull of v_1, v_2 and v_3.

Proof. Let $x \in \mathbb{TP}^{n-1}$ and $T = \text{type}(x)$. First, suppose x is in X_S. Then $S \subseteq T$. For every i, j, k such that $i \in S_j$, we also have $i \in T_j$, and so by definition we have $v_{ij} - x_j \leq v_{ik} - x_k$, or $x_k - x_j \leq v_{ik} - v_{ij}$. Hence x lies in the set on the right hand side of (6). For the proof of the reverse inclusion, suppose that x lies in the right hand side of (6). Then, for all i, j with $i \in S_j$, and for all k, we have $v_{ij} - x_j \leq v_{ik} - x_k$. This means that $v_{ij} - x_j = \min(v_{i1} - x_1, \ldots, v_{in} - x_n)$ and hence $i \in T_j$. Consequently, for all j, we have $S_j \subset T_j$, and so $x \in X_S$. $\square$

As an example for Lemma 10, we consider the region $X_{(2,1,3)}$ in the tropical convex hull of $v_1 = (0, 0, 2)$, $v_2 = (0, 2, 0)$, and $v_3 = (0, 1, -2)$. This region is defined by six linear inequalities, one of which is redundant, as depicted in Figure 3. Lemma 10 has the following immediate corollaries.

Corollary 11. *The intersection $X_S \cap X_T$ is equal to the polyhedron $X_{S \cup T}$.*

Proof. The inequalities defining $X_{S \cup T}$ are precisely the union of the inequalities defining X_S and X_T, and points satisfying these inequalities are precisely those in $X_S \cap X_T$. $\square$

Corollary 12. *The polyhedron X_S is bounded if and only if $S_j \neq \emptyset$ for all $j = 1, 2, \ldots, n$.*

Proof. Suppose that $S_j \neq \emptyset$ for all $j = 1, 2, \ldots, n$. Then for every j and k, we can find $i \in S_j$ and $m \in S_k$, which via Lemma 10 yield the inequalities $v_{mk} - v_{mj} \leq x_k - x_j \leq v_{ik} - v_{ij}$. This implies that each $x_k - x_j$ is bounded on X_S, which means that X_S is a bounded subset of $\mathbb{TP}^{n-1}$.
Conversely, suppose some S_j is empty. Then the only inequalities involving x_j are of the form $x_j - x_k \leq c_{jk}$. Consequently, if any point x is in S_j, so too is $x - ke_j$ for $k > 0$, where e_j is the j-th basis vector. Therefore, in this case, X_S is unbounded. $\square$

Corollary 13. *Suppose we have $S = (S_1, \ldots, S_n)$, with $S_1 \cup \cdots \cup S_n = \{1, \ldots, r\}$. Then if $S \subseteq T$, X_T is a face of X_S, and furthermore all faces of X_S are of this form.*

Proof. For the first part, it suffices to prove that the statement is true when T covers S in the poset of containment, i.e. when $T_j = S_j \cup \{i\}$ for some $j \in \{1, \ldots, n\}$ and $i \notin S_j$, and $T_k = S_k$ for $k \neq j$.
We have the inequality presentation of X_S given by Lemma 10. By the same lemma, the inequality presentation of X_T consists of the inequalities defining X_S together with the inequalities

$$\{x_k - x_j \leq v_{ik} - v_{ij} \mid k \in \{1, \ldots, n\}\}. \tag{7}$$

By assumption, i is in some S_m. We claim that X_T is the face of S defined by the equality

$$x_m - x_j = v_{im} - v_{ij}. \tag{8}$$

Since X_S satisfies the inequality $x_j - x_m \leq v_{ij} - v_{im}$, (8) defines a face F of S. The inequality $x_m - x_j \leq v_{im} - v_{ij}$ is in the set (7), so (8) is valid on X_T and $X_T \subseteq F$. However, any point in F, being in X_S, satisfies $x_k - x_m \leq v_{ik} - v_{im}$ for all $k \in \{1, \ldots, n\}$. Adding (8) to these inequalities proves that the inequalities (7) are valid on F, and hence $F \subseteq X_T$. So $X_T = F$ as desired.
By the discussion in the proof of the first part, prescribing equality in the facet-defining inequality $x_k - x_j \leq v_{ik} - v_{ij}$ yields X_T, where $T_k = S_k \cup \{i\}$ and $T_j = S_j$ for $j \neq k$. Therefore, all facets of X_S can be obtained as regions X_T, and it follows recursively that all faces of X_S are of this form. □

Corollary 14. *Suppose that $S = (S_1, \ldots, S_n)$ is an n-tuple of indices satisfying $S_1 \cup \cdots \cup S_n = \{1, \ldots, r\}$. Then X_S is equal to X_T for some type T.*

Proof. Let x be a point in the relative interior of X_S, and let $T = \mathrm{type}(x)$. Since $x \in X_S$, T contains S, and by Lemma 13, X_T is a face of X_S. However, since x is in the relative interior of X_S, the only face of X_S containing x is X_S itself, so we must have $X_S = X_T$ as desired. □

We are now prepared to state our main theorem in this section.

Theorem 15. *The collection of convex polyhedra X_S, where S ranges over all types, defines a cell decomposition $\mathcal{C}_V$ of $\mathbb{TP}^{n-1}$. The tropical polytope $P = \mathrm{tconv}(V)$ equals the union of all bounded cells X_S in this decomposition.*

Proof. Since each point has a type, it is clear that the union of the X_S is equal to $\mathbb{TP}^{n-1}$. By Corollary 13, the faces of X_S are equal to X_U for $S \subseteq U$, and by Corollary 14, $X_U = X_W$ for some type W, and hence X_U is in our collection. The only thing remaining to check to show that this collection defines a cell decomposition is that $X_S \cap X_T$ is a face of both X_S and X_T,

but $X_S \cap X_T = X_{S \cup T}$ by Corollary 11, and $X_{S\cup T}$ is a face of X_S and X_T by Corollary 13.
For the second assertion consider any point $x \in \mathbb{TP}^{n-1}$ and let $S = \text{type}(x)$. We have seen in the proof of the Tropical Farkas Lemma (Proposition 9) that x lies in P if and only if no S_j is empty. By Corollary 12, this is equivalent to the polyhedron X_S being bounded. □

The collection of bounded cells X_S is referred to as the tropical complex generated by V; thus, Theorem 15 states that this provides a polyhedral decomposition of the polytope $P = \text{tconv}(V)$. Different sets V may have the same tropical polytope as their convex hull, but generate different tropical complexes; the decomposition of a tropical polytope depends on the chosen generating set, although we will see later (Proposition 21) that there is a unique minimal generating set.
Here is a nice geometric construction of the cell decomposition $\mathcal{C}_V$ of $\mathbb{TP}^{n-1}$ induced by $V = \{v_1, \ldots, v_r\}$. Let $\mathcal{F}$ be the fan in $\mathbb{TP}^{n-1}$ defined by the tropical hyperplane (3) with $a_1 = \cdots = a_n = 0$. Two vectors x and y lie in the same relatively open cone of the fan $\mathcal{F}$ if and only if

$$\{\, j \,:\, x_j \,=\, \min(x_1, \ldots, x_n)\,\} \quad = \quad \{\, j \,:\, y_j \,=\, \min(y_1, \ldots, y_n)\,\}.$$

If we translate the negative of $\mathcal{F}$ by the vector v_i then we get a new fan which we denote by $v_i - \mathcal{F}$. Two vectors x and y lie in the same relatively open cone of the fan $v_i - \mathcal{F}$ if and only if

$$\begin{aligned} &\{\, j \,:\, x_j - v_{ij} \,=\, \max(x_1 - v_{i1}, \ldots, x_n - v_{in})\,\} \\ &\quad = \quad \{\, j \,:\, y_j - v_{ij} \,=\, \max(y_1 - v_{i1}, \ldots, y_n - v_{in})\,\}. \end{aligned}$$

Proposition 16. *The cell decomposition $\mathcal{C}_V$ is the common refinement of the r fans $v_i - \mathcal{F}$.*

Proof. We need to show that the cells of this common refinement are precisely the convex polyhedra X_S. Take a point x, with $T = \text{type}(x)$ and define $S_x = (S_{x1}, \ldots, S_{xn})$ by letting $i \in S_{xj}$ whenever

$$x_j - v_{ij} \quad = \quad \max(x_1 - v_{i1}, \ldots, x_n - v_{in}). \tag{9}$$

Two points x and y are in the relative interior of the same cell of the common refinement if and only if they are in the same relatively open cone of each fan; this is tantamount to saying that $S_x = S_y$. However, we claim that $S_x = T$. Indeed, taking the negative of both sides of (9) yields exactly the condition for i being in T_j, by the definition of type. Consequently, the condition for two points having the same type is the same as the condition for the two points being in the relative interior of the same chamber of the common refinement of the fans $v_1 - \mathcal{F}$, $v_2 - \mathcal{F}$, $\ldots, v_r - \mathcal{F}$. □

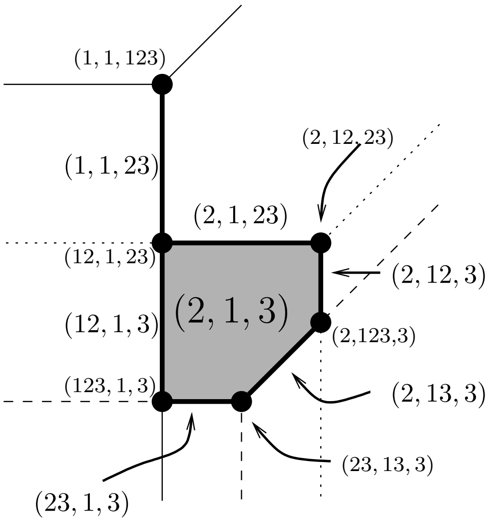


Figure 4: A tropical complex expressed as the bounded cells in the common refinement of the fans $v_1 - \mathcal{F}$, $v_2 - \mathcal{F}$ and $v_3 - \mathcal{F}$. Cells are labeled with their types.

An example of this construction is shown for our usual example, where $v_1 = (0, 0, 2)$, $v_2 = (0, 2, 0)$, and $v_3 = (0, 1, -2)$, in Figure 4.
The next few results provide additional information about the polyhedron X_S. Let G_S denote the undirected graph with vertices $\{1, \ldots, n\}$, where $\{j, k\}$ is an edge if and only if $S_j \cap S_k \neq \emptyset$.

PROPOSITION 17. *The dimension d of the polyhedron X_S is one less than the number of connected components of G_S, and X_S is affinely and tropically isomorphic to some polyhedron X_T in $\mathbb{TP}^d$.*

Proof. The proof is by induction on n. Suppose we have $i \in S_j \cap S_k$. Then X_S satisfies the linear equation $x_k - x_j = c$ where $c = v_{ik} - v_{ij}$. Eliminating the variable x_k (projecting onto $\mathbb{TP}^{n-2}$), X_S is affinely and tropically isomorphic to X_T where the type T is defined by $T_r = S_r$ for $r \neq j$ and $T_j = S_j \cup S_k$. The region X_T exists in the cell decomposition of $\mathbb{TP}^{n-2}$ induced by the vectors $w_1, \ldots, w_n$ with $w_{ir} = v_{ir}$ for $r \neq j$, and $w_{ij} = \max(v_{ij}, v_{ik} - c)$. The graph G_T is obtained from the graph G_S by contracting the edge $\{j, k\}$, and thus has the same number of connected components.
This induction on n reduces us to the case where all of the S_j are pairwise disjoint. We must show that X_S has dimension $n - 1$. Suppose not. Then X_S lies in $\mathbb{TP}^{n-1}$ but has dimension less than $n - 1$. Therefore, one of the inequalities in (6) holds with equality, say $x_k - x_j = v_{ik} - v_{ij}$ for all $x \in X_S$. The inequality "$\leq$" implies $i \in S_j$ and the inequality "$\geq$" implies $i \in S_k$. Hence

S_j and S_k are not disjoint, a contradiction. $\square$

The following proposition can be regarded as a converse to Lemma 10.

Proposition 18. *Let R be any polytope in $\mathbb{TP}^{n-1}$ defined by inequalities of the form $x_k - x_j \leq c_{jk}$. Then R arises as a cell X_S in the decomposition $\mathcal{C}_V$ of $\mathbb{TP}^{n-1}$ defined by some set $V = \{v_1, \ldots, v_n\}$.*

Proof. Define the vectors v_i to have coordinates $v_{ij} = c_{ij}$ for $i \neq j$, and $v_{ii} = 0$. (If c_{ij} did not appear in the given inequality presentation then simply take it to be a very large positive number.) Then by Lemma 10, the polytope in $\mathbb{TP}^{n-1}$ defined by the inequalities $x_k - x_j \leq c_{jk}$ is precisely the unique cell of type $(1, 2, \ldots, n)$ in the tropical convex hull of $\{v_1, \ldots, v_n\}$. $\square$

The region X_S is a polytope both in the ordinary sense and in the tropical sense.

Proposition 19. *Every bounded cell X_S in the tropical complex generated by V is itself a tropical polytope, equal to the tropical convex hull of its vertices. The number of vertices of the polytope X_S is at most $\binom{2n-2}{n-1}$, and this bound is tight for all positive integers n.*

Proof. By Proposition 17, if X_S has dimension d, it is affinely and tropically isomorphic to a region in the convex hull of a set of points in $\mathbb{TP}^d$, so it suffices to consider the full-dimensional case.
The inequality presentation of Lemma 10 demonstrates that X_S is tropically convex for all S, since if two points satisfy an inequality of that form, so does any tropical linear combination thereof. Therefore, it suffices to show that X_S is contained in the tropical convex hull of its vertices.
The proof is by induction on the dimension of X_S. All proper faces of X_S are polytopes X_T of lower dimension, and by induction are contained in the tropical convex hull of their vertices. These vertices are a subset of the vertices of X_S, and so this face is in the tropical convex hull.
Take any point $x = (x_1, \ldots, x_n)$ in the interior of X_S. Since X_S has dimension n, we can travel in any direction from x while remaining in X_S. Let us travel in the $(1, 0, \ldots, 0)$ direction until we hit the boundary, to obtain points $y_1 = (x_1 + b, x_2, \ldots, x_n)$ and $y_2 = (x_1 - c, x_2, \ldots, x_n)$ in the boundary of X_S. These points are contained in the tropical convex hull by the inductive hypothesis, which means that $x = y_1 \oplus c \odot y_2$ is also, completing the proof of the first assertion.
For the second assertion, we consider the convex hull of all differences of unit vectors, $e_i - e_j$. This is a lattice polytope of dimension $n-1$ and normalized volume $\binom{2n-2}{n-1}$. To see this, we observe that this polytope is tiled by n copies of the convex hull of the origin and the $\binom{n}{2}$ vectors $e_i - e_j$ with $i < j$. The other $n-1$ copies are gotten by cyclic permutation of the coordinates. This latter polytope was studied by Gel'fand, Graev and Postnikov, who showed in

[4, Theorem 2.3 (2)] that the normalized volume of this polytope equals the Catalan number $\frac{1}{n}\binom{2n-2}{n-1}$.
We conclude that every complete fan whose rays are among the vectors $e_i - e_j$ has at most $\binom{2n-2}{n-1}$ maximal cones. This applies in particular to the normal fan of X_S, hence X_S has at most $\binom{2n-2}{n-1}$ vertices. Since the configuration $\{e_i - e_j\}$ is unimodular, the bound is tight whenever the fan is simplicial and uses all the rays $e_i - e_j$. □

We close this section with two more results about arbitrary tropical polytopes in $\mathbb{TP}^{n-1}$.

PROPOSITION 20. *If P and Q are tropical polytopes in $\mathbb{TP}^{n-1}$ then $P \cap Q$ is also a tropical polytope.*

Proof. Since P and Q are both tropically convex, $P \cap Q$ must also be. Consequently, if we can find a finite set of points in $P \cap Q$ whose convex hull contains all of $P \cap Q$, we will be done. By Theorem 15, P and Q are the finite unions of bounded cells $\{X_S\}$ and $\{X_T\}$ respectively, so $P \cap Q$ is the finite union of the cells $X_S \cap X_T$. Consider any $X_S \cap X_T$. Using Lemma 10 to obtain the inequality representations of X_S and X_T, we see that this region is of the form dictated by Proposition 18, and therefore obtainable as a cell X_W in some tropical complex. By Proposition 19, X_W is itself a tropical polytope, and we can therefore find a finite set whose convex hull is equal to $X_S \cap X_T$. Taking the union of these sets over all choices of S and T then gives us the desired set of points whose convex hull contains all of $P \cap Q$. □

PROPOSITION 21. *Let $P \subset \mathbb{TP}^{n-1}$ be a tropical polytope. Then there exists a unique minimal set V such that $P = \mathrm{tconv}(V)$.*

Proof. Suppose that P has two minimal generating sets, $V = \{v_1, \ldots, v_m\}$ and $W = \{w_1, \ldots, w_r\}$. Write each element of W as $w_i = \oplus_{j=1}^m c_{ij} \odot v_j$. We claim that $V \subseteq W$. Consider $v_1 \in V$ and write

$$v_1 \;=\; \bigoplus_{i=1}^{r} d_i \odot w_i \;=\; \bigoplus_{j=1}^{m} f_j \odot v_j \qquad \text{where } f_j = \min_i(d_i + c_{ij}). \tag{10}$$

If the term $f_1 \odot v_1$ does not minimize any coordinate in the right-hand side of (10), then v_1 is a linear combination of $v_2, \ldots, v_m$, contradicting the minimality of V. However, if $f_1 \odot v_1$ minimizes any coordinate in this expression, it must minimize all of them, since $(v_1)_j - (v_1)_k = (f_1 \odot v_1)_j - (f_1 \odot v_1)_k$. In this case we get $v_1 = f_1 \odot v_1$, or $f_1 = 0$. Pick any i for which $f_1 = d_i + c_{i1}$; we claim that $w_i = c_{i1} \odot v_1$. Indeed, if any other term in $w_i = \oplus_{j=1}^m c_{ij} \odot v_j$ contributed nontrivially to w_i, that term would also contribute to the expression on the right-hand side of (10), which is a contradiction. Consequently, $V \subseteq W$, which means $V = W$ since both sets are minimal by hypothesis. □

Like many of the results presented in this section, Propositions 20 and 21 parallel results on ordinary polytopes. We have already mentioned the tropical analogues of the Farkas Lemma and of Carathéodory's Theorem (Propositions 5 and 9); Proposition 17 is analogous to the result that a polytope $P \subset \mathbb{R}^n$ of dimension d is affinely isomorphic to some $Q \subset \mathbb{R}^d$. Proposition 19 hints at a duality between an inequality representation and a vertex representation of a tropical polytope; this duality has been studied in greater detail by Michael Joswig [11].

4 Subdividing products of simplices

Every set $V = \{v_1, \ldots, v_r\}$ of r points in $\mathbb{TP}^{n-1}$ begets a tropical polytope $P = \text{tconv}(V)$ equipped with a cell decomposition into the tropical complex generated by V. Each cell of this tropical complex is labelled by its type, which is an n-vector of finite subsets of $\{1, \ldots, r\}$. Two configurations (and their corresponding tropical complexes) V and W have the same *combinatorial type* if the types occurring in their tropical complexes are identical; note that by Lemma 13, this implies that the face posets of these polyhedral complexes are isomorphic.

With the definition in the previous paragraph, the statement of Theorem 1 has now finally been made precise. We will prove this correspondence between tropical complexes and subdivisions of products of simplices by constructing the polyhedral complex $\mathcal{C}_P$ in a higher-dimensional space.

Let W denote the $(r + n - 1)$-dimensional real vector space $\mathbb{R}^{r+n}/(1, \ldots, 1, -1, \ldots, -1)$. The natural coordinates on W are denoted $(y, z) = (y_1, \ldots, y_r, z_1, \ldots, z_n)$. As before, we fix an ordered subset $V = \{v_1, \ldots, v_r\}$ of $\mathbb{TP}^{n-1}$ where $v_i = (v_{i1}, \ldots, v_{in})$. This defines the unbounded polyhedron

$$\mathcal{P}_V = \left\{ (y, z) \in W \,:\, y_i + z_j \leq v_{ij} \text{ for all } i \in \{1, \ldots, r\} \text{ and } j \in \{1, \ldots, n\} \right\}. \tag{11}$$

Lemma 22. *There is a piecewise-linear isomorphism between the tropical complex generated by V and the complex of bounded faces of the $(r + n - 1)$-dimensional polyhedron $\mathcal{P}_V$. The image of a cell X_S of $\mathcal{C}_P$ under this isomorphism is the bounded face $\{y_i + z_j = v_{ij} \,:\, i \in S_j\}$ of the polyhedron $\mathcal{P}_V$. That bounded face maps isomorphically to X_S via projection onto the z-coordinates.*

Proof. Let F be a bounded face of $\mathcal{P}_V$, and define S_j via $i \in S_j$ if $y_i + z_j = v_{ij}$ is valid on all of F. If some y_i or z_j appears in no equality, then we can subtract arbitrary positive multiples of that basis vector to obtain elements of F, contradicting the assumption that F is bounded. Therefore, each i must appear in some S_j, and each S_j must be nonempty.

Since every y_i appears in some equality, given a specific z in the projection of F onto the z-coordinates, there exists a unique y for which $(y, z) \in F$, so this projection is an affine isomorphism from F to its image. We need to show that this image is equal to X_S.

Let z be a point in the image of this projection, coming from a point (y,z) in the relative interior of F. We claim that $z \in X_S$. Indeed, looking at the jth coordinate of z, we find

$$-y_i + v_{ij} \;\geq\; z_j \quad \text{for all } i, \tag{12}$$

$$-y_i + v_{ij} \;=\; z_j \quad \text{for } i \in S_j. \tag{13}$$

The defining inequalities of X_S are $x_j - x_k \leq v_{ij} - v_{ik}$ with $i \in S_j$. Subtracting the inequality $-y_i + v_{ik} \geq z_k$ from the equality in (13) yields that this inequality is valid on z as well. Therefore, $z \in X_S$. Similar reasoning shows that $S = \mathrm{type}(z)$. We note that the relations (12) and (13) can be rewritten elegantly in terms of the tropical product of a row vector and a matrix:

$$z \quad = \quad (-y) \odot V \quad = \quad \bigoplus_{i=1}^{r} (-y_i) \odot v_i. \tag{14}$$

For the reverse inclusion, suppose that $z \in X_S$. We define $y = V \odot (-z)$. This means that

$$y_i \;=\; \min(v_{i1} - z_1, v_{i2} - z_2, \ldots, v_{in} - z_n). \tag{15}$$

We claim that $(y,z) \in F$. Indeed, we certainly have $y_i + z_j \leq v_{ij}$ for all i and j, so $(y,z) \in \mathcal{P}_V$. Furthermore, when $i \in S_j$, we know that $v_{ij} - z_j$ achieves the minimum in the right-hand side of (15), so that $v_{ij} - z_j = y_i$ and $y_i + z_j = v_{ij}$ is satisfied. Consequently, $(y,z) \in F$ as desired.

It follows immediately that the two complexes are isomorphic: if F is a face corresponding to X_S and G is a face corresponding to X_T, where S and T are both types, then X_S is a face of X_T if and only if $T \subseteq S$. However, by the discussion above, this is equivalent to saying that the equalities G satisfies (which correspond to T) are a subset of the equalities F satisfies (which correspond to S); this is true if and only if F is a face of G. So X_S is a face of X_T if and only if F is a face of G, which implies the isomorphism of complexes. ☐

The boundary complex of the polyhedron $\mathcal{P}_V$ is polar to the regular subdivision of the product of simplices $\Delta_{r-1} \times \Delta_{n-1}$ defined by the weights v_{ij}. We denote this regular polyhedral subdivision by $(\partial \mathcal{P}_V)^*$. Explicitly, a subset of vertices (e_i, e_j) of $\Delta_{r-1} \times \Delta_{n-1}$ forms a cell of $(\partial \mathcal{P}_V)^*$ if and only if the equations $y_i + z_j = v_{ij}$ indexed by these vertices specify a face of the polyhedron $\mathcal{P}_V$. We refer to the book of De Loera, Rambau and Santos [5] for basics on polyhedral subdivisions.

We now present the proof of the result stated in the introduction.

Proof of Theorem 1: The poset of bounded faces of $\mathcal{P}_V$ is antiisomorphic to the poset of interior cells of the subdivision $(\partial \mathcal{P}_V)^*$ of $\Delta_{r-1} \times \Delta_{n-1}$. Since every full-dimensional cell of $(\partial \mathcal{P}_V)^*$ is interior, the subdivision is uniquely determined by its interior cells. In other words, the combinatorial type of $\mathcal{P}_V$

is uniquely determined by the lists of facets containing each bounded face of $\mathcal{P}_V$. These lists are precisely the types of regions in $\mathcal{C}_P$ by Lemma 22. This completes the proof. $\square$

Theorem 1, which establishes a bijection between the tropical complexes generated by r points in $\mathbb{TP}^{n-1}$ and the regular subdivisions of a product of simplices $\Delta_{r-1} \times \Delta_{n-1}$, has many striking consequences. First of all, we can pick off the types present in a tropical complex simply by looking at the cells present in the corresponding regular subdivision. In particular, if we have an interior cell M, the corresponding type appearing in the tropical complex is defined via $S_j = \{i \in [n] \mid (j,i) \in M\}$.
It is worth noting that via the Cayley Trick [19], Theorem 1 is equivalent to saying that tropical complexes generated by r points in $\mathbb{TP}^{n-1}$ are in bijection with the regular mixed subdivisions of the dilated simplex $r\Delta^{n-1}$. This connection is expanded upon and employed in a paper with Francisco Santos [6]. Another astonishing consequence of Theorem 1 is the identification of the row span and column span of a matrix. This result can also be derived from [3, Theorem 42].

Theorem 23. *Given any matrix $M \in \mathbb{R}^{r\times n}$, the tropical complex generated by its column vectors is isomorphic to the tropical complex generated by its row vectors. This isomorphism is gotten by restricting the piecewise linear maps $\mathbb{R}^n \to \mathbb{R}^r$, $z \mapsto M \odot (-z)$ and $\mathbb{R}^r \to \mathbb{R}^n$, $y \mapsto (-y) \odot M$.*

Proof. By Theorem 1, the matrix M corresponds via the polyhedron $\mathcal{P}_M$ to a regular subdivision of $\Delta_{r-1} \times \Delta_{n-1}$, and the complex of interior faces of this regular subdivision is combinatorially isomorphic to both the tropical complex generated by its row vectors, which are r points in $\mathbb{TP}^{n-1}$, and the tropical complex generated by its column vectors, which are n points in $\mathbb{TP}^{r-1}$. Furthermore, Lemma 22 tells us that the cell in $\mathcal{P}_M$ is affinely isomorphic to its corresponding cell in both tropical complexes. Finally, in the proof of Lemma 22, we showed that the point (y,z) in a bounded face F of $\mathcal{P}_M$ satisfies $y = M \odot (-z)$ and $z = (-y) \odot M$. This point projects to y and z, and so the piecewise-linear isomorphism mapping these two complexes to each other is defined by the stated maps. $\square$

The common tropical complex of these two tropical polytopes is given by the complex of bounded faces of the common polyhedron $\mathcal{P}_M$, which lives in a space of dimension $r+n-1$; the tropical polytopes are unfoldings of this complex into dimensions $r-1$ and $n-1$. Theorem 23 also gives a natural bijection between the combinatorial types of tropical convex hulls of r points in $\mathbb{TP}^{n-1}$ and the combinatorial types of tropical convex hulls of n points in $\mathbb{TP}^{r-1}$, incidentally proving that there are the same number of each. This duality statement extends a similar statement in [3].
Figure 5 shows the dual of the convex hull of $\{(0,0,2),(0,2,0),(0,1,-2)\}$, also

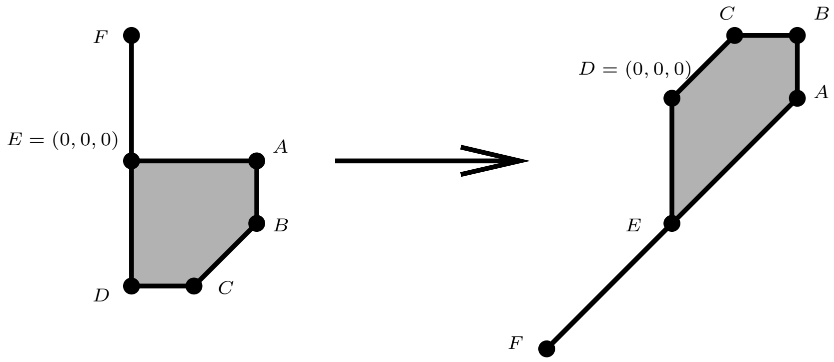


Figure 5: A demonstration of tropical polytope duality.

a tropical triangle (here $r = n = 3$). For instance, we compute:

$$\begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -4 \end{pmatrix}.$$

This point is the image of the point $(0, 0, 2)$ under this duality map. Note that duality does not preserve the generating set; the polytope on the right is the convex hull of points $\{F, D, B\}$, while the polytope on the left is the convex hull of points $\{F, A, C\}$. This is necessary, of course, since in general a polytope with r vertices is mapped to a polytope with n vertices, and r need not equal n as it does in our example.

We now discuss the generic case when the subdivision $(\partial\mathcal{P}_V)^*$ is a regular triangulation of $\Delta_{r-1} \times \Delta_{n-1}$. We refer to [18, §5] for the geometric interpretation of the *tropical determinant*.

PROPOSITION 24. *For a configuration V of r points in $\mathbb{TP}^{n-1}$ with $r \geq n$ the following are equivalent:*

1. *The regular subdivision $(\partial\mathcal{P}_V)^*$ is a triangulation of $\Delta_{r-1} \times \Delta_{n-1}$.*
2. *No k of the points in V have projections onto a k-dimensional coordinate subspace which lie in a tropical hyperplane, for any $2 \leq k \leq n$.*
3. *No $k \times k$-submatrix of the $r \times n$-matrix (v_{ij}) is tropically singular, i.e. has vanishing tropical determinant, for any $2 \leq k \leq n$.*

Proof. The last equivalence is proven in [18, Lemma 5.1]. We will prove that (1) and (3) are equivalent. The tropical determinant of a k by k matrix M is the tropical polynomial $\oplus_{\sigma \in S_k}(\odot_{i=1}^k M_{i\sigma(i)})$. The matrix M is tropically singular if the minimum $\min_{\sigma \in S_k}\left(\sum_{i=1}^k M_{i\sigma(i)}\right)$ is achieved twice.

The regular subdivision $(\partial\mathcal{P}_V)^*$ is a triangulation if and only if the polyhedron $\mathcal{P}_V$ is simple, which is to say if and only if no $r+n$ of the facets $y_i + z_j \leq v_{ij}$ meet at a single vertex. For each vertex v, consider the bipartite graph G_v consisting of vertices $y_1, \ldots, y_n$ and $z_1, \ldots, z_j$ with an edge connecting y_i and z_j if v lies on the corresponding facet. This graph is connected, since each y_i and z_j appears in some such inequality, and thus it will have a cycle if and only if it has at least $r+n$ edges. Consequently, $\mathcal{P}_V$ is not simple if and only there exists some G_v with a cycle.

If there is a cycle, without loss of generality it reads $y_1, z_1, y_2, z_2, \ldots, y_k, z_k$. Consider the submatrix M of (v_{ij}) given by $1 \leq i \leq k$ and $1 \leq j \leq k$. We have $y_1 + z_1 = M_{11}$, $y_2 + z_2 = M_{22}$, and so on, and also $z_1 + y_2 = M_{12}, \ldots, z_k + y_1 = M_{k1}$. Adding up all of these equalities yields $y_1 + \cdots + y_k + z_1 + \cdots + z_k = M_{11} + \cdots + M_{kk} = M_{12} + \cdots + M_{k1}$. But consider any permutation σ in the symmetric group S_k. Since we have $M_{i\sigma(i)} = v_{i\sigma(i)} \geq y_i + z_{\sigma(i)}$, we have $\sum M_{i\sigma(i)} \geq x_1 + \cdots + x_k + y_1 + \cdots + y_k$. Consequently, the permutations equal to the identity and to $(12\cdots k)$ simultaneously minimize the determinant of the minor M. This logic is reversible, proving the equivalence of (1) and (3). □

If the r points of V are in general position, the tropical complex they generate is called a *generic tropical complex*. These polyhedral complexes are then polar to the complexes of interior faces of regular triangulations of $\Delta_{r-1} \times \Delta_{n-1}$.

Corollary 25. *All tropical complexes generated by r points in general position in $\mathbb{TP}^{n-1}$ have the same f-vector. Specifically, the number of faces of dimension k is equal to the multinomial coefficient*

$$\binom{r+n-k-2}{r-k-1,\, n-k-1,\, k} = \frac{(r+n-k-2)!}{(r-k-1)! \cdot (n-k-1)! \cdot k!}.$$

Proof. By Proposition 24, these objects are in bijection with regular triangulations of $P = \Delta_{r-1} \times \Delta_{n-1}$. The polytope P is equidecomposable [1], meaning that all of its triangulations have the same f-vector. The number of faces of dimension k of the tropical complex generated by given r points is equal to the number of interior faces of codimension k in the corresponding triangulation. Since all triangulations of all products of simplices have the same f-vector, they must also have the same interior f-vector, which can be computed by taking the f-vector and subtracting off the f-vectors of the induced triangulations on the proper faces of P. These proper faces are all products of simplices and hence equidecomposable, so all of these induced triangulations have f-vectors independent of the original triangulation as well.

To compute this number, we therefore need only compute it for one tropical complex. Let the vectors v_i, $1 \leq i \leq r$, be given by $v_i = (i, 2i, \cdots, ni)$. By Theorem 10, to count the faces of dimension k in this tropical complex, we enumerate the existing types with k degrees of freedom. Consider any index i. We claim that for any x in the tropical convex hull of $\{v_i\}$, the set $\{j \mid i \in S_j\}$

is an interval I_i, and that if $i < m$, the intervals I_m and I_i meet in at most one point, which in that case is the largest element of I_m and the smallest element of I_i.

Suppose we have $i \in S_j$ and $m \in S_l$ with $i < m$. Then we have by definition $v_{ij} - x_j \leq v_{il} - x_l$ and $v_{ml} - x_l \leq v_{mj} - x_j$. Adding these inequalities yields $v_{ij} + v_{ml} \leq v_{il} + v_{mj}$, or $ij + ml \leq il + mj$. Since $i < m$, it follows that we must have $l \leq j$. Therefore, we can never have $i \in S_j$ and $m \in S_l$ with $i < m$ and $j < l$. The claim follows immediately, since the I_i cover $[1, n]$.

The number of degrees of freedom of an interval set $(I_1, \ldots, I_r)$ is easily seen to be the number of i for which I_i and I_{i+1} are disjoint. Given this, it follows from a simple combinatorial counting argument that the number of interval sets with k degrees of freedom is the multinomial coefficient given above. Finally, a representative for every interval set is given by $x_j = x_{j+1} - c_j$, where if S_j and S_{j+1} have an element i in common (they can have at most one), $c_j = i$, and if not then $c_j = (\min(S_j) + \max(S_{j+1}))/2$. Therefore, each interval set is in fact a valid type, and our enumeration is complete. □

Corollary 26. *The number of combinatorially distinct generic tropical complexes generated by r points in $\mathbb{TP}^{n-1}$ equals the number of distinct regular triangulations of $\Delta_{r-1} \times \Delta_{n-1}$. The number of respective symmetry classes under the natural action of the product of symmetric groups $G = S_r \times S_n$ on both spaces is also the same.*

The symmetries in the group G correspond to a natural action on $\Delta_{r-1} \times \Delta_{n-1}$ given by permuting the vertices of the two component simplices; the symmetries in the symmetric group S_r correspond to permuting the points in a tropical polytope, while those in the symmetric group S_n correspond to permuting the coordinates. (These are dual, as per Corollary 23.) The number of symmetry classes of regular triangulations of the polytope $\Delta_{r-1} \times \Delta_{n-1}$ is computable via Jörg Rambau's TOPCOM [17] for small r and n:

	2	3
2	5	35
3	35	7,955
4	530	
5	13,631	

For example, the $(2, 3)$ entry of the table divulges that there are 35 symmetry classes of regular triangulations of $\Delta_2 \times \Delta_3$. These correspond to the 35 combinatorial types of four-point configurations in $\mathbb{TP}^2$, or to the 35 combinatorial types of three-point configurations in $\mathbb{TP}^3$. These 35 configurations (with the tropical complexes they generate) are shown in Figure 6; the labeling corresponds to Rambau's labeling (see [17]) of the regular triangulations of $\Delta_3 \times \Delta_2$.

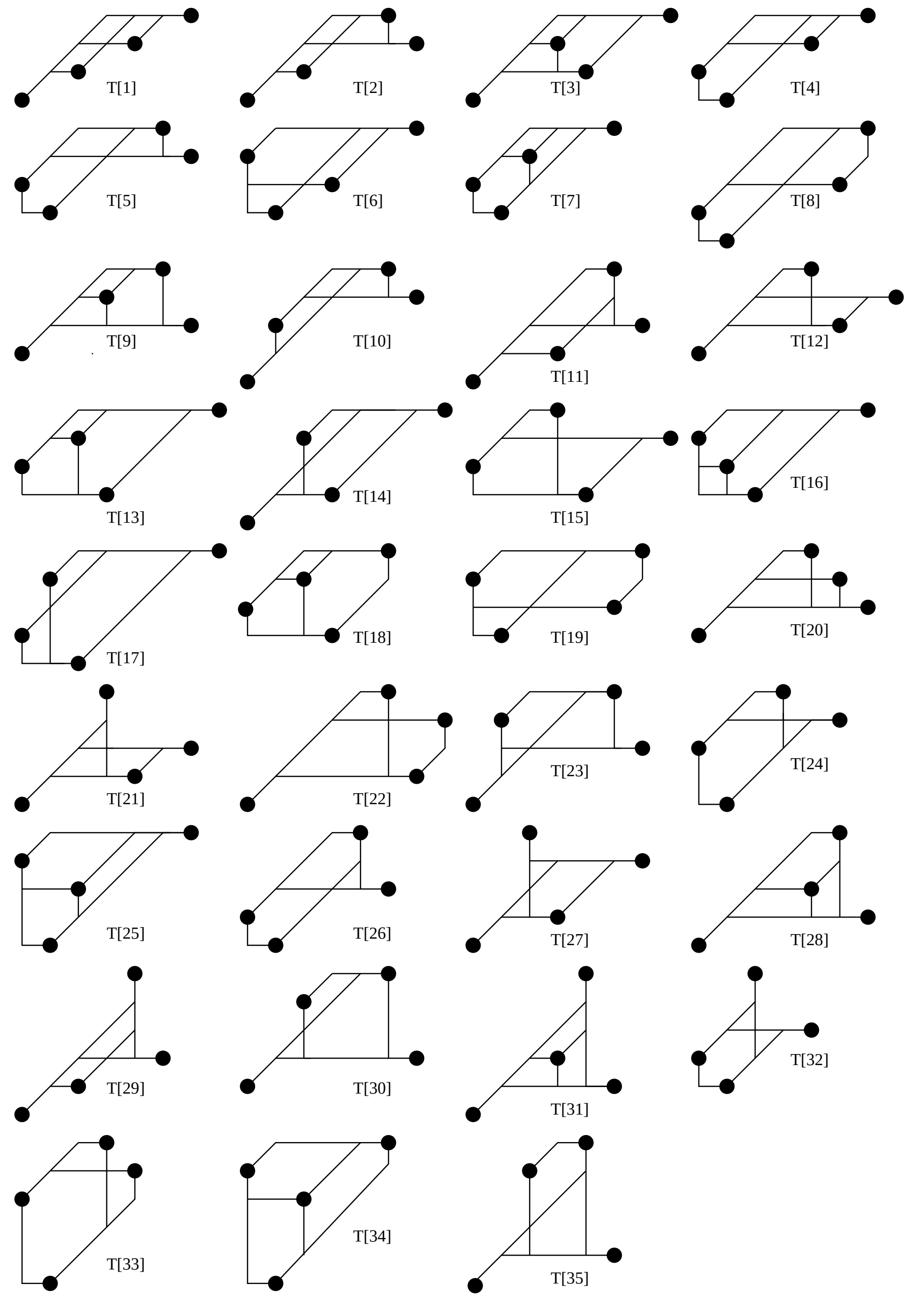


Figure 6: The 35 symmetry classes of tropical complexes generated by four points in $\mathbb{TP}^2$.

5 Phylogenetic analysis using tropical polytopes

A fundamental problem in bioinformatics is the reconstruction of phylogenetic trees from approximate distance data. In this section we show how tropical convexity might help provide new algorithmic tools for this problem. Our approach augments the results in [20, §4] and it provides a tropical interpretation of the work on T-theory by Andreas Dress and his collaborators [7], [8], [9].
Consider a symmetric $n \times n$-matrix $D = (d_{ij})$ whose entries d_{ij} are nonnegative real numbers and whose diagonal entries d_{ii} are all zero. We say that D is a (finite) *metric* if the triangle inequality $d_{ij} \leq d_{ik} + d_{jk}$ holds for all indices i, j, k. Our starting point is the following easy observation:

Proposition 27. *The symmetric matrix D is a metric if and only if all principal 3×3-minors of the negated symmetric matrix $-D = (-d_{ij})$ are tropically singular.*

Proof. Both properties involve only three points, so we may assume $n = 3$, in which case

$$-D \quad = \quad \begin{pmatrix} 0 & -d_{12} & -d_{13} \\ -d_{12} & 0 & -d_{23} \\ -d_{13} & -d_{23} & 0 \end{pmatrix}.$$

The tropical determinant of this matrix is the minimum of the six expressions

$$0,\ -2d_{12},\ -2d_{13},\ -2d_{23},\ -d_{12} - d_{13} - d_{23} \ \text{ and } \ -d_{12} - d_{13} - d_{23}.$$

This minimum is attained twice if and only if it is attained by the last two (identical) expressions, which occurs if and only if the three triangle inequalities are satisfied. □

In what follows we assume that $D = (d_{ij})$ is a metric. Let P_D denote the tropical convex hull in $\mathbb{TP}^{n-1}$ of the n row vectors (or column vectors) of the negated matrix $-D = (-d_{ij})$. Proposition 27 tells us that the tropical polytope P_D is always one-dimensional for $n = 3$.
The finite metric $D = (d_{ij})$ is said to be a *tree metric* if there exists a weighted tree T with n leaves such that d_{ij} denotes the distance between the i-th leaf and the j-th leaf along the unique path between these leaves in T. The next theorem characterizes tree metrics among all metrics by the dimension of the tropical polytope P_D. It is the tropical interpretation of results that are quite classical and well-known in the phylogenetics literature.

Theorem 28. *For a given finite metric $D = (d_{ij})$ the following conditions are equivalent:*

1. *D is a tree metric,*
2. *the tropical polytope P_D has dimension one,*
3. *all 4×4-minors of the matrix $-D$ are tropically singular,*

4. *all principal* 4×4*-minors of the matrix* $-D$ *are tropically singular,*

5. *For any choice of four indices* $i, j, k, l \in \{1, 2, \ldots, n\}$*, the maximum of the three numbers* $d_{ij} + d_{kl}$*,* $d_{ik} + d_{jl}$ *and* $d_{il} + d_{ik}$ *is attained at least twice.*

Proof. The condition (5) is the familiar *Four Point Condition* for tree metrics. The equivalence of (1) and (5) is a classical result due to various authors, including Buneman [2] and Zaretsky [21]. See equation (B3) on page 57 in [7]. Suppose that the condition (5) holds. By the discussion in [20, §4], this means that $-D$ is a point in the tropical Grassmannian of lines, in symbols $-D \in \mathrm{Gr}(2,n) \subset \mathbb{TP}^{\binom{n}{2}}$. By [20, Theorem 3.8], the point $-D$ corresponds to a tropical line L_D in $\mathbb{TP}^{n-1}$. The n distinguished points whose coordinates are the rows of $-D$ lie on the line L_D. By Corollary 7, it follows that their tropical convex hull P_D is contained in L_D. This means that P_D has dimension one, that is, (2) holds.
Suppose that (2) holds. Then the tropical rank of the matrix $-D$ is equal to two, by [6, Theorem 4.2]. This means that all $r \times r$-minors of $-D$ are tropically singular for $r \geq 3$. The case $r = 4$ is precisely the statement (3).
Obviously, the condition (3) implies the condition (4). What remains is to prove the implication from (4) to (5). For this we note that the tropical determinant of the 4×4-matrix

$$\begin{pmatrix} 0 & -d_{12} & -d_{13} & -d_{14} \\ -d_{12} & 0 & -d_{23} & -d_{24} \\ -d_{13} & -d_{23} & 0 & -d_{34} \\ -d_{14} & -d_{24} & -d_{34} & 0 \end{pmatrix}$$

equals twice the minimum of $-d_{12} - d_{34}$, $-d_{13} - d_{24}$ and $-d_{14} - d_{23}$. (It's the tropicalization of a 4×4-Pfaffian). The matrix is tropically singular if and only if the minimum is attained twice. $\square$

If the five equivalent conditions of Theorem 28 are satisfied then the metric tree T coincides with the one-dimensional tropical polytope P_D. To make sense of this statement, we regard tropical projective space $\mathbb{TP}^{n-1}$ as a metric space with respect to the infinity norm induced from $\mathbb{R}^n$,

$$||x - y|| \quad = \quad \max\big\{ \, |x_i + y_j - x_j - y_i| \; : \; 1 \leq i < j \leq n \, \big\},$$

and we note that the finite metric D embeds isometrically into P_D via the rows of $-\frac{1}{2}D$:

$$i \quad \mapsto \quad \frac{1}{2} \cdot (-d_{i1}, -d_{i2}, -d_{i3}, \ldots, -d_{in}) \qquad \text{for } i = 1, 2, \ldots, n$$

We learned from [8] that the tropical polytope P_D first appeared in the 1964 paper [10] by John Isbell. For the proof of the following result we assume familiarity with results from [7] and [8].

THEOREM 29. *The tropical polytope P_D equals Isbell's injective hull of the metric D.*

Proof. According to Lemma 22, the tropical polytope P_D is the bounded complex of the following unbounded polyhedron in the $(2n-1)$-dimensional space $W = \mathbb{R}^{2n}/\mathbb{R}(1,\ldots,1,-1,\ldots,-1)$:

$$\mathcal{P}_{-D} \quad = \quad \big\{ (y,z) \in W \;:\; y_i + z_j \leq -d_{ij} \text{ for all } 1 \leq i,j \leq n \big\}.$$

Dress et al. [7] showed that the injective hull $T(D)$ of the finite metric D coincides with the complex of bounded faces of the following n-dimensional unbounded polyhedron:

$$\mathcal{Q}_{-D} \quad = \quad \big\{ x \in \mathbb{R}^n \;:\; x_i + x_j \geq d_{ij} \text{ for all } 1 \leq i,j \leq n \big\}.$$

What we need to show is that the two polyhedra have the same bounded complex.

The metric D satisfies the tropical matrix identity $-D = D \odot (-D)$, because $-d_{ij} = \min_k(d_{ik} - d_{kj})$. This implies that any column vector y of $-D$ satisfies $y = (-y) \odot (-D)$.

Consider any vertex (y,z) of $\mathcal{P}_{-D}$. Then y is a column vector of $-D$. Equation (14) implies $z = (-y) \odot (-D) = y$. Hence every vertex of $\mathcal{P}_{-D}$ lies in the subspace defined by $y = z$, and so does the complex of bounded faces of $\mathcal{P}_{-D}$. Therefore the linear map $(y,z) \mapsto -y$ induces an isomorphism between the bounded complex of $\mathcal{P}_{-D}$ and the bounded complex of $\mathcal{Q}_{-D}$. □

Theorem 23 specifies an involution on the set of all tropical complexes. We are interested in the fixed points of this canonical involution. A necessary condition is that $r = n$ and V is a symmetric matrix. The previous result and its proof can be reinterpreted as follows:

COROLLARY 30. *A tropical complex P is pointwise fixed under the canonical involution (on the set of all tropical complexes) if and only if P is the injective hull of a metric on $\{1,2,\ldots,n\}$.*

Proof. In order for P to be fixed under the canonical involution, it is necessary that $n = d$. Hence we can write $P = \text{tconv}(-D)$ for some non-negative square matrix D. Now, P is fixed under the involution if and only if the identity $-D = D \odot (-D)$ holds. This identity is equivalent to D being a metric. □

Dress, Huber and Moulton [7] emphasize that the tropical polytope P_D records many important invariants of a given finite metric D. For instance, the dimension of P_D gives information about how far the metric is from being a tree metric. In practical biological applications of phylogenetic reconstruction, the distances d_{ij} are not known exactly, and P_D appears to contain many of the various trees which are found by existing software for phylogenetic reconstruction.

The dimension of the tropical complex $P_D = \mathrm{tconv}(-D)$ can be characterized combinatorially by tropicalizing the sub-Pfaffians of a skew-symmetric $n \times n$-matrix. The tropical Pfaffians of format 4×4 specify the four point condition (5) in Theorem 28, while the tropical sub-Pfaffians of format 6×6 specify the six-point condition which is discussed in [7, page 25]. The combinatorial study of *k-compatible split systems* can be interpreted in the setting of tropical algebraic geometry (cf. [15], [18], [20]) as the study of the k-th *secant variety* in the Grassmannian $\mathrm{Gr}(2,n) \subset \mathbb{TP}^{\binom{n}{2}}$.

Tropical convexity provides a convenient language to study numerous extensions of the classical problem of tree reconstruction. As an example, imagine the following scenario, which would correspond to the Grassmannian of planes in $\mathbb{TP}^{n-1}$, denoted $\mathrm{Gr}(3,n)$.

Suppose there are n taxa, labeled $1,2,\ldots,n$, and rather than having a distance for any pair i,j, we are now given a proximity measure d_{ijk} for any triple $i,j,k \in \{1,2,\ldots,n\}$. We can then construct a tropical polytope by taking the tropical convex hull of $\binom{n}{2}$ points as follows:

$$P \quad = \quad \mathrm{tconv}\,\bigl\{\, \bigl(-d_{ij1}, -d_{ij2}, -d_{ij3}, \ldots - d_{ijn}\bigr) \in \mathbb{TP}^{n-1} \;:\; 1 \leq i < j \leq n \,\bigr\}.$$

Under certain hypotheses, the tropical polytope P can be realized as the complex of bounded faces of the polyhedron in $\mathbb{R}^n$ defined by the inequalities $x_i + x_j + x_k \geq d_{ijk}$. It provides a polyhedral model for the tree-like nature of the data (d_{ijk}). The case of most interest is when P is two-dimensional in which case it plays the role of a *two-dimensional phylogenetic tree*.

The construction of this particular tropical polytope P was pioneered by Dress and Terhalle in the important paper [9]. There they discuss *valuated matroids*, which are essentially the points on the tropical Grassmannian of [20], and they call P the *tight span of a valuated matroid*. We share their view that these tropical polytopes constitute a promising tool for phylogenetic analysis.

Acknowledgements

This work was conducted while Mike Develin held the AIM Postdoctoral Fellowship 2003-2008 and Bernd Sturmfels held the MSRI-Hewlett Packard Professorship 2003/2004. Sturmfels also acknowledges partial support from the National Science Foundation (DMS-0200729).

References

[1] L.J. Billera, R. Cushman, and J.A. Sanders: The Stanley decomposition of the harmonic oscillator, *Nederl. Akad. Wetensch. Indag. Math.* 50 (1988), 375–393.

[2] P. Buneman: A note on metric properties of trees, *Journal of Combinatorial Theory, Ser. B* 17 (1974) 48–50.

[3] G. Cohen, S. Gaubert, and J.-P. Quadrat: Duality and separation theorems in idempotent semimodules, `arXiv:math.FA/0212294`.

[4] I.M. Gel'fand, M.I. Graev and A. Postnikov: Combinatorics of hypergeometric functions associated with positive roots, in *Arnold-Gel'fand Mathematical Seminars, Geometry and Singularity Theory*, (Eds. V.I. Arnold, I.M. Gel'fand, M. Smirnov and V.S. Retakh), Birkhäuser, Boston, 1997, pp. 205–221.

[5] J. De Loera, J. Rambau and F. Santos: *Triangulations of Point Configurations*, Algorithms and Computation in Mathematics, Springer Verlag, Heidelberg, to appear.

[6] M. Develin, F. Santos and B. Sturmfels: On the rank of a tropical matrix, `arXiv:math.CO/0312114`.

[7] A. Dress, K.T. Huber and V. Moulton: An explicit computation of the injective hull of certain finite metric spaces in terms of their associated Buneman complex, *Advances in Mathematics* 168 (2002) 1–28.

[8] A. Dress, V. Moulton and W. Terhalle: T-Theory – an overview, *European Journal of Combinatorics* 17 (1996) 161–175.

[9] A. Dress and W. Terhalle: The tree of life and other affine buildings. Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998), *Documenta Mathematica*, Extra Vol. III, 1998, 565–574.

[10] J. Isbell: Six theorems about metric spaces, *Comment. Math. Helv.* 39 (1964) 65–74.

[11] M. Joswig: Tropical halfspaces, `arXiv:math.CO/0312068`.

[12] A.N. Kirillov: Introduction to tropical combinatorics, in *Physics and Combinatorics 2000*, Proceedings of the Nagoya 2000 International Workshop, (Eds. A.N. Kirillov and N. Liskova), pp. 82–150, World Scientific, 2001.

[13] G. Litvinov, V. Maslov, and G. Shpiz: Idempotent functional analysis: an algebraic approach, *Mathematical Notes* 69 (2001) 696–729, `arXiv:math.FA/0009128`.

[14] J. Matoušek, On directional convexity, *Discrete and Computational Geometry* 25 (2001) 389–403.

[15] G. Mikhalkin: Counting curves via lattice paths in polygons, *C. R. Acad. Sci. Paris* 336 (2003), no. 8, 629–634.

[16] M. Noumi and Y. Yamada: Tropical Robinson-Schensted-Knuth correspondence and birational Weyl group actions, `arXiv:math-ph/0203030`.

[17] J. Rambau: TOPCOM (Triangulations of Point Configurations and Oriented Matroids), software available at `http://www.zib.de/rambau/TOPCOM/`. The numbers cited for regular triangulations of products of simplices can be found at `http://www.zib.de/rambau/TOPCOM/TOPCOM-examples/`.

[18] J. Richter-Gebert, B. Sturmfels, and T. Theobald: First steps in tropical geometry, `arXiv:math.AG/0306366`, to appear in *Idempotent Mathematics and Mathematical Physics*, Proceedings Vienna 2003, (editors G.L. Litvinov and V.P. Maslov), American Mathematical Society, 2004.

[19] F. Santos: The Cayley Trick and triangulations of products of simplices, `arXiv:math.CO/0312069`.

[20] D. Speyer and B. Sturmfels: The tropical Grassmannian, `arXiv:math.AG/0304218`, to appear in *Advances in Geometry*.

[21] K.A. Zaretsky: Reconstruction of a tree from the distances between its pendant vertices, *Uspekhi Math. Nauk. Russian Mathematical Surveys* 20 (1965) 90–92.

Mike Develin
American Inst. of Math
360 Portage Ave.
Palo Alto, CA 94306-2244
USA
develin@post.harvard.edu

Bernd Sturmfels
Dept. of Mathematics
UC-Berkeley
Berkeley, CA 94720
USA
bernd@math.berkeley.edu

Rigidity II: Non-Orientable Case

Serge Yagunov[1]

Received: August 3, 2003
Revised: March 4, 2004

Communicated by Ulf Rehmann

Abstract. The following paper is devoted to the construction of transfer maps (Becker-Gottlieb transfers) for non-orientable cohomology theories on the category of smooth algebraic varieties. Since non-orientability makes obstruction to the existence of transfer structure, we define transfers for a specially constructed class of morphisms. Being rather small, this class is yet big enough for application purposes. As an application of the developed transfer technique we get the proof of rigidity theorem for all cohomology theories represented by T-spectra.

2000 Mathematics Subject Classification: 14Fxx
Keywords and Phrases: non-orientable cohomology theories, rigidity, Becker-Gottlieb transfers, hermitian K-theory, higher Witt groups.

Introduction

The purpose of the current paper is to generalize the results obtained in [PY] to non-orientable cohomology theories on the category of smooth varieties over an algebraically closed field. This generalization seems to be especially interesting after a recent paper of Hornbostel [Ho] who proved T-representability of higher Witt groups and Hermitian K-theory. These results give us two good examples of non-orientable theories, which have important algebraic and arithmetic meaning.

In our proof we mostly follow the strategy described in [PY]. This, roughly speaking, includes constructing of transfer maps for a given theory, checking such fundamental properties as commutativity of base-change diagrams for transversal squares, finite additivity, and normalization. Finally, we use these properties to establish the main result (Rigidity theorem). Employing further the technique of Suslin [Su1], one can obtain the results similar to ones obtained in Suslin's paper for the extension of algebraically closed fields. Slightly

[1]Partially supported by TMR ERB FMRX CT-97-0107, INTAS, and RFFI 00-01-00116 grants.

adapting methods of Gabber [Ga] we may generalize the rigidity property for Hensel local rings to an arbitrary cohomology theory.

Certainly, this program would fail already on the first step, because of the result of Panin [Pa] showing that there exists an one-to-one correspondence between orientations and transfer structures. However, shrinking the class of morphisms for which we define transfers to some smaller class $\mathcal{C}_{\mathrm{triv}}$ we may construct a satisfactory transfer structure. On the other hand, the class $\mathcal{C}_{\mathrm{triv}}$ is still big enough to be used in the proof of the rigidity theorem.

Acknowledgements. I'm very grateful to Ivan Panin for really inspiring discussions during the work. Probably, this paper would never appear without his help. I'm very grateful to Ulf Rehmann, who invited me to stay in the very nice working environment of the University of Bielefeld, where the text was mostly written. Also I would like to thank Jens Hornbostel for proofreading of the draft version of this paper and really valuable discussions concerning rigidity for the Henselian case.

Notation remarks. We use the standard 'support' notation for cohomology of pairs and denote $A(X,U)$ by $A_Z(X)$, provided that U is an open subscheme of X and $Z = X - U$. Moreover, in this case we often denote the pair (X,U) by $(X)_Z$.

We omit grading of cohomology groups whenever it is possible. However, to make the T-suspension isomorphism compatible to the usual notation, we write $A^{[d]}$ for cohomology shifted by d. For example, if A denotes a cohomology theory $A^{*,*}$ represented by a T-spectrum, we set $A^{[d]} = A^{*+2d,*+d}$.

For a closed smooth subscheme $Z \subset X \in Sm/k$ we denote by $B(X,Z)$ the deformation to the normal cone of Z in X. Namely, we set $B(X,Z)$ to be the blow-up of $X \times \mathbb{A}^1$ with center at $Z \times \{0\}$. More details for this well-known construction may be found in [Fu, Chapter 5],[MV, Theorem 3.2.23], or [Pa].

The notation pt is reserved for the final object $\operatorname{Spec} k$ in Sm/k.

1. Rigidity Theorem

Denote by $\mathcal{C}_{\mathrm{triv}}$ the class of equipped morphisms (f,Θ) where f is decomposed as $f\colon X \overset{\tau}{\hookrightarrow} Y \times \mathbb{A}^n \xrightarrow{p} Y$ such that τ is a closed embedding with trivial normal bundle $\mathcal{N}_{Y\times\mathbb{A}^n/X}$, p is a projection morphism, and $\Theta\colon \mathcal{N}_{Y\times\mathbb{A}^n/X} \cong X \times \mathbb{A}^N$ is a trivialization isomorphism. Abusing the notation we often omit Θ if the trivialization is clear from the context. The main purpose of this paper is to show that the class $\mathcal{C}_{\mathrm{triv}}$ may be endowed with a transfer structure, which makes given cohomology theory A a functor with weak transfers (see [PY, Definition 1.8]) with respect to this class. We also show that $\mathcal{C}_{\mathrm{triv}}$ is still big enough to fit all the requirements of constructions used in [PY] to prove the Rigidity Theorem. This, finally, yields Theorem 1.10, which may be applied to concrete examples of theories.

Let Sm/k be a category of smooth varieties over an algebraically closed field k. Denote by Sm^2/k a category whose objects are pairs (X,Y), where $X,Y \in Sm/k$, the scheme Y is a locally closed subscheme in X and morphisms are defined in a usual way as morphisms of pairs. A functor $\mathcal{E}\colon X \mapsto (X, \emptyset)$ identifies Sm/k with a full subcategory of Sm^2/k.

Definition 1.1. *We say that a functor $F\colon Sm/k \to G$ admits extension to pairs by a functor $\mathcal{F}\colon Sm^2/k \to G$ if $F = \mathcal{F} \circ \mathcal{E}$.*

Definition 1.2. *We call a contravariant functor $\mathcal{A}\colon Sm^2/k \to Gr\text{-}Ab$ to the category of graded abelian groups a cohomology theory if it satisfies the following four properties:*

(1) Suspension Isomorphism. *For a scheme $X \in Sm/k$ and its open subscheme U set $W = X - U$. Then, we are given a functorial isomorphism*

$$\mathcal{A}_W(X) \overset{\Sigma}{\cong} \mathcal{A}^{[1]}_{W \times \{0\}}(X \times \mathbb{A}^1)$$

induced by the T-suspension morphism.

(2) Zariski Excision. *Let $X \overset{i}{\supseteq} X_0 \supseteq Z$ be objects of Sm/k such that X_0 is open in X and Z is closed in X. Then, the induced map $i^*\colon \mathcal{A}_Z(X) \overset{\cong}{\to} \mathcal{A}_Z(X_0)$ is an isomorphism.*

(3) Homotopy Invariance. *For every $(X, Y) \in Sm^2/k$ the map $p^*\colon \mathcal{A}(X, Y) \to \mathcal{A}(X \times \mathbb{A}^1, Y \times \mathbb{A}^1)$ induced by the projection is an isomorphism.*

(4) Homotopy purity. *Let $Z \subset Y \subset X \in Sm/k$ be closed embeddings of smooth varieties. Let $\mathcal{N}$ be the corresponding normal bundle over Y, $i_0\colon \mathcal{N} \hookrightarrow B(X, Y)$ and $i_1\colon X \hookrightarrow B(X, Y)$ be canonical embeddings over 0 and 1, respectively. Then, the induced maps:*

$$\mathcal{A}_Z(\mathcal{N}) \underset{\cong}{\overset{i_0^*}{\leftarrow}} \mathcal{A}_{Z \times \mathbb{A}^1}(B(X, Y)) \underset{\cong}{\overset{i_1^*}{\to}} \mathcal{A}_Z(X)$$

are isomorphisms.

Definition 1.3. *We call a contravariant functor $A\colon Sm/k \to Gr\text{-}Ab$ a cohomology theory if it admits an extension to pairs by the functor $\mathcal{A}$ which is a cohomology theory.*

In what follows we often use the same notation for functors and their extensions to pairs. We also usually identify objects X and $(X, \emptyset)$.
Most important examples of cohomology theories may be obtained in the following way.

Example 1.4. *Every functor represented by a T-spectrum in the sense of Voevodsky (see* [Vo]*) is a cohomology theory.*
Since the category of spaces, introduced by Voevodsky [Vo, p.583]*, has fibred coproducts, we can extend any functor $A\colon Sm/k \to G$ to Sm^2/k setting $\mathcal{A}(X, Y) = A(X/Y)$. All functors represented by T-spectra satisfy conditions (1)–(3) of Definition 1.2 (see* [MV, PY, Pa]*). Condition (4) is actually Theorem 2.2.8 from* [Pa].

Theorem 1.5. *Every cohomology theory A given on the category Sm/k of smooth varieties over an algebraically closed field k may be endowed with the structure of a functor with weak transfers for the class $\mathcal{C}_{\mathrm{triv}}$, i.e. for every $f\colon X \to Y \in \mathcal{C}_{\mathrm{triv}}$ we assign the transfer map $f_!\colon A(X) \to A(Y)$, which satisfy properties 3.1–3.3 below.*

We postpone the proof of this theorem till the last section and show, first, that the class $\mathcal{C}_{\mathrm{triv}}$ is big enough to make the proof of the Rigidity Theorem given at [PY] running. We reproduce here some constructions and arguments from [PY]. From now on we consider the case of algebraically closed base field k. Let $A\colon (Sm/k) \to \text{Gr-Ab}$ be a cohomology theory and X be a smooth curve over k. We can construct a map $\Phi\colon \mathrm{Div}(X) \to \mathrm{Hom}(A(X), A(k))$ defined on canonical generators as: $[x] \mapsto x^*$, where $x^*\colon A(X) \to A(k)$ is the pull-back map, corresponding to the point $x \in X(k)$.

Theorem 1.6. *Let X be a smooth affine curve with trivial tangent bundle, $\bar{X}$ be its projective completion, and $X_\infty = \bar{X} - X$. Let also A be a homotopy invariant contravariant functor with weak transfers for the class $\mathcal{C}_{\mathrm{triv}}$. Then, the map Φ can be decomposed in the following way:*

$$\begin{tikzcd}
\mathrm{Div}(X) \arrow[rr, "\Phi"] \arrow[dr, "\Omega"'] & & \mathrm{Hom}(A(X), A(k)) \\
& \mathrm{Pic}(\bar{X}, X_\infty), \arrow[ur, "\Psi"'] &
\end{tikzcd}$$

where $\mathrm{Pic}(\bar{X}, X_\infty)$ *is the relative Picard group (see* [SV]*) and the map Ω is the canonical homomorphism.*

Proof. Let us recall that a divisor $\mathcal{D}$ lies in the kernel of Ω if and only if there exists a function $f \in k(\bar{X})$ such that $f|_{X_\infty} = 1$ and $\mathcal{D} = [f]$. We denote zero and pole locuses of f by $\mathrm{div}_0(f) = D$, and $\mathrm{div}_\infty(f) = D'$, respectively. It is now sufficient to check that $\Phi(D) = \Phi(D')$.
Denote by X^0 the open locus $f \neq 1$ on $\bar{X}$. By the choice of the function f, we have: $i\colon X^0 \subset X$ and the morphism $f\colon X^0 \to \mathbb{P}^1 - \{1\} = \mathbb{A}^1$ is finite. This shows the existence of a decomposition $f\colon X^0 \xrightarrow{\tau} \mathbb{A}^n \xrightarrow{p} \mathbb{A}^1$, with closed embedding τ and projection p. For the normal bundle $\mathcal{N}_{\mathbb{A}^n/X^0}$ we have a short exact sequence:

$$0 \longrightarrow T_{X^0} \longrightarrow \tau^*(T_{\mathbb{A}^n}) \longrightarrow \mathcal{N}_{\mathbb{A}^n/X^0} \longrightarrow 0. \tag{1.1}$$

Since the tangent bundles T_{X^0} and $T_{\mathbb{A}^n}$ are trivial, the normal bundle $\mathcal{N}_{\mathbb{A}^n/X^0}$ is stably trivial. Finally, because every stably trivial bundle over a curve is trivial, we have: $f \in \mathcal{C}_{\mathrm{triv}}$.
Moreover, since we may assume (as well as in [PY, Proof of Theorem 1.11]) that the corresponding divisors are unramified, the map f is étale over the

points $\{0\}$ and $\{\infty\}$. Consider now the diagram:

$$
\begin{tikzcd}
 & & A(X) \arrow[d, "i^*"] & & \\
\bigoplus\limits_{x/\infty} A(x) \arrow[dr, "\sum\limits_{x/\infty} \tilde{f}_{x,!}"'] & A(D') \arrow[l] \arrow[d, "\tilde{f}_{\infty,!}"] & A(X^0) \arrow[l] \arrow[r] \arrow[d, "\tilde{f}_!"] & A(D) \arrow[r] \arrow[d, "\tilde{f}_{0,!}"] & \bigoplus\limits_{x/0} A(x) \arrow[dl, "\sum\limits_{x/0} \tilde{f}_{x,!}"] \\
 & A(pt) & A(\mathbb{A}^1) \arrow[l, "i^*_\infty"'] \arrow[r, "i^*_0"] & A(pt) &
\end{tikzcd}
\tag{1.2}
$$

where $\tilde{f}_{0,!}$ ($\tilde{f}_{\infty,!}$) denotes the transfer map corresponding to the morphism $\tilde{f}_0$ (resp. $\tilde{f}_\infty$), which is a restriction of the morphism $\tilde{f}$ to the divisor D (D', respectively). Due to the discussion above, all vertical arrows are well defined, since all corresponding morphisms belong to $\mathcal{C}_{\mathrm{triv}}$. (For morphisms $\tilde{f}_0$ and $\tilde{f}_\infty$ we choose trivialization maps as restrictions of Θ.)

Using the standard properties of the functor with weak transfers (see 3.1–3.3), one can see that the diagram above is commutative.

On the other hand, it is easy to check that going from $A(X)$ to two different copies of $A(pt)$ one obtains the maps $\Phi(D)$ and $\Phi(D')$, respectively.

Finally, using the homotopy invariance of the functor A, one has:

$$\Phi(D) = i_0^* \tilde{f}_! i^* = i_\infty^* \tilde{f}_! i^* = \Phi(D'). \tag{1.3}$$

$\square$

Since the group $\mathrm{Pic}(\bar{X}, X_\infty)^\circ$ of relative divisors of degree 0 is n-divisible over an algebraically closed field of characteristic relatively prime to n, we get the following corollary:

COROLLARY 1.7. *Let us assume, in addition to the hypothesis of Theorem 1.6, that there exists an integer n coprime to the exponential characteristic* $\mathrm{Char}(k)$ *such that $nA(Y) = 0$ for any $Y \in Sm/k$. Then, the map Ψ can be passed through the degree map* $\mathrm{Pic}(\bar{X}, X_\infty) \overset{deg}{\to} \mathbb{Z}$. *Namely, if D, D' are two divisors of the same degree, one has:* $\Phi(D) = \Phi(D') \colon A(X) \to A(\mathrm{Spec}\, k)$.

Now we want to get rid of the normal bundle triviality assumption. For this end, we need the following simple geometric observation.

LEMMA 1.8. *For a smooth curve X and a divisor D on X one can choose such an open neighborhood X^0 of* $\mathrm{Supp}\, D$ *that the tangent bundle T_{X^0} is trivial.*

Proof. Let Υ be an invertible sheaf on X corresponding to the tangent bundle T_X. Denote by $\mathcal{O}_{X,D}$ a localization of $\mathcal{O}_X$ at the support of the divisor D. The scheme $\mathrm{Spec}\, \mathcal{O}_{X,D}$ is a spectrum of a regular semi-local ring endowed with a natural morphism $j \colon \mathrm{Spec}\, \mathcal{O}_{X,D} \to X$. Therefore, the sheaf $j^*\Upsilon$ is free. This means there exists an open neighborhood $X^0 \supset \mathrm{Supp}\, D$ such that the restriction $\Upsilon|_{X^0}$ is free as well. $\square$

The proofs of following two theorems are same, word by word, to ones of Theorems 1.13 and 2.17 from [PY].

Theorem 1.9 (The Rigidity Theorem). *Let $A\colon Sm/k \to Gr\text{-}Ab$ be a homotopy invariant functor with weak transfers for the class $\mathcal{C}_{\mathrm{triv}}$. Assume that the field k is algebraically closed and $nA = 0$ for some integer n, coprime to* $\operatorname{Char} k$. *Then, for every smooth affine variety V and any two k-rational points $t_1, t_2 \in V(k)$ the induced maps $t_1^*, t_2^*\colon A(V) \to A(\operatorname{Spec} k)$ coincide.*

Theorem 1.10. *Let $k \subset K$ be an extension of algebraically closed fields. Let also A be a cohomology theory vanishing after multiplication by n, coprime to the exponential field characteristic. Then, for any $X \in Sm/k$, we have:*

$$A(X) \xrightarrow{\cong} A(X_K).$$

Besides Theorem 1.10 we would like to mention briefly the following nice application of the developed technique [2]

Theorem 1.11. *Let A and k be as above, $M \in Sm/k$, and R be the henselization of M at some closed point. Then, the map*

$$A(R) \xrightarrow{\cong} A(\operatorname{Spec} k)$$

is an isomorphism.

The proof of the theorem may be achieved as a direct compilation of Gabber [Ga], Suslin-Voevodsky [Su2, SV] (see also an approach of [GT]) results and Theorem 1.6. By general strategy, one reduces the above statement to a form of the Rigidity Theorem. Namely, it is possible to construct such a curve M over the field k with some special divisor D that the statement of Theorem 1.11 would follow from the fact that $\Phi(D) = 0$. The divisor D, by its construction, can be written in the form $D = n \cdot \tilde{D} + [f]$, for some divisor $\tilde{D}$ and rational function f on M (as follows from the proper base change theorem [Mi, SGA4]). Finally, it is sufficient to apply the statement of Theorem 1.6 to complete the proof.

2. Becker–Gottlieb Transfers

In this section we construct transfer maps required in Theorem 1.5. First of all, we build transfers with support for closed embeddings. Let $W \hookrightarrow X \overset{f}{\hookrightarrow} Y$ be closed embeddings such that $W, X, Y \in Sm/k$ and $(f, \Theta) \in \mathcal{C}_{\mathrm{triv}}$ of codimension n. We now define a map $(f, \Theta)_!\colon A_W(X) \to A_W^{[n]}(Y)$. Consider, first, following isomorphisms:

$$\varphi_W(\Theta)\colon A_W(X) \xrightarrow[\cong]{\Sigma^n} A^{[n]}_{W\times\{0\}}(X \times \mathbb{A}^n) \xrightarrow[\cong]{\Theta^*} A_W^{[n]}(\mathcal{N}_{Y/X}). \tag{2.1}$$

The next step involves Homotopy Purity property. Consider the map:

$$\chi_W : A_W(\mathcal{N}_{Y/X}) \xrightarrow[\cong]{(i_0^*)^{-1}} A_{W\times\mathbb{A}^1}(B(Y,X)) \xrightarrow[\cong]{i_1^*} A_W(Y), \tag{2.2}$$

[2] This observation was obtained jointly with Jens Hornbostel.

Definition 2.1. *The composite map:* $(f,\Theta)_!^W = \chi_W \circ \varphi_W(\Theta)\colon A_W(X) \to A_W^{[n]}(Y)$ *is called Becker–Gottlieb transfer for the closed embedding* f *with support* W.

In case $W = X$ we often omit any mentioning of the support. One can easily verify that the defined transfer map commutes with support extension homomorphisms. Namely, the following lemma holds.

Lemma 2.2. *Suppose we have a chain of closed embeddings:* $W_2 \hookrightarrow W_1 \hookrightarrow X \xhookrightarrow{f} Y$, *with* $f \in \mathcal{C}_{\mathrm{triv}}$. *Then, the diagram*

$$\begin{tikzcd} A_{W_2}(X) \arrow[r, "f_!^{W_2}"] \arrow[d, "ext."'] & A_{W_2}(Y) \arrow[d, "ext."] \\ A_{W_1}(X) \arrow[r, "f_!^{W_1}"] & A_{W_1}(Y) \end{tikzcd}$$

commutes. (Here ext. denotes the support extension homomorphism.)

Construction-Definition 2.3. *Let now* $(f\colon X \to Y, \Theta) \in \mathcal{C}_{\mathrm{triv}}$ *be a morphism of relative dimension* d *endowed with a decomposition* $X \xhookrightarrow{\tau} Y \times \mathbb{A}^n \xrightarrow{p} Y$, *with closed embedding* τ *and projection* p. *We define Becker–Gottlieb*[3] *transfer map* $(f,\Theta)_!$ *in the following way. Consider the standard open embedding* $\mathbb{A}^n \xhookrightarrow{j} \mathbb{P}^n$ *and denote the complement of* $\mathbb{A}^n$ *by* $\mathbb{P}_\infty$. *The following morphisms of pairs are induced by standard embeddings:*

(2.3)

$$(Y \times \mathbb{A}^n)_X \xhookrightarrow{j_X} (Y \times \mathbb{P}^n)_X \xleftarrow{\alpha} (Y \times \mathbb{P}^n, Y \times \mathbb{P}_\infty) \xrightarrow{\beta} (Y \times \mathbb{P}^n)_{Y\times\{0\}} \xhookleftarrow{j_Y} (Y \times \mathbb{A}^n)_{Y\times\{0\}}.$$

Since the morphism β *identifies* $\mathbb{P}_\infty$ *with zero-section of the line bundle* $\mathbb{P}^n - \{0\}$ *over* $\mathbb{P}^{n-1}$, *it induces an isomorphism of cohomology groups* $\beta^*\colon A_Y(Y \times \mathbb{P}^n) \xrightarrow{\cong} A(Y \times \mathbb{P}^n, Y \times \mathbb{P}_\infty)$. *The morphism* j_X *gives us an excision isomorphism* $j_X^*\colon A_X(Y \times \mathbb{P}^n) \xrightarrow{\cong} A_X(Y \times \mathbb{A}^n)$.
We define $(f,\Theta)_!$ *as a the following composite map:*

(2.4)

$$A(X) \xrightarrow{(\tau,\Theta)_!} A_X^{[d+n]}(Y \times \mathbb{A}^n) \xrightarrow{j_Y^* \circ (\beta^*)^{-1} \circ \alpha^* \circ (j_X^*)^{-1}} A_{Y\times\{0\}}^{[d+n]}(Y \times \mathbb{A}^n) \xrightarrow[\cong]{\Sigma^{-n}} A^{[d]}(Y),$$

where Σ^{-n} *denotes the* n*-fold* T*-desuspension. We usually denote the map* $\Sigma^{-n} \circ j_Y^* \circ (\beta^*)^{-1} \circ \alpha^* \circ (j_X^*)^{-1}$ *by* $p_!$.

3. Proof of Theorem 1.5

We now prove Theorem 1.5 checking consequently all necessary properties of a functor with weak transfers.

[3] We call the transfer maps under construction Becker–Gottlieb transfers, since we generally follow the philosophy of their paper [BG]. However, our algebraic construction is much more restrictive than the original topological one.

PROPOSITION 3.1 (BASE CHANGE PROPERTY). *Given a diagram with Cartesian squares*

$$\begin{tikzcd}
X' \arrow[r, hook, "\tau'"] \arrow[d, "g'"] \arrow[rr, bend left, "f'"] & Y' \times \mathbb{A}^n \arrow[r] \arrow[d] & Y' \arrow[d, "g"] \\
X \arrow[r, hook, "\tau"] \arrow[rr, bend right, "f"'] & Y \times \mathbb{A}^n \arrow[r] & Y
\end{tikzcd}$$

where $f \in \mathcal{C}_{\mathrm{triv}}$ of codimension d, and morphisms τ, τ' are closed embeddings such that the left-hand-side square is transversal. We also require Θ' to be a base-change of Θ in the sense that the square:

$$\begin{tikzcd}
X' \times \mathbb{A}^{d+n} \arrow[d, "g' \times \mathrm{id}"'] & \mathcal{N}_{Y' \times \mathbb{A}^n / X'} \arrow[l, "\Theta'"'] \arrow[d, "N(g')"] \arrow[r, equal] & \mathcal{N}_{Y \times \mathbb{A}^n / X} \underset{X}{\times} X' \\
X \times \mathbb{A}^{d+n} & \mathcal{N}_{Y \times \mathbb{A}^n / X} \arrow[l, "\Theta"'] &
\end{tikzcd} \tag{3.1}$$

is Cartesian. Then, the diagram:

$$\begin{array}{ccc}
A(X') & \xrightarrow{(f', \Theta')_!} & A^{[d]}(Y') \\
{\scriptstyle g'^*}\big\uparrow & & {\scriptstyle g^*}\big\uparrow \\
A(X) & \xrightarrow{(f, \Theta)_!} & A^{[d]}(Y).
\end{array}$$

commutes.

Proof. Let us look at the diagram appearing on the first step of the computation of $f_!$.

$$\begin{tikzcd}
A(X') \arrow[r, "\Sigma^{d+n}"] & A^{[d+n]}_{X' \times \{0\}}(X' \times \mathbb{A}^{d+n}) \arrow[r, "\Theta'^*"] & A^{[d+n]}_{X'}(\mathcal{N}_{Y' \times \mathbb{A}^n / X'}) \\
A(X) \arrow[u, "g'^*"] \arrow[r, "\Sigma^{d+n}"] & A^{[d+n]}_{X \times \{0\}}(X \times \mathbb{A}^{d+n}) \arrow[u, "\Sigma^n (g')^*"'] \arrow[r, "\Theta^*"] & A^{[d+n]}_{X}(\mathcal{N}_{Y \times \mathbb{A}^n / X}) \arrow[u, "N(g')^*"']
\end{tikzcd} \tag{3.2}$$

The left-hand-side square commutes because of the suspension functoriality. Commutativity of the right-hand-side one follows from Diagram 3.1. Going further along the construction of $f_!$, we may see that all other squares whose commutativity has to be checked are either commute already in the category Sm^2/k or include (de-)suspension isomorphisms like the very left one. This shows the required base-change diagram commutes. □

PROPOSITION 3.2 (ADDITIVITY). *Let $X = X_0 \sqcup X_1 \in Sm/k$ be a disjoint union of subvarieties X_0 and X_1, $e_m \colon X_m \hookrightarrow X$ $(m = 0, 1)$ be embedding maps, and*

$(f\colon X \to Y, \Theta) \in \mathcal{C}_{\mathrm{triv}}$ $(\operatorname{codim} f = d)$. Setting $f_m = f \circ e_m$, we have:

$$f_{0,!}e_0^* + f_{1,!}e_1^* = f_!.$$

(All necessary decompositions and trivializations for morphisms f_0, f_1 are assumed to be corresponding restrictions of ones for f.)

Proof. As it follows from the proof of [PY, Proposition 4.3], in order to show the additivity property it is sufficient to check the commutativity of the following pentagonal diagram:

(3.3)
$$\begin{tikzcd}
A(X_0) \arrow[d, "\tau_{0,!}"'] & A_{X_0}(X) \arrow[l, "\psi^*"', "\cong"] \arrow[r, "\varphi^*"] \arrow[ld, dotted] & A(X) \arrow[d, "\tau_!"] \\
A_{X_0}^{[d+n]}(Y \times \mathbb{A}^n) \arrow[rr, "\chi^*"] \arrow[rd, "p_{0,!}"'] & & A_X^{[d+n]}(Y \times \mathbb{A}^n) \arrow[ld, "p_!"] \\
& A^{[d]}(Y), &
\end{tikzcd}$$

where ψ^* is the excision homomorphism, φ^* and χ^* are extension of support maps, and the composites $p_{0,!} \circ \tau_{0,!}$ ($p_! \circ \tau_!$, resp.) form the transfer maps $f_{0,!}$ ($f_!$, respectively). We prove, first, the commutativity of the bottom triangle. Both oblique arrows may be factored through the group $A(Y \times \mathbb{P}^n, Y \times \mathbb{P}_\infty)$. Since the diagram:

(3.4)
$$\begin{tikzcd}
(Y \times \mathbb{A}^n)_{X_0} \arrow[d, "j_{X_0}"'] & & (Y \times \mathbb{A}^n)_X \arrow[ll, hook', "\chi"'] \arrow[d, "j_X"] \\
(Y \times \mathbb{P}^n)_{X_0} & & (Y \times \mathbb{P}^n)_X \arrow[ll, hook'] \\
& (Y \times \mathbb{P}^n, Y \times \mathbb{P}_\infty) \arrow[lu, hook'] \arrow[ru, hook] &
\end{tikzcd}$$

commutes already in Sm^2/k, the required triangle commutes as well. We now show the commutativity of the rectangular part of Diagram 3.3. Define the dotted map as a transfer with the support X_0 corresponding to the embedding $X \hookrightarrow Y \times \mathbb{A}^n = \mathcal{Y}$. This (due to Lemma 2.2) makes the right trapezium part of the diagram commutative. Commutativity of the upper-left triangle is equivalent, by the definition, to a claim that the following diagram commutes.
(3.5)

$$\begin{tikzcd}
A_{X_0}(X) \arrow[r, "\Sigma^{d+n}"] \arrow[d, "\psi^*"', "\cong"] \arrow[dr, phantom, "(1)"] & A_{X_0}^{[d+n]}(X \times \mathbb{A}^{d+n}) \arrow[r] \arrow[d, "\cong"] \arrow[dr, phantom, "(2)"] & A_{X_0}^{[d+n]}(\mathcal{N}_{\mathcal{Y}/X}) \arrow[d, "\cong"] \arrow[dr, phantom, "(3)"] & A_{X_0 \times \mathbb{A}^1}^{[d+n]}(B(\mathcal{Y}, X)) \arrow[l, "\cong"'] \arrow[r] \arrow[dr, phantom, "(4)"] & A_{X_0}^{[d+n]}(\mathcal{Y}) \arrow[d, equal] \\
A(X_0) \arrow[r, "\Sigma^{d+n}"] & A_{X_0}^{[d+n]}(X_0 \times \mathbb{A}^{d+n}) \arrow[r] & A_{X_0}^{[d+n]}(\mathcal{N}_{\mathcal{Y}/X_0}) & A_{X_0 \times \mathbb{A}^1}^{[d+n]}(B(\mathcal{Y}, X_0)) \arrow[l, "\cong"'] \arrow[r] \arrow[u, "\sigma^*"] & A_{X_0}^{[d+n]}(\mathcal{Y})
\end{tikzcd}$$

(Here the map σ^* is induced by a morphism blowing down a component X_1. See [PY, Section 4] for more details.) Simple arguments utilized at the end of the proof of Proposition 3.1 may be used here as well. Namely, the suspension isomorphism functoriality implies commutativity of square (1). Square (2) commutes, because its bottom horizontal arrow is the restriction of the top one. Squares (3) and (4) are induced by commutative diagrams of varieties. The proposition follows. $\square$

PROPOSITION 3.3 (NORMALIZATION). *For the morphism* $\mathrm{id}\colon pt \to pt = \mathrm{Spec}(k)$ *endowed with an arbitrary decomposition* $pt \overset{\tau}{\hookrightarrow} \mathbb{A}^n \to pt$ *the map* $\mathrm{id}_!\colon A(pt) \to A(pt)$ *is identical.*

In order to prove this statement we show, first, that in case of the identity map $pt \to pt$ the constructed transfer map does not depend on the choice of normal bundle trivialization.

LEMMA 3.4. *Let* $X \in Sm^2/k$ *be a smooth pair endowed with a linear action of* $SL_n(k)$. *Then, for any matrix* $\alpha \in SL_n(k)$ *the induced isomorphism* $A(X) \overset{\alpha^*}{\to} A(X)$ *is the identity map.*

Proof. Every matrix of $SL_n(k)$ may be written as a product of elementary matrices. Every elementary matrix acts trivially on X, because of existence of a canonical contracting homotopy $H(e_{ij}(a), t) = e_{ij}(at)$. $\square$

LEMMA 3.5. *The standard homothety action of* $\lambda \in k^*$ *on the affine line* $\mathbb{A}^1$ *induces the identity isomorphism* $A_{\{0\}}(\mathbb{A}^1) \overset{\lambda^*}{\to} A_{\{0\}}(\mathbb{A}^1)$.

Proof. Consider a diagram:

(3.6)
$$\begin{tikzcd} \mathbb{A}^1 \arrow[r, "\lambda"] \arrow[d, hook] & \mathbb{A}^1 \arrow[d, hook] \\ \mathbb{P}^1 \arrow[r, "\Lambda"] & \mathbb{P}^1, \end{tikzcd}$$

where $\Lambda = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$ and the vertical arrows are standard open embeddings given by: $a \mapsto (a\colon 1)$. Due to the excision axiom, this diagram yields the following commutative diagram of cohomology groups:

(3.7)
$$\begin{tikzcd} A_{\{0\}}(\mathbb{A}^1) & A_{\{0\}}(\mathbb{A}^1) \arrow[l, "\lambda^*"'] \\ A_{\{0\}}(\mathbb{P}^1) \arrow[u, "\cong"] & A_{\{0\}}(\mathbb{P}^1). \arrow[l, "\Lambda^*"'] \arrow[u, "\cong"'] \end{tikzcd}$$

Let us get rid of the support. Since the natural map $A(\mathbb{P}^1) \to A(\mathbb{A}^1) = A(pt)$ is split by the projection $\mathbb{P}^1 \to pt$, the cohomology long exact sequence shows that the support extension map $A_{\{0\}}(\mathbb{P}^1) \overset{ext}{\to} A(\mathbb{P}^1)$ is a monomorphism. The action of diagonal matrices clearly commutes with the map *ext*. Therefore, it is sufficient to show that the matrix Λ acts trivially on $A(\mathbb{P}^1)$. This matrix is SL_2-equivalent to the matrix $\begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\lambda} \end{pmatrix}$. (Let us recall that the field k is algebraically closed.) Due to Lemma 3.4 two SL_2-equivalent matrices induce the same action in cohomology and the latter matrix obviously acts trivially on $\mathbb{P}^1$. $\square$

LEMMA 3.6. *Any matrix of $GL_n(k)$ acting on $\mathbb{A}^n$ by left multiplication induces trivial action on cohomology groups $A_{\{0\}}(\mathbb{A}^n)$.*

Proof. Changing, if necessary, the acting matrix by its SL_n-equivalent, we may assume that the action is given by the diagonal matrix $\Lambda = diag(\lambda, 1, \dots, 1)$. Let us also mention that the pair $(\mathbb{A}^n, \mathbb{A}^n - \{0\})$ is the n-fold T-suspension of $T = (\mathbb{A}^1, \mathbb{A}^1 - \{0\})$. Since we have chosen the matrix Λ in a special way (acting only on the first factor), the suspension isomorphism and Lemma 3.5 complete the proof. $\square$

Proof of Proposition 3.3 Let us consider the chain of maps giving the transfer map: $\mathrm{id}_! : A(pt) \to A(pt)$. We take into account that the normal bundle to pt in $\mathbb{A}^n$ is canonically isomorphic to $\mathbb{A}^n$.

(3.8)

$$\begin{tikzcd}
A(pt) \arrow[d, "\Sigma^n"] \\
A_{\{0\}}(\mathbb{A}^n) \arrow[r, "(i_0^*)^{-1}"] \arrow[rr, dotted, bend right, "\delta"'] & A_{\mathbb{A}^1}(B(\mathbb{A}^n, pt)) \arrow[r, "i_1^*"] & A_{\{0\}}(\mathbb{A}^n) \arrow[r, "(j^*)^{-1}"] \arrow[dr, dotted, "\gamma"] & A_{\{0\}}(\mathbb{P}^n) \arrow[r, "\alpha^*"] & A(\mathbb{P}^n, \mathbb{P}_\infty) \arrow[d, "(\beta^*)^{-1}"'] \\
 & & & A_{\{0\}}(\mathbb{A}^n) \arrow[d, "\Sigma^{-n}"] & A_{\{0\}}(\mathbb{P}^n) \arrow[l, "j^*"'] \\
 & & & A(pt)
\end{tikzcd}$$

In this diagram j^* denotes the excision isomorphism and maps γ and δ are just set to be composites of the fitting arrows. As it was shown in [PY, Lemma 5.8], the map δ is identical. Since in the considered case both maps α^* and β^* are induced by the same embedding $(\mathbb{P}^n, \mathbb{P}_\infty) \hookrightarrow (\mathbb{P}^n, \mathbb{P}^n - \{0\})$, the map γ is also identical. This finishes the proof of Normalization property. $\square$

The latter three propositions actually check all the conditions required by the definition of a functor with weak transfers. This completes the proof of Theorem 1.5 as well.

REFERENCES

[BG] Becker, J. C.; Gottlieb, D. H. *The transfer map and fiber bundles.* Topology 14 (1975), pp. 1–12.

[Fu] Fulton, W. *Intersection theory.* Springer-Verlag, Berlin, 1998, xiv+470 pp.

[Ga] Gabber, O. *K-theory of Henselian Local Rings and Henselian Pairs.* Contemp. Math. 126 (1992), pp. 59-70.

[GT] H. Gillet and R. Thomason: *The K-theory of strict hensel local rings and a theorem of Suslin.* JPAA 34 (1984), pp. 241-254

[Ho] Hornbostel, J. $\mathbb{A}^1$*-representability of hermitian K-theory and Witt groups.* (preprint) `www.math.uiuc.edu/K-theory/0578/` 2002.

[Mi] Milne, J. *Étale cohomology.* Princeton University Press 1980, xiii+323 pp.

[MV] Morel, F.; Voevodsky, V. A^1*-homotopy theory of schemes.* Inst. Hautes Études Sci. Publ. Math. No. 90, pp. 45–143 (2001).

[Pa] Panin, I. *Push-forwards in Oriented Cohomology Theories of Algebraic Varieties: II* (preprint) `www.math.uiuc.edu/K-theory/0619/` 2003.

[PY] Panin, I.; Yagunov S. *Rigidity for Orientable Functors.* Journal of Pure and Applied Algebra, vol.172, (2002) pp. 49–77.

[SGA4] *Théorie des topos et cohomologie étale des schémas*(SGA 4) dirigé par M.Artin, A.Grothendieck et J.-L.Verdier, Tome III, Lecture Notes in Mathematics 305, Springer-Verlag, 1973.

[Su1] Suslin, A. *On the K-theory of algebraically closed fields.* Invent. Math. 73 (1983), no. 2, pp. 241–245.

[Su2] Suslin, A. *Algebraic K-Theory of Fields.* Proceedings of the ICM 1986 (Berkeley), Vol.1, pp. 222-244.

[SV] Suslin, A.; Voevodsky, V. *Singular homology of abstract algebraic varieties.* Invent. Math. 123 (1996), no. 1, pp. 61–94.

[Vo] Voevodsky,V. $\mathbb{A}^1$*-homotopy theory.* Documenta Mathematica, Extra volume ICM 1998, pp. 579-604.

Serge Yagunov
Steklov Mathematical Institute (St.Petersburg)
Fontanka 27, 191023 St.Petersburg, Russia
Max-Planck-Institut für Mathematik
Vivatsgasse 7, 53111, Bonn, Germany
yagunov@pdmi.ras.ru
yagunov@mathematik.uni-bielefeld.de

Comparison of Abelian Categories Recollements

Vincent Franjou[1] and Teimuraz Pirashvili[2]

Received: June 25, 2003

Communicated by Ulf Rehmann

Abstract. We give a necessary and sufficient condition for a morphism between recollements of abelian categories to be an equivalence.

2000 Mathematics Subject Classification: 18F, 18E40, 16G, 16D90, 16E

Keywords and Phrases: recollement, abelian category, functor

1 Introduction

Recollements of abelian and triangulated categories play an important role in geometry of singular spaces [3], in representation theory [4, 12], in polynomial functors theory [8, 9, 14] and in ring theory, where recollements are known as torsion, torsion-free theories [6]. A fundamental example of recollement of abelian categories is due to MacPherson and Vilonen [10]. It first appeared as an inductive step in the construction of perverse sheaves. The main motivation for our work was to understand when a recollement can be obtained through the construction of MacPherson and Vilonen.

A *recollement situation* consists of three abelian categories $\mathcal{A}'$, $\mathcal{A}$, $\mathcal{A}''$ together with additive functors:

$$\begin{array}{ccccc} & \xleftarrow{i^*} & & \xleftarrow{j_!} & \\ \mathcal{A}' & \xrightarrow{i_*} & \mathcal{A} & \xrightarrow{j^*} & \mathcal{A}'' \\ & \xleftarrow{i^!} & & \xleftarrow{j_*} & \end{array}$$

which satisfy the following conditions:

[1]membre du laboratoire Jean-Leray, UMR 6629 UN/CNRS

[2]supported by the grants INTAS-99-00817 and RTN-Network “K-theory, linear algebraic groups and related structures” HPRN-CT-2002- 00287

i. $j_!$ is left adjoint to j^* and j^* is left adjoint to j_*

ii. the unit $Id_{\mathcal{A}''} \to j^*j_!$ and the counit $j^*j_* \to Id_{\mathcal{A}''}$ are isomorphisms

iii. i^* is left adjoint to i_* and i_* is left adjoint to $i^!$

iv. the unit $Id_{\mathcal{A}'} \to i^!i_*$ and the counit $i^*i_* \to Id_{\mathcal{A}'}$ are isomorphisms

v. i_* is an embedding onto the full subcategory of $\mathcal{A}$ with objects A such that $j^*A = 0$.

In this case one says that $\mathcal{A}$ is a *recollement of* $\mathcal{A}''$ *and* $\mathcal{A}'$. *These notations will be kept throughout the paper.* Thus in any recollement situation, the category $i_*\mathcal{A}'$ is a localizing and colocalizing subcategory of $\mathcal{A}$ in the sense of [5], and the category $\mathcal{A}''$ is equivalent to the quotient category of $\mathcal{A}$ by $i_*\mathcal{A}'$.

If $\mathcal{B}$ is also a recollement of $\mathcal{A}''$ and $\mathcal{A}'$, then a *comparison functor* $\mathcal{A} \to \mathcal{B}$ is an exact functor which commutes with all the structural functors $i^*, i_*, i^!, j_!, j^*, j_*$. According to [12, Theorem 2.5], a comparison functor between recollements of triangulated categories is an equivalence of categories. Our example in Section 2.2 shows that this is not necessarily the case for recollements of abelian categories.
Our main result, Theorem 7.2, characterizes which comparisons of recollements are equivalences of categories. As an application, we give a homological criterion deciding when a recollement can be obtained through the construction of MacPherson and Vilonen.
Theorem. *A recollement situation of categories with enough projectives is isomorphic to a MacPherson-Vilonen construction if and only if the following two conditions hold.*

i. There exists an exact functor $r\colon \mathcal{A} \to \mathcal{A}'$ *such that* $r \circ i_* = Id_{\mathcal{A}'}$ *.*

ii. For any projective object V *of the category* $\mathcal{A}'$, $(\mathrm{L}_2 i^*)(i_*V) = 0$.

2 Examples

Our examples are related to polynomial functors. The relevance of this formalism to polynomial functors was stressed by N. Kuhn [8].
We let $\mathcal{A}'$ be the category of finite vector spaces over the field with two elements $\mathbb{F}_2$, and we let $\mathcal{A}''$ be the category of finite vector spaces over $\mathbb{F}_2$ with involution, or finite representations of Σ_2 over $\mathbb{F}_2$.

2.1

In the first example, the category $\mathcal{A}$ is a category of diagrams of finite vector spaces over $\mathbb{F}_2$:

$$(V_1, H, V_2, P): \; V_1 \rightleftarrows V_2 \; ,$$

where H: $V_1 \to V_2$ and P: $V_2 \to V_1$ are linear maps which satisfy: $PHP = 0$ and $HPH = 0$. The category $\mathcal{A}$ is equivalent to the category of quadratic functors from finitely generated free abelian groups to vector spaces over $\mathbb{F}_2$. It is a recollement for the following functors:

$$i^*(V_1, H, V_2, P) = \text{Coker }(P),\ j_!(V, T) = (V_T, 1+T, V, p)$$

$$i_*(V) = (V, 0, 0, 0), \qquad j^*(V_1, H, V_2, P) = (V_2, HP-1)$$

$$i^!(V_1, H, V_2, P) = \text{Ker }(H)\ ,\ \ j_*(V, T) = (V^T, h, V, 1+T)\ ,$$

where $V^T = \text{Ker }(1-T)$, $V_T = \text{Coker }(1-T)$, h is the inclusion and p is the quotient map. Note that the functor i_* admits an obvious exact retraction r: $(V_1, H, V_2, P) \mapsto V_1$.

2.2 Comparison fails for abelian categories recollements

We now consider the full subcategory of the category $\mathcal{A}$ in Example 2.1, whose objects satisfy the relation: $PH = 0$. This category is equivalent to the category of quadratic functors from finite vector spaces to vector spaces over $\mathbb{F}_2$. The same formulae define a recollement as well. As a result, the inclusion of categories is a comparison functor. It is not, however, an equivalence of categories.

3 The construction of MacPherson and Vilonen [10]

3.1

Let $\mathcal{A}'$ and $\mathcal{A}''$ be abelian categories. Let F: $\mathcal{A}'' \to \mathcal{A}'$ be a right exact functor, let G: $\mathcal{A}'' \to \mathcal{A}'$ be a left exact functor and let ξ: $F \to G$ be a natural transformation. Define the category $\mathcal{A}(\xi)$ as follows. The objects of $\mathcal{A}(\xi)$ are tuples (X, V, α, β), where X is in $\mathcal{A}''$, V is in $\mathcal{A}'$, $\alpha : F(X) \to V$ and $\beta : V \to G(X)$ are morphisms in $\mathcal{A}'$ such that the following diagram commutes:

$$\begin{tikzcd} F(X) \arrow[rr, "\xi_X"] \arrow[dr, "\alpha"'] & & G(X) \\ & V \arrow[ur, "\beta"'] & \end{tikzcd} .$$

A morphism from (X, V, α, β) to $(X', V', \alpha', \beta')$ is a pair (f, φ), where $f : X \to X'$ is a morphism in $\mathcal{A}''$ and φ: $V \to V'$ is a morphism in $\mathcal{A}'$, such that the following diagram commutes:

$$\begin{tikzcd} F(X) \arrow[r, "\alpha"] \arrow[d, "F(f)"] & V \arrow[r, "\beta"] \arrow[d, "\varphi"] & G(X) \arrow[d, "G(f)"] \\ F(X') \arrow[r, "\alpha'"] & V' \arrow[r, "\beta'"] & G(X') \end{tikzcd} .$$

The category $\mathcal{A}(\xi)$ comes with functors:

$$i^*(X,V,\alpha,\beta) = \text{Coker }(\alpha)\ , \qquad j_!(X) = (X, F(X), Id_{F(X)}, \xi_X)\ ,$$

$$i_*(V) = (0,V,0,0)\ , \qquad j^*(X,V,\alpha,\beta) = X\ ,$$

$$i^!(X,V,\alpha,\beta) = \text{Ker }(\beta)\ , \qquad j_*(X) = (X, G(X), \xi_X, Id_{G(X)})\ .$$

The functor i_* has a retraction functor r:

$$r(X,V,\alpha,\beta) = V\ .$$

The category $\mathcal{A}(\xi)$ is abelian in such a way that the functors r and j^* are exact. The above data define a recollement. Note that we recover the natural transformation ξ from the retraction r and the recollement data as:

$$F = rj_! \qquad G = rj_* \qquad \xi \simeq rN\ .$$

The category $\mathcal{A}$ depends only [10, Proposition 1.2] on the class of the extension

$$0 \to i^!j_! \to F \xrightarrow{\xi} G \to i^*j_* \to 0\ ,$$

image by r of the exact sequence (4).

3.2

We now consider two particular cases of this construction, already known to Grothendieck (see [1]). Let F: $\mathcal{A}'' \to \mathcal{A}'$ be a right exact functor. Take ξ: $F \to 0$ to be the transformation into the trivial functor. The corresponding construction is denoted by $\mathcal{A}' \rtimes_F \mathcal{A}''$. Thus objects of this category are triples (V,X,α), where V and X are objects of $\mathcal{A}'$ and $\mathcal{A}''$ respectively and α is a morphism α: $F(X) \to V$ of the category $\mathcal{A}'$. Note that $i^*j_* = 0$ and $i^!j_! \cong F$. Moreover, $i^!$ and j_* are exact functors.
Similarly, let $\mathcal{B}'$ and $\mathcal{B}''$ be abelian categories and let G: $\mathcal{B}'' \to \mathcal{B}'$ be a left exact functor. We take ξ: $0 \to G$ to be the natural transformation from the trivial functor. The corresponding recollement is denoted by $\mathcal{B}' \ltimes_G \mathcal{B}''$. Objects of this category are triples $(B'', B', \beta : B' \to G(B''))$. Assuming now $\mathcal{B}' = \mathcal{A}''$, $\mathcal{B}'' = \mathcal{A}'$ and $G : \mathcal{A}' \to \mathcal{A}''$ is right adjoint to F, the category $\mathcal{A}' \rtimes_F \mathcal{A}'' = \mathcal{A}'' \ltimes_G \mathcal{A}'$ fits into two different recollement situations.

4 General properties of recollements

Most of the properties in this section can probably be found in [3] or other references. We list them for convenience. Note however that, when they are not a consequence of [5], they are usually stated and proved in the context of triangulated categories. We consistently provide statements (and a few proofs) in the context of abelian categories and derived functors.

4.1 First properties

We remark as usual that taking opposite categories results in the exchange of $j_!$ and i^* with j_* and $i^!$ respectively. This is referred to as duality. For instance, the relation $j^*i_* = 0$ - a consequence of (v) - yields the dual relation $i^!j_* = 0$.

Proposition 4.1 *In any recollement situation:*

$$i^*j_! = 0\ ,\quad i^!j_* = 0.$$

Proposition 4.2 *The units and counits of adjonction give rise to exact sequences of natural transformations:*

$$j_!j^* \xrightarrow{\epsilon} Id_{\mathcal{A}} \to i_*i^* \to 0 \tag{1}$$

$$0 \to i_*i^! \to Id_{\mathcal{A}} \xrightarrow{\eta} j_*j^*\ . \tag{2}$$

We now recall the definition of the norm N: $j_! \to j_*$. For any X, Y in $\mathcal{A}''$, there are natural isomorphisms:

$$\mathrm{Hom}_{\mathcal{A}}(j_!X, j_*Y) \cong \mathrm{Hom}_{\mathcal{A}''}(X, j^*j_*Y) \cong \mathrm{Hom}_{\mathcal{A}''}(X, Y).$$

For $Y = X$, let N_X: $j_!X \to j_*X$ be the map corresponding to the identity of X. It is a natural transformation [3, 1.4.6.2]. The norm N is thus defined so that: $Nj^* = \eta \circ \epsilon$. Hence:

$$N \cong N(j^*j_*) = (Nj^*)j_* \cong (\eta\circ\epsilon)j_* = \eta j_*\circ\epsilon j_* \cong \epsilon j_* \text{ and, dually } N \cong \eta j_!\ . \tag{3}$$

The image of the norm is a functor

$$j_{!*} := \mathrm{Im}\ (N : j_! \to j_*) :\ \mathcal{A}'' \to \mathcal{A}\ .$$

Proposition 4.3 *In any recollement situation:* $i^*j_{!*} = 0\ ,\ \ i^!j_{!*} = 0\ .$

Proof. Use Proposition 4.1 and apply i^* to the epi $j_! \to j_{!*}$. □

Proposition 4.4 *In any recollement situation, there is a short exact sequence of natural transformations*

$$0 \to i_*i^!j_! \to j_! \xrightarrow{N} j_* \to i_*i^*j_* \to 0\ . \tag{4}$$

Proof. Precompose the exact sequence (1) with j_*. Precomposition is exact, hence one gets the following exact sequence:

$$j_! \to j_* \to i_*i^*j_* \to 0\ ,$$

where the left arrow is the norm N according to (3). Dually, there is an exact sequence:

$$0 \to i_*i^!j_! \to j_! \xrightarrow{N} j_*\ .$$

Splicing the two sequences together gives the result. □

Applying the snake lemma, one gets the following strong restriction on the functors $i^!j_!$ and i^*j_* of a recollement situation.

Corollary 4.5 *For any short exact sequence in $\mathcal{A}''$:*

$$0 \to X \to Y \to Z \to 0$$

there is an exact sequence in $\mathcal{A}'$:

$$i^! j_!(X) \to i^! j_!(Y) \to i^! j_!(Z) \to i^* j_*(X) \to i^* j_*(Y) \to i^* j_*(Z)$$

4.2 Homological properties

In this section we investigate the derived functors of the functors in a recollement situation. We use the following convention throughout this section: When mentioning left derived functors $\mathrm{L}-$, the category $\mathcal{A}$, and thus the categories $\mathcal{A}'$ and $\mathcal{A}''$, have enough projectives, and, similarly, when mentioning right derived functors $\mathrm{R}-$, the categories $\mathcal{A}$, $\mathcal{A}'$ and $\mathcal{A}''$ have enough injectives. Most of the proofs consist in applying long exact sequences for derived functors to Section 4.1's exact sequences.

Proposition 4.6 *For each integer $n \geq 1$:*

$$j^*(\mathrm{L}_n j_!) = 0 \quad , \quad j^*(\mathrm{R}^n j_*) = 0 \ .$$

Proposition 4.7

$$(\mathrm{L}_1 i^*) i_* = 0 \quad , \qquad (\mathrm{R}^1 i^!) i_* = 0 \tag{5}$$

$$(\mathrm{L}_1 i^*) j_! = 0 \quad , \qquad (\mathrm{R}^1 i^!) j_* = 0 \tag{6}$$

$$(\mathrm{L}_1 i^*) j_{!*} = i^! j_! \quad , \qquad (\mathrm{R}^1 i^!) j_{!*} = i^* j_* \tag{7}$$

Proposition 4.8 *There is a natural exact sequence:*

$$\begin{aligned} 0 \to \mathrm{Ext}^1_{\mathcal{A}'}(i^* A, V) \to \mathrm{Ext}^1_{\mathcal{A}}(A, i_* V) \xrightarrow{\eta} \mathrm{Hom}_{\mathcal{A}'}((\mathrm{L}_1 i^*) A, V) \to \\ \to \mathrm{Ext}^2_{\mathcal{A}'}(i^* A, V) \to \mathrm{Ext}^2_{\mathcal{A}}(A, i_* V) \ . \end{aligned}$$

Proof. This follows from the spectral sequence for the derived functors of the composite functors:

$$E^2_{pq} = \mathrm{Ext}^p_{\mathcal{A}'}(\mathrm{L}_q i^*(A), V) \Longrightarrow \mathrm{Ext}^{p+q}_{\mathcal{A}}(A, i_* V) \ . \tag{8}$$

□

Proposition 4.9 *Let A be an object in* Ker i^*. *The counit ϵ_A: $j_! j^* A \to A$ is epi and its kernel is in $i_* \mathcal{A}'$. Indeed, if $\mathcal{A}$ has enough projectives, there is a short exact sequence:*

$$0 \to i_*(\mathrm{L}_1 i^*) A \to j_! j^* A \xrightarrow{\epsilon_A} A \to 0 \ . \tag{9}$$

We prove the dual statement:

Proposition 4.10 *Let A be an object in* $\mathrm{Ker}\ i^!$. *The unit* η_A: $A \to j_*j^*A$ *is mono and its cokernel is in* $i_*\mathcal{A}'$. *Indeed, if* $\mathcal{A}$ *has enough injectives, there is a short exact sequence:*

$$0 \to A \xrightarrow{\eta_A} j_*j^*A \to i_*(\mathrm{R}^1 i^!)A \to 0\ . \tag{10}$$

Proof. When $i^!A = 0$, the exact sequence (2) simplifies to a short exact sequence:

$$0 \to A \xrightarrow{\eta_A} j_*j^*A \to \mathrm{Coker}\ \eta_A \to 0\ . \tag{11}$$

First applying the exact functor j^*, and using that $j^*\eta$ is an iso, we see that $j^*(\mathrm{Coker}\ \eta_A) = 0$. Thus $\mathrm{Coker}\ \eta_A$ is in $i_*\mathcal{A}'$. Suppose that $\mathcal{A}$ has enough injectives. Applying now the left exact functor $i^!$, the long exact sequence for right derived functors gives an exact sequence:

$$0 \to i^!A \to i^!j_*j^*A \to i^!\mathrm{Coker}\ \eta_A \to (\mathrm{R}^1 i^!)A \to (\mathrm{R}^1 i^!)j_*j^*A\ .$$

Proposition 4.1 and (6) give an isomorphism $i^!\mathrm{Coker}\ (\eta_A) \cong \mathrm{R}^1 i^!(A)$. □

4.3 Description of the image of j_*, $j_{!*}$, $j_!$

Since $j^*j_! \cong j^*j_* \cong j^*j_{!*} \cong Id_{\mathcal{A}''}$, the functors $j_!, j_*, j_{!*}$: $\mathcal{A}'' \to \mathcal{A}$ are full embeddings. The next result describes the essential image of each of them.

Proposition 4.11 *The functors* $j_!, j_*, j_{!*} : \mathcal{A}'' \to \mathcal{A}$ *induce the following equivalences of categories:*

$$j_{!*} : \mathcal{A}'' \to \{A \in \mathcal{A} \mid i^*(A) = 0 = i^!(A)\},$$

$$j_! : \mathcal{A}'' \to \{A \in \mathcal{A} \mid i^*(A) = 0 = \mathrm{L}_1 i^*(A)\},$$
$$j_* : \mathcal{A}'' \to \{A \in \mathcal{A} \mid i^!(A) = 0 = \mathrm{R}^1 i^!(A)\}.$$

4.4 A monomorphism on Ext-groups

Since $j^* : \mathcal{A} \to A''$ is an exact functor, it induces an homomorphism

$$\mathrm{Ext}^n_{\mathcal{A}}(A, B) \to \mathrm{Ext}^n_{\mathcal{A}''}(j^*A, j^*B),\ n \geq 0.$$

It is well-known that when A and B are simple objects, this map is injective for $n = 1$ (see for example [8, Proposition 4.12]). The following more general result holds.

Proposition 4.12 *Let* $A, B \in \mathcal{A}$ *be objects for which* $i^*A = 0$ *and* $i^!B = 0$. *Suppose* $j^*A \neq 0$ *and* $j^*B \neq 0$. *Then*

$$\mathrm{Ext}^1_{\mathcal{A}}(A, B) \to \mathrm{Ext}^1_{\mathcal{A}''}(j^*A, j^*B)$$

is a monomorphism.

5 Description of Ker i^* and Ker $i^!$

Let Ker $i^!$ be the full subcategory of objects A of $\mathcal{A}$ such that $i^!A = 0$, and let Ker i^* be the full subcategory of objects A of $\mathcal{A}$ such that $i^*A = 0$. In this section, we describe these subcategories of $\mathcal{A}$ in terms of the categories $\mathcal{A}'$, $\mathcal{A}''$, and the functors i^*j_*, $i^!j_!$ between them, through the following construction:

Definition 5.1 *Let $T\colon \mathcal{A}'' \to \mathcal{A}'$ be an additive functor between abelian categories. The category $\mathcal{M}(T)$ has objects triples (X, V, α) where X is in $\mathcal{A}''$, V is in $\mathcal{A}'$, and $\alpha\colon V \to TX$ is a monomorphism. A map from (X, V, α) to (X', V', α') is a pair of morphisms (f, φ) such that the following diagram commutes:*

$$\begin{tikzcd} V \arrow[r, "\alpha"] \arrow[d, "\varphi"] & T(X) \arrow[d, "T(f)"] \\ V' \arrow[r, "\alpha'"] & T(X') \end{tikzcd} .$$

The following theorem is inspired by [13].

Theorem 5.2 *In a recollement with enough projectives, the functor $A \mapsto (j^*A,\ i^*A,\ i^*\eta_A : i^*A \to i^*j_*j^*A)$ is an equivalence from the category* Ker $i^!$ *to the category $\mathcal{M}(i^*j_*)$.*

Proof. First, we show that the functor is well defined. Apply the functor i^* on the short exact sequence (11). There results an exact sequence:

$$\mathrm{L}_1 i^*(\mathrm{Coker}\ \eta_A) \to i^*A \to i^*j_*j^*(A) \to i^*\mathrm{Coker}\ \eta_A \to 0\ .$$

whose left term cancels by Proposition 4.10 and (5). The map $i^*\eta_A$ is thus mono.

Next, we define the quasi-inverse: $\mathcal{M}(i^*j_*) \to$ Ker $i^!$. To an object (X, V, α), it associates the kernel $A(X, V, \alpha)$ of the composite of epis:

$$j_*X \overset{\epsilon j_*}{\to} i_*i^*j_*X \to \mathrm{Coker}\ i_*\alpha\ .$$

That is, $A(X, V, \alpha)$ fits in the following map of extensions:

$$\begin{tikzcd} 0 \arrow[r] & j_{!*}X \arrow[r] & j_*X \arrow[r] & i_*i^*j_*X \arrow[r] & 0 \\ 0 \arrow[r] & j_{!*}X \arrow[r] \arrow[u, "="] & A(X,V,\alpha) \arrow[r] \arrow[u] & i_*V \arrow[r] \arrow[u, "i_*\alpha"] & 0 \end{tikzcd} .$$

To a map (f, φ), it associates the map induced by $j_*(f)$.

We leave the verifications to the reader, with the help of the isomorphism $Nj^* \cong \epsilon \circ \eta$. $\square$

The dual study of the category Ker i^* leads to the following.

Theorem 5.3 *In a recollement with enough injectives, the functor* $A \mapsto (j^*A,\ i^!\mathrm{Ker}\ \epsilon_A,\ i^!\mathrm{Ker}\ \epsilon_A \to i^! j_! j^* A)$ *is an equivalence from the category* $\mathrm{Ker}\ i^*$ *to the category* $\mathcal{M}(i^! j_!)$.

This time, the quasi-inverse fits in the following map of extensions:

$$\begin{tikzcd} 0 \arrow[r] & i_* i^! j_! X \arrow[r] \arrow[d] & j_! X \arrow[r] \arrow[d] & j_{!*} X \arrow[r] \arrow[d, "="] & 0 \\ 0 \arrow[r] & \mathrm{Coker}\ (i_*\alpha) \arrow[r] & A(X,V,\alpha) \arrow[r] & j_{!*} X \arrow[r] & 0\ . \end{tikzcd}$$

Note (Proposition 4.9) that when the recollement has enough projectives, $i^!\mathrm{Ker}\ \epsilon_A$ is nothing but $(\mathrm{L}_1 i^*)A$.

6 Recollements as linear extensions

The exact sequence (2) tells that every object A in $\mathcal{A}$ sits in a short exact sequence:

$$0 \to \mathrm{Ker}\ \eta_A \to A \overset{\eta_A}{\to} \mathrm{Im}\ \eta_A \to 0\ .$$

where $\mathrm{Ker}\ \eta_A \cong i_* i^! A$ is in $i_*\mathcal{A}'$ and $\mathrm{Im}\ \eta_A \cong A/i_* i^! A$ is in $\mathrm{Ker}\ i^!$. We denote by $\mathcal{G}$ the category encoding these data from the recollement situation. That is, objects of the category $\mathcal{G}$ are triples (A, U, e) of an object A in $\mathrm{Ker}\ i^!$, an object U in $\mathcal{A}'$ and an extension class e in the group $\mathrm{Ext}^1_{\mathcal{A}}(A, i_*U)$. A map from (A, U, e) to (A', U', e') is a pair of morphism $(\alpha : A \to A', \beta : U \to U')$ such that: $\alpha^* e' = (i_*\beta)_* e$ in the group $\mathrm{Ext}^1_{\mathcal{A}}(A', i_*U)$. It comes with a functor:

$$\mathcal{A} \to \mathcal{G} \qquad B \mapsto (\mathrm{Im}\ \eta_B,\ i^! B,\ [0 \to \mathrm{Ker}\ \eta_B \to B \xrightarrow{\eta} \mathrm{Im}\ \eta_B \to 0])\ .$$

Because of the Yoneda correspondence between extensions and elements in Ext^1, this functor induces an equivalence of categories to $\mathcal{G}$ from the following category $\mathcal{B}$. The objects of $\mathcal{B}$ are those of $\mathcal{A}$, and a map in $\mathrm{Hom}_{\mathcal{B}}(B, B')$ is a class of maps in $\mathrm{Hom}_{\mathcal{A}}(B, B')$ inducing the same map in $\mathcal{G}$.

We claim that $\mathcal{A} \to \mathcal{B}$ defines a linear extension of categories in the sense of Baues and Wirsching. For completeness, we now recall what we need from this theory (however, the following defining properties might be better understood by just looking at our example).

Definition 6.1 [2, IV.3] Let $\mathcal{B}$ be a category and let $D : \mathcal{B}^{op} \times \mathcal{B} \to \mathcal{A}b$ be a bifunctor with abelian groups values. We say that

$$0 \longrightarrow D \longrightarrow \mathcal{C} \xrightarrow{p} \mathcal{B} \longrightarrow 0 \tag{12}$$

is a linear extension of the category $\mathcal{B}$ by D if the following conditions hold:

i. $\mathcal{C}$ is a category and p is a functor. Moreover $\mathcal{C}$ and $\mathcal{B}$ have the same objects, p is the identity on objects and p is surjective on morphisms.

ii. For any objects c and d in $\mathcal{B}$, the abelian group $D(c,d)$ acts on the set $\mathrm{Hom}_{\mathcal{C}}(c,d)$. Moreover $p(f_0) = p(g_0)$ if and only if there is unique α in $D(c,d)$ such that: $g_0 = f_0 + \alpha$. Here for each $f_0 : c \to d$ in $\mathcal{C}$ and $\alpha \in D(c,d)$ we write $f_0 + \alpha$ for the action of α on f_0.

iii. The action satisfies the linear distributivity law: for two composable maps f_0 and g_0 in $\mathcal{C}$

$$(f_0 + \alpha)(g_0 + \beta) = f_0 g_0 + f_* \beta + g^* \alpha \ ,$$

where $f = p(f_0)$ and $g = p(g_0)$.

A morphism between two linear extensions

$$\begin{tikzcd} 0 \arrow[r] & D \arrow[r] \arrow[d, "\phi_1"] & \mathcal{C} \arrow[r, "p"] \arrow[d, "\phi_0"] & \mathcal{B} \arrow[r] \arrow[d, "\phi"] & 0 \\ 0 \arrow[r] & D' \arrow[r] & \mathcal{C}' \arrow[r, "p'"] & \mathcal{B}' \arrow[r] & 0 \end{tikzcd}$$

consists of functors ϕ and ϕ_0, such that $\phi p = p' \phi_0$, together with a natural transformation $\phi_1 : D \to D' \circ (\phi^{op} \times \phi)$ such that:

$$\phi_0(f_0 + \alpha) = \phi_0(f_0) + \phi_1(\alpha)$$

for all $f_0 : c \to d$ in $\mathcal{C}$ and α in $D(c,d)$.

The following properties of linear extensions are relevant to our problem.

i. If $\mathcal{B}$ is a small category, there is [2, IV.6] a canonical bijection

$$M(\mathcal{B}, D) \cong H^2(\mathcal{B}, D).$$

from the set of equivalence classes of linear extensions of $\mathcal{B}$ by D and the second cohomology group $H^2(\mathcal{B}, D)$ of $\mathcal{B}$ with coefficients in D.

ii. The functor p reflects isomorphisms and yields a bijection on the sets of isomorphism classes $\mathrm{Iso}(\mathcal{C}) \cong \mathrm{Iso}(\mathcal{B})$.

iii. Let (ϕ_1, ϕ_0, ϕ) be a morphism of linear extensions. Suppose that $\phi_1(c,d)$ is an isomorphism for any c and d in $\mathcal{B}$. Then ϕ is an equivalence of categories if and only if ϕ_0 is an equivalence of categories.

iv. If $\mathcal{B}$ is an additive category and D is a biadditive bifunctor, then the category $\mathcal{C}$ is additive [7, Proposition 3.4].

We now describe recollements in terms of linear extensions.

Proposition 6.2 *Let D be the bifunctor defined on $\mathcal{B}$ by:*

$$D(B, B') := \mathrm{Hom}_{\mathcal{A}}(B/i_* i^! B, i_* i^! B') \ .$$

The category $\mathcal{A}$ is a linear extension of $\mathcal{B}$ by D.

Proof. It reduces to the following. Two maps of extensions:

$$\begin{tikzcd} 0 \arrow[r] & U \arrow[r] \arrow[d] & A \arrow[r] \arrow[d, shift right, "f"'] \arrow[d, shift left, "g"] & X \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & U' \arrow[r] & A' \arrow[r] & X' \arrow[r] & 0 \end{tikzcd}$$

agree on the side vertical arrows if and only if their difference $f-g$ factors through a map in the group $\mathrm{Hom}(X,U')$. □

The results of Section 5 shows that the categories $\mathcal{A}'$, $\mathcal{A}''$ and the functors i^*j_*, $i^!j_!$ of the recollement situation determine the category Ker $i^!$. We now show that it does determine the bifunctor D as well. For an object B in $\mathcal{B}$, let $((X,V,\alpha),U)$ be its image under the composite:

$$\mathcal{B} \simeq \mathcal{G} \to \text{Ker } i^! \times \mathcal{A}' \simeq \mathcal{M}(i^*j_*) \times \mathcal{A}' \ .$$

That is: $X = j^*A$, $V = i^*A$, for $A = B/i_*i^!B$, $U = i^!B$. Then:

$$D(B,B') := \mathrm{Hom}_{\mathcal{A}}(A, i_*U') = \mathrm{Hom}_{\mathcal{A}'}(i^*A, U') = \mathrm{Hom}_{\mathcal{A}'}(V, U') \ . \tag{13}$$

7 A comparison theorem

We have seen in Section 2.2 an example of a comparison functor which is not an equivalence of categories. However, a comparison functor E indeed yields an equivalence from Ker $(i^* : \mathcal{A}_1 \to \mathcal{A}')$ to Ker $(i^* : \mathcal{A}_2 \to \mathcal{A}')$, and similarly for Ker $i^!$. If E is an equivalence of categories, then clearly E commutes with the derived functors $\mathrm{R}^\bullet i^!$ and $\mathrm{L}_\bullet i^*$. This observation leads to the following definition.

Definition 7.1 *Let $(\mathcal{A}', \mathcal{A}_1, \mathcal{A}'')$ and $(\mathcal{A}', \mathcal{A}_2, \mathcal{A}'')$ be two recollement situations. Assume that the categories $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}', \mathcal{A}''$ have enough projective objects. A comparison functor E: $\mathcal{A}_1 \to \mathcal{A}_2$ is left admissible if the following diagram commutes.*

$$\begin{tikzcd} \mathcal{A}' \arrow[d, "="'] & \text{Ker } i^! \arrow[l, "\mathrm{L}_1 i^*"'] \arrow[d, "E"] \\ \mathcal{A}' & \text{Ker } i^! \arrow[l, "\mathrm{L}_1 i^*"] \end{tikzcd}$$

A right admissible comparison functor is defined similarly by using the functors $\mathrm{R}^1 i^!$ *and the categories* Ker i^*.

Theorem 7.2 *Let E be a comparison functor between categories with enough injectives and projectives. The following conditions are equivalent*

i. E is right admissible

ii. E is left admissible

iii. E is an equivalence of categories.

Proof. It is clear that iii) implies both conditions i) and ii). We only show that ii) implies iii). A dual argument shows that i) implies iii). By Section 6, the functor E yields a commutative diagram of linear extensions

$$\begin{array}{ccccccccc} 0 & \longrightarrow & D_1 & \longrightarrow & \mathcal{A}_1 & \longrightarrow & \mathcal{B}_1 & \longrightarrow & 0 \\ & & \downarrow & & \downarrow{\scriptstyle E} & & \downarrow & & \\ 0 & \longrightarrow & D_2 & \longrightarrow & \mathcal{A}_2 & \longrightarrow & \mathcal{B}_2 & \longrightarrow & 0 \end{array}$$

First we show that E yields an equivalence of categories $\mathcal{B}_1 \to \mathcal{B}_2$. By Section 6 it suffices to show that E yields an equivalence $\mathcal{G}_1 \to \mathcal{G}_2$. When there are enough projectives, E yields an equivalence on Ker $i^!$ (Theorem 5.2). The induced map

$$\mathrm{Ext}^1_{\mathcal{A}_1}(A, i_*U) \to \mathrm{Ext}^1_{\mathcal{A}_2}(E(A), i_*U)$$

is an isomorphism for U in $\mathcal{A}'$ and A in Ker $i^!$, thanks to Proposition 4.8 and the five-lemma. Once $\mathcal{B}_1$ and $\mathcal{B}_2$ are identified, we use the computation (13) to conclude that the morphism of bifunctors $D_1 \to D_2$ is an isomorphism. The rest is a consequence of the properties of linear extensions of categories. □

8 Recollement pré-héréditaire

8.1 pre-hereditary recollement

Definition 8.1 *A recollement situation with enough projectives is pre-hereditary if for any projective object V of the category $\mathcal{A}'$:*

$$(\mathrm{L}_2 i^*)(i_*V) = 0 \ .$$

Proposition 8.2 *In a pre-hereditary recollement situation:* $(\mathrm{L}_2 i^*)i_* = 0$.

Proof. By (5) the functor $(\mathrm{L}_2 i^*)i_*$ is right exact. If it vanishes on projective objects, it vanishes on all objects. □

Lemma 8.3 *In a pre-hereditary recollement situation there is an isomorphism of functors*

$$(\mathrm{L}_1 i^*)j_* \cong i^! j_!.$$

Proof. Apply the functor i^* to the short exact sequence:

$$0 \to j_{!*} \to j_* \to i_* i^* j_* \to 0 \ .$$

By (5), $\mathrm{L}_1 i^*$ vanishes on $i_* i^* j_*$, and by hypothesis $\mathrm{L}_2 i^*$ vanishes on $i_* i^* j_*$. Hence the long exact sequence for left derived functors yields an isomorphism: $(\mathrm{L}_1 i^*)j_{!*} \cong (\mathrm{L}_1 i^*)j_*$. The result follows by (7). □

Theorem 8.4 *Let $(\mathcal{A}', \mathcal{A}, \mathcal{A}'')$ and $(\mathcal{A}', \mathcal{B}, \mathcal{A}'')$ be two pre-hereditary recollement situations and let $E : \mathcal{A} \to \mathcal{B}$ be a comparison functor. Then E is admissible and hence is an equivalence of categories.*

Proof. We have to prove that $\mathrm{L}_1 i^*$ has the same value on A and EA, provided that $i^! A = 0$. For such an A, there is a short exact sequence (11). Applying the functor i^* results in an exact sequence:

$$\mathrm{L}_2 i^*(\mathrm{Coker}\ \eta_A) \to \mathrm{L}_1 i^*(A) \to \mathrm{L}_1 i^*(j_* j^* A) \to \mathrm{L}_1 i^*(\mathrm{Coker}\ \eta_A)$$

whose right term cancels by Proposition 4.10 and (5), and whose left term cancels by Proposition 8.2. This gives an isomorphism: $\mathrm{L}_1 i^*(A) \cong (\mathrm{L}_1 i^*) j_* j^*(A)$. Lemma 8.3 finishes the proof. $\square$

8.2 MacPherson-Vilonen recollements

The following proposition is a formalized version of the construction of projectives in [11, Proposition 2.5].

Proposition 8.5 *Let $\mathcal{A}(F \xrightarrow{\xi} G)$ be a MacPherson-Vilonen recollement. Assume further that the left exact functor G has a left adjoint G^*. Then the exact functor r has a left adjoint r^* defined by:*

$$r^* V = (G^* V, FG^* V \oplus V, (1,0), \xi_{G^* V} \oplus \eta_V)$$

where in this formula η denotes the unit of adjonction: $id_{\mathcal{A}'} \to GG^*$. In particular, there is a short exact sequence:

$$0 \to j_! G^* \to r^* \to i_* \to 0\ . \tag{14}$$

Proof. Necessarily, $j^* r^* = (r j_!)^* = G^*$. Then check. $\square$

Proposition 8.6 *Every MacPherson-Vilonen recollement with enough projectives is pre-hereditary .*

Proof. Apply the functor i^* to the short exact sequence (14). Part of the resulting long exact sequence is an exact sequence:

$$(\mathrm{L}_2 i^*) r^* \to (\mathrm{L}_2 i^*) i_* \to (\mathrm{L}_1 i^*) j_! G^*\ ,$$

whose right term cancels by (6). To conclude, if P is a projective in $\mathcal{A}'$, then $r^* P$ is a projective in $\mathcal{A}$, because r^* is left adjoint to an exact functor. $\square$

This leads to the following characterization of MacPherson-Vilonen recollements. A special case appeared in [15, Proposition 2.6]

Theorem 8.7 *A recollement situation of categories with enough projectives is isomorphic to a MacPherson-Vilonen construction if and only if the recollement is pre-hereditary and there exists an exact functor* $r\colon \mathcal{A} \to \mathcal{A}'$ *such that* $r \circ i_* = Id_{\mathcal{A}'}$.

Proof. Consider a recollement with such an exact retraction functor r. The natural transformation $N\colon j_! \to j_*$ yields a transformation rN from the right exact functor $rj_!$ to the left exact functor rj_*. Thus we can form the MacPherson-Vilonen construction $\mathcal{A}(rj_! \stackrel{rN}{\to} rj_*)$. We define a functor $E\colon \mathcal{A} \to \mathcal{A}(rj_! \stackrel{rN}{\to} rj_*)$ by:

$$E(A) = (j^*(A), r(A), r(\epsilon_A), r(\eta_A)).$$

One checks with Section 3 and (3) that E is a comparison functor. By Proposition 8.6, $\mathcal{A}(rN)$ is pre-hereditary. If $\mathcal{A}$ is also pre-hereditary, Theorem 8.4 applies. □

Remark. Similarly one can define *pre-cohereditary* recollements by the condition $\mathrm{R}^2 i^!(i_* V) = 0$ for any injective V in $\mathcal{A}'$. We leave to the reader to dualize the above results.

8.3 The case when $i^* j_* = 0$ or $i^! j_! = 0$

In this section, we characterize the recollements $\mathcal{A} = \mathcal{A}' \rtimes_F \mathcal{A}''$ of Section 3.2.

Proposition 8.8 *For a recollement with enough projectives, the following are equivalent:*

i. The functor i^* *is exact.*

ii. $i^! j_! = 0$.

Dually, for a recollement with enough injectives, the following are equivalent:

i. The functor $i^!$ *is exact.*

ii. $i^* j_* = 0$.

Proof. We prove the second assertion. Assume that $i^!$ is exact. Applying $i^!$ to the epimorphism $j_* \to i_* i^* j_*$ gets an epimorphism $0 = i^! j_* \to i^! i_* i^* j_* \cong i^* j_*$. Assume conversely that $i^* j_* = 0$ and suppose that the recollement has enough injectives. We first prove that $\mathrm{R}^1 i^!(A) = 0$ when $i^! A = 0$. By Proposition 4.10, if $i^! A = 0$, there is an epimorphism $j_* j^* A \to \operatorname{Coker} \eta_A \cong i_*(\mathrm{R}^1 i^!)(A)$. Applying the right exact functor i^*, we get an epimorphism $i^* j_* j^*(A) \to (\mathrm{R}^1 i^!)(A)$. Next, we apply $i^!$ to the short exact sequence (2). It yields an exact sequence:

$$0 \to i^! \stackrel{\simeq}{\to} i^! \to i^! \mathrm{Im}\ \eta \to (\mathrm{R}^1 i^!) i_* i^! \to \mathrm{R}^1 i^! \to (\mathrm{R}^1 i^!) \mathrm{Im}\ \eta\ .$$

By (5), $(\mathrm{R}^1 i^!) i_* i^! = 0$, so that: $i^! \mathrm{Im}\ \eta = 0$. It results that $(\mathrm{R}^1 i^!) \mathrm{Im}\ \eta = 0$ as well, and finally that $R^1 i^! = 0$. □

As an application we recover [1, Proposition 2.4].

Proposition 8.9 *Every recollement situation with enough projectives, such that: $i^! j_! = 0$, is equivalent to $\mathcal{A}' \ltimes_{i^* j_*} \mathcal{A}''$. Dually, every recollement situation with enough injectives, such that: $i^* j_* = 0$, is equivalent to $\mathcal{A}' \rtimes_{i^! j_!} \mathcal{A}''$.*

Proof. When the recollement has enough projectives, Theorem 8.7 applies for $r = i^*$. □

Corollary 8.10 *Let $\mathcal{A}', \mathcal{A}, \mathcal{A}''$ be a recollement situation with enough projective or enough injectives. If the norm $N : j_! \to j_*$ is an isomorphism, then $\mathcal{A} \cong \mathcal{A}' \times \mathcal{A}''$.*

Proof. By Proposition 4.4: $i^* j_* = i^! j_! = 0$. Then we apply Proposition 8.9. □

Acknowledgements. The second author would like to thank University of Nantes for hospitality and support.

References

[1] M. Artin. Grothendieck Topologies. Harvard University, 1962.

[2] H.-J. Baues and G. Wirsching. *Cohomology of small categories.* J. Pure and Appl. Algebra 38 (1985), 187–211.

[3] A.A. Beilinson, J. Bernstein, P. Deligne. Faisceaux pervers. Analysis and topology on singular spaces, I (Luminy, 1981), 5–171, Astérisque 100, Soc. Math. France, Paris, 1982.

[4] E. Cline, B. Parshall, L. Scott. *Finite-dimensional algebras and highest weight categories.* J. Reine Angew. Math. 391 (1988), 85–99.

[5] P. Gabriel. *Des catégories abéliennes.* Bull. S. M. F. 90 (1962), 323-448.

[6] J.P. Jans. *Some aspects of torsion.* Pacific J. Math. 15 (1965), 1249-1259.

[7] M. Jibladze, T. Pirashvili. *Cohomology of algebraic theories.* J. of Algebra 137 (1991), 253-296.

[8] N. J. Kuhn. *Generic representations of the finite general linear groups and the Steenrod algebra. II.* K-Theory 8 (1994), no. 4, 395–428.

[9] N. J. Kuhn. *A stratification of generic representation theory and generalized Schur algebras.* K-Theory 26 (2002), no. 1, 15–49.

[10] R. MacPherson, K. Vilonen. *Elementary construction of perverse sheaves.* Invent. Math. 84 (1986),403-436.

[11] R. Mirollo, K. Vilonen. *Bernstein-Gelfand-Gelfand reciprocity on perverse sheaves.* Ann. scient. Éc. Norm. Sup. 20 (1987), 311–324.

[12] B. Parshall, L. Scott. Derived Categories, Quasi-Hereditary Algebras and Algebraic Groups. Carlton University Mathematical notes 3 (1988), 1-104.

[13] T. Pirashvili. *On quadtratic functors.* Bull. Ac. Sc. Georgian SSR 129 (1988), 485–488.

[14] T. Pirashvili. *Polynomial functors.* Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR 91 (1988), 55–66.

[15] K. Vilonen. *Perverse sheaves and finite dimensional algebras.* Transactions A.M.S. 341 (1994), 665–676

Vincent Franjou
Université de Nantes
dép. de mathématiques
2, rue de la Houssinière
BP 92208 44322 Nantes cedex 3
France

Teimuraz Pirashvili
A. M. Razmadze Mathematical
Institute
Aleksidze str.1
Tbilisi 380093
Republic of Georgia

On the Scattering Theory of the Laplacian with a Periodic Boundary Condition. II. Additional Channels of Scattering

Rupert L. Frank and Roman G. Shterenberg

Received: February 12, 2004
Revised: March 4, 2004

Communicated by Heinz Siedentop

Abstract. We study spectral and scattering properties of the Laplacian $H^{(\sigma)} = -\Delta$ in $L_2(\mathbb{R}_+^2)$ corresponding to the boundary condition $\frac{\partial u}{\partial \nu} + \sigma u = 0$ for a wide class of periodic functions σ. For non-negative σ we prove that $H^{(\sigma)}$ is unitarily equivalent to the Neumann Laplacian $H^{(0)}$. In general, there appear additional channels of scattering which are analyzed in detail.

2000 Mathematics Subject Classification: Primary 35J10; Secondary 35J25, 35P05, 35P25.
Keywords and Phrases: Scattering theory, periodic operator, Schrödinger operator, singular potential.

Introduction

0.1 Setting of the problem

The present paper is a continuation of [Fr], but can be read independently. It studies the Laplacian

$$H^{(\sigma)}u = -\Delta u \qquad \text{on } \mathbb{R}_+^2 \tag{0.1}$$

together with a boundary condition of the third type

$$\frac{\partial u}{\partial \nu} + \sigma u = 0 \qquad \text{on } \mathbb{R} \times \{0\}, \tag{0.2}$$

where ν denotes the exterior unit normal and where the function $\sigma : \mathbb{R} \to \mathbb{R}$ is assumed to be 2π-periodic. Moreover, let

$$\sigma \in L_{q,loc}(\mathbb{R}) \qquad \text{for some } q > 1.$$

Under this condition $H^{(\sigma)}$ can be defined as a self-adjoint operator in $L_2(\mathbb{R}^2_+)$ by means of the lower semibounded and closed quadratic form

$$\int_{\mathbb{R}^2_+} |\nabla u(x)|^2\,dx + \int_{\mathbb{R}} \sigma(x_1)|u(x_1,0)|^2\,dx_1, \qquad u \in H^1(\mathbb{R}^2_+).$$

We analyze the spectrum of $H^{(\sigma)}$ and develop a scattering theory viewing $H^{(\sigma)}$ as a (rather singular) perturbation of $H^{(0)}$, the Neumann Laplacian on $\mathbb{R}^2_+$. (For the abstract mathematical scattering theory see, e.g., [Ya1].)
By means of the Bloch-Floquet theory we represent $H^{(\sigma)}$ as a direct integral

$$\int_{-1/2}^{1/2} \oplus H^{(\sigma)}(k)\,dk \tag{0.3}$$

with fiber operators $H^{(\sigma)}(k)$ acting in $L_2(\Pi)$ where $\Pi := (-\pi,\pi)\times\mathbb{R}_+$ is the halfstrip. Due to the relation (0.3) the investigation of the operator $H^{(\sigma)}$ reduces to the study of the operators $H^{(\sigma)}(k)$.

0.2 The main results

It was shown in [Fr] that the wave operators

$$W_\pm^{(\sigma)}(k) := W_\pm(H^{(\sigma)}(k), H^{(0)}(k))$$

on the halfstrip exist and are complete. This immediately implies the existence of the wave operators

$$W_\pm^{(\sigma)} := W_\pm(H^{(\sigma)}, H^{(0)})$$

on the halfplane and the coincidence of the ranges

$$\mathcal{R}(W_+^{(\sigma)}) = \mathcal{R}(W_-^{(\sigma)}).$$

(Of course, the existence of the wave operators can also be obtained by a modification of the Cook method, see Section 17 in [Ya2].) Moreover, it was shown in [Fr] that the singular continuous spectrum of the operators $H^{(\sigma)}(k)$ is empty.
In the present paper we will study the point spectrum of the operators $H^{(\sigma)}(k)$. In general, there will be (discrete or embedded) eigenvalues which may produce bands in the spectrum of the operator $H^{(\sigma)}$ on the halfplane. In this case, the wave operators are not complete and there appear additional channels of scattering. For the additional bands in the spectrum we give some quantitative estimates and we construct an example where a gap in the spectrum appears. Moreover, we prove that the spectrum of the operator $H^{(\sigma)}$ is purely absolutely continuous.
Under the additional assumption

$$\sigma(x_1) \geq 0, \qquad \text{a.e. } x_1 \in \mathbb{R}, \tag{0.4}$$

we prove that the operators $H^{(\sigma)}(k)$ have no eigenvalues. This implies that the wave operators $W_\pm^{(\sigma)}$ are unitary and provide a unitary equivalence between the operators $H^{(\sigma)}$ and $H^{(0)}$.

0.3 Additional channels of scattering

Additional channels of scattering were already discovered in a number of other problems that exhibit periodicity with respect to some but not all space directions. Without aiming at completeness we mention the papers [DaSi], [Sa] concerning the scattering theory of problems of this type, [GrHoMe], [Ka] concerning Schrödinger operators with periodic point interactions and [BeBrPa] concerning the case of discrete Schrödinger operators.
In the present paper, using the specific properties of the operator under consideration we are able not only to show the appearance of additional channels of scattering but also to develop a more detailed analysis of these channels. In particular, we give some sufficient conditions for existence and non-existence of additional channels and prove that the spectrum of the operator is purely absolutely continuous.
The problem of absolute continuity in a case with partial periodicity is also investigated in [FiKl], where the Schrödinger operator with an electric potential is considered.

0.4 Outline of the paper

Let us explain the structure of this paper. In Section 1 we recall the precise definition of the operators $H^{(\sigma)}$ and $H^{(\sigma)}(k)$ in terms of quadratic forms and the direct integral decomposition. In Subsection 1.2 we state the main result in the case of non-negative σ (Theorem 1.1) and the main result about absolute continuity (Theorem 1.2).
In Section 2 we transform the eigenvalue problem for $H^{(\sigma)}(k)$ and $\lambda \in \mathbb{R}$ in the spirit of the Birman-Schwinger principle to the problem whether 0 is an eigenvalue of a certain "discrete pseudo-differential operator" of order one in $L_2(\mathbb{T})$. In this way we reduce the problem of (possibly embedded) eigenvalues to the study of operators with compact resolvent. In Section 3 we prove the absence of eigenvalues of $H^{(\sigma)}(k)$ under the condition (0.4), which implies Theorem 1.1. The general case is treated in Section 4 and the proof of Theorem 1.2 is given in Subsection 4.3. We supplement this in Section 5 with a more detailed analysis in the case when σ is a trigonometric polynomial. Finally, in Section 6 we describe and discuss the additional channels of scattering that appear in the general case. In Subsection 6.2 we construct an example of an open gap.

0.5 Acknowledgements

The authors are deeply grateful to Prof. M. Sh. Birman for the setting of the problem, useful discussions and constant attention to the work. The authors also thank Prof. T. A. Suslina for useful consultations and Prof. N. Filonov and Prof. F. Klopp for making us [FiKl] accessible before its publication. The authors thank Prof. A. Laptev, the KTH Stockholm, the Mittag-Leffler Institute as well as Prof. H. Siedentop and the LMU Munich for hospitality. The first author acknowledges gratefully the partial financial support of the German Merit Foundation and of the European Union through the IHP network HPRN-CT-2002-00277. The second author acknowledges gratefully the partial financial support of RFBR (grant no. 02-01-00798) and of the German Academic Exchange Service through the IQN network.

1 Setting of the problem. The main result

1.1 Notation

We introduce the *halfplane*

$$\mathbb{R}^2_+ := \{x = (x_1, x_2) \in \mathbb{R}^2 : \ x_2 > 0\} = \mathbb{R} \times \mathbb{R}_+,$$

and the *halfstrip*

$$\Pi := \{x = (x_1, x_2) \in \mathbb{R}^2 : \ -\pi < x_1 < \pi, \ x_2 > 0\} = (-\pi, \pi) \times \mathbb{R}_+,$$

where $\mathbb{R}_+ := (0, +\infty)$. Moreover, we need the lattice $2\pi\mathbb{Z}$. Unless stated otherwise, periodicity conditions are understood with repect to this lattice. We think of the corresponding torus $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ as the interval $[-\pi, \pi]$ with endpoints identified.

We use the notation $D = (D_1, D_2) = -i\nabla$ in $\mathbb{R}^2$.

For a measurable set $\Lambda \subset \mathbb{R}$ we denote by $\operatorname{meas}\Lambda$ its Lebesgue measure.

For an open set $\Omega \subset \mathbb{R}^d$, $d = 1, 2$, the index in the notation of the norm $\|.\|_{L_2(\Omega)}$ is usually dropped. The space $L_2(\mathbb{T})$ may be formally identified with $L_2(-\pi, \pi)$. We denote the Fourier coefficients of a function $f \in L_2(\mathbb{T})$ by $\hat{f}_n := \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x_1) e^{-inx_1}\, dx_1$, $n \in \mathbb{Z}$.

Next, $H^s(\Omega)$ is the Sobolev space of order $s \in \mathbb{R}$ (with integrability index 2). By $H^s(\mathbb{T})$ we denote the closure of $C^\infty(\mathbb{T})$ in $H^s(-\pi, \pi)$. Here $C^\infty(\mathbb{T})$ is the space of functions in $C^\infty(-\pi, \pi)$ which can be extended 2π-periodically to functions in $C^\infty(\mathbb{R})$. The space $H^s(\mathbb{T})$ is endowed with the norm

$$\|f\|^2_{H^s(\mathbb{T})} := \sum_{n\in\mathbb{Z}} (1 + n^2)^s |\hat{f}_n|^2, \qquad f \in H^s(\mathbb{T}).$$

By $\tilde{H}^s(\Pi)$ we denote the closure of $\tilde{C}^\infty(\Pi) \cap H^s(\Pi)$ in $H^s(\Pi)$. Here $\tilde{C}^\infty(\Pi)$ is the space of functions in $C^\infty(\Pi)$ which can be extended 2π-periodically with respect to x_1 to functions in $C^\infty(\mathbb{R}^2_+)$.

Statements and formulae which contain the double index "$\pm$" are understood as two independent assertions.

1.2 The operators $H^{(\sigma)}$ on the halfplane. Main results

Before describing the main results we recall the definition of the operators $H^{(\sigma)}$ from [Fr]. Let σ be a real-valued periodic function satisfying

$$\sigma \in L_q(\mathbb{T}) \qquad \text{for some } q > 1. \tag{1.1}$$

It is easy to see (cf. [Fr]) that under this condition the quadratic form

$$\begin{aligned} &\mathcal{D}[h^{(\sigma)}] := H^1(\mathbb{R}^2_+), \\ &h^{(\sigma)}[u] := \int_{\mathbb{R}^2_+} |Du(x)|^2\,dx + \int_{\mathbb{R}} \sigma(x_1)|u(x_1,0)|^2\,dx_1 \end{aligned} \tag{1.2}$$

is lower semibounded and closed in the Hilbert space $L_2(\mathbb{R}^2_+)$, so it generates a self-adjoint operator which will be denoted by $H^{(\sigma)}$. The case $\sigma = 0$ corresponds to the *Neumann Laplacian* on the halfplane, whereas the case $\sigma \neq 0$ implements a (generalized) *boundary condition of the third type*.
The spectrum of the "unperturbed" operator $H^{(0)}$ coincides with $[0, +\infty)$ and is purely absolutely continuous of infinite multiplicity.
In [Fr] we proved the existence of the wave operators

$$W_\pm^{(\sigma)} := W_\pm(H^{(\sigma)}, H^{(0)}) = s - \lim_{t\to\pm\infty} \exp(itH^{(\sigma)})\exp(-itH^{(0)}).$$

We state now the main results of the present part. An especially complete result can be obtained under the additional assumption

$$\sigma(x_1) \geq 0, \qquad \text{a.e. } x_1 \in \mathbb{R}. \tag{1.3}$$

Theorem 1.1. *Assume that σ satisfies* (1.1) *and* (1.3)*. Then the wave operators $W_\pm^{(\sigma)}$ exist, are unitary and satisfy*

$$H^{(\sigma)} = W_\pm^{(\sigma)} H^{(0)} W_\pm^{(\sigma)*}. \tag{1.4}$$

In particular, under the condition (1.3) the spectrum of the operator $H^{(\sigma)}$ is purely absolutely continuous. This is also true for general σ.

Theorem 1.2. *Assume that σ satisfies* (1.1)*. Then the operator $H^{(\sigma)}$ has purely absolutely continuous spectrum.*

However, in contrast to the case of non-negative σ now the operator $H^{(\sigma)}$ may be not unitarily equivalent to $H^{(0)}$ and then the wave operators $W_\pm^{(\sigma)}$ are not complete. This is connected with the existence of *additional channels of scattering*. The discussion of this phenomenon is conveniently postponed to Section 6.

1.3 Definition of the operators $H^{(\sigma)}(k)$ on the halfstrip. Direct Integral Decomposition

Let σ be a real-valued periodic function satisfying (1.1) and let $k \in [-\frac{1}{2}, \frac{1}{2}]$. It follows (cf. [Fr]) that the quadratic form

$$\begin{aligned}
&\mathcal{D}[h^{(\sigma)}(k)] := \tilde{H}^1(\Pi),\\
&h^{(\sigma)}(k)[u] := \int_\Pi \left(|(D_1+k)u(x)|^2 + |D_2 u(x)|^2 \right) \, dx + \int_{-\pi}^{\pi} \sigma(x_1)|u(x_1,0)|^2 \, dx_1
\end{aligned} \tag{1.5}$$

is lower semibounded and closed in the Hilbert space $L_2(\Pi)$, so it generates a self-adjoint operator which will be denoted by $H^{(\sigma)}(k)$. In addition to the Neumann (if $\sigma = 0$) or third type (if $\sigma \neq 0$) boundary condition at $\{x_2 = 0\}$, the functions in $\mathcal{D}(H^{(\sigma)})$ satisfy periodic boundary conditions at $\{x_1 \in \{-\pi, \pi\}\}$. The operator $H^{(\sigma)}$ on the halfplane can be partially diagonalized by means of the *Gelfand transformation.* This operator is initially defined for $u \in \mathcal{S}(\mathbb{R}^2_+)$, the Schwartz class on $\mathbb{R}^2_+$, by

$$(\mathcal{U}u)(k,x) := \sum_{n\in\mathbb{Z}} e^{-ik(x_1+2\pi n)} u(x_1+2\pi n, x_2), \qquad k \in [-\tfrac{1}{2}, \tfrac{1}{2}],\, x \in \Pi,$$

and extended by continuity to a *unitary* operator

$$\mathcal{U} : L_2(\mathbb{R}^2_+) \to \int_{-1/2}^{1/2} \oplus L_2(\Pi) \, dk. \tag{1.6}$$

One finds (cf. [Fr]) that

$$\mathcal{U}\, H^{(\sigma)} \,\mathcal{U}^* = \int_{-1/2}^{1/2} \oplus H^{(\sigma)}(k) \, dk. \tag{1.7}$$

This relation allows us to investigate the operator $H^{(\sigma)}$ by studying the fibers $H^{(\sigma)}(k)$.

In [Fr] it was shown that

$$\sigma_{ac}\left(H^{(\sigma)}(k)\right) = [k^2, +\infty), \qquad \sigma_{sc}\left(H^{(\sigma)}(k)\right) = \emptyset. \tag{1.8}$$

In the present part we give a detailed analysis of the point spectrum of $H^{(\sigma)}(k)$.

2 Characterization of eigenvalues of the operator $H^{(\sigma)}(k)$

Let σ be a real-valued periodic function satisfying (1.1) and let $k \in [-\frac{1}{2}, \frac{1}{2}]$, $\lambda \in \mathbb{R}$. In the Hilbert space $L_2(\mathbb{T})$ we consider the quadratic forms

$$\begin{aligned}
&\mathcal{D}[b^{(\sigma)}(\lambda,k)] := H^{1/2}(\mathbb{T}),\\
&b^{(\sigma)}(\lambda,k)[f] := \sum_{n\in\mathbb{Z}} \beta_n(\lambda,k)\, |\hat{f}_n|^2 + \int_{-\pi}^{\pi} \sigma(x_1)|f(x_1)|^2 \, dx_1,
\end{aligned} \tag{2.1}$$

where

$$\beta_n(\lambda,k) := \begin{cases} \sqrt{(n+k)^2-\lambda} & \text{if } (n+k)^2 > \lambda, \\ -\sqrt{\lambda-(n+k)^2} & \text{if } (n+k)^2 \le \lambda. \end{cases} \tag{2.2}$$

It follows from the Sobolev embedding theorems that the forms $b^{(\sigma)}(\lambda,k)$ are lower semibounded and closed, so they generate self-adjoint operators which will be denoted by $B^{(\sigma)}(\lambda,k)$.
The compactness of the embedding of $H^{1/2}(\mathbb{T})$ in $L_2(\mathbb{T})$ implies that the operators $B^{(\sigma)}(\lambda,k)$ have compact resolvent.
Now we characterize the eigenvalues of the operator $H^{(\sigma)}(k)$ as the values λ for which 0 is an eigenvalue of the operators $B^{(\sigma)}(\lambda,k)$. More precisely, we have

Proposition 2.1. *Let $k \in [-\frac{1}{2},\frac{1}{2}]$ and $\lambda \in \mathbb{R}$.*

1. *Let $u \in \mathcal{N}(H^{(\sigma)}(k)-\lambda I)$ and define*

$$f(x_1) := u(x_1,0), \qquad x_1 \in \mathbb{T}. \tag{2.3}$$

Then $f \in \mathcal{N}(B^{(\sigma)}(\lambda,k))$, $\hat{f}_n = 0$ if $(n+k)^2 \le \lambda$ and, moreover,

$$u(x) = \frac{1}{\sqrt{2\pi}} \sum_{(n+k)^2>\lambda} \hat{f}_n\, e^{inx_1}\, e^{-\beta_n(\lambda,k)\,x_2}, \qquad x \in \Pi. \tag{2.4}$$

2. *Let $f \in \mathcal{N}(B^{(\sigma)}(\lambda,k))$ such that $\hat{f}_n = 0$ if $(n+k)^2 \le \lambda$ and define u by* (2.4).
Then $u \in \mathcal{N}(H^{(\sigma)}(k)-\lambda I)$ and, moreover, (2.3) *holds.*

For the proof of Proposition 2.1 we use the following notation. For $u \in L_2(\Pi)$ and $n \in \mathbb{Z}$ we define

$$\hat{u}_n(x_2) := \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} u(x) e^{-inx_1}\, dx_1, \qquad x_2 \in \mathbb{R}_+,$$

so that, with respect to convergence in $L_2(\Pi)$,

$$u(x) = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{Z}} e^{inx_1}\, \hat{u}_n(x_2), \qquad x \in \Pi.$$

Moreover, one finds that $u \in \tilde{H}^1(\Pi)$ iff

$$\hat{u}_n \in H^1(\mathbb{R}_+),\ n \in \mathbb{Z}, \quad \text{and} \quad \sum_{n\in\mathbb{Z}} \big((1+n^2)\|\hat{u}_n\|^2 + \|D_2\hat{u}_n\|^2\big) < \infty.$$

The proof of the following observation is straightforward.

Lemma 2.2. *Let $k \in [-\frac{1}{2},\frac{1}{2}]$ and $\lambda \in \mathbb{R}$.*

1. *Let $u \in \tilde{H}^1(\Pi)$, then the following are equivalent:*

(i) $u \in \mathcal{N}(H^{(\sigma)}(k) - \lambda I)$,

(ii) $\int_0^\infty D_2\hat{u}_n \overline{D_2\varphi}\, dx_2 + \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \sigma(x_1) u(x_1, 0) e^{-inx_1}\, dx_1\, \overline{\varphi(0)} =$
$= (\lambda - (n+k)^2) \int_0^\infty \hat{u}_n \overline{\varphi}\, dx_2, \qquad n \in \mathbb{Z},\ \varphi \in H^1(\mathbb{R}_+)$.

2. Let $f \in H^{1/2}(\mathbb{T})$*, then the following are equivalent:*

(i) $f \in \mathcal{N}(B^{(\sigma)}(\lambda, k))$,

(ii) $\beta_n(\lambda, k)\, \hat{f}_n + \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \sigma(x_1) f(x_1) e^{-inx_1}\, dx_1 = 0, \qquad n \in \mathbb{Z}$.

Proof of Proposition 2.1. The proof follows easily from Lemma 2.2. Note that if $u \in \mathcal{N}(H^{(\sigma)}(k) - \lambda I)$, then $D_2^2 \hat{u}_n = ((\lambda - (n+k)^2)\hat{u}_n$. Therefore

$$\hat{u}_n(x_2) = \begin{cases} 0 & \text{if} \quad (n+k)^2 \leq \lambda, \\ \hat{f}_n\, e^{-\beta_n(\lambda,k)\, x_2} & \text{if} \quad (n+k)^2 > \lambda, \end{cases}$$

with f defined by (2.3). □

Remark 2.3. Obviously, the statement of Proposition 2.1 does not depend on the definition of $\beta_n(\lambda, k)$ for $(n+k)^2 \leq \lambda$. The reason for our choice (2.2) is of technical nature and will become clear in Subsection 4.2 below.

3 The case of non-negative σ

Proposition 2.1 allows us to deduce easily the main result if σ is non-negative. We start with the operators $H^{(\sigma)}(k)$ on the halfstrip.

Theorem 3.1. *Assume that* σ *satisfies* (1.1) *and* (1.3) *and let* $k \in [-\frac{1}{2}, \frac{1}{2}]$*. Then the operator* $H^{(\sigma)}(k)$ *has purely absolutely continuous spectrum.*

Proof. In view of (1.8) it suffices to prove that $H^{(\sigma)}(k)$ has no eigenvalues. For this we use Proposition 2.1. Let $\lambda \in \mathbb{R}$ and $f \in \mathcal{N}(B^{(\sigma)}(\lambda, k))$ such that $\hat{f}_n = 0$ if $(n+k)^2 \leq \lambda$. It follows that

$$b^{(\sigma)}(\lambda, k)[f] \geq \gamma \|f\|^2$$

where $\gamma := \min\{\beta_n(\lambda, k) :\ n \in \mathbb{Z},\ \hat{f}_n \neq 0\} > 0$. Together with $b^{(\sigma)}(\lambda, k)[f] = 0$ this implies $f = 0$. So by Proposition 2.1 (1), λ is not an eigenvalue of $H^{(\sigma)}(k)$. □

Concerning the operator $H^{(\sigma)}$ on the halfplane we obtain immediately the

Proof of Theorem 1.1. In [Fr] we showed that $W_\pm^{(\sigma)}$ is unitarily equivalent to the direct integral of the operators $W_\pm^{(\sigma)}(k)$, $k \in [-\frac{1}{2}, \frac{1}{2}]$. The latter were shown to be complete, and by Theorem 3.1 they are actually unitary. Thus $W_\pm^{(\sigma)}$ is unitary and (1.4) follows from the intertwining property of wave operators. □

4 The general case

4.1 The point spectrum of the operators $H^{(\sigma)}(k)$

If we impose no additional condition on σ we have the following result on the point spectrum of the operators $H^{(\sigma)}(k)$.

Theorem 4.1. *Assume that σ satisfies* (1.1) *and let $k \in [-\frac{1}{2}, \frac{1}{2}]$. Then $\sigma_p\left(H^{(\sigma)}(k)\right)$ (if non-empty) consists of eigenvalues of finite multiplicities which may accumulate at $+\infty$ only.*

Note that the case of an infinite sequence of (embedded) eigenvalues actually occurs.

Example 4.2. Let $\sigma \equiv \sigma_0 < 0$ be a negative constant and $k \in [-\frac{1}{2}, \frac{1}{2}]$. Then

$$\sigma_p\Big(H^{(\sigma)}(k)\Big) = \{-\sigma_0^2 + (n+k)^2 :\ n \in \mathbb{Z}\}.$$

This follows easily by Proposition 2.1 or directly by separation of variables.

For the proof of Theorem 4.1 we need an auxiliary result. For $k \in [-\frac{1}{2}, \frac{1}{2}]$, $\lambda \in \mathbb{R}$ we denote by $\mu_m(\lambda, k)$, $m \in \mathbb{N}$, the eigenvalues of $B^{(\sigma)}(\lambda, k)$ arranged in non-decreasing order and repeated according to their multiplicities. Then we have

Lemma 4.3. *Let $k \in [-\frac{1}{2}, \frac{1}{2}]$, then the functions $\mu_m(., k)$, $m \in \mathbb{N}$, are continuous and strictly decreasing on $\mathbb{R}$.*

The proof (of *strict* monotonicity) uses an analyticity argument and is conveniently postponed to Subsection 4.2.

Proof of Theorem 4.1. Proposition 2.1 (1) implies for $\lambda \in \mathbb{R}$

$$\dim \mathcal{N}(H^{(\sigma)}(k) - \lambda I) \leq \dim \mathcal{N}(B^{(\sigma)}(\lambda, k)). \tag{4.1}$$

Since $B^{(\sigma)}(\lambda, k)$ has compact resolvent, it follows that eigenvalues λ of $H^{(\sigma)}(k)$ have finite multiplicities.
To prove that the only possible accumulation point of $\sigma_p\left(H^{(\sigma)}(k)\right)$ is $+\infty$, let $\Lambda = (\lambda_-, \lambda_+)$ be an open interval. It follows from (4.1) and Lemma 4.3 that

$$\begin{aligned}
&\sharp_{cm}\{\lambda \in (\lambda_-, \lambda_+) :\ \lambda \text{ is eigenvalue of } H^{(\sigma)}(k)\} \leq \\
&\quad \leq \sum_{\lambda \in (\lambda_-, \lambda_+)} \dim \mathcal{N}(B^{(\sigma)}(\lambda, k)) = \\
&\quad = \sharp\{m \in \mathbb{N} :\ \mu_m(\lambda, k) = 0 \text{ for some } \lambda \in (\lambda_-, \lambda_+)\} = \\
&\quad = \sharp\{m \in \mathbb{N} :\ \mu_m(\lambda_-, k) > 0 \text{ and } \mu_m(\lambda_+, k) < 0\} = \\
&\quad = \sharp_{cm}\{\mu < 0 :\ \mu \text{ is eigenvalue of } B^{(\sigma)}(\lambda_+, k)\} - \\
&\qquad - \sharp_{cm}\{\mu \leq 0 :\ \mu \text{ is eigenvalue of } B^{(\sigma)}(\lambda_-, k)\},
\end{aligned} \tag{4.2}$$

where $\sharp_{cm}\{...\}$ means that the cardinality of $\{...\}$ is determined according to multiplicities. The RHS of (4.2) is finite since $B^{(\sigma)}(\lambda_+, k)$, $B^{(\sigma)}(\lambda_-, k)$ are lower semibounded and have compact resolvent. This completes the proof of the theorem. $\square$

Remark 4.4. We emphasize the equality

$$\begin{aligned}\sharp_{cm}\{\lambda \in (-\infty, k^2) :\ \lambda \text{ is eigenvalue of } H^{(\sigma)}(k)\} = \\ = \sharp_{cm}\{\mu < 0 :\ \mu \text{ is eigenvalue of } B^{(\sigma)}(k^2, k)\}.\end{aligned} \tag{4.3}$$

Indeed, it follows from Proposition 2.1 (2) that the estimate (4.1) becomes an equality for $\lambda < k^2$, therefore also (4.2) for $\lambda_+ = k^2$, and we obtain (4.3) by choosing $-\lambda_-$ so large that $B^{(\sigma)}(\lambda_-, k)$ is positive.
The equality (4.3) can be used to obtain estimates on the number of eigenvalues of $H^{(\sigma)}(k)$ below k^2 and on its asymptotics in the limit of large coupling constant. Such calculations for the operators $B^{(\sigma)}(k^2, k)$ are rather standard, so we do not go into details.

4.2 Complexification

Now we extend the operator family $B^{(\sigma)}(\lambda, k)$ to complex values of λ and k. For this construction we fix $k_0 \in [-\frac{1}{2}, \frac{1}{2}]$, $\lambda_0 \in \mathbb{R} \setminus \{(n+k_0)^2 : n \in \mathbb{Z}\}$. We can choose $\delta_0 > 0$ (depending on λ_0, k_0) such that

$$(n+\kappa)^2 - z \neq 0, \qquad n \in \mathbb{Z},$$

for all $z, \kappa \in \mathbb{C}$ such that $|z - \lambda_0| < \delta_0$, $|\kappa - k_0| < \delta_0$. Therefore, if we put

$$\tilde{U} := \{z \in \mathbb{C} : |z - \lambda_0| < \delta\}, \qquad \tilde{V} := \{\kappa \in \mathbb{C} : |\kappa - k_0| < \delta\},$$

the functions β_n, $n \in \mathbb{Z}$, admit a unique analytic continuation to $\tilde{U} \times \tilde{V}$, and we can define sectorial and closed forms $b^{(\sigma)}(z, \kappa)$ for $z \in \tilde{U}$, $\kappa \in \tilde{V}$ by (2.1) with $\beta_n(\lambda, k)$ replaced by $\beta_n(z, \kappa)$. The corresponding m-sectorial operators will be denoted by $B^{(\sigma)}(z, \kappa)$. For fixed $\kappa \in \tilde{V}$ ($z \in \tilde{U}$, respectively) they form an analytic family of type (B) with respect to $z \in \tilde{U}$ ($\kappa \in \tilde{V}$, respectively) (see, e.g., Section VII.4 in [K]).
From this construction we obtain

Lemma 4.5. *Let $k_0 \in [-\frac{1}{2}, \frac{1}{2}]$, $\lambda_0 \in \mathbb{R} \setminus \{(n+k_0)^2 : n \in \mathbb{Z}\}$ such that 0 is an eigenvalue of $B^{(\sigma)}(\lambda_0, k_0)$. Then there exist open neighbourhoods U, $V \subset \mathbb{R}$ of λ_0, k_0 and a real-analytic function $h : U \times V \to \mathbb{C}$ such that for all $\lambda \in U$, $k \in V \cap [-\frac{1}{2}, \frac{1}{2}]$ one has*

$$0 \in \sigma_p\left(B^{(\sigma)}(\lambda, k)\right) \qquad \textit{iff} \qquad h(\lambda, k) = 0.$$

Proof. The proof is rather standard, so we only sketch the major steps. We consider the family $B^{(\sigma)}(z,\kappa)$, $z \in \tilde{U}$, $\kappa \in \tilde{V}$, constructed above. Since these operators have compact resolvent, we can use a Riesz projection to separate the eigenvalues around 0 from the rest of the spectrum. The resulting operator has finite rank and is analytic with respect to z and κ, so its determinant h has the desired properties. □

Our next goal is to show that for every $\lambda \in U$ the function $h(\lambda, .)$ constructed above is not identically zero. For the proof of this we need to consider quasimomenta $\kappa = k + iy$ with large imaginary part y. So fix $k \in [-\frac{1}{2}, \frac{1}{2}]$, $\lambda \in \mathbb{R} \setminus \{(n+k)^2 : n \in \mathbb{Z}\}$, then the above construction (with λ_0, k_0 replaced by λ, k) yields a $\delta > 0$ and an analytic family $B^{(\sigma)}(\lambda,\kappa)$, $|\kappa - k| < \delta$. If we assume in addition that $k \neq 0$ and choose $\delta \in (0, |k|)$, we find that

$$(n+\kappa)^2 - \lambda \neq 0, \qquad n \in \mathbb{Z},$$

holds for all $\kappa \in \mathbb{C}$ such that $|\mathrm{Re}\,\kappa - k| < \delta$. Therefore $B^{(\sigma)}(\lambda,\kappa)$ admits a further analytic extension to

$$\tilde{\tilde{V}} := \{\kappa \in \mathbb{C} : |\mathrm{Re}\,\kappa - k| < \delta\}.$$

Concerning quasimomenta with large imaginary part we have the technical

Lemma 4.6. *Let $k \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$, $\lambda \in \mathbb{R} \setminus \{(n+k)^2 : n \in \mathbb{Z}\}$ and $\delta \in (0, |k|)$ as above. Then there exist constants $y_0 = y_0(\lambda, k, \delta)$, $C = C(\lambda, k, \delta)$ such that for all $k' \in [-\frac{1}{2}, \frac{1}{2}]$, $y \in \mathbb{R}$ satisfying $|k' - k| < \delta$, $|y| > y_0$ the operator $B^{(\sigma)}(\lambda, k' + iy)$ is boundedly invertible with*

$$\left\| \left(B^{(\sigma)}(\lambda, k' + iy) \right)^{-1} \right\| \leq \frac{C}{1 + |y|}.$$

Proof. It suffices to find constants y_0, $\tilde{C} = \tilde{C}(\lambda, k, \delta) > 0$ such that for all $0 \neq f \in H^{1/2}(\mathbb{T})$, $k' \in [-\frac{1}{2}, \frac{1}{2}]$, $y \in \mathbb{R}$ satisfying $|k' - k| < \delta$, $|y| > y_0$ there exists $0 \neq g \in H^{1/2}(\mathbb{T})$ such that

$$|b^{(\sigma)}(\lambda, k' + iy)[f, g]| \geq \tilde{C}\,(1 + |y|)\,\|f\|\|g\|.$$

For given $0 \neq f \in H^{1/2}(\mathbb{T})$, $k' \in [-\frac{1}{2}, \frac{1}{2}] \cap (k - \delta, k + \delta)$, $y \in \mathbb{R}$ we define g by its Fourier coefficients

$$\hat{g}_n := \frac{\beta_n(\lambda, k' + iy)}{|\beta_n(\lambda, k' + iy)|}\,\hat{f}_n, \qquad n \in \mathbb{Z}.$$

(Note that $\beta_n(\lambda, k' + iy) \neq 0$ by the choice of δ.) Then we have $0 \neq g \in H^{1/2}(\mathbb{T})$, $\|g\| = \|f\|$ and

$$|b^{(\sigma)}(\lambda, k' + iy)[f, g]| \geq \sum_{n \in \mathbb{Z}} |\beta_n(\lambda, k' + iy)||\hat{f}_n|^2 - \tfrac{1}{2}\|\sqrt{|\sigma|}f\|^2 - \tfrac{1}{2}\|\sqrt{|\sigma|}g\|^2. \quad (4.4)$$

Using the elementary estimates

$$\begin{aligned} |\beta_n(\lambda, k'+iy)| &\geq c_1\,(1+|y|), \qquad n\in\mathbb{Z},\, |k'-k|<\delta, \\ |\beta_n(\lambda, k'+iy)| &\geq c_2\,(1+|n|), \qquad n\in\mathbb{Z},\, |k'-k|<\delta, \end{aligned} \tag{4.5}$$

(with some constants $c_1 = c_1(\lambda, k, \delta) > 0, c_2 = c_2(\lambda, k, \delta) > 0$) and the Sobolev embedding theorem we find that for sufficiently large y_0

$$\|\sqrt{|\sigma|}f\|^2 \leq \tfrac{1}{2}\sum_{n\in\mathbb{Z}} |\beta_n(\lambda, k'+iy)||\hat{f}_n|^2, \qquad |k'-k|<\delta,\, |y|>y_0.$$

Using a similar estimate for $\|\sqrt{|\sigma|}g\|^2$ and (4.4), (4.5) we obtain

$$|b^{(\sigma)}(\lambda, k'+iy)[f,g]| \geq \tfrac{1}{2}\,c_1\,(1+|y|)\,\|f\|\|g\|, \qquad |k'-k|<\delta, |y|>y_0,$$

which concludes the proof. □

As announced above, we have

Lemma 4.7. *Let k_0, λ_0, h, U and V be as in Lemma 4.5. Then for all $\lambda \in U$ one has $h(\lambda, .) \not\equiv 0$.*

Proof. To arrive at a contradiction we assume that $h(\lambda, .) \equiv 0$ for some $\lambda \in U$. We choose $k \in V \setminus \{0\}$ and consider the family $B^{(\sigma)}(\lambda, \kappa)$, $\kappa \in \tilde{\tilde{V}}$ constructed above. It follows from the Analytic Fredholm Alternative (see, e.g., Theorem VII.1.10 in [K]) that all operators of this family have 0 as an eigenvalue. But this contradicts Lemma 4.6. □

As an immediate consequence of Lemmas 4.5 and 4.7 and relation (4.1) we obtain the following result which will be needed in Subsection 4.3 to prove that the spectrum of the operator $H^{(\sigma)}$ is purely absolutely continuous.

Corollary 4.8. *There exists a countable number of open intervals U_j, $V_j \subset \mathbb{R}$ and real-analytic functions $h_j : U_j \times V_j \to \mathbb{C}$ satisfying*

1. *for all $k \in [-\frac{1}{2}, \frac{1}{2}]$ and $\lambda \in \mathbb{R} \setminus \{(n+k)^2 : n \in \mathbb{Z}\}$ such that $\lambda \in \sigma_p\left(H^{(\sigma)}(k)\right)$ there is a j such that $(\lambda, k) \in U_j \times V_j$ and $h_j(\lambda, k) = 0$, and*
2. *for all j and all $\lambda \in U_j$ one has $h_j(\lambda, .) \not\equiv 0$.*

To complete this subsection we prove Lemma 4.3 which was used in the proof of Theorem 4.1.

Proof of Lemma 4.3. That $\mu_m(., k)$ is a continuous, non-increasing function follows from the variational principle and the continuity and monotonicity of the operators $B^{(\sigma)}(\lambda, k)$ with respect to λ.
To prove the strict monotonicity we assume to the contrary that for some $m \in \mathbb{N}$ the function $\mu_m(., k)$ coincides on an interval Λ with a constant $\mu_0 \in \mathbb{R}$. We choose $\lambda_0 \in \Lambda \setminus \{(n+k)^2 : n \in \mathbb{Z}\}$ and consider the family $B^{(\sigma)}(z, k)$, $z \in \tilde{U}$,

constructed at the beginning of this subsection (with λ_0, k_0 replaced by λ_0, k). It follows from the Analytic Fredholm Alternative (see, e.g., Theorem VII.1.10 in [K]) that μ_0 is an eigenvalue of $B^{(\sigma)}(z,k)$ also for complex $z \in \tilde{U}$.
However, let $z \in \tilde{U} \cap \mathbb{C}_\pm$ and $f \in \mathcal{N}(B^{(\sigma)}(z,k) - \mu_0 I)$. We have $\mp \operatorname{Im} \beta_n(z,k) > 0$, $n \in \mathbb{Z}$, so $\operatorname{Im} b^{(\sigma)}(\lambda,k)[f] = 0$ implies that $\hat{f}_n = 0$, $n \in \mathbb{Z}$, i.e., $f = 0$. So μ_0 is not an eigenvalue of $B^{(\sigma)}(z,k)$. □

4.3 Proof of Theorem 1.2

Now we prove Theorem 1.2 following the method suggested in [FiKl]. We need the following result from Complex Analysis of Several Variables which can be proved by means of the Implicit Function Theorem (see [FiKl]).

Lemma 4.9. *Let U, $V \subset \mathbb{R}$ be open intervals and $h : U \times V \to \mathbb{C}$ be real-analytic. Let $\Lambda \subset U$ with* $\operatorname{meas} \Lambda = 0$ *such that for all $\lambda \in \Lambda$ one has $h(\lambda, .) \not\equiv 0$. Then*

$$\operatorname{meas} \{k \in V : \; h(\lambda,k) = 0 \quad \textit{for some } \lambda \in \Lambda\} = 0.$$

Proof of Theorem 1.2. Let $\Lambda \subset \mathbb{R}$ with $\operatorname{meas} \Lambda = 0$. We denote the spectral projection of $H^{(\sigma)}$ ($H^{(\sigma)}(k)$, respectively) corresponding to Λ by $E^{(\sigma)}(\Lambda)$ ($E^{(\sigma)}(\Lambda,k)$, respectively). Then it follows from (1.7) that

$$\mathcal{U}\, E^{(\sigma)}(\Lambda)\, \mathcal{U}^* = \int_{-1/2}^{1/2} \oplus E^{(\sigma)}(\Lambda,k)\, dk$$

and we have to prove that this operator is equal to 0.
For this we write $[-\frac{1}{2}, \frac{1}{2}] = K_1 \cup K_2 \cup K_3$ where

$$\begin{aligned} K_1 &= \left\{ k \in [-\tfrac{1}{2}, \tfrac{1}{2}] : \; \sigma_p \left(H^{(\sigma)}(k) \right) \cap \Lambda = \emptyset \right\}, \\ K_2 &= \left\{ k \in [-\tfrac{1}{2}, \tfrac{1}{2}] : \; \sigma_p \left(H^{(\sigma)}(k) \right) \cap \Lambda \cap \{(n+k)^2 : \; n \in \mathbb{Z}\} \neq \emptyset \right\}, \\ K_3 &= \left\{ k \in [-\tfrac{1}{2}, \tfrac{1}{2}] : \; \emptyset \neq \sigma_p \left(H^{(\sigma)}(k) \right) \cap \Lambda \subset \left(\mathbb{R} \setminus \{(n+k)^2 : \; n \in \mathbb{Z}\} \right) \right\}. \end{aligned}$$

Since $\sigma_{sc}\left(H^{(\sigma)}(k)\right) = \emptyset$ we immediately obtain $E^{(\sigma)}(\Lambda,k) = 0$ for $k \in K_1$. Now

$$K_2 \subset \bigcup_{n \in \mathbb{Z}} \{k \in [-\tfrac{1}{2}, \tfrac{1}{2}] : \; (n+k)^2 - \lambda = 0 \quad \text{for some } \lambda \in \Lambda\}, \tag{4.6}$$

and with the notation of Corollary 4.8

$$K_3 \subset \bigcup_j \{k \in V_j \cap [-\tfrac{1}{2}, \tfrac{1}{2}] : \; h_j(\lambda,k) = 0 \quad \text{for some } \lambda \in U_j \cap \Lambda\}. \tag{4.7}$$

It follows from Lemma 4.9 that $\operatorname{meas} K_2 = \operatorname{meas} K_3 = 0$, which concludes the proof. □

5 The case of a trigonometric polynomial σ

We have seen in Example 4.2 that the operators $H^{(\sigma)}(k)$ may have embedded eigenvalues. Let us investigate this phenomenon under the additional assumption that only finitely many Fourier coefficients of σ are non-zero. Note that in this case the operator $B^{(\sigma)}(\lambda,k)$ acts in Fourier space as a finite-diagonal matrix. This allows us to exclude the existence of large embedded eigenvalues.

Proposition 5.1. *Assume that σ is a trigonometric polynomial of degree $N > 0$ and let $k \in [-\frac{1}{2},\frac{1}{2}]$. Then*

$$\sigma_p\left(H^{(\sigma)}(k)\right) \subset \left(-\|\sigma_-\|_\infty^2 + k^2, (N-|k|)^2\right).$$

Here $\sigma_- := \frac{1}{2}(|\sigma|-\sigma)$ denotes the negative part of σ.

Proof. The proof of $\sigma_p\left(H^{(\sigma)}(k)\right) \subset [-\|\sigma_-\|_\infty^2 + k^2, +\infty)$ is similar to the proof of Theorem 3.1. Moreover, it is easy to see that $-\|\sigma_-\|_\infty^2 + k^2 \in \sigma_p\left(H^{(\sigma)}(k)\right)$ only if σ coincides a.e. with a negative constant, which is excluded by the assumption $N > 0$.
Let us prove now that $\sigma_p\left(H^{(\sigma)}(k)\right) \subset (-\infty,(N-|k|)^2)$. For this we use Proposition 2.1. Let $\lambda \geq (N-|k|)^2$ and $f \in \mathcal{N}(B^{(\sigma)}(\lambda,k))$ such that

$$\hat{f}_n = 0 \qquad \text{if } (n+k)^2 \leq \lambda. \tag{5.1}$$

In particular, we see from $B^{(\sigma)}(\lambda,k)f = 0$ that

$$\sqrt{(n+k)^2-\lambda}\,\hat{f}_n + \frac{1}{\sqrt{2\pi}}\sum_{m=-N}^{N}\hat{\sigma}_m\hat{f}_{n-m} = 0 \qquad \text{if } (n+k)^2 \geq \lambda. \tag{5.2}$$

The estimate

$$\sharp\{n\in\mathbb{Z} :\ (n+k)^2 \leq \lambda\} \geq \sharp\{n\in\mathbb{Z} :\ (n+k)^2 \leq (N-|k|)^2\} \geq 2N$$

and (5.1) imply that $\hat{f}_n = 0$ for at least $2N$ consecutive n. Using $\hat{\sigma}_N = \overline{\hat{\sigma}_{-N}} \neq 0$ it is easy to see from (5.2) that $\hat{f}_n = 0$ for all n, i.e. $f = 0$. So by Proposition 2.1 (1), λ is not an eigenvalue of $H^{(\sigma)}(k)$. □

We show now that embedded eigenvalues in the interval $[(N-1+|k|)^2,(N-|k|)^2)$ can occur but are "rare".

Proposition 5.2. *Assume that σ is a trigonometric polynomial of degree $N > 0$ and let $k \in (-\frac{1}{2},\frac{1}{2})$. Then $H^{(\sigma)}(k)$ may have only simple eigenvalues in $[(N-1+|k|)^2,(N-|k|)^2)$ and the set*

$$\{(\lambda,k)\in\mathbb{R}\times(-\tfrac{1}{2},\tfrac{1}{2}) :\ \lambda\in\sigma_p\left(H^{(\sigma)}(k)\right)\cap[(N-1+|k|)^2,(N-|k|)^2)\} \tag{5.3}$$

is finite.

For the proof of this proposition we introduce the following auxiliary operators in the Hilbert space $l_2(\mathbb{N})$.

$$\mathcal{D}(A^{(\sigma)}(\lambda,k)) := \{\alpha \in l_2(\mathbb{N}) : \sum_{n=1}^{\infty}(1+n^2)|\alpha_n|^2 < \infty\},$$

$$\Big(A^{(\sigma)}(\lambda,k)\,\alpha\Big)_n := \begin{cases} \beta_n(\lambda,k)\,\alpha_n + \frac{1}{\sqrt{2\pi}}\sum_{m=-N}^{n-1}\hat{\sigma}_m\,\alpha_{n-m} & \text{if } n \le N, \\ \beta_n(\lambda,k)\,\alpha_n + \frac{1}{\sqrt{2\pi}}\sum_{m=-N}^{N}\hat{\sigma}_m\,\alpha_{n-m} & \text{if } n > N. \end{cases} \tag{5.4}$$

The operators $A^{(\sigma)}(\lambda,k)$ are self-adjoint and have compact resolvent.

Lemma 5.3. *Let $k \in (-\frac{1}{2},\frac{1}{2})$ and $\lambda \in [(N-1+|k|)^2,(N-|k|)^2)$. Then λ is an eigenvalue of $H^{(\sigma)}(k)$ iff there exist $0 \ne \alpha^+$, $\alpha^- \in \mathcal{D}(A^{(\sigma)}(\lambda,k))$ such that $A^{(\sigma)}(\lambda,k)\,\alpha^+ = A^{(\sigma)}(\lambda,-k)\,\alpha^- = 0$ and $\alpha_n^+ = \alpha_n^- = 0$ for $n < N$. In this case, λ is a simple eigenvalue.*

Proof. We use Proposition 2.1. If λ is an eigenvalue of $H^{(\sigma)}(k)$, there exists a $0 \ne f \in \mathcal{N}(B^{(\sigma)}(\lambda,k))$ such that $\hat{f}_n = 0$ if $|n| < N$. We note that the only relation between the positive and the negative Fourier coefficients of f is the equation

$$\hat{\sigma}_N\hat{f}_{-N} + \hat{\sigma}_{-N}\hat{f}_N = 0.$$

Therefore f is unique up to multiples. We put

$$\alpha_n^+ := \hat{f}_n, \qquad \alpha_n^- := \overline{\hat{f}_{-n}}, \qquad n \in \mathbb{N}, \tag{5.5}$$

and find (using $\hat{\sigma}_n = \overline{\hat{\sigma}_{-n}}$, $n \in \mathbb{Z}$) that α^+, α^- are as claimed.
Conversely, let α^+, α^- have the properties of the lemma. Then $\alpha_N^+\alpha_N^- \ne 0$ and multiplying α^+ by a constant if necessary, we can assume that $\hat{\sigma}_N\overline{\alpha_N^-} + \hat{\sigma}_{-N}\alpha_N^+ = 0$. Defining f by (5.5) and $\hat{f}_n := 0$ if $|n| < N$ we find that $0 \ne f \in \mathcal{N}(B^{(\sigma)}(\lambda,k))$, so λ is an eigenvalue of $H^{(\sigma)}(k)$ by Proposition 2.1 (2). This completes the proof. ☐

The reason for introducing the operators $A^{(\sigma)}(\lambda,k)$ is that they are not only monotone with respect to λ but also with respect to k. This is essentially used in the

Proof of Proposition 5.2. It remains to prove that the set (5.3) is finite. We denote by $\nu_m(\lambda,k)$, $m \in \mathbb{N}$, the eigenvalues of the operator $A^{(\sigma)}(\lambda,k)$, arranged in non-decreasing order and repeated according to their multiplicities. By the same arguments as in the proof of Lemma 4.3 one finds that the functions $\nu_m(\lambda,.)$ ($\nu_m(.,k)$, respectively) are continuous and strictly increasing (strictly decreasing, respectively) for fixed λ (k, respectively).
Now Lemma 5.3 implies that if λ is an eigenvalue of $H^{(\sigma)}(k)$ in $[(N-1+|k|)^2,(N-|k|)^2)$ then there exist m, $m' \in \mathbb{N}$ such that

$$\nu_m(\lambda,k) = \nu_{m'}(\lambda,-k) = 0. \tag{5.6}$$

It follows easily from the monotonicity properties mentioned above that for each pair $(m,m') \in \mathbb{N}\times\mathbb{N}$ there exists at most one pair (λ,k) with $\lambda \in [(N-1+|k|)^2,(N-|k|)^2)$ such that (5.6) holds. Since the functions ν_m are strictly positive for sufficiently large m we conclude that the set (5.3) is finite. $\square$

Example 5.4. In the case $N=1$ it is convenient to write σ as

$$\sigma(x_1) := -\alpha + \mathrm{Re}\beta\cos x_1 + \mathrm{Im}\beta\sin x_1, \qquad x_1 \in \mathbb{T},$$

with $\alpha\in\mathbb{R}$, $\beta\in\mathbb{C}$. Under the conditions

$$0<\alpha<1, \qquad 0<|\beta|\le 1-\alpha, \tag{5.7}$$

one finds that

$$\nu_m(\lambda,k)>0 \qquad \text{for } m\ge 2,\ k\in(-\tfrac{1}{2},\tfrac{1}{2}),\ \lambda\in[k^2,(1-|k|)^2).$$

Thus it follows from (5.6) and the strict monotonicity of $\nu_1(\lambda,.)$ that the operator $H^{(\sigma)}(k)$ has no eigenvalues in $[k^2,(1-|k|)^2)$ for $k\neq 0$. We consider the case $k=0$. Under condition (5.7) one easily derives the estimates

$$\begin{aligned} \nu_1(\lambda,0)&\ge 0 &&\text{for } \lambda\in[0,1-(\alpha+|\beta|)^2],\\ \nu_1(\lambda,0)&< 0 &&\text{for } \lambda\in(1-\alpha^2,1), \end{aligned}$$

which imply that $H^{(\sigma)}(0)$ has a (unique) embedded eigenvalue in $[0,1)$. It can be shown (see Remark 5.5 below) that it depends real-analytically on the "coupling parameter" $|\beta|>0$.
Let us emphasize that if $0<\alpha<1$ and $\beta=0$, the operator $H^{(\sigma)}(0)$ has embedded eigenvalues $-\alpha^2+m^2$, $m\in\mathbb{N}$, each double degenerate (see Example 4.2). As soon as the coupling is turned on (i.e., $|\beta|>0$) all the eigenvalues above 1 as well as one of the eigenvalues in $(0,1)$ dissolve, whereas the other one of the eigenvalues in $(0,1)$ depends smoothly on $|\beta|\in[0,1-\alpha]$.

Remark 5.5. Let us mention that the eigenvalue in the above example is due to the following symmetry. Since the operator is (up to unitary equivalence) invariant under a shift with respect to x_1 we may assume that $\beta\in\mathbb{R}$. Then σ is even with respect to $x_1=0$ and so for $k=0$ the decomposition into even and odd functions reduces the operator $H^{(\sigma)}(0)$. It remains to notice that the essential spectrum of the part of the operator acting on odd functions starts at the point $\lambda=1$.

6 Additional Channels of Scattering of the operators $H^{(\sigma)}$

6.1 Additional Channels due to discrete eigenvalues

Here we construct the additional channels of scattering of $H^{(\sigma)}$ which arise from the discrete eigenvalues of the operators $H^{(\sigma)}(k)$.
For $k\in[-\frac{1}{2},\frac{1}{2}]$ we denote by

$$\lambda_1(k)\le\lambda_2(k)\le\cdots\lambda_{l(k)}(k)<k^2 \tag{6.1}$$

the discrete eigenvalues of $H^{(\sigma)}(k)$, arranged in non-decreasing order and repeated according to their multiplicities. By Theorem 4.1 $l(k)$ is a finite number, possibly equal to 0. It is convenient to set $\lambda_l(k) := k^2$ if $l > l(k)$. The functions λ_l are continuous on $[-\frac{1}{2}, \frac{1}{2}]$ for each $l \in \mathbb{N}$. Combining this with (1.7) we find

$$\sigma\left(H^{(\sigma)}\right) = \bigcup_{l \in \mathbb{N}} \lambda_l([-\tfrac{1}{2}, \tfrac{1}{2}]) \cup [0, +\infty), \tag{6.2}$$

i.e., the spectrum of $H^{(\sigma)}$ has band structure.
According to Theorem 1.2 none of the functions λ_l is constant (since this would correspond to an eigenvalue of $H^{(\sigma)}$).
To construct the additional channels of scattering we introduce some notation. We put

$$\mathcal{K}_l := \{k \in [-\tfrac{1}{2}, \tfrac{1}{2}] : \, l \le l(k)\}, \qquad l \in \mathbb{N}_0.$$

These sets are open in $[-\frac{1}{2}, \frac{1}{2}]$ and $\mathcal{K}_l = \emptyset$ for sufficiently large l. We define

$$l_0 := \max\{l \in \mathbb{N}_0 : \, \mathcal{K}_l \neq \emptyset\}.$$

Now assume $l_0 > 0$ (which means that some of the operators $H^{(\sigma)}(k)$ have discrete eigenvalues). For each $k \in [-\frac{1}{2}, \frac{1}{2}]$ we can choose orthonormal eigenfunctions $\psi_l(.,k)$, $1 \le l \le l(k)$, corresponding to the eigenvalues (6.1),

$$H^{(\sigma)}(k)\, \psi_l(.,k) = \lambda_l(k)\, \psi_l(.,k),$$

such that the mappings

$$\mathcal{K}_l \to L_2(\Pi), \; k \mapsto \psi_l(.,k), \qquad 1 \le l \le l_0,$$

are piecewise analytic. Recall that the functions $\psi_l(.,k)$ are of the form (2.4). It is convenient to define $\psi_l(.,k) := 0$ if $k \notin \mathcal{K}_l$ and to extend the functions $\psi_l(.,k)$ periodically with respect to the variable x_1 to functions on $\mathbb{R}^2_+$ for all $k \in [-\frac{1}{2}, \frac{1}{2}]$.
For $1 \le l \le l_0$ we denote by $P_l(k)$, $k \in [-\frac{1}{2}, \frac{1}{2}]$, the projection in $L_2(\Pi)$ onto the subspace spanned by $\psi_l(.,k)$. With this notation, we call the subspaces

$$\mathfrak{C}_l := \mathcal{R}\left(\mathcal{U}^* \left(\int_{-1/2}^{1/2} \oplus P_l(k)\, dk\right) \mathcal{U}\right), \qquad 1 \le l \le l_0,$$

additional channels of scattering (ACS) of the operator $H^{(\sigma)}$. Here $\mathcal{U}$ is the Gelfand transformation (1.6). Thus the functions $u \in \mathfrak{C}_l$ are precisely the functions of the form

$$u(x) = \int_{-1/2}^{1/2} f(k) \psi_l(x,k) e^{ikx_1}\, dk, \qquad x \in \mathbb{R}^2_+, \tag{6.3}$$

with $f \in L_2(\mathcal{K}_l)$ arbitrary. In particular, it follows from the form (2.4) of the eigenfunction $\psi_l(.,k)$ that functions $u \in \mathfrak{C}_l$ decay exponentially with respect to

the variable x_2 provided $\mathcal{K}_l = [-\frac{1}{2}, \frac{1}{2}]$.
Let us list some more properties of the spaces $\mathfrak{C}_l$. One has for all $1 \le l, j \le l_0$

$$\mathfrak{C}_l \perp \mathfrak{C}_j, \ \ j \neq l, \qquad \text{and} \qquad \mathfrak{C}_l \perp \mathcal{R}(W_\pm^{(\sigma)}).$$

Indeed, this follows from the fact that $\psi_l(.,k)$ is orthogonal to $\psi_j(.,k)$, $j \neq l$, and to the subspace $\mathcal{R}(W_\pm^{(\sigma)}(k))$ for all $k \in [-\frac{1}{2}, \frac{1}{2}]$. In particular, Theorem 1.2 implies that the wave operators $W_\pm^{(\sigma)}$ are not complete if there exists an ACS (i.e., $l_0 > 0$). Moreover, the spaces $\mathfrak{C}_l$ reduce the operator $H^{(\sigma)}$, and on functions $u \in \mathfrak{C}_l$ of the form (6.3) $H^{(\sigma)}$ acts by multiplying the function f with the function λ_l. Thus, the part of $H^{(\sigma)}$ on $\mathfrak{C}_l$ is unitarily equivalent to multiplication with the function λ_l on $L_2(\mathcal{K}_l)$.
Remark 1.10 of [Fr] shows that functions $u \in \mathfrak{C}_l$ correspond to states which propagate along the boundary.

6.2 Existence of ACS. Existence of gaps

It is clear from Theorem 1.1 that there are no ACS if σ is non-negative. Let us give an easy sufficient condition for the existence of ACS. It requires σ to be "negative in mean".

Proposition 6.1. *Assume that* $\hat\sigma_0 := \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} \sigma(x_1)\, dx_1 < 0$. *Then*

$$\sigma\left(H^{(\sigma)}\right) \cap (-\infty, 0) \neq \emptyset.$$

Proof. Indeed, we prove that $H^{(\sigma)}(k)$ has an eigenvalue smaller or equal to $k^2 - \frac{1}{2\pi}\hat\sigma_0^2$ for all $k \in [-\frac{1}{2}, \frac{1}{2}]$. For this we consider the trial function defined by $u(x) := e^{\hat\sigma_0 x_2/\sqrt{2\pi}}$, $x \in \Pi$, which satisfies

$$h^{(\sigma)}(k)[u] = (k^2 - \tfrac{1}{2\pi}\hat\sigma_0^2)\|u\|^2.$$

The assertion follows now from the variational principle. □

Remark 6.2. With more elaborate techniques one can show that the conclusion of Proposition 6.1 remains valid under the assumption $\hat\sigma_0 = 0$, $\sigma \not\equiv 0$.

We give now an example where the first gap of $H^{(\sigma)}$ is open, i.e. where

$$\max_{k \in [-\frac{1}{2}, \frac{1}{2}]} \lambda_1(k) < \min_{k \in [-\frac{1}{2}, \frac{1}{2}]} \lambda_2(k). \tag{6.4}$$

We start with a more general construction. Let $-\pi \le c \le \pi$ be given. For $k \in [-\frac{1}{2}, \frac{1}{2}]$ we consider the self-adjoint operators $H_D^{(\sigma)}(k)$, $H_N^{(\sigma)}(k)$ in $L_2(\Pi)$ which differ from $H^{(\sigma)}(k)$ only by Dirichlet and natural boundary conditions, respectively, at $\{x_1 \in \{-\pi, c, \pi\}\}$. More precisely, the operators $H_\nu^{(\sigma)}(k)$, $\nu \in$

$\{D, N\}$, are defined by the quadratic forms $h_\nu^{(\sigma)}(k)$ given by the same formal expression (1.5) as $h^{(\sigma)}(k)$ but with domains

$$\begin{aligned}\mathcal{D}[h_D^{(\sigma)}(k)] &:= \{u \in H^1(\Pi):\ u(.,-\pi) = u(.,c) = u(.,\pi) = 0\},\\ \mathcal{D}[h_N^{(\sigma)}(k)] &:= \{u \in L_2(\Pi):\ u|_{(-\pi,c)\times\mathbb{R}_+} \in H^1((-\pi,c)\times\mathbb{R}_+),\\ &\qquad u|_{(c,\pi)\times\mathbb{R}_+} \in H^1((c,\pi)\times\mathbb{R}_+)\}.\end{aligned}$$

It follows that

$$H_N^{(\sigma)}(k) \le H^{(\sigma)}(k) \le H_D^{(\sigma)}(k). \tag{6.5}$$

Moreover, it is easy to see that for each ν all the operators $H_\nu^{(\sigma)}(k)$, $k \in [-\frac{1}{2}, \frac{1}{2}]$, are unitarily equivalent. Their essential spectrum starts at $(\frac{\pi}{\pi+|c|})^2$ if $\nu = D$ and at 0 if $\nu = N$. We define the numbers λ_l^ν, $l \in \mathbb{N}$, as the successive infima of the variational quotient

$$\frac{h_\nu^{(\sigma)}(k)[u]}{\|u\|^2}, \qquad 0 \neq u \in \mathcal{D}[h_\nu^{(\sigma)}(k)].$$

By the variational principle the $\lambda_l^N < 0$ coincide with the discrete eigenvalues of the operator $H_N^{(\sigma)}(k)$, and similarly for $\nu = D$. It follows from (6.5) together with the variational principle that for all $l \in \mathbb{N}$

$$\lambda_l^N \le \lambda_l(k) \le \lambda_l^D, \qquad k \in [-\tfrac{1}{2}, \tfrac{1}{2}]. \tag{6.6}$$

Let us give now an example of an open gap.

Example 6.3. Let a, $b \in \mathbb{R}$ and

$$\sigma(x_1) := \begin{cases} -a & \text{if } x_1 \in [-\pi, c],\\ b & \text{if } x_1 \in (c, \pi).\end{cases}$$

We claim that under the assumptions

$$a > \tfrac{\pi}{\pi+c}, \qquad b \ge 0, \qquad -\pi < c < \pi, \tag{6.7}$$

the inequality (6.4) holds.

Indeed, one easily finds that $\lambda_1^D = \lambda_2^N = -a^2 + (\frac{\pi}{\pi+c})^2$, so because of (6.6) and the continuity of λ_1 it suffices to prove that

$$\lambda_1(k) < -a^2 + (\tfrac{\pi}{\pi+c})^2, \qquad k \in [-\tfrac{1}{2}, \tfrac{1}{2}].$$

To arrive at a contradiction we assume that we have equality for some k. Consider the eigenfunction u of $H_D^{(\sigma)}(k)$ corresponding to the eigenvalue $-a^2 + (\frac{\pi}{\pi+c})^2$,

$$u(x) := \begin{cases} 2\sqrt{\frac{a}{\pi+c}}\, e^{-ikx_1} \sin(\frac{\pi}{\pi+c}(x_1+\pi))\, e^{-ax_2}, & x_1 \in [-\pi, c],\\ 0, & x_1 \in (c, \pi).\end{cases}$$

Then $u \in \tilde{H}^1(\Pi)$, $\|u\| = 1$ and $h^{(\sigma)}(k)[u] = \inf\{h^{(\sigma)}(k)[v] :\ v \in \tilde{H}^1(\Pi),\ \|v\| = 1\}$. It follows from general principles that $u \in \mathcal{D}(H^{(\sigma)}(k))$ and $H^{(\sigma)}(k)u = \lambda_1(k)u$. By Elliptic Regularity we must have $u \in H^2_{loc}(\Pi)$, which is obviously not true. This contradiction completes the proof of (6.4).

Remark 6.4. It follows from (6.6) that the condition $\lambda_l^D < \lambda_{l+1}^N$ for some $l \in \mathbb{N}$ is sufficient for an open gap. This can be used to construct further examples.

Remark 6.5. By an argument similar to the one in Example 6.3 we find that if there exists a non-empty connected open subset Λ of the torus such that $\sigma(x_1) \le -\frac{\pi}{\operatorname{meas}\Lambda}$, $x_1 \in \Lambda$, then $\sigma\left(H^{(\sigma)}\right) \cap (-\infty, 0) \neq \emptyset$, so there exist ACS.

To conclude this subsection we note that the number of ACS (due to discrete eigenvalues) can be estimated using (4.3).

6.3 Additional Channels due to embedded eigenvalues

In general, the embedded eigenvalues of the operators $H^{(\sigma)}(k)$, $k \in [-\frac{1}{2}, \frac{1}{2}]$, also contribute to the spectrum of the operator $H^{(\sigma)}$. Therefore the subspace

$$\mathfrak{C}_* := \mathcal{R}\left(\mathcal{U}^*\left(\int_{-1/2}^{1/2} \oplus E^{(\sigma)}\left(\sigma_p\left(H^{(\sigma)}(k)\right) \cap \left[k^2, +\infty\right), k\right)\, dk\right)\mathcal{U}\right)$$

may be non-trivial and, in this case, will be called an ACS. We have

$$L_2(\mathbb{R}^2_+) = \mathcal{R}(W_\pm^{(\sigma)}) \oplus (\textstyle\sum_{l=1}^{l_0} \oplus \mathfrak{C}_l) \oplus \mathfrak{C}_*.$$

The subspace $\mathfrak{C}_*$ reduces the operator $H^{(\sigma)}$ and is orthogonal to the ACS $\mathfrak{C}_l$, $1 \le l \le l_0$, and to $\mathcal{R}(W_\pm^{(\sigma)})$.

Let us consider some examples. If $\sigma \equiv \sigma_0 < 0$ is a negative constant, we know from Example 4.2 that the embedded eigenvalues of $H^{(\sigma)}(k)$ depend piecewise analytically on k and all of them contribute to the spectrum of $H^{(\sigma)}$. We note that in this case the part of $H^{(\sigma)}$ on $\mathfrak{C}_*$ is an unbounded operator.

If σ is a trigonometric polynomial of degree $N > 0$, we know from Proposition 5.1 that $H^{(\sigma)}(k)$ has no embedded eigenvalues greater or equal to $(N - |k|)^2$. Moreover, we know from Proposition 5.2 that the embedded eigenvalues in the interval $[(N - 1 + |k|)^2, (N - |k|)^2)$ do not contribute to the spectrum of the operator $H^{(\sigma)}$. So the part of $H^{(\sigma)}$ on $\mathfrak{C}_*$ is a bounded operator with spectrum contained in $[0, (N - \frac{1}{2})^2]$.

In the special case when σ is a trigonometric polynomial of degree one, it follows again from Proposition 5.1 and Proposition 5.2 that $\mathfrak{C}_* = \{0\}$. We emphasize (see Example 5.4) that embedded eigenvalues of the operators $H^{(\sigma)}(k)$ actually occur in this case.

The question whether $\mathfrak{C}_*$ can be non-trivial for non-constant σ remains open.

References

[BeBrPa] F. Bentosela, Ph. Briet, L. Pastur, *On the spectral and wave propagation properties of the surface Maryland model*, J. Math. Phys. 44, no. 1 (2003), 1-35.

[DaSi] E. B. Davies, B. Simon, *Scattering Theory for Systems with Different Spatial Asymptotics on the Left and Right*, Commun. Math. Phys. 63 (1978), 277-301.

[FiKl] N. Filonov, F. Klopp, *Absolute continuity of Schrödinger operator with potential which is periodic in some directions and decreases in anothers*, preprint, mp_arc 04-26.

[Fr] R. L. Frank, *On the scattering theory of the Laplacian with a periodic boundary condition. I. Existence of wave operators*, Documenta Math. 8 (2003), 547-565.

[GrHoMe] A. Grossmann, R. Hoegh-Krohn, M. Mebkhout, *The one particle theory of periodic point interactions. Polymers, mono-molecular layers, and crystals*, Commun. Math. Phys. 77 (1980), 87-110.

[K] T. Kato, *Perturbation Theory for Linear Operators*, Springer, 1976.

[Ka] Yu. E. Karpeshina, *Spectrum and eigenfunctions of Schrödinger operator with the zero-range potential of the homogeneous two-dimensional lattice type in three-dimensional space*, Theor. Math. Phys. 57 (1983), 1231-1237.

[Sa] A. W. Saenz, *Quantum-mechanical scattering by impenetrable periodic surfaces*, J. Math. Phys. 22, no. 12 (1981), 2872-2884.

[Ya1] D. R. Yafaev, *Mathematical Scattering Theory*, Amer. Math. Soc., 1992.

[Ya2] D. R. Yafaev, *Scattering Theory: Some Old and New Problems*, Lecture Notes in Mathematics 1735, Springer, 2000.

Rupert L. Frank
Royal Institute of Technology
Department of Mathematics
Lindstedtsvägen 25
10044 Stockholm, Sweden
rupert@math.kth.se

Roman G. Shterenberg
St. Petersburg State University
Department of Physics
Ul'yanovskaya 1, Petrodvorets,
198504 St. Petersburg, Russia
roman@rs3759.spb.edu

Adding Tails to C^*-Correspondences

Paul S. Muhly and Mark Tomforde

Received: September 22, 2003
Revised: March 9, 2004

Communicated by Joachim Cuntz

Abstract. We describe a method of adding tails to C^*-correspondences which generalizes the process used in the study of graph C^*-algebras. We show how this technique can be used to extend results for augmented Cuntz-Pimsner algebras to C^*-algebras associated to general C^*-correspondences, and as an application we prove a gauge-invariant uniqueness theorem for these algebras. We also define a notion of relative graph C^*-algebras and show that properties of these C^*-algebras can provide insight and motivation for results about relative Cuntz-Pimsner algebras.

2000 Mathematics Subject Classification: 46L08, 46L55
Keywords and Phrases: C^*-correspondence, Cuntz-Pimsner algebra, relative Cuntz-Pimsner algebra, graph C^*-algebra, adding tails, gauge-invariant uniqueness

1 Introduction

In [18] Pimsner introduced a way to construct a C^*-algebra $\mathcal{O}_X$ from a pair (A, X), where A is a C^*-algebra and X is a C^*-correspondence (sometimes called a Hilbert bimodule) over A. Throughout his analysis Pimsner assumed that his correspondence was full and that the left action of A on X was injective. These Cuntz-Pimsner algebras have been found to compose a class of C^*-algebras that is extraordinarily rich and includes numerous C^*-algebras found in the literature: crossed products by automorphisms, crossed products by endomorphisms, partial crossed products, Cuntz-Krieger algebras, C^*-algebras of graphs with no sinks, Exel-Laca algebras, and many more. Consequently, the study of Cuntz-Pimsner algebras has received a fair amount of attention by the operator algebra community in recent years, and because information about $\mathcal{O}_X$ is very densely codified in (A, X), determining how to extract it has been the focus of much current effort.

One interesting consequence of this effort has been the introduction of the so-called *relative Cuntz-Pimsner algebras*, denoted $\mathcal{O}(K, X)$, that have Cuntz-Pimsner algebras as quotients. Very roughly speaking, a relative Cuntz-Pimsner algebra arises by relaxing some of the relations that must hold among the generators of a Cuntz-Pimsner algebra. These relations are codified in an ideal K of A. (The precise definition will be given shortly.) Relative Cuntz-Pimsner algebras arise quite naturally, particularly when trying to understand the ideal structure of a Cuntz-Pimsner algebra (See, e.g., [15, 6]). It turns out, in fact, that not only are Cuntz-Pimsner algebras quotients of relative Cuntz-Pimsner algebras, but quotients of Cuntz-Pimsner algebras are often relative Cuntz-Pimsner algebras [6, Theorem 3.1].

Although in his initial work Pimsner assumed that his C^*-correspondences were full and had injective left action, in recent years there have been efforts to remove these restrictions. Pimsner himself described how to deal with the case when X was not full, defining the so-called *augmented Cuntz-Pimsner algebras* [18, Remark 1.2(3)]. However, the case when the left action is not injective has been more elusive. In [6] it was shown that for any C^*-correspondence X and for any ideal K of A consisting of elements that act as compact operators on the left of X, one may define $\mathcal{O}(K, X)$ to be a C^*-algebra which satisfies a certain universal property [6, Proposition 1.3]. In the case that X is full with injective left action, this definition agrees with previously defined notions of relative Cuntz-Pimsner algebras, and the Cuntz-Pimsner algebra $\mathcal{O}_X$ is equal to $\mathcal{O}(J(X), X)$, where $J(X)$ denotes the ideal consisting of all elements of A which act on the left of X as compact operators.

In [6] it was proposed that for a general C^*-correspondence X, the C^*-algebra $\mathcal{O}(J(X), X)$ is the proper analogue of the Cuntz-Pimsner algebra. However, upon further analysis it seems that this is not exactly correct. To see why, consider the case of graph C^*-algebras. If $E = (E^0, E^1, r, s)$ is a graph, then there is a natural C^*-correspondence $X(E)$ over $C_0(E^0)$ associated to E (see [7, Example 1.2]). If E has no sinks, then the C^*-algebra $\mathcal{O}(J(X(E)), X(E))$ is isomorphic to the graph C^*-algebra $C^*(E)$. However, when E has sinks this will not necessarily be the case.

It is worth mentioning that graphs with sinks play an important role in the study of graph C^*-algebras. Even if one begins with a graph E containing no sinks, an analysis of $C^*(E)$ will often necessitate considering C^*-algebras of graphs with sinks. For example, quotients of $C^*(E)$ will often be isomorphic to C^*-algebras of graphs with sinks even when E has no sinks. Consequently, one needs a theory that incorporates these objects.

This deficiency in the generalization of Cuntz-Pimsner algebras was addressed by Katsura in [10] and [11]. If X is a C^*-correspondence over a C^*-algebra A with left action $\phi : A \to \mathcal{L}(X)$, then Katsura proposed that the appropriate analogue of the Cuntz-Pimsner algebra is $\mathcal{O}_X := \mathcal{O}(J_X, X)$, where

$$J_X := \{a \in J(X) : ab = 0 \text{ for all } b \in \ker \phi\}.$$

(Note that when ϕ is injective $J_X = J(X)$.) It turns out that when ϕ is injective, $\mathcal{O}_X$ is equal to the augmented Cuntz-Pimsner algebra of X, and when X is also full $\mathcal{O}_X$ coincides with the Cuntz-Pimsner algebra of X. Furthermore, if E is a graph (possibly containing sinks), then $\mathcal{O}_{X(E)}$ is isomorphic to $C^*(E)$. In addition, as with graph algebras, the class of $\mathcal{O}_X$'s is closed under quotients by gauge-invariant ideals. These facts, together with the analysis described in [11] and [12], provide strong arguments for using $\mathcal{O}_X := \mathcal{O}(J_X, X)$ as the analogue of the Cuntz-Pimsner algebra. We shall adopt this viewpoint here, and for a general C^*-correspondence X we define $\mathcal{O}_X := \mathcal{O}(J_X, X)$ to be the C^*-algebra associated to X.

In this paper we shall describe a method which will allow one to "bootstrap" many results for augmented Cuntz-Pimsner algebras to C^*-algebras associated to general correspondences. This method is inspired by a technique from the theory of graph C^*-algebras, where one can often reduce to the sinkless case by the process of "adding tails to sinks". Specifically, if E is a graph and v is a vertex of E, then by *adding a tail to* v we mean attaching a graph of the form

$$v \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \cdots$$

to E. It is well known that if F is the graph formed by adding a tail to every sink of E, then F is a graph with no sinks and $C^*(E)$ is canonically isomorphic to a full corner of $C^*(F)$. Thus in the proofs of many theorems about graph C^*-algebras, one can reduce to the case of no sinks.

In this paper we describe a generalization of this process for C^*-correspondences. More specifically, if X is a C^*-correspondence over a C^*-algebra A, then we describe how to construct a C^*-algebra B and a C^*-correspondence Y over B with the property that the left action of Y is injective and $\mathcal{O}_X$ is canonically isomorphic to a full corner of $\mathcal{O}_Y$. Thus many questions about C^*-algebras associated to correspondences can be reduced to questions about augmented Cuntz-Pimsner algebras, and many results characterizing properties of augmented Cuntz-Pimsner algebras may be easily generalized to C^*-algebras associated to general correspondences. As an application of this technique, we use it in the proof of Theorem 5.1 to extend the Gauge-Invariant Uniqueness Theorem for augmented Cuntz-Pimsner algebras to C^*-algebras of general correspondences.

This paper is organized as follows. We begin in Section 2 with some preliminaries. In Section 3 we analyze graph C^*-algebras in the context of Cuntz-Pimsner and relative Cuntz-Pimsner algebras, and describe a notion of a *relative graph C^*-algebra*. Since graph algebras provide much of the impetus for our analysis of C^*-correspondences, we examine these objects carefully in order to provide a framework which will motivate and illuminate the results of subsequent sections. In Section 4 we describe our main result — a process of "adding tails" to general C^*-correspondences. We also prove that this process preserves the Morita equivalence class of the associated C^*-algebra. In Section 5 we provide an application of our technique of "adding tails" by using it to extend the Gauge-Invariant Uniqueness Theorem for augmented Cuntz-Pimsner algebras to C^*-algebras associated to general correspondences. We also interpret this theorem in the context of relative Cuntz-Pimsner algebras, and in Section 6 we use it to classify the gauge-invariant ideals in C^*-algebras associated to certain correspondences. Finally, we conclude in Section 7 by discussing other possible applications of our technique.

The authors would like to thank Takeshi Katsura for pointing out an error in a previous draft of this paper, and for many useful conversations regarding these topics.

2 Preliminaries

For the most part we will use the notation and conventions of [6], augmenting them when necessary with the innovations of [10] and [11].

Definition 2.1. If A is a C^*-algebra, then a *right Hilbert A-module* is a Banach space X together with a right action of A on X and an A-valued inner product $\langle \cdot, \cdot \rangle_A$ satisfying

(i) $\langle \xi, \eta a \rangle_A = \langle \xi, \eta \rangle_A a$

(ii) $\langle \xi, \eta \rangle_A = \langle \eta, \xi \rangle_A^*$

(iii) $\langle \xi, \xi \rangle_A \geq 0$ and $\|\xi\| = \langle \xi, \xi \rangle_A^{1/2}$

for all $\xi, \eta \in X$ and $a \in A$. For a Hilbert A-module X we let $\mathcal{L}(X)$ denote the C^*-algebra of adjointable operators on X, and we let $\mathcal{K}(X)$ denote the closed two-sided ideal of compact operators given by

$$\mathcal{K}(X) := \overline{\operatorname{span}}\{\Theta^X_{\xi,\eta} : \xi, \eta \in X\}$$

where $\Theta^X_{\xi,\eta}$ is defined by $\Theta^X_{\xi,\eta}(\zeta) := \xi\langle \eta, \zeta \rangle_A$. When no confusion arises we shall often omit the superscript and write $\Theta_{\xi,\eta}$ in place of $\Theta^X_{\xi,\eta}$.

Definition 2.2. If A is a C^*-algebra, then a *C^*-correspondence* is a right Hilbert A-module X together with a $*$-homomorphism $\phi : A \to \mathcal{L}(X)$. We consider ϕ as giving a left action of A on X by setting $a \cdot x := \phi(a)x$.

Definition 2.3. If X is a C^*-correspondence over A, then a *representation* of X into a C^*-algebra B is a pair (π, t) consisting of a $*$-homomorphism $\pi : A \to B$ and a linear map $t : X \to B$ satisfying

(i) $t(\xi)^* t(\eta) = \pi(\langle \xi, \eta \rangle_A)$

(ii) $t(\phi(a)\xi) = \pi(a)t(\xi)$

(iii) $t(\xi a) = t(\xi)\pi(a)$

for all $\xi, \eta \in X$ and $a \in A$.

Note that Condition (iii) follows from Condition (i) due to the equation

$$\|t(\xi)\pi(a) - t(\xi a)\|^2 = \|(t(\xi)\pi(a) - t(\xi a))^*(t(\xi)\pi(a) - t(\xi a))\| = 0.$$

If (π, t) is a representation of X into a C^*-algebra B, we let $C^*(\pi, t)$ denote the C^*-subalgebra of B generated by $\pi(A) \cup t(X)$.
A representation (π, t) is said to be *injective* if π is injective. Note that in this case t will also be isometric since

$$\|t(\xi)\|^2 = \|t(\xi)^* t(\xi)\| = \|\pi(\langle \xi, \xi \rangle_A)\| = \|\langle \xi, \xi \rangle_A\| = \|\xi\|^2.$$

When (π, t) is a representation of X into $\mathcal{B}(\mathcal{H})$ for a Hilbert space $\mathcal{H}$, we say that (π, t) is *a representation of X on $\mathcal{H}$*.

In the literature a representation (π, t) is sometimes referred to as a *Toeplitz representation* (See, e.g., [7] and [6].), and as an *isometric representation* [15]. However, here, all representations considered will be at least Toeplitz or isometric and so we drop the additional adjective. We note that in [7] the authors show that given a correspondence X over a C^*-algebra A, there is a C^*-algebra, denoted $\mathcal{T}_X$ and a representation (π_X, t_X) of X in $\mathcal{T}_X$ that is universal in the following sense: $\mathcal{T}_X$ is generated as a C^*-algebra by the ranges of π_X and t_X, and given any representation (π, t) in a C^*-algebra B, then there is a C^*-homomorphism of $\mathcal{T}_X$ into B, denoted $\rho_{(\pi,t)}$, that is unique up to an inner automorphism of B, such that $\pi = \rho_{(\pi,t)} \circ \pi_X$ and $t = \rho_{(\pi,t)} \circ t_X$. The C^*-algebra $\mathcal{T}_X$ and the representation (π_X, t_X) are unique up to an obvious notion of isomorphism. We call $\mathcal{T}_X$ *the* Toeplitz algebra of the correspondence X, but we call (π_X, t_X) *a universal representation* of X in $\mathcal{T}_X$, with emphasis on the indefinite article, because at times we want to consider more than one.

Definition 2.4. For a representation (π, t) of a C^*-correspondence X on B there exists a $*$-homomorphism $\pi^{(1)} : \mathcal{K}(X) \to B$ with the property that

$$\pi^{(1)}(\Theta_{\xi,\eta}) = t(\xi)t(\eta)^*.$$

See [18, p. 202], [9, Lemma 2.2], and [7, Remark 1.7] for details on the existence of this $*$-homomorphism. Also note that if (π, t) is an injective representation, then $\pi^{(1)}$ will be injective as well.

Definition 2.5. For an ideal I in a C^*-algebra A we define

$$I^\perp := \{a \in A : ab = 0 \text{ for all } b \in I\}.$$

If X is a C^*-correspondence over A, we define an ideal $J(X)$ of A by $J(X) := \phi^{-1}(\mathcal{K}(X))$. We also define an ideal J_X of A by

$$J_X := J(X) \cap (\ker \phi)^\perp.$$

Note that $J_X = J(X)$ when ϕ is injective, and that J_X is the maximal ideal on which the restriction of ϕ is an injection into $\mathcal{K}(X)$.

Definition 2.6. If X is a C^*-correspondence over A and K is an ideal in $J(X)$, then we say that a representation (π, t) is *coisometric on K*, or is *K-coisometric* if

$$\pi^{(1)}(\phi(a)) = \pi(a) \qquad \text{for all } a \in K.$$

In [6, Proposition 1.3] the authors show that given a correspondence X over a C^*-algebra A, and an ideal K of A contained in $J(X)$, there is a C^*-algebra, denoted $\mathcal{O}(K,X)$, and a representation (π_X, t_X) of X in $\mathcal{O}(K,X)$ that is coisometric on K and is universal with this property, in the following sense: $\mathcal{O}(K,X)$ is generated as a C^*-algebra by the ranges of π_X and t_X, and given any representation (π,t) of X in a C^*-algebra B that is K-coisometric, then there is a C^*-homomorphism of $\mathcal{O}(K,X)$ into B, denoted $\rho_{(\pi,t)}$, that is unique up to an inner automorphism of B, such that $\pi = \rho_{(\pi,t)} \circ \pi_X$ and $t = \rho_{(\pi,t)} \circ t_X$.

Definition 2.7. The algebra $\mathcal{O}(K,X)$, associated with an ideal K in $J(X)$, is called the *relative Cuntz-Pimsner algebra* determined by X and the ideal K. Further, a representation (π_X, t_X) that is coisometric on K and has the universal property just described is called a *universal* K-coisometric representation of X.

Remark 2.8. When the ideal K is the zero ideal in $J(X)$, then the algebra $\mathcal{O}(K,X)$ becomes $\mathcal{T}_X$ and a universal 0-coisometric representation of X is simply a representation of X. Furthermore, if X is a C^*-correspondence in which ϕ is injective, then $\mathcal{O}_X := \mathcal{O}(J_X, X)$ is precisely the augmented Cuntz-Pimsner algebra of X defined in [18]. If X is full, i.e., if $\overline{\operatorname{span}}\{\langle \xi, \eta \rangle_A : \xi, \eta \in X\} = A$, then the augmented Cuntz-Pimsner algebra of X and the Cuntz-Pimsner algebra of X coincide. Thus $\mathcal{O}_X$ coincides with the Cuntz-Pimsner algebra of [18] when ϕ is injective and X is full. Whether or not ϕ is injective, a universal $J(X)$-coisometric representation is sometimes called a universal Cuntz-Pimsner covariant representation [6, Definition 1.1].

Remark 2.9. If $\mathcal{O}(K,X)$ is a relative Cuntz-Pimsner algebra associated to a C^*-correspondence X, and if (π,t) is a universal K-coisometric representation of X, then for any $z \in \mathbb{T}$ (π, zt) is also a universal K-coisometric representation. Hence by the universal property, there exists a homomorphism $\gamma_z : \mathcal{O}(K,X) \to \mathcal{O}(K,X)$ such that $\gamma_z(\pi(a)) = \pi(a)$ for all $a \in A$ and $\gamma_z(t(\xi)) = zt(\xi)$ for all $\xi \in X$. Since $\gamma_{z^{-1}}$ is an inverse for this homomorphism, we see that γ_z is an automorphism. Thus we have an action $\gamma : \mathbb{T} \to \operatorname{Aut} \mathcal{O}(K,X)$ with the property that $\gamma_z(\pi(a)) = \pi(a)$ and $\gamma_z(t(\xi)) = zt(\xi)$. Furthermore, a routine $\epsilon/3$ argument shows that γ is strongly continuous. We call γ the *gauge action on* $\mathcal{O}(K,X)$.

3 Viewing graph C^*-algebras as Cuntz-Pimsner algebras

Let $E := (E^0, E^1, r, s)$ be a directed graph with countable vertex set E^0, countable edge set E^1, and range and source maps $r, s : E^1 \to E^0$. A *Cuntz-Krieger E-family* is a collection of partial isometries $\{s_e : e \in E^1\}$ with commuting range projections together with a collection of mutually orthogonal projections $\{p_v : v \in E^0\}$ that satisfy

1. $s_e^* s_e = p_{r(e)}$ for all $e \in E^1$
2. $s_e s_e^* \leq p_{s(e)}$ for all $e \in E^1$

3. $p_v = \sum_{\{e:s(e)=v\}} s_e s_e^*$ for all $v \in E^0$ with $0 < |s^{-1}(v)| < \infty$

The *graph algebra* $C^*(E)$ is the C^*-algebra generated by a universal Cuntz-Krieger E-family (see [14, 13, 2, 5, 1]).

Example 3.1 *(The Graph C^*-correspondence).* If $E = (E^0, E^1, r, s)$ is a graph, we define $A := C_0(E^0)$ and

$$X(E) := \{x : E^1 \to \mathbb{C} : \text{ the function } v \mapsto \sum_{\{f \in E^1 : r(f)=v\}} |x(f)|^2 \text{ is in } C_0(E^0)\ \}.$$

Then $X(E)$ is a C^*-correspondence over A with the operations

$$\begin{aligned}(x \cdot a)(f) &:= x(f)a(r(f)) \text{ for } f \in E^1 \\ \langle x, y\rangle_A(v) &:= \sum_{\{f \in E^1 : r(f)=v\}} \overline{x(f)}y(f) \text{ for } f \in E^1 \\ (a \cdot x)(f) &:= a(s(f))x(f) \text{ for } f \in E^1\end{aligned}$$

and we call $X(E)$ the *graph C^*-correspondence* associated to E. Note that we could write $X(E) = \bigoplus^0_{v \in E^0} \ell^2(r^{-1}(v))$ where this denotes the C_0 direct sum (sometimes called the restricted sum) of the $\ell^2(r^{-1}(v))$'s. Also note that $X(E)$ and A are spanned by the point masses $\{\delta_f : f \in E^1\}$ and $\{\delta_v : v \in E^0\}$, respectively.

Theorem 3.2 ([5, Proposition 12]). *If E is a graph with no sinks, and $X(E)$ is the associated graph C^*-correspondence, then $\mathcal{O}(J(X(E)), X(E)) \cong C^*(E)$. Furthermore, if (π_X, t_X) is a universal $J(X(E))$-coisometric representation, then $\{t_X(\delta_e), \pi_X(\delta_v)\}$ is a universal Cuntz-Krieger E-family in $\mathcal{O}(J(X(E)), X(E))$.*

It was shown in [7, Proposition 4.4] that

$$J(X(E)) = \overline{\operatorname{span}}\{\delta_v : |s^{-1}(v)| < \infty\}$$

and if v emits finitely many edges, then

$$\phi(\delta_v) = \sum_{\{f \in E^1 : s(f)=v\}} \Theta_{\delta_f, \delta_f} \quad \text{and} \quad \pi_X(\phi(\delta_v)) = \sum_{\{f \in E^1 : s(f)=v\}} t_X(\delta_f)t_X(\delta_f)^*.$$

Furthermore, one can see that $\delta_v \in \ker \phi$ if and only if v is a sink in E. Also $\delta_v \in \overline{\operatorname{span}}\{\langle x, y\rangle_A\}$ if and only if v is a source, and since $\delta_{s(f)} \cdot \delta_f = \delta_f$ we see that $\overline{\operatorname{span}}\, A \cdot X = X$ and $X(E)$ is essential. These observations show that we have the following correspondences between the properties of the graph E and the properties of the graph C^*-correspondence $X(E)$.

Property of $X(E)$	Property of E
$\phi(\delta_v) \in \mathcal{K}(X(E))$	v emits a finite number of edges
$\phi(A) \subseteq \mathcal{K}(X(E))$	E is row-finite
ϕ is injective	E has no sinks
$X(E)$ is full	E has no sources
$X(E)$ is essential	always

Remark 3.3. If E is a graph with no sinks, then $\mathcal{O}(J(X(E)), X(E))$ is canonically isomorphic to $C^*(E)$. When E has sinks, this will not be the case. If (π, t) is the universal $J(X(E))$-coisometric representation of $X(E)$, then it will be the case that $\{t(\delta_e), \pi(\delta_v)\}$ is a Cuntz-Krieger E-family. However, when v is a sink in E, $\phi(\delta_v) = 0$ and thus $\pi(\delta_v) = \pi^{(1)}(\phi(\delta_v)) = 0$. Consequently, $\{t(\delta_e), \pi(\delta_v)\}$ will not be a universal Cuntz-Krieger E-family when E has sinks. However, if E is a graph with sinks, then we see that $\phi(\delta_v) = 0$ if and only if v is a sink, and $\delta_v \in (\ker \phi)^\perp$ if and only if v is not a sink. Thus

$$J_{X(E)} = \overline{\operatorname{span}}\{\delta_v : 0 < |s^{-1}(v)| < \infty\}$$

and a proof similar to that in [7, Proposition 4.4] shows that $\mathcal{O}_{X(E)} := \mathcal{O}(J_{X(E)}, X(E))$ is isomorphic to $C^*(E)$. Furthermore, if (π_X, t_X) is a universal $J(X(E))$-coisometric representation of $X(E)$, then $\{t_X(\delta_e), \pi_X(\delta_v)\}$ is a universal Cuntz-Krieger E-family in $\mathcal{O}_{X(E)}$.

3.1 Relative Graph Algebras

We shall now examine relative Cuntz-Pimsner algebras in the context of graph algebras. If E is a graph and $X(E)$ is the associated graph C^*-correspondence, then $J_{X(E)} := \overline{\operatorname{span}}\{\delta_v : 0 < |s^{-1}(v)| < \infty\}$. If K is an ideal in $J_{X(E)}$, then $K = \overline{\operatorname{span}}\{\delta_v : v \in V\}$ for some subset V of vertices which emit a finite and nonzero number of edges. If $(\mathcal{O}(K, X(E)), t_X, \pi_X)$ is the relative Cuntz-Pimsner algebra determined by K, then the relation $\pi_X(\delta_v) = \sum_{s(e)=v} t_X(\delta_e) t_X(\delta_e)^*$ will hold only for vertices $v \in V$. This motivates the following definition.

Definition 3.4. Let $E = (E^0, E^1, r, s)$ be a graph and define $R(E) := \{v \in E^0 : 0 < |s^{-1}(v)| < \infty\}$. For any $V \subseteq R(E)$ we define a *Cuntz-Krieger* (E, V)-*family* to be a collection of mutually orthogonal projections $\{p_v : v \in E^0\}$ together with a collection of partial isometries $\{s_e : e \in E^1\}$ that satisfy

1. $s_e^* s_e = p_{r(e)}$ for $e \in E^1$
2. $s_e s_e^* < p_{s(e)}$ for $e \in E^1$
3. $p_v = \sum_{s(e)=v} s_e s_e^*$ for all $v \in V$

We refer to a Cuntz-Krieger $(E, R(E))$-family as simply a Cuntz-Krieger E-family, and we refer to a Cuntz-Krieger $(E, \emptyset)$-family as a Toeplitz-Cuntz-Krieger family.

Definition 3.5. If E is a graph and $V \subseteq R(E)$, then we define the *relative graph algebra* $C^*(E, V)$ to be the C^*-algebra generated by a universal Cuntz-Krieger (E, V)-family.

The existence of $C^*(E, V)$ can be proven by adapting the argument for the existence of graph algebras in [13], or by realizing $C^*(E, V)$ as a relative Cuntz-Pimsner algebra.

Note that $C^*(E, R(E))$ is the graph algebra $C^*(E)$, and $C^*(E, \emptyset)$ is the Toeplitz algebra defined in [7, Theorem 4.1] (but different from the Toeplitz algebra defined in [4]). It is also the case that if $\{s_e, p_v\}$ is a universal Cuntz-Krieger (E, V)-family, then whenever $v \in R(E) \backslash V$ we have $p_v > \sum_{s(e)=v} s_e s_e^*$.

Definition 3.6. Let $E = (E^0, E^1, r, s)$ be a graph and $V \subseteq R(E)$. We define the graph E_V to be the graph with vertex set $E_V^0 := E^0 \cup \{v' : v \in R(E) \backslash V\}$, edge set $E^1 \cup \{e' : e \in E^1 \text{ and } r(e) \in R(E) \backslash V\}$, and r and s extended to E_V^1 by defining $s(e') := s(e)$ and $r(e') := r(e)'$.

Roughly speaking, when forming E_V one takes E and adds a sink for each element $v \in R(E) \backslash V$ as well as edges to this sink from each vertex that feeds into v.

Theorem 3.7. *If E is a graph and $V \subseteq R(E)$, then the relative graph algebra $C^*(E, V)$ is canonically isomorphic to the graph algebra $C^*(E_V)$.*

Proof. Let $\{s_e, p_v : e \in E^1, v \in E^0\}$ be a generating Cuntz-Krieger (E, V)-family in $C^*(E, V)$. For $w \in E_V^0$ and $f \in E_V^1$ define

$$q_w := \begin{cases} p_v & \text{if } w \notin R(E) \backslash V \\ \sum_{\{e \in E^1 : s(e) = w\}} s_e s_e^* & \text{if } w \in R(E) \backslash V \\ p_v - \sum_{\{e \in E^1 : s(e) = v\}} s_e s_e^* & \text{if } w = v' \text{ for some } v \in R(E) \backslash V. \end{cases}$$

$$t_f := \begin{cases} s_f q_{r(f)} & \text{if } f \in E^1 \\ s_e q_{r(e)'} & \text{if } f = e' \text{ for some } e \in E^1. \end{cases}$$

It is straightforward to check that $\{t_f, q_w : f \in E_V^1, w \in E_V^0\}$ is a Cuntz-Krieger E_V-family in $C^*(E, V)$. Thus by the universal property there exists a homomorphism $\alpha : C^*(E_V) \to C^*(E, V)$ taking the generators of $C^*(E_V)$ to $\{t_f, q_w\}$. By the gauge-invariant uniqueness theorem [1, Theorem 2.1] α is injective. Furthermore, whenever $v \in R(E) \backslash V$ we see that $p_v = q_v + q_{v'}$ and whenever $r(e) \in R(E) \backslash V$ we see that $s_e = t_e + t_{e'}$. Thus $\{q_w, t_f\}$ generates $C^*(E, V)$ and α is surjective. Consequently α is an isomorphism. □

This theorem shows that the class of relative graph algebras is the same as the class of graph algebras. Thus we gain no new C^*-algebras by considering relative graph algebras in place of graph algebras. However, we maintain that relative graph algebras are still useful and arise naturally in the study of graph algebras. In particular, we give three examples of common situations in which relative graph algebras prove convenient.

Example 3.8 *(Subalgebras of Graph Algebras).* Let $E = (E^0, E^1, r, s)$ be a graph and let $\{s_e, p_v : e \in E^1, v \in E^0\}$ be a generating Cuntz-Krieger E-family in $C^*(E)$. If $F = (F^0, F^1, r_F, s_F)$ is a subgraph of E, and A denotes the C^*-subalgebra of $C^*(E)$ generated by $\{s_e, p_v : e \in F^1, v \in F^0\}$, then it is well-known that A is a graph algebra (but not necessarily the C^*-algebra associated to F). In fact, we see that for any $v \in F^0$, the sum $\sum_{\{e \in F^1 : s_F(e) = v\}} s_e s_e^*$ may

not add up to p_v because some of the edges in $s^{-1}(v)$ may not be in F. However, if we let $V := \{v \in R(F) : s_F^{-1}(v) = s^{-1}(v)\}$. Then $\{s_e, p_v : e \in F^1, v \in F^0\}$ is a Cuntz-Krieger (F, V)-family and $A \cong C^*(F, V)$.
These subalgebras arise often in the study of graph algebras. In [8, Lemma 2.4] they were realized as graph algebras by the method shown in the proof of Theorem 3.7, and in [19, Lemma 1.2] these subalgebras were realized as graph algebras by using the notion of a dual graph. In both of these instances it would have been convenient to have used relative graph algebras. Realizing the subalgebra as $C^*(F, V)$ would have provided an economy of notation as well as a more direct analysis of the subalgebras under consideration.

Example 3.9 *(Spielberg's Toeplitz Graph Algebras).* In [21] Spielberg introduced a notion of a Toeplitz graph groupoid and a Toeplitz graph algebra. The Toeplitz graph algebras defined in [21, Definition 2.17] are relative graph algebras as defined in Definition 3.5 (see [21, Theorem 2.9]). Spielberg also made use of his Toeplitz graph algebras in [22] to construct graph algebras with a specified K-theory.

Example 3.10 *(Quotients of Graph Algebras).* If $E = (E^0, E^1, r, s)$ is a row-finite graph and H is a saturated hereditary subset of vertices of E, then it follows from [2, Theorem 4.1(b)] that $C^*(E)/I_H \cong C^*(F)$ where F is the subgraph defined by

$$F^0 := E^0 \backslash H \qquad\qquad F^1 := \{e \in E^1 : r(e) \notin H\}.$$

If E is not row-finite, then this is not necessarily the case. The obstruction is due to the vertices in the set

$$B_H = \{v \in E^0 \mid v \text{ is an infinite emitter and } 0 < |s^{-1}(v) \cap r^{-1}(E^0 \setminus H)| < \infty\}.$$

In fact, if $\{s_e, p_v\}$ is a generating Cuntz-Krieger E-family in $C^*(E)$, then the cosets $\{s_e + I_H, p_v + I_H : v \notin H, r(e) \notin H\}$ will have the property that $p_v + I_H \geq \sum_{e \in E \backslash H : s(e) = v\}} (s_e + I_H)(s_e + I_H)^*$ with equality occurring if and only if $v \in R(F) \backslash B_H$. Thus it turns out that $\{s_e + I_H, p_v + I_H : v \notin H, r(e) \notin H\}$ will be a Cuntz-Krieger $(F, R(F) \backslash B_H)$-family and $C^*(E)/I_H \cong C^*(F, R(F) \backslash B_H)$. The quotient $C^*(E)/I_H$ was realized as a graph algebra in [1, Proposition 3.4] by a technique similar to that used in the proof of Theorem 3.7. However, relative graph algebras provide a more natural context for describing these quotients.

In addition to their applications in the situations mentioned above, relative graph algebras can be useful for another reason. Since any relative graph algebra is canonically isomorphic to a graph algebra, we see that for every theorem about graph algebras there will be a corresponding theorem for relative graph algebras. Thus the relative graph algebras provide a class of relative Cuntz-Pimsner algebras that are well understood. With this in mind, we shall now state a version of the Gauge-Invariant Uniqueness Theorem for relative graph algebras.

Theorem 3.11 (Gauge-Invariant Uniqueness for Relative Graph Algebras). *Let $E = (E^0, E^1, r, s)$ be a graph and $V \subseteq R(E)$. Also let $\{s_e, p_v : e \in E^1, v \in E^0\}$ and let $\gamma : \mathbb{T} \to \operatorname{Aut} C^*(E, V)$ denote the gauge action on $C^*(E, V)$. If $\rho : C^*(E, V) \to A$ is a $*$-homomorphism between C^*-algebras that satisfies*

1. *$\rho(p_v) \neq 0$ for all $v \in E^0$*
2. *$\rho(p_v - \sum_{s(e)=v} s_e s_e^*) \neq 0$ for all $v \in V$*
3. *there exists a strongly continuous action $\beta : \mathbb{T} \to \operatorname{Aut} A$ such that $\beta_z \circ \rho = \rho \circ \gamma_z$ for all $z \in \mathbb{T}$.*

then ρ is injective.

Proof. By Theorem 3.7 there exists an isomorphism $\alpha : C^*(E_V) \to C^*(E, V)$ and a generating Cuntz-Krieger E_V-family $\{t_e, q_w\}$ for which

$$\alpha(q_w) := \begin{cases} p_v & \text{if } w \notin R(E)\backslash V \\ \sum_{\{e \in E^1 : s(e)=w\}} s_e s_e^* & \text{if } w \in R(E)\backslash V \\ p_v - \sum_{\{e \in E^1 : s(e)=v\}} s_e s_e^* & \text{if } w = v' \text{ for some } v \in R(E)\backslash V. \end{cases}$$

$$\alpha(t_f) := \begin{cases} s_f q_{r(f)} & \text{if } f \in E^1 \\ s_e q_{r(e)'} & \text{if } f = e' \text{ for some } e \in E^1. \end{cases}$$

To show that ρ is injective, it suffices to show that $\rho \circ \alpha$ is injective. We shall do this by applying the gauge-invariant uniqueness theorem for graph algebras [1, Theorem 2.1] to $\rho \circ \alpha$. Now clearly if $w \notin R(E)\backslash V$, then $\rho \circ \alpha(q_w) \neq 0$ by (1). If $w = v'$, then $\rho \circ \alpha(q_w) \neq 0$ by (2). Furthermore, if $w \in R(E)\backslash V$ then $\rho \circ \alpha(q_w) = 0$ implies that $\rho(\sum_{s(e)=w} s_e s_e^*) = 0$ and thus for any $f \in s^{-1}(v)$ we have

$$\rho(s_f) = \rho(\sum_{s(e)=w} s_e s_e^*)\rho(s_f) = 0.$$

But then $\rho(p_{r(f)}) = \rho(s_f^* s_f) = 0$ which contradicts (1). Hence we must have $\rho \circ \alpha(q_w) \neq 0$. Finally, if γ' denotes the gauge action on $C^*(E_V)$, then by checking on generators we see that $\beta_z \circ (\rho \circ \alpha) = (\rho \circ \alpha) \circ \gamma'_z$. Therefore, $\rho \circ \alpha$ is injective by the gauge invariant uniqueness theorem for graph algebras, and consequently ρ is injective. $\square$

We have shown in Theorem 3.7 that every relative graph algebra is isomorphic to a graph algebra. More generally, Katsura has shown in [12] that every relative Cuntz-Pimsner algebra is isomorphic to the C^*-algebra associated to a correspondence; that is, if $\mathcal{O}(K, X)$ is a relative Cuntz-Pimsner algebra, then there exists a C^*-correspondence X' such that $\mathcal{O}_{X'} := \mathcal{O}(J_{X'}, X')$ is isomorphic to $\mathcal{O}(K, X)$. In Theorem 5.1 we shall prove a gauge-invariant uniqueness theorem for C^*-algebras associated to correspondences. Afterwards, in Remark 5.3, we shall use Katsura's analysis in [12] to give an interpretation of Theorem 3.11 in the context of relative Cuntz-Pimsner algebras.

4 Adding Tails to C^*-correspondences

If E is a graph and v is a vertex of E, then by *adding a tail to* v we mean attaching a graph of the form

$$v \xrightarrow{e_1} v_1 \xrightarrow{e_2} v_2 \xrightarrow{e_3} v_3 \xrightarrow{e_4} \cdots$$

to E. It was shown in [2, §1] that if F is the graph formed by adding a tail to every sink of E, then F is a graph with no sinks and $C^*(E)$ is canonically isomorphic to a full corner of $C^*(F)$. The technique of adding tails to sinks is a simple but powerful tool in the analysis of graph algebras. In the proofs of many results it allows one to reduce to the case in which the graph has no sinks and thereby avoid certain complications and technicalities.

Our goal in this section is to develop a process of "adding tails to sinks" for C^*-correspondences, so that given any C^*-correspondence X we may form a C^*-correspondence Y with the property that the left action of Y is injective and $\mathcal{O}_X$ is canonically isomorphic to a full corner in $\mathcal{O}_Y$.

Definition 4.1. Let X be a C^*-correspondence over A with left action $\phi : A \to \mathcal{L}(X)$, and let I be an ideal in A. We define *the tail determined by* I to be the C^*-algebra

$$T := I^{(\mathbb{N})}$$

where $I^{(\mathbb{N})}$ denotes the c_0-direct sum of countably many copies of the ideal I. We shall denote the elements of T by

$$\vec{f} := (f_1, f_2, f_3, \ldots)$$

where each f_i is an element of I. We shall consider T as a right Hilbert C^*-module over itself (see [20, Example 2.10]). We define $Y := X \oplus T$ and $B := A \oplus T$. Then Y is a right Hilbert B-module in the usual way; that is, the right action is given by

$$(\xi, \vec{f}) \cdot (a, \vec{g}) := (\xi \cdot a, \vec{f}\vec{g}) \qquad \text{for } \xi \in X,\ a \in A, \text{ and } \vec{f}, \vec{g} \in T$$

and the inner product is given by

$$\langle (\xi, \vec{f}), (\nu, \vec{g}) \rangle_B := (\langle \xi, \nu \rangle_A, \vec{f}^* \vec{g}) \qquad \text{for } \xi, \nu \in X \text{ and } \vec{f}, \vec{g} \in T.$$

Furthermore, we shall make Y into a C^*-correspondence over B by defining a left action $\phi_B : B \to \mathcal{L}(Y)$ as

$$\phi_B(a, \vec{f})(\xi, \vec{g}) := (\phi(a)(\xi), (ag_1, f_1 g_2, f_2 g_3, \ldots)) \text{ for } a \in A,\ \xi \in X, \text{ and } \vec{f}, \vec{g} \in T.$$

We call Y *the C^*-correspondence formed by adding the tail T to X.*

Lemma 4.2. *Let X be a C^*-correspondence over A, and let $T := (\ker \phi)^{(\mathbb{N})}$ be the tail determined by* $\ker \phi$. *If $Y := X \oplus T$ is the C^*-correspondence over $B := A \oplus T$ formed by adding the tail T to X, then the left action $\phi_B : B \to \mathcal{L}(Y)$ is injective. Consequently, $J_Y = J(Y)$ and $\mathcal{O}_Y = \mathcal{O}(J(Y), Y)$ is equal to the C^*-algebra defined by Pimsner in [18].*

Proof. If $(a, \vec{f}) \in \ker \phi_B$, then for all $\xi \in X$ we have

$$(\phi(a)\xi, \vec{0}) = \phi_B(a, \vec{f})(\xi, \vec{0}) = (0, \vec{0})$$

so that $\phi(a)\xi = 0$ and $a \in \ker \phi$. Thus $(0, (a, f_1, f_2, \ldots)) \in X \oplus T$ and

$$(0, (aa^*, f_1 f_1^*, f_2 f_2^*, \ldots)) = \phi_B(a, \vec{f})(0, (a^*, f_1, f_2, \ldots)) = (0, \vec{0})$$

so that $\|a\|^2 = \|aa^*\| = 0$ and $\|f_i\|^2 = \|f_i f_i^*\| = 0$ for all $i \in \mathbb{N}$. Consequently, $a = 0$ and $\vec{f} = \vec{0}$ so that ϕ_B is injective. □

Theorem 4.3. *Let X be a C^*-correspondence over A, and let $T := (\ker \phi)^{(\mathbb{N})}$ be the tail determined by $\ker \phi$. Also let $Y := X \oplus T$ be the C^*-correspondence over $B := A \oplus T$ formed by adding the tail T to X.*

(a) If (π, t) is a J_X-coisometric representation of X on a Hilbert space $\mathcal{H}_X$, then there is a Hilbert space $\mathcal{H}_Y = \mathcal{H}_X \oplus \mathcal{H}_T$ and a $J(Y)$-coisometric representation $(\tilde{\pi}, \tilde{t})$ of Y on $\mathcal{H}_Y$ with the property that $\tilde{\pi}|_X = \pi$ and $\tilde{t}|_A = t$.

(b) If $(\tilde{\pi}, \tilde{t})$ is a $J(Y)$-coisometric representation of Y into a C^*-algebra C, then $(\tilde{\pi}|_A, \tilde{t}|_X)$ is a J_X-coisometric representation of X into C. Furthermore, if $\tilde{\pi}|_A$ is injective, then $\tilde{\pi}$ is injective.

(c) Let (π_Y, t_Y) be a universal $J(Y)$-coisometric representation of Y. Then $(\pi, t) := (\pi_Y|_A, t_Y|_X)$ is a J_X-coisometric representation of X in $C^*(\pi_Y, t_Y)$. Furthermore, $\rho_{(\pi,t)} : \mathcal{O}_X \to C^*(\pi_X, t_X) \subseteq \mathcal{O}_Y$ is an isomorphism onto the C^*-subalgebra of $\mathcal{O}_Y$ generated by

$$\{\pi_Y(a, \vec{0}), t_Y(\xi, \vec{0}) : a \in A \text{ and } \xi \in X\}$$

and this C^*-subalgebra is a full corner of $\mathcal{O}_Y$. Consequently, $\mathcal{O}_X$ is naturally isomorphic to a full corner of $\mathcal{O}_Y$.

Corollary 4.4. *If X is a C^*-correspondence and (π_X, t_X) is a universal $J(X)$-coisometric representation of X, then (π_X, t_X) is injective.*

Proof. By the theorem (π_X, t_X) extends to a universal $J(Y)$-coisometric representation (π_Y, t_Y) of Y. Since ϕ_B is injective by Lemma 4.2 it follows from [6, Corollary 6.2] that (π_Y, t_Y) is injective. Consequently, $\pi_X = \pi_Y|_A$ is injective. □

To prove this theorem we shall need a number of lemmas.

Lemma 4.5. *Let X be a C^*-correspondence and let $T := (\ker \phi)^{(\mathbb{N})}$ be the tail determined by $\ker \phi$. Also let $Y := X \oplus T$ be the C^*-correspondence over $B := A \oplus T$ formed by adding the tail T to X. Then for any $(a, \vec{f}) \in Y$ we have that $(a, \vec{f}) \in J(Y)$ if and only if $a = a_1 + a_2$ with $a_1 \in J_X$ and $a_2 \in \ker \phi$.*

Proof. Suppose $a = a_1 + a_2$ with $a_1 \in J_X$ and $a_2 \in \ker \phi$. Then we may write $\phi(a_1) = \lim_n \sum_{k=1}^{N_n} \Theta^X_{\xi_{n,k},\eta_{n,k}}$ for some $\xi_{n,k}, \eta_{n,k} \in X$. But then

$$\phi_B(a_1, \vec{0}) = \lim_n \sum_{k=1}^{N_n} \Theta^Y_{(\xi_{n,k},\vec{0}),(\eta_{n,k},\vec{0})} \in \mathcal{K}(Y).$$

In addition, since $a_2 \in \ker \phi$ we see that if we let $\{\vec{e}_\lambda\}_{\lambda \in \Lambda}$ be an approximate unit for T with $\vec{e}_\lambda = (e^1_\lambda, e^2_\lambda, \ldots)$ for each λ, then

$$\phi_B(a_2, \vec{f}) = \lim_\lambda \Theta^Y_{(0,(a_2,f_1,f_2,\ldots)),(0,(e^1_\lambda,e^2_\lambda,\ldots))} \in \mathcal{K}(Y).$$

Thus $\phi_B(a, \vec{f}) = \phi_B(a_1, \vec{0}) + \phi_B(a_2, \vec{f})) \in \mathcal{K}(Y)$.

Conversely, suppose that $\phi_B(a, \vec{f}) \in \mathcal{K}(Y)$. Then we may write

$$\phi_B(a, \vec{f}) = \lim_n \sum_{k=1}^{N_n} \Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})}.$$

If we write $\vec{f}_{n,k} = (f^1_{n,k}, f^2_{n,k}, \ldots)$ and $\vec{g}_{n,k} = (g^1_{n,k}, g^2_{n,k}, \ldots)$ then for any $(\xi, \vec{g}) \in X \oplus T$ we have that

$$\begin{aligned}(\phi(a)\xi, (ag_1, f_1 g_2, \ldots)) &= \phi_B(a, \vec{f})(\xi, \vec{g}) \\ &= \lim_n \sum_{k=1}^{N_n} \Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})}(\xi, \vec{g}) \\ &= \lim_n \sum_{k=1}^{N_n} ((\xi_{n,k}\langle \eta_{n,k}, \xi\rangle_A, (f^1_{n,k} {g^1_{n,k}}^* g_1, f^2_{n,k} {g^2_{n,k}}^* g_2, \ldots)) \\ &= \lim_n \sum_{k=1}^{N_n} ((\Theta^X_{\xi_{n,k},\eta_{n,k}} \xi, (f^1_{n,k} {g^1_{n,k}}^* g_1, f^2_{n,k} {g^2_{n,k}}^* g_2, \ldots)).\end{aligned} \tag{1}$$

Now since the operator norm on $\mathcal{L}(Y)$ dominates the operator norm on $\mathcal{L}(X)$, we see that $\lim_n \sum_{k=1}^{N_n} \Theta^X_{\xi_{n,k},\eta_{n,k}}$ converges and $\phi(a) = \lim_n \sum_{k=1}^{N_n} \Theta^X_{\xi_{n,k},\eta_{n,k}}$. Thus $a \in \mathcal{K}(X)$.

Furthermore, if $\{e_\lambda\}_{\lambda \in \Lambda}$ is an approximate unit for $\ker \phi$, then for any $n, m \in \mathbb{N}$

we have

$$\begin{aligned}
&\Big\|\big(\sum_{k=1}^{N_n} f^1_{n,k}{g^1_{n,k}}^* - \sum_{k=1}^{N_m} f^1_{m,k}{g^1_{m,k}}^*\big)e_\lambda\Big\| \\
&= \Big\|\big(\sum_{k=1}^{N_n}\Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})} - \sum_{k=1}^{N_m}\Theta^Y_{(\xi_{m,k},\vec{f}_{m,k}),(\eta_{m,k},\vec{g}_{m,k})}\big)(0,(e_\lambda,0,0,\ldots))\Big\| \\
&\leq \Big\|\sum_{k=1}^{N_n}\Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})} - \sum_{k=1}^{N_m}\Theta^Y_{(\xi_{m,k},\vec{f}_{m,k}),(\eta_{m,k},\vec{g}_{m,k})}\Big\|\ \Big\|(0,(e_\lambda,0,0,\ldots))\Big\| \\
&= \Big\|\sum_{k=1}^{N_n}\Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})} - \sum_{k=1}^{N_m}\Theta^Y_{(\xi_{m,k},\vec{f}_{m,k}),(\eta_{m,k},\vec{g}_{m,k})}\Big\|
\end{aligned}$$

for all $\lambda \in \Lambda$. Taking the limit with respect to λ shows that

$$\begin{aligned}
&\|\sum_{k=1}^{N_n} f^1_{n,k}{g^1_{n,k}}^* - \sum_{k=1}^{N_m} f^1_{m,k}{g^1_{m,k}}^*\| \\
&\qquad\qquad \leq \|\sum_{k=1}^{N_n}\Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})} - \sum_{k=1}^{N_m}\Theta^Y_{(\xi_{m,k},\vec{f}_{m,k}),(\eta_{m,k},\vec{g}_{m,k})})\|.
\end{aligned}$$

Since the $\sum_{k=1}^{N_n}\Theta^Y_{(\xi_{n,k},\vec{f}_{n,k}),(\eta_{n,k},\vec{g}_{n,k})}$'s converge in the operator norm on $\mathcal{L}(Y)$, this inequality implies that $\sum_{k=1}^{N_n} f^1_{n,k}{g^1_{n,k}}^*$ converges to an element in $\ker\phi$. If we let $a_2 = \lim_n \sum_{k=1}^{N_n} f^1_{n,k}{g^1_{n,k}}^* \in \ker\phi$, then Eq.(1) shows that $ag = a_2 g$ for all $g \in \ker\phi$. But then $a_1 := a - a_2 \in (\ker\phi)^\perp$, and consequently $a_1 \in J_X$. Since $a = a_1 + a_2$ the proof is complete. □

Lemma 4.6. *Let $(\tilde{\pi}, \tilde{t})$ be a representation of Y which is coisometric on* $\ker\phi \oplus T$*, and suppose that $\tilde{\pi}|_A$ is injective. For any $f \in \ker\phi$ we define $\epsilon_i(f) := (0,\ldots,0,f,0,\ldots) \in T$ where f appears in the i^{th} position. Then for every $i \in \mathbb{N}$ and for every $f \in \ker\phi$, the equation $\tilde{\pi}(0,\epsilon_i(f)) = 0$ implies that $f = 0$.*

Proof. First note that it suffices to prove the lemma for $f \geq 0$, because if $\tilde{\pi}(0,\epsilon_i(f)) = 0$ then $\tilde{\pi}(0,\epsilon_i(ff^*)) = \tilde{\pi}(0,\epsilon_i(f))\tilde{\pi}(0,\epsilon_i(f))^* = 0$, and $ff^* = 0$ if and only if $f = 0$.

If $\tilde{\pi}(0,\epsilon_i(f)) = 0$ and $f \geq 0$, then

$$\begin{aligned}
\|\tilde{t}(0,\epsilon_i(\sqrt{f}))\|^2 &= \|\tilde{t}(0,\epsilon_i(\sqrt{f}))^*\tilde{t}(0,\epsilon_i(\sqrt{f}))\| \\
&= \|\tilde{\pi}(\langle(0,\epsilon_i(\sqrt{f})),(0,\epsilon_i(\sqrt{f}))\rangle_B\| \\
&= \|\tilde{\pi}(0,\epsilon_i(f))\| \\
&= 0
\end{aligned}$$

so that $\tilde{t}(0, \epsilon_i(\sqrt{f})) = 0$ and consequently

$$\begin{aligned}0 &= \tilde{t}(0, \epsilon_i(\sqrt{f}))\tilde{t}(0, \epsilon_i(\sqrt{f}))^* = \tilde{\pi}^{(1)}(\Theta^Y_{(0,\epsilon_i(\sqrt{f})),(0,\epsilon_i(\sqrt{f}))}) \\ &= \begin{cases} \tilde{\pi}^{(1)}(\phi_B(f, \vec{0})) & \text{if } i = 1 \\ \tilde{\pi}^{(1)}(\phi_B(0, \epsilon_{i-1}(f))) & \text{if } i \geq 2 \end{cases} = \begin{cases} \tilde{\pi}(f, \vec{0}) & \text{if } i = 1 \\ \tilde{\pi}(0, \epsilon_{i-1}(f)) & \text{if } i \geq 2. \end{cases}\end{aligned}$$

If $i = 1$, the fact that $\tilde{\pi}|_A$ is injective implies that $f = 0$. Furthermore, an inductive argument combined with the above equality shows that for all $i \in \mathbb{N}$ we have $f = 0$. □

LEMMA 4.7. *Let $(\tilde{\pi}, \tilde{t})$ be a representation of Y which is coisometric on $J_Y = J(Y)$. If $\vec{f} = (f_1, f_2, \ldots) \in T$ and $\vec{g} = (g_1, g_2, \ldots) \in T$, then*

$$\tilde{t}(0, \vec{f})\tilde{t}(0, \vec{g})^* = \tilde{\pi}(f_1 g_1^*, (f_2 g_2^*, f_3 g_3^*, \ldots)).$$

Proof. For any $(\xi, \vec{h}) \in Y = X \oplus T$ we have

$$\begin{aligned}\Theta_{(0,\vec{f}),(0,\vec{g})}(\xi, \vec{h}) &= (0, \vec{f})\langle(0, \vec{g}), (\xi, \vec{h})\rangle_B = (0, \vec{f}\vec{g}^*\vec{h}) \\ &= \phi_B(f_1 g_1^*, ((f_2 g_2^*, \ldots))(\xi, \vec{h})\end{aligned}$$

so that $\Theta_{(0,\vec{f}),(0,\vec{g})} = \phi_B(f_1 g_1^*, (f_2 g_2^*, \ldots))$. Thus

$$\begin{aligned}\tilde{t}(0, \vec{f})\tilde{t}(0, \vec{g})^* &= \tilde{\pi}^{(1)}(\Theta_{(0,\vec{f}),(0,\vec{g})}) = \tilde{\pi}^{(1)}(\phi_B(f_1 g_1^*, (f_2 g_2^*, \ldots)) \\ &= \tilde{\pi}(f_1 g_1^*, (f_2 g_2^*, f_3 g_3^*, \ldots)).\end{aligned}$$

□

LEMMA 4.8. *Let $(\tilde{\pi}, \tilde{t})$ be a representation of Y. If $\xi \in X$, $a \in A$, and $\vec{f} \in T$, then the following relations hold:*

$$\begin{aligned}&(1)\quad \tilde{t}(0, \vec{f})\tilde{\pi}(a, \vec{0}) = 0 \\ &(2)\quad \tilde{t}(0, \vec{f})\tilde{t}(\xi, \vec{0}) = 0 \\ &(3)\quad \tilde{t}(0, \vec{f})\tilde{t}(\xi, \vec{0})^* = 0\end{aligned}$$

Proof. To see (1) we note that $\tilde{t}(0, \vec{f})\tilde{\pi}(a, \vec{0}) = \tilde{t}((0, \vec{f})(a, \vec{0})) = \tilde{t}(0, 0) = 0$. To see (2) and (3) let $\{\vec{e}_\lambda\}_{\lambda \in \Lambda}$ be an approximate unit for T. Then

$$\tilde{t}(0, \vec{f})\tilde{t}(\xi, \vec{0}) = \lim_\lambda \tilde{t}(0, \vec{f}\vec{e}_\lambda)\tilde{t}(\xi, \vec{0}) = \lim_\lambda \tilde{t}(0, \vec{f})\tilde{\pi}(0, \vec{e}_\lambda)\tilde{t}(\xi, \vec{0}) = 0$$

which shows that (2) holds, and

$$\begin{aligned}\tilde{t}(0, \vec{f})\tilde{t}(\xi, \vec{0})^* &= \lim_\lambda \tilde{t}(0, \vec{f}\vec{e}_\lambda)\tilde{t}(\xi, \vec{0})^* = \lim_\lambda \tilde{t}(0, \vec{f})\tilde{\pi}(0, \vec{e}_\lambda)\tilde{t}(\xi, \vec{0})^* \\ &= \lim_\lambda \tilde{t}(0, \vec{f})(\tilde{t}(\xi, \vec{0})\tilde{\pi}(0, \vec{e}_\lambda))^* = 0\end{aligned}$$

which shows that (3) holds. □

Lemma 4.9. *Let $(\tilde{\pi}, \tilde{t})$ be a representation of Y, and define $(\pi, t) := (\tilde{\pi}|_A, \tilde{t}|_X)$. If $c \in C^*(\pi, t)$ and $\vec{f} \in T$, then*

$$\tilde{t}(0, \vec{f})c = 0.$$

Proof. Since $C^*(\pi, t)$ is generated by elements of the form $\tilde{\pi}(a, \vec{0})$ and $\tilde{t}(\xi, \vec{0})$, the result follows from the relations in Lemma 4.8. □

Lemma 4.10. *Let $(\tilde{\pi}, \tilde{t})$ be a representation of Y which is coisometric on $J_Y = J(Y)$, and define $(\pi, t) := (\tilde{\pi}|_A, \tilde{t}|_X)$. If $n \in \{0, 1, 2, \ldots\}$, then any element of the form*

$$\tilde{t}(\xi_1, \vec{f_1}) \ldots \tilde{t}(\xi_n, \vec{f_n})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_n, \vec{g_n})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^*$$

will be equal to $c + \tilde{\pi}(0, \vec{k})$ for some $c \in C^*(\pi, t)$ and some $\vec{k} \in T$.

Proof. We shall prove this by induction on n.
Base Case: $n = 0$. Then the term above is equal to $\tilde{\pi}(a, \vec{h}) = \tilde{\pi}(a, \vec{0}) + \tilde{\pi}(0, \vec{h})$ and the claim holds trivially.
Inductive Step: Assume the claim holds for n. Given an element

$$\tilde{t}(\xi_1, \vec{f_1}) \ldots \tilde{t}(\xi_{n+1}, \vec{f}_{n+1})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_{n+1}, \vec{g}_{n+1})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^*$$

it follows from the inductive hypothesis that

$$\tilde{t}(\xi_2, \vec{f_2}) \ldots \tilde{t}(\xi_{n+1}, \vec{f}_{n+1})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_{n+1}, \vec{g}_{n+1})^* \ldots \tilde{t}(\eta_2, \vec{g_2})^*$$

has the form $c + \tilde{\pi}(0, \vec{k})$ for $c \in C^*(\pi, t)$ and $\vec{k} \in T$. Thus using Lemma 4.9 gives

$$\begin{aligned}
&\tilde{t}(\xi_1, \vec{f_1}) \ldots \tilde{t}(\xi_{n+1}, \vec{f}_{n+1})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_{n+1}, \vec{g}_{n+1})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^* \\
&= \tilde{t}(\xi_1, \vec{f_1})(c + \tilde{\pi}(0, \vec{k}))\tilde{t}(\eta_1, \vec{g_1})^* \\
&= (\tilde{t}(\xi_1, \vec{0}) + \tilde{t}(0, \vec{f_1}))(c + \tilde{\pi}(0, \vec{k}))(\tilde{t}(\eta_1, \vec{0}) + \tilde{t}(0, \vec{g_1})^*) \\
&= \tilde{t}(\xi_1, \vec{0})c\tilde{t}(\eta_1, \vec{0})^* + \tilde{t}(0, \vec{f_1})\tilde{\pi}(0, \vec{k})\tilde{t}(0, \vec{g_1})^* \\
&= \tilde{t}(\xi_1, \vec{0})c\tilde{t}(\eta_1, \vec{0})^* + \tilde{t}(0, \vec{f_1}\vec{k})\tilde{t}(0, \vec{g_1})^*.
\end{aligned}$$

It follows from Lemma 4.7 that $\tilde{t}(0, \vec{f_1}\vec{k})\tilde{t}(0, \vec{g_1})^*$ is of the form $c' + \tilde{\pi}(0, \vec{k'})$ with $c' \in \operatorname{im} \pi \subseteq C^*(\pi, t)$. Since $\tilde{t}(\xi_1, \vec{0})c\tilde{t}(\eta_1, \vec{0})^*$ is also in $C^*(\pi, t)$ the proof is complete. □

We wish to show that if $(\tilde{\pi}, \tilde{t})$ is a representation of Y and if we restrict to obtain $(\pi, t) := (\tilde{\pi}|_X, \tilde{t}|_A)$, then $C^*(\pi, t)$ is a corner of $C^*(\tilde{\pi}, \tilde{t})$. If A is unital and X is left essential, then this corner will be determined by the projection $\pi(1, \vec{0})$. However, in the following lemma we wish to consider the general case and must make use of approximate units to define the projection that determines the corner.

LEMMA 4.11. *Let X be a C^*-correspondence over A and let $T := (\ker \phi)^{\mathbb{N}}$ be the tail determined by $\ker \phi$. If $Y := X \oplus T$ is the C^*-correspondence over $B := A \oplus T$ formed by adding the tail T to X, and if $(\tilde{\pi}, \tilde{t})$ is a representation of Y, then there exists a projection $p \in \mathcal{M}(C^*(\tilde{\pi}, \tilde{t}))$ with the property that for all $a \in A$, $\xi \in X$, and $\vec{f} \in T$ the following relations hold:*

(1) $p\tilde{t}(\xi, \vec{f}) = \tilde{t}(\xi, (f_1, 0, 0, \ldots))$

(2) $\tilde{t}(\xi, \vec{f})p = \tilde{t}(\xi, \vec{0})$

(3) $p\tilde{\pi}(a, \vec{f}) = \tilde{\pi}(a, \vec{f})p = \tilde{\pi}(a, \vec{0})$

Proof. Let $\{\vec{e}_\lambda\}_{\lambda \in \Lambda}$ be an approximate unit for T, and for each $\lambda \in \Lambda$ let $\vec{e}_\lambda = (e_\lambda^1, e_\lambda^2, \ldots)$. Consider $\{\tilde{\pi}(0, \vec{e}_\lambda)\}_{\lambda \in \Lambda}$. For any element

$$\tilde{t}(\xi_1, \vec{f_1}) \ldots \tilde{t}(\xi_n, \vec{f_n})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_m, \vec{g_m})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^* \tag{2}$$

we have

$$\begin{aligned} &\lim_\lambda \tilde{\pi}(0, \vec{e}_\lambda)\tilde{t}(\xi_1, \vec{f_1}) \ldots \tilde{t}(\xi_n, \vec{f_n})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_m, \vec{g_m})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^* \\ = &\lim_\lambda \tilde{t}(0, (0, e_\lambda^1 f_{12}, e_\lambda^2 f_{13}, \ldots) \ldots \tilde{t}(\xi_n, \vec{f_n})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_m, \vec{g_m})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^* \\ = &\, \tilde{t}(0, (0, f_{12}, f_{13}, \ldots)) \ldots \tilde{t}(\xi_n, \vec{f_n})\tilde{\pi}(a, \vec{h})\tilde{t}(\eta_m, \vec{g_m})^* \ldots \tilde{t}(\eta_1, \vec{g_1})^* \end{aligned}$$

so this limit exists.
Now since any $c \in C^*(\tilde{\pi}, \tilde{t})$ can be approximated by a finite sum of elements of the form shown in (2), it follows that $\lim_\lambda \tilde{\pi}(0, \vec{e}_\lambda)c$ exists for all $c \in C^*(\tilde{\pi}, \tilde{t})$. Let us view $C^*(\tilde{\pi}, \tilde{t})$ as a C^*-correspondence over itself (see [20, Example 2.10]). If we define $q : C^*(\tilde{\pi}, \tilde{t}) \to C^*(\tilde{\pi}, \tilde{t})$ by $q(c) = \lim_\lambda \tilde{\pi}(0, \vec{e}_\lambda)c$ then we see that for any $c, d \in C^*(\tilde{\pi}, \tilde{t})$ we have

$$d^*q(c) = \lim_\lambda d^*\tilde{\pi}(0, \vec{e}_\lambda)c = \lim_\lambda (\tilde{\pi}(0, \vec{e}_\lambda)d)^*c = q(d)^*c$$

and hence q is an adjointable operator on $C^*(\tilde{\pi}, \tilde{t})$. Therefore q defines (left multiplication by) an element in the multiplier algebra $\mathcal{M}(C^*(\tilde{\pi}, \tilde{t}))$ [20, Theorem 2.47]. It is easy to check that $q^2 = q^* = q$ so that q is a projection. Now if we let $p := 1 - q$ in $\mathcal{M}(C^*(\tilde{\pi}, \tilde{t}))$, then it is easy to check that relations (1), (2), and (3) follow from the definition of q. $\square$

Proof of Theorem 4.3. (a) Let $I := \ker \phi$, set $\mathcal{H}_0 := \pi(I)\mathcal{H}_X$, and define $\mathcal{H}_T := \bigoplus_{i=1}^\infty \mathcal{H}_i$ where $\mathcal{H}_i = \mathcal{H}_0$ for all $i = 1, 2, \ldots$. We define $\tilde{t} : Y \to \mathcal{B}(\mathcal{H}_X \oplus \mathcal{H}_T)$ and $\tilde{\pi} : B \to \mathcal{B}(\mathcal{H}_X \oplus \mathcal{H}_T)$ as follows: Viewing Y as $Y = X \oplus T$ and B as $B = A \oplus T$, for any $(h, (h_1, h_2, \ldots)) \in \mathcal{H}_{\mathcal{Q}} \oplus \mathcal{H}_T$ we define

$$\tilde{t}(\xi, (f_1, f_2, \ldots))(h, (h_1, h_2, \ldots)) = (t(\xi)h + \pi(f_1)h_1, (\pi(f_2)h_2, \pi(f_3)h_3, \ldots))$$

and

$$\tilde{\pi}(a, (f_1, f_2, \ldots))(h, (h_1, h_2, \ldots)) = (\pi(a)h, (\pi(f_1)h_1, \pi(f_2)h_2, \ldots)).$$

Then it is straightforward to show that $(\tilde{\pi}, \tilde{t})$ is a representation of Y on $\mathcal{H}_Q \oplus \mathcal{H}_T$. To see that $(\tilde{\pi}, t)$ is coisometric on $J(Y)$, choose an element $(a, (f_1, f_2, \ldots)) \in J(Y)$. By Lemma 4.5 we know that $a = a_1 + a_2$ for $a_1 \in J_X$ and $a_2 \in \ker \phi$. Furthermore, since $a_1 \in J(X)$ we may write $\phi(a_1) = \lim_n \sum_{k=1}^{N_n} \Theta^X_{\xi_{n,k}, \eta_{n,k}}$ for some $\xi_{n,k}, \eta_{n,k} \in X$. It follows that

$$\phi_B(a_1, \vec{0}) = \lim_n \sum_{k=1}^{N_n} \Theta^Y_{(\xi_{n,k}, \vec{0}), (\eta_{n,k}, \vec{0})} \in \mathcal{K}(Y).$$

In addition, since $a_2 \in \ker \phi$ we see that if we let $\{\vec{e}_\lambda\}_{\lambda \in \Lambda}$ be an approximate unit for T with $\vec{e}_\lambda = (e^1_\lambda, e^2_\lambda, \ldots)$ for each λ, then

$$\phi_B(a_2, \vec{f}) = \lim_\lambda \Theta^Y_{(0,(a,f_1,f_2,\ldots)),(0,(e^1_\lambda, e^2_\lambda, \ldots))} \in \mathcal{K}(Y).$$

Now for any $n \in \mathbb{N}$ we see that $\{e^n_\lambda\}_{\lambda \in \Lambda}$ is an approximate unit for $\ker \phi$. Furthermore, we see that for all $(\xi, \vec{f}), (\eta, \vec{g}) \in Y = X \oplus T$ we have

$$\tilde{t}(\xi, \vec{f})\tilde{t}(\eta, \vec{g})^* = (t(\xi)t(\eta)^* + \pi(f_1 g_1^*), (\pi(f_2 g_2^*), \pi(f_3 g_3^*), \ldots))$$

and thus

$$\begin{aligned}
&\tilde{\pi}^{(1)}(\phi_B(a, \vec{f})) \\
&= \tilde{\pi}^{(1)}(\phi_B(a_1, \vec{0})) + \tilde{\pi}^{(1)}(\phi_B(a_2, \vec{f})) \\
&= \lim_n \sum_{k=1}^{N_n} \tilde{t}(\xi_{n,k}, \vec{0})\tilde{t}(\eta_{n,k}, \vec{0})^* + \lim_\lambda \tilde{t}(0, (a_2, f_1, f_2, \ldots))\tilde{t}(0, (e^1_\lambda, e^2_\lambda, \ldots))^* \\
&= \lim_n \sum_{k=1}^{N_n} (t(\xi_{n,k})t(\eta_{n,k})^*, \vec{0}) + \lim_\lambda (\pi(a_2 e^1_\lambda), (\pi(f_1 e^2_\lambda), \pi(f_2 e^3_\lambda), \ldots)) \\
&= (\pi^{(1)}(\phi(a_1)), \vec{0}) + (\pi(a_2), (\pi(f_1), \pi(f_2), \ldots)) \\
&= (\pi(a_1), \vec{0}) + (\pi(a_2), (\pi(f_1), \pi(f_2), \ldots)) \\
&= \tilde{\pi}(a, \vec{f})
\end{aligned}$$

so $(\tilde{\pi}, \tilde{t})$ is coisometric on $J(Y)$.

(b) If $(\tilde{\pi}, \tilde{t})$ is a representation of Y in a C^*-algebra C which is coisometric on $J(Y)$, then it is straightforward to see that the restriction $(\tilde{\pi}|_A, \tilde{t}|_X)$ is a representation. To see that $(\tilde{\pi}|_A, \tilde{t}|_X)$ is coisometric on J_X, choose an element $a \in J_X$. Since $J_X \subseteq J(X)$ we may write $\phi(a) = \lim_n \sum_{k=1}^{N_n} \Theta^X_{\xi_{n,k}, \eta_{n,k}}$ for some $\xi_{n,k}, \eta_{n,k} \in X$. In addition, since $a \in (\ker \phi)^\perp \subseteq J_X$ we have that

$$\phi_B(a, \vec{0}) = \lim_n \sum_{k=1}^{N_n} \Theta^Y_{(\xi_{n,k}, \vec{0}), (\eta_{n,k}, \vec{0})} \in \mathcal{K}(Y)$$

and we have

$$\begin{aligned}\tilde{\pi}|_A^{(1)}(\phi_A(a)) &= \lim_n \sum_{k=1}^{N_n} \tilde{t}|_X(\xi_{n,k})\tilde{t}|_X(\eta_{n,k})^* = \lim_n \sum_{k=1}^{N_n} \tilde{t}(\xi_{n,k},\vec{0})\tilde{t}(\eta_{n,k},\vec{0})^* \\ &= \tilde{\pi}^{(1)}(\phi_B(a,\vec{0})) = \tilde{\pi}(a,\vec{0}) = \tilde{\pi}|_A(a)\end{aligned}$$

so $(\tilde{\pi}|_A, \tilde{t}|_X)$ is coisometric on J_X.

Furthermore, suppose that the restriction $\tilde{\pi}|_A$ is injective. If $(a,\vec{f}) \in B := A \oplus T$ and $\tilde{\pi}(a,\vec{f}) = 0$, let $\{g_\lambda\}_{\lambda\in\Lambda}$ be an approximate unit for $\ker\phi$, and for any $f \in \ker\phi$ and $i \in \mathbb{N}$ let $\epsilon_i(f) := (0,\ldots,0,f,0,\ldots)$ where f is in the i^{th} position. Since $\tilde{\pi}(a,\vec{f}) = 0$ we see that if we write $\vec{f} = (f_1, f_2, \ldots)$, then for all $i \in \mathbb{N}$ we have

$$\tilde{\pi}(0,\epsilon_i(g_\lambda f_i)) = \tilde{\pi}(0,\epsilon_i(g_\lambda))\tilde{\pi}(a,\vec{f}) = 0,$$

and taking limits with respect to λ shows that $\tilde{\pi}(0,\epsilon_i(f_i)) = 0$ for all $i \in \mathbb{N}$. From Lemma 4.6 it follows that $f_i = 0$ for all $i \in \mathbb{N}$. Thus $\vec{f} = 0$, and since $\tilde{\pi}|_A$ is injective we also have that $a = 0$. Hence $\tilde{\pi}$ is injective.

(c) The fact that $(\pi,t) := (\pi_Y|_A, t_Y|_X)$ is a representation which is coisometric on J_X follows from Part (b). Furthermore, the fact that $\rho_{(\pi,t)}$ is injective follows from Part (a) which shows that any $*$-representation of $\mathcal{O}_X$ factors through a $*$-representation of $\mathcal{O}_Y$. All that remains is to show that $\operatorname{im}\rho_{(\pi,t)} = C^*(\pi,t)$ is a full corner of $\mathcal{O}_Y$.

Let $p \in \mathcal{M}(\mathcal{O}_Y)$ be the projection described in Lemma 4.11. We shall first show that $C^*(\pi,t) = p\mathcal{O}_Y p$. To begin, we see from the relations in Lemma 4.11 that for all $a \in A$ we have $p\pi(a)p = p\pi_Y(a,\vec{0})p = \pi_Y(a,\vec{0}) = \pi(a)$ and for all $\xi \in X$ we have $pt(\xi)p = p(t_Y(\xi,\vec{0}))p = t_Y(\xi,\vec{0}) = t(\xi)$. Thus $C^*(\pi,t) \subseteq p\mathcal{O}_Y p$.

To see the reverse inclusion, note that any element in $\mathcal{O}_Y$ is the limit of sums of elements of the form

$$t_Y(\xi_1,\vec{f_1})\ldots t_Y(\xi_n,\vec{f_n})\pi_Y(a,\vec{h})t_Y(\eta_m,\vec{g_m})^*\ldots t_Y(\eta_1,\vec{g_1})^*$$

and thus any element of $p\mathcal{O}_Y p$ is the limit of sums of elements of the form

$$pt_Y(\xi_1,\vec{f_1})\ldots t_Y(\xi_n,\vec{f_n})\pi_Y(a,\vec{h})t_Y(\eta_m,\vec{g_m})^*\ldots t_Y(\eta_1,\vec{g_1})^*p$$

Therefore, it suffices to show that each of these elements is in $C^*(\pi,t)$. Now if $n \geq m$, then we may use Lemma 4.10 to write

$$t_Y(\xi_{n-m+1},\vec{f}_{n-m+1})\ldots t_Y(\xi_n,\vec{f_n})\pi_Y(a,\vec{h})t_Y(\eta_m,\vec{g_m})^*\ldots t_Y(\eta_1,\vec{g_1})^*$$

as $c + \pi_Y(0, \vec{k})$ for $c \in C^*(\pi, t)$ and $\vec{k} \in T$. Then

$$\begin{aligned}
&pt_Y(\xi_1, \vec{f_1}) \dots t_Y(\xi_n, \vec{f_n})\pi_Y(a, \vec{h})t_Y(\eta_m, \vec{g_m})^* \dots t_Y(\eta_1, \vec{g_1})^* p \\
&= pt_Y(\xi_1, \vec{f_1}) \dots t_Y(\xi_{n-m}, \vec{f}_{n-m})(c + \pi_Y(0, \vec{k}))p \\
&= pt_Y(\xi_1, \vec{f_1}) \dots t_Y(\xi_{n-m}, \vec{f}_{n-m})cp \\
&= pt_Y(\xi_1, \vec{f_1}) \dots t_Y(\xi_{n-m}, \vec{f}_{n-m})pcp \\
&= pt_Y(\xi_1, \vec{f_1}) \dots t_Y(\xi_{n-m-1}, \vec{f}_{n-m-1})pt_Y(\xi_{n-m}, \vec{0})pcp \\
&\quad \vdots \\
&= pt_Y(\xi_1, \vec{0})p \dots pt_Y(\xi_{n-m}, \vec{0})pcp \\
&= t_Y(\xi_1, \vec{0}) \dots t_Y(\xi_{n-m}, \vec{0})c \\
&= t(\xi_1) \dots t(\xi_{n-m})c
\end{aligned}$$

which is in $C^*(\pi, t)$. The case when $n \leq m$ is similar. Hence $p\mathcal{O}_Y p \subseteq C^*(\pi, t)$. To see that the corner $C^*(\pi, t) = p\mathcal{O}_Y p$ is full, suppose that $\mathcal{I}$ is an ideal in $\mathcal{O}_Y$ that contains $C^*(\pi, t)$. For $f \in \ker\phi$ and $n \in N$ define $\epsilon_n(f) := (0, \dots, 0, f, 0, \dots) \in T$, where the term f is in the n^{th} position. Let $\{e_\lambda\}_{\lambda\in\Lambda}$ be an approximate unit for $\ker\phi$. Now $t_Y(\xi, \vec{0}), \pi_Y(a, \vec{0}) \in C^*(\pi, t) \subseteq \mathcal{I}$ for all $a \in A$ and $\xi \in X$, and since T is the c_0-direct sum of countably many copies of $\ker\phi$ in order to show that $\mathcal{I}$ is all of $\mathcal{O}_Y$ it suffices to prove that for all $n \in N$ and $\lambda \in \Lambda$ we have $t_Y(0, \epsilon_n(e_\lambda)) \in \mathcal{I}$ and $\pi_Y(0, \epsilon_n(e_\lambda)) \in \mathcal{I}$. We shall prove this by induction on n.

Base Case: For any $\beta, \lambda \in \Lambda$ we have from Lemma 4.7 that

$$t_Y(0, \epsilon_1(e_\lambda))t_Y(0, \epsilon_1(e_\beta))^* = \pi_Y(e_\lambda e_\beta, \vec{0}) \in \mathcal{I}.$$

Also for any $\alpha \in \Lambda$ we have

$$\begin{aligned}
t_Y(0, (\epsilon_1(e_\lambda e_\beta e_\alpha)) &= t_Y(0, \epsilon_1(e_\lambda))\pi_Y(0, \epsilon_1(e_\beta^* e_\alpha)) \\
&= t_Y(0, \epsilon_1(e_\lambda))t_Y(0, \epsilon_1(e_\beta))^* t_Y(0, \epsilon_1(e_\alpha))
\end{aligned}$$

which is in $\mathcal{I}$. Taking limits with respect to α and β gives

$$t_Y(0, \epsilon_1(e_\lambda)) = \lim_\beta \lim_\alpha t_Y(0, \epsilon_1(e_\lambda e_\beta e_\alpha)) \in \mathcal{I}.$$

Furthermore, since $t_Y(0, \epsilon_1(e_\lambda)) \in \mathcal{I}$ for all $\lambda \in \Lambda$, we see that

$$\pi_Y(0, \epsilon_1(e_\lambda)) = \lim_\beta \pi_Y(0, \epsilon_1(e_\lambda e_\beta)) = \lim_\beta t_Y(0, \epsilon_1(e_\lambda))^* t_Y(0, \epsilon_1(e_\beta)) \in \mathcal{I}.$$

Inductive step: Suppose that $t_Y(0, \epsilon_n(e_\lambda)), \pi_Y(0, \epsilon_n(e_\lambda)) \in \mathcal{I}$ for any $\lambda \in \Lambda$. Then for all $\lambda, \beta \in \Lambda$ we have

$$\begin{aligned}
t_Y(0, \epsilon_{n+1}(e_\lambda))t_Y(0, \epsilon_{n+1}(e_\beta))^* &= \pi_Y^{(1)}(\Theta^Y_{(0,\epsilon_{n+1}(e_\lambda)),(0,\epsilon_{n+1}(e_\beta))}) \\
&= \pi_Y^{(1)}(\phi_B(0, \epsilon_n(e_\beta e_\lambda))) \\
&= \pi_Y(0, \epsilon_n(e_\beta e_\lambda)) \\
&= \pi_Y(0, \epsilon_n(e_\beta))\pi_Y(0, \epsilon_n(e_\lambda))
\end{aligned}$$

which is in $\mathcal{I}$. Thus for any $\alpha \in \Lambda$ we have that

$$\begin{aligned} t_Y(0, \epsilon_{n+1}(e_\lambda e_\beta e_\alpha)) &= t_Y(0, \epsilon_{n+1}(e_\lambda))\pi_Y(0, \epsilon_{n+1}(e_\beta e_\alpha)) \\ &= t_Y(o, \epsilon_{n+1}(e_\lambda))t_Y(0, \epsilon_{n+1}(e_\beta))^* t_Y(0, \epsilon_{n+1}(e_\alpha)) \end{aligned}$$

is in $\mathcal{I}$. Taking limits with respect to α and β gives

$$t_Y(0, \epsilon_{n+1}(e_\lambda)) = \lim_\beta \lim_\alpha t_Y(0, \epsilon_{n+1}(e_\lambda e_\beta e_\alpha)) \in \mathcal{I}.$$

Furthermore, since $t_Y(0, \epsilon_{n+1}(e_\lambda)) \in \mathcal{I}$ for all $\lambda \in \Lambda$, we have

$$\pi_Y(0, \epsilon_{n+1}(e_\lambda)) = \lim_\beta \pi_Y(0, \epsilon_{n+1}(e_\beta e_\lambda)) = \lim_\beta t_Y(0, \epsilon_{n+1}(e_\beta))^* t_Y(0, \epsilon_{n+1}(e_\lambda))$$

which is in the ideal $\mathcal{I}$. □

5 Gauge-Invariant Uniqueness

Recall that we let γ denote the gauge action of $\mathbb{T}$ on $\mathcal{O}_X$. A gauge-invariant uniqueness was proven in [6, Theorem 4.1] for (augmented) Cuntz-Pimsner algebras. Our method of adding tails, together with Theorem 4.3, will allow us to extend this theorem to the case when ϕ is not injective, and ultimately to all relative Cuntz-Pimsner algebras.

The following Gauge-Invariant Uniqueness Theorem was proven by Katsura using direct methods in [11, Theorem 6.4]. We shall now give an alternate proof, showing how the method of adding tails can be used to bootstrap [6, Theorem 4.1] to the general case.

Theorem 5.1 (Gauge-Invariant Uniqueness). *Let X be a C^*-correspondence over A, and let (π_X, t_X) be a universal $J(X)$-coisometric representation of X. If $\rho : \mathcal{O}_X \to C$ is a homomorphism between C^*-algebras which satisfies the following two conditions:*

1. *the restriction of ρ to $\pi_X(A)$ is injective*
2. *there is a strongly continuous action $\beta : \mathbb{T} \to \mathrm{Aut}(\rho(\mathcal{O}_X))$ such that $\beta_z \circ \rho = \rho \circ \gamma_z$ for all $z \in \mathbb{T}$*

then ρ is injective.

Remark 5.2. When ϕ is injective, the statement above is actually an equivalent reformulation of [6, Theorem 4.1]. The equivalence relies on the fact that for any C^*-correspondence X, the universal $J(X)$-coisometric representation (i_A, i_X) has the property that i_A is injective if and only if the left action ϕ is injective.

Proof of Theorem 5.1. Let $T := (\ker \phi)^{\mathbb{N}}$ be the tail determined by $\ker \phi$, and let $Y := X \oplus T$ be the C^*-correspondence over $B := A \oplus T$ formed by adding the tail T to X. By Theorem 4.3(c) we may identify $(\mathcal{O}_X, \pi_X, t_X)$ with $(S, \pi_Y|_A, t_Y|_X)$ where S is the C^*-subalgebra of $\mathcal{O}_Y$ generated by

$$\{\pi_Y(a, \vec{0}), t_Y(\xi, \vec{0}) : a \in A \text{ and } \xi \in X\}.$$

Since $\beta : \mathbb{T} \to \operatorname{Aut}(\operatorname{im}\rho)$ is an action of $\mathbb{T}$ on $\operatorname{im}\rho$, there exists a Hilbert space $\mathcal{H}_X$, a faithful representation $\kappa : \operatorname{im}\rho \to \mathcal{B}(\mathcal{H}_X)$, and a unitary representation $U : \mathbb{T} \to \mathcal{U}(\mathcal{H}_X)$ such that

$$\kappa(\beta_z(x)) = U_z \kappa(x) U_z^* \qquad \text{for all } x \in \operatorname{im}\rho \text{ and } z \in \mathbb{T}.$$

In addition, since $\tau := \kappa \circ \rho$ is a $*$-homomorphism from S into $\mathcal{B}(\mathcal{H}_X)$ which is faithful on $\pi_X(A)$, it follows from Theorem 4.3(a) that τ may be extended to a $*$-homomorphism $\tilde{\tau} : \mathcal{O}_Y \to \mathcal{B}(\mathcal{H}_X \oplus \mathcal{H}_T)$ with $\tilde{\tau}$ faithful on $\pi_Y(B)$.
We shall now define a unitary representation $W : \mathbb{T} \to \mathcal{B}(\mathcal{H}_X \oplus \mathcal{H}_T)$ as follows. We see from the proof of Theorem 4.3(a) that $\mathcal{H}_T := \bigoplus_{i=1}^\infty \mathcal{H}_i$. Thus for $(h, (h_1, h_2, \ldots)) \in \mathcal{H}_\mathcal{Q} \oplus \mathcal{H}_T$ we define

$$W_z(h, (h_1, h_2, \ldots)) := (U_z h, (z^{-1}h_1, z^{-2}h_2, \ldots)) \qquad \text{for } z \in \mathbb{T}.$$

We may then define $\tilde{\beta} : \mathbb{T} \to \operatorname{Aut}(\mathcal{B}(\mathcal{H}_X \oplus \mathcal{H}_T))$ by $\tilde{\beta}_z(T_0) := W_z T_0 W_z^*$, and we see that $\tilde{\beta}$ is a strongly continuous gauge action. Furthermore, if γ' denotes the gauge action of $\mathbb{T}$ on $\mathcal{O}_Y$, then $\tilde{\beta}_z \circ \tilde{\tau} = \tilde{\tau} \circ \gamma'_z$ (to see this recall how the extension $\tilde{\tau}$ is defined in the proof of Theorem 4.3(a) and then simply check on the generators $\{t_Y(\xi, \vec{f}), \pi_Y(a, \vec{g}) : \xi \in X, a \in A, \text{ and } \vec{f}, \vec{g} \in T\}$). Thus by [6, Theorem 4.1] we have that $\tilde{\tau}$ is injective. Hence $\tilde{\tau}|_S = \tau = \kappa \circ \rho$ is injective, and ρ is injective. $\square$

To conclude this section we shall interpret our result in the relative Cuntz-Pimsner setting.

Remark 5.3. Katsura has shown in [12] that if $\mathcal{O}(K, X)$ is a relative Cuntz-Pimsner algebra, then there exists a C^*-correspondence X' with the property that $\mathcal{O}_{X'}$ is naturally isomorphic to $\mathcal{O}(K, X)$. Using this analysis one can obtain the following interpretation of Theorem 5.1 for relative Cuntz-Pimsner algebras.

Interpretation of Theorem 5.1 for Relative Cuntz-Pimsner Algebras: *Let X be a C^*-correspondence with left action $\phi : X \to \mathcal{L}(X)$, let K be an ideal in $J(X) := \phi^{-1}(\mathcal{K}(X))$, and let (π_X, t_X) be a universal K-coisometric representation of X. If $\rho : \mathcal{O}_X \to C$ is a homomorphism between C^*-algebras which satisfies the following three conditions:*

(1) the restriction of ρ to $\pi_X(A)$ is injective

(2) if $\rho(\pi_X(a)) \in \rho(\pi_X^{(1)}(\mathcal{K}(X)))$, then $\pi_X(a) \in \pi_X(K)$

(3) there is a strongly continuous action $\beta : \mathbb{T} \to \operatorname{Aut}(\rho(\mathcal{O}_X))$ such that $\beta_z \circ \rho = \rho \circ \gamma_z$ for all $z \in \mathbb{T}$

then ρ is injective.

Finally, we mention that if we define a map $T_K : J(X) \to \mathcal{O}(K, X)$ by

$$T_K(a) := \pi_X(a) - \pi_X^{(1)}(\phi(a))$$

then the equation

$$\begin{aligned} T_K(a)T_K(b) &= (\pi_X(a) - \pi_X^{(1)}(\phi(a)))(\pi_X(b) - \pi_X^{(1)}(\phi(b))) \\ &= \pi_X(ab) - \pi_X^{(1)}(\phi(a))\pi_X(b) - \pi_X(a)\pi_X^{(1)}(\phi(b)) + \pi_X^{(1)}(\phi(ab)) \\ &= \pi_X(ab) - \pi_X^{(1)}(\phi(ab)) \\ &= T_K(ab) \end{aligned}$$

shows that this map is a homomorphism. If π_X is injective (which by [15, Proposition 2.21] occurs if and only if $K \cap \ker\phi = \emptyset$), then we may replace Condition (2) in the above statement by the condition

(2′) the restriction of ρ to $T_K(J(X))$ is injective.

6 Gauge-Invariant Ideals

In this section we use Theorem 5.1 to characterize the gauge-invariant ideals in C^*-algebras associated to certain correspondences.

Definition 6.1. Let X be a C^*-correspondence over A. We say that an ideal $I \lhd A$ is *X-invariant* if $\phi(I)X \subseteq XI$. We say that an X-invariant ideal $I \lhd A$ is *X-saturated* if

$$a \in J_X \text{ and } \phi(a)X \subseteq XI \implies a \in I.$$

Remark 6.2. In [9] the authors only considered Hilbert bimodules (i.e. C^*-correspondences) for which ϕ is injective and $\phi(A) \subseteq \mathcal{K}(X)$, and thus the definition of X-saturated that they gave was that $a \in A$ and $\phi(a)X \subseteq XI$ implies $a \in I$. Since $J_X = A$ throughout their paper, this notion is equivalent to the one defined in Definition 6.1. In [6, Remark 3.11] it was suggested that the definition of X-saturated for general C^*-correspondences should also be that $a \in A$ and $\phi(a)X \subseteq XI$ implies $a \in I$. However, after considering how the definition of *saturated* was extended to (or rather modified for) non-row-finite graphs in [1, §3] and [3, §3] we believe that Definition 6.1 is the appropriate generalization.

Recall that if I is an ideal of A, then

$$X_I := \{x \in X : \langle x, y\rangle_A \in I \text{ for all } y \in X\}$$

is a right Hilbert A-module, and by the Hewitt-Cohen Factorization Theorem $X_I = XI := \{x \cdot i : x \in X \text{ and } i \in I\}$ (see [6, §2]). Furthermore, X/XI is a right Hilbert A/I-module in the obvious way [6, Lemma 2.1]. In order for

X/XI to be a C^*-correspondence, we need the ideal I to be X-invariant. Let $q^I : A \to A/I$ and $q^{XI} : X \to X/XI$ be the appropriate quotient maps. If I is X-invariant, then one may define $\phi_{A/I} : A/I \to \mathcal{L}(X/XI)$ by

$$\phi_{A/I}(q^I(a))(q^{XI}(x)) := q^{XI}(\phi(a)(x))$$

and with this action X/XI is a C^*-correspondence over A/I [6, Lemma 3.2].

LEMMA 6.3. *Let X be a C^*-correspondence over a C^*-algebra A, and let I be an X-saturated X-invariant ideal in A. If $q^I : A \to A/I$ denotes the quotient map, then*

$$q^I(J_X) \subseteq J_{X/XI}.$$

Furthermore, if X has the following two properties:

1. $\phi(A) \subseteq \mathcal{K}(X)$
2. $\ker \phi$ *is complemented in A (i.e. there exists an ideal J of A with the property that $A = J \oplus \ker \phi$),*

then

$$q^I(J_X) = J_{X/XI}.$$

Proof. Let $a \in J_X$. Then $a \in J(X)$, and it follows from [6, Lemma 2.7] that $q^I(a) \in J(X/XI)$. Also, if $q^I(b) \in \ker \phi_{A/I}$, then $q^I(ab) \in \ker \phi_{A/I}$ and for all $x \in X$ we have

$$q^{XI}(\phi(ab)(x)) = \phi_{A/I}(ab)q^{XI}(x) = 0$$

and thus

$$\phi(ab)XI \subseteq XI. \tag{3}$$

Since $a \in J_X$ and J_X is an ideal, we see that $ab \in J_X$. Now since I is X-saturated, (3) implies that $ab \in I$ and $q^I(a)q^I(b) = q^I(ab) = 0$. Thus $q^I(a) \in (\ker \phi_{A/I})^\perp$ and $q^I(a) \in J_{X/XI}$.

Now suppose that Conditions (1) and (2) in the statement of the lemma hold. Since $\phi(A) \subseteq \mathcal{K}(X)$ it follows that $J(X) = A$. In addition, [6, Lemma 2.7] shows that $q^I(J(X)) = J(X/XI)$. From Condition (2) we know that $A = J \oplus \ker \phi$ for some ideal J of A. However, the definition of J_X then implies that $J = J_X$. Thus if $a \in A$ and $q^I(a) \in J_{X/XI}$, then we may write $a = b + c$ for $b \in J_X$ and $c \in \ker \phi$. But then $q^I(b) \in J_{X/XI}$ by the first part of the lemma, and $q^I(c) = q^I(a) - q^I(b) \in J_{X/XI}$. Since $c \in \ker \phi$ it follows that for all $x \in X$ we have

$$\phi_{A/I}(q^I(c))q^{XI}(x) = q^{XI}(\phi(c)(x)) = 0$$

and thus $q^I(c) \in \ker \phi_{A/I}$. Thus $q^I(c) \in J_{X/XI} \cap \ker \phi_{A/I} = \{0\}$ so $q^I(c) = 0$ and $q^I(a) = q^I(b) \in q^I(J_X)$. Thus $J_{X/XI} \subseteq q^I(J_X)$. □

The following theorem was proven in [9, Theorem 4.3] under the hypotheses that ϕ is injective, A is unital, and X is full and finite projective as a right A-module (so in particular, $\phi(A) \subseteq \mathcal{K}(X)$). However, Theorem 5.1 allows us to give a fairly simple proof of the result for much more general C^*-correspondences.

THEOREM 6.4. *Let X be a C^*-correspondence with the following two properties:*

1. $\phi(A) \subseteq \mathcal{K}(X)$
2. $\ker \phi$ *is complemented in A (i.e. there exists an ideal J of A with the property that $A = J \oplus \ker \phi$),*

and let (π_X, t_X) be a universal $J(X)$-coisometric representation of X. Then there is a lattice isomorphism from the X-saturated X-invariant ideals of A onto the gauge-invariant ideals of $\mathcal{O}_X$ given by

$$I \mapsto \mathcal{I}(I) := \textit{the ideal in } \mathcal{O}_X \textit{ generated by } \pi_X(I)$$

Proof. To begin we see that $\mathcal{I}(I)$ is in fact gauge invariant since

$$\mathcal{I}(I) = \overline{\text{span}}\{t_X(x_1)\dots t_X(x_n)\pi_X(a)t_X(y_1)^* \dots t_X(y_m)^* \\ : a \in I, x_1 \dots x_n \in X, y_1 \dots y_m \in X, \text{and } n, m \geq 0\}.$$

In addition, the map $I \mapsto \mathcal{I}(I)$ is certainly inclusion preserving.

To see that the map is surjective, let $\mathcal{I}$ be a gauge-invariant ideal in $\mathcal{O}_X$. If we define $I := \pi_X^{-1}(\mathcal{I})$, then it is straightforward to show that I is X-invariant and X-saturated. Now clearly $\mathcal{I}(I) \subseteq \mathcal{I}$ so there exists a quotient map $q : \mathcal{O}_X/\mathcal{I}(I) \to \mathcal{O}_X/\mathcal{I}$. Furthermore, by [6, Theorem 3.1] we have that $\mathcal{O}_X/\mathcal{I}(I)$ is canonically isomorphic to $\mathcal{O}(q^I(J_X), X/XI)$, which by Lemma 6.3 is equal to $\mathcal{O}_{X/XI} := \mathcal{O}(J_{X/XI}, X/XI)$. If we identify $\mathcal{O}_X/\mathcal{I}(I)$ with $\mathcal{O}_{X/XI}$, then we see that $q(\pi_{X/XI}(q^I(a))) = 0$ implies that $\pi_X(a) \in \mathcal{I}$ so that $a \in I$ and $q^I(a) = 0$. Thus q is faithful on $\pi_{X/XI}(A/I)$. Furthermore, since $\mathcal{I}$ is gauge invariant, the gauge action on $\mathcal{O}_X$ descends to an action on the quotient $\mathcal{O}_X/\mathcal{I}$, and q intertwines this action and the action on $\mathcal{O}_{X/XI}$. Therefore Theorem 5.1 implies that q is injective and consequently $\mathcal{I}(I) = \mathcal{I}$.

To see that the above map is injective it suffices to prove that $\pi_X(a) \in \mathcal{I}(I)$ if and only if $a \in I$. Now $\mathcal{O}_X/\mathcal{I}(I)$ is canonically isomorphic to $\mathcal{O}_{X/XI}$ as in the previous paragraph. Hence $\pi_X(a) \in \mathcal{I}(I)$ implies $\pi_{A/I}(q^I(a)) = 0$, but since $\pi_{X/XI}$ is injective by Corollary 4.4 it follows that $q^I(a) = 0$ and $a \in I$. □

Remark 6.5. We mention that in [17] we have constructed examples which show that the above theorem does not hold if either of the hypotheses (1) or (2) are removed. We also mention that Katsura [12] has given a description of the gauge-invariant ideals in C^*-algebras associated to general C^*-correspondences in terms certain pairs of ideals in A.

7 CONCLUDING REMARKS

In Section 4 we gave a method for "adding tails to sinks" in C^*-correspondences; that is, given a C^*-correspondence X we described how to form a C^*-correspondence Y with the property that the left action of Y is injective and $\mathcal{O}_X$ is canonically isomorphic to a full corner in $\mathcal{O}_Y$. The process of adding tails to

C^*-correspondences provides a useful tool for extending results for augmented Cuntz-Pimsner algebras (i.e. C^*-algebras associated to C^*-correspondences in which ϕ is injective) to C^*-algebras associated to general C^*-correspondences. We used this idea in Section 5 to extend the Gauge-Invariant Uniqueness Theorem for augmented Cuntz-Pimsner algebras to the general case. More generally, however, we see that many questions about C^*-algebras associated to correspondences may be reduced to the corresponding questions for augmented Cuntz-Pimsner algebras. For example, we see that for any property that is preserved by Morita equivalence (e.g. simplicity, AF-ness, pure infiniteness), one need only characterize when augmented Cuntz-Pimsner algebras will have this property, and then by adding tails one may easily deduce a theorem for C^*-algebras associated to general C^*-correspondences.

In addition, if $p \in \mathcal{M}(\mathcal{O}_Y)$ is the projection that determines $\mathcal{O}_X$ as a full corner of $\mathcal{O}_Y$ (so that $\mathcal{O}_X \cong p\mathcal{O}_Y p$), then the Rieffel correspondence from the lattice of ideals of $\mathcal{O}_Y$ to the lattice of ideals of $\mathcal{O}_X$ takes the form $I \mapsto pIp$. Furthermore, we see from Lemma 4.11 that p is gauge invariant, and consequently the Rieffel correspondence preserves gauge invariance of ideals. Thus questions about the ideal structure of $\mathcal{O}_X$, or about gauge-invariant ideals of $\mathcal{O}_X$, may be reduced to the corresponding questions for ideals in the augmented Cuntz-Pimsner algebra $\mathcal{O}_Y$.

Finally, we mention that in [17, §4] the method of adding tails has proven very useful in the analysis of topological quivers. Topological quivers, which were first introduced in [16, Example 5.4], are generalizations of graphs in which the sets of vertices and edges are replaced by topological spaces. By adding tails to topological quivers in [17] the authors are able to reduce their analyses to the case when there are no sinks, or equivalently, to the case when the left action of the associated C^*-correspondence is injective. This simplifies the proofs of many results for topological quivers and allows one to avoid a number of technicalities.

References

[1] T. Bates, J. H. Hong, I. Raeburn, and W. Szymański, *The ideal structure of C^*-algebras of infinite graphs*, Illinois J. Math 46 (2002), 1159–1176.

[2] T. Bates, D. Pask, I. Raeburn, and W. Szymański, *The C^*-algebras of row-finite graphs*, New York J. Math. 6 (2000), 307–324.

[3] D. Drinen and M. Tomforde, *The C^*-algebras of arbitrary graphs*, Rocky Mountain J. Math., to appear.

[4] R. Exel and M. Laca, *Cuntz-Krieger algebras for infinite matrices*, J. reine angew. Math. 512 (1999), 119–172.

[5] N. Fowler, M. Laca, and I. Raeburn, *The C^*-algebras of infinite graphs*, Proc. Amer. Math. Soc. 8 (2000), 2319–2327.

[6] N. Fowler, P. Muhly, and I. Raeburn, *Representations of Cuntz-Pimsner algebras*, Indiana Univ. Math. J. 52 (2003), 569–605.

[7] N. Fowler and I. Raeburn, *The Toeplitz algebra of a Hilbert bimodule*, Indiana Univ. Math. J. 48 (1999), 155–181.

[8] J. H. Hong and W. Szymański, *Purely infinite Cuntz-Krieger algebras of directed graphs*, preprint (2002).

[9] T. Kajiwara, C. Pinzari, Y. Watatani, *Ideal structure and simplicity of the C^*-algebras generated by Hilbert bimodules*, J. Funct. Anal. 159 (1998), 295–322.

[10] T. Katsura, *A construction of C^*-algebras from C^*-correspondences*, Contemp. Math., Proceedings of the AMS conference "Advances in Quantum Dynamics", to appear.

[11] T. Katsura, *On C^*-algebras associated with C^*-correspondences*, preprint (2003).

[12] T. Katsura, *Ideal structure of C^*-algebras associated with C^*-correspondences*, preprint (2003).

[13] A. Kumjian, D. Pask, and I. Raeburn, *Cuntz-Krieger algebras of directed graphs*, Pacific J. Math. 184 (1998), 161–174.

[14] A. Kumjian, D. Pask, I. Raeburn, and J. Renault, *Graphs, groupoids, and Cuntz-Krieger algebras*, J. Funct. Anal. 144 (1997), 505–541.

[15] P. Muhly and B. Solel, *Tensor algebras over C^*-correspondences (representations, dilations, and C^*-envelopes)*, J. Funct. Anal. 158 (1998), 389–457.

[16] P. Muhly and B. Solel, *On the Morita Equivalence of Tensor algebras*, Proc. London Math. Soc. 81 (2000), 113–168.

[17] P. Muhly and M. Tomforde, *Topological quivers*, preprint (2003).

[18] M. V. Pimsner, *A class of C^*-algebras generalizing both Cuntz-Krieger algebras and crossed products by $\mathbb{Z}$*, Fields Institute Comm. 12 (1997), 189–212.

[19] I. Raeburn and W. Szymański, *Cuntz-Krieger algebras of infinite graphs and matrices*, preprint (2000).

[20] I. Raeburn and D. P. Williams, Morita Equivalence and Continuous-Trace C^*-Algebras, Math. Surveys & Monographs, vol. 60, Amer. Math. Soc., Providence, 1998.

[21] J. Spielberg, *A functorial approach to the C^*-algebras of a graph*, Internat. J. Math. 13 (2002), 245–277.

[22] J. Spielberg, *Semiprojectivity for certain purely infinite C^*-algebras*, preprint (2001).

Paul S. Muhly
Department of Mathematics
University of Iowa
Iowa City
IA 52242-1419
USA
muhly@math.uiowa.edu

Mark Tomforde
Department of Mathematics
University of Iowa
Iowa City
IA 52242-1419
USA
tomforde@math.uiowa.edu

The first author was supported by NSF Grant DMS-0070405 and the second author was supported by NSF Postdoctoral Fellowship DMS-0201960.

Absolute Continuity of the Spectrum of a Schrödinger Operator with a Potential Which is Periodic in Some Directions and Decays in Others

N. Filonov[1] and F. Klopp[2]

Received: January 31, 2004
Revised: March 30, 2004

Communicated by Heinz Siedentop

Abstract. We prove that the spectrum of a Schrödinger operator with a potential which is periodic in certain directions and super-exponentially decaying in the others is purely absolutely continuous. Therefore, we reduce the operator using the Bloch-Floquet-Gelfand transform in the periodic variables, and show that, except for at most a set of quasi-momenta of measure zero, the reduced operators satisfies a limiting absorption principle.

2000 Mathematics Subject Classification: 35J10, 35Q40, 81C10

1 Formulation of the result

There are many papers (see, for example, [1, 9]) devoted to the question of the absolute continuity of the spectrum of differential operators with coefficients periodic in the whole space. In the present article, we consider the situation where the coefficients are periodic in some variables and decay very fast (super-exponentially) when the other variables tend to infinity. The corresponding operator describes the scattering of waves on an infinite membrane or filament. Recently, quite a few studies have been devoted to similar problems, for periodic, quasi-periodic or random surface Hamiltonians (see, e.g. [3, 7, 2]).

[1]N.F.'s research was partially supported by the FNS 2000 "Programme Jeunes Chercheurs".
[2]F.K.'s research was partially supported by the program RIAC 160 at Université Paris 13 and by the FNS 2000 "Programme Jeunes Chercheurs".

Let (x,y) denote the points of the space $\mathbb{R}^{m+d}$. Define $\Omega = \mathbb{R}^m \times (0,2\pi)^d$ and $\langle x\rangle = \sqrt{x^2+1}$. For $a \in \mathbb{R}$, introduce the spaces

$$L_{p,a} = \{f : e^{a\langle x\rangle} f \in L_p(\Omega)\}, \qquad H^2_a = \{f : e^{a\langle x\rangle} f \in H^2(\Omega)\},$$

where $1 \le p \le \infty$ and $H^2(\Omega)$ is the Sobolev space. Our main result is

Theorem 1.1. *Consider in $L_2(\mathbb{R}^{m+d})$ the self-adjoint operator*

$$Hu = -\operatorname{div}(g\nabla u) + Vu \tag{1}$$

and assume that the functions $g: \mathbb{R}^{m+d} \to \mathbb{R}$ and $V: \mathbb{R}^{m+d} \to \mathbb{R}$ satisfy following conditions:

1. *$\forall l \in \mathbb{Z}^d$, $\forall (x,y) \in \mathbb{R}^{m+d}$,*
$$g(x,y+2\pi l) = g(x,y), \quad V(x,y+2\pi l) = V(x,y);$$
2. *there exists $g_0 > 0$ such that $(g-g_0),\ \Delta g,\ V \in L_{\infty,a}$ for any $a > 0$;*
3. *there exists $c_0 > 0$ such that $\forall (x,y) \in \mathbb{R}^{m+d}$, $g(x,y) \ge c_0$.*

Then, the spectrum of H is purely absolutely continuous.

Remark 1.1. *Operators with different values of g_0 differ from one another only by multiplication by a constant; so, without loss of generality, we can and, from now on, do assume that $g_0 = 1$.*

Remark 1.2. *If $V \equiv 0$, (1) is the acoustic operator. If $g \equiv 1$, it is the Schrödinger operator with electric potential V.*

The basic philosophy of our proof is the following. To prove the absolute continuity of the spectrum for periodic operators (i.e., periodic with respect to a non degenerate lattice in $\mathbb{R}^d$), one applies the Floquet-Bloch-Gelfand reduction to the operator and one is left with proving that the Bloch-Floquet-Gelfand eigenvalues must vary with the quasi-momentum i.e., that they cannot be constant on sets of positive measure (see e.g. [9]). If one tries to follow the same line in the case of operators that are only periodic with respect to a sub-lattice, the problem one encounters is that, as the resolvent of the Bloch-Floquet-Gelfand reduction of the operator is not compact, its spectrum may contain continuous components and some Bloch-Floquet-Gelfand eigenvalues may be embedded in these continuous components. The perturbation theory of such embedded eigenvalues (needed to control their behavior in the Bloch quasi-momentum) is more complicated than that of isolated eigenvalues. To obtain a control on these eigenvalues, we use an idea of the theory of resonances (see e.g. [13]): if one analytically dilates Bloch-Floquet-Gelfand reduction of the operator, these embedded eigenvalues become isolated eigenvalues, and thus can be controlled in the usual way.

Let us now briefly sketch our proof. We make the Bloch-Floquet-Gelfand transformation with respect to the periodic variables (see section 3) and get a family of operators $H(k)$ in the cylinder Ω. Then, we consider the corresponding resolvent in suitable weighted spaces. It analytically depends on the quasi-momentum k and the spectral (non real) parameter λ. It turns out that we can extend it analytically with respect to λ from the upper half-plane to the lower one (see Theorem 5.1 below) and thus establish the limit absorption principle. This suffices to prove the absolute continuity of the initial operator (see section 7).

Note that an analytic extension of the resolvent of the operator (1) with coefficients g and V which decay in all directions is constructed in the paper [4] (with $m = 3$, $d = 0$; see also [10] for $g \equiv 1$). In the case of a potential decaying in all directions but one (i.e., if $d = 1$), the analytic extension of the resolvent of the whole operator (1) (not only for the operator $H(k)$ (see section 3)) is investigated in [6] when $g \equiv 1$. Note also that our approach has shown to be useful in the investigation of the perturbation of free operator in the half-plane by δ-like potential concentrated on a line (see [5]); the wave operators are also constructed there.

In section 2, we establish some auxiliary inequalities. In section 3, we define the Floquet-Gelfand transformation and construct an analytic extension of the resolvent of free operator in the cylinder Ω. In sections 4 and 5, we prove a limiting absorption principle for the initial operator in the cylinder. An auxiliary fact from theory of functions is established in section 6. Finally, the proof of Theorem 1.1 is completed in section 7.

We denote by $B_\delta(k_0)$ a ball in real space

$$B_\delta(k_0) = \{k \in \mathbb{R}^d : |k - k_0| < \delta\}$$

and by k_1 the first coordinate of k, $k = (k_1, k')$. We will use the spaces of function in Ω with periodic boundary conditions,

$$\tilde{H}^2 = \left\{ f \in H^2(\Omega) : f\,|_{y_i=0} = f\,|_{y_i=2\pi},\ \frac{\partial f}{\partial y_i}\,|_{y_i=0} = \frac{\partial f}{\partial y_i}\,|_{y_i=2\pi},\ i = 1, \dots, d \right\},$$
$$\tilde{H}^2_{loc} = \left\{ f \in H^2_{loc}(\Omega) : f\,|_{y_i=0} = f\,|_{y_i=2\pi}, \frac{\partial f}{\partial y_i}\,|_{y_i=0} = \frac{\partial f}{\partial y_i}\,|_{y_i=2\pi},\ i = 1, \dots, d \right\}.$$

Finally $B(X, Y)$ is the space of all bounded operators from X to Y, and $B(X) = B(X, X)$, both endowed with their natural topology.

Thanks: the authors are grateful to Prof. P. Kuchment for drawing their attention to the question addressed in the present paper, and to Prof. T. Suslina for useful discussions.

2 Auxiliary estimations

In this section, we assume that the pair $(k_0, \lambda_0) \in \mathbb{R}^{d+1}$ satisfies

$$(k_0 + n)^2 \neq \lambda_0 \qquad \forall n \in \mathbb{Z}^d. \tag{2}$$

The constants in all the inequalities in this section may depend on (k_0, λ_0). The set

$$J = \{n \in \mathbb{Z}^d : (k_0 + n)^2 < \lambda_0\} \tag{3}$$

is finite. In a neighborhood of (k_0, λ_0), the partition of $\mathbb{Z}^d$ into J and $(\mathbb{Z}^d \setminus J)$ is clearly the same. In other words, there exists $\delta = \delta(k_0, \lambda_0) > 0$ such that

$$\text{if } k \in B_\delta(k_0) \text{ and } \lambda \in B_\delta(\lambda_0), \text{ then } \ (k + n)^2 < \lambda \Leftrightarrow n \in J. \tag{4}$$

Choose $\tilde{k} \in B_\delta(k_0)$ with $\tilde{k}_1 \notin \mathbb{Z}$ and put

$$k(\tau) := (\tilde{k}_1 + i\tau, \tilde{k}') \in \mathbb{C}^d, \quad \tau \in \mathbb{R},$$

and

$$M_1 = M_1(k_0, \lambda_0) := (B_\delta(k_0) \cup \{k(\tau)\}_{\tau \in \mathbb{R}}) \times B_\delta(\lambda_0). \tag{5}$$

Lemma 2.1. *There exists $c > 0$ such that, for all $\zeta \in \mathbb{R}^m$, $(k, \lambda) \in M_1$, $n \in \mathbb{Z}^d \setminus J$ and $\tau \in \mathbb{R}$, we have*

$$|\zeta^2 + (k + n)^2 - \lambda| \geq c,$$
$$|\zeta^2 + (k(\tau) + n)^2 - \lambda| \geq c|\tau|.$$

Proof. By virtue of (4), there exists $c > 0$ such that, for $n \in \mathbb{Z} \setminus J$,

$$\forall k \in B_\delta(k_0), \ \forall \lambda \in B_\delta(\lambda_0), \quad (k + n)^2 - \lambda > c.$$

Hence, for $\zeta \in \mathbb{R}^m$, $n \in \mathbb{Z} \setminus J$,

$$\forall k \in B_\delta(k_0), \ \forall \lambda \in B_\delta(\lambda_0), \quad \zeta^2 + (k + n)^2 - \lambda > c.$$

The second inequality is an immediate corollary of our choice of $\tilde{k}_1$ and the equality

$$\mathrm{Im}(\zeta^2 + (k(\tau) + n)^2 - \lambda) = 2(\tilde{k}_1 + n_1)\tau.$$

This completes the proof of Lemma 2.1. □

In the remaining part of this section, we assume $\lambda_0 > 0$. In this case, we will need to change the integration path in the Fourier transformation; we now describe the contour deformation. Fix $\eta > \sqrt{\lambda_0}$ and, let γ be the contour in the complex plane defined as

$$\gamma = \{-\xi + i\eta\}_{\xi \in [\eta, \infty)} \cup \{\alpha(1 - i)\}_{\alpha \in [-\eta, \eta]} \cup \{\xi - i\eta\}_{\xi \in [\eta, \infty)}. \tag{6}$$

The following two assertions are clear.

Lemma 2.2. *If $g \in L_2(\gamma)$ and $\eta_0 > \eta$ then the function*

$$h(t) = e^{-\eta_0|t|} \int_\gamma e^{itz} g(z) dz$$

belongs to $L_2(\mathbb{R})$.

Lemma 2.3. *Let Γ denote the open set between real axis and γ (it consists of two connected components). Let g be an analytic function in Γ such that $g \in C(\overline{\Gamma})$ and $|g(z)| \leq C(1 + |\operatorname{Re} z|)^{-2}$. Then,*

$$\int_{\mathbb{R}} e^{itz} g(z) dz = \int_\gamma e^{itz} g(z) dz \qquad \forall t \in \mathbb{R}.$$

Establish an analogue of Lemma 2.1 for $n \in J$ and $\zeta \in \gamma^m$ i.e., $\zeta = (\zeta_1, \dots, \zeta_m) \in \mathbb{C}^m$, $\zeta_j \in \gamma$.

Lemma 2.4. *Let $\lambda_0 > 0$, $\eta > \sqrt{\lambda_0}$ and γ be defined by (6). There exists $c > 0$ such that, for all $\zeta \in \gamma^m$, $(k, \lambda) \in M_1$, $n \in J$ and $\tau \in \mathbb{R}$, we have*

$$\begin{gathered} |\zeta^2 + (k+n)^2 - \lambda| \geq c, \\ |\zeta^2 + (k(\tau)+n)^2 - \lambda| \geq c|\tau|. \end{gathered} \tag{7}$$

Proof. By virtue of (4), there exists $c > 0$ such that, for $n \in J$,

$$\forall k \in B_\delta(k_0),\ \forall \lambda \in B_\delta(\lambda_0), \quad (k+n)^2 - \lambda < -2c.$$

Hence, for $\zeta \in \gamma^m$ such that $|\zeta| \leq \sqrt{c}$, one has

$$\forall k \in B_\delta(k_0),\ \forall \lambda \in B_\delta(\lambda_0), \quad \operatorname{Re}(\zeta^2 + (k+n)^2 - \lambda) < -c.$$

On the other hand, for $\zeta \in \gamma^m$ such that $|\zeta| \geq \sqrt{c}$, one has

$$\forall k \in B_\delta(k_0),\ \forall \lambda \in B_\delta(\lambda_0), \quad \operatorname{Im}(\zeta^2 + (k+n)^2 - \lambda) < -c$$

if one chooses c sufficiently small. Thus, it remains to prove the second inequality. Therefore, we write

$$\zeta^2 = -2i \sum_p \alpha_p^2 + \sum_q (\xi_q - i\eta)^2,$$

where the indexes p correspond to the coordinates of ζ which are in the middle part of γ (i.e., $|\operatorname{Re} \zeta_p| < \eta$) and the indexes q correspond to the extreme parts of γ (i.e., $|\operatorname{Re} \zeta_q| \geq \eta$); it is possible that there are only indexes p or only q. Without loss of generality, we suppose that, for all q, $\xi_q \geq 0$. Thus,

$$\begin{aligned} \zeta^2 + (k(\tau)+n)^2 - \lambda = \sum_q (\xi_q^2 - \eta^2) + (\tilde{k}+n)^2 - \tau^2 - \lambda \\ + 2i \left(-\sum_p \alpha_p^2 - \sum_q \xi_q \eta + (\tilde{k}_1 + n_1)\tau \right). \end{aligned}$$

Fix some $\sigma \in (\eta^{-1}\sqrt{\lambda_0}, 1)$. If $\sum_q \xi_q \geq \sigma|\tau|$ then,

$$\left|\mathrm{Im}(\zeta^2 + (k(\tau)+n)^2 - \lambda)\right| \geq 2\left(\sigma\eta - |\tilde{k}_1 + n_1|\right)|\tau| > 2(\sigma\eta - \sqrt{\lambda_0})|\tau|,$$

as $(\tilde{k}+n)^2 < \lambda_0$. If $\sum_q \xi_q \leq \sigma|\tau|$ then $\sum_q \xi_q^2 \leq \sigma^2\tau^2$ and

$$\left|\mathrm{Re}(\zeta^2 + (k(\tau)+n)^2 - \lambda)\right| \geq \tau^2 + \lambda - (\tilde{k}+n)^2 - \sigma^2\tau^2 > (1-\sigma^2)\tau^2$$

again by virtue of (4). This completes the proof of Lemma 2.4. □

3 The resolvent of free operator in the cylinder

Let us consider the Floquet-Gelfand transformation

$$(Uf)(k,x,y) = \sum_{l\in\mathbb{Z}^d} e^{i\langle k, y+2\pi l\rangle} f(x, y+2\pi l).$$

It is a unitary operator

$$U : L_2(\mathbb{R}^{m+d}) \to \int_{[0,1)^d}^{\oplus} L_2(\Omega)dk.$$

Introduce the family of operators $(H(k))_{k\in\mathbb{C}^d}$ on the cylinder Ω where for $k \in \mathbb{C}^d$, $\operatorname{Dom} H(k) = \tilde{H}^2$ and

$$H(k) = \left(i\nabla - (0,\overline{k})\right)^* g(x,y)\left(i\nabla - (0,k)\right) + V(x,y). \tag{8}$$

Then, the Schrödinger operator (1) is unitarily equivalent to the direct integral of these operators in Ω:

$$UHU^* = \int_{[0,1)^d}^{\oplus} H(k)dk.$$

In this section, we investigate the free operator

$$A(k) = -\Delta_x + \left(i\nabla_y - \overline{k}\right)^*(i\nabla_y - k) \tag{9}$$

(which corresponds $H(k)$ with $g \equiv 1$, $V \equiv 0$). For $k \in \mathbb{R}^d$ and $\lambda \notin \mathbb{R}$, its resolvent can be expressed as

$$\left((A(k)-\lambda)^{-1}f\right)(x,y) = \sum_{n\in\mathbb{Z}^d}\int_{\mathbb{R}^m} \frac{e^{i\zeta x + iny}(Ff)(\zeta,n)d\zeta}{\zeta^2 + (k+n)^2 - \lambda}, \tag{10}$$

where F denotes the Fourier transformation in the cylinder

$$(Ff)(\zeta,n) = (2\pi)^{-m-d}\int_\Omega e^{-i\zeta x - iny} f(x,y)\,dx\,dy.$$

Let $(k_0,\lambda_0) \in \mathbb{R}^{d+1}$ satisfy (2) and, J and M_1 be defined respectively by formulas (3) and (5) in the previous section.

Lemma 3.1. *There exists $\mathcal{V}_1$, a neighborhood of the set M_1 in $\mathbb{C}^{d+1}$ such that, for $(k,\lambda)\in\mathcal{V}_1$, the operator $R_1(k,\lambda)$ given by*

$$(R_1(k,\lambda)f)(x,y)=\sum_{n\in\mathbb{Z}^d\setminus J}\int_{\mathbb{R}^m}\frac{e^{i\zeta x+iny}(Ff)(\zeta,n)d\zeta}{\zeta^2+(k+n)^2-\lambda}$$

is well defined and is bounded from $L_2(\Omega)$ to $H^2(\Omega)$. The $B(L_2(\Omega),H^2(\Omega))$-valued function $(k,\lambda)\mapsto R_1(k,\lambda)$ is analytic in $\mathcal{V}_1$. For $\tau\neq 0$, the estimate

$$\|R_1(k(\tau),\lambda)\|_{B(L_2(\Omega))}\le C|\tau|^{-1}$$

holds.

Proof. It immediately follows from Lemma 2.1. □

Lemma 3.2. *Let $\lambda_0>0$, $\eta>\sqrt{\lambda_0}$, $a>\eta\sqrt{m}$ and the contour γ be defined by (6). Then, there exists a neighborhood of the set M_1, say $\mathcal{V}_2$, such that, for $(k,\lambda)\in\mathcal{V}_2$, the operator $R_2(k,\lambda)$ given by*

$$(R_2(k,\lambda)f)(x,y)=\sum_{n\in J}\int_\gamma\cdots\int_\gamma\frac{e^{i\zeta x+iny}(Ff)(\zeta,n)}{\zeta^2+(k+n)^2-\lambda}\,d\zeta_1\cdots d\zeta_m \tag{11}$$

is well defined as a bounded operator from $L_{2,a}$ to H^2_{-a}. The $B(L_{2,a},H^2_{-a})$-valued function $(k,\lambda)\mapsto R_2(k,\lambda)$ is analytic in $\mathcal{V}_2$. For $\tau\neq 0$, the estimate

$$\|R_2(k(\tau),\lambda)\|_{B(L_{2,a},\,L_{2,-a})}\le C|\tau|^{-1}$$

holds.

Proof. If $f\in L_{2,a}$ then the function $(Ff)(\cdot,n)$ is square integrable on γ^m. By Lemma 2.4, the denominator in (11) never vanishes for $(k,\lambda)\in M_1$; therefore, in some neighborhood of M_1. So

$$\left|(\zeta^2+(k+n)^2-\lambda)^{-1}e^{i\zeta x+iny}\right|\le C|e^{i\zeta x}|$$

where the constant does not depend on $\zeta\in\gamma^m$ and on x; the same is true for the second derivatives of $(\zeta^2+(k+n)^2-\lambda)^{-1}e^{i\zeta x+iny}$ with respect to (x,y). Hence, $R_2(k,\lambda)\in B(L_{2,a},H^2_{-a})$ by virtue of Lemma 2.2. Estimation (7) yields the estimation for the norm of $R_2(k(\tau),\lambda)$. □

Now, we construct an analytic extension of the resolvent of $A(k)$.

Theorem 3.1. *Let $(k_0,\lambda_0)\in\mathbb{R}^{d+1}$ satisfy (2) and the set M_1 be defined in (5). Then, there exists a neighborhood of M_1in $\mathbb{C}^{d+1}$, say M_0, a real number a and a $B(L_{2,a},H^2_{-a})$-valued function, say $(k,\lambda)\mapsto R_A(k,\lambda)$, defined and analytic in M_0, such that, for $(k,\lambda)\in M_0$, $k\in\mathbb{R}^d$, $\operatorname{Im}\lambda>0$ and $f\in L_{2,a}$, one has*

$$R_A(k,\lambda)f=(A(k)-\lambda)^{-1}f \tag{12}$$

and

$$\|R_A(k(\tau),\lambda)\|_{B(L_{2,a},\,L_{2,-a})}\le C|\tau|^{-1}. \tag{13}$$

Proof. If $\lambda_0 \leq 0$, we can take $R_A = R_1$ (R_1 is constructed in Lemma 3.1; here, $J = \emptyset$ and $a = 0$).
If $\lambda_0 > 0$ then, we put $R_A = R_1 + R_2$, where R_1, R_2 and a are defined in Lemmas 3.1 and 3.2, and M_0 is the intersection of $\mathcal{V}_1$ and $\mathcal{V}_2$ defined respectively in Lemma 3.1 and Lemma 3.2. If $f \in L_{2,a}$ then $(Ff)(\cdot, n)$ is an analytic function in the domain $\{\zeta : |\operatorname{Im}\zeta| < a\}$ and is uniformly bounded on $\{\zeta : |\operatorname{Im}\zeta| \leq \eta\sqrt{m}\}$. If $\zeta \in \overline{\Gamma}^m$ where Γ is the open set between $\mathbb{R}$ and γ (see Lemma 2.3), then, $\operatorname{Im}\zeta^2 \leq 0$; therefore, the integrand in (11) has no poles when $\operatorname{Im}\lambda > 0$. Hence, the integral in right hand side of (10) for $n \in J$ coincides with the corresponding integral in (11) due to Lemma 2.3, and (12) holds.
The estimate (13) is a simple corollary of the estimations of Lemmas 3.1 and 3.2. □

4 Invertibility of operators of type $(I + WR_A)$

Lemma 4.1. *Let $W \in L_{\infty,b}$ for $b > 2a > 0$. Then, the operator of multiplication by W (we will denote it by the same letter) is*

1. *bounded as an operator from $L_{2,-a}$ to $L_{2,a}$;*
2. *compact as an operator from H^2_{-a} to $L_{2,a}$.*

Proof. The first assertion is evident. In order to prove the second it is enough to introduce functions

$$W_\rho(x, y) = \begin{cases} W(x, y), & |x| < \rho, \\ 0, & |x| \geq \rho, \end{cases}$$

and note that the multiplication by W_ρ is a compact operator from H^2_{-a} to $L_{2,a}$ and that

$$\|W - W_\rho\|_{B(L_{2,-a}, L_{2,a})} \to 0$$

when $\rho \to \infty$. □

The next lemma is a well known result from analytic Fredholm theory (see, e.g., [8, 11]).

Lemma 4.2. *Let U be a domain in $\mathbb{C}^p$, $z_0 \in U$. Let $z \mapsto T(z)$ be an analytic function with values in the set of compact operators in some Hilbert space $\mathcal{H}$. Then, there exists a neighborhood U_0 of the point z_0 and an analytic function $h : U_0 \to \mathbb{C}$ such that, for $z \in U_0$,*

$$(I + T(z))^{-1} \text{ exists if and only if } h(z) = 0.$$

Now, we can establish the existence of the inverse of $(I + WR_A)$.

THEOREM 4.1. *Let (k_0, λ_0) satisfy (2), $R_A(k, \lambda)$ and a be defined as in Theorem 3.1. Pick $b > 2a$, and let $(x, y, \lambda) \mapsto W(x, y, \lambda)$ be a function which belongs to $L_{\infty,b}$ for all λ, and is analytic with respect to λ i.e., $\lambda \mapsto W(\cdot, \cdot, \lambda) \in Hol(\mathbb{C}, L_{\infty,b})$.*
Then, there exists $\varepsilon > 0$, an open set $U \subset \mathbb{C}^{d+1}$ such that $B_\varepsilon(k_0) \times B_\varepsilon(\lambda_0) \subset U$, and an analytic function $h : U \to \mathbb{C}$ such that

$$\forall \lambda \in B_\varepsilon(\lambda_0), \quad \exists k \in B_\varepsilon(k_0) \quad \textit{such that} \quad h(k, \lambda) \neq 0, \tag{14}$$

and, for any $(k, \lambda) \in U$, the operator $(I + W(\lambda) R_A(k, \lambda))$ is invertible in $L_{2,a}$ if and only if $h(k, \lambda) \neq 0$.

Proof. Due to Theorem 3.1 and Lemma 4.1, the operator $W(\lambda) R_A(k(\tau), \lambda)$ is compact in $L_{2,a}$ and satisfies the inequality

$$\|W(\lambda) R_A(k(\tau), \lambda)\|_{B(L_{2,a})} \le C|\tau|^{-1}, \quad \forall \lambda \in B_\varepsilon(\lambda_0).$$

Therefore, for $|\tau|$ large enough, the operator $(I + W(\lambda) R_A(k(\tau), \lambda))^{-1}$ exists and is bounded on $L_{2,a}$. The operator-valued function $\lambda \mapsto W(\lambda) R_A(k, \lambda)$ is analytic in M_0 (defined in Theorem 3.1). The analytic Fredholm alternative yields that, for each $\lambda \in B_\varepsilon(\lambda_0)$, one can find $k \in B_\varepsilon(k_0)$ such that the operator $(I + W(\lambda) R_A(k, \lambda))^{-1}$ exists. Now, applying Lemma 4.2 with $\mathcal{H} = L_{2,a}$, $z = (k, \lambda)$ and $T(z) = W R_A$, completes the proof of Theorem 4.1. □

5 THE RESOLVENT OF THE OPERATOR H

We can reduce the general case of operator (1) with a "metric" g to the case of "pure" Schrödinger operator due to the following lemma. This identity (for the totally periodic case) is known (see [1]). We include the proof for the convenience of the reader.

LEMMA 5.1. *Let the operators $H(k)$ and $A(k)$ be defined by (8) and (9) respectively, and let the conditions of Theorem 1.1 be fulfilled with $g_0 = 1$. If $u \in \tilde{H}^2$ then,*

$$(H(k) - \lambda)\, g^{-1/2} u = g^{1/2}\, (A(k) + W(\lambda) - \lambda)\, u,$$

where

$$W(\lambda) = \frac{1}{g} \left(\frac{\Delta g}{2} - \frac{|\nabla g|^2}{4g} + V + \lambda(g - 1) \right). \tag{15}$$

REMARK 5.1. *If $g \equiv 1$ then $W(\lambda) \equiv V$.*

Proof. It is enough to prove the equality

$$\left(i\nabla - (0, \overline{k})\right)^* g\, (i\nabla - (0, k))\, (g^{-1/2} u) = g^{1/2} \left(A(k) + \frac{\Delta g}{2g} - \frac{|\nabla g|^2}{4g^2} \right) u. \tag{16}$$

We have

$$(i\nabla - (0, k))\, (g^{-1/2} u) = i g^{-1/2} \nabla u - \frac{i}{2} g^{-3/2} \nabla g u - (0, k)(g^{-1/2} u).$$

Therefore, the left hand side of (16) is equal to

$$
\begin{aligned}
&(i\nabla - (0,\overline{k}))^* \left(ig^{1/2}\nabla u - \frac{i}{2} g^{-1/2}\nabla g u - (0,k)(g^{1/2}u) \right) \\
&\quad = -g^{1/2}\Delta u + \frac{1}{2}\operatorname{div}(g^{-1/2}\nabla g)u - i\langle k, \nabla_y (g^{1/2}\overline{u})\rangle_{\mathbb{C}} \\
&\qquad\qquad\qquad - ig^{1/2}\langle \nabla_y u, \overline{k}\rangle_{\mathbb{C}} + \frac{i}{2} g^{-1/2}\langle \nabla_y g, \overline{k}\rangle_{\mathbb{C}} u + k^2 g^{1/2} u \\
&\quad = g^{1/2}\left(-\Delta_x u + (i\nabla_y - \overline{k})^*(i\nabla_y - k)u + \frac{1}{2} g^{-1/2}\operatorname{div}(g^{-1/2}\nabla g)u \right).
\end{aligned}
$$

This completes the proof of Lemma 5.1. □

In the following theorem, we describe the meromorphic extension of the resolvent of $H(k)$.

Theorem 5.1. *Let the conditions of Theorem 1.1 be fulfilled, the operator $H(k)$ be defined by (8) and $(k_0,\lambda_0) \in \mathbb{R}^{d+1}$ satisfy (2). Then, there exists numbers $a \geq 0$, $\varepsilon > 0$, a neighborhood U of (k_0,λ_0) in $\mathbb{C}^{d+1}$ containing the set $B_\varepsilon(k_0) \times B_\varepsilon(\lambda_0)$, a function $h \in Hol(U)$ satisfying (14) and an operator-valued function $(k,\lambda) \mapsto R_H(k,\lambda)$ having the following properties:*

1. *R_H is defined on the set $\{(k,\lambda) \in U : h(k,\lambda) \neq 0\}$ and is analytic there;*
2. *for $(k,\lambda) \in U$ such that $h(k,\lambda) \neq 0$, one has $R_H(k,\lambda) \in B(L_{2,a}, L_{2,-a})$;*
3. *for $(k,\lambda) \in U$, $k \in \mathbb{R}^d$, $\operatorname{Im}\lambda > 0$, $f \in L_{2,a}$, one has*

$$
R_H(k,\lambda)f = (H(k) - \lambda)^{-1} f. \tag{17}
$$

Remark 5.2. *It will be seen from the proof that $R_H(k,\lambda) \in B(L_{2,a}, H^2_{-a})$ though we do not need this fact.*

Proof. By the assumptions of Theorem 1.1, for any $b > 0$, $\nabla g \in L_{\infty,b}$. So, if we define $W(\lambda)$ by (15), for any $b > 0$, $W(\lambda) \in L_{\infty,b}$. We can thus apply Theorem 4.1. Let U, h, a and R_A be as in this theorem. On the set where $h(k,\lambda) \neq 0$, we put

$$
R_H(k,\lambda) = g^{-1/2} R_A(k,\lambda) \left(I + W(\lambda)R_A(k,\lambda)\right)^{-1} g^{-1/2}.
$$

By Theorem 4.1, $R_H(k,\lambda) \in B(L_{2,a}, H^2_{-a})$. Let $f \in L_{2,a}$. Then,

$$
\left(I + W(\lambda)R_A(k,\lambda)\right)^{-1} g^{-1/2} f \in L_{2,a} \tag{18}
$$

and we can apply Lemma 5.1 to the function

$$
u = R_A(k,\lambda) \left(I + W(\lambda)R_A(k,\lambda)\right)^{-1} g^{-1/2} f \in H^2_{-a}, \tag{19}
$$

so

$$(H(k)-\lambda)R_H(k,\lambda)f = g^{1/2}\left(A(k)+W(\lambda)-\lambda\right)u. \tag{20}$$

For real k and non real λ, we have by (12) and (18)

$$(A(k)-\lambda)u = \left(I+W(\lambda)R_A(k,\lambda)\right)^{-1} g^{-1/2}f,$$

hence, by (19),

$$\left(A(k)+W(\lambda)-\lambda\right)u = g^{-1/2}f,$$

and, finally, by (20)

$$(H(k)-\lambda)R_H(k,\lambda)f = f. \tag{21}$$

For $\operatorname{Im}\lambda > 0$, the operators $(H(k)-\lambda)^{-1}$ and $(A(k)-\lambda)^{-1}$ are well defined in $L_2(\Omega)$. As $R_H(k,\lambda)f \in L_2(\Omega)$, (21) gives $R_H(k,\lambda)f = (H(k)-\lambda)^{-1}f$. This completes the proof of Theorem 5.1. $\square$

6 One fact from the theory of functions

Lemma 6.1. *Let V be an open subset of $\mathbb{R}^d$. Let f be a real-analytic function in a box $(c,d)\times V$. Let Λ be a subset of V of measure zero,* $\operatorname{mes}\Lambda = 0$. *Then,*

$$\operatorname{mes}\{k\in(c,d) : \exists\lambda\in\Lambda \textit{ s.t. } f(k,\lambda)=0 \textit{ and } \partial_k f(k,\lambda)\neq 0\} = 0. \tag{22}$$

Proof. The Implicit Function Theorem implies that, for any point (k_0,λ_0) such that $f(k_0,\lambda_0)=0\neq\partial_k f(k_0,\lambda_0)$, we can find rational numbers $\tilde k_0$, $\tilde r_0>0$, a vector $\tilde\lambda_0 = (\tilde\lambda_0^1,\cdots,\tilde\lambda_0^d)$ with rational coordinates, and a cube $C_{\tilde r_0}(\tilde k_0,\tilde\lambda_0)$ where

$$\begin{gathered}(k_0,\lambda_0)\in C_{\tilde r_0}(\tilde k_0,\tilde\lambda_0) = (\tilde k_0-\tilde r_0,\tilde k_0+\tilde r_0)\times C_{\tilde r_0}(\tilde\lambda_0)\subset (c,d)\times V\\ C_{\tilde r_0}(\tilde\lambda_0) = (\tilde\lambda_0^1-\tilde r_0,\tilde\lambda_0^1+\tilde r_0)\times\cdots\times(\tilde\lambda_0^d-\tilde r_0,\tilde\lambda_0^d+\tilde r_0)\end{gathered}$$

and a real analytic function $\theta:\ C_{\tilde r_0}(\tilde\lambda_0)\to(\tilde k_0-\tilde r_0,\tilde k_0+\tilde r_0)$ such that

1. $\theta(\lambda_0)=k_0$;
2. $f(k,\lambda)=0\Leftrightarrow\theta(\lambda)=k$ if $(k,\lambda)\in C_{\tilde r_0}(\tilde k_0,\tilde\lambda_0)$.

Therefore,

$$\begin{aligned}\operatorname{mes}\{k:\exists\lambda\in\Lambda \text{ s.t. } (k,\lambda)\in C_{\tilde r_0}(\tilde k_0,\tilde\lambda_0) \text{ and } f(k,\lambda)=0\}\\ \le \operatorname{mes}\theta(\Lambda\cap C_{\tilde r_0}(\tilde\lambda_0)) = 0.\end{aligned}$$

The set

$$\{(k,\lambda): f(k,\lambda)=0 \text{ and } \partial_k f(k,\lambda)\neq 0\}$$

can be covered by a countable number of cubes $C_{\tilde r}(\tilde k,\tilde\lambda)$ constructed as above, say $(C_{\tilde r_i}(\tilde k_i,\tilde\lambda_i))_{i\in\mathbb{N}}$; hence, the measure of the set under consideration in (22) is also equal to zero as

$$\begin{aligned}\{k\in(c,d):\exists\lambda\in\Lambda \text{ s.t. } f(k,\lambda)=0 \text{ and } \partial_k f(k,\lambda)\neq 0\}\\ \subset\bigcup_{i\in\mathbb{N}}\operatorname{mes}\{k:\exists\lambda\in\Lambda \text{ s.t. } (k,\lambda)\in C_{\tilde r_i}(\tilde k_i,\tilde\lambda_i) \text{ and } f(k,\lambda)=0\}.\end{aligned}$$

This completes the proof of Lemma 6.1. □

Lemma 6.1 has a multidimensional analogue.

Lemma 6.2. *Let U be an open subset of $\mathbb{R}^d$, and V be an open subset of $\mathbb{R}^{d'}$. Let f be a real-analytic function on the set $U\times V$, and pick $\Lambda\subset V$ such that* $\operatorname{mes}\Lambda=0$. *For $k\in U$, we write $k=(k_1,k')$ where k_1 is real and $k'\in\mathbb{R}^{d-1}$. Then,*

$$\operatorname{mes}\{k\in U:\ \exists\lambda\in\Lambda \textit{ s.t. } f(k,\lambda)=0 \textit{ and } \partial_{k_1}f(k,\lambda)\neq 0\}=0. \tag{23}$$

Proof. Cover U with countably many open sets of the form $(a,b)\times\tilde U$ i.e.,

$$U=\bigcup_{i\in\mathbb{N}}(a_i,b_i)\times\tilde U_i.$$

For $i\in\mathbb{N}$, one has

$$\begin{aligned}\{k\in(a_i,b_i)\times\tilde U_i:\ \exists\lambda\in\Lambda \text{ s.t. } f(k,\lambda)=0 \text{ and } \partial_{k_1}f(k,\lambda)\neq 0\}\\ \subset\{k_1\in(a_i,b_i):\ \exists\lambda\in\Lambda \text{ s.t. } f(k_1,k',\lambda)=0 \text{ and}\\ \partial_{k_1}f(k_1,k',\lambda)\neq 0\}\times\tilde U_i.\end{aligned} \tag{24}$$

By Lemma 6.1, the set in the right hand side of equation (24) has measure 0 (as $\tilde U_i\times\Lambda$ has measure zero in $\mathbb{R}^{d+d'-1}$). As

$$\begin{aligned}\{k\in U:\ \exists\lambda\in\Lambda \text{ s.t. } f(k,\lambda)=0 \text{ and } \partial_{k_1}f(k,\lambda)\neq 0\}\\ =\bigcup_{i\in\mathbb{N}}\{k\in(a_i,b_i)\times\tilde U_i:\ \exists\lambda\in\Lambda \text{ s.t. } f(k,\lambda)=0 \text{ and } \partial_{k_1}f(k,\lambda)\neq 0\},\end{aligned}$$

(23) holds, which completes the proof of Lemma 6.2. □

Finally, we prove

Theorem 6.1. *Let U be a region in $\mathbb{R}^d$, Λ be a subset of an interval (a,b) such that* $\operatorname{mes}\Lambda=0$. *Let h be a real-analytic function defined on the set $U\times(a,b)$ and suppose that*

$$\forall\lambda\in\Lambda\quad\exists k\in U\quad\textit{such that}\quad h(k,\lambda)\neq 0. \tag{25}$$

Then,

$$\operatorname{mes}\{k\in U:\ \exists\lambda\in\Lambda \textit{ s.t. } h(k,\lambda)=0\}=0.$$

Proof. For any $k \in U$ and $\lambda \in \Lambda$, by assumption (25), there exists a multi-index $\alpha \in \mathbb{Z}_+^d$ such that $\partial_k^\alpha h(k,\lambda) \neq 0$. Therefore,

$$\begin{aligned}&\{k \in U : h(k,\lambda) = 0 \text{ for some } \lambda \in \Lambda\}\\ &\qquad \subset \bigcup_{j=1}^{d} \bigcup_{\alpha \in \mathbb{Z}_+^d} \left\{k \in U : \partial_k^\alpha h(k,\lambda) = 0, \partial_{k_j}\partial_k^\alpha h(k,\lambda) \neq 0 \text{ for some } \lambda \in \Lambda\right\}.\end{aligned}$$

Reference to Lemma 6.2 then completes the proof of Theorem 6.1. □

7 The proof of Theorem 1.1

The following lemma is well known (see for example [12]).

Lemma 7.1. *Fix $b > 0$. Let B be a self-adjoint operator in $L_2(\Omega)$. Suppose that R_B is an analytic function defined in a complex neighborhood of an interval $[\alpha,\beta]$ except at a finite number of points $\{\mu_1,\ldots,\mu_N\}$, that the values of R_B are in $B(L_{2,b}, L_{2,-b})$ and that*

$$R_B(\lambda)\varphi = (B-\lambda)^{-1}\varphi \quad \textit{if } \operatorname{Im}\lambda > 0, \varphi \in L_{2,b}.$$

Then, the spectrum of B in the set $[\alpha,\beta] \setminus \{\mu_1,\ldots,\mu_N\}$ is absolutely continuous. If $\Lambda \subset [\alpha,\beta]$, $\operatorname{mes}\Lambda = 0$ and $\mu_j \notin \Lambda$, $j = 1,\ldots,N$, then $E_B(\Lambda) = 0$, where E_B is the spectral projector of B.

Proof of Theorem 1.1. By Theorem 5.1, the set of all points $(k,\lambda) \in \mathbb{R}^{d+1}$ satisfying (2) can be represented as the following union

$$\{(k,\lambda) \in \mathbb{R}^{d+1} \text{ s.t. (2) be satisfied}\} = \bigcup_{j=1}^{\infty} B_{\varepsilon_j}(k_j) \times B_{\varepsilon_j}(\lambda_j), \tag{26}$$

where, for every j, there exists

- a number $a_j \geq 0$,
- an analytic scalar function h_j defined in a complex neighborhood of $\overline{B_{\varepsilon_j}(k_j) \times B_{\varepsilon_j}(\lambda_j)}$ with the property
$$\forall \lambda \in B_{\varepsilon_j}(\lambda_j) \quad \exists k \in B_{\varepsilon_j}(k_j) \quad \text{such that} \quad h_j(k,\lambda) \neq 0,$$
- an analytic $B(L_{2,a_j}, L_{2,-a_j})$-valued function $R_H^{(j)}$ defined on the set where $h_j(k,\lambda) \neq 0$ and satisfying (17).

Now, pick $\Lambda \subset \mathbb{R}$ such that $\operatorname{mes}\Lambda = 0$. Set

$$\begin{aligned} K_0 &= \{k \in [0,1]^d : (k+n)^2 = \lambda \text{ for some } n \in \mathbb{Z}^d, \lambda \in \Lambda\},\\ K_1 &= \{k \in [0,1]^d : h_j(k,\lambda) = 0 \text{ for some } j \in \mathbb{N}, \lambda \in \Lambda\}.\end{aligned}$$

Thanks to Theorem 6.1, we know

$$\operatorname{mes} K_0 = \operatorname{mes} K_1 = 0. \tag{27}$$

For $k \notin K_0$, denote

$$\Lambda_j(k) = \{\lambda \in \Lambda : (k, \lambda) \in B_{\varepsilon_j}(k_j) \times B_{\varepsilon_j}(\lambda_j)\}.$$

It is clear that $\Lambda_j(k) \subset (\lambda_j - \varepsilon_j, \lambda_j + \varepsilon_j)$, $\operatorname{mes} \Lambda_j(k) = 0$, and, by (26),

$$\Lambda = \bigcup_{j=1}^{\infty} \Lambda_j(k) \quad \forall k \notin K_0. \tag{28}$$

If $k \notin (K_0 \cup K_1)$ and $\Lambda_j(k) \neq \emptyset$ then $h_j(k, \lambda) \neq 0$ for $\lambda \in \Lambda_j(k)$ and $\lambda \mapsto h_j(k, \lambda)$ has at most a finite number of zeros in $[\lambda_j - \varepsilon_j, \lambda_j + \varepsilon_j]$. So we can apply Lemma 7.1; therefore,

$$E_{H(k)}(\Lambda_j(k)) = 0 \quad \forall j.$$

This and (28) implies that

$$E_{H(k)}(\Lambda) = 0.$$

Finally, one computes

$$E_H(\Lambda) = \int_{[0,1]^d} E_{H(k)}(\Lambda)\, dk = \int_{[0,1]^d \setminus K_0 \setminus K_1} E_{H(k)}(\Lambda)\, dk = 0$$

by virtue of (27). So, we proved that the spectral resolution of H vanishes on any set of Lebesgue measure 0, which means, by definition, that the spectrum of the operator H is purely absolutely continuous.

References

[1] M. Sh. Birman, T. A. Suslina, *Periodic magnetic Hamiltonian with variable metrics. Problem of absolute continuity*, Algebra i Analiz, vol. 11 (1999), 2, pp. 1–40 (Russian). English translation in St. Petersburg Math. J. 11 (2000), no. 2, pp. 203–232.

[2] A. Boutet de Monvel, P. Stollmann, *Dynamical localization for continuum random surface models*, Arch. Math., 80 (2003), pp. 87–97.

[3] A. Chahrour, J. Sahbani, *On the spectral and scattering theory of the Schrödinger operator with surface potential*, Rev. Math. Phys., 12 (2000), pp. 561–573.

[4] D. M. Eidus, *Some boundary value problems in infinite domains*, Izvestiya AN SSSR 27 (1963), 5, pp. 1055–1080 (Russian).

[5] R. Frank, R. Shterenberg, *On the scattering theory of the Laplacian with a periodic boundary condition. Additional channels of scattering*, Doc. Math., Vol. 9 (2004), pp. 57–77.

[6] C. Gérard, *Resonance theory in atomic-surface scattering*, Comm. Math. Phys., 126 (1989), pp. 263–290.

[7] V. Jakšić, Y. Last, *Corrugated surfaces and a.c. spectrum*, Rev. Math. Phys., 12 (2000), pp. 1465–1503.

[8] T. Kato. *Perturbation Theory for Linear Operators.* Springer Verlag, Berlin, 1980.

[9] P. Kuchment, *Floquet theory for partial differential equations*, Birkhäuser, Basel (1993).

[10] O. A. Ladyzhenskaya, *On the principle of limit amplitude*, Uspekhi Mat. Nauk, XII N3, 75 (1957), pp. 161–164 (Russian).

[11] M. Reed and B. Simon. *Methods of Modern Mathematical Physics, Vol I: Functional Analysis.* Academic Press, New-York, 1980.

[12] D. R. Yafaev, *Mathematical scattering theory*, Providence, Rhode Island, AMS, 1992.

[13] M. Zworski, *Quantum resonances and partial differential equations*, Proceedings of the ICM, Beijing 2002, vol. 3, pp. 243–254.

N. Filonov
Department of
Mathematical Physics
St Petersburg
State University
198504 St Petersburg-
Petrodvorets
Russia
filonov@mph.phys.spbu.ru

F. Klopp
LAGA, Institut Galilée
Université Paris-Nord
F-93430 Villetaneuse
France
klopp@math.univ-paris13.fr

Dihedral Galois Representations and Katz Modular Forms

Gabor Wiese

Received: February 23, 2004
Revised: March 19, 2004

Communicated by Don Blasius

Abstract. We show that any two-dimensional odd dihedral representation ρ over a finite field of characteristic $p > 0$ of the absolute Galois group of the rational numbers can be obtained from a Katz modular form of level N, character ϵ and weight k, where N is the conductor, ϵ is the prime-to-p part of the determinant and k is the so-called minimal weight of ρ. In particular, $k = 1$ if and only if ρ is unramified at p. Direct arguments are used in the exceptional cases, where general results on weight and level lowering are not available.

2000 Mathematics Subject Classification: 11F11, 11F80, 14G35

1 Introduction

In [S1] Serre conjectured that any odd irreducible continuous Galois representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ for a prime p comes from a modular form in characteristic p of a certain level N_ρ, weight $k_\rho \geq 2$ and character ϵ_ρ. Later Edixhoven discussed in [E2] a slightly modified definition of weight, the so-called *minimal weight*, denoted $k(\rho)$, by invoking Katz' theory of modular forms. In particular, one has that $k(\rho) = 1$ if and only if ρ is unramified at p.
The present note contains a proof of this conjecture for *dihedral representations*. We define those to be the continuous irreducible Galois representations that are induced from a character of the absolute Galois group of a quadratic number field. Let us mention that this is equivalent to imposing that the projective image is isomorphic to a dihedral group D_n with $n \geq 3$.

Theorem 1 *Let p be a prime and $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ an odd dihedral representation. As in [S1] define N_ρ to be the conductor of ρ and ϵ_ρ to be the prime-to-p part of* $\det \circ \rho$ *(considered as a character of $(\mathbb{Z}/(N_\rho p)\mathbb{Z})^*$). Define $k(\rho)$ as in [E2].*
Then there exists a normalised Katz eigenform $f \in \mathcal{S}_{k(\rho)}(\Gamma_1(N_\rho), \epsilon_\rho, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$, whose associated Galois representation ρ_f is isomorphic to ρ.

We will on the one hand show directly that ρ comes from a Katz modular form of level N_ρ, character ϵ_ρ and minimal weight $k(\rho) = 1$, if ρ is unramified at p. If on the other hand ρ is ramified at p, we will finish the proof by applying the fundamental work by Ribet, Edixhoven, Diamond, Buzzard and others on "weight and level lowering" (see Theorem 10).
Let us recall that in weight at least 2 every Katz modular form on Γ_1 is classical, i.e. a reduction from a characteristic zero form of the same level and weight. Hence multiplying by the Hasse invariant, if necessary, it follows from Theorem 1 that every odd dihedral representation as above also comes from a classical modular form of level N_ρ and Serre's weight k_ρ. However, if one also wants the character to be ϵ_ρ, one has to exclude in case $p = 2$ that ρ is induced from $\mathbb{Q}(i)$ and in case $p = 3$ that ρ is induced from $\mathbb{Q}(\sqrt{-3})$ (see [B], Corollary 2.7, and [D], Corollary 1.2).
Edixhoven's theorem on weight lowering ([E2], Theorem 4.5) states that modularity in level N_ρ and the modified weight $k(\rho)$ follows from modularity in level N_ρ and Serre's weight k_ρ, unless one is in a so-called *exceptional case*. A representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ is called *exceptional* if the semi-simplification of its restriction to a decomposition group at p is the sum of two copies of an unramified character. Because of work by Coleman and Voloch the only open case left is that of characteristic 2 (see the introduction of [E2]).
Exceptionality at 2 is a common phenomenon for mod 2 dihedral representations. One way to construct examples is to consider the Hilbert class field H of a quadratic field K that is unramified at 2 and has a non-trivial class group. One lets ρ_K be the dihedral representation obtained by induction to $G_{\mathbb{Q}}$ of a mod 2 character of the Galois group of $H|K$. If the prime 2 stays inert in $\mathcal{O}_K$, then $2\mathcal{O}_K$ splits completely in H and the order of $\rho_K(\mathrm{Frob}_2)$ is 2, where Frob_2 is a Frobenius element at 2. Consequently, ρ_K is exceptional. An example for this behaviour is provided by $K = \mathbb{Q}(\sqrt{229})$. If the prime 2 splits in $\mathcal{O}_K$ and the primes of $\mathcal{O}_K$ lying above 2 are principal, then $\rho_K(\mathrm{Frob}_2)$ is the identity and hence ρ_K is exceptional. This happens for example for $K = \mathbb{Q}(\sqrt{2089})$.
Let us point out that some of the weight one forms that we obtain cannot be lifted to characteristic zero forms of weight one and the same level, so that the theory of modular forms by Katz becomes necessary. Namely, if $p = 2$ and the dihedral representation in question has odd conductor N and is induced from a real quadratic field K of discriminant N, whose fundamental units have norm -1, then there does not exist an odd characteristic zero representation with conductor dividing N that reduces to ρ. The representation coming from the quadratic field $\mathbb{Q}(\sqrt{229})$ used above, can also here serve as an example.

The fact that dihedral representations come from *some* modular form is well-known (apparently already due to Hecke). So the subtle issue is to adjust the level, character and weight. It should be noted that Rohrlich and Tunnell solved many cases for $p = 2$ with Serre's weight k_ρ by rather elementary means in [R-T], however, with the more restrictive definition of a dihedral representation to be such that its image in $\mathrm{GL}_2(\overline{\mathbb{F}_2})$, and not in $\mathrm{PGL}_2(\overline{\mathbb{F}_2})$, is isomorphic to a dihedral group.
Let us also mention that it is possible to do computations of weight one forms in positive characteristic on a computer (see [W]) and thus to collect evidence for Serre's conjecture in some cases.
This note is organised as follows. The number theoretic ingredients on dihedral representations are provided in Section 2. In Section 3 some results on oldforms, also in positive characteristic, are collected. Section 4 is devoted to the proof of Theorem 1. Finally, in Section 5 we include a result on the irreducibility of certain mod p representations.

I wish to thank Peter Stevenhagen for helpful discussions and comments and especially Bas Edixhoven for invaluable explanations and his constant support.

2 Dihedral representations

We shall first recall some facts on Galois representations. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}(V)$ be a continuous representation with V a 2-dimensional vector space over an algebraically closed discrete field k.
Let L be the number field such that $\mathrm{Ker}(\rho) = G_L$ (by the notation G_L we always mean the absolute Galois group of L). Given a prime Λ of L dividing the rational prime l, we denote by $G_{\Lambda,i}$ the i-th ramification group in lower numbering of the local extension $L_\Lambda|\mathbb{Q}_l$. Furthermore, one sets

$$n_l(\rho) = \sum_{i\geq 0} \frac{\dim(V/V^{G_{\Lambda,i}})}{(G_{\Lambda,0} : G_{\Lambda,i})}.$$

This number is an integer, which is independent of the choice of the prime Λ above l. With this one defines the *conductor* of ρ to be $\mathfrak{f}(\rho) = \prod_l l^{n_l(\rho)}$, where the product runs over all primes l different from the characteristic of k. If k is the field of complex numbers, $\mathfrak{f}(\rho)$ coincides with the *Artin conductor*.
Let ρ be a dihedral representation. Then ρ is induced from a character $\chi : G_K \to k^*$ for a quadratic number field K such that $\chi \neq \chi^\sigma$, with $\chi^\sigma(g) = \chi(\sigma^{-1} g \sigma)$ for all $g \in G_K$, where σ is a lift to $G_{\mathbb{Q}}$ of the non-trivial element of $G_{K|\mathbb{Q}}$. For a suitable choice of basis we then have the following explicit description of ρ: If an unramified prime l splits in K as $\Lambda\sigma(\Lambda)$, then $\rho(\mathrm{Frob}_l) = \begin{pmatrix} \chi(\mathrm{Frob}_\Lambda) & 0 \\ 0 & \chi^\sigma(\mathrm{Frob}_\Lambda) \end{pmatrix}$. Moreover, $\rho(\sigma)$ is represented by the matrix $\begin{pmatrix} 0 & 1 \\ \chi(\sigma^2) & 0 \end{pmatrix}$. As ρ is continuous, its image is a finite group, say, of order m.

Lemma 2 *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ be an odd dihedral representation that is unramified at p. Define K, χ, σ and m as above. Let N be the conductor of ρ. Let ζ_m a primitive m-th root of unity and $\mathfrak{P}$ a prime of $\mathbb{Q}(\zeta_m)$ above p.*
Then one of the following two statements holds.

(a) There exists an odd dihedral representation $\widehat{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}[\zeta_m])$, which has Artin conductor N and reduces to ρ modulo $\mathfrak{P}$.

(b) One has that $p = 2$ and K is real quadratic. Moreover, there is an infinite set S of primes such that for each $l \in S$ the trace of $\rho(\mathrm{Frob}_l)$ is zero, and there exists an odd dihedral representation $\widehat{\rho} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}[\zeta_m])$, which has Artin conductor Nl and reduces to ρ modulo $\mathfrak{P}$.

Proof. Suppose that the quadratic field K equals $\mathbb{Q}(\sqrt{D})$ with D square-free. The character $\chi : G_K \to k^*$ can be uniquely lifted to a character $\widetilde{\chi} : G_K \to \mathbb{Z}[\zeta_m]^*$ of the same order, which reduces to χ modulo $\mathfrak{P}$. Denote by $\widetilde{\rho}$ the continuous representation $\mathrm{Ind}_{G_K}^{G_{\mathbb{Q}}} \widetilde{\chi}$. For the choice of basis discussed above the matrices representing ρ can be lifted to matrices representing $\widetilde{\rho}$, whose non-zero entries are in the m-th roots of unity. Then for any open subgroup H of $G_{\mathbb{Q}}$, one has that $(\overline{\mathbb{F}_p}^2)^{\rho(H)}$ is isomorphic to $(\mathbb{Z}[\zeta_m]^2)^{\widetilde{\rho}(H)} \otimes \overline{\mathbb{F}_p}$. Hence the conductor of ρ equals the Artin conductor of $\widetilde{\rho}$, as $\widetilde{\rho}$ is unramified at p. Alternatively, one can first remark that the conductor of χ equals the conductor of $\widetilde{\chi}$ and then use the formulae $\mathfrak{f}(\rho) = \mathrm{Norm}_{K|\mathbb{Q}}(\mathfrak{f}(\chi))D$ and $\mathfrak{f}(\widetilde{\rho}) = \mathrm{Norm}_{K|\mathbb{Q}}(\mathfrak{f}(\widetilde{\chi}))D$.
Thus condition (a) is satisfied if $\widetilde{\rho}$ is odd. Let us now consider the case when $\widetilde{\rho}$ is even. This immediately implies $p = 2$ and that the quadratic field K is real, as is the number field L whose absolute Galois group G_L equals the kernel of ρ, and hence also the kernel of $\widetilde{\chi}$. We shall now adapt "Serre's trick" from [R-T], p. 307, to our situation.
Let $\mathfrak{f}$ be the conductor of $\widetilde{\chi}$. As L is totally real, $\mathfrak{f}$ is a finite ideal of $\mathcal{O}_K$. Via class field theory, $\widetilde{\chi}$ can be identified with a complex character of $\mathrm{CL}_K^{\mathfrak{f}}$, the ray class group modulo $\mathfrak{f}$. Let ∞_1, ∞_2 be the infinite places of K. Consider the class

$$c = [\{(\lambda) \in \mathrm{CL}_K^{4D\mathfrak{f}\infty_1\infty_2} \mid \mathrm{Norm}(\lambda) < 0, \lambda \equiv 1 \bmod 4D\mathfrak{f}\}]$$

in the ray class group of K modulo $4D\mathfrak{f}\infty_1\infty_2$. By Cebotarev's density theorem the primes of $\mathcal{O}_K$ are uniformly distributed over the conjugacy classes of $\mathrm{CL}_K^{4D\mathfrak{f}\infty_1\infty_2}$. Hence, there are infinitely many primes Λ of degree 1 in the class c. Take S to be the set of rational primes lying under them. Let a prime Λ from the class c be given. It is principal, say $\Lambda = (\lambda)$, and coprime to $4D\mathfrak{f}$. By construction we have $c^2 = [\Lambda^2] = 1$. As $\mathrm{CL}_K^{\mathfrak{f}}$ is a quotient of $\mathrm{CL}_K^{4D\mathfrak{f}\infty_1\infty_2}$, the class of Λ in $\mathrm{CL}_K^{\mathfrak{f}}$ has order 1 or 2. Since $p = 2$, the character χ has odd order and we conclude that $\chi(\Lambda) = 1$.
We have $\lambda \equiv 1 \bmod 4D\mathfrak{f}$ and $\mathrm{Norm}(\lambda) = -l$ for some odd prime l. Hence, the extension $K(\sqrt{\lambda})$ has two real and two complex embeddings and is unramified at 2 and at the primes dividing $D\mathfrak{f}$. We represent $K(\sqrt{\lambda})$ by the quadratic character $\xi : G_K \to \{\pm 1\}$. For the complex conjugation, the "infinite Frobenius

element", Frob_{∞_1}, we have that $\xi(\mathrm{Frob}_{\infty_1})\xi^\sigma(\mathrm{Frob}_{\infty_1}) = -1$. We now consider the representation $\widehat{\rho}$ obtained by induction from the character $\widehat{\chi} = \widetilde{\chi}\xi$. Using the same basis as in the discussion at the beginning of this section, an element g of G_K is represented by the matrix $\begin{pmatrix} \widetilde{\chi}(g)\xi(g) & 0 \\ 0 & \widetilde{\chi}^\sigma(g)\xi^\sigma(g) \end{pmatrix}$. In particular, we obtain that the determinant of Frob_∞ over $\mathbb{Q}$ equals -1, whence $\widehat{\rho}$ is odd. Moreover, as l splits in K, one has that $\rho(\mathrm{Frob}_l)$ is the identity matrix, so that the trace of $\rho(\mathrm{Frob}_l)$ is zero.
The reduction of $\widehat{\rho}$ equals ρ, as ξ is trivial in characteristic 2. Moreover, outside Λ the conductor of $\widehat{\chi}$ equals the conductor of $\widetilde{\chi}$. At the prime Λ the local conductor of $\widehat{\chi}$ is Λ, as the ramification is tame. Consequently, the Artin conductor of $\widehat{\rho}$ equals Nl. $\square$

Also without the condition that it is unramified at p, one can lift a dihedral representation to characteristic zero, however, losing control of the Artin conductor.

Lemma 3 *Let $\rho : G_\mathbb{Q} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ be an odd dihedral representation. Define K, χ, m, ζ_m and $\mathfrak{P}$ as in the previous lemma.*
There exists an odd dihedral representation $\widehat{\rho} : G_\mathbb{Q} \to \mathrm{GL}_2(\mathbb{Z}[\zeta_m])$, whose reduction modulo $\mathfrak{P}$ is isomorphic to ρ.

Proof. We proceed as in the preceding lemma for the definitions of $\widetilde{\chi}$ and $\widetilde{\rho}$. If $\widetilde{\rho}$ is even, then $p = 2$ and K is real. In that case we choose some $\lambda \in \mathcal{O}_K - \mathbb{Z}$, which satisfies $\mathrm{Norm}(\lambda) < 0$. The field $K(\sqrt{\lambda})$ then has two real and two complex embeddings and gives a character $\xi : G_K \to \mathbb{Z}[\zeta_m]^*$. As in the proof of the preceding lemma one obtains that the representation $\widehat{\rho} = \mathrm{Ind}_{G_K}^{G_\mathbb{Q}} \widetilde{\chi}\xi$ is odd and reduces to ρ modulo $\mathfrak{P}$. $\square$

3 On oldforms

In this section we collect some results on oldforms. We try to stay as much as possible in the characteristic zero setting. However, we also need a result on Katz modular forms.

Proposition 4 *Let N, k, r be positive integers, p a prime and ϵ a Dirichlet character of modulus N. The homomorphism*

$$\phi_{p^r}^N : \big(\mathcal{S}_k(\Gamma_1(N), \epsilon, \mathbb{C})\big)^{r+1} \hookrightarrow \mathcal{S}_k(\Gamma_1(Np^r), \epsilon, \mathbb{C}), \quad (f_0, f_1, \ldots, f_r) \mapsto \sum_{i=0}^{r} f_i(q^{p^i})$$

is compatible with all Hecke operators T_n with $(n, p) = 1$.
Let $f \in \mathcal{S}_k(\Gamma_1(N), \epsilon, \mathbb{C})$ be a normalised eigenform for all Hecke operators. Then the forms $f(q), f(q^{p^2}), \ldots, f(q^{p^r})$ in the image of $\phi_{p^r}^N$ are linearly independent, and on their span the action of the operator T_p in level Np^r is given

by the matrix

$$\begin{pmatrix} a_p(f) & 1 & 0 & 0 & \dots & 0 \\ -\delta p^{k-1}\epsilon(p) & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ & & \vdots & & & \\ 0 & \dots & 0 & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 & 0 \end{pmatrix},$$

where $\delta = 1$ *if* $p \nmid N$ *and* $\delta = 0$ *otherwise.*

Proof. The embedding map and its compatibility with the Hecke action away from p is explained in [D-I], Section 6.1. The linear independence can be checked on q-expansions. Finally, the matrix can be elementarily computed. □

Corollary 5 *Let* p *be a prime,* $r \geq 0$ *some integer and* $f \in \mathcal{S}_k(\Gamma_1(Np^r), \epsilon, \mathbb{C})$ *an eigenform for all Hecke operators. Then there exists an eigenform for all Hecke operators* $\tilde{f} \in \mathcal{S}_k(\Gamma_1(Np^{r+2}), \epsilon, \mathbb{C})$*, which satisfies* $a_l(\tilde{f}) = a_l(f)$ *for all primes* $l \neq p$ *and* $a_p(\tilde{f}) = 0$*.*

Proof. One computes the characteristic polynomial of the operator T_p of Proposition 4 and sees that it has 0 as a root if the dimension of the matrix is at least 3. Hence one can choose the desired eigenform $\tilde{f}$ in the image of $\phi_{p^2}^{Np^r}$. □

As explained in the introduction, Katz' theory of modular forms ought to be used in the study of Serre's conjecture. Following [E3], we briefly recall this concept, which was introduced by Katz in [K]. However, we shall use a "non-compactified" version.

Let $N \geq 1$ be an integer and R a ring, in which N is invertible. One defines the category $[\Gamma_1(N)]_R$, whose objects are pairs $(E/S/R, \alpha)$, where S is an R-scheme, E/S an elliptic curve (i.e. a proper smooth morphism of R-schemes, whose geometric fibres are connected smooth curves of genus one, together with a section, the "zero section", $0 : S \to E$) and $\alpha : (\mathbb{Z}/N\mathbb{Z})_S \to E[N]$, the *level structure*, is an embedding of S-group schemes. The morphisms in the category are cartesian diagrams

$$\begin{array}{ccc} E' & \longrightarrow & E \\ \downarrow & \square & \downarrow \\ S' & \longrightarrow & S, \end{array}$$

which are compatible with the zero sections and the level structures. For every such elliptic curve $E/S/R$ we let $\underline{\omega}_{E/S} = 0^*\underline{\Omega}_{E/S}$. For every morphism $\pi : E'/S'/R \to E/S/R$ the induced map $\underline{\omega}_{E'/S'} \to \pi^*\underline{\omega}_{E/S}$ is an isomorphism.

A *Katz cusp form* $f \in \mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ assigns to every object $(E/S/R, \alpha)$ of $[\Gamma_1(N)]_R$ an element $f(E/S/R, \alpha) \in \underline{\omega}_{E/S}^{\otimes k}(S)$, compatibly for the morphisms in

the category, subject to the condition that all q-expansions (which one obtains by adjoining all N-th roots of unity and plugging in a suitable Tate curve) only have positive terms.
For the following definition let us remark that if $m \geq 1$ is coprime to N and is invertible in R, then any morphism of group schemes of the form $\phi_{Nm} : (\mathbb{Z}/Nm\mathbb{Z})_S \to E[Nm]$ can be uniquely written as $\phi_N \times_S \phi_m$ with $\phi_N : (\mathbb{Z}/N\mathbb{Z})_S \to E[N]$ and $\phi_m : (\mathbb{Z}/m\mathbb{Z})_S \to E[m]$.

DEFINITION 6 *A Katz modular form* $f \in \mathcal{S}_k(\Gamma_1(Nm), R)_{\mathrm{Katz}}$ *is called* independent of m *if for all elliptic curves* $E/S/R$, *all* $\phi_N : (\mathbb{Z}/N)_S \hookrightarrow E[N]$ *and all* $\phi_m, \phi'_m : (\mathbb{Z}/m)_S \hookrightarrow E[m]$ *one has the equality*

$$f(E/S/R, \phi_N \times_S \phi_m) = f(E/S/R, \phi_N \times_S \phi'_m) \in \underline{\omega}_{E/S}^{\otimes k}(S).$$

PROPOSITION 7 *Let* N, m *be coprime positive integers and* R *a ring, which contains the* Nm*-th roots of unity and* $\frac{1}{Nm}$. *A Katz modular form* $f \in \mathcal{S}_k(\Gamma_1(Nm), R)_{\mathrm{Katz}}$ *is independent of* m *if and only if there exists a Katz modular form* $g \in \mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ *such that*

$$f(E/S/R, \phi_{Nm}) = g(E/S/R, \phi_{Nm} \circ \psi)$$

for all elliptic curves $E/S/R$ *and all* $\phi_{Nm} : (\mathbb{Z}/Nm\mathbb{Z})_S \hookrightarrow E[Nm]$. *Here* ψ *denotes the canonical embedding* $(\mathbb{Z}/N\mathbb{Z})_S \hookrightarrow (\mathbb{Z}/Nm\mathbb{Z})_S$ *of* S*-group schemes. In that case,* f *and* g *have the same* q*-expansion at* ∞.

PROOF. If $m = 1$, there is nothing to do. If necessary replacing m by m^2, we can hence assume that m is at least 3.
Let us now consider the category $[\Gamma_1(N; m)]_R$, whose objects are triples $(E/S/R, \phi_N, \psi_m)$, where S is an R scheme, E/S an elliptic curve, $\phi_N : (\mathbb{Z}/N\mathbb{Z})_S \hookrightarrow E[N]$ an embedding of group schemes and $\psi_m(\mathbb{Z}/m\mathbb{Z})_S^2 \cong E[m]$ an isomorphism of group schemes. The morphisms are cartesian diagrams compatible with the zero sections, the ϕ_N and the ψ_m as before.
We can pull back the form $f \in \mathcal{S}_k(\Gamma_1(Nm), R)_{\mathrm{Katz}}$ to a Katz form h on $[\Gamma_1(N; m)]_R$ as follows. First let $\beta : (\mathbb{Z}/m\mathbb{Z})_S \hookrightarrow (\mathbb{Z}/m\mathbb{Z})_S^2$ be the embedding of S-group schemes defined by mapping onto the first factor. Using this, f gives rise to h by setting

$$h((E/S/R, \phi_N, \psi_m)) = f((E/S/R, \phi_N, \psi_m \circ \beta)) \in \underline{\omega}_{E/S}^{\otimes k}(S).$$

As f is independent of m, it is clear that h is independent of ψ_m and thus invariant under the natural $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$-action.
As $m \geq 3$, one knows that the category $[\Gamma_1(N; m)]_R$ has a final object $(E^{\mathrm{univ}}/Y_1(N; m)_R/R, \alpha^{\mathrm{univ}})$. In other words, h is an $\mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z})$-invariant global section of $\underline{\omega}_{E^{\mathrm{univ}}/Y_1(N;m)_R}^{\otimes k}$. Since this R-module is equal to $\mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ (see e.g. Equation 1.2 of [E3], p. 210), we find some $g \in \mathcal{S}_k(\Gamma_1(N), R)_{\mathrm{Katz}}$ such that $f(E/S/R, \phi_{Nm}) = g(E/S/R, \phi_{Nm} \circ \psi)$ for all $(E/S/R, \phi_{Nm})$.

Plugging in the Tate curve, one sees that the standard q-expansions of f and g coincide. $\square$

Corollary 8 *Let N, m be coprime positive integers, p a prime not dividing Nm and $\epsilon : (\mathbb{Z}/N\mathbb{Z})^* \to \overline{\mathbb{F}_p}$ a character. Let $f \in \mathcal{S}_k(\Gamma_1(Nm), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ be a Katz cuspidal eigenform for all Hecke operators.*
If f is independent of m, then there exists an eigenform for all Hecke operators $g \in \mathcal{S}_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ such that the associated Galois representations ρ_f and ρ_g are isomorphic.

Proof. From the preceding proposition we get a modular form $g \in \mathcal{S}_k(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$, noting that the character is automatically good. Because of the compatibility of the embedding map with the operators T_l for primes $l \nmid m$, we find that g is an eigenform for these operators. As the operators T_l for primes $l \nmid m$ commute with the others, we can choose a form of the desired type. $\square$

4 Proof of the principal result

We first cover the weight one case.

Theorem 9 *Let p be a prime and $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ an odd dihedral representation of conductor N, which is unramified at p. Let ϵ denote the character $\det \circ \rho$.*
Then there exists a Katz eigenform f in $\mathcal{S}_1(\Gamma_1(N), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$, whose associated Galois representation is isomorphic to ρ.

Proof. Assume first that part (a) of Lemma 2 applies to ρ, and let $\widehat{\rho}$ be a lift provided by that lemma. A theorem by Weil-Langlands (Theorem 1 of [S2]) implies the existence of a newform g in $\mathcal{S}_1(\Gamma_1(N), \det \circ \widehat{\rho}, \mathbb{C})$, whose associated Galois representation is isomorphic to $\widehat{\rho}$. Now reduction modulo a suitable prime above p yields the desired modular form. In particular, one does not need Katz' theory in this case.
If part (a) of Lemma 2 does not apply, then part (b) does, and we let S be the infinite set of primes provided. For each $l \in S$ the theorem of Weil-Langlands yields a newform $f^{(l)}$ in $\mathcal{S}_1(\Gamma_1(Nl), \mathbb{C})$, whose associated Galois representation reduces to ρ modulo $\mathfrak{P}$, where $\mathfrak{P}$ is the ideal from the lemma. Moreover, the congruence $a_q(f^{(l)}) \equiv 0 \bmod \mathfrak{P}$ holds for all primes $q \in S$ different from l.
From Corollary 5 we obtain Hecke eigenforms $\widetilde{f}^{(l)} \in \mathcal{S}_1(\Gamma_1(Nl^3), \mathbb{C})$ such that $a_l(\widetilde{f}^{(l)}) = 0$ and $a_q(\widetilde{f}^{(l)}) = a_q(f^{(l)}) \equiv 0 \bmod \mathfrak{P}$ for all primes $q \in S$, $q \neq l$. Reducing modulo the prime ideal $\mathfrak{P}$, we get eigenforms $g^{(l)} \in \mathcal{S}_1(\Gamma_1(Nl^3), \epsilon, \overline{\mathbb{F}_p})$, whose associated Galois representations are isomorphic to ρ. One also has $a_q(g^{(l)}) = 0$ for all $q \in S$.
The coefficients $a_q(f^{(l)})$ for all primes $q \mid N$ appear in the L-series of the complex representation $\rho_{f^{(l)}}$ associated to $f^{(l)}$. As the image of $\rho_{f^{(l)}}$ is isomorphic

to a fixed finite group G, not depending on l, there are only finitely many possibilities for the value of $a_q(f^{(l)})$. Hence the same holds for the $g^{(l)}$. Consequently, there are two forms $g_1 = g^{(l_1)}$ and $g_2 = g^{(l_2)}$ for $l_1 \neq l_2$ that have the same coefficients at all primes $q \mid N$. For primes $q \nmid Nl_1l_2$ one has that the trace of $\rho_{f^{(l_1)}}(\mathrm{Frob}_q)$ is congruent to the trace of $\rho_{f^{(l_2)}}(\mathrm{Frob}_q)$, whence $a_q(g_1) = a_q(g_2)$. Let us point out that this includes the case $q = p = 2$, as the complex representation is unramified at p.
In the next step we embed g_1 and g_2 into $\mathcal{S}_1(\Gamma_1(Nl_1^3l_2^3), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ via the method in the statement of Proposition 7. As the q-expansions coincide, g_1 and g_2 are mapped to the same form h. But as h comes from g_2, it is independent of l_1 and analogously also of l_2. Since $\rho_h = \rho$, Theorem 9 follows immediately from Corollary 8. $\square$

We will deduce the cases of weight at least two from general results. The current state of the art in "level and weight lowering" seems to be the following theorem.

Theorem 10 *[Ribet, Edixhoven, Diamond, Buzzard,...] Let p be a prime and $\rho : G_\mathbb{Q} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ a continuous irreducible representation, which is assumed to come from some modular form. Define k_ρ and N_ρ as in [S1]. If $p = 2$, additionally assume either (i) that the restriction of ρ to a decomposition group at 2 is not contained within the scalar matrices or (ii) that ρ is ramified at 2. Then there exists a normalised eigenform $f \in \mathcal{S}_{k_\rho}(\Gamma_1(N_\rho), \overline{\mathbb{F}_p})$ giving rise to ρ.*

Proof. The case $p \neq 2$ is Theorem 1.1 of [D], and the case $p = 2$ with condition (i) follows from Propositions 1.3 and 2.4 and Theorem 3.2 of [B], multiplying by the Hasse invariant if necessary.
We now show that if $p = 2$ and ρ restricted to a decomposition group $G_{\mathbb{Q}_2}$ at 2 is contained within the scalar matrices, then ρ is unramified at 2. Let $\phi : G_\mathbb{Q} \to \overline{\mathbb{F}_2}^*$ be the character such that $\phi^2 = \det \circ \rho$. As ϕ has odd order, it is unramified at 2 because of the Kronecker-Weber theorem. If ρ restricted to $G_{\mathbb{Q}_2}$ is contained within the scalar matrices, then we have that $\rho|_{G_{\mathbb{Q}_2}}$ is $\begin{pmatrix} \phi|_{G_{\mathbb{Q}_2}} & 0 \\ 0 & \phi|_{G_{\mathbb{Q}_2}} \end{pmatrix}$, whence ρ is unramified at 2. $\square$

Proof of theorem 1. Let ρ be the dihedral representation from the assertion. If ρ is unramified at p, one has $k(\rho) = 1$, and Theorem 1 follows from Theorem 9.
If ρ is ramified at p, then let $\widehat{\rho}$ be a characteristic zero representation lifting ρ, as provided by Lemma 3. The theorem by Weil-Langlands already used above (Theorem 1 of [S2]) implies the existence of a newform in weight one and characteristic zero giving rise to $\widehat{\rho}$. So from Theorem 10 we obtain that ρ comes from a modular form of Serre's weight k_ρ and level N_ρ. Let us note that using Katz modular forms the character is automatically the conjectured one ϵ_ρ.
The weights k_ρ and $k(\rho)$ only differ in two cases (see [E2], remark 4.4). The first case is when $k(\rho) = 1$. The other case is when $p = 2$ and ρ is not finite

at 2. Then one has $k(\rho) = 3$ and $k_\rho = 4$. In that case one applies Theorem 3.4 of [E2] to obtain an eigenform of the same level and character in weight 3, or one applies Theorem 3.2 of [B] directly. $\square$

5 An irreducibility result

We first study the relation between the level of an eigenform in characteristic p and the conductor of the associated Galois representation.

Lemma 11 *Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\overline{\mathbb{F}_p})$ be a continuous representation of conductor N, and let k be a positive integer. If $f \in \mathcal{S}_k(\Gamma_1(M), \epsilon, \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ is a Hecke eigenform giving rise to ρ, then N divides M.*

Proof. By multiplying with the Hasse invariant, if necessary, we can assume that the weight is at least 2. Hence the form f can be lifted to characteristic zero (see e.g. [D-I], Theorem 12.3.2) in the same level. Thus there exists a newform g, say of level L, whose Galois representation ρ_g reduces to ρ. Now Proposition 0.1 of [L] yields that N divides L. As L divides M, the lemma follows. $\square$

We can derive the following proposition, which is of independent interest.

Proposition 12 *Let $f \in \mathcal{S}_k(\Gamma_0(N), \overline{\mathbb{F}_p})_{\mathrm{Katz}}$ be a normalised Hecke eigenform for a square-free level N with $p \nmid N$ in some weight $k \geq 1$.*

(a) If $p = 2$, the associated Galois representation is either irreducible or trivial.

(b) For any prime p the associated Galois representation is either irreducible or corresponds to a direct sum $\alpha \oplus \chi_p^{k-1}\alpha^{-1}$, where χ_p is the mod p cyclotomic character and α is a character factoring through $G(\mathbb{Q}(\zeta_p)|\mathbb{Q})$ for a primitive p-th root of unity ζ_p.

Proof. Let us assume that the representation ρ associated to f is reducible. Since ρ is semi-simple, it is isomorphic to the direct sum of two characters $\alpha \oplus \beta$. As the determinant is the $(k-1)$-th power of the mod p cyclotomic character χ_p, we have that $\beta = \chi_p^{k-1}\alpha^{-1}$. Since the conductor of χ_p^{k-1} is 1, it follows that the conductor of α equals that of β. Consequently, the conductor of ρ is the square of the conductor of α. Lemma 11 implies that the conductor of ρ divides N. As we have assumed this number to be square-free, we have that ρ can only ramify at p.
The number field L with $G_L = \mathrm{Ker}(\rho)$ is abelian. As only p can be ramified, it follows that L is contained in $\mathbb{Q}(\zeta_{p^n})$ for some p^n-th root of unity. Since the order of α is prime to p, we conclude that α factors through $G(\mathbb{Q}(\zeta_p)|\mathbb{Q})$. In characteristic $p = 2$ this implies that ρ is the trivial representation, as χ_2 is the trivial character. $\square$

References

[B] K. Buzzard. *On level lowering for mod 2 representations.* Mathematical Research Letters 7 (2000). 95-110.

[D] F. Diamond *The refined conjecture of Serre.* Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993), Internat. Press, Cambridge, MA, 1995. 22-37.

[D-I] F. Diamond and J. Im. *Modular forms and modular curves.* Seminar on Fermat's Last Theorem (Toronto, ON, 1993-1994), CMS Conf. Proc., 17, Amer. Math. Soc., Providence, RI, 1995. 39-133.

[E1] S. J. Edixhoven. *Comparison of integral structures on spaces of modular forms of weight two, and computation of spaces of forms mod* 2 *of weight one.* Accepted for publication in the Journal de l'Institut de Mathmatiques de Jussieu. `http://arxiv.org/abs/math.NT/0312019`

[E2] S. J. Edixhoven. *The weight in Serre's conjectures on modular forms.* Invent. Math. 109 (1992). 563-594.

[E3] S. J. Edixhoven. *Serre's conjecture.* Modular Forms and Fermat's Last Theorem (Gary Cornell, Joseph Silverman and Glenn Stevens, editors). Springer-Verlag, 1997. 209-242.

[K] N. M. Katz. *p-adic properties of modular schemes and modular forms.* Modular functions of one variable, III (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972). Lecture Notes in Mathematics, Vol. 350, Springer, Berlin, 1973. 69-190.

[L] R. Livn. *On the conductors of mod l Galois representations coming from modular forms.* J. of Number Theory 31 (1989). 133-141.

[R-T] D. E. Rohrlich, J. B. Tunnell. *An elementary case of Serre's conjecture.* Pacific Journal of Mathematics 181, No. 3 (1997). 299-309.

[S1] J.-P. Serre. *Sur les reprsentations de degr* 2 *de* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Duke Mathematical Journal 54, No. 1 (1987). 179-230.

[S2] J.-P. Serre. *Modular forms of weight one and Galois representations.* Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975), Academic Press, London, 1977. 193-268.

[W] G. Wiese. *Computing Hecke algebras of weight* 1 *in MAGMA.* Appendix B of [E1].

Gabor Wiese
Mathematisch Instituut
Universiteit Leiden
Postbus 9512
2300 RA Leiden
The Netherlands
gabor@math.leidenuniv.nl

Erratum to the paper "Absolute Continuity of the Spectrum of a Schrödinger Operator with a Potential Which is Periodic in Some Directions and Decays in Others"

Documenta Math Vol. 9 (2004), 107-121

N. Filonov[1] and F. Klopp[2]

Received: April 9, 2004

Communicated by Heinz Siedentop

Abstract. Lemma 6.1 and 6.2 in [1] are false as stated there. Below, we correct the proof of Theorem 6.1 accordingly.

2000 Mathematics Subject Classification: 35J10, 35Q40, 81C10

6 One fact from the theory of functions

Lemma 6.1. *Let U be an open subset of $\mathbb{R}^d$. Let f be a real-analytic function on the set $U \times (a,b)$, and pick $\Lambda \subset (a,b)$ such that* $\operatorname{mes} \Lambda = 0$. *Then*

$$\operatorname{mes}\{k \in U : \ \exists \lambda \in \Lambda \ s.t. \ f(k,\lambda) = 0 \ and \ \partial_{k_1} f(k,\lambda) \neq 0\} = 0. \tag{1}$$

Proof. The Implicit Function Theorem implies that, for any point (k^*,λ^*) such that $f(k^*,\lambda^*) = 0 \neq \partial_{k_1} f(k,\lambda^*)$, we can find rational numbers $\tilde{r} > 0$, $\tilde{\lambda}$, a vector $\tilde{k}$ with rational coordinates, and a real analytic function θ defined in $B_{\tilde{r}}(\tilde{k}', \tilde{\lambda})$ such that

1. $(k^*,\lambda^*) \in B_{\tilde{r}}(\tilde{k},\tilde{\lambda})$;

[1]N.F.'s research was partially supported by the FNS 2000 "Programme Jeunes Chercheurs".

[2]F.K.'s research was partially supported by the program RIAC 160 at Université Paris 13 and by the FNS 2000 "Programme Jeunes Chercheurs".

2. $\theta((k^*)', \lambda) = k_1^*$;

3. $f(k, \lambda) = 0 \Leftrightarrow \theta(k', \lambda) = k_1$ if $(k, \lambda) \in B_{\tilde{r}}(\tilde{k}, \tilde{\lambda})$.

The Jacobian of the map

$$(k', \lambda) \mapsto (\theta(k', \lambda), k')$$

is bounded, so

$$\text{mes}\{(\theta(k', \lambda), k') : (k', \lambda) \in B_{\tilde{r}}(\tilde{k}', \tilde{\lambda}), \lambda \in \Lambda\} = 0,$$

and therefore,

$$\text{mes}\{k : \exists \lambda \in \Lambda \text{ s.t. } (k, \lambda) \in B_{\tilde{r}}(\tilde{k}, \tilde{\lambda}) \text{ and } f(k, \lambda) = 0\} = 0.$$

The set

$$\{(k, \lambda) : f(k, \lambda) = 0 \text{ and } \partial_{k_1} f(k, \lambda) \neq 0\}$$

can be covered by a countable number of balls $B_{\tilde{r}_i}(\tilde{k}_i, \tilde{\lambda}_i)$ constructed as above, hence the measure of the set in (1) is also equal to zero. □

Theorem 6.1. *Let U be a region in $\mathbb{R}^d$, Λ be a subset of an interval (a, b) such that* $\text{mes}\, \Lambda = 0$. *Let h be a real-analytic function defined on the set $U \times (a, b)$ and suppose that*

$$\forall \lambda \in \Lambda \quad \exists k \in U \quad \textit{such that} \quad h(k, \lambda) \neq 0. \tag{2}$$

Then,

$$\text{mes}\{k \in U : \ \exists \lambda \in \Lambda \textit{ s.t. } h(k, \lambda) = 0\} = 0.$$

Proof. The proof of Theorem 6.1 is that given in [1] except that one uses Lemma 6.1. □

References

[1] N. Filonov and F. Klopp. Absolute continuity of the spectrum of a Schrödinger operator with a potential which is periodic in some directions and decays in others. *Documenta Mathematica*, 9:107–121, 2004.

N. Filonov
Department of Mathematical Physics
St Petersburg State University
198504 St Petersburg-Petrodvorets
Russia
filonov@mph.phys.spbu.ru

F. Klopp
LAGA, Institut Galilée
Université Paris-Nord
F-93430 Villetaneuse
France
klopp@math.univ-paris13.fr

The Free Cover of a Row Contraction

William Arveson

Received: March 30, 2004

Communicated by Joachim Cuntz

Abstract. We establish the existence and uniqueness of finite free resolutions - and their attendant Betti numbers - for graded commuting d-tuples of Hilbert space operators. Our approach is based on the notion of *free cover* of a (perhaps noncommutative) row contraction. Free covers provide a flexible replacement for minimal dilations that is better suited for higher-dimensional operator theory.

For example, every graded d-contraction that is finitely multi-cyclic has a unique free cover of finite type - whose kernel is a Hilbert module inheriting the same properties. This contrasts sharply with what can be achieved by way of dilation theory (see Remark 2.5).

2000 Mathematics Subject Classification: 46L07, 47A99
Keywords and Phrases: Free Resolutions, Multivariable Operator Theory

1. Introduction

The central result of this paper establishes the existence and uniqueness of finite free resolutions for commuting d-tuples of operators acting on a common Hilbert space (Theorem 2.6). Commutativity is essential for that result, since finite resolutions do not exist for noncommuting d-tuples.

On the other hand, we base the existence of free resolutions on a general notion of *free cover* that is effective in a broader noncommutative context. Since free covers have applications that go beyond the immediate needs of this paper, and since we intend to take up such applications elsewhere, we present the general version below (Theorem 2.4). In the following section we give precise statements of these two results, we comment on how one passes from the larger noncommutative category to the commutative one, and we relate these results to previous work that has appeared in the literature. Section 3 concerns generators for Hilbert modules, in which we show that the examples of primary interest are properly generated. The next two sections are devoted to proofs of the main results - the existence and uniqueness of free covers and of finite free resolutions. In Section 6 we discuss examples.

In large part, this paper was composed during a visit to Kyoto University that was supported by MEXT, the Japanese Ministry of Education, Culture, Science and Technology. It is a pleasure to thank Masaki Izumi for providing warm hospitality and a stimulating intellectual environment during this period.

2. Statement of results

A *row contraction* of dimension d is a d-tuple of operators $(T_1, \dots, T_d)$ acting on a common Hilbert space H that has norm at most 1 when viewed as an operator from $H \oplus \cdots \oplus H$ to H. A *d-contraction* is a row contraction whose operators mutually commute, $T_j T_k = T_k T_j$, $1 \le j, k \le d$. In either case, one can view H as a module over the noncommutative polynomial algebra $\mathbb{C}\langle z_d, \dots, z_d\rangle$ by way of

$$f \cdot \xi = f(T_1, \dots, T_d)\xi, \qquad f \in \mathbb{C}\langle z_1, \dots, z_d\rangle,$$

and H becomes a *contractive* Hilbert module in the sense that

$$\|z_1 \cdot \xi_1 + \cdots + z_d \cdot \xi_d\|^2 \le \|\xi_1\|^2 + \cdots + \|\xi_d\|^2, \qquad \xi_1, \dots, \xi_d \in H.$$

The maps of this category are linear operators $A \in \mathcal{B}(H_1, H_2)$ that are homomorphisms of the module structure and satisfy $\|A\| \le 1$. It will be convenient to refer to a Hilbert space endowed with such a module structure simply as a *Hilbert module.*

Associated with every Hilbert module H there is an integer invariant that we shall call the defect, defined as follows. Let $Z \cdot H$ denote the closure of the range of the coordinate operators

$$Z \cdot H = \{z_1\xi_1 + \cdots + z_d\xi_d : \xi_1, \dots, \xi_d \in H\}^-.$$

$Z \cdot H$ is a closed submodule of H, hence the quotient $H/(Z \cdot H)$ is a Hilbert module whose row contraction is $(0, \dots, 0)$. One can identify $H/(Z \cdot H)$ more concretely as a subspace of H in terms of the ambient operators $T_1, \dots, T_d$,

$$H/(Z \cdot H) \sim H \ominus (Z \cdot H) = \ker T_1^* \cap \cdots \cap \ker T_d^*.$$

Definition 2.1. A Hilbert module H is said to be *properly generated* if $H \ominus (Z \cdot H)$ is a generator:

$$H = \overline{\operatorname{span}}\{f \cdot \zeta : f \in \mathbb{C}\langle z_1, \dots, z_d\rangle \quad \zeta \in H \ominus (Z \cdot H)\}.$$

In general, the quotient Hilbert space $H/(Z \cdot H)$ is called the *defect space* of H and its dimension $\dim(H/(Z \cdot H))$ is called the *defect.*

The defect space of a finitely generated Hilbert module must be finite-dimensional. Indeed, it is not hard to see that the defect is dominated by the smallest possible number of generators. A fuller discussion of properly generated Hilbert modules of finite defect will be found in Section 3.

The free objects of this category are defined as follows. Let Z be a Hilbert space of dimension $d = 1, 2, \dots$ and let $F^2(Z)$ be the Fock space over Z,

$$F^2(Z) = \mathbb{C} \oplus Z \oplus Z^{\otimes 2} \oplus Z^{\otimes 3} \oplus \cdots$$

where $Z^{\otimes n}$ denotes the full tensor product of n copies of Z. One can view $F^2(Z)$ as the completion of the tensor algebra over Z in a natural Hilbert space norm; in turn, if we fix an orthonormal basis $e_1, \dots, e_d$ for Z then we can define an isomorphism of the noncommutative polynomial algebra $\mathbb{C}\langle z_1, \dots, z_d\rangle$ onto the tensor algebra by sending z_k to e_k, $k = 1, \dots, d$. This allows us to realize the Fock space as a completion of $\mathbb{C}\langle z_1, \dots, z_d\rangle$, on which the multiplication operators associated with the coordinates $z_1, \dots, z_d$ act as a row contraction. We write this Hilbert module as $F^2\langle z_1, \dots, z_d\rangle$; and when there is no possibility of confusion about the dimension or choice of basis, we often use the more compact F^2.

One forms free Hilbert modules of higher multiplicity by taking direct sums of copies of F^2. More explicitly, let C be a Hilbert space of dimension $r = 1, 2, \dots, \infty$ and consider the Hilbert space $F^2 \otimes C$. There is a unique Hilbert module structure on $F^2 \otimes C$ satisfying

$$f \cdot (\xi \otimes \zeta) = (f \cdot \xi) \otimes \zeta, \qquad f \in \mathbb{C}\langle z_1, \dots, z_d\rangle, \quad \xi \in F^2, \quad \zeta \in C,$$

making $F^2 \otimes C$ into a properly generated Hilbert module of defect r.

More generally, it is apparent that any homomorphism of Hilbert modules $A : H_1 \to H_2$ induces a contraction

$$\dot{A} : H_1/(Z \cdot H_1) \to H_2/(Z \cdot H_2)$$

that maps one defect space into the other.

Definition 2.2. Let H be a Hilbert module. By a *cover* of H we mean a contractive homomorphism of Hilbert modules $A : F \to H$ that has dense range and induces a unitary operator $\dot{A} : F/(Z \cdot F) \to H/(Z \cdot H)$ from one defect space onto the other. A *free cover* of H is a cover $A : F \to H$ in which $F = F^2\langle z_1, \dots, z_d\rangle \otimes C$ is a free Hilbert module.

Remark 2.3 (The Extremal Property of Covers). In general, if one is given a contractive homomorphism with dense range $A : F \to H$, there is no way of relating the image $A(F \ominus (Z \cdot F))$ to $H \ominus (Z \cdot H)$, even when A induces a bijection $\dot{A}$ of one defect space onto the other. But since a cover is a contraction that induces a *unitary* map of defect spaces, it follows that a cover must map $F \ominus (Z \cdot F)$ isometrically onto $H \ominus (Z \cdot H)$ (see Lemma 4.1). This extremal property is critical, leading for example to the uniqueness assertion of Theorem 2.4 below.

It is not hard to give examples of finitely generated Hilbert modules H that are degenerate in the sense that $Z \cdot H = H$ (see the proof of Proposition 3.4), and in such cases, free covers $A : F \to H$ cannot exist when $H \neq \{0\}$. As we will see momentarily, the notion of free cover is effective for Hilbert modules that are *properly* generated. We emphasize that in a free cover $A : F \to H$ of a finitely generated Hilbert module H with $F = F^2 \otimes C$,

$$\dim C = \operatorname{defect}(F^2 \otimes C) = \operatorname{defect} H < \infty,$$

so that for finitely generated Hilbert modules for which a free cover exists, *the free module associated with a free cover must be of finite defect.* More generally, we say that a diagram of Hilbert modules

$$F \underset{A}{\longrightarrow} G \underset{B}{\longrightarrow} H$$

is *weakly exact* at G if $AF \subset \ker B$ and the map $A : F \to \ker B$ defines a cover of $\ker B$. This implies that AF is dense in $\ker B$, but of course it asserts somewhat more.

Any cover $A : F \to H$ of H can be converted into another one by composing it with a unitary automorphism of F on the right. Two covers $A : F_A \to H$ and $B : F_B \to H$ are said to be *equivalent* if there is a unitary isomorphism of Hilbert modules $U : F_A \to F_B$ such that $B = AU$. Notice the one-sided nature of this relation; in particular, two equivalent covers of a Hilbert module H must have identical (non-closed) ranges. When combined with Proposition 3.2 below, the following result gives an effective characterization of the existence of free covers.

Theorem 2.4. *A contractive Hilbert module H over the noncommutative polynomial algebra $\mathbb{C}\langle z_1, \dots, z_d\rangle$ has a free cover if, and only if, it is properly generated; and in that case all free covers of H are equivalent.*

Remark 2.5 (The Rigidity of Dilation Theory). Let H be a pure, finitely generated, contractive Hilbert module over $\mathbb{C}\langle z_1, \dots, z_d\rangle$ (see [Arv98]). The methods of dilation theory lead to the fact that, up to unitary equivalence of Hilbert modules, H can be realized as a quotient of a free Hilbert module

$$H = (F^2\langle z_1, \dots, z_d\rangle \otimes C)/M$$

where M is an invariant subspace of $F^2 \otimes C$. In more explicit terms, there is a contractive homomorphism $L : F^2 \otimes C \to H$ of Hilbert modules such that $LL^* = \mathbf{1}_H$. When such a realization is minimal, there is an appropriate sense in which it is unique.

The problem with this realization of H as a quotient of a free Hilbert module is that the coefficient space C is often infinite-dimensional; moreover, the connecting map L is only rarely a cover. Indeed, in order for C to be finite-dimensional it is necessary and sufficient that the "defect operator" of H, namely

$$\Delta = (\mathbf{1}_H - T_1T_1^* - \cdots - T_dT_d^*)^{1/2}, \tag{2.1}$$

should be of finite rank. The fact is that this finiteness condition often fails, even when the underlying operators of H commute.

For example, any invariant subspace $K \subseteq H^2$ of the rank-one free commutative Hilbert module H^2, that is also invariant under the gauge group Γ_0 (see the following paragraphs), becomes a finitely generated graded Hilbert module whose operators $T_1, \dots, T_d$ are the restrictions of the d-shift to K. However, the defect operator of such a K is of infinite rank in every nontrivial case - namely, whenever K is nonzero and of infinite codimension in H^2. Thus, even though dilation theory provides a realization of K as the quotient of another free commutative Hilbert module $K \cong (H^2 \otimes C)/M$, the free Hilbert module $H^2 \otimes C$ must have *infinite defect*.

One may conclude from these observations that dilation theory is too rigid to provide an effective representation of finitely generated Hilbert modules as quotients of free modules of finite defect, and a straightforward application of dilation theory cannot lead to finite free resolutions in multivariable operator theory. Our purpose below is to initiate an approach to the existence of free resolutions that is based on free covers.

We first discuss grading in the general noncommutative context. By a *grading* on a Hilbert module H we mean a strongly continuous unitary representation of the circle group $\Gamma : \mathbb{T} \to \mathcal{B}(H)$ that is related to the ambient row contraction as follows

$$\Gamma(\lambda)T_k\Gamma(\lambda)^* = \lambda T_k, \qquad \lambda \in \mathbb{T}, \quad k = 1, \dots, d. \tag{2.2}$$

Thus we are restricting ourselves to gradings in which the given operators $T_1, \dots, T_d$ are all of degree one. The group Γ is called the *gauge group* of the Hilbert module H. While there are many gradings of H that satisfy (2.2), when we refer to H as a graded Hilbert module it is implicit that a particular gauge group has been singled out. A graded morphism $A : H_1 \to H_2$ of graded Hilbert modules is a homomorphism $A \in \hom(H_1, H_2)$ that is of degree zero in the sense that

$$A\Gamma_1(\lambda) = \Gamma_2(\lambda)A, \qquad \lambda \in \mathbb{T},$$

Γ_k denoting the gauge group of H_k.

The natural gauge group of $F^2(Z)$ is defined by

$$\Gamma_0(\lambda) = \sum_{n=0}^{\infty} \lambda^n E_n$$

where E_n is the projection onto $Z^{\otimes n}$. Thus, $F^2 = F^2\langle z_1, \dots, z_d\rangle$ *becomes a graded contractive Hilbert module over* $\mathbb{C}\langle z_1, \dots, z_d\rangle$ *of defect* 1. More generally, let $F = F^2 \otimes C$ be a free Hilbert module of higher defect. Since the ambient operators $U_1, \dots, U_d$ of F^2 generate an irreducible C^*-algebra, one readily verifies that the most general strongly continuous unitary representation Γ of the circle group on F that satisfies $\Gamma(\lambda)(U_k \otimes \mathbf{1}_C)\Gamma(\lambda)^* = \lambda U_k \otimes \mathbf{1}_C$ for $k = 1, \dots, d$ must decompose into a tensor product of representations

$$\Gamma(\lambda) = \Gamma_0(\lambda) \otimes W(\lambda), \qquad \lambda \in \mathbb{T},$$

where W is an arbitrary strongly continuous unitary representation of $\mathbb{T}$ on the coefficient space C. It will be convenient to refer to a Hilbert space C that has been endowed with such a group W as a *graded Hilbert space*.

In order to discuss free resolutions, we shift attention to the more restricted category whose objects are graded Hilbert modules over the commutative polynomial algebra $\mathbb{C}[z_1, \dots, z_d]$ and whose maps are graded morphisms. In this context, one replaces the noncommutative free module $F^2 = F^2\langle z_1, \dots, z_d\rangle$ with its commutative counterpart $H^2 = H^2[z_1, \dots, z_d]$, namely the completion of $\mathbb{C}[z_1, \dots, z_d]$ in its natural norm. While this notation differs from the notation $H^2(\mathbb{C}^d)$ used in [Arv98] and [Arv00], it is more useful for our purposes here. The commutative free Hilbert module H^2 is realized as a quotient of F^2 as follows. Consider the the operator $A \in \mathcal{B}(F^2, H^2)$ obtained by closing the map that sends a noncommutative polynomial $f \in \mathbb{C}\langle z_1, \dots, z_d\rangle$ to its commutative image $\tilde{f} \in \mathbb{C}[z_d, \dots, z_d]$. This operator is a graded partial isometry with range H^2, whose kernel is the closure of the commutator ideal in $\mathbb{C}\langle z_1, \dots, z_d\rangle$,

$$K = \overline{\operatorname{span}}\{f \cdot (z_j z_k - z_k z_j) \cdot g : 1 \le j, k \le d, \quad f, g \in \mathbb{C}\langle z_1, \dots, z_d\rangle\}. \tag{2.3}$$

One sees this in more concrete terms after one identifies $H^2 \subseteq F^2$ with the completion of the symmetric tensor algebra in the norm inherited from F^2. In that realization one has $H^2 = K^\perp$, and A can be taken as the projection with range $K^\perp = H^2$. The situation is similar for graded free modules having multiplicity; indeed, for any graded coefficient space C the map

$$A \otimes 1_C : F^2 \otimes C \to H^2 \otimes C$$

defines a graded *cover* of the commutative free Hilbert module $H^2 \otimes C$.

The most general graded Hilbert module over the commutative polynomial algebra $\mathbb{C}[z_1, \dots, z_d]$ is a graded Hilbert module over $\mathbb{C}\langle z_1, \dots, z_d\rangle$ whose underlying row contraction $(T_1, \dots, T_d)$ satisfies $T_j T_k = T_k T_j$ for all j, k. Any vector ζ in such a module H has a unique decomposition into a Fourier series relative to the spectral subspaces of the gauge group,

$$\zeta = \sum_{n=-\infty}^{\infty} \zeta_n,$$

where $\Gamma(\lambda)\zeta_n = \lambda^n \zeta_n$, $n \in \mathbb{Z}$, $\lambda \in \mathbb{T}$. ζ is said to have finite Γ-spectrum if all but a finite number of the terms ζ_n of this series are zero. Finally, a graded contractive module H is said to be *finitely generated* if there is a finite set of vectors $\zeta_1, \dots, \zeta_s \in H$, each of which has finite Γ-spectrum, such that sums of the form

$$f_1 \cdot \zeta_1 + \cdots + f_s \cdot \zeta_s, \qquad f_1, \dots, f_s \in \mathbb{C}\langle z_1, \dots, z_d\rangle$$

are dense in H.

Our main result is the following counterpart of Hilbert's syzygy theorem.

Theorem 2.6. *For every finitely generated graded contractive Hilbert module H over the commutative polynomial algebra $\mathbb{C}[z_1, \dots, z_d]$ there is a weakly exact finite sequence of graded Hilbert modules*

$$0 \longrightarrow F_n \longrightarrow \cdots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow H \longrightarrow 0 \tag{2.4}$$

in which each $F_k = H^2 \otimes C_k$ is a free graded commutative Hilbert module of finite defect. The sequence (2.4) is unique up to a unitary isomorphism of diagrams and it terminates after at most $n = d$ steps.

Definition 2.7. The sequence (2.4) is called the *free* resolution of H.

Remark 2.8 (Betti numbers, Euler characteristic). The sequence (2.4) gives rise to a sequence of d numerical invariants of H

$$\beta_k(H) = \begin{cases} \text{defect}(F_k), & 1 \le k \le n, \\ 0, & n < k \le d. \end{cases}$$

and their alternating sum

$$\chi(H) = \sum_{k=1}^{d} (-1)^{k+1} \beta_k(H)$$

is called the *Euler characteristic* of H. Notice that this definition makes sense for any finitely generated (graded contractive) Hilbert module over $\mathbb{C}[z_1, \dots, z_d]$, and generalizes the Euler characteristic of [Arv00] that was restricted to Hilbert modules of *finite rank* as in Remark 2.5.

Remark 2.9 (Curvature and Index). The curvature invariant of [Arv00] is defined only in the context of finite rank contractive Hilbert modules, hence the index formula of [Arv02] that relates the curvature invariant to the index of a Dirac operator is not meaningful in the broader context of Theorem 2.6. On the other hand, the proof of that formula included an argument showing that the Euler characteristic can be calculated in terms of the Koszul complex associated with the Dirac operator, and that part of the proof is easily adapted to this context to yield the following more general variation of the index theorem. *For any finitely generated graded Hilbert module H with Dirac operator D, both* $\ker D_+$ *and* $\ker D_+^*$ *are finite-dimensional, and*

$$(-1)^d \chi(H) = \dim \ker D_+ - \dim \ker D_+^*.$$

Remark 2.10 (Relation to Localized Dilation-Theoretic Resolutions). We have pointed out in Remark 2.5 that for pure d-contractions $(T_1, \dots, T_d)$ acting on a Hilbert space, dilation-theoretic techniques give rise to an exact sequence of contractive Hilbert modules and partially isometric maps

$$\cdots \longrightarrow H^2 \otimes C_2 \longrightarrow H^2 \otimes C_1 \longrightarrow H \longrightarrow 0,$$

in which the coefficient spaces C_k of the free Hilbert modules are typically infinite-dimensional, and which apparently fails to terminate in a finite number of steps. However, in a recent paper [Gre03], Greene studied "localizations" of the above exact sequence at various points of the unit ball, and he has shown that when one localizes at the origin of $\mathbb{C}^d$, the homology of his localized complex agrees with the homology of Taylor's Koszul complex (see [Tay70a],[Tay70b]) of the underlying operator d-tuple $(T_1, \dots, T_d)$ in all cases. Interesting as these local results are, they appear unrelated to the global methods and results of this paper.

Remark 2.11 (Resolutions of modules over function algebras). We also point out that our use of the terms *resolution* and *free resolution* differs substantially from usage of similar terms in work of Douglas, Misra and Varughese [DMV00], [DMV01], [DM03a], [DM03b]. For example, in [DM03b], the authors consider

Hilbert modules over certain algebras $\mathcal{A}(\Omega)$ of analytic functions on bounded domains $\Omega \subseteq \mathbb{C}^d$. They introduce a notion of *quasi-free* Hilbert module that is related to localization, and is characterized as follows. Consider an inner product on the algebraic tensor product $\mathcal{A}(\Omega) \otimes \ell^2$ of vector spaces with three properties: a) evaluations at points of Ω should be locally uniformly bounded, b) module multiplication from $\mathcal{A}(\Omega) \times (\mathcal{A}(\Omega) \otimes \ell^2)$ to $\mathcal{A}(\Omega) \otimes \ell^2$ should be continuous, and c) it satisfies a technical condition relating Hilbert norm convergence to pointwise convergence throughout Ω. The completion of $\mathcal{A}(\Omega) \otimes \ell^2$ in that inner product gives rise to a Hilbert module over $\mathcal{A}(\Omega)$, and such Hilbert modules are called *quasi-free*.

The main result of [DM03b] asserts that "weak" quasi-free resolutions

$$\cdots \longrightarrow Q_2 \longrightarrow Q_1 \longrightarrow H \longrightarrow 0$$

exist for certain Hilbert modules H over $\mathcal{A}(\Omega)$, namely those that are higher-dimensional generalizations of the Hilbert modules studied by Cowen and Douglas in [CD78] for domains $\Omega \subseteq \mathbb{C}$. The modules Q_k are quasi-free in the sense above, but their ranks may be infinite and such sequences may fail to terminate in a finite number of steps.

3. Generators

Throughout this section we consider contractive Hilbert modules over the noncommutative polynomial algebra $\mathbb{C}\langle z_1, \dots, z_d\rangle$, perhaps graded.

Definition 3.1. Let H be a Hilbert module over $\mathbb{C}\langle z_1, \dots, z_d\rangle$. By a *generator* for H we mean a linear subspace $G \subseteq H$ such that

$$H = \overline{\operatorname{span}}\,\{f \cdot \zeta : f \in \mathbb{C}\langle z_1, \dots, z_d\rangle, \quad \zeta \in G\}.$$

We also say that H is *finitely generated* if it has a finite-dimensional generator, and in the category of graded Hilbert modules the term means a bit more, namely, that there is a finite-dimensional *graded* generator.

According to Definition 2.1, a finitely generated Hilbert module H is properly generated precisely when the defect subspace $H \ominus (Z \cdot H)$ is a finite-dimensional generator. In general, the defect subspace $H \ominus (Z \cdot H)$ of a finitely generated Hilbert module is necessarily finite-dimensional, but it can fail to generate and is sometimes $\{0\}$ (for examples, see the proof of Proposition 3.4). In particular, *finitely generated Hilbert modules need not be properly generated.* The purpose of this section is to show that many important examples *are* properly generated, and that many others are related to properly generated Hilbert modules in a simple way.

The following result can be viewed as a noncommutative operator theoretic counterpart of Nakayama's Lemma ([Eis04], Lemma 1.4).

Proposition 3.2. *Every finitely generated graded Hilbert module is properly generated.*

Proof. The space $G = H \ominus (Z \cdot H)$ is obviously a graded subspace of H, and it is finite-dimensional because $\dim G = \dim(H/(Z \cdot H))$ is dominated by the cardinality of any generating set. We have to show that G is a generator.

For that, we claim that the spectrum of the gauge group Γ is bounded below. Indeed, the hypothesis implies that there is a finite set of elements $\zeta_1, \dots, \zeta_s$ of H, each having finite Γ-spectrum, which generate H. By enlarging the set of generators appropriately and adjusting notation, we can assume that each ζ_k is an eigenvector of Γ,

$$\Gamma(\lambda)\zeta_k = \lambda^{n_k}\zeta_k, \qquad \lambda \in \mathbb{T}, \quad 1 \le k \le s.$$

Let n_0 be the minimum of $n_1, n_2, \dots, n_s$. For any monomial f of respective degrees $p_1, \dots, p_d$ in the noncommuting variables $z_1, \dots, z_d$ and every $k = 1, \dots, s$, $f \cdot \zeta_k$ is an eigenvector of Γ satisfying

$$\Gamma(\lambda)(f \cdot \zeta_k) = \lambda^N f \cdot \zeta_k$$

with $N = p_1 + \cdots + p_d + n_k \ge n_0$. Since elements of this form have H as their closed linear span, the spectrum of Γ is bounded below by n_0.

Setting $H_n = \{\xi \in H : \Gamma(\lambda)\xi = \lambda^n \xi, \lambda \in \mathbb{T}\}$ for $n \in \mathbb{Z}$, we conclude that

$$H = H_{n_0} \oplus H_{n_0+1} \oplus \cdots,$$

and one has $Z \cdot H_n \subseteq H_{n+1}$ for all $n \ge n_0$.

Since $G = H \ominus (Z \cdot H)$ is gauge-invariant it has a decomposition

$$G = G_{n_0} \oplus G_{n_0+1} \oplus \cdots,$$

in which $G_{n_0} = H_{n_0}$, $G_n = H_n \ominus (Z \cdot H_{n-1})$ for $n > n_0$, and where only a finite number of G_k are nonzero. Thus, each eigenspace H_n decomposes into a direct sum

$$H_n = G_n \oplus (Z \cdot H_{n-1}), \qquad n > n_0.$$

Setting $n = n_0 + 1$ we have $H_{n_0+1} = \operatorname{span}(G_{n_0+1} + Z \cdot G_{n_0})$ and, continuing inductively, we find that for all $n > n_0$,

$$H_n = \operatorname{span}(G_n + Z \cdot G_{n-1} + Z^{\otimes 2} \cdot G_{n-2} + \cdots + Z^{\otimes(n-n_0)} \cdot G_{n_0}),$$

where $Z^{\otimes r}$ denotes the space of homogeneous polynomials of total degree r. Since H is spanned by the subspaces H_n, it follows that G is a generator. $\square$

One obtains the most general examples of graded Hilbert submodules of the rank-one free commutative Hilbert module H^2 in explicit terms by choosing a (finite or infinite) sequence of homogeneous polynomials $\phi_1, \phi_2, \dots$ and considering the closure in H^2 of the set of all finite linear combinations $f_1 \cdot \phi_1 + \cdots + f_s \cdot \phi_s$, where $f_1, \dots, f_s$ are arbitrary polynomials and $s = 1, 2, \dots$. In Remark 2.5 above, we alluded to the fact that in all nontrivial cases, graded submodules of H^2 are Hilbert modules of infinite rank. However, Proposition 5.3 below implies that these examples are properly finitely generated, so they have free covers of finite defect by Theorem 2.4.

Remark 3.3 (Examples of Higher Degree). We now describe a class of infinite rank ungraded examples with substantially different properties. Perhaps we should point out that there is a more general notion of grading with respect to which the ambient operators $T_1, \dots, T_d$ in these examples are graded with various degrees larger than one. For brevity, we retain the simpler definition of grading (2.2) by viewing these examples as ungraded. Fix a d-tuple of positive integers $N_1, \dots, N_d$ and consider the following d-contraction acting on the Hilbert space $H = H^2(\mathbb{C}^d)$

$$(T_1, \dots, T_d) = (S_1^{N_1}, \dots, S_d^{N_d}),$$

where $(S_1, \dots, S_d)$ is the d-shift. The defect space of this Hilbert module

$$G = H \ominus (T_1 H + \dots + T_d H)^-$$

coincides with the intersection of the kernels $\ker T_1^* \cap \dots \cap \ker T_d^*$; and in this case one can compute these kernels explicitly, with the result

$$G = \operatorname{span}\{z_1^{n_1} \cdots z_d^{n_d} : 0 \le n_k < N_k, \quad 1 \le k \le d\}.$$

Moreover, for every set of nonnegative integers $\ell_1, \dots, \ell_d$, the set of vectors $T_1^{\ell_1} \cdots T_d^{\ell_d} G$ contains all monomials of the form

$$z_1^{\ell_1 N_1 + n_1} \cdots z_d^{\ell_d N_d + n_d}, \qquad 0 \le n_k < N_k, \quad 1 \le k \le d.$$

It follows from these observations that G is a proper generator for H, and Theorem 2.4 provides a free cover of the form $A : H^2 \otimes G \to H$.

Another straightforward computation with coefficients shows that the defect operator of this Hilbert module is of infinite rank whenever at least one of the integers $N_1, \dots, N_d$ is larger than 1. In more detail, each monomial $z^n = z_1^{n_1} \cdots z_d^{n_d}$, $n_1, \dots, n_d \ge 0$, is an eigenvector of the defect operator $\Delta = (\mathbf{1} - T_1 T_1^* - \dots - T_d T_d^*)^{1/2}$; and when $n_k \ge N_k$ for all k, a straightforward application of the formulas on pp. 178–179 of [Arv98] shows that

$$\Delta\, z_1^{n_1} \cdots z_d^{n_d} = c(n) z_1^{n_1} \cdots z_d^{n_d}$$

where the eigenvalues $c(n) = c(n_1, \dots, n_d)$ satisfy $0 < c(n) < 1$. Hence the defect operator has infinite rank. We conclude that, while dilation theory provides a coisometry $B : H^2 \otimes C \to H$ from another free Hilbert module to H, it is necessary that C be an infinite dimensional Hilbert space. Needless to say, such a B cannot define a free cover.

The preceding examples are all of infinite rank, and it is natural to ask about finite rank d-contractions – which were the focus of [Arv98], [Arv00], [Arv02]. Significantly, while the Hilbert module associated with a finite rank d-contraction is frequently not properly generated, it can always be extended to a properly generated one by way of a finite-dimensional perturbation.

Proposition 3.4. *Every pure Hilbert module H of finite rank can be extended trivially to a properly generated one in the sense that there is an exact sequence of Hilbert modules*

$$0 \longrightarrow K \longrightarrow \tilde{H} \underset{A}{\longrightarrow} H \longrightarrow 0$$

in which $\tilde{H}$ is is a properly generated pure Hilbert module of the same rank, K is a finite-dimensional Hilbert submodule of $\tilde{H}$, and A is a coisometry.

Proof. A standard dilation-theoretic technique (see [Arv98] for the commutative case, the proof of which works as well in general) shows that a pure Hilbert module of rank r is unitarily equivalent to a quotient of the form

$$H \cong (F^2 \otimes C)/M$$

where F^2 is the noncommutative free module of rank 1, C is an r-dimensional coefficient space, and M is a closed submodule of $F^2 \otimes C$. We identify H with the orthocomplement $M^\perp$ of M in $F^2 \otimes C$, with operators $T_1, \dots, T_d$ obtained by compressing to $M^\perp$ the natural operators $U_1 \otimes \mathbf{1}_C, \dots, U_d \otimes \mathbf{1}_C$ of $F^2 \otimes C$.

Consider $\tilde{H} = M^\perp + 1 \otimes C$. This is a finite-dimensional extension of $M^\perp$ that is also invariant under $U_k^* \otimes \mathbf{1}_C$, hence it defines a pure Hilbert module of rank r by compressing the natural operators in the same way to obtain $\tilde{T}_1, \dots, \tilde{T}_d \in \mathcal{B}(\tilde{H})$. Since $\tilde{H}$ contains H, the projection $P_{M^\perp}$ restricts to a homomorphism of Hilbert modules $A : \tilde{H} \to H$. A is a coisometry, and the kernel of A is finite-dimensional because $\dim(\tilde{H}/H) < \infty$.

To see that $\tilde{H}$ is properly generated, one computes the defect operator Δ of $\tilde{H}$. Indeed, $\Delta = (\mathbf{1}_{\tilde{H}} - \tilde{T}_1\tilde{T}_1^* - \cdots - \tilde{T}_d\tilde{T}_d^*)^{1/2}$ is seen to be the compression of the defect operator of $U_1 \otimes \mathbf{1}_C, \dots, U_d \otimes \mathbf{1}_C$ to $\tilde{H}$, and the latter defect operator is the projection onto $1 \otimes C$. Since $\tilde{H}$ contains $1 \otimes C$, the defect operator of $\tilde{H}$ is the projection on $1 \otimes C$.

Finally, we make use of the observation that a pure finite rank d-tuple is properly generated whenever its defect operator is a projection. Indeed, the range of the defect operator Δ is always a generator, and when Δ is a projection its range coincides with $\ker \tilde{T}_1^* \cap \cdots \cap \tilde{\ker} T_d^*$. □

4. Existence of Free Covers

We now establish the existence and uniqueness of free covers for properly generated Hilbert modules over $\mathbb{C}\langle z_1, \dots, z_d\rangle$. A cover $A : F \to H$ induces a unitary map of defect spaces; the following result implies that this isometry of quotients lifts to an isometry of the corresponding subspaces.

Lemma 4.1. *Let H be a Hilbert module and let $A : F \to H$ be a cover. Then A restricts to a unitary operator from $F \ominus (Z \cdot F)$ to $H \ominus (Z \cdot H)$.*

Proof. Let $Q \in \mathcal{B}(H)$ be the projection onto $H \ominus (Z \cdot H)$. The natural map of H onto the quotient Hilbert space $H/(Z \cdot H)$ is a partial isometry whose adjoint is the isometry

$$\eta + Z \cdot H \in H/(Z \cdot H) \mapsto Q\eta \in H \ominus (Z \cdot H), \qquad \eta \in H.$$

Thus we can define a unitary map $\tilde{A}$ from $F \ominus (Z \cdot F)$ onto $H \ominus (Z \cdot H)$ by composing the three unitary operators

$$\begin{aligned} \zeta \in F \ominus (Z \cdot F) &\mapsto \zeta + Z \cdot F \in F/(Z \cdot F), \\ \dot{A} : F/(Z \cdot F) &\to H/(Z \cdot H), \\ \eta + Z \cdot H \in H/(Z \cdot H) &\mapsto Q\eta \in H \ominus (Z \cdot H), \qquad \eta \in H, \end{aligned}$$

to obtain $\tilde{A}\zeta = QA\zeta$, $\zeta \in F \ominus (Z \cdot F)$. We claim now that $QA\zeta = A\zeta$ for all $\zeta \in F \ominus (Z \cdot F)$. To see that, note that $Q^\perp$ is the projection onto $Z \cdot H$, so that for all $\zeta \in F \ominus (Z \cdot F)$ one has

$$\begin{aligned} \|QA\zeta\| = \|A\zeta - Q^\perp A\zeta\| &= \inf_{\eta \in Z \cdot H} \|A\zeta - \eta\| \\ &= \|\dot{A}(\zeta + Z \cdot F)\|_{H/(Z \cdot H)} = \|\zeta + Z \cdot F\|_{F/(Z \cdot F)} = \|\zeta\|. \end{aligned}$$

Hence, $\|A\zeta - QA\zeta\|^2 = \|A\zeta\|^2 - \|QA\zeta\|^2 = \|A\zeta\|^2 - \|\zeta\|^2 \leq 0$, and the claim follows. We conclude that the restriction of A to $F \ominus (Z \cdot F)$ is an isometry with range $H \ominus (Z \cdot H)$. $\square$

Proof of Theorem 2.4. Let H be a properly generated Hilbert module over $\mathbb{C}\langle z_1, \dots, z_d\rangle$ and set $C = H \ominus (Z \cdot H)$. The hypothesis asserts that C is a generator. We will show that there is a (necessarily unique) contraction $A : F^2 \otimes C \to H$ satisfying

$$A(f \otimes \zeta) = f \cdot \zeta, \qquad f \in \mathbb{C}\langle z_1, \dots, z_d\rangle, \quad \zeta \in C, \tag{4.1}$$

and that such an operator A defines a free cover. For that, consider the completely positive map defined on $\mathcal{B}(H)$ by $\phi(X) = T_1 X T_1^* + \cdots + T_d X T_d^*$, and let $\Delta = (\mathbf{1} - \phi(\mathbf{1}))^{1/2}$ be the defect operator of (2.1). Since $H \ominus (Z \cdot H)$ is the intersection of kernels $\ker T_1^* \cap \cdots \cap \ker T_d^* = \ker \phi(\mathbf{1})$, it follows that

$$C = H \ominus (Z \cdot H) = \{\zeta \in H : \Delta\zeta = \zeta\}.$$

Thus, C is a subspace of the range of Δ on which Δ restricts to the identity operator, and which generates H. We now use the "dilation telescope" to show that there is a unique contraction $L : F^2 \otimes \overline{\Delta H} \to H$ such that

$$L(f \otimes \zeta) = f \cdot \Delta\zeta, \qquad f \in \mathbb{C}\langle z_1, \dots, z_d\rangle, \quad \zeta \in \overline{\Delta H}. \tag{4.2}$$

Indeed, since the monomials $\{z_{i_1} \otimes \cdots \otimes z_{i_n} : i_1, \dots, i_n \in \{1, \dots, d\}\}$, $n = 1, 2, \dots$, together with the constant polynomial 1, form an orthonormal basis for F^2, the formal adjoint of L is easily computed and found to be

$$L^*\xi = 1 \otimes \Delta\xi + \sum_{n=1}^{\infty} \sum_{i_1, \dots, i_n = 1}^{d} z_{i_1} \otimes \cdots \otimes z_{i_n} \otimes \Delta T_{i_n}^* \cdots T_{i_1}^* \xi, \qquad \xi \in H.$$

One calculates norms in the obvious way to obtain

$$\begin{aligned}\|L^*\xi\|^2 &= \|\Delta\xi\|^2 + \sum_{n=1}^{\infty}\sum_{i_1,\dots,i_n=1}^{d} \|\Delta T_{i_n}^* \cdots T_{i_1}^*\xi\|^2 \\ &= \langle(\mathbf{1}-\phi(\mathbf{1}))\xi,\xi\rangle + \sum_{n=1}^{\infty}\langle(\phi^n(\mathbf{1}-\phi(\mathbf{1}))\xi,\xi\rangle \\ &= \langle(\mathbf{1}-\phi(\mathbf{1}))\xi,\xi\rangle + \sum_{n=1}^{\infty}\langle(\phi^n(\mathbf{1})-\phi^{n+1}(\mathbf{1}))\xi,\xi\rangle \\ &= \|\xi\|^2 - \lim_{n\to\infty}\langle\phi^n(\mathbf{1})\xi,\xi\rangle \le \|\xi\|^2.\end{aligned}$$

Hence $\|L\| = \|L^*\| \le 1$. Finally, let A be the restriction of L to the submodule $F^2 \otimes C \subseteq F^2 \otimes \overline{\Delta H}$, where we now consider $F^2 \otimes C$ as a free Hilbert module of possibly smaller defect. Since Δ restricts to the identity on C, (4.1) follows from (4.2).

By its definition, the restriction of A to $1 \otimes C$ is an isometry with range $C = H \ominus (Z \cdot H)$, hence A induces a unitary operator of defect spaces

$$\dot{A} : (F^2 \otimes C)/(Z \cdot (F^2 \otimes C)) \cong 1 \otimes C \to C = H \ominus (Z \cdot H) \cong H/(Z \cdot H).$$

The range of A is dense, since it contains

$$\{f \cdot \zeta : f \in \mathbb{C}\langle z_1, \dots, z_d\rangle,\ \zeta \in H \ominus (Z \cdot H)\}$$

and H is properly generated. It follows that $A : F^2 \otimes C \to H$ is a free cover.

For uniqueness, let $B : \tilde{F} = F^2 \otimes \tilde{C} \to H$ be another free cover of H. We exhibit a unitary isomorphism of Hilbert modules $V \in \mathcal{B}(F^2 \otimes \tilde{C}, F^2 \otimes C)$ such that $BV = A$ as follows. We have already pointed out that the defect space of $\tilde{F} = F^2 \otimes \tilde{C}$ (resp. $F = F^2 \otimes C$)) is identified with $1 \otimes \tilde{C}$ (resp. $1 \otimes C$). Similarly, the defect space of H is identified with $H \ominus (Z \cdot H)$. Since both A and B are covers of H, Lemma 4.1 implies that they restrict to unitary operators, from the respective spaces $1 \otimes C$ and $1 \otimes \tilde{C}$, onto the *same* subspace $H \ominus (Z \cdot H)$ of H. Thus there is a unique unitary operator $V_0 : C \to \tilde{C}$ that satisfies

$$A(1 \otimes \zeta) = B(1 \otimes V_0\zeta), \qquad \zeta \in C.$$

Let $V = \mathbf{1}_{F^2} \otimes V_0 \in \mathcal{B}(F^2 \otimes \tilde{C}, F^2 \otimes C)$. Obviously V is a unitary operator, and it satisfies $BV = A$ since for every polynomial $f \in \mathbb{C}\langle z_1, \dots, z_d\rangle$

$$BV(f \otimes \zeta) = B(f \cdot (1 \otimes V_0\zeta)) = f \cdot B(1 \otimes V_0\zeta) = f \cdot A(1 \otimes \zeta) = A(f \otimes \zeta),$$

and one can take the closed linear span on both sides. V must implement an isomorphism of modules since for any polynomials $f, g \in \mathbb{C}\langle z_1, \dots, z_d\rangle$ and every $\zeta \in \tilde{C}$ we have

$$V(f \cdot (g \otimes \zeta)) = (\mathbf{1} \otimes V_0)(f \cdot g \otimes \zeta) = f \cdot (g \otimes V_0\zeta) = f \cdot V(g \otimes \zeta).$$

Conversely, if a Hilbert module H has a free cover $A : F^2 \otimes C \to H$, then since $1 \otimes C$ is the orthocomplement of $Z \cdot (F^2 \otimes C)$, Lemma 4.1 implies that

$A(1 \otimes C) = H \ominus (Z \cdot H))$. Since A is a module homomorphism, we see that

$$A(\operatorname{span}\{f \otimes \zeta : f \in \mathbb{C}, \quad \zeta \in C\}) = \operatorname{span}\{f \cdot \zeta : f \in \mathbb{C}, \quad \zeta \in H \ominus (Z \cdot H)\}.$$

The closure of the left side is H because A has dense range, and we conclude that $H \ominus (Z \cdot H)$ is a generator of H. □

We require the following consequence of Theorem 2.4 for finitely generated graded Hilbert modules.

Theorem 4.2. *Every finitely generated graded Hilbert module H over the non-commutative polynomial algebra has a graded free cover $A : F^2 \otimes C \to H$, and any two graded graded free covers are equivalent.*

If the underlying operators of H commute, then this free cover descends naturally to a commutative graded free cover $B : H^2 \otimes C \to H$.

Proof. Proposition 3.2 implies that the space $C = H \ominus (Z \cdot H)$ is a finite-dimensional generator. Moreover, since $Z \cdot H$ is invariant under the gauge group, so is C, and the restriction of the gauge group to C gives rise to a unitary representation $W : \mathbb{T} \to \mathcal{B}(C)$ of the circle group on C.

If we make the free Hilbert module $F^2 \otimes C$ into a graded one by introducing the gauge group

$$\Gamma(\lambda) = \Gamma_0(\lambda) \otimes W(\lambda), \qquad \lambda \in \mathbb{T},$$

then we claim that the map $A : F^2 \otimes C \to H$ defined in the proof of Theorem 2.4 must intertwine Γ and Γ_H. Indeed, this follows from the fact that for every polynomial $f \in \mathbb{C}\langle z_1, \dots, z_d\rangle$, every $\zeta \in C$, and every $\lambda \in \mathbb{T}$, one has

$$\begin{aligned}\Gamma_H(\lambda)A(f \otimes \zeta) &= \Gamma_H(\lambda)(f \cdot \zeta) = f(\lambda z_1, \dots, \lambda z_d) \cdot \Gamma_H(\lambda)\zeta \\ &= A(\Gamma_0(\lambda)f \otimes W(\lambda)\zeta) = A\Gamma(\lambda)(f \otimes \zeta).\end{aligned}$$

The proof of uniqueness in the graded context is now a straightforward variation of the uniqueness proof of Theorem 2.4. Finally, since H^2 is naturally identified with the quotient F^2/K where K is the closure of the commutator ideal in $\mathbb{C}\langle z_1, \dots, z_d\rangle$, it follows that when the underlying operators commute, the cover $A : F^2 \otimes C \to H$ factors naturally through $(F^2/K) \otimes C \sim H^2 \otimes C$ and one can promote A to a graded commutative free cover $B : H^2 \otimes C \to H$. □

5. Existence of Free Resolutions

We turn now to the proof of existence of finite resolutions for graded Hilbert modules over the commutative polynomial algebra $\mathbb{C}[z_1, \dots, z_d]$. We require some algebraic results obtained by Hilbert at the end of the century before last [Hil90], [Hil93]. While Hilbert's theorems have been extensively generalized, what we require are the most concrete versions of a) the basis theorem and b) the syzygy theorem. We now describe these classical results in a formulation that is convenient for our purposes, referring the reader to [Nor76], [Eis94] and [Ser00] for more detail on the underlying linear algebra.

Let $T_1, \dots, T_d$ be a set of commuting linear operators acting on a complex vector space M. We view M as a module over $\mathbb{C}[z_1, \dots, z_d]$ in the usual way, with $f \cdot \xi = f(T_1, \dots, T_d)\xi$, $f \in \mathbb{C}[z_1, \dots, z_d]$, $\xi \in M$. Such a module is said to be *graded* if there is a specified sequence M_n, $n \in \mathbb{Z}$, of subspaces that gives rise to an algebraic direct sum decomposition

$$M = \sum_{n=-\infty}^{\infty} M_n$$

with the property $T_k M_n \subseteq M_{n+1}$, for all $k = 1, \dots, d$, $n \in \mathbb{Z}$. Thus, every element ξ of M admits a unique decomposition $\xi = \sum_n \xi_n$, where ξ_n belongs to M_n and $\xi_n = 0$ for all but a finite number of n. We confine ourselves to the standard grading on $\mathbb{C}[z_1, \dots, z_d]$ in which the generators $z_1, \dots, z_d$ are all of degree 1. Finally, M is said to be *finitely generated* if there is a finite set $\zeta_1, \dots, \zeta_s \in M$ such that

$$M = \{f_1 \cdot \zeta_1 + \cdots + f_s \cdot \zeta_s : f_1, \dots, f_d \in \mathbb{C}[z_1, \dots, z_d]\}.$$

A free module is a module of the form $F = \mathbb{C}[z_1, \dots, z_d] \otimes C$ where C is a complex vector space, the module action being defined in the usual way by $f \cdot (g \otimes \zeta) = (f \cdot g) \otimes \zeta$. The *rank* of F is the dimension of C. A free module can be graded in many ways, and for our purposes the most general grading on $F = \mathbb{C}[z_1, \dots, z_d] \otimes C$ is defined as follows. Given an arbitrary grading on the "coefficient" vector space C

$$C = \sum_{n=-\infty}^{\infty} C_n,$$

there is a corresponding grading of the tensor product $F = \sum_n F_n$ in which

$$F_n = \sum_{k=0}^{\infty} Z^k \otimes C_{n-k},$$

where Z^k denotes the space of all homogeneous polynomials of degree k in $\mathbb{C}[z_1, \dots, z_d]$, and where the sum on the right denotes the linear subspace of F spanned by $\cup\{Z^k \otimes C_{n-k} : k \in \mathbb{Z}\}$. If C is finite-dimensional, then there are integers $n_1 \leq n_2$ such that

$$C = C_{n_1} + C_{n_1+1} + \cdots + C_{n_2},$$

so that

$$F_n = \sum_{k=0}^{\infty} Z^k \otimes C_{n-k} = \sum_{k=\max(n-n_2,0)}^{\max(n-n_1,0)} Z^k \otimes C_{n-k} \tag{5.1}$$

is finite-dimensional for each $n \in \mathbb{Z}$, $F_n = \{0\}$ for $n < n_1$, and F_n is spanned by $Z^{n-n_2} \cdot F_{n_2}$ for $n \geq n_2$.

Homomorphisms of graded modules $u : M \to N$ are required to be of degree zero

$$u(M_n) \subseteq N_n, \qquad n \in \mathbb{Z}.$$

It will also be convenient to adapt Serre's definition of minimality for homomorphisms of modules over local rings (page 84 of [Ser00]) to homomorphisms of graded modules over $\mathbb{C}[z_1,\dots,z_d]$, as follows. A homomorphism $u : M \to N$ of modules is said to be *minimal* if it induces an isomorphism of vector spaces

$$\dot{u} : M/(z_1 \cdot M + \cdots + z_d \cdot M) \to u(M)/u(z_1 \cdot M + \cdots + z_d \cdot M).$$

Equivalently, u is minimal iff $\ker u \subseteq z_1 \cdot M + \cdots + z_d \cdot M$.

A *free resolution* of an algebraic graded module M is a (perhaps infinite) exact sequence of graded modules

$$\cdots \longrightarrow F_n \longrightarrow \cdots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow M \longrightarrow 0,$$

where each F_r is free or 0. Such a resolution is said to be *finite* if each F_r is of finite rank and $F_r = 0$ for sufficiently large r, and *minimal* if for every $r = 1, 2, \dots$, the arrow emanating from F_r denotes a minimal homomorphism.

Theorem 5.1 (Basis Theorem). *Every submodule of a finitely generated module over $\mathbb{C}[z_1,\dots,z_d]$ is finitely generated.*

Theorem 5.2 (Syzygy Theorem). *Every finitely generated graded module M over $\mathbb{C}[z_1,\dots,z_d]$ has a finite free resolution*

$$0 \longrightarrow F_n \longrightarrow \cdots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow M \longrightarrow 0$$

that is minimal with length n at most d, and any two minimal resolutions are isomorphic.

While we have stated the ungraded version of the basis theorem, all we require is the special case for graded modules. We base the proof of Theorem 2.6 on two operator-theoretic results, the first of which is a Hilbert space counterpart of the basis theorem for graded modules.

Proposition 5.3. *Let H be a finitely generated graded Hilbert module over $\mathbb{C}[z_1,\dots,z_d]$ and let $K \subseteq H$ be a closed gauge-invariant submodule. Then K is a properly generated graded Hilbert module of finite defect.*

Proof. We first collect some structural information about H itself. Let Γ be the gauge group of H and consider the spectral subspaces of Γ

$$H_n = \{\xi \in H : \Gamma(\lambda)\xi = \lambda^n \xi\}, \qquad n \in \mathbb{Z}.$$

The finite-dimensional subspace $G = H \ominus (Z \cdot H)$ is invariant under the action of Γ, and Proposition 3.2 implies that G is a generator. Writing $G_n = G \cap H_n$, $n \in \mathbb{Z}$, it follows that G decomposes into a finite sum of mutually orthogonal subspaces

$$G = G_{n_1} + G_{n_1+1} + \cdots + G_{n_2},$$

where $n_1 \le n_2$ are fixed integers. A computation similar to that of (5.1) shows that $H_n = \{0\}$ for $n < n_1$, and for $n \ge n_1$, H_n can be expressed in terms of the G_k by way of

$$H_n = \sum_{k=\max(n-n_2,0)}^{\max(n-n_1,0)} Z^k \cdot G_{n-k}. \qquad n \in \mathbb{Z}; \tag{5.2}$$

in particular, each H_n is finite-dimensional.

Consider the algebraic module

$$H^0 = \operatorname{span}\{f \cdot \zeta : f \in \mathbb{C}[z_1, \dots, z_d], \quad \zeta \in G\}$$

generated by G. Formula (5.2) shows that H^0 is linearly spanned by the spectral subspaces of H,

$$H^0 = H_{n_1} + H_{n_1+1} + \cdots . \tag{5.3}$$

Now let $K \subseteq H$ be a closed invariant subspace that is also invariant under the action of Γ. Letting $K_n = H_n \cap K$ be the corresponding spectral subspace of K, then we have a decomposition of K into mutually orthogonal finite-dimensional subspaces

$$K = K_{n_1} \oplus K_{n_1+1} \oplus \cdots ,$$

such that $z_k K_n \subseteq K_{n+1}$, for $1 \leq k \leq d$, $n \geq n_1$. Let K^0 be the (nonclosed) linear span

$$K^0 = K_{n_1} + K_{n_1+1} + \cdots .$$

Obviously, K^0 is dense in K and it is a submodule of the finitely generated algebraic module H^0. Theorem 5.1 implies that there is a finite set of vectors $\zeta_1, \dots, \zeta_s \in K^0$ such that

$$K^0 = \{f_1 \cdot \zeta_1 + \cdots + f_s \cdot \zeta_s : f_1, \dots, f_s \in \mathbb{C}[z_1, \dots, z_d]\}.$$

Choosing p large enough that $\zeta_1, \dots, \zeta_s \in K_{n_1} + \cdots + K_p$, we find that $K_{n_1} + \cdots + K_p$ is a graded finite-dimensional generator for K. An application of Proposition 3.2 now completes the proof. □

5.1. From Hilbert Modules to Algebraic Modules. A finitely generated graded Hilbert module H over $\mathbb{C}[z_1, \dots, z_d]$ has many finite-dimensional graded generators G; if one fixes such a G then there is an associated algebraic graded module $M(H, G)$ over $\mathbb{C}[z_1, \dots, z_d]$, namely

$$M(H, G) = \operatorname{span}\{f \cdot \zeta : f \in \mathbb{C}[z_1, \dots, z_d], \quad \zeta \in G\}.$$

The second result that we require is that it is possible to make appropriate choices of G so as to obtain a functor from Hilbert modules to algebraic modules. We now define this functor and collect its basic properties.

Consider the category $\mathcal{H}_d$ whose objects are graded finitely generated Hilbert modules over $\mathbb{C}[z_1, \dots, z_d]$, with covers as maps. Thus, $\hom(H, K)$ consists of graded homomorphisms $A : H \to K$ satisfying $\|A\| \leq 1$, such that AH is dense in K, and which induce unitary operators of defect spaces

$$\dot{A} : H/(Z \cdot H) \to K/(Z \cdot K).$$

Since we are requiring maps in $\hom(H, K)$ to have dense range, a straightforward argument (that we omit) shows that $\hom(\cdot, \cdot)$ is closed under composition.

The corresponding algebraic category $\mathcal{A}_d$ has objects consisting of graded finitely generated modules over $\mathbb{C}[z_1, \dots, z_d]$, in which $u \in \hom(M, N)$ means that u is a *minimal* graded homomorphism satisfying $u(M) = N$.

Proposition 5.4. *For every Hilbert module H in $\mathcal{H}_d$ let H^0 be the algebraic module over $\mathbb{C}[z_1, \dots, z_d]$ defined by*

$$H^0 = \operatorname{span}\{f \cdot \zeta : f \in \mathbb{C}[z_1, \dots, z_d], \quad \zeta \in H \ominus (Z \cdot H)\}.$$

Then H^0 belongs to $\mathcal{A}_d$. Moreover, for every $A \in \hom(H, K)$ one has $AH^0 = K^0$, and the restriction A^0 of A to H^0 defines an element of $\hom(H^0, K^0)$. The association $H \to H^0$, $A \to A^0$ is a functor satisfying:

(i) *For every $H \in \mathcal{H}_d$, $H^0 = \{0\} \implies H = \{0\}$.*
(ii) *For every $A \in \hom(H, K)$, $A^0 = 0 \implies A = 0$.*
(iii) *For every free graded Hilbert module F of defect r, F^0 is a free algebraic graded module of rank r.*

Proof. Since the defect subspace $H \ominus (Z \cdot H)$ is finite-dimensional and invariant under the action of the gauge group Γ, H^0 is a finitely generated module over the polynomial algebra that is invariant under the action of the gauge group. Thus it acquires an algebraic grading $H^0 = \sum_n H_n^0$ by setting

$$H_n^0 = H^0 \cap H_n = \{\xi \in H^0 : \Gamma(\lambda)\xi = \lambda^n \xi, \quad \lambda \in \mathbb{T}\}, \qquad n \in \mathbb{Z}.$$

Let $H, K \in \mathcal{H}_d$ and let $A \in \hom(H, K)$. Lemma 4.1 implies that

$$A(H \ominus (Z \cdot H)) = K \ominus (Z \cdot K),$$

so that A restricts to a surjective graded homomorphism of modules $A^0 : H^0 \to K^0$. We claim that A^0 is minimal, i.e., $\ker A^0 \subseteq z_1 \cdot H^0 + \cdots + z_d \cdot H^0$. To see that, choose $\xi \in H^0$ such that $A\xi = 0$. Since H^0 decomposes into a sum

$$H^0 = H \ominus (Z \cdot H) + z_1 \cdot H^0 + \cdots + z_d \cdot H^0,$$

we can decompose ξ correspondingly

$$\xi = \zeta + z_1 \cdot \eta_1 + \cdots + z_d \cdot \eta_d,$$

where $\zeta \in H \ominus (Z \cdot H)$ and $\eta_j \in H^0$. Since $\dot{A}$ is an injective operator defined on $H/(Z \cdot H)$, $\ker A$ must be contained in $Z \cdot H$. It follows that $\xi \in Z \cdot H$, and therefore $\zeta = \xi - z_1 \cdot \eta_1 - \cdots - z_d \cdot \eta_d \in Z \cdot H = (H \ominus (Z \cdot H))^\perp$ is orthogonal to itself. Hence $\zeta = 0$, and we have the desired conclusion

$$\xi = z_1 \cdot \eta_1 + \cdots + z_d \cdot \eta_k \in z_1 \cdot H^0 + \cdots + z_d \cdot H^0.$$

The restriction A^0 of A to H^0 is therefore a minimal homomorphism, whence $A^0 \in \hom(H^0, K^0)$.

The composition rule $(AB)^0 = A^0 B^0$ follows immediately, so that we have defined a functor. Finally, both properties (i) and (ii) are consequences of the fact that, by Proposition 3.2, H^0 is dense in H, while (iii) is obvious. $\square$

Proof of Theorem 2.6. Given a graded finitely generated Hilbert module H, we claim that there is a weakly exact sequence

$$\cdots \longrightarrow F_n \longrightarrow \cdots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow H \longrightarrow 0, \tag{5.4}$$

in which each F_r is a free graded Hilbert module of finite defect. Indeed, Proposition 5.3 implies that H is properly generated, and by Theorem 2.6, it has a graded free cover $A : F_1 \to H$ in which $F_1 = H^2 \otimes C_1$ is a graded free

Hilbert module with $\dim C_1 = \text{defect}(F_1) = \text{defect}(H) < \infty$. This gives a sequence of graded Hilbert modules

$$F_1 \underset{A}{\longrightarrow} H \longrightarrow 0 \tag{5.5}$$

that is weakly exact at H. Proposition 5.3 implies that $\ker A$ is a properly generated graded Hilbert module of finite defect, so that another application of Theorem 2.6 produces a graded free cover $B : F_2 \to \ker A$ in which F_2 is a graded free Hilbert module of finite defect. Thus we can extend (5.5) to a longer sequence

$$F_2 \longrightarrow F_1 \longrightarrow H \longrightarrow 0$$

that is weakly exact at F_1 and H. Continuing inductively, we obtain (5.4).

Another application of Theorem 2.6 implies that the sequence (5.4) is uniquely determined by H up to a unitary isomorphism of diagrams. The only issue remaining is whether its length is finite. To see that (5.4) must terminate, consider the associated sequence of graded algebraic modules provided by Proposition 5.4

$$\cdots \longrightarrow F_n^0 \longrightarrow \cdots \longrightarrow F_2^0 \longrightarrow F_1^0 \longrightarrow H^0 \longrightarrow 0.$$

Proposition 5.4 implies that this is a *minimal* free resolution of H^0 into graded free modules F_r^0 of finite rank. The uniqueness assertion of Theorem 5.2 implies that there is an integer $n \leq d$ such that $F_r^0 = 0$ for all $r > n$. By Proposition 5.4 (i), we have $F_r = 0$ for $r > n$. □

Remark 5.5 (Noncommutative Generalizations). Perhaps it is worth pointing out that there is no possibility of generalizing Theorem 2.6 to the noncommutative setting, the root cause being that Hilbert's basis theorem fails for modules over the noncommutative algebra $\mathbb{C}\langle z_1, \dots, z_d\rangle$. More precisely, there are finitely generated graded Hilbert modules H over $\mathbb{C}\langle z_1, \dots, z_d\rangle$ that do not have finite free resolutions. Indeed, while Theorem 2.4 implies that for any such Hilbert module H there is a graded free Hilbert module $F_1 = F^2 \otimes C$ with $\dim C < \infty$ and a weakly exact sequence of graded Hilbert modules

$$F_1 \underset{A}{\longrightarrow} H \longrightarrow 0,$$

and while the kernel of A is a certainly a graded submodule of $F^2 \otimes C$, the kernel of A need not be finitely generated. For such a Hilbert module H, this sequence cannot be continued beyond F_1 within the category of Hilbert modules of finite defect.

As a concrete example of this phenomenon, let $N \geq 2$ be an integer, let $Z = \mathbb{C}^d$ for some $d \geq 2$, and consider the free graded noncommutative Hilbert module

$$F^2 = \mathbb{C} \oplus Z \oplus Z^{\otimes 2} \oplus Z^{\otimes 3} \oplus \cdots .$$

We claim that there is an infinite sequence of unit vectors $\zeta_N, \zeta_{N+1}, \dots \in F^2$ such that $\zeta_n \in Z^{\otimes n}$ and, for all $n \geq N$,

$$\zeta_{n+1} \perp M_n = \{f_N \cdot \zeta_N + f_{N+1} \cdot \zeta_{N+1} + \cdots + f_n \cdot \zeta_n : f_N, \dots, f_n \in \mathbb{C}\langle z_1, \dots, z_d\rangle\}.$$

Indeed, choose a unit vector ζ_N arbitrarily in $Z^{\otimes N}$ and, assuming that $\zeta_N, \dots, \zeta_n$ have been defined with the stated properties, note that M_n is a graded submodule of F^2 such that

$$M_n \cap Z^{\otimes(n+1)} = Z^{\otimes(n+1-N)} \cdot \zeta_N + Z^{\otimes(n-N)} \cdot \zeta_{N+1} + \cdots + Z \cdot \zeta_n.$$

Recalling that $\dim Z^k = d^k$, an obvious dimension estimate implies that

$$\dim(M_n \cap Z^{\otimes(n+1)}) \le d^{n+1-N} + \cdots + d = d\frac{d^{n-N+1}-1}{d-1}$$
$$< \frac{d^{n-N+2}}{d-1} \le d^{n-N+2} < d^{n+1} = \dim(Z^{\otimes(n+1)}).$$

Hence there is a unit vector $\zeta_{n+1} \in Z^{\otimes(n+1)}$ that is orthogonal to M_n. Now let M be the closure of $M_N \cup M_{N+1} \cup \cdots$. M is a graded invariant subspace of F^2 with the property that $M \ominus (Z \cdot M)$ contains the orthonormal set $\zeta_N, \zeta_{N+1}, \dots$, so that M cannot be finitely generated.

Finally, if we take H to be the Hilbert space quotient F^2/M, then H is a graded Hilbert module over $\mathbb{C}\langle z_1, \dots, z_d\rangle$ having a single gauge-invariant cyclic vector $1 + M$, such that the natural projection $A : F^2 \to H = F^2/M$ is a graded free cover of H where $\ker A = M$ is not finitely generated.

Notice that the preceding construction used the fact that the dimensions of the spaces $Z^{\otimes k}$ of noncommutative homogeneous polynomials grow exponentially in k. If one attempts to carry out this construction in the commutative setting, in which F^2 is replaced by H^2, one will find that the construction of the sequence $\zeta_N, \zeta_{N+1}, \dots$ fails at some point because the dimensions of the spaces Z^k of homogeneous polynomials grow too slowly. Indeed, as reformulated in Proposition 5.3, Hilbert's remarkable basis theorem implies that this construction *must* fail in the commutative setting, since every graded submodule of H^2 is finitely generated.

6. Examples of Free Resolutions

In this section we discuss some examples of free resolutions and their associated Betti numbers. There are two simple - and closely related - procedures for converting a free Hilbert module into one that is no longer free, by changing its ambient operators as follows.

(1) Append a number r of zero operators to the d-shift $(S_1, \dots, S_d)$ to obtain a $(d + r)$-contraction acting on $H^2[z_1, \dots, z_d]$ that is not the $(d + r)$-shift.
(2) Pass from $H^2[z_1, \dots, z_d]$ to a quotient $H^2[z_1, \dots, z_d]/K$ where K is the closed submodule generated by some of the coordinates $z_1, \dots, z_d$.

We begin by pointing out that one can understand either of these examples (1) or (2) by analyzing the other. We then calculate the Betti numbers of the Hilbert modules of (1) in the case where one appends three zero operators to the d-shift. In order to calculate the Betti numbers of a graded Hilbert module one has to calculate its free resolution, and that is the route we follow.

To see that (1) and (2) are equivalent constructions, consider the operator $(d+r)$-tuple $\bar{T} = (S_1, \dots, S_d, 0, \dots, 0)$ obtained from the d-shift $(S_1, \dots, S_d)$ acting on $H^2[z_1, \dots, z_d]$ by adjoining r zero operators. Let K be the closed invariant subspace of $H^2[z_1, \dots, z_{d+r}]$ generated by $z_{d+1}, z_{d+2}, \dots, z_{d+r}$. Recalling that $H^2[z_1, \dots, z_d]$ embeds isometrically in $H^2[z_1, \dots, z_{d+r}]$ with orthocomplement K,

$$H^2[z_1, \dots, z_{d+r}] = H^2[z_1, \dots, z_d] \oplus K,$$

one finds that the quotient Hilbert module $H^2[z_1, \dots, z_{d+r}]/K$ is identified with $H^2[z_1, \dots, z_d]$ in such a way that the natural $(d+r)$-contraction defined by the quotient is unitarily equivalent to $\bar{T}$.

Before turning to explicit computations we point out that, in order to calculate free resolutions, one has to iteratively calculate free covers. The procedure is summarized as follows.

Remark 6.1 (Free Covers and Free Resolutions). Let H be a finitely generated graded Hilbert module over $\mathbb{C}[z_1, \dots, z_d]$. In order to calculate the free resolution of H one has to iterate the following procedure.

(1) One first calculates the free cover $A_1 : H^2[z_1, \dots, z_d] \otimes G_1 \to H$ of H, following the proof of Theorem 2.4. To carry that out, one must calculate the unique proper generator $G_1 \subseteq H$

$$G_1 = H \ominus (Z \cdot H),$$

the connecting map A_1 being the closure of the multiplication map

$$A(f \otimes \zeta) = f \cdot \zeta, \qquad f \in \mathbb{C}[z_1, \dots, z_d], \quad \zeta \in G_1,$$

where the free Hilbert module $H^2[z_1, \dots, z_d] \otimes G_1$ is endowed with the grading $\Gamma_0 \otimes W$, W being the unitary representation of the circle group on G defined by restricting the grading Γ_H of H,

$$W(\lambda) = \Gamma_H(\lambda) \restriction_G, \qquad \lambda \in \mathbb{T}.$$

Notice that in order to carry out this step, one basically has to identify $Z \cdot H$ and its orthocomplement in concrete terms.

(2) One then replaces H with the finitely generated graded Hilbert module $\ker A_1 \subseteq H^2[z_1, \dots, z_d] \otimes G_1$ and repeats the procedure. It is significant that in order to continue, one must identify the kernel of A_1 and its proper generator $G_2 = \ker A_1 \ominus (Z \cdot \ker A_1)$.

According to Theorems 2.4 and 2.6, this process will terminate in the zero Hilbert module after at most d steps, and the resulting sequence

$$0 \longrightarrow H^2[z_1, \dots, z_d] \otimes G_n \underset{A_n}{\longrightarrow} \cdots \underset{A_2}{\longrightarrow} H^2[z_1, \dots, z_d] \otimes G_1 \underset{A_1}{\longrightarrow} H \longrightarrow 0$$

is the free resolution of H. Once one has the free resolution, one can read off the Betti numbers of H as the multiplicities of the various free Hilbert modules that have appeared in the sequence, *in their order of appearance*.

We now discuss the examples of (1) for the case $r = 3$ and arbitrary d.

Proposition 6.2. *The Hilbert module associated with the* $(d+3)$*-contraction* $(S_1,\dots,S_d,0,0,0)$ *acting on* $H^2[z_1,\dots,z_d]$ *has Euler characteristic zero, and its sequence of Betti numbers is*

$$(\beta_1,\dots,\beta_{d+3}) = (1,3,3,1,0,\dots,0).$$

Sketch of Proof. We show that the free resolution of H has the form

$$0 \to F_4 \to F_3 \to F_2 \to F_1 \to H \to 0$$

where $F_k = H^2[z_1,\dots,z_{d+3}] \otimes G_k$, G_1, G_2, G_3, G_4 being graded coefficient spaces of respective dimensions $1,3,3,1$. We will exhibit the modules F_k and the connecting maps explicitly, but we omit the details of computations with polynomials.

We first compute the proper generator $H \ominus (Z\cdot H)$ of H. Writing

$$T_1T_1^* + \cdots + T_{d+3}T_{d+3}^* = S_1S_1^* + \cdots + S_dS_d^*,$$

one sees that the defect operator $(\mathbf{1} - \sum_k T_kT_k^*)^{1/2}$ is the one-dimensional projection [1] onto the constant polynomials. It follows that H has defect 1, and its proper generator is the one-dimensional space $\mathbb{C}\cdot 1$.

Hence the first term in the free resolution of H is given by the free cover $A_1 : H^2[z_1,\dots,z_{d+3}] \to H$, where A_1 is the closure of the map defined on polynomials $f \in \mathbb{C}[z_1,\dots,z_{d+3}]$ by

$$A_1 f = f(S_1,\dots,S_d,0,0,0)\cdot 1 = f(z_1,\dots,z_d,0,0,0).$$

A_1 is a coisometry, and further computation with polynomials shows that its kernel is the closure $K_1 = \overline{(z_{d+1}, z_{d+2}, z_{d+3})}$ of the ideal in $\mathbb{C}[z_1,\dots,z_{d+3}]$ generated by $z_{d+1}, z_{d+2}, z_{d+3}$. This gives a sequence of contractive homomorphisms of degree zero

$$0 \longrightarrow K_1 \longrightarrow H^2[z_1,\dots,z_{d+3}] \longrightarrow H \longrightarrow 0$$

that is exact at $H^2[z_1,\dots,z_{d+3}]$.

The kernel K_1 is a graded submodule of $H^2[z_1,\dots,z_{d+3}]$, but the rank of its defect operator is typically *infinite*. However, by Proposition 5.3, it has a unique finite-dimensional proper generator G, given by

$$G = K_1 \ominus (Z\cdot K_1) = K_1 \ominus \overline{(z_1\cdot K_1 + \cdots + z_{d+3}\cdot K_1)}.$$

To compute G, note that each of the elements z_{d+1}, z_{d+2}, z_{d+3} is of degree one, while any homogeneous polynomial of $Z\cdot K_1$ is of degree at least two. It follows that $K_1 = \operatorname{span}\{z_{d+1}, z_{d+2}, z_{d+3}\} \oplus (Z\cdot K_1)$, and this identifies G as the 3-dimensional Hilbert space

$$G = \operatorname{span}\{z_{d+1}, z_{d+2}, z_{d+3}\}.$$

The multiplication map $A_2 : F\otimes G \to F$

$$A_2(f\otimes\zeta) = f\cdot\zeta, \qquad f \in \mathbb{C}[z_1,\dots,z_{d+3}], \quad \zeta \in G$$

is a contractive morphism that defines a free cover of K_1; and A_2 becomes a degree zero map with respect to the gauge group Γ on $H^2[z_1,\dots,z_{d+3}]\otimes G$ defined by $\Gamma = \Gamma_0 \otimes W$ where W is the restriction of the gauge group of

$H^2[z_1, \dots, z_{d+3}]$ to its subspace G, namely $W(\lambda) = \lambda \mathbf{1}_G$, $\lambda \in \mathbb{T}$. It follows that the sequence

$$H^2[z_1, \dots, z_{d+3}] \otimes G \underset{A_2}{\longrightarrow} H^2[z_1, \dots, z_{d+3}] \underset{A_1}{\longrightarrow} H \longrightarrow 0$$

is weakly exact at $H^2[z_1, \dots, z_{d+3}]$ and H.

Now consider $K_2 = \ker A_2 \subseteq H^2[z_1, \dots, z_{d+3}] \otimes G$. Since every element of $H^2[z_1, \dots, z_{d+3}] \otimes G$ can be written uniquely in the form

$$\xi_1 \otimes z_{d+1} + \xi_2 \otimes z_{d+2} + \xi_3 \otimes z_{d+3}, \qquad \xi_k \in H^2[z_1, \dots, z_{d+3}]$$

we have

$$K_2 = \{\xi_1 \otimes z_{d+1} + \xi_2 \otimes z_{d+2} + \xi_3 \otimes z_{d+3} : z_{d+1} \cdot \xi_1 + z_{d+2} \cdot \xi_2 + z_{d+3} \cdot \xi_3 = 0\}.$$

A nontrivial calculation with polynomials now shows that K_2 is the closed submodule of of $H^2 \otimes G$ generated by the three "commutators" $\zeta_1, \zeta_2, \zeta_3$

$$\begin{aligned} \zeta_1 &= z_{d+2} \otimes z_{d+3} - z_{d+3} \otimes z_{d+2} = z_{d+2} \wedge z_{d+3}, \\ \zeta_2 &= z_{d+1} \otimes z_{d+3} - z_{d+3} \otimes z_{d+1} = z_{d+1} \wedge z_{d+3} \\ \zeta_3 &= z_{d+1} \otimes z_{d+2} - z_{d+2} \otimes z_{d+1} = z_{d+1} \wedge z_{d+2}. \end{aligned}$$

Note, for example, that

$$f \cdot \zeta_1 + g \cdot \zeta_2 = -g z_{d+3} \otimes z_{d+1} - f z_{d+3} \otimes z_{d+2} + (g z_{d+1} + f z_{d+2}) \otimes z_{d+3}.$$

These elements ζ_k are all homogeneous of degree two. Since any homogeneous element of $Z \cdot K_2$ has degree at most three, it must be orthogonal to $\zeta_1, \zeta_2, \zeta_3$. It follows that

$$K_2 \ominus (Z \cdot K_2) = \operatorname{span}\{\zeta_2, \zeta_2, \zeta_3\}$$

is 3-dimensional, having $2^{-1/2}\zeta_1, 2^{-1/2}\zeta_2, 2^{-1/2}\zeta_3$ as an orthonormal basis.

Set $\tilde{G} = \operatorname{span}\{\zeta_2, \zeta_2, \zeta_3\}$, with its grading (in this case homogeneous of degree 2) as inherited from the grading of $H^2[z_1, \dots, z_{d+3}] \otimes G$. The corresponding free cover $A_3 : H^2[z_1, \dots, z_{d+3}] \otimes \tilde{G} \to K_2$ is given by

$$A_3(f_1 \otimes \zeta_1 + f_2 \otimes \zeta_2 + f_3 \otimes \zeta_3) = f_1 \cdot \zeta_1 + f_2 \cdot \zeta_2 + f_2 \cdot \zeta_3,$$

for polynomials f_1, f_2, f_3, and the grading of $H^2[z_1, \dots, z_{d+3}] \otimes \tilde{G}$ is given by $\Gamma(\lambda)(f \otimes \zeta) = \lambda^2(\Gamma_0(\lambda) f \otimes \zeta)$, $\lambda \in \mathbb{T}$.

Finally, consider the submodule $K_3 = \ker A_3 \subseteq H^2[z_1, \dots, z_{d+3}] \otimes \tilde{G}$. Another computation with polynomials shows that K_3 has a single generator

$$\begin{aligned} \eta &= z_{d+1} \otimes \zeta_1 - z_{d+2} \otimes \zeta_2 + z_{d+3} \otimes \zeta_3 \\ &= z_{d+1} \otimes (z_{d+2} \wedge z_{d+3}) - z_{d+2} \otimes (z_{d+1} \wedge z_{d+3}) + z_{d+3} \otimes (z_{d+1} \wedge z_{d+2}), \end{aligned}$$

where as above, $z_j \wedge z_k$ denotes $z_j \otimes z_k - z_k \otimes z_j$. The homogeneous element η has degree 3, so that after appropriate renormalization it becomes a unit vector spanning $K_3 \ominus (Z \cdot K_3)$. Thus, we obtain a free cover $A_4 : H^2[z_1, \dots, z_{d+3}] \to K_3$ by closing the map of polynomials

$$A_4(f) = f \cdot \eta, \qquad f \in \mathbb{C}[z_1, \dots, z_{d+3}].$$

Notice that the grading that $H^2[z_1, \dots, z_{d+3}]$ acquires by this construction is not the standard grading Γ_0, but rather $\Gamma(\lambda) = \lambda^3\Gamma_0(\lambda)$, $\lambda \in \mathbb{T}$.

Since the kernel of A_4 is obviously $\{0\}$, we have obtained a free resolution

$$0 \to F \xrightarrow{A_4} F \otimes \tilde{G} \xrightarrow{A_3} F \otimes G \xrightarrow{A_2} F \xrightarrow{A_1} H \to 0$$

in which $F = H^2[z_1, \dots, z_{d+3}]$.

This shows that H is a Hilbert module over $\mathbb{C}[z_1, \dots, z_{d+3}]$ whose Betti numbers $(\beta_1, \cdots, \beta_{d+3})$ are given by a nontrivial sequence $(1, 3, 3, 1, 0, \dots, 0)$ with alternating sum zero. $\square$

References

[Arv98] W. Arveson. Subalgebras of C^*-algebras III: Multivariable operator theory. *Acta Math.*, 181:159–228, 1998. arXiv:funct-an/9705007.

[Arv00] W. Arveson. The curvature invariant of a Hilbert module over $\mathbb{C}[z_1, \dots, z_d]$. *J. Reine Angew. Mat.*, 522:173–236, 2000. arXiv:math.OA/9808100.

[Arv02] W. Arveson. The Dirac operator of a commuting d-tuple. *Jour. Funct. Anal.*, 189:53–79, 2002. arXiv:math.OA/0005285.

[CD78] M. Cowen and R. G. Douglas. Complex geometry and operator theory. *Acta Math*, 141:187–261, 1978.

[DM03a] R. G. Douglas and G. Misra. Equivalence of quotient Hilbert modules. *Proc. Indian Acad. Sci. (Math. Sci.)*, 113(3):281–291, August 2003.

[DM03b] R. G. Douglas and G. Misra. Quasi-free resolutions of Hilbert modules. *Int. Eq. Op. Th.*, 47:435–456, 2003.

[DMV00] R. G. Douglas, G. Misra, and C. Varughese. On quotient modules - the case of arbitrary multiplicity. *J. Funct. Anal.*, 174:364–398, 2000.

[DMV01] R. G. Douglas, G. Misra, and C. Varughese. Some geometric invariants from resolutions of Hilbert modules. In *Operator Theory: Advances and Applications*, volume 129, pages 241–270. Birkhauser, Basel, 2001.

[Eis94] D. Eisenbud. *Commutative Algebra with a view toward algebraic geometry*, volume 150 of *Graduate Texts in Mathematics*. Springer Verlag, 1994.

[Eis04] D. Eisenbud. *The Geometry of Syzygies: A Second Course in Commutative Algebra and Algebraic geometry.* Springer Verlag, 2004. manuscript in preparation, `http://www.msri.org/people/staff/de/ready.pdf`.

[Gre03] D. Greene. On free resolutions in multivariable operator theory. *J. Funct. Anal.*, 200(2):429–450, 2003. arXiv:math.FA/0109111 v2.

[Hil90] D. Hilbert. Über die Theorie von algebraischen Formen. *Math. Ann.*, 36:473–534, 1890.

[Hil93] D. Hilbert. Über die vollen Invariantensysteme. *Math. Ann.*, 42:313–373, 1893.

[Nor76] D. G. Northcott. *Finite Free Resolutions*, volume 71 of *Cambridge Tracts in Mathematics*. Cambridge, Cambridge-New York-Melbourne, 1976.

[Ser00] J.-P. Serre. *Local Algebra*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000.

[Tay70a] J. L. Taylor. The analytic functional calculus for several commuting operators. *Acta Math*, 125:1–38, 1970. MR 42 6622.

[Tay70b] J. L. Taylor. A joint spectrum for several commuting operators. *J. Funct. Anal.*, 6:172–191, 1970. MR 42 3603.

William Arveson
Department of Mathematics
University of California
Berkeley, CA 94720
arveson@math.berkeley.edu

Classification of Holomorphic Vector Bundles on Noncommutative Two-Tori

A. Polishchuk[1]

Received: August 27, 2003
Revised: May 18, 2004

Communicated by Joachim Cuntz

Abstract. We prove that every holomorphic vector bundle on a noncommutative two-torus T can be obtained by successive extensions from standard holomorphic bundles considered in [2]. This implies that the category of holomorphic bundles on T is equivalent to the heart of a certain t-structure on the derived category of coherent sheaves on an elliptic curve.

1. Introduction

In this paper we continue the study of holomorphic bundles on noncommutative two-tori that was begun in [2]. Recall that for every $\theta \in \mathbb{R} \setminus \mathbb{Q}$ and $\tau \in \mathbb{C} \setminus \mathbb{R}$ we considered in [2] holomorphic vector bundles on a noncommutative complex torus $T = T_{\theta,\tau}$. By definition, the algebra A_θ of smooth functions on T consists of series $\sum_{(m,n)\in\mathbb{Z}^2} a_{m,n}U_1^m U_2^n$ where the coefficients $a_{m,n} \in \mathbb{C}$ decrease rapidly at infinity and the multiplication is defined using the rule

$$U_1U_2 = \exp(2\pi i\theta)U_2U_1.$$

We consider the derivation $\delta = \delta_\tau : A_\theta \to A_\theta$ defined by

$$\delta(\sum a_{m,n}U_1^m U_2^n) = 2\pi i \sum_{m,n}(m\tau + n)a_{m,n}U_1^m U_2^n$$

as an analogue of the $\overline{\partial}$-operator. A holomorphic bundle over T is a pair $(E, \overline{\nabla})$ consisting of a finitely generated projective right A_θ-module E and an operator $\overline{\nabla} : E \to E$ satisfying the Leibnitz identity

$$\overline{\nabla}(ea) = \overline{\nabla}(e)a + e\delta(a),$$

where $e \in E$, $a \in A_\theta$. There is an obvious definition of a holomorphic map between holomorphic bundles, so we can define the category $\mathcal{C}(T)$ of holomorphic bundles on T.

[1]This work was partially supported by NSF grant DMS-0302215

For every pair of relatively prime integers (c,d) such that $c\theta+d>0$ and a complex number z we define a *standard holomorphic bundle* $(E_{d,c}(\theta),\overline{\nabla}_z)$ as follows. If $c\neq 0$ then

$$E_{d,c}(\theta)=\mathcal{S}(\mathbb{R}\times\mathbb{Z}/c\mathbb{Z})=\mathcal{S}(\mathbb{R})^{|c|},$$

where $\mathcal{S}(\mathbb{R})$ is the Schwartz space of functions on $\mathbb{R}$, with the A_θ-action defined by

$$fU_1(x,\alpha)=f(x-\frac{1}{\mu},\alpha-1),\ fU_2(x,\alpha)=\exp(2\pi i(x-\frac{\alpha d}{c}))f(x,\alpha),$$

where $x\in\mathbb{R}$, $\alpha\in\mathbb{Z}/c\mathbb{Z}$, $\mu=\frac{c}{c\theta+d}$. The operator $\overline{\nabla}_z$ on this space is given by

$$\overline{\nabla}_z(f)=\frac{\partial f}{\partial x}+2\pi i(\tau\mu x+z)f. \tag{1.1}$$

For $c=0$ and $d=1$ we set $E_{1,0}(\theta)=A_\theta$ with the natural right A_θ-action and the operator $\overline{\nabla}_z$ is given by

$$\overline{\nabla}_z(a)=\delta(a)+2\pi iza.$$

We define degree, rank and slope of a bundle $E=E_{d,c}(\theta)$ by setting $\deg(E)=c$, $\operatorname{rk}(E)=c\theta+d$ and $\mu(E)=\deg(E)/\operatorname{rk}(E)$. Note that $\operatorname{rk}(E)>0$ and $\mu=\mu(E)$ in the formulae above.

According to the theorem of Rieffel (see [5]) every finitely generated projective right A_θ-module is isomorphic to $E=E_{d,c}(\theta)^{\oplus n}$ for some (c,d) as above and $n\geq 0$. Moreover, the degree and rank defined above extend to additive functions on the category of finitely generated projective A_θ-modules.

The category of holomorphic bundles $\mathcal{C}=\mathcal{C}(T)$ has a natural structure of a $\mathbb{C}$-linear exact category. In particular, for every pair of holomorphic bundles E_1 and E_2 we can form the vector space $\operatorname{Ext}^1_{\mathcal{C}}(E_1,E_2)$ parametrizing extensions of E_1 by E_2. Sometimes we will also use the notation $\operatorname{Ext}^0_{\mathcal{C}}:=\operatorname{Hom}_{\mathcal{C}}$. Let $\mathcal{C}'\subset\mathcal{C}$ be the minimal full subcategory of $\mathcal{C}$ containing all standard holomorphic bundles and closed under extensions. Our main result is the following theorem.

Theorem 1.1. *One has $\mathcal{C}'=\mathcal{C}$.*

Combining this theorem with the study of the category $\mathcal{C}'$ in [2] we obtain the following result.

Corollary 1.2. *The category $\mathcal{C}$ is abelian. It is equivalent to the heart $\mathcal{C}^\theta$ of the t-structure on the derived category of coherent sheaves on the elliptic curve $\mathbb{C}/\mathbb{Z}+\mathbb{Z}\tau$, associated with θ (see section 3 of* [2] *or section 3.1 below).*

Remark. Recall that we always assume θ to be irrational. For rational θ the category $\mathcal{C}$ will not be abelian.

Corollary 1.3. *For every indecomposable holomorphic bundle E on T there exists a standard holomorphic bundle $\overline{E}$ and a filtration $0=E_0\subset E_1\subset\ldots\subset E_n=E$ by holomorphic subbundles such that all quotients E_i/E_{i-1} are isomorphic to $\overline{E}$.*

The proof of Theorem 1.1 consists of two steps. First, we develop the cohomology theory for holomorphic bundles on T and prove the analogues of the standard theorems for them (such as finiteness, Riemann-Roch and Serre duality). Then we combine these results with the techniques of [4] where the category $\mathcal{C}'$ was described in terms of coherent modules over a certain algebra.

Acknowledgments. Parts of this paper were written during the author's visits to Max-Planck-Institut für Mathematik in Bonn and the Institut des Hautes Études Scientifiques. I'd like to thank these institutions for hospitality and support.

2. Cohomology of holomorphic bundles on noncommutative two-tori

2.1. Cohomology and Ext-spaces. Let $(E,\overline{\nabla})$ be a holomorphic bundle on $T=T_{\theta,\tau}$. Then the *cohomology* of E is defined by

$$H^i(E)=H^i(E,\overline{\nabla})=H^i(E\stackrel{\overline{\nabla}}{\to}E)$$

where $i=0$ or $i=1$. Thus, $H^0(E)=\ker(\overline{\nabla})$, $H^1(E)=\operatorname{coker}(\overline{\nabla})$. These spaces are closely related to Ext^i-spaces in the category of holomorphic bundles (where $i=0$ or $i=1$). To explain this connection we have to use Morita equivalences between noncommutative tori. Recall that for every standard bundle $E_0=E_{d,c}(\theta)$ the algebra of endomorphisms $\operatorname{End}_{A_\theta}(E_0)$ can be identified with the algebra $A_{\theta'}$ for some $\theta'\in\mathbb{R}$. In fact, $\theta'=\frac{a\theta+b}{c\theta+d}$, where a and b are chosen in such a way that $\begin{pmatrix}a & b\\ c & d\end{pmatrix}\in\mathrm{SL}_2(\mathbb{Z})$. Furthermore, if E_0 is equipped with a standard holomorphic structure $\overline{\nabla}$ then the formula $\phi\mapsto[\overline{\nabla},\phi]$ defines a derivation of $\operatorname{End}_{A_\theta}(E_0)\simeq A_{\theta'}$, hence the corresponding torus $T_{\theta'}$ is equipped with a complex structure. In fact, this derivation on $A_{\theta'}$ is equal to $\delta_\tau/\operatorname{rk}(E_0)$, where τ is the same parameter that was used to define the complex structure on T_θ (see Proposition 2.1 of [2]). Now one can define the Morita equivalence

$$\mathcal{C}(T_{\theta',\tau})\to\mathcal{C}(T_{\theta,\tau}):E\mapsto E\otimes_{A_{\theta'}}E_0,$$

where the tensor product is equipped with the complex structure

$$\overline{\nabla}(e\otimes e_0)=\frac{1}{\operatorname{rk}(E_0)}\overline{\nabla}_E(e)\otimes e_0+e\otimes\overline{\nabla}_{E_0}(e_0)$$

(see Propositions 2.1 and 3.2 of [2]). This functor sends standard holomorphic bundles on $T_{\theta',\tau}$ to standard holomorphic bundles on $T_{\theta,\tau}$. The inverse functor is

$$\mathcal{C}(T_{\theta,\tau})\to\mathcal{C}(T_{\theta',\tau}):E\mapsto\operatorname{Hom}_{A_\theta}(E_0,E),\tag{2.1}$$

where the latter space has a natural right action of $A_{\theta'}\simeq\operatorname{End}_{A_\theta}(E_0)$. Now we can formulate the connection between the cohomology and Ext-groups. For every holomorphic bundle E and a standard holomorphic bundle E_0 on $T=$

$T_{\theta,\tau}$ one has a natural isomorphism

$$\mathrm{Ext}^i_{\mathcal{C}}(E_0, E) \simeq H^i(\mathrm{Hom}_{A_\theta}(E_0, E)), \tag{2.2}$$

where $\mathrm{Hom}_{A_\theta}(E_0, E)$ is viewed as a holomorphic bundle on $T_{\theta'}$ (the proof is similar to Proposition 2.4 of [2]). Note that for an arbitrary pair of holomorphic bundles E_1 and E_2 one can still define an operator $\overline{\nabla}$ on $\mathrm{Hom}_{A_\theta}(E_1, E_2)$ such that the analogue of isomorphism (2.2) holds. However, we have a natural interpretation of $\mathrm{Hom}_{A_\theta}(E_1, E_2)$ as a holomorphic bundle on some noncommutative two-torus only in the case when one of the bundles E_1 or E_2 is standard (see (2.3) below).

2.2. Duality and metrics. One can define the category of *left holomorphic bundles* on T by replacing right A_θ-modules with left ones and changing the Leibnitz identity appropriately. The definition of cohomology for these bundles remains the same. There is a natural duality functor $E \mapsto E^\vee$ that associates to a (right) holomorphic bundle E the left holomorphic bundle $\mathrm{Hom}_{A_\theta}(E, A_\theta)$. More generally, for every standard holomorphic bundle E_0 we can consider $\mathrm{Hom}_{A_\theta}(E, E_0)$ as a left module over $\mathrm{End}_{A_\theta}(E_0) \simeq A_{\theta'}$ equipped with an induced holomorphic structure. Then the natural isomorphism

$$\mathrm{Ext}^i_{\mathcal{C}}(E, E_0) \simeq H^i(\mathrm{Hom}_{A_\theta}(E, E_0)) \tag{2.3}$$

allows to view $\mathrm{Ext}^i_{\mathcal{C}}(E, E_0)$ as cohomology of a holomorphic bundle on $T_{\theta'}$. Using duality, the functor (2.1) for a standard holomorphic bundle E_0 can be rewritten as the usual Morita functor due to the isomorphism

$$\mathrm{Hom}_{A_\theta}(E_0, E) \simeq E \otimes_{A_\theta} E_0^\vee. \tag{2.4}$$

For a standard holomorphic bundle E_0 on $T_{\theta,\tau}$ the dual bundle $E_0^\vee$ can also be considered as a right holomorphic bundle on $T_{\theta',\tau}$, where $\mathrm{End}_{A_\theta}(E_0) \simeq A_{\theta'}$. In fact, it is again a standard holomorphic bundle (see Corollary 2.3 of [2]). More precisely, for $E_0 = E_{d,c}(\theta)$ we have $\theta' = \frac{a\theta+b}{c\theta+d}$, where $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. The left action of $A_{\theta'}$ on $E_{d,c}(\theta) = \mathcal{S}(\mathbb{R} \times \mathbb{Z}/c\mathbb{Z})$ (where $c \neq 0$) is defined by the formulae

$$U_1 f(x,\alpha) = f(x - \frac{1}{c}, \alpha - a),\ U_2 f(x,\alpha) = \exp(2\pi i(\frac{x}{c\theta+d} - \frac{\alpha}{c}))f(x,\alpha),$$

where $x \in \mathbb{R}$, $\alpha \in \mathbb{Z}/c\mathbb{Z}$. We can identify $E_{d,c}(\theta)^\vee$ considered as a right $A_{\theta'}$-module with $E_{a,-c}(\theta')$ using the natural pairing

$$t : E_{a,-c}(\theta') \otimes E_{d,c}(\theta) \to A_\theta$$

constructed as follows (see Proposition 1.2 of [2]). First, we define the map

$$b : E_{a,-c}(\theta') \otimes E_{d,c}(\theta) \to \mathbb{C}$$

by the formula

$$b(f_1, f_2) = \sum_{\alpha \in \mathbb{Z}/c\mathbb{Z}} \int_{x \in \mathbb{R}} f_1(\frac{x}{c\theta+d}, \alpha) f_2(x, -a\alpha) dx.$$

Then t is given by

$$t(f_1, f_2) = \sum_{(m,n)\in\mathbb{Z}^2} U_1^m U_2^n b(U_2^{-n} U_1^{-m} f_1 \otimes f_2).$$

The corresponding isomorphism

$$E_{d,c}(\theta)^{\vee} \simeq E_{a,-c}(\theta')$$

is compatible with the $A_\theta - A_{\theta'}$-bimodule structures and with holomorphic structures (see Corollary 2.3 of [2]). Note that $b = \mathrm{tr} \circ t$, where $\mathrm{tr} : A_\theta \to \mathbb{C}$ is the trace functional sending $\sum a_{m,n} U_1^m U_2^n$ to $a_{0,0}$.
On the other hand, we can define a $\mathbb{C}$-antilinear isomorphism

$$\sigma : E_{d,c}(\theta) \to E_{a,-c}(\theta')$$

by the formula

$$\sigma(f)(x, \alpha) = \overline{f((c\theta + d)x, -a\alpha)}.$$

This isomorphism satisfies

$$\sigma(bea) = a^* \sigma(e) b^* \tag{2.5}$$

for $e \in E_{d,c}(\theta)$, $a \in A_\theta$, $b \in A_{\theta'}$, where $* : A_\theta \to A_\theta$ is the $\mathbb{C}$-antilinear anti-involution sending U_i to U_i^{-1}. In view of the identification of $E_{a,-c}(\theta')$ with the dual bundle to $E_{d,c}(\theta)$ the isomorphism σ should be considered as an analogue of the Hermitian metric on $E_{d,c}(\theta)$. The corresponding analogue of the scalar product on global sections is simply the Hermitian form on $E_{d,c}(\theta)$ given by the formula

$$\langle f_1, f_2 \rangle = b(\sigma(f_2), f_1) = \sum_{\alpha\in\mathbb{Z}/c\mathbb{Z}} \int_{x\in\mathbb{R}} f_1(x, \alpha)\overline{f_2(x, \alpha)} dx, \tag{2.6}$$

where $f_1, f_2 \in E_{d,c}(\theta)$. We can also define the corresponding L^2-norm: $||f||_0^2 = \langle f, f \rangle$. The above Hermitian form is related to the structure of $A_{\theta'} - A_\theta$-bimodule on $E_{d,c}(\theta)$ in the following way:

$$\langle f_1, a f_2 b \rangle = \langle a^* f_1 b^*, f_2 \rangle,$$

where $a \in A_{\theta'}$, $b \in A_\theta$ (this is a consequence of (2.5) and of Lemma 1.1 of [2]).
In the case of the trivial bundle $E_{1,0}(\theta) = A_\theta$ we can easily modify the above definitions. First of all, $\theta' = \theta$ and the dual bundle is still A_θ. The role of σ is played by $* : A_\theta \to A_\theta$ and the Hermitian form on A_θ is given by $\langle a, b \rangle = \mathrm{tr}(ab^*)$. The corresponding L^2-norm is

$$|| \sum a_{m,n} U_1^m U_2^n ||_0^2 = \sum |a_{m,n}|^2.$$

Note that the operator $\overline{\nabla}_z$ on $E_{d,c}(\theta)$ admits an adjoint operator $\overline{\nabla}_z^*$ with respect to the Hermitian metrics introduced above. Namely, for $c \neq 0$ it is given by

$$\overline{\nabla}_z^*(f) = -\frac{\partial f}{\partial x} - 2\pi i(\overline{\tau}\mu x + \overline{z}) f,$$

while for $c = 0$ we have $\overline{\nabla}_z^* = -\delta_{\overline{\tau}} - 2\pi i z \,\mathrm{id}$ on A_θ. In either case we have

$$\overline{\nabla}_z \overline{\nabla}_z^* - \overline{\nabla}_z^* \overline{\nabla}_z = \lambda \cdot \mathrm{id} \tag{2.7}$$

for some constant $\lambda \in \mathbb{R}$.

It follows that for an arbitrary holomorphic structure $\overline{\nabla}$ on $E = E_{d,c}(\theta)^{\oplus n}$ there exists an adjoint operator $\overline{\nabla}^* : E \to E$ with respect to the above Hermitian metric. Indeed, we can write $\overline{\nabla} = \overline{\nabla}_0 + \phi$, where $\overline{\nabla}_0$ is the standard holomorphic structure and set $\overline{\nabla}^* = \overline{\nabla}_0^* + \phi^*$.

2.3. Sobolev spaces. The idea to consider Sobolev spaces for bundles on noncommutative tori is due to M. Spera (see [7], [8]). Let $(E, \overline{\nabla})$ be a standard holomorphic bundle on $T_{\theta,\tau}$. For $s \in \mathbb{Z}$, $s \geq 0$ we define the s-th Sobolev norm on E by setting

$$||e||_s^2 = \sum_{i=0}^{s} ||\overline{\nabla}^i e||_0^2,$$

where $||e||_0$ is the L^2-norm on E. We define $W_s(E)$ to be the completion of E with respect to this norm. Note that there is a natural embedding $W_{s+1}(E) \subset W_s(E)$. We can define analogous spaces for $E^{\oplus n}$ in an obvious way.

All the definitions above make sense also for rational θ. Moreover, for $\theta \in \mathbb{Z}$ the space E can be identified with the space of smooth section of a holomorphic vector bundle V on an elliptic curve $\mathbb{C}/\mathbb{Z} + \mathbb{Z}\tau$ in such a way that $\overline{\nabla}$ corresponds to the $\overline{\partial}$-operator. Furthermore, the L^2-norm above corresponds to the L^2-norm with respect to a Hermitian metric on V that has constant curvature. This implies that in this case the Sobolev spaces $W_s(E)$ coincide with the corresponding Sobolev spaces constructed for the holomorphic bundle V. Indeed, using the equation (2.7) it is easy to see that the norm $||e||_s$ is equivalent to the norm given by

$$(||e||_s')^2 = \sum_{i=0}^{s} \langle e, (\overline{\nabla}^* \overline{\nabla})^i e \rangle$$

which is equivalent to the standard Sobolev norm.

An important observation is that the operator $\overline{\nabla}_z$ defined by (1.1) depends only on $\tau\mu$, where μ is the slope of the bundle, so it is the same for the bundle $E_{d,c}(\theta)$ on $T_{\theta,\tau}$ and the bundle $E_{d,c}(\mathrm{sign}(c)N)$ on the commutative torus $T_{\mathrm{sign}(c)N,\tau'}$, where N is a large enough integer so that $|c|N + d > 0$, $\tau' = (|c|N + d)/(c\theta + d)$. Therefore, the sequences of spaces $(W_s(E))$ in these two cases are the same. Hence, the following standard results about Sobolev spaces in the commutative case extend immediately to our situation (the first two are analogues of Rellich's lemma and Sobolev's lemma). In all these results E is a direct sum of a finite number of copies of a standard holomorphic bundle.

Lemma 2.1. *The embedding $W_s(E) \subset W_{s-1}(E)$ is a compact operator.*

Lemma 2.2. *One has $E = \cap_{s \geq 0} W_s(E)$.*

Lemma 2.3. *The operator $\overline{\nabla}$ extends to a bounded operator $W_s(E) \to W_{s-1}(E)$*

The following result is the only noncommutative contribution to the techniques of Sobolev spaces, however, it is quite easy.

Lemma 2.4. *For every $\phi \in \mathrm{End}_{A_\theta}(E)$ the operator $\phi : E \to E$ extends to a bounded operator $W_s(E) \to W_s(E)$.*

Proof. It suffices to prove that for every $s \geq 0$ one has

$$||\phi e||_s \leq C \cdot ||e||_s$$

for some constant $C > 0$. By our assumption $E \simeq E_0^{\oplus N}$ for some standard bundle E_0. Identifying $\mathrm{End}_{A_\theta}(E_0)$ with $A_{\theta'}$ for some $\theta' \in \mathbb{R}$ we can write $\phi = \sum a_{m,n} U_1^m U_2^n$ where U_1 and U_2 are unitary generators of $A_{\theta'}$, $a_{m,n}$ are complex $N \times N$ matrices. Since U_1 and U_2 act on E by unitary operators, it follows that

$$||\phi e||_0 \leq C(\phi) \cdot ||e||_0$$

for $e \in E$, where $C(\phi) = \sum_{m,n} ||a_{m,n}||$ (here $||a||$ denote the norm of a matrix a). Applying the Leibnitz rule repeatedly we derive similarly that

$$\sum_{i=0}^{s} ||\overline{\nabla}^i(\phi e)||_0^2 \leq \sum_{i=0}^{s} c_i \cdot ||\overline{\nabla}^i e||_0^2$$

for some constants $c_i > 0$ which implies the result. □

It is convenient to extend the definition of the chain $\ldots \subset W_1(E) \subset W_0(S)$ to the chain of embedded spaces

$$\ldots \subset W_1(E) \subset W_0(S) \subset W_{-1}(E) \subset \ldots$$

by setting $W_{-s}(E) = \overline{W_s(E)}^*$ (the space of $\mathbb{C}$-antilinear functionals) and using the natural Hermitian form of $W_0(E)$. It is easy to see that the results of this section hold for all integer values of s.

Lemma 2.5. *Let $\overline{\nabla} : E \to E$ be a (not necessarily standard) holomorphic structure on E. Then the operators $\overline{\nabla}$ and $\overline{\nabla}^*$ can be extended to bounded operators $W_s(E) \to W_{s-1}(E)$ for every $s \in \mathbb{Z}$.*

Proof. Let $\overline{\nabla}_0$ be a standard holomorphic structure on E. Then $\overline{\nabla} = \overline{\nabla}_0 + \phi$ for some $\phi \in \mathrm{End}_{A_\theta}(E)$. By Lemma 2.3 (resp., Lemma 2.4) there exist a continuous extension $\overline{\nabla}_0 : W_s(E) \to W_{s-1}(E)$ (resp., $\phi : W_s(E) \to W_s(E)$). Hence, $\overline{\nabla}$ extends to a family of continuous operators $\overline{\nabla}(s) : W_s(E) \to W_{s-1}(E)$ for $s \in \mathbb{Z}$. The extensions of $\overline{\nabla}^*$ are given by the adjoint operators $\overline{\nabla}(-s+1)^* : W_s(E) \to W_{s-1}(E)$. □

2.4. Applications to cohomology. We begin our study of cohomology with standard holomorphic bundles.

Proposition 2.6. *Let $(E, \overline{\nabla})$ be a direct sum of several copies of a standard holomorphic bundle on $T_{\theta,\tau}$.*

(i) The cohomology spaces $H^0(E)$ *and* $H^1(E)$ *are finite-dimensional and for* $\mathrm{Im}(\tau)<0$ *one has*

$$\chi(E)=\dim H^0(E)-\dim H^1(E)=\deg(E).$$

(ii) There exists an operator $Q:E\to E$ *such that*

$$\mathrm{id}-Q\overline{\nabla}=\pi_{\ker\overline{\nabla}},$$

$$\mathrm{id}-\overline{\nabla}Q=\pi_{\overline{\nabla}(E)^{\perp}},$$

where $\overline{\nabla}(E)^{\perp}\subset E$ *is the orthogonal complement to* $\overline{\nabla}(E)\subset E$*, for a finite-dimensional subspace* $V\subset E$ *we denote by* $\pi_V:E\to V$ *the orthogonal projection.*
(iii) If $\mathrm{Im}(\tau)<0$ *and* $\deg(E)>0$ *then* $H^1(E)=0$.
(iv) The operator $Q:E\to E$ *extends to a bounded operator* $W_s(E)\to W_{s+1}(E)$.
(v) For every $e\in W_0(E)$ *one has*

$$||Qe||_0\le\frac{1}{2\sqrt{\pi|\,\mathrm{Im}(\tau)\mu(E)|}}||e||_0.$$

Proof. In the commutative case the assertions (i)-(iii) are well known. For example, the operator Q is given by $\overline{\partial}^*G$, where G is the Green operator for the $\overline{\partial}$-Laplacian. The condition $\mathrm{Im}(\tau)<0$ corresponds to the way we define the operator δ_τ on A_θ (see Proposition 3.1 of [2]). As before we can deduce (i)-(iii) in general from the commutative case. One can also prove these assertions directly in the noncommutative case (see Proposition 2.5 of [2] for the proofs of (i) and (iii)). The assertion (iv) follows immediately from the identity

$$\overline{\nabla}^nQ=\overline{\nabla}^{n-1}(\mathrm{id}-\pi_{\overline{\nabla}(E)^{\perp}}).$$

To prove (v) we can assume that $E=E_{d,c}(\theta)$, where $c\ne0$ and $\overline{\nabla}=\overline{\nabla}_z$ for some $z\in\mathbb{C}$. Then the space $W_0(E)$ is the orthogonal sum of $|c|$ copies of $L^2(\mathbb{R})$. Moreover, the operator $\overline{\nabla}$ respects this decomposition and restricts to the operator

$$f\mapsto f'+(ax+z)f$$

on each copy, where $a=2\pi i\tau\mu(E)$. Hence, the operator Q also respects this decomposition and it suffices to consider its restriction to one copy of $L^2(\mathbb{R})$. Since $\mathrm{Re}(a)\ne0$, by making the unitary transformation of the form $\widetilde{f}(x)=\exp(itx)f(x+t')$ for some $t,t'\in\mathbb{R}$ we can reduce ourselves to the case $z=0$. Furthermore, the transformation of the form $\widetilde{f}=\exp(i\,\mathrm{Im}(a)x^2/2)f$ gives a unitary equivalence with the operator $\overline{\nabla}:f\mapsto f'+\lambda xf$ where $\lambda=\mathrm{Re}(a)$. Consider the following complete orthogonal system of functions in $L^2(\mathbb{R})$:

$$f_n(x)=H_n(\sqrt{|\lambda|}x)\exp(-|\lambda|\frac{x^2}{2}),\ n=0,1,2,\dots,$$

where $H_n(x) = (-1)^n \exp(x^2)\frac{d^n}{dx^n}(\exp(-x^2))$ are Hermite polynomials ((f_n) is an eigenbasis of the operator $f \mapsto -f'' + \lambda^2 x^2$). Note that

$$||f_n||_0^2 = \frac{1}{\sqrt{\lambda}} \int_{\mathbb{R}} H_n(x)^2 \exp(-x^2)dx = \frac{2^n \cdot n! \cdot \sqrt{\pi}}{\sqrt{\lambda}}.$$

Assume first that $\lambda > 0$. Then using the formula $H_n'(x) = 2nH_{n-1}(x)$ we obtain

$$\overline{\nabla}(f_n) = 2n\sqrt{\lambda} f_{n-1}$$

for $n > 0$ and $\overline{\nabla}(f_0) = 0$. Therefore, in this case

$$Q(f_n) = \frac{1}{(2n+2)\sqrt{\lambda}} f_{n+1}$$

for all $n \geq 0$. Hence,

$$\frac{||Q(f_n)||_0}{||f_n||_0} = \frac{||f_{n+1}||_0}{(2n+2)\sqrt{\lambda}||f_n||_0} = \frac{1}{\sqrt{(2n+2)\lambda}} \leq \frac{1}{\sqrt{2\lambda}}$$

which implies (v) in this case.
Now assume that $\lambda < 0$. Then using the formula $H_n'(x) = 2xH_n(x) - H_{n+1}(x)$ we find

$$\overline{\nabla}(f_n) = -\sqrt{|\lambda|} f_{n+1}.$$

Hence,

$$Q(f_n) = -\frac{1}{\sqrt{|\lambda|}} f_{n-1}$$

for $n > 0$ and $Q(f_0) = 0$. It follows that

$$\frac{||Q(f_n)||_0}{||f_n||_0} = \frac{||f_{n-1}||_0}{\sqrt{|\lambda|}||f_n||_0} = \frac{1}{\sqrt{2n|\lambda|}} \leq \frac{1}{\sqrt{2|\lambda|}}$$

for $n \geq 1$, which again implies our statement. □

Now we are ready to prove results about cohomology of arbitrary holomorphic bundles. We will use the following well known lemma.

Lemma 2.7. *Let $L : W \to W'$ and $L' : W' \to W$ be bounded operators between Banach spaces such that $L'L = \mathrm{id} + C$, $LL' = \mathrm{id} + C'$ for some compact operators $C : W \to W$ and $C' : W' \to W'$. Then the operator L is Fredholm.*

Theorem 2.8. *(i) For every holomorphic bundle $(E, \overline{\nabla})$ on $T_{\theta,\tau}$ the spaces $H^0(E)$ and $H^1(E)$ are finite-dimensional.*
(ii) If $\operatorname{Im} \tau < 0$ then

$$\chi(E) = \dim H^0(E) - \dim H^1(E) = \deg(E).$$

(iii) Let us equip E with a metric by identifying it with the direct sum of several copies of a standard bundle. Then one has the following orthogonal decompositions

$$E = \ker(\overline{\nabla}) \oplus \overline{\nabla}^*(E),$$
$$E = \ker(\overline{\nabla}^*) \oplus \overline{\nabla}(E),$$

where $\overline{\nabla}^* : E \to E$ is the adjoint operator to $\overline{\nabla}$.

Proof. Let us write the holomorphic structure on E in the form

$$\overline{\nabla} = \overline{\nabla}_0 + \phi,$$

where $(E, \overline{\nabla}_0)$ is holomorphically isomorphic to the direct sum of several copies of a standard holomorphic bundle, $\phi \in \operatorname{End}_{A_\theta}(E)$. By Lemma 2.5, $\overline{\nabla}$ has a bounded extension to an operator $\overline{\nabla} : W_s(E) \to W_{s-1}(E)$ for every $s \in \mathbb{Z}$.
Consider the operator $Q : E \to E$ constructed in Proposition 2.6 for the holomorphic structure $\overline{\nabla}_0$. Then Q extends to a bounded operator $W_s(E) \to W_{s+1}(E)$ for every $s \in \mathbb{Z}$. We have

$$Q\overline{\nabla} = Q\overline{\nabla}_0 + Q\phi = \mathrm{id} - \pi_0 + Q\phi,$$

where π_0 is the orthogonal projection to the finite-dimensional space $\ker(\overline{\nabla}_0) \subset E$. Clearly, π_0 defines a bounded operator $W_0(E) \to \ker(\overline{\nabla}_0)$. Hence, the operator $C = Q\overline{\nabla} - \mathrm{id} : W_s(E) \to W_s(E)$ factors as a composition of some bounded operator $W_s(E) \to W_{s+1}(E)$ with the embedding $W_{s+1}(E) \to W_s(E)$. By Lemma 2.1 this implies that C is a compact operator. Similarly, the operator $C' = \overline{\nabla}Q - \mathrm{id} : W_s(E) \to W_s(E)$ is compact. Applying Lemma 2.7 we deduce that $\overline{\nabla} : W_s(E) \to W_{s-1}(E)$ is a Fredholm operator. This immediately implies that $H^0(E)$ is finite-dimensional. Moreover, we claim that

$$\ker(\overline{\nabla} : W_s(E) \to W_{s-1}(E)) = H^0(E) \subset E$$

for any $s \in \mathbb{Z}$. Indeed, it suffices to check that if $\overline{\nabla}(e) = 0$ for $e \in W_s(E)$ then $e \in E$. Let us prove by induction in $t \geq s$ that $e \in W_t(E)$. Assume that this is true for some t. Then

$$e = Q\overline{\nabla}(e) - C(e) = -C(e) \in W_{t+1}(E).$$

Since $\cap_t W_t(E) = E$ by Lemma 2.2 we conclude that $e \in E$.
Let $Q^* : W_s(E) \to W_{s+1}(E)$ be the adjoint operator to $Q : W_{-s-1}(E) \to W_{-s}(E)$, where $s \in \mathbb{Z}$. Then the operators $Q^*\overline{\nabla}^* - \mathrm{id} = (C')^*$ and $\overline{\nabla}^* Q^* - \mathrm{id} = C^*$ are compact. Thus, the same argument as before shows that for every $s \in \mathbb{Z}$ the operator $\overline{\nabla}^* : W_s(E) \to W_{s-1}(E)$ is Fredholm and one has

$$\ker(\overline{\nabla}^* : W_s(E) \to W_{s-1}(E)) \subset E.$$

Next we claim that

$$E = \ker(\overline{\nabla}^*) \oplus \overline{\nabla}(E).$$

Since the orthogonal complement to $\ker(\overline{\nabla}^*)$ in $W_0(E)$ coincides with the image of $\overline{\nabla} : W_1(E) \to W_0(E)$, it suffices to prove that $E \cap \overline{\nabla}(W_1(E)) \subset \overline{\nabla}(E)$. But if $e = \overline{\nabla}(e_1)$ for some $e \in E$, $e_1 \in W_1(E)$, then we can easily prove by induction in $s \geq 1$ that $e_1 \in W_s(E)$. Indeed, assuming that $e_1 \in W_s(E)$ we have

$$e_1 = Q\overline{\nabla}(e_1) - C(e_1) = Q(e) - C(e_1) \in W_{s+1}(E).$$

A similar argument using the operator Q^* shows that

$$E = \ker(\overline{\nabla}) \oplus \overline{\nabla}^*(E).$$

Thus, we checked that $H^0(E)$ and $H^1(E)$ are finite-dimensional and that $\chi(E)$ coincides with the index of the Fredholm operator $\overline{\nabla} = \overline{\nabla}_0 + \phi : W_1(E) \to W_0(E)$. Note that

$$\overline{\nabla}_t := \overline{\nabla}_0 + t\phi : W_1(E) \to W_0(E)$$

is a continuous family of Fredholm operators depending on $t \in [0,1]$. It follows that the index of $\overline{\nabla} = \overline{\nabla}_1$ is equal to the index of $\overline{\nabla}_0$ computed in Proposition 2.6. □

Corollary 2.9. *If one of the holomorphic bundles E_1 and E_2 is standard then the spaces $\mathrm{Hom}_{\mathcal{C}}(E_1, E_2)$ and $\mathrm{Ext}^1_{\mathcal{C}}(E_1, E_2)$ are finite-dimensional and*

$$\begin{aligned}\chi(E_1, E_2) &:= \dim \mathrm{Hom}_{\mathcal{C}}(E_1, E_2) - \dim \mathrm{Ext}^1_{\mathcal{C}}(E_1, E_2) \\ &= \mathrm{rk}(E_1)\deg(E_2) - \mathrm{rk}(E_2)\deg(E_1).\end{aligned}$$

The following vanishing result will play a crucial role in the proof of Theorem 1.1.

Theorem 2.10. *Assume that $\mathrm{Im}(\tau) < 0$. For every holomorphic bundle E on $T = T_{\theta,\tau}$ there exists a constant $C = C(E) \in \mathbb{R}$ such that for every standard holomorphic bundle E_0 on T with $\mu(E_0) < C$ one has $\mathrm{Ext}^1_{\mathcal{C}}(E_0, E) = 0$.*

Proof. Let us choose a (non-holomorphic) isomorphism $E \simeq E_1^{\oplus N}$, where E_1 is a standard holomorphic bundle on $T_{\theta,\tau}$. Then we can write the holomorphic structure on E as

$$\overline{\nabla}_E = \overline{\nabla}_0 + \phi,$$

where $\overline{\nabla}_0$ comes from the standard holomorphic structure on E_1, $\phi \in \mathrm{End}_{A_\theta}(E)$. Then for every standard holomorphic bundle E_0 we can consider the holomorphic bundle $E' = \mathrm{Hom}_{A_\theta}(E_0, E)$ on $T_{\theta',\tau}$, where $\mathrm{End}_{A_\theta}(E_0) = A_{\theta'}$. Note that $\mathrm{Ext}^1_{\mathcal{C}}(E_0, E) \simeq H^1(E')$, so we want to prove that the latter group vanishes for $\mu(E_0) << 0$. Recall that the holomorphic structure $\overline{\nabla}'$ on E' is given by

$$\overline{\nabla}'(f)(e_0) = \mathrm{rk}(E_0) \cdot [\overline{\nabla}(f(e_0)) - f(\overline{\nabla}_{E_0}(e_0))],$$

where $e_0 \in E_0$, $f \in E'$ (see section 2.2 of [2]). The isomorphism $E \simeq E_1^{\oplus N}$ induces an isomorphism $E' \simeq (E_1')^{\oplus N}$, where E_1' is the standard bundle $\mathrm{Hom}_{A_\theta}(E_0, E_1)$ on $T_{\theta',\tau}$. Therefore, we have

$$\overline{\nabla}' = \overline{\nabla}'_0 + \mathrm{rk}(E_0)\phi,$$

where $\overline{\nabla}'_0$ corresponds to the standard holomorphic structure on $(E_1')^{\oplus N}$ and ϕ is now considered as an $A_{\theta'}$-linear endomorphism of E'. Note that by Proposition 2.6(iii) we have $H^1(E', \overline{\nabla}'_0) = 0$ as long as $\mu(E') > 0$. It is easy to compute that

$$\mathrm{rk}(E') = \mathrm{rk}(E)/\mathrm{rk}(E_0),\ \deg(E') = \mathrm{rk}(E)\,\mathrm{rk}(E_0)(\mu(E) - \mu(E_0)),$$

hence

$$\mu(E') = \mathrm{rk}(E_0)^2(\mu(E) - \mu(E_0)).$$

Therefore, $\mu(E') > 0$ provided that $\mu(E_0) < \mu(E)$. In this case the operator Q on E' constructed in Proposition 2.6 for the standard holomorphic structure $\overline{\nabla}'_0$ satisfies $\overline{\nabla}'_0 Q = \mathrm{id}$. Let C_0 be a constant such that

$$||Qe'||_0 \le \frac{C_0}{\sqrt{\mu(E')}}||e'||_0$$

for $e' \in W_0(E')$ (see Proposition 2.6(v)). Also, let us write $\phi \in \mathrm{End}_{A_\theta}(E) \simeq \mathrm{End}_{A_\theta}(E_1^{\oplus N})$ in the form $\phi = \sum a_{m,n}U_1^m U_2^n$, where U_1 and U_2 are unitary generators of $\mathrm{End}_{A_\theta}(E_1)$ and $a_{m,n}$ are $N \times N$ complex matrices. Then we set $C(\phi) = \sum_{m,n} ||a_{m,n}||$ (the sum of the matrix norms of all coefficients). Now we choose the constant $C < \mu(E)$ in such a way that

$$\frac{C_0 C(\phi)}{\sqrt{\mu(E) - C}} < 1.$$

Then for $\mu(E_0) < C$ we will have

$$\mathrm{rk}(E_0) \cdot ||\phi Q e'||_0 \le \frac{\mathrm{rk}(E_0) C_0 C(\phi)}{\sqrt{\mu(E')}}||e'||_0 < \frac{C_0 C(\phi)}{\sqrt{\mu(E) - C}}||e'||_0 < r \cdot ||e'||_0$$

for some $0 < r < 1$. It follows from the above estimate that the operator $\mathrm{id} + \mathrm{rk}(E_0)\phi Q : W_0(E') \to W_0(E')$ is invertible. Therefore, we can define the operator

$$\widetilde{Q} = Q(\mathrm{id} + \mathrm{rk}(E_0)\phi Q)^{-1} : W_0(E') \to W_1(E')$$

that satisfies

$$(\overline{\nabla}'_0 + \mathrm{rk}(E_0)\phi)\widetilde{Q} = \mathrm{id}\,.$$

Hence, the operator $\overline{\nabla}' = \overline{\nabla}'_0 + \mathrm{rk}(E_0)\phi : W_1(E') \to W_0(E')$ is surjective. But $H^1(E')$ can be identified with the cokernel of this operator (see the proof of Theorem 2.8), so $H^1(E') = 0$. □

2.5. Serre duality. For every holomorphic bundle E we have a natural pairing

$$E \otimes_{A_\theta} E^\vee \to A_\theta : e \otimes f \mapsto f(e).$$

It is compatible with the $\overline{\partial}$-operators, so it induces a pairing

$$H^{1-i}(E) \otimes H^i(E^\vee) \to H^1(A_\theta) \simeq \mathbb{C} \tag{2.8}$$

for $i = 0, 1$.

Theorem 2.11. *The pairing (2.8) is perfect.*

Proof. Since we can switch E and $E^\vee$, it suffices to consider the case $i = 0$. We choose an isomorphism of E with a direct sum of several copies of a standard holomorphic bundle E_0, so that we can talk about standard metrics and Sobolev spaces. Note that the isomorphism $H^1(A_\theta)$ is induced by the trace functional $\mathrm{tr} : A_\theta \to \mathbb{C} : \sum a_{m,n}U_1^m U_2^n \mapsto a_{0,0}$. Hence, the pairing (2.8) is induced by the pairing

$$b : E \otimes E^\vee \to \mathbb{C} : e \otimes f \mapsto \mathrm{tr}(f(e))$$

that satisfies the identity

$$\mathrm{rk}(E_0)b(\overline{\nabla}_E(e), e^\vee) + b(e, \overline{\nabla}_{E^\vee}(e^\vee)) = 0, \tag{2.9}$$

where $e \in E$, $e^\vee \in E^\vee$ (see Proposition 2.2 of [2]). According to Theorem 2.8(iii) we have orthogonal decompositions

$$E^\vee = \ker(\overline{\nabla}_{E^\vee}) \oplus \overline{\nabla}^*_{E^\vee}(E^\vee),$$

$$E = \ker(\overline{\nabla}^*_E) \oplus \overline{\nabla}_E(E).$$

Therefore, it suffices to check that b induces a perfect pairing between $\ker(\overline{\nabla}^*_E)$ and $\ker(\overline{\nabla}_{E^\vee})$. Let $\sigma : E \to E^\vee$ be the $\mathbb{C}$-antilinear isomorphism defined in section 2.2. We claim that σ maps $\ker(\overline{\nabla}^*_E)$ isomorphically onto $\ker(\overline{\nabla}_{E^\vee})$. Since $b(e_1, \sigma(e_2)) = \langle e_1, e_2\rangle$ for $e_1, e_2 \in E$, the theorem would immediately follow this. To prove the claim it is enough to check that $\overline{\nabla}_{E^\vee} = -\mathrm{rk}(E_0)\sigma\overline{\nabla}^*_E\sigma^{-1}$. To this end let us rewrite (2.9) as follows:

$$\mathrm{rk}(E_0)\langle\overline{\nabla}_E(e), \sigma^{-1}e^\vee\rangle = -\langle e, \sigma^{-1}\overline{\nabla}_{E^\vee}(e^\vee)\rangle.$$

Since the left-hand side is equal to $\mathrm{rk}(E_0)\langle e, \overline{\nabla}^*_E\sigma^{-1}e^\vee\rangle$ we conclude that $\mathrm{rk}(E_0)\overline{\nabla}^*_E\sigma^{-1} = -\sigma^{-1}\overline{\nabla}_{E^\vee}$ as required. □

Recall that we denote by $\mathcal{C}' \subset \mathcal{C}$ the full subcategory consisting of all successive extensions of standard holomorphic bundles. Theorem 3.8 of [2] implies that the derived category of $\mathcal{C}'$ is equivalent to the derived category of coherent sheaves on an elliptic curve. Therefore, the standard Serre duality gives a functorial isomorphism

$$\mathrm{Ext}^1_{\mathcal{C}'}(E_1, E_2) \simeq \mathrm{Hom}_{\mathcal{C}'}(E_2, E_1)^*$$

for $E_1, E_2 \in \mathcal{C}'$. Note that we can replace here $\mathrm{Ext}^i_{\mathcal{C}'}$ with $\mathrm{Ext}^i_{\mathcal{C}}$. Now using the above theorem we can extend this isomorphism to the case when only one of the objects E_1, E_2 belongs to $\mathcal{C}'$.

Corollary 2.12. *For every holomorphic bundles E and E_0 such that $E_0 \in \mathcal{C}'$ the natural pairings*

$$\mathrm{Ext}^i_{\mathcal{C}}(E_0, E) \otimes \mathrm{Ext}^{1-i}_{\mathcal{C}}(E, E_0) \to \mathrm{Ext}^1_{\mathcal{C}}(E_0, E_0) \to \mathbb{C}$$

for $i = 0, 1$ are perfect. Here the functional $\mathrm{Ext}^1_{\mathcal{C}}(E_0, E_0) \to \mathbb{C}$ is induced by the Serre duality on $\mathcal{C}'$. Therefore, we have functorial isomorphisms

$$\mathrm{Ext}^{1-i}_{\mathcal{C}}(E, E_0) \widetilde{\to} \mathrm{Ext}^i_{\mathcal{C}}(E_0, E)^* \tag{2.10}$$

for $E \in \mathcal{C}$, $E_0 \in \mathcal{C}'$.

Proof. If E_0 is standard the assertion follows from Theorem 2.11. It remains to observe that if for fixed $E \in \mathcal{C}$ the map (2.10) is an isomorphism for some $E_0, E_0' \in \mathcal{C}'$ then it is also an isomorphism for any extension of E_0 by E_0'. □

3. Ampleness

3.1. Ample sequences of standard holomorphic bundles. Let us start by recalling some basic notions concerning ample sequences in abelian categories and associated $\mathbb{Z}$-algebras. The reader can consult [3] for more details.

Definition.(see [9],[3]): Let $(E_n)_{n\in\mathbb{Z}}$ be a sequence of objects in a $\mathbb{C}$-linear abelian category $\mathcal{A}$ such that $\mathrm{Hom}(E_n, E)$ is finite-dimensional for every $E \in \mathcal{A}$ and every $n \in \mathbb{Z}$. Then (E_n) is called *ample* if the following two conditions hold:
(i) for every surjection $E \to E'$ in $\mathcal{A}$ the induced map $\mathrm{Hom}(E_n, E) \to \mathrm{Hom}(E_n, E')$ is surjective for all $n << 0$;
(ii) for every object $E \in \mathcal{A}$ and every $N \in \mathbb{Z}$ there exists a surjection $\oplus_{i=1}^{s} E_{n_i} \to E$ where $n_i < N$ for all i.

To a sequence $(E_n)_{n\in\mathbb{Z}}$ one can associate a so called $\mathbb{Z}$-algebra $A = \oplus_{i\le j} A_{ij}$, where $A_{ii} = \mathbb{C}$, $A_{ij} = \mathrm{Hom}(E_i, E_j)$ for $i < j$, the multiplications $A_{jk} \otimes A_{ij} \to A_{ik}$ are induced by the composition in $\mathcal{C}$. One can define for $\mathbb{Z}$-algebras all the standard notions associated with graded algebras (see [3]). In particular, we can talk about right A-modules: these have form $M = \oplus_{i\in\mathbb{Z}} M_i$ and the right A-action is given by the maps $M_j \otimes A_{ij} \to M_i$. The analogues of free A-modules are direct sums of the modules P_n, $n \in \mathbb{Z}$, defined by $(P_n)_i = A_{ni}$. We say that an A-module M is *finitely generated* if there exists a surjection $\oplus_{i=1}^{s} P_{n_i} \to M$. A finitely generated A-module M is called *coherent* if for every morphism $f : P \to M$, where P is a finitely generated free module, the module $\ker(f)$ is finitely generated. Finally, a $\mathbb{Z}$-algebra A is called *coherent* if all the modules P_n are coherent and in addition all one-dimensional A-modules are coherent.
The main theorem of [3] asserts that if (E_n) is ample then the $\mathbb{Z}$-algebra A is coherent and the natural functor $E \mapsto \oplus_{i<0} \mathrm{Hom}(E_i, E)$ gives an equivalence of categories

$$\mathcal{A} \simeq \mathrm{cohproj}\, A \tag{3.1}$$

where $\mathrm{cohproj}\, A$ is the quotient of the category of coherent right A-modules by the subcategory of finite-dimensional modules. We are going to apply this theorem to the category $\mathcal{C}'$ generated by standard holomorphic bundles on $T = T_{\theta,\tau}$. Recall that in [2] we identified this category with a certain abelian subcategory $\mathcal{C}^\theta$ of the derived category $D^b(X)$ of coherent sheaves on the elliptic curve $X = \mathbb{C}/\mathbb{Z} + \mathbb{Z}\tau$. To define $\mathcal{C}^\theta$ one has to consider two full subcategories in the category $\mathrm{Coh}(X)$ of coherent sheaves on X: $\mathrm{Coh}_{<\theta}$ (resp., $\mathrm{Coh}_{>\theta}$) is the minimal subcategory of $\mathrm{Coh}(X)$ closed under extensions and containing all stable bundles of slope $< \theta$ (resp., all stable bundles of slope $> \theta$ and all torsion sheaves). Then by the definition

$$\begin{aligned} \mathcal{C}^\theta = \{K \in D^b(X) \quad : \quad & H^{>0}(K) = 0, H^0(K) \in \mathrm{Coh}_{>\theta}, \\ & H^{-1}(K) \in \mathrm{Coh}_{<\theta}, H^{<-1}(K) = 0\}. \end{aligned}$$

Thus, $\mathcal{C}^\theta$ contains $\mathrm{Coh}_{>\theta}$ and $\mathrm{Coh}_{<\theta}[1]$ and these two subcategories generate $\mathcal{C}^\theta$ in an appropriate sense. The fact that $\mathcal{C}^\theta$ is abelian follows from the torsion theory (see [1]). Note that the vectors $(\deg(K), \mathrm{rk}(K)) \in \mathbb{Z}^2$ for $K \in \mathcal{C}^\theta$ are characterized by the inequality

$$\deg(K) - \theta \, \mathrm{rk}(K) > 0.$$

In [2] we showed that a version of Fourier-Mukai transform gives an equivalence $\mathcal{S} : \mathcal{C}' \widetilde{\to} \mathcal{C}^\theta$ (this $\mathcal{S}$ differs from the transform studied in section 3.3 of [2] by the shift $K \mapsto K[1]$). Standard holomorphic bundles correspond under $\mathcal{S}$ to *stable* objects of $\mathcal{C}^\theta$: the latter are structure sheaves of points and objects of the form $V[n]$ where V is a stable bundle, $n \in \{0, 1\}$. Moreover, one has

$$\deg \mathcal{S}(E_{d,c}(\theta)) = d, \ \mathrm{rk}\, \mathcal{S}(E_{d,c}(\theta)) = -c.$$

It follows that

$$\mathrm{rk}(\mathcal{S}(E)) = -\deg(E), \ \mu(\mathcal{S}(E)) = \theta - \mu(E)^{-1}.$$

The following criterion of ampleness in $\mathcal{C}'$ is essentially contained in the proof of Theorem 3.5 of [4], where we showed the existence of ample sequences in $\mathcal{C}'$.

Theorem 3.1. *Let (E_n) be a sequence of standard holomorphic bundles on T such that $\mu(E_n) \to -\infty$ as $n \to -\infty$ and $\mathrm{rk}(E_n) > c$ for all $n << 0$ for some constant $c > 0$. Then (E_n) is an ample sequence in $\mathcal{C}'$. Moreover, for every $E \in \mathcal{C}'$ the natural morphism $\mathrm{Hom}(E_n, E) \otimes_{\mathbb{C}} E_n \to E$ in $\mathcal{C}$ is surjective for $n << 0$.*

Proof. Let $\mathcal{F}_n = \mathcal{S}(E_n)$ be the corresponding sequence of stable objects of $\mathcal{C}^\theta$. Then $\mathrm{rk}(\mathcal{F}_n) = -\deg(E_n) \to +\infty$ and $\mu(\mathcal{F}_n) = \theta - \mu(E_n)^{-1} \to \theta$ as $n \to -\infty$. Moreover, we have

$$\mu(\mathcal{F}_n) - \theta = \frac{\mathrm{rk}(E_n)}{\mathrm{rk}(\mathcal{F}_n)} > \frac{c}{\mathrm{rk}(\mathcal{F}_n)}.$$

Therefore, the same proof as in Theorem 3.5 of [4] (where we considered only the special case $c = 1$) shows that the sequence $(\mathcal{F}_n)$ is ample in $\mathcal{C}^\theta$ and that for every $\mathcal{F} \in \mathcal{C}^\theta$ the morphism $\mathrm{Hom}(\mathcal{F}_n, \mathcal{F}) \otimes_{\mathbb{C}} \mathcal{F}_n \to \mathcal{F}$ is surjective for $n << 0$. Hence, the same assertions hold for the sequence (E_n) in $\mathcal{C}'$. □

Theorem 3.2. *Let (E_n) be a sequence as in Theorem 3.1 and let $A = \oplus_{i \le j} \mathrm{Hom}_{\mathcal{C}}(E_i, E_j)$ be the corresponding $\mathbb{Z}$-algebra. Then for every holomorphic bundle E on T the A-module $M(E) = \oplus_{i<0} \mathrm{Hom}_{\mathcal{C}}(E_i, E)$ is coherent. Also, for every sufficiently small i_0 the canonical morphism of A-modules*

$$\mathrm{Hom}_{\mathcal{C}}(E_{i_0}, E) \otimes P_{i_0} \to M(E)$$

has finite-dimensional cokernel.

First, we need a criterion for finite generation of modules $M(E)$ with a weaker assumption on (E_n).

Lemma 3.3. *Let E be a holomorphic bundle on T and let $C = C(E)$ be the corresponding constant from Theorem 2.10. Let (E_i) be a sequence of standard holomorphic bundles such that $\mu(E_i) \to \mu$ as $i \to -\infty$, where $\mu \in \mathbb{R} \cup \{-\infty\}$. Assume that for some $\epsilon > 0$ one has $\mathrm{rk}(E_i)^2(\mu(E_i) - \mu) > 1 + \epsilon$ for all $i << 0$ (this condition is vacuous if $\mu = -\infty$). Then*
(i) for every sufficiently small $i_0 \in \mathbb{Z}$ there exists $i_1 = i_1(i_0)$ such that for all $i < i_1$ there exists a standard holomorphic bundle F_i fitting into the following short exact sequence in $\mathcal{C}'$:

$$0 \to E_i \overset{\mathrm{can}}{\to} \mathrm{Hom}_{\mathcal{C}}(E_i, E_{i_0})^* \otimes E_{i_0} \to F_i \to 0, \tag{3.2}$$

where can *is the canonical morphism.*
(ii) Assume in addition that $\mu < C$ and if $\mu = -\infty$ then for some $c > 0$ one has $\mathrm{rk}(E_i) > c$ for all $i << 0$. Then for all sufficiently small i_0 there exists i_1 such that for $i < i_1$ one has $\mathrm{Ext}^1(F_i, E) = 0$, where F_i is defined by (3.2). Under the same assumptions the A-module $M(E)$ is finitely generated.

Proof. (i) Let us denote $r_i = \mathrm{rk}(E_i)$, $\mu_i = \mu(E_i)$. If μ is finite then for every sufficiently small i_0 one has $r_{i_0}^2(\mu_{i_0} - \mu) > 1$. Therefore, we can find $i_1 < i_0$ such that for $i < i_1$ one has $r_{i_0}^2(\mu_{i_0} - \mu_i) > 1$. If $\mu = -\infty$ then we can take any i_0 and then still find $i_1 < i_0$ such that for $i < i_1$ the above inequality holds. We are going to construct F_i in this situation. Using the equivalence $\mathcal{S} : \mathcal{C}' \widetilde{\to} \mathcal{C}^\theta \subset D^b(X)$ we can first define $\widetilde{F_i} \in D^b(X)$ from the exact triangle

$$\mathcal{S}(E_i) \to \mathrm{Hom}_{\mathcal{C}}(E_i, E_{i_0})^* \otimes \mathcal{S}(E_{i_0}) \to \widetilde{F_i} \to \mathcal{S}(E_i)[1].$$

In other words, $\widetilde{F_i}$ is the image of E_i under the equivalence $R_{E_{i_0}} : D^b(X) \to D^b(X)$ given by the right twist with respect to E_{i_0} (see [4], sec.2.3, or [6]; our functor differs from that of [6] by a shift). It follows that $\mathrm{Hom}_{D^b(E)}(\widetilde{F_i}, \widetilde{F_i}) \simeq \mathrm{Hom}_{\mathcal{C}}(E_i, E_i) \simeq \mathbb{C}$, so $\widetilde{F_i}$ is a stable object and either $\widetilde{F_i} \in \mathcal{C}^\theta$ or $\widetilde{F_i} \in \mathcal{C}^\theta[1]$. To prove that $\widetilde{F_i} \in \mathcal{C}^\theta$ it suffices to check that $\deg(\widetilde{F_i}) - \theta\,\mathrm{rk}(\widetilde{F_i}) > 0$. But

$$\deg(\widetilde{F_i}) - \theta\,\mathrm{rk}(\widetilde{F_i}) = \chi(E_i, E_{i_0}) r_{i_0} - r_i = ((\mu_{i_0} - \mu_i) r_{i_0}^2 - 1) r_i > 0$$

by our choice of i. Hence, we have $\widetilde{F_i} \in \mathcal{C}^\theta$ and we can set $F_i = \mathcal{S}^{-1}(\widetilde{F_i})$.
(ii) If μ is finite then we can choose i_0 such that $r_{i_0}^2(\mu_{i_0} - \mu) > 1 + \epsilon$ and $\mu_{i_0} + \epsilon^{-1}(\mu_{i_0} - \mu) < C$. If $\mu = -\infty$ then we choose i_0 such that $\mu_{i_0} + 2r_{i_0}^{-2} < C$ (here we use the assumption that $r_i > c$ for $i << 0$). In either case for sufficiently small i we have short exact sequence (3.2). Applying to it the functor $\mathrm{Hom}(?, E)$ we get the long exact sequence

$$\begin{array}{rcl} 0 \to \mathrm{Hom}_{\mathcal{C}}(F_i, E) & \to & \mathrm{Hom}_{\mathcal{C}}(E_{i_0}, E) \otimes \mathrm{Hom}_{\mathcal{C}}(E_i, E_{i_0}) \\ & \to & \mathrm{Hom}_{\mathcal{C}}(E_i, E) \to \mathrm{Ext}^1_{\mathcal{C}}(F_i, E). \end{array}$$

Thus, vanishing of $\mathrm{Ext}^1_{\mathcal{C}}(F_i, E)$ for all $i << 0$ would imply that the A-module $M(E)$ is finitely generated. By the definition of the constant C this vanishing would follow from the inequality $\mu(F_i) < C$. In the case $\mu \neq -\infty$ we have for $i << 0$

$$\mu(F_i) = \mu_{i_0} + \frac{\mu_{i_0} - \mu_i}{r_{i_0}^2(\mu_{i_0} - \mu_i) - 1} < \mu_{i_0} + \epsilon^{-1}(\mu_{i_0} - \mu_i),$$

so the required inequality follows for $i << 0$ from our choice of i_0. In the case $\mu = -\infty$ we can finish the proof similarly using the inequality

$$\mu(F_i) < \mu_{i_0} + \frac{2}{r_{i_0}^2}$$

that holds for all $i << 0$. □

Proof of Theorem 3.2. By Lemma 3.3 for any sufficiently small i_0 there exists i_1 (depending on i_0) such that we have short exact sequence (3.2) and the induced sequence of A-modules

$$0 \to \oplus_{i<i_1} \mathrm{Hom}_{\mathcal{C}}(F_i, E) \to \mathrm{Hom}_{\mathcal{C}}(E_{i_0}, E) \otimes (P_{i_0})_{<i_1} \to M(E)_{<i_1} \to 0$$

is exact, where for every A-module $M = \oplus M_i$ we set $M_{<n} = \oplus_{i<n} M_i$. This immediately implies the last assertion of the theorem. Note that the structure of the A-module on $\oplus \mathrm{Hom}_{\mathcal{C}}(F_i, E)$ is defined using the natural isomorphisms $\mathrm{Hom}_{\mathcal{C}}(F_i, F_j) \simeq \mathrm{Hom}_{\mathcal{C}}(E_i, E_j)$ coming from the equality $F_i = R_{E_{i_0}}(E_i)$, where $R_{E_{i_0}}$ is the right twist with respect to E_{i_0}. It suffices to prove that the module $M'(E) := \oplus_{i<i_1} \mathrm{Hom}_{\mathcal{C}}(F_i, E)$ is finitely generated. Indeed, this would imply that the module $M(E)_{<i_1}$ is finitely presented and hence coherent (since A is coherent), therefore, the module $M(E)$ is also coherent. To check that $M'(E)$ is finitely generated we will use the criterion of Lemma 3.3 for the sequence (F_i). We have

$$\mu(F_i) \to \mu = \mu_{i_0} + 1/r_{i_0}^2$$

as $i \to -\infty$. Also,

$$\mathrm{rk}(F_i)^2(\mu(F_i) - \mu) = \frac{((\mu_{i_0} - \mu_i)r_{i_0}^2 - 1)r_i^2}{r_{i_0}^2} > \frac{((\mu_{i_0} - \mu_i)r_{i_0}^2 - 1)c^2}{r_{i_0}^2} \to +\infty$$

as $i \to -\infty$. Hence, the conditions of Lemma 3.3 will be satisfied once we show that $\mu = \mu_{i_0} + 1/r_{i_0}^2$ can be made smaller than any given constant by an appropriate choice of i_0. But this is of course true since $\mu_{i_0} + 1/r_{i_0}^2 < \mu_{i_0} + 1/c^2$ and $\mu_{i_0} \to -\infty$ as $i_0 \to -\infty$. □

3.2. Proof of Theorem 1.1. Let us pick a sequence $(E_n)_{n\in\mathbb{Z}}$ of stable holomorphic bundles satisfying conditions of Theorem 3.1 (it is easy to see that such a sequence exists, see the proof of Theorem 3.5 in [4]). Let E be a holomorphic bundle on T. Then by Theorem 3.2 the module $M = \oplus_i \mathrm{Hom}_{\mathcal{C}}(E_i, E)$ is coherent, hence, we can consider the object $E' \in \mathcal{C}'$ corresponding to this module via the equivalence (3.1). By the definition this means that there is an isomorphism of A-modules

$$M(E')_{<i_0} \simeq M(E)_{<i_0} \tag{3.3}$$

for some i_0. By Theorem 3.2 assuming that i_0 is small enough we can ensure that for all $i < i_0$ the canonical morphism $M(E')_i \otimes P_i \to M(E')$ has finite-dimensional cokernel. We claim that there exists a morphism $f : E' \to E$ in $\mathcal{C}$ that induces the same isomorphism of A-modules $M(E')_{<i_1} \simeq M(E)_{<i_1}$ for some $i_1 < i_0$ as the isomorphism (3.3). Indeed, by Theorem 3.1 we can find a resolution for E' in $\mathcal{C}'$ of the form

$$\ldots \to V_1 \otimes E_{n_1} \to V_0 \otimes E_{n_0} \to E' \to 0 \tag{3.4}$$

where $V_0 = \mathrm{Hom}(E_{n_0}, E')$ and $n_j < i_0$ for all $j \geq 0$. Using this resolution we can compute $\mathrm{Hom}_{\mathcal{C}}(E', E)$:

$$\mathrm{Hom}_{\mathcal{C}}(E', E) \simeq \ker(V_0^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_0}, E) \to V_1^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_1}, E)).$$

Using isomorphism (3.3) we can identify this space with

$$\ker(V_0^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_0}, E') \to V_1^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_1}, E')) \simeq \mathrm{Hom}_{\mathcal{C}}(E', E').$$

Thus, we obtain an isomorphism $\mathrm{Hom}_{\mathcal{C}}(E', E) \simeq \mathrm{Hom}_{\mathcal{C}}(E', E')$. We define $f \in \mathrm{Hom}_{\mathcal{C}}(E', E)$ to be the element corresponding to the identity in $\mathrm{Hom}_{\mathcal{C}}(E', E')$. Let us check that f induces the same isomorphism as (3.3) on some truncations of the modules $M(E')$ and $M(E)$. The definition of f implies that the composition of the induced morphism $f_* : M(E') \to M(E)$ with the natural morphism $V_0 \otimes P_{n_0} = M(V_0 \otimes E_{n_0}) \to M(E')$ coincides with the morphism $V_0 \otimes P_{n_0} \to M(E)$ induced by the isomorphism $V_0 = \mathrm{Hom}(E_{n_0}, E') \simeq \mathrm{Hom}(E_{n_0}, E)$ induced by (3.3). Therefore, our claim follows from the fact that the above morphism $V_0 \otimes P_{n_0} \to M(E')$ induces a surjective morphism on appropriate truncations.

Thus, we can assume from the beginning that the isomorphism (3.3) is induced by a morphism $f : E' \to E$. Next, we are going to construct a morphism $g : E' \to E$ such that $g \circ f = \mathrm{id}_{E'}$. To do this we note that by Serre duality $\mathrm{Hom}_{\mathcal{C}}(E, E') \simeq \mathrm{Ext}^1_{\mathcal{C}}(E', E)^*$ (see Corollary 2.12). Let us make n_0 smaller if needed so that $\mathrm{Ext}^1_{\mathcal{C}}(E_{n_0}, E) = \mathrm{Ext}^1_{\mathcal{C}}(E_{n_0}, E') = 0$. Then the space $\mathrm{Ext}^1_{\mathcal{C}}(E', E)$ can be computed using resolution (3.4):

$$\begin{aligned} \mathrm{Ext}^1_{\mathcal{C}}(E', E) \quad \simeq \quad & H^1[V_0^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_0}, E) \\ & \to V_1^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_1}, E) \to V_2^* \otimes \mathrm{Hom}_{\mathcal{C}}(E_{n_2}, E)]. \end{aligned} \tag{3.5}$$

Indeed, let us define $K_1 \in \mathcal{C}'$ from the short exact sequence

$$0 \to K_1 \to V_0 \otimes E_{n_0} \to E' \to 0,$$

so that we have the following resolution for K_1:

$$\ldots \to V_2 \otimes E_{n_2} \to V_1 \otimes E_{n_1} \to K_1 \to 0.$$

Then the isomorphism (3.5) can be derived from the induced exact sequences

$$\mathrm{Hom}_{\mathcal{C}}(V_0 \otimes, E_{n_0}, E) \to \mathrm{Hom}_{\mathcal{C}}(K_1, E) \to \mathrm{Ext}^1_{\mathcal{C}}(E', E) \to \mathrm{Ext}^1_{\mathcal{C}}(V_0 \otimes E_{n_0}, E) = 0,$$

$$0 \to \mathrm{Hom}_{\mathcal{C}}(K_1, E) \to \mathrm{Hom}_{\mathcal{C}}(V_1 \otimes E_{n_1}, E) \to \mathrm{Hom}_{\mathcal{C}}(V_2 \otimes E_{n_2}, E).$$

Using the fact that isomorphism (3.5) is functorial in E such that $\mathrm{Ext}^1_{\mathcal{C}}(E_{n_0}, E) = 0$ we derive that the morphism $\mathrm{Ext}^1_{\mathcal{C}}(E', E') \to \mathrm{Ext}^1_{\mathcal{C}}(E', E)$

induced by f is an isomorphism. But there is a natural functional $\phi' \in \mathrm{Ext}^1_{\mathcal{C}}(E', E')^*$ given by Serre duality. Let $\phi \in \mathrm{Ext}^1_{\mathcal{C}}(E', E)^*$ be the corresponding functional. The isomorphism $\mathrm{Ext}^1_{\mathcal{C}}(E', E)^* \widetilde{\rightarrow} \mathrm{Hom}_{\mathcal{C}}(E, E')$ maps ϕ to some element $g \in \mathrm{Hom}_{\mathcal{C}}(E, E')$. By functoriality of the Serre duality the following diagram is commutative:

$$\begin{tikzcd}
\mathrm{Hom}_{\mathcal{C}}(E, E') \arrow[r, "\alpha"] \arrow[d, "f^*"'] & \mathrm{Ext}^1_{\mathcal{C}}(E', E)^* \arrow[d, "f^*"'] \\
\mathrm{Hom}_{\mathcal{C}}(E', E') \arrow[r, "\alpha'"] & \mathrm{Ext}^1_{\mathcal{C}}(E', E')^*
\end{tikzcd}$$

where the vertical arrows are induced by f. Since $\phi' = \alpha'(\mathrm{id}_{E'})$, $f^*(\phi) = \phi'$ and $\alpha(g) = \phi$ we deduce that $f^*(g) = \mathrm{id}_{E'}$, i.e. $g \circ f = \mathrm{id}_{E'}$. Therefore, we have $E \simeq E' \oplus E''$ for some holomorphic bundle E'' such that $\mathrm{Hom}_{\mathcal{C}}(E_i, E'') = 0$ for $i < i_0$. But Theorem 2.10 implies that $\mathrm{Ext}^1_{\mathcal{C}}(E_i, E'') = 0$ for all sufficiently negative i. Together with Corollary 2.9 this implies that $E'' = 0$. □

References

[1] D. Happel, I. Reiten, S. O. Smalo, *Tilting in abelian categories and quasitilted algebras*, Memoirs AMS 575, 1996.

[2] A. Polishchuk, A. Schwarz, *Categories of holomorphic bundles on noncommutative two-tori*, Comm. Math. Phys. 236 (2003), 135–159.

[3] A. Polishchuk, *Noncommutative* Proj *and coherent algebras*, preprint math.AG/0212182.

[4] A. Polishchuk, *Noncommutative two-tori with real multiplication as noncommutative projective varieties*, Journal of Geometry and Physics 50 (2004), 162–187.

[5] M. A. Rieffel, The cancellation theorem for projective modules over irrational rotation $\mathbb{C}^*$-algebras, Proc. London Math. Soc. (3) 47 (1983), 285–302.

[6] P. Seidel, R. Thomas, *Braid group actions on derived categories of coherent sheaves*, Duke Math. J. 108 (2001), 37–108.

[7] M. Spera, *Sobolev theory for noncommutative tori*, Rend. Sem. Mat. Univ. Padova 86 (1991), 143–156.

[8] M. Spera, *A symplectic approach to Yang-Mills theory for noncommutative tori*, Canad. J. Math. 44 (1992), 368–387.

[9] J. T. Stafford, M. Van den Berg, *Noncommutative curves and noncommutative surfaces*, Bull. AMS 38 (2001), 171–216.

Alexander Polishchuk
Department of Mathematics
University of Oregon
Eugene, OR 97403
apolishuoregon.edu

Multipliers of Improper Similitudes

R. Preeti[1] and J.-P. Tignol[2]

Received: February 11, 2004
Revised: May 5, 2004

Communicated by Alexander Merkurev

Abstract. For a central simple algebra with an orthogonal involution (A, σ) over a field k of characteristic different from 2, we relate the multipliers of similitudes of (A, σ) with the Clifford algebra $C(A, \sigma)$. We also give a complete description of the group of multipliers of similitudes when $\deg A \le 6$ or when the virtual cohomological dimension of k is at most 2.

2000 Mathematics Subject Classification: 11E72.
Keywords and Phrases: Central simple algebra with involution, hermitian form, Clifford algebra, similitude.

Introduction

A. Weil has shown in [22] how to obtain all the simple linear algebraic groups of adjoint type D_n over an arbitrary field k of characteristic different from 2: every such group is the connected component of the identity in the group of automorphisms of a pair (A, σ) where A is a central simple k-algebra of degree $2n$ and $\sigma\colon A \to A$ is an involution of orthogonal type, i.e., a linear map which over a splitting field of A is the adjoint involution of a symmetric bilinear form. (See [7] for background material on involutions on central simple algebras and classical groups.) Every automorphism of (A, σ) is inner, and induced by an element $g \in A^\times$ which satisfies $\sigma(g)g \in k^\times$. The group of *similitudes* of (A, σ) is defined by that condition,

$$\mathrm{GO}(A, \sigma) = \{g \in A^\times \mid \sigma(g)g \in k^\times\}.$$

[1]The first author gratefully acknowledges the generous support of the Université catholique de Louvain, Belgium and the ETH-Z, Switzerland.
[2]Work supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00287, KTAGS. The second author is supported in part by the National Fund for Scientific Research (Belgium).

The map which carries $g \in \mathrm{GO}(A,\sigma)$ to $\sigma(g)g \in k^\times$ is a homomorphism

$$\mu\colon \mathrm{GO}(A,\sigma) \to k^\times$$

called the *multiplier map.* Taking the reduced norm of each side of the equation $\sigma(g)g = \mu(g)$, we obtain

$$\mathrm{Nrd}_A(g)^2 = \mu(g)^{2n},$$

hence $\mathrm{Nrd}_A(g) = \pm\mu(g)^n$. The similitude g is called *proper* if $\mathrm{Nrd}_A(g) = \mu(g)^n$, and *improper* if $\mathrm{Nrd}_A(g) = -\mu(g)^n$. The proper similitudes form a subgroup $\mathrm{GO}_+(A,\sigma) \subset \mathrm{GO}(A,\sigma)$. (As an algebraic group, $\mathrm{GO}_+(A,\sigma)$ is the connected component of the identity in $\mathrm{GO}(A,\sigma)$.)
Our purpose in this work is to study the multipliers of similitudes of a central simple k-algebra with orthogonal involution (A,σ). We denote by $G(A,\sigma)$ (resp. $G_+(A,\sigma)$, resp. $G_-(A,\sigma)$) the group of multipliers of similitudes of (A,σ) (resp. the group of multipliers of proper similitudes, resp. the coset of multipliers of improper similitudes),

$$\begin{aligned} G(A,\sigma) &= \{\mu(g) \mid g \in \mathrm{GO}(A,\sigma)\},\\ G_+(A,\sigma) &= \{\mu(g) \mid g \in \mathrm{GO}_+(A,\sigma)\},\\ G_-(A,\sigma) &= \{\mu(g) \mid g \in \mathrm{GO}(A,\sigma) \setminus \mathrm{GO}_+(A,\sigma)\}. \end{aligned}$$

When A is split ($A = \mathrm{End}_k V$ for some k-vector space V), hyperplane reflections are improper similitudes with multiplier 1, hence

$$G(A,\sigma) = G_+(A,\sigma) = G_-(A,\sigma).$$

When A is not split however, we may have $G(A,\sigma) \neq G_+(A,\sigma)$.
Multipliers of similitudes were investigated in relation with the discriminant $\operatorname{disc}\sigma$ by Merkurjev–Tignol [14]. Our goal is to obtain similar results relating multipliers of similitudes to the next invariant of σ, which is the Clifford algebra $C(A,\sigma)$ (see [7, §8]). As an application, we obtain a complete description of $G(A,\sigma)$ when $\deg A \le 6$ or when the virtual cohomological dimension of k is at most 2.

To give a more precise description of our results, we introduce some more notation. Throughout the paper, k denotes a field of characteristic different from 2. For any integers n, $d \ge 1$, let μ_{2^n} be the group of 2^n-th roots of unity in a separable closure of k and let $H^d(k, \mu_{2^n}^{\otimes(d-1)})$ be the d-th cohomology group of the absolute Galois group with coefficients in $\mu_{2^n}^{\otimes(d-1)}$ ($= \mathbf{Z}/2^n\mathbf{Z}$ if $d = 1$). Denote simply

$$H^d k = \varinjlim_n H^d(k, \mu_{2^n}^{\otimes(d-1)}),$$

so H^1k and H^2k may be identified with the 2-primary part of the character group of the absolute Galois group and with the 2-primary part of the Brauer group of k, respectively,

$$H^1k = X_2(k), \qquad H^2k = \mathrm{Br}_2(k).$$

In particular, the isomorphism $k^\times/k^{\times 2} \simeq H^1(k, \mathbf{Z}/2\mathbf{Z})$ derived from the Kummer sequence (see for instance [7, (30.1)]) yields a canonical embedding

$$k^\times/k^{\times 2} \hookrightarrow H^1k. \tag{1}$$

The Brauer class (or the corresponding element in H^2k) of a central simple k-algebra E of 2-primary exponent is denoted by $[E]$.
If K/k is a finite separable field extension, we denote by $N_{K/k}\colon H^dK \to H^dk$ the norm (or corestriction) map. We extend the notation above to the case where $K \simeq k \times k$ by letting $H^d(k\times k) = H^dk \times H^dk$ and

$$N_{(k\times k)/k}(\xi_1, \xi_2) = \xi_1 + \xi_2 \qquad \text{for } (\xi_1, \xi_2) \in H^d(k\times k).$$

Our results use the product

$$\cdot\colon\ k^\times \times H^dk \to H^{d+1}k \qquad \text{for } d = 1 \text{ or } 2$$

induced as follows by the cup-product: for $x \in k^\times$ and $\xi \in H^dk$, choose n such that $\xi \in H^d(k, \mu_{2^n}^{\otimes(d-1)})$ and consider the cohomology class $(x)_n \in H^1(k, \mu_{2^n})$ corresponding to the 2^n-th power class of x under the isomorphism $H^1(k,\mu_{2^n}) = k^\times/k^{\times 2^n}$ induced by the Kummer sequence; let then

$$x\cdot\xi = (x)_n \cup \xi \in H^{d+1}(k, \mu_{2^n}^{\otimes d}) \subset H^{d+1}k.$$

In particular, if $d = 1$ and ξ is the square class of $y \in k^\times$ under the embedding (1), then $x\cdot\xi$ is the Brauer class of the quaternion algebra $(x, y)_k$.
Throughout the paper, we denote by A a central simple k-algebra of even degree $2n$, and by σ an orthogonal involution of A. Recall from [7, (7.2)] that $\operatorname{disc}\sigma \in k^\times/k^{\times 2} \subset H^1k$ is the square class of $(-1)^n \operatorname{Nrd}_A(a)$ where $a \in A^\times$ is an arbitrary skew-symmetric element. Let Z be the center of the Clifford algebra $C(A,\sigma)$; thus, Z is a quadratic étale k-algebra, $Z = k[\sqrt{\operatorname{disc}\sigma}]$, see [7, (8.10)]. The following relation between similitudes and the discriminant is proved in [14, Theorem A] (see also [7, (13.38)]):

Theorem 1. *Let (A,σ) be a central simple k-algebra with orthogonal involution of even degree. For $\lambda \in G(A,\sigma)$,*

$$\lambda\cdot\operatorname{disc}\sigma = \begin{cases} 0 & \text{if } \lambda \in G_+(A,\sigma), \\ [A] & \text{if } \lambda \in G_-(A,\sigma). \end{cases}$$

For $d = 2$ (resp. 3), let $(H^dk)/A$ be the factor group of H^dk by the subgroup $\{0, [A]\}$ (resp. by the subgroup $k^\times\cdot[A]$). Theorem 1 thus shows that for $\lambda \in G(A,\sigma)$

$$\lambda\cdot\operatorname{disc}\sigma = 0 \quad \text{in } (H^2k)/A.$$

Our main results are Theorems 2, 3, 4, and 5 below.

Theorem 2. *Suppose A is split by Z. There exists an element $\gamma(\sigma) \in H^2k$ such that $\gamma(\sigma)_Z = [C(A,\sigma)]$ in H^2Z. For $\lambda \in G(A,\sigma)$,*

$$\lambda \cdot \gamma(\sigma) = 0 \quad \text{in } (H^3k)/A.$$

Remark 1. In the conditions of the theorem, the element $\gamma(\sigma) \in H^2k$ is not uniquely determined if $Z \not\simeq k \times k$. Nevertheless, if $\lambda \cdot \operatorname{disc}\sigma = 0$ in $(H^2k)/A$, then $\lambda \cdot \gamma(\sigma) \in (H^3k)/A$ is uniquely determined. Indeed, if γ, $\gamma' \in H^2k$ are such that $\gamma_Z = \gamma'_Z$, then there exists $u \in k^\times$ such that $\gamma' = \gamma + u \cdot \operatorname{disc}\sigma$, hence

$$\lambda \cdot \gamma' = \lambda \cdot \gamma + \lambda \cdot u \cdot \operatorname{disc}\sigma.$$

The last term vanishes in $(H^3k)/A$ since $\lambda \cdot \operatorname{disc}\sigma = 0$ in $(H^2k)/A$.

The proof of Theorem 2 is given in Section 1. It shows that in the split case, where $A = \operatorname{End}_k V$ and σ is adjoint to some quadratic form q on V, we may take for $\gamma(\sigma)$ the Brauer class of the full Clifford algebra $C(V,q)$. Note that the statement of Theorem 2 does not discriminate between multipliers of proper and improper similitudes, but Theorem 1 may be used to distinguish between them. Slight variations of the arguments in the proof of Theorem 2 also yield the following result on multipliers of *proper* similitudes:

Theorem 3. *Suppose the Schur index of A is at most 4. If $\lambda \in G_+(A,\sigma)$, then there exists $z \in Z^\times$ such that $\lambda = N_{Z/k}(z)$ and*

$$N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) = 0 \quad \text{in } (H^3k)/A.$$

The proof is given in Section 1. Note however that the theorem holds without the hypothesis that $\operatorname{ind} A \leq 4$, as follows from Corollaries 1.20 and 1.21 in [12]. Using the Rost invariant of Spin groups, these corollaries actually yield an explicit element z as in Theorem 3 from any proper similitude with multiplier λ.

Remark 2. The element $N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) \in (H^3k)/A$ depends only on $N_{Z/k}(z)$ and not on the specific choice of $z \in Z$. Indeed, if z, $z' \in Z^\times$ are such that $N_{Z/k}(z) = N_{Z/k}(z')$, then Hilbert's Theorem 90 yields an element $u \in Z^\times$ such that, denoting by ι the nontrivial automorphism of Z/k,

$$z' = zu\iota(u)^{-1},$$

hence

$$N_{Z/k}\big(z' \cdot [C(A,\sigma)]\big) = \\ N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) + N_{Z/k}\big(u \cdot [C(A,\sigma)]\big) - N_{Z/k}\big(\iota(u) \cdot [C(A,\sigma)]\big).$$

Since $N_{Z/k} \circ \iota = N_{Z/k}$ and since the properties of the Clifford algebra (see [7, (9.12)]) yield

$$[C(A,\sigma)] - \iota[C(A,\sigma)] = [A]_Z,$$

it follows that

$$N_{Z/k}\big(u\cdot[C(A,\sigma)]\big)-N_{Z/k}\big(\iota(u)\cdot[C(A,\sigma)]\big)=N_{Z/k}\big(u\cdot[A]_Z\big).$$

By the projection formula, the right side is equal to $N_{Z/k}(u)\cdot[A]$. The claim follows.

Remark 3. Theorems 2 and 3 coincide when they both apply, i.e., if A is split by Z (hence $\operatorname{ind} A=1$ or 2), and $\lambda\in G_+(A,\sigma)$. Indeed, if $\lambda=N_{Z/k}(z)$ and $\gamma(\sigma)_Z=[C(A,\sigma)]$ then the projection formula yields

$$N_{Z/k}\big(z\cdot[C(A,\sigma)]\big)=\lambda\cdot\gamma(\sigma).$$

Remarkably, the conditions in Theorems 1 and 2 turn out to be sufficient for λ to be the multiplier of a similitude when $\deg A\leq 6$ or when the virtual cohomological 2-dimension[3] of k is at most 2.

THEOREM 4. *Suppose $n\leq 3$, i.e., $\deg A\leq 6$.*

- *If A is not split by Z , then every similitude is proper,*

  $$G(A,\sigma)=G_+(A,\sigma),\qquad G_-(A,\sigma)=\varnothing.$$

  Moreover, for $\lambda\in k^\times$, we have $\lambda\in G(A,\sigma)$ if and only if there exists $z\in Z^\times$ such that $\lambda=N_{Z/k}(z)$ and

  $$N_{Z/k}\big(z\cdot[C(A,\sigma)]\big)=0\quad\text{in } (H^3k)/A.$$

- *If A is split by Z, let $\gamma(\sigma)\in H^2k$ be as in Theorem 2. For $\lambda\in k^\times$, we have $\lambda\in G(A,\sigma)$ if and only if*

  $$\lambda\cdot\operatorname{disc}\sigma=0\ \text{in } (H^2k)/A\qquad\text{and}\qquad\lambda\cdot\gamma(\sigma)=0\ \text{in } (H^3k)/A.$$

The proof is given in Section 2.

Note that if $\deg A=2$, then A is necessarily split by Z and we may choose $\gamma(\sigma)=0$, hence Theorem 4 simplifies to

$$\lambda\in G(A,\sigma)\quad\text{if and only if}\quad\lambda\cdot\operatorname{disc}\sigma=0\text{ in }(H^2k)/A,$$

a statement which is easily proved directly. (See [14, p. 15] or [7, (12.25)].)

If $\deg A=4$, multipliers of similitudes can also be described up to squares as reduced norms from a central simple algebra E of degree 4 such that $[E]=\gamma(\sigma)$ if A is split by Z (see Corollary 4.5) or as norms of reduced norms of $C(A,\sigma)$ if A is not split by Z (see Corollary 2.1).

[3] The authors are grateful to Parimala for her suggestion to investigate the case of low cohomological dimension.

For the next statement, recall that the virtual cohomological 2-dimension of k (denoted $\operatorname{vcd}_2 k$) is the cohomological 2-dimension of $k(\sqrt{-1})$. If v is an ordering of k, we let k_v be a real closure of k for v and denote simply by $(A,\sigma)_v$ the algebra with involution $(A \otimes_k k_v, \sigma \otimes \mathrm{Id}_{k_v})$.

Theorem 5. *Suppose* $\operatorname{vcd}_2 k \leq 2$, *and* A *is split by* Z. *For* $\lambda \in k^\times$, *we have* $\lambda \in G(A,\sigma)$ *if and only if*

$$\lambda > 0 \quad \textit{at every ordering } v \textit{ of } k \textit{ such that } (A,\sigma)_v \textit{ is not hyperbolic},$$
$$\lambda \cdot \operatorname{disc}\sigma = 0 \textit{ in } (H^2k)/A \qquad \textit{and} \qquad \lambda \cdot \gamma(\sigma) = 0 \textit{ in } (H^3k)/A.$$

The proof is given in Section 3.

1 Proofs of Theorems 2 and 3

Theorems 2 and 3 are proved by reduction to the split case, which we consider first. We thus assume $A = \operatorname{End}_k V$ for some k-vector space V of dimension $2n$, and σ is adjoint to a quadratic form q on V. Then $\operatorname{disc}\sigma = \operatorname{disc} q$ and $C(A,\sigma)$ is the even Clifford algebra $C(A,\sigma) = C_0(V,q)$. We denote by $C(V,q)$ the full Clifford algebra of q, which is a central simple k-algebra, and by $I^m k$ the m-th power of the fundamental ideal Ik of the Witt ring Wk.

Lemma 1.1. *For* $\lambda \in k^\times$, *the following conditions are equivalent:*

(a) $\lambda \cdot \operatorname{disc} q = 0$ *in* H^2k *and* $\lambda \cdot [C(V,q)] = 0$ *in* H^3k*;*

(b) $\langle\lambda\rangle \cdot q \equiv q \bmod I^4k$.

Proof. For $\alpha_1, \ldots, \alpha_m \in k^\times$, let

$$\langle\langle \alpha_1, \ldots, \alpha_m \rangle\rangle = \langle 1, -\alpha_1\rangle \otimes \cdots \otimes \langle 1, -\alpha_m\rangle.$$

Let $e_2\colon I^2k \to H^2k$ be the Witt invariant and $e_3\colon I^3k \to H^3k$ be the Arason invariant. By a theorem of Merkurjev [9] (resp. of Merkurjev–Suslin [13] and Rost [17]), we have $\ker e_2 = I^3k$ and $\ker e_3 = I^4k$. Therefore, the lemma follows if we prove

$$\lambda \cdot \operatorname{disc} q = 0 \quad \text{if and only if} \quad \langle\langle\lambda\rangle\rangle \cdot q \in I^3k, \tag{2}$$

and that, assuming that condition holds,

$$e_3(\langle\langle\lambda\rangle\rangle \cdot q) = \lambda \cdot [C(V,q)]. \tag{3}$$

Let $\delta \in k^\times$ be such that $\operatorname{disc} q = (\delta)_1 \in H^1(k, \mathbf{Z}/2\mathbf{Z}) \subset H^1k$. Then

$$q \equiv \langle\langle\delta\rangle\rangle \bmod I^2k, \tag{4}$$

hence

$$e_2(\langle\langle\lambda\rangle\rangle \cdot q) = e_2(\langle\langle\lambda,\delta\rangle\rangle) = \lambda \cdot \operatorname{disc} q,$$

proving (2). Now, assuming $\lambda \cdot \operatorname{disc} q = 0$, we have $\langle\langle \lambda, \delta \rangle\rangle = 0$ in Wk, hence

$$\langle\langle \lambda \rangle\rangle \cdot q = \langle\langle \lambda \rangle\rangle \cdot (q \perp \langle\langle \delta \rangle\rangle).$$

By (4), we have $q \perp \langle\langle \delta \rangle\rangle \in I^2 k$, hence

$$e_3(\langle\langle \lambda \rangle\rangle \cdot q) = \lambda \cdot e_2(q \perp \langle\langle \delta \rangle\rangle). \tag{5}$$

The computation of Witt invariants in [8, Chapter 5] yields

$$e_2(q \perp \langle\langle \delta \rangle\rangle) = [C(V, q)] + (-1) \cdot \operatorname{disc} q. \tag{6}$$

Since $\lambda \cdot \operatorname{disc} q = 0$ by hypothesis, (3) follows from (5) and (6). □

Proof of Theorem 2. If A is split, then using the same notation as in Lemma 1.1 we may take $\gamma(\sigma) = [C(V, q)]$, and Theorem 2 readily follows from Lemma 1.1. For the rest of the proof, we may thus assume A is not split, hence $\operatorname{disc} \sigma \neq 0$ since Z is assumed to split A. Let $G = \{\mathrm{Id}, \iota\}$ be the Galois group of Z/k. The properties of the Clifford algebra (see for instance [7, (9.12)]) yield

$$[C(A, \sigma)] - \iota[C(A, \sigma)] = [A]_Z = 0.$$

Therefore, $[C(A, \sigma)]$ lies in the subgroup $(\operatorname{Br} Z)^G$ of $\operatorname{Br} Z$ fixed under the action of G. The "Teichmüller cocycle" theory [6] (or the spectral sequence of group extensions, see [19, Remarque, p. 126]) yields an exact sequence

$$\operatorname{Br} k \to (\operatorname{Br} Z)^G \to H^3(G, Z^\times).$$

Since G is cyclic, $H^3(G, Z^\times) = H^1(G, Z^\times)$. By Hilbert's Theorem 90, $H^1(G, Z^\times) = 1$, hence $(\operatorname{Br} Z)^G$ is the image of the scalar extension map $\operatorname{Br} k \to \operatorname{Br} Z$, and there exists $\gamma(\sigma) \in \operatorname{Br} k$ such that $\gamma(\sigma)_Z = [C(A, \sigma)]$. Then, by [7, (9.12)],

$$2\gamma(\sigma) = N_{Z/k}\big([C(A, \sigma)]\big) = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ [A] & \text{if } n \text{ is even,} \end{cases} \tag{7}$$

hence $4\gamma(\sigma) = 0$. Therefore, $\gamma(\sigma) \in \operatorname{Br}_2(k) = H^2 k$.
Note that $\operatorname{ind} A = 2$, since A is split by the quadratic extension Z/k, hence A is Brauer-equivalent to a quaternion algebra Q. Let X be the conic associated with Q; the function field $k(X)$ splits A. Since Theorem 2 holds in the split case, we have

$$\lambda \cdot \gamma(\sigma) \in \ker\big(H^3 k \to H^3 k(X)\big).$$

By a theorem of (Arason–) Peyre [16, Proposition 4.4], the kernel on the right side is the subgroup $k^\times \cdot [A] \subset H^3 k$, hence

$$\lambda \cdot \gamma(\sigma) = 0 \quad \text{in } (H^3 k)/A.$$

□

Proof of Theorem 3. Suppose first A is split, and use the same notation as in Lemma 1.1. If $\lambda \in G(A,\sigma)$, then $\langle\lambda\rangle \cdot q \simeq q$ and Lemma 1.1 yields

$$\lambda \cdot \operatorname{disc} q = 0 \text{ in } H^2k \qquad \text{and} \qquad \lambda \cdot [C(V,q)] = 0 \text{ in } H^3k.$$

The first equation implies that $\lambda = N_{Z/k}(z)$ for some $z \in Z^\times$. Since

$$[C(A,\sigma)] = [C_0(V,q)] = [C(V,q)]_Z,$$

the projection formula yields

$$N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) = N_{Z/k}(z) \cdot [C(V,q)] = \lambda \cdot [C(V,q)] = 0,$$

proving the theorem if A is split.
If A is not split, we extend scalars to the function field $k(X)$ of the Severi–Brauer variety of A. For $\lambda \in G_+(A,\sigma)$, there still exists $z \in Z^\times$ such that $\lambda = N_{Z/k}(z)$, by Theorem 1. Since Theorem 3 holds in the split case, we have

$$N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) \in \ker\big(H^3k \to H^3k(X)\big),$$

and Peyre's theorem concludes the proof. (Note that applying Peyre's theorem requires the hypothesis that $\operatorname{ind} A \leq 4$.) $\square$

2 Algebras of low degree

We prove Theorem 4 by considering separately the cases $\operatorname{ind} A = 1$, 2, and 4.

2.1 Case 1: A is split

Let $A = \operatorname{End}_k V$, $\dim V \leq 6$, and let σ be adjoint to a quadratic form q on V. Since $C(A,\sigma) = C_0(V,q)$, we may choose $\gamma(\sigma) = [C(V,q)]$. The equations

$$\lambda \cdot \operatorname{disc} \sigma = 0 \text{ in } (H^2k)/A \qquad \text{and} \qquad \lambda \cdot \gamma(\sigma) = 0 \text{ in } (H^3k)/A$$

are then equivalent to

$$\lambda \cdot \operatorname{disc} q = 0 \text{ in } H^2k \qquad \text{and} \qquad \lambda \cdot [C(V,q)] = 0 \text{ in } H^3k,$$

hence, by Lemma 1.1, to $\langle\langle\lambda\rangle\rangle \cdot q \in I^4k$. Since $\dim q = 6$, the Arason–Pfister Hauptsatz [8, Chapter 10, Theorem 3.1] shows that this relation holds if and only if $\langle\langle\lambda\rangle\rangle \cdot q = 0$, i.e., $\lambda \in G(V,q) = G(A,\sigma)$, and the proof is complete.

2.2 Case 2: $\operatorname{ind} A = 2$

Let Q be a quaternion (division) algebra Brauer-equivalent to A. We represent A as $A = \operatorname{End}_Q U$ for some 3-dimensional (right) Q-vector space. The involution σ is then adjoint to a skew-hermitian form h on U (with respect to the conjugation involution on Q), which defines an element in the Witt group

$W^{-1}(Q)$. Let X be the conic associated with Q. The function field $k(X)$ splits Q, hence Morita equivalence yields an isomorphism

$$W^{-1}\big(Q \otimes k(X)\big) \simeq Wk(X).$$

Moreover, Dejaiffe [4] and Parimala–Sridharan–Suresh [15] have shown that the scalar extension map

$$W^{-1}(Q) \to W^{-1}\big(Q \otimes k(X)\big) \simeq Wk(X) \tag{8}$$

is injective. Let (V, q) be a quadratic space over $k(X)$ representing the image of (U, h) under (8). We may assume $\dim V = \deg A \le 6$ and σ is adjoint to q after scalar extension to $k(X)$. An element $\lambda \in k^\times$ lies in $G(V, q)$ if and only if $\langle\langle\lambda\rangle\rangle \cdot q = 0$; by the injectivity of (8), this condition is also equivalent to $\langle\langle\lambda\rangle\rangle \cdot h = 0$ in $W^{-1}(Q)$, i.e., to $\lambda \in G(A, \sigma)$. Therefore,

$$G(V, q) \cap k^\times = G(A, \sigma). \tag{9}$$

Suppose first A is not split by Z. Theorem 1 then shows that every similitude of (A, σ) is proper, and it only remains to show that if $\lambda = N_{Z/k}(z)$ for some $z \in Z^\times$ such that

$$N_{Z/k}\big(z \cdot [C(A, \sigma)]\big) = 0 \quad \text{in } (H^3k)/A,$$

then $\lambda \in G(A, \sigma)$. Extending scalars to $k(X)$, we derive from the last equation by the projection formula

$$N_{Z(X)/k(X)}(z) \cdot [C(V, q)] = 0 \quad \text{in } H^3k(X).$$

Therefore, by Lemma 1.1, $\langle\lambda\rangle \cdot q \equiv q \bmod I^4k(X)$, i.e.,

$$\langle\langle\lambda\rangle\rangle \cdot q \in I^4k(X).$$

Since $\dim q \le 6$, the Arason–Pfister Hauptsatz implies $\langle\langle\lambda\rangle\rangle \cdot q = 0$, hence $\lambda \in G(V, q)$ and therefore $\lambda \in G(A, \sigma)$ by (9). Theorem 4 is thus proved when $\operatorname{ind} A = 2$ and A is not split by Z.

Suppose next A is split by Z. In view of Theorems 1 and 2, it suffices to show that if $\lambda \in k^\times$ satisfies

$$\lambda \cdot \operatorname{disc} \sigma = 0 \text{ in } (H^2k)/A \qquad \text{and} \qquad \lambda \cdot \gamma(\sigma) = 0 \text{ in } (H^3k)/A,$$

then $\lambda \in G(A, \sigma)$. Again, extending scalars to $k(X)$, the conditions become

$$\lambda \cdot \operatorname{disc} q = 0 \text{ in } H^2k(X) \qquad \text{and} \qquad \lambda \cdot [C(V, q)] = 0 \text{ in } H^3k(X).$$

By Lemma 1.1, these equations imply $\langle\langle\lambda\rangle\rangle \cdot q \in I^4k(X)$, hence $\langle\langle\lambda\rangle\rangle \cdot q = 0$ by the Arason–Pfister Hauptsatz since $\dim q \le 6$. It follows that $\lambda \in G(V, q)$, hence $\lambda \in G(A, \sigma)$ by (9).

2.3 Case 3: $\operatorname{ind} A = 4$

Since $\deg A \leq 6$, this case arises only if $\deg A = 4$, i.e., A is a division algebra. This division algebra cannot be split by the quadratic k-algebra Z, hence all the similitudes are proper, by Theorem 1. Theorem 3 shows that if $\lambda \in G(A,\sigma)$, then there exists $z \in Z^\times$ such that $\lambda = N_{Z/k}(z)$ and $N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) = 0$ in $(H^3k)/A$, and it only remains to prove the converse.
Let $z \in Z^\times$ be such that $N_{Z/k}\big(z \cdot [C(A,\sigma)]\big) = u \cdot [A]$ for some $u \in k^\times$. Since by [7, (9.12)], $N_{Z/k}\big([C(A,\sigma)]\big) = [A]$, it follows that

$$N_{Z/k}\big(u^{-1}z \cdot [C(A,\sigma)]\big) = 0 \quad \text{in } H^3k. \tag{10}$$

Since $\deg A = 4$, the Clifford algebra $C(A,\sigma)$ is a quaternion algebra over Z. Let

$$C(A,\sigma) = (z_1, z_2)_Z.$$

Suppose first $\operatorname{disc}\sigma \neq 0$, i.e., Z is a field. Let $s\colon Z \to k$ be a k-linear map such that $s(1) = 0$, and let $s_*\colon WZ \to Wk$ be the corresponding (Scharlau) transfer map. By [2, Satz 3.3, Satz 4.18], Equation (10) yields

$$s_*\big(\langle\!\langle u^{-1}z, z_1, z_2\rangle\!\rangle\big) \in I^4k.$$

However, the form $s_*\big(\langle\!\langle u^{-1}z, z_1, z_2\rangle\!\rangle\big)$ is isotropic since $\langle\!\langle u^{-1}z, z_1, z_2\rangle\!\rangle$ represents 1 and $s(1) = 0$. Moreover, its dimension is 2^4, hence the Arason–Pfister Hauptsatz implies

$$s_*\big(\langle\!\langle u^{-1}z, z_1, z_2\rangle\!\rangle\big) = 0 \quad \text{in } Wk.$$

It follows that

$$s_*\big(\langle u^{-1}z\rangle \cdot \langle\!\langle z_1, z_2\rangle\!\rangle\big) = s_*\big(\langle\!\langle z_1, z_2\rangle\!\rangle\big),$$

hence the form on the left side is isotropic. Therefore, the form $\langle u^{-1}z\rangle \cdot \langle\!\langle z_1, z_2\rangle\!\rangle$ represents an element $v \in k^\times$. Then $v^{-1}u^{-1}z$ is represented by $\langle\!\langle z_1, z_2\rangle\!\rangle$, which is the reduced norm form of $C(A,\sigma)$, hence $z \in k^\times \operatorname{Nrd}(C(A,\sigma)^\times)$, and

$$N_{Z/k}(z) \in k^{\times 2} N_{Z/k}\big(\operatorname{Nrd}(C(A,\sigma)^\times)\big).$$

By [7, (15.11)], the group on the right is $G_+(A,\sigma)$. We have thus proved $N_{Z/k}(z) \in G(A,\sigma)$, and the proof is complete when Z is a field.
Suppose finally $\operatorname{disc}\sigma = 0$, i.e., $Z \simeq k \times k$. Then $C(A,\sigma) \simeq C' \times C''$ for some quaternion k-algebras $C' = (z_1', z_2')_k$ and $C'' = (z_1'', z_2'')_k$, and [7, (15.13)] shows

$$G(A,\sigma) = \operatorname{Nrd}(C'^\times)\operatorname{Nrd}(C''^\times).$$

We also have $z = (z', z'')$ for some z', $z'' \in k^\times$, and (10) becomes

$$u^{-1}z' \cdot [C'] + u^{-1}z'' \cdot [C''] = 0 \quad \text{in } H^3k.$$

It follows that

$$\langle\!\langle u^{-1}z', z_1', z_2'\rangle\!\rangle \simeq \langle\!\langle u^{-1}z'', z_1'', z_2''\rangle\!\rangle.$$

By [2, Lemma 1.7], there exists $v \in k^\times$ such that

$$\langle\langle u^{-1}z', z_1', z_2'\rangle\rangle \simeq \langle\langle v, z_1', z_2'\rangle\rangle \simeq \langle\langle v, z_1'', z_2''\rangle\rangle \simeq \langle\langle u^{-1}z'', z_1'', z_2''\rangle\rangle,$$

hence $v^{-1}u^{-1}z' \in \operatorname{Nrd}(C')$ and $v^{-1}u^{-1}z'' \in \operatorname{Nrd}(C'')$. Therefore,

$$N_{Z/k}(z) = z'z'' \in \operatorname{Nrd}(C'^\times)\operatorname{Nrd}(C''^\times),$$

and the proof of Theorem 4 is complete.
To finish this section, we compare the descriptions of $G_+(A,\sigma)$ for $\deg A = 4$ or 6 in [7] with those which follow from Theorem 4 (and Remark 3).

Corollary 2.1. *Suppose* $\deg A = 4$. *If* $\operatorname{disc}\sigma \neq 0$, *then*

$$\begin{aligned} G_+(A,\sigma) &= k^{\times 2} N_{Z/k}\bigl(\operatorname{Nrd}(C(A,\sigma)^\times)\bigr) \\ &= \{N_{Z/k}(z) \mid N_{Z/k}\bigl(z\cdot[C(A,\sigma)]\bigr) = 0 \textit{ in } (H^3k)/A\}. \end{aligned}$$

If $\operatorname{disc}\sigma = 0$, *then* $C(A,\sigma) \simeq C' \times C''$ *for some quaternion* k*-algebras* C', C'', *and*

$$\begin{aligned} G_+(A,\sigma) &= \operatorname{Nrd}(C'^\times)\operatorname{Nrd}(C''^\times) \\ &= \{z'z'' \mid z'\cdot[C'] + z''\cdot[C''] = 0 \textit{ in } (H^3k)/A\}. \end{aligned}$$

Proof. See [7, (15.11)] for the case $\operatorname{disc}\sigma \neq 0$ and [7, (15.13)] for the case $\operatorname{disc}\sigma = 0$. □

Corollary 2.2. *Suppose* $\deg A = 6$. *If* $\operatorname{disc}\sigma \neq 0$, *let* ι *be the nontrivial automorphism of the field extension* Z/k *and let* $\underline{\sigma}$ *be the canonical (unitary) involution of* $C(A,\sigma)$. *Let also*

$$\operatorname{GU}(C(A,\sigma),\underline{\sigma}) = \{g \in C(A,\sigma) \mid \underline{\sigma}(g)g \in k^\times\}.$$

Then

$$\begin{aligned} &G_+(A,\sigma) = \\ &\quad \{N_{Z/k}(z) \mid z\iota(z)^{-1} = (\underline{\sigma}(g)g)^{-2}\operatorname{Nrd}(g) \textit{ for some } g \in \operatorname{GU}(C(A,\sigma),\underline{\sigma})\} \\ &\qquad\qquad = \{N_{Z/k}(z) \mid N_{Z/k}\bigl(z\cdot[C(A,\sigma)]\bigr) = 0 \textit{ in } (H^3k)/A\}. \end{aligned}$$

If $\operatorname{disc}\sigma = 0$, *then* $C(A,\sigma) \simeq C \times C^{\mathrm{op}}$ *for some central simple* k*-algebra* C *of degree* 4, *and*

$$\begin{aligned} G_+(A,\sigma) &= k^{\times 2}\operatorname{Nrd}(C^\times) \\ &= \{z \in k^\times \mid z\cdot[C] = 0 \textit{ in } (H^3k)/A\}. \end{aligned}$$

Proof. See [7, (15.31)] for the case $\operatorname{disc}\sigma \neq 0$ and [7, (15.34)] for the case $\operatorname{disc}\sigma = 0$. In the latter case, Theorem 3 shows that $G_+(A,\sigma)$ consists of products $z'z''$ where z', $z'' \in k^\times$ are such that

$$z' \cdot [C] + z'' \cdot [C^{\mathrm{op}}] = 0 \quad \text{in } (H^3 k)/A.$$

However, $[C^{\mathrm{op}}] = -[C]$, and $2[C] = [A]$ by [7, (9.15)], hence

$$z' \cdot [C] + z'' \cdot [C^{\mathrm{op}}] = z'z'' \cdot [C] \quad \text{in } (H^3 k)/A.$$

Note that the equation

$$k^{\times 2} \operatorname{Nrd}(C^\times) = \{z \in k^\times \mid z \cdot [C] = 0 \text{ in } (H^3 k)/A\}$$

can also be proved directly by a theorem of Merkurjev [11, Proposition 1.15]. □

3 Fields of low virtual cohomological dimension

Our goal in this section is to prove Theorem 5. Together with Theorem 2, the following lemma completes the proof of the "only if" part:

Lemma 3.1. *If $\lambda \in G(A,\sigma)$, then $\lambda > 0$ at every ordering v such that $(A,\sigma)_v$ is not hyperbolic.*

Proof. If $(A,\sigma)_v$ is not hyperbolic, then A_v is split, by [18, Chapter 10, Theorem 3.7]. We may thus represent $A_v = \operatorname{End}_{k_v} V$ for some k_v-vector space V, and $\sigma \otimes \operatorname{Id}_{k_v}$ is adjoint to a non-hyperbolic quadratic form q. If $\lambda \in G(A,\sigma)$, then $\lambda \in G(V,q)$, hence

$$\langle\lambda\rangle \cdot q \simeq q.$$

Comparing the signatures of each side, we obtain $\lambda > 0$. □

For the "if" part, we use the following lemma:

Lemma 3.2. *Let F be an arbitrary field of characteristic different from 2. If $\operatorname{vcd}_2 F \leq 3$, then the torsion part of the 4-th power of IF is trivial,*

$$I_t^4 F = 0.$$

Proof. Our proof uses the existence of the cohomological invariants $e_n \colon I^n F \to H^n(F,\mu_2)$, and the fact that $\ker e_n = I^{n+1}F$, proved for fields of virtual cohomological 2-dimension at most 3 by Arason–Elman–Jacob [3].
Suppose first $-1 \notin F^{\times 2}$. From $\operatorname{vcd}_2 F \leq 3$, it follows that $H^n(F(\sqrt{-1}),\mu_2) = 0$ for $n \geq 4$, hence the Arason exact sequence

$$H^n(F(\sqrt{-1}),\mu_2) \xrightarrow{N} H^n(F,\mu_2) \xrightarrow{(-1)_1\cup} H^{n+1}(F,\mu_2) \to H^{n+1}(F(\sqrt{-1}),\mu_2)$$

(see [2, Corollar 4.6] or [7, (30.12)]) shows that the cup-product with $(-1)_1$ is an isomorphism $H^n(F,\mu_2) \simeq H^{n+1}(F,\mu_2)$ for $n \geq 4$. If $q \in I_t^4F$, there is an integer ℓ such that $2^\ell q = 0$, hence the 4-th invariant $e_4(q) \in H^4(F,\mu_2)$ satisfies

$$\underbrace{(-1)_1 \cup \cdots \cup (-1)_1}_{\ell} \cup e_4(q) = 0 \quad \text{in } H^{\ell+4}(F,\mu_2).$$

Since $(-1)_1\cup$ is an isomorphism, it follows that $e_4(q) = 0$, hence $q \in I_t^5F$. Repeating the argument with e_5, e_6, ..., we obtain $q \in \bigcap_n I^nF$, hence $q = 0$ by the Arason–Pfister Hauptsatz [8, p. 290].
If $-1 \in F^{\times 2}$, then the hypothesis implies that $H^n(F,\mu_2) = 0$ for $n \geq 4$, hence for $q \in I^4F$ we get successively $e_4(q) = 0$, $e_5(q) = 0$, etc., and we conclude as before. □

Proof of Theorem 5. As observed above, the "only if" part follows from Theorem 2 and Lemma 3.1. The proof of the "if" part uses the same arguments as the proof of Theorem 2 in the case where $\operatorname{ind} A = 2$.
We first consider the split case. If $A = \operatorname{End}_k V$ and σ is adjoint to a quadratic form q on V, then we may choose $\gamma(\sigma) = C(V,q)$, and the conditions

$$\lambda \cdot \operatorname{disc}\sigma = 0 \text{ in } (H^2k)/A \qquad \text{and} \qquad \lambda \cdot \gamma(\sigma) = 0 \text{ in } (H^3k)/A$$

imply, by Lemma 1.1, that $\langle\langle\lambda\rangle\rangle \cdot q \in I^4k$. Moreover, for every ordering v on k, the signature $\operatorname{sgn}_v(\langle\langle\lambda\rangle\rangle \cdot q)$ vanishes, since $\lambda > 0$ at every v such that $\operatorname{sgn}_v(q) \neq 0$. Therefore, by Pfister's local-global principle [8, Chapter 8, Theorem 4.1], $\langle\langle\lambda\rangle\rangle \cdot q$ is torsion. Since the hypothesis on k implies, by Lemma 3.2, that $I_t^4k = 0$, we have $\langle\langle\lambda\rangle\rangle \cdot q = 0$, hence $\lambda \in G(V,q) = G(A,\sigma)$. Note that Lemma 3.2 yields $I_t^4k = 0$ under the weaker hypothesis $\operatorname{vcd}_2 k \leq 3$. Therefore, the split case of Theorem 5 holds when $\operatorname{vcd}_2 k \leq 3$.
Now, suppose A is not split. Since A is split by Z, it is Brauer-equivalent to a quaternion algebra Q. Let $k(X)$ be the function field of the conic X associated with Q. This field splits A, hence there is a quadratic space (V,q) over $k(X)$ such that $A \otimes k(X)$ may be identified with $\operatorname{End}_{k(X)} V$ and $\sigma \otimes \operatorname{Id}_{k(X)}$ with the adjoint involution with respect to q. As in Section 2 (see Equation (9)), we have

$$G(V,q) \cap k^\times = G(A,\sigma).$$

Therefore, it suffices to show that the conditions on λ imply $\lambda \in G(V,q)$.
If v is an ordering of k such that $(A,\sigma)_v$ is hyperbolic, then q_w is hyperbolic for any ordering w of $k(X)$ extending v, since hyperbolic involutions remain hyperbolic over scalar extensions. Therefore, $\lambda > 0$ at every ordering w of $k(X)$ such that q_w is not hyperbolic. Moreover, the conditions

$$\lambda \cdot \operatorname{disc}\sigma = 0 \text{ in } (H^2k)/A \qquad \text{and} \qquad \lambda \cdot \gamma(\sigma) = 0 \text{ in } (H^3k)/A$$

imply

$$\lambda \cdot \operatorname{disc} q = 0 \text{ in } H^2k(X) \qquad \text{and} \qquad \lambda \cdot [C(V,q)] = 0 \text{ in } H^3k(X).$$

Since X is a conic, Proposition 11, p. 93 of [20] implies

$$\operatorname{vcd}_2 k(X) = 1 + \operatorname{vcd}_2 k \leq 3.$$

As Theorem 5 holds in the split case over fields of virtual cohomological 2-dimension at most 3, it follows that $\lambda \in G(V, q)$. □

Remark. The same arguments show that if $\operatorname{vcd}_2 k \leq 2$ and $\operatorname{ind} A = 2$, then $G_+(A, \sigma)$ consists of the elements $N_{Z/k}(z)$ where $z \in Z^\times$ is such that

$$N_{Z/k}\big(z \cdot [C(A, \sigma)]\big) = 0 \quad \text{in } (H^3 k)/A.$$

4 Examples

In this section, we give an explicit description of the element $\gamma(\sigma)$ of Theorem 2 in some special cases. Throughout this section, we assume the algebra A is not split, and is split by Z (hence Z is a field and $\operatorname{disc} \sigma \neq 0$). Our first result is easy:

Proposition 4.1. *If A is split by Z and σ becomes hyperbolic after scalar extension to Z, then we may choose $\gamma(\sigma) = 0$.*

Proof. Let ι be the nontrivial automorphism of Z/k. Since Z is the center of $C(A, \sigma)$,

$$C(A, \sigma) \otimes_k Z \simeq C(A, \sigma) \times {}^{\iota}C(A, \sigma). \tag{11}$$

On the other hand, $C(A, \sigma) \otimes_k Z \simeq C(A \otimes_k Z, \sigma \otimes \operatorname{Id}_Z)$, and since σ becomes hyperbolic over Z, one of the components of $C(A \otimes_k Z, \sigma \otimes \operatorname{Id}_Z)$ is split, by [7, (8.31)]. Therefore,

$$[C(A, \sigma)] = [{}^{\iota}C(A, \sigma)] = 0 \qquad \text{in } \operatorname{Br} Z.$$

□

Corollary 4.2. *In the conditions of Proposition 4.1, if* $\deg A \leq 6$ *or* $\operatorname{vcd}_2 k \leq 2$, *then*

$$G_+(A, \sigma) = \{\lambda \in k^\times \mid \lambda \cdot \operatorname{disc} \sigma = 0 \ \textit{in } H^2 k\}$$

and

$$G_-(A, \sigma) = \{\lambda \in k^\times \mid \lambda \cdot \operatorname{disc} \sigma = [A] \ \textit{in } H^2 k\}.$$

Proof. This readily follows from Proposition 4.1 and Theorem 2 or 5. □

To give further examples where $\gamma(\sigma)$ can be computed, we fix a particular representation of A as follows. Since A is assumed to be split by Z, it is Brauer-equivalent to a quaternion k-algebra Q containing Z. We choose a quaternion basis 1, i, j, ij of Q such that $Z = k(i)$. Let $A = \operatorname{End}_Q U$ for some right Q-vector space U, and let σ be the adjoint involution of a skew-hermitian form h on U with respect to the conjugation involution on Q. For x, $y \in U$, we decompose

$$h(x,y) = f(x,y) + jg(x,y) \qquad \text{with } f(x,y),\ g(x,y) \in Z.$$

It is easily verified that f (resp. g) is a skew-hermitian (resp. symmetric bilinear) form on U viewed as a Z-vector space. (See [18, Chapter 10, Lemma 3.1].) We have

$$A \otimes_k Z = (\operatorname{End}_Q U) \otimes_k Z = \operatorname{End}_Z U.$$

Moreover, for x, $y \in U$ and $\varphi \in \operatorname{End}_Q U$, the equation

$$h\bigl(x, \varphi(y)\bigr) = h\bigl(\sigma(\varphi)(x), y\bigr)$$

implies

$$g\bigl(x, \varphi(y)\bigr) = g\bigl(\sigma(\varphi)(x), y\bigr),$$

hence $\sigma \otimes_k \operatorname{Id}_Z$ is adjoint to g.

Proposition 4.3. *With the notation above,*

$$[C(A,\sigma)] = [C(U,g)] \qquad \text{in } \operatorname{Br} Z.$$

Proof. Since $\sigma \otimes \operatorname{Id}_Z$ is the adjoint involution of g,

$$C(A \otimes_k Z, \sigma \otimes \operatorname{Id}_Z) \simeq C_0(U,g). \tag{12}$$

Now, $\operatorname{disc}\sigma$ is a square in Z, hence $C_0(U,g)$ decomposes into a direct product

$$C_0(U,g) \simeq C' \times C'' \tag{13}$$

where C', C'' are central simple Z-algebras Brauer-equivalent to $C(U,g)$. The proposition follows from (11), (12), and (13). □

To give an explicit description of g, consider an h-orthogonal basis $(e_1,\dots,e_n)$ of U. In the corresponding diagonalization of h,

$$h \simeq \langle u_1,\dots,u_n\rangle,$$

each $u_\ell \in Q$ is a pure quaternion, since h is skew-hermitian. Let $u_\ell^2 = a_\ell \in k^\times$ for $\ell = 1, \dots, n$. Then

$$\operatorname{disc}\sigma = (-1)^n \operatorname{Nrd}(u_1)\dots\operatorname{Nrd}(u_n) = a_1\dots a_n,$$

so we may assume $i^2 = a_1 \dots a_n$. Write

$$u_\ell = \mu_\ell i + jv_\ell \qquad \text{where } \mu_\ell \in k \text{ and } v_\ell \in Z. \tag{14}$$

Each $e_\ell Q$ is a 2-dimensional Z-vector space, and we have a g-orthogonal decomposition

$$U = e_1Q \oplus \dots \oplus e_nQ.$$

If $v_\ell = 0$, then $g(e_\ell, e_\ell) = 0$, hence $e_\ell Q$ is hyperbolic. If $v_\ell \neq 0$, then $(e_\ell, e_\ell u_\ell)$ is a g-orthogonal basis of $e_\ell Q$, which yields the following diagonalization of the restriction of g:

$$\langle v_\ell, -a_\ell v_\ell\rangle.$$

Therefore,

$$g = g_1 + \cdots + g_n \tag{15}$$

where

$$g_\ell = \begin{cases} 0 & \text{if } v_\ell = 0, \\ \langle v_\ell\rangle\langle 1, -a_\ell\rangle & \text{if } v_\ell \neq 0. \end{cases} \tag{16}$$

We now consider in more detail the cases $n = 2$ and $n = 3$.

4.1 Algebras of degree 4

Suppose $\deg A = 4$, i.e., $n = 2$, and use the same notation as above. If $v_1 = 0$, then squaring each side of (14) yields $a_1 = \mu_1^2 a_1 a_2$, hence $a_2 \in k^{\times 2}$, a contradiction since Q is assumed to be a division algebra. The case $v_2 = 0$ leads to the same contradiction. Therefore, we necessarily have $v_1 \neq 0$ and $v_2 \neq 0$. By (15) and (16),

$$g = \langle v_1\rangle\langle 1, -a_1\rangle + \langle v_2\rangle\langle 1, -a_2\rangle,$$

hence by [8, p. 121],

$$\begin{aligned} [C(A,\sigma)] &= (a_1, v_1)_Z + (a_2, v_2)_Z + (a_1, a_2)_Z \\ &= (a_1, -v_1 v_2)_Z. \end{aligned} \tag{17}$$

Since the division algebra Q contains the pure quaternions u_1, u_2 and i with $u_1^2 = a_1$, $u_2^2 = a_2$ and $i^2 = a_1 a_2$, we have a_1, a_2, $a_1 a_2 \notin k^{\times 2}$ and we may consider the field extension

$$L = k(\sqrt{a_1}, \sqrt{a_2}).$$

We identify Z with a subfield of L by choosing in L a square root of $a_1 a_2$, and denote by ρ_1, ρ_2 the automorphisms of L/k defined by

$$\rho_1(\sqrt{a_1}) = -\sqrt{a_1}, \qquad \rho_2(\sqrt{a_1}) = \sqrt{a_1},$$
$$\rho_1(\sqrt{a_2}) = \sqrt{a_2}, \qquad \rho_2(\sqrt{a_2}) = -\sqrt{a_2}.$$

Thus, $Z \subset L$ is the subfield of $\rho_1 \circ \rho_2$-invariant elements. Let $j^2 = b$. Then (14) yields

$$a_1 = \mu_1^2 a_1 a_2 + b N_{Z/k}(v_1), \qquad a_2 = \mu_2^2 a_1 a_2 + b N_{Z/k}(v_2),$$

hence $N_{Z/k}(-v_1 v_2) = a_1 a_2 b^{-2}(1 - \mu_1^2 a_2)(1 - \mu_2^2 a_1)$ and

$$\frac{-v_1 v_2}{\rho_1(-v_1 v_2)} = \frac{-v_1 v_2}{\rho_2(-v_1 v_2)} = \frac{a_1 a_2}{b^2 \rho_1(-v_1 v_2)^2}(1 - \mu_1^2 a_2)(1 - \mu_2^2 a_1).$$

Since $L = Z(\sqrt{a_1}) = Z(\sqrt{a_2})$, it follows that $1 - \mu_1^2 a_2$ and $1 - \mu_2^2 a_1$ are norms from L/Z. Therefore, the preceding equation yields

$$\frac{-v_1 v_2}{\rho_1(-v_1 v_2)} = \frac{-v_1 v_2}{\rho_2(-v_1 v_2)} = N_{L/Z}(\ell) \qquad \text{for some } \ell \in L^\times.$$

Since $N_{Z/k}(-v_1 v_2 \rho_1(-v_1 v_2)^{-1}) = 1$, we have $N_{L/k}(\ell) = 1$. By Hilbert's Theorem 90, there exists $b_1 \in L^\times$ such that

$$\rho_1(b_1) = b_1 \qquad \text{and} \qquad b_1 \rho_2(b_1)^{-1} = \ell \rho_1(\ell). \tag{18}$$

Set $b_2 = -v_1 v_2 \rho_1(\ell) b_1^{-1}$. Computation yields

$$\rho_2(b_2) = b_2 \qquad \text{and} \qquad \rho_1(b_2) b_2^{-1} = \ell \rho_2(\ell). \tag{19}$$

Define an algebra E over k by

$$E = L \oplus L r_1 \oplus L r_2 \oplus L r_1 r_2$$

where the multiplication is defined by

$$\begin{aligned} r_1 x = \rho_1(x) r_1, \qquad r_2 x = \rho_2(x) r_2 \qquad &\text{for } x \in L, \\ r_1^2 = b_1, \qquad r_2^2 = b_2, \qquad \text{and} \qquad &r_1 r_2 = \ell r_2 r_1. \end{aligned}$$

Since b_1, b_2 and ℓ satisfy (18) and (19), the algebra E is a crossed product, see [1]. It is thus a central simple k-algebra of degree 4.

PROPOSITION 4.4. *With the notation above, we may choose* $\gamma(\sigma) = [E] \in \operatorname{Br} k$.

Proof. The centralizer $C_E Z$ of Z in E is $L \oplus L r_1 r_2$. Computation shows that

$$(r_1 r_2)^2 = -v_1 v_2.$$

Since conjugation by $r_1 r_2$ maps $\sqrt{a_1} \in L$ to its opposite, it follows that

$$C_E Z = (a_1, -v_1 v_2)_Z.$$

Since $[C_E Z] = [E]_Z$, the proposition follows from (17). □

COROLLARY 4.5. *Let*

$$E_+ = C_E Z = \{x \in E^\times \mid xz = zx \textit{ for all } z \in Z\}$$

and

$$E_- = \{x \in E^\times \mid xz = \rho_1(z) x \textit{ for all } z \in Z\}.$$

Then

$$G_+(A,\sigma) = k^{\times 2} \operatorname{Nrd}_E(E_+) \qquad \textit{and} \qquad G_-(A,\sigma) = k^{\times 2} \operatorname{Nrd}_E(E_-).$$

Proof. As observed in the proof of Proposition 4.4, $C_E Z \simeq C(A,\sigma)$. Since, by [5, Corollary 5, p. 150],

$$\operatorname{Nrd}_E(x) = N_{Z/k}(\operatorname{Nrd}_{C_E Z} x) \qquad \text{for } x \in C_E Z,$$

the description of $G_+(A,\sigma)$ above follows from [7, (15.11)] (see also Corollary 2.1).
To prove $k^{\times 2}\operatorname{Nrd}_E(E_-) \subset G_-(A,\sigma)$, it obviously suffices to prove $\operatorname{Nrd}_E(E_-) \subset G_-(A,\sigma)$. From the definition of E, it follows that $r_1 \in E_-$. By [10, p. 80],

$$\operatorname{Nrd}_E(r_1)\cdot[E] = 0 \qquad \text{in } H^3k. \tag{20}$$

Let $L_1 \subset L$ be the subfield fixed under ρ_1. We have $r_1^2 = b_1 \in L_1$, hence

$$\operatorname{Nrd}_E(r_1) = N_{L_1/k}(b_1).$$

On the other hand, the centralizer of L_1 is

$$C_E L_1 = L \oplus L r_1 \simeq (a_1a_2, b_1)_{L_1},$$

hence

$$[N_{L_1/k}(C_E L_1)] = \bigl(a_1a_2, N_{L_1/k}(b_1)\bigr)_k = \operatorname{Nrd}_E(r_1)\cdot\operatorname{disc}\sigma \quad \text{in } H^2k. \tag{21}$$

Since $[C_E L_1] = [E_{L_1}]$, we have $[N_{L_1/k}(C_E L_1)] = 2[E]$. But $2[E] = 2\gamma(\sigma) = [A]$ by (7), hence (21) yields

$$\operatorname{Nrd}_E(r_1)\cdot\operatorname{disc}\sigma = [A] \qquad \text{in } H^2k. \tag{22}$$

From (20), (22) and Theorems 1, 2 it follows that $\operatorname{Nrd}_E(r_1) \in G_-(A,\sigma)$.
Now, suppose $x \in E_-$. Then $r_1x \in E_+$, hence $\operatorname{Nrd}_E(r_1x) \in G_+(A,\sigma)$ by the first part of the corollary. Since

$$G_+(A,\sigma)G_-(A,\sigma) = G_-(A,\sigma)$$

it follows that

$$\operatorname{Nrd}_E(x) \in \operatorname{Nrd}_E(r_1)G_+(A,\sigma) = G_-(A,\sigma).$$

We have thus proved $k^{\times 2}\operatorname{Nrd}_E(E_-) \subset G_-(A,\sigma)$.
To prove the reverse inclusion, consider $\lambda \in G_-(A,\sigma)$. Since

$$G_-(A,\sigma)G_-(A,\sigma) = G_+(A,\sigma),$$

we have $\lambda\operatorname{Nrd}_E(r_1) \in G_+(A,\sigma)$, hence by the first part of the corollary,

$$\lambda\operatorname{Nrd}_E(r_1) \in k^{\times 2}\operatorname{Nrd}_E(E_+).$$

It follows that

$$\lambda \in k^{\times 2}\operatorname{Nrd}_E(r_1E_+) = k^{\times 2}\operatorname{Nrd}_E(E_-).$$

□

4.2 Algebras of degree 6

Suppose $\deg A = 6$, i.e., $n = 3$, and use the same notation as in the beginning of this section. If σ (i.e., h) is isotropic, then h is Witt-equivalent to a rank 1 skew-hermitian form, say $\langle u\rangle$. Hence $i^2 = \operatorname{disc}\sigma = u^2 \in k^\times$. Hence we may assume that h is Witt-equivalent to the rank 1 skew-hermitian form $\langle \mu i\rangle$ for some $\mu \in k^\times$. This implies that the form g is hyperbolic and $C(U,g)$ is split. Hence we may choose $\gamma(\sigma) = 0$. By Theorem 4, we then have $\lambda \in G(A,\sigma)$ if and only if $\lambda.\operatorname{disc}\sigma = 0$ in $(H^2k)/A$. If σ becomes isotropic over Z, the form g is isotropic, hence we may choose a diagonalization of h

$$h \simeq \langle u_1, u_2, u_3\rangle$$

such that $g(u_3, u_3) = 0$, i.e., in the notation of (14), $u_3 = \mu_3 i$. Raising each side to the square, we obtain

$$a_3 = \mu_3^2 a_1 a_2 a_3,$$

hence $a_1 \equiv a_2 \bmod k^{\times 2}$. It follows that u_2 is conjugate to a scalar multiple of u_1, i.e., there exists $x \in Q^\times$ and $\theta \in k^\times$ such that

$$u_2 = \theta x u_1 x^{-1} = \theta \operatorname{Nrd}_Q(x)^{-1} x u_1 \overline{x}.$$

Since $\langle u_1\rangle \simeq \langle x u_1 \overline{x}\rangle$, we may let $\nu = -\theta \operatorname{Nrd}(x)^{-1} \in k^\times$ to obtain

$$h \simeq \langle u_1, -\nu u_1, \mu_3 i\rangle.$$

If $v_1 = 0$, then g is hyperbolic, hence we may choose $\gamma(\sigma) = 0$ by Proposition 4.1. If $v_1 \neq 0$, then (15) and (16) yield

$$g = \langle v_1\rangle\langle 1, -a_1\rangle + \langle -\nu v_1\rangle\langle 1, -a_1\rangle = \langle v_1\rangle\langle\langle a_1, \nu\rangle\rangle.$$

The Clifford algebra of g is the quaternion algebra $(a_1, \nu)_Z$, hence we may choose

$$\gamma(\sigma) = (a_1, \nu)_k.$$

Suppose finally that σ does not become isotropic over Z, hence v_1, v_2, $v_3 \neq 0$. Then

$$g = \langle v_1\rangle\langle 1, -a_1\rangle + \langle v_2\rangle\langle 1, -a_2\rangle + \langle v_3\rangle\langle 1, -a_3\rangle$$

and, by Proposition 4.3,

$$[C(A,\sigma)] = (a_1, v_1)_Z + (a_2, v_2)_Z + (a_3, v_3)_Z + (a_1, a_2)_Z + (a_1, a_3)_Z + (a_2, a_3)_Z.$$

Since $Z = k(\sqrt{a_1 a_2 a_3})$, the right side simplifies to

$$[C(A,\sigma)] = (a_1, v_1 v_3)_Z + (a_2, v_2 v_3)_Z + (a_1, a_2)_Z + (a_1 a_2, -1)_Z. \tag{23}$$

By [7, (9.16)], $N_{Z/k}C(A,\sigma)$ is split, hence

$$\big(a_1, N_{Z/k}(v_1 v_3)\big)_k = \big(a_2, N_{Z/k}(v_2 v_3)\big)_k \qquad \text{in } \operatorname{Br} k.$$

By the "common slot lemma" (see for instance [2, Lemma 1.7]), there exists $\alpha \in k^\times$ such that

$$\big(a_1, N_{Z/k}(v_1v_3)\big)_k = \big(\alpha, N_{Z/k}(v_1v_3)\big)_k = \big(\alpha, N_{Z/k}(v_2v_3)\big)_k = \big(a_2, N_{Z/k}(v_2v_3)\big)_k.$$

Then

$$\big(\alpha a_1, N_{Z/k}(v_1v_3)\big)_k = \big(\alpha a_2, N_{Z/k}(v_2v_3)\big)_k = \big(\alpha, N_{Z/k}(v_1v_2)\big)_k = 0.$$

By [21, (2.6)], there exist β_1, β_2, $\beta_3 \in k^\times$ such that

$$(\alpha a_1, v_1v_3)_Z = (\alpha a_1, \beta_1)_Z, \qquad (\alpha a_2, v_2v_3)_Z = (\alpha a_2, \beta_2)_Z,$$
$$(\alpha, v_1v_2)_Z = (\alpha, \beta_3)_Z.$$

Since

$$(a_1, v_1v_3)_Z + (a_2, v_2v_3)_Z = (\alpha a_1, v_1v_3)_Z + (\alpha a_2, v_2v_3)_Z + (\alpha, v_1v_2)_Z,$$

it follows from (23) that

$$[C(A,\sigma)] = (\alpha a_1, \beta_1)_Z + (\alpha a_2, \beta_2)_Z + (\alpha, \beta_3)_Z + (a_1, a_2)_Z + (a_1a_2, -1)_Z.$$

We may thus take

$$\begin{aligned}\gamma(\sigma) &= (a_1, \beta_1)_k + (a_2, \beta_2)_k + (\alpha, \beta_1\beta_2\beta_3)_k + (a_1, a_2)_k + (a_1a_2, -1)_k\\ &= (a_1, -a_2\beta_1)_k + (a_2, -\beta_2)_k + (\alpha, \beta_1\beta_2\beta_3)_k.\end{aligned}$$

References

[1] S. A. Amitsur and D. Saltman, Generic Abelian crossed products and p-algebras, J. Algebra 51 (1978), no. 1, 76–87. MR0491789 (58 #10988)

[2] J. Kr. Arason, Cohomologische invarianten quadratischer Formen, J. Algebra 36 (1975), no. 3, 448–491. MR0389761 (52 #10592)

[3] J. Kr. Arason, R. Elman and B. Jacob, Fields of cohomological 2-dimension three, Math. Ann. 274 (1986), no. 4, 649–657. MR0848510 (87m:12006)

[4] I. Dejaiffe, Formes antihermitiennes devenant hyperboliques sur un corps de déploiement, C. R. Acad. Sci. Paris Sér. I Math. 332 (2001), no. 2, 105–108. MR1813765 (2001m:11054)

[5] P. K. Draxl, *Skew fields*, Cambridge Univ. Press, Cambridge, 1983. MR0696937 (85a:16022)

[6] S. Eilenberg and S. MacLane, Cohomology and Galois theory. I. Normality of algebras and Teichmüller's cocycle, Trans. Amer. Math. Soc. 64 (1948), 1–20. MR0025443 (10,5e)

[7] M.-A. Knus et al., *The book of involutions*, Amer. Math. Soc., Providence, RI, 1998. MR1632779 (2000a:16031)

[8] T. Y. Lam, *The algebraic theory of quadratic forms*, W. A. Benjamin, Inc., Reading, Mass., 1973. MR0396410 (53 #277)

[9] A. S. Merkurjev, On the norm residue symbol of degree 2, Dokl. Akad. Nauk SSSR 261 (1981), no. 3, 542–547. MR0638926 (83h:12015)

[10] A. S. Merkurjev, K-theory of simple algebras, in *K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992)*, 65–83, Proc. Sympos. Pure Math., Part 1, Amer. Math. Soc., Providence, RI. MR1327281 (96f:19004)

[11] A. S. Merkurjev, Certain K-cohomology groups of Severi-Brauer varieties, in *K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992)*, 319–331, Proc. Sympos. Pure Math., Part 2, Amer. Math. Soc., Providence, RI. MR1327307 (96g:19004)

[12] A. S. Merkurjev, R. Parimala and J.-P. Tignol, Invariants of quasi-trivial tori and the Rost invariant, Algebra i Analiz 14 (2002) 110–151; St. Petersburg Math. J. 14 (2003) 791–821.

[13] A. S. Merkurjev and A. A. Suslin, Norm residue homomorphism of degree three, Izv. Akad. Nauk SSSR Ser. Mat. 54 (1990), no. 2, 339–356; translation in Math. USSR-Izv. 36 (1991), no. 2, 349–367. MR1062517 (91f:11083)

[14] A. S. Merkurjev and J.-P. Tignol, The multipliers of similitudes and the Brauer group of homogeneous varieties, J. Reine Angew. Math. 461 (1995), 13–47. MR1324207 (96c:20083)

[15] R. Parimala, R. Sridharan and V. Suresh, Hermitian analogue of a theorem of Springer, J. Algebra 243 (2001), no. 2, 780–789. MR1850658 (2002g:11043)

[16] E. Peyre, Products of Severi-Brauer varieties and Galois cohomology, in *K-theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992)*, 369–401, Proc. Sympos. Pure Math., Part 2, Amer. Math. Soc., Providence, RI. MR1327310 (96d:19008)

[17] M. Rost, On Hilbert Satz 90 for K_3 for degree-two extensions, preprint, Regensburg 1986.
`http://www.mathematik.uni-bielefeld.de/~rost/K3-86.html`

[18] W. Scharlau, *Quadratic and Hermitian forms*, Springer, Berlin, 1985. MR0770063 (86k:11022)

[19] J.-P. Serre, *Corps locaux*, Actualités Sci. Indust., No. 1296. Hermann, Paris, 1962. MR0150130 (27 #133)

[20] J-P. Serre, *Cohomologie galoisienne*, Fifth edition, Springer, Berlin, 1994. MR1324577 (96b:12010)

[21] J.-P. Tignol, Corps à involution neutralisés par une extension abélienne élémentaire, in *The Brauer group (Sem., Les Plans-sur-Bex, 1980)*, 1–34, Lecture Notes in Math., 844, Springer, Berlin. MR0611863 (82h:16013)

[22] A. Weil, Algebras with involutions and the classical groups, J. Indian Math. Soc. (N.S.) 24 (1960), 589–623 (1961). MR0136682 (25 #147)

R. Preeti
Université de Rennes 1
F-35042 Rennes, France
preeti.raman@math.univ-rennes1.fr

J.-P. Tignol
Université catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium
tignol@math.ucl.ac.be

Erratum for "Tropical Convexity"

Mike Develin and Bernd Sturmfels

Received: July 10, 2004

Communicated by Günter Ziegler

Abstract. Theorem 29 and Corollary 30 of [1] are incorrect. This only concerns the application of tropical convexity to phylogenetic trees. None of the results on tropical convexity itself is affected.

2000 Mathematics Subject Classification: 52A30; 92B10

Theorem 29 and Corollary 30 of [1] are not correct. If D is a symmetric matrix which represents a finite metric, then the tropical polytope P_D always contains Isbell's injective hull of the metric, but in general these two polyhedral spaces are not equal. The flaw lies in the statement (made in the proof of Theorem 29) that for any vertex (y, z) of P_{-D}, the vector y is a column of $-D$. The tropical polytope P_{-D} can have vertices for which this is not the case, corresponding to vertices in the tropical convex hull which are not in the generating set. The injective hull of D is the intersection of P_{-D} with the linear space $\{y = z\}$. If the metric D is a tree metric, then the tropical polytope P_D is one-dimensional and is indeed equal to the given tree, so Theorem 28 is correct as stated.

For instance, for a generic metric D on four points, the tropical tetrahedron P_D given by the tropical convex hull of the negated columns of the matrix is three-dimensional, while the injective hull is a two-dimensional complex consisting of four edges emanating from a quadrangle [2, Figure A3]. Even in this case, there does not seem be a straightforward relationship between the combinatorial structure of the tropical tetrahedron P_D and that of the injective hull of D.

While the connection between metrics and regular subdivisions of products of simplices via tropical convex hulls is invalid, metrics are intimately related to subdivisions of other polytopes, namely hypersimplices, as shown in [3].

References

[1] M. Develin and B. Sturmfels: Tropical Convexity, *Documenta Math.* 9 (2004) 1–27.

[2] A. Dress: Trees, Tight Extensions of Metric Spaces, and the Cohomological Dimension of Certain Groups, *Advances in Mathematics* 53 (1984) 321–402.

[3] B. Sturmfels and J. Yu: Classification of Six-Point Metrics, *Electronic Journal of Combinatorics* 11 (2004/05) R44.

Mike Develin
American Inst. of Math
360 Portage Ave.
Palo Alto, CA 94306-2244
USA
develin@post.harvard.edu

Bernd Sturmfels
Dept. of Mathematics
UC-Berkeley
Berkeley, CA 94720
USA
bernd@math.berkeley.edu

Asymptotic Expansions for Bounded Solutions to Semilinear Fuchsian Equations

Xiaochun Liu and Ingo Witt

Received: Jun 12, 2002
Revised: July 15, 2004

Communicated by Bernold Fielder

Abstract. It is shown that bounded solutions to semilinear elliptic Fuchsian equations obey complete asymptotic expansions in terms of powers and logarithms in the distance to the boundary. For that purpose, Schulze's notion of asymptotic type for conormal asymptotic expansions near a conical point is refined. This in turn allows to perform explicit computations on asymptotic types — modulo the resolution of the spectral problem for determining the singular exponents in the asymptotic expansions.

2000 Mathematics Subject Classification: Primary: 35J70; Secondary: 35B40, 35J60
Keywords and Phrases: Calculus of conormal symbols, conormal asymptotic expansions, discrete asymptotic types, weighted Sobolev spaces with discrete asymptotics, semilinear Fuchsian equations

Contents

1 Introduction

In this paper, we study solutions $u = u(x)$ to semilinear elliptic equations of the form

$$Au = F(x, B_1u, \dots, B_Ku) \quad \text{on } X^\circ = X \setminus \partial X. \tag{1.1}$$

Here, X is a smooth compact manifold with boundary, ∂X, and of dimension $n+1$, A, $B_1, \dots, B_K$ are Fuchsian differential operators on X°, see Definition 2.1, with real-valued coefficients and of orders μ, $\mu_1, \dots, \mu_K$, respectively, where $\mu_J < \mu$ for $1 \le J \le K$, and $F = F(x, \nu)\colon X^\circ \times \mathbb{R}^K \to \mathbb{R}$ is a smooth function subject to further conditions as $x \to \partial X$. In case A is elliptic in the sense of Definition 2.2 (a) we shall prove that bounded solutions $u\colon X^\circ \to \mathbb{R}$ to Eq. (1.1) possess complete conormal asymptotic expansion of the form

$$u(t,y) \sim \sum_{j=0}^{\infty} \sum_{k=0}^{m_j} t^{-p_j} \log^k t \, c_{jk}(y) \ \text{ as } t \to +0. \tag{1.2}$$

Here, $(t, y) \in [0, 1) \times Y$ are normal coordinates in a neighborhood $\mathcal{U}$ of ∂X, Y is diffeomorphic to ∂X, and the exponents $p_j \in \mathbb{C}$ appear in conjugated pairs, $\operatorname{Re} p_j \to -\infty$ as $j \to \infty$, $m_j \in \mathbb{N}$, and $c_{jk}(y) \in C^\infty(Y)$. Note that such conormal asymptotic expansions are typical of solutions u to linear equations of the form (1.1), i.e., in case $F(x) = F(x, \nu)$ is independent of $\nu \in \mathbb{R}^K$.

The general form (1.2) of asymptotics was first thoroughly investigated by Kondrat'ev in his nowadays classical paper [9]. After that to assign asymptotic types to conormal asymptotic expansions of the form (1.2) has proved to be very fruitful. In its consequence, it provides a functional-analytic framework for treating singular problems, both linear and non-linear ones, of the kind (1.1). Function spaces with asymptotics will be discussed in Sections 2.4, 3.1. In its standard setting, going back to Rempel–Schulze [14] in case $n = 0$ (when Y is always assumed be a point) and Schulze [15] in the general case, an asymptotic type P for conormal asymptotic expansions of the form (1.2) is given by a sequence $\{(p_j, m_j, L_j)\}_{j=0}^{\infty}$, where $p_j \in \mathbb{C}$, $m_j \in \mathbb{N}$ are as in (1.2), and L_j is a finite-dimensional linear subspace of $C^\infty(Y)$ to which the coefficients $c_{jk}(y)$ for $0 \le k \le m_j$ are required to belong. (In case $n = 0$, the spaces $L_j = \mathbb{C}$ disappear.) A function $u(x)$ is said to have conormal asymptotics of type P as $x \to \partial X$ if $u(x)$ obeys a conormal asymptotic expansion of the form (1.2), with the data given by P.

When treating semilinear equations we shall encounter asymptotic types belonging to bounded functions $u(x)$, i.e., asymptotic types P for which

$$\begin{cases} p_0 = 0,\, m_0 = 0,\, L_0 = \operatorname{span}\{1\}, \\ \operatorname{Re} p_j < 0 \text{ for all } j \geq 1, \end{cases} \tag{1.3}$$

where $1 \in L_0$ is the function on Y being constant 1.
It turns out that this notion of asymptotic type resolves asymptotics not fine enough to suit a treatment of semilinear problems. The difficulty with it is that only the aspect of the production of asymptotics is emphasized — via the finite-dimensionality of the spaces L_j — but not the aspect of their annihilation. For semilinear problems, however, the latter affair becomes crucial. Therefore, in Section 2, we shall introduce a refined notion of asymptotic type, where additionally linear relations between the various coefficients $c_{jk}(y) \in L_j$, even for different j, are taken into account.

Let $\underline{\mathrm{As}}(Y)$ be the set of all these refined asymptotic types, while $\underline{\mathrm{As}}^\sharp(Y) \subset \underline{\mathrm{As}}(Y)$ denotes the set of asymptotic types belonging to bounded functions according to (1.3). For $R \in \underline{\mathrm{As}}(Y)$, let $C_R^\infty(X)$ be the space of smooth functions $u \in C^\infty(X^\circ)$ having conormal asymptotic expansions of type R, and $C_R^\infty(X \times \mathbb{R}^K) = C^\infty(\mathbb{R}^K; C_R^\infty(X))$, where $C_R^\infty(X)$ is equipped with its natural (nuclear) Fréchet topology. In the formulation of Theorem 1.1, below, we will assume that $F \in C_R^\infty(X \times \mathbb{R}^K)$, where

$$\omega(t) t^{\mu - \bar{\mu} - \varepsilon} C_R^\infty(X) \subset L^\infty(X) \tag{1.4}$$

for some $\varepsilon > 0$. Here, $\bar{\mu} = \max_{1 \leq J \leq K} \mu_J < \mu$ and $\omega = \omega(t)$ is a cut-off function supported in $\mathcal{U}$, i.e., $\omega \in C^\infty(X)$, $\operatorname{supp} \omega \Subset \mathcal{U}$. Here and in the sequel, we always assume that $\omega = \omega(t)$ depends only on t for $0 < t < 1$ and $\omega(t) = 1$ for $0 < t \leq 1/2$. Condition (1.4) means that, given the operator A and then compared to the operators $B_1, \dots, B_K$, functions in $C_R^\infty(X)$ cannot be too singular as $t \to +0$.
There is a small difference between the set $\underline{\mathrm{As}}^b(Y)$ of all bounded asymptotic types and the set $\underline{\mathrm{As}}^\sharp(Y)$ of asymptotic types as described by (1.3); $\underline{\mathrm{As}}^\sharp(Y) \subsetneq \underline{\mathrm{As}}^b(Y)$. The set $\underline{\mathrm{As}}^\sharp(Y)$ actually appears as the set of multiplicatively closable asymptotic types, see Lemma 3.4. This shows up in the fact that when only boundedness is presumed asymptotic types belonging to $\underline{\mathrm{As}}^b(Y)$ — but not to $\underline{\mathrm{As}}^\sharp(Y)$ — need to be excluded from the considerations by the following non-resonance type condition (1.5), below:
Let $\mathcal{H}^{-\infty,\delta}(X) = \bigcup_{s \in \mathbb{R}} \mathcal{H}^{s,\delta}(X)$ for $\delta \in \mathbb{R}$ be the space of distributions $u = u(x)$ on X° having conormal order at least δ. (The weighted Sobolev space $\mathcal{H}^{s,\delta}(X)$, where $s \in \mathbb{R}$ is Sobolev regularity, is introduced in (2.31).) Note that $\bigcup_{\delta \in \mathbb{R}} \mathcal{H}^{-\infty,\delta}(X)$ is the space of all extendable distributions on X° that in turn is dual to the space $C_{\mathcal{O}}^\infty(X)$ of all smooth functions on X vanishing to infinite order at ∂X. Note also that the conormal order δ for $\delta \to \infty$ is the parameter in which the asymptotics (1.2) are understood.

Now, fix $\delta \in \mathbb{R}$ and suppose that a real-valued $u \in \mathcal{H}^{-\infty,\delta}(X)$ satisfying $Au \in C^\infty_{\mathcal{O}}(X)$ has an asymptotic expansion of the form

$$u(x) \sim \operatorname{Re}\left(\sum_{j=0}^{\infty}\sum_{k=0}^{m_j} t^{l+j+i\beta}\log^k t\, c_{jk}(y)\right) \quad \text{as } t \to +0,$$

where $l \in \mathbb{Z}$, $\beta \in \mathbb{R}$, $\beta \neq 0$ (and $l > \delta - 1/2$ provided that $c_{0m_0}(y) \not\equiv 0$ due to the assumption $u \in \mathcal{H}^{-\infty,\delta}(X)$). Then, for each $1 \leq J \leq K$, it is additional required that

$$B_J u = O(1) \text{ as } t \to +0 \text{ implies } B_J u = o(1) \text{ as } t \to +0, \tag{1.5}$$

where O and o are Landau's symbols. Condition (1.5) means that there is no real-valued $u \in \mathcal{H}^{-\infty,\delta}(X)$ with $Au \in C^\infty_{\mathcal{O}}(X)$ such that $B_J u$ admits an asymptotic series starting with the term $\operatorname{Re}(t^{i\beta}d(y))$ for some $\beta \in \mathbb{R} \setminus \{0\}$, $d(y) \in C^\infty(Y)$. This condition is void if $\delta \geq 1/2 + \bar{\mu}$.
Our main theorem states:

Theorem 1.1. *Let $\delta \in \mathbb{R}$ and $A \in \mathrm{Diff}^{\mu}_{\mathrm{Fuchs}}(X)$ be elliptic in the sense of* Definition 2.2 (a)*, $B_J \in \mathrm{Diff}^{\mu_J}_{\mathrm{Fuchs}}(X)$ for $1 \leq J \leq K$, where $\mu_J < \mu$, and $F \in C^\infty_R(X \times \mathbb{R}^k)$ for some asymptotic type $R \in \underline{\mathrm{As}}(Y)$ satisfying* (1.4)*. Further, let the non-resonance type condition* (1.5) *be satisfied. Then there exists an asymptotic type $P \in \underline{\mathrm{As}}(Y)$ expressible in terms of A, $B_1, \dots, B_K$, R, and δ such that each solution $u \in \mathcal{H}^{-\infty,\delta}(X)$ to* Eq. (1.1) *satisfying $B_J u \in L^\infty(X)$ for $1 \leq J \leq K$ belongs to the space $C^\infty_P(X)$.*

Under the conditions of Theorem 1.1, interior elliptic regularity already implies $u \in C^\infty(X^\circ)$. Thus, the statement concerns the fact that u possesses a complete conormal asymptotic expansion of type P near ∂X. Furthermore, the asymptotic type P can at least in principle be calculated once A, $B_1, \dots, B_K$, R, and δ are known.
Some remarks about Theorem 1.1 are in order: First, the solution u is asked to belong to the space $\mathcal{H}^{-\infty,\delta}(X)$. Thus, if the non-resonance type condition (1.5) is satisfied for all $\delta \in \mathbb{R}$ — which is generically true — then the foregoing requirement can be replaced by the requirement for u being an extendable distribution. In this case, $P_\delta \preccurlyeq P_{\delta'}$ for $\delta \geq \delta'$ in the natural ordering of asymptotic types, where P_δ denotes the asymptotic type associated with the conormal order δ. Moreover, jumps in this relation occur only for a discrete set of values of $\delta \in \mathbb{R}$ and, generically, P_δ eventually stabilizes as $\delta \to -\infty$.
Secondly, for a solution $u \in C^\infty_P(X)$ to Eq. (1.1), neither u nor the right-hand side $F(x, B_1 u(x), \dots, B_K u(x))$ need be bounded. Unboundedness of u, however, requires that, up to a certain extent, asymptotics governed by the elliptic operator A are canceled jointly by the operators $B_1, \dots, B_K$. Again, this is a non-generic situation. Furthermore, in applications one often has that one of the operators B_J, say B_1, is the identity — belonging to $\mathrm{Diff}^0_{\mathrm{Fuchs}}(X)$ — i.e., $B_1 u = u$ for all u. Then this leads to $u \in L^\infty(X)$ and explains the term "bounded solutions" in the paper's title.

Remark 1.2. Theorem 1.1 continues to hold for sectional solutions in vector bundles over X. Let E_0, E_1, E_2 be smooth vector bundles over X, $A \in \mathrm{Diff}^{\mu}_{\mathrm{Fuchs}}(X; E_0, E_1)$ be elliptic in the above sense, $B \in \mathrm{Diff}^{\mu-1}_{\mathrm{Fuchs}}(X; E_0, E_2)$, and $F \in C^{\infty}_{R}(X, E_2; E_1)$. Then, under the same technical assumptions as above, each solution u to $Au = F(x, Bu)$ in the class of extendable distributions with $Bu \in L^{\infty}(X; E_2)$ belongs to the space $C^{\infty}_{P}(X; E_0)$ for some resulting asymptotic type P.

Theorem 1.1 has actually been stated as one, though basic example for a more general method for deriving — and then justifying — conormal asymptotic expansions for solutions to semilinear elliptic Fuchsian equations. This method always works if one has boundedness assumptions as made above, but boundedness can often successfully be replaced by *structural assumptions* on the nonlinearity. An example is provided in Section 3.4. The proposed method works indeed not only for elliptic Fuchsian equations, but for other Fuchsian equations as well. In technical terms, what counts is the invertible of the complete sequence of conormal symbols in the algebra of complete Mellin symbols under the Mellin translation product, and this is equivalent to the ellipticity of the principal conormal symbol (which, in fact, is a substitute for the non-characteristic boundary in boundary problems). For elliptic Fuchsian differential operator, this latter condition is always fulfilled.
The derivation of conormal asymptotic expansions for solutions to semilinear Fuchsian equations is a purely algebraic business once the singular exponents and their multiplicities for the linear part are known. However, a strict justification of these conormal asymptotic expansions — in the generality supplied in this paper — requires the introduction of the refined notion of asymptotic type and corresponding function spaces with asymptotics. For this reason, from a technical point of view the main result of this paper is Theorem 2.42 which states the existence of a complete sequence of holomorphic Mellin symbols realizing a given proper asymptotic type in the sense of exactly annihilating asymptotics of that given type. (The term "proper" is introduced in Definition 2.22.) The construction of such Mellin symbols relies on the factorization result of Witt [21].

Remark 1.3. Behind part of the linear theory, there is Schulze's cone pseudodifferential calculus. The interested reader should consult Schulze [15, 16]. We do not go much into the details, since for most of the arguments this is not needed. Indeed, the algebra of complete Mellin symbols controls the production and annihilation of asymptotics, and it is this algebra that is detailed discussed.
The relation with conical points is as follows: A conical point leads — via blow-up, i.e., the introduction of polar coordinates — to a manifold with boundary. Vice versa, each manifold with boundary gives rise to a space with a conical point — via shrinking the boundary to a point. Since in both situations the analysis is taken place over the interior of the underlying configuration, i.e., away from the conical point and the boundary, respectively, there is no essential

difference between these two situations. Thus, the geometric situation is given by the kind of degeneracy admitted for, say, differential operators. In the case considered in this paper, this degeneracy is of Fuchsian type.

The first part of this paper, Section 2, is devoted to the linear theory and the introduction of the refined notion of asymptotic type. Then, in a second part, Theorem 1.1 is proved in Section 3.

2 Asymptotic types

In this section, we introduce the notion of discrete asymptotic type. A comparison of this notion with the formerly known notions of weakly discrete asymptotic type and strongly discrete asymptotic type, respectively, can be found in Figure 1. The definition of discrete asymptotic type is modeled on part of the Gohberg-Sigal theory of the inversion of finitely meromorphic, operator-valued functions at a point, see Gohberg-Sigal [4]. See also Witt [18] for the corresponding notion of local asymptotic type, i.e., asymptotic types at one singular exponent $p \in \mathbb{C}$ in (1.2) only. Finally, in Section 2.4, function spaces with asymptotics are introduced. The definition of these function spaces relies on the existence of complete (holomorphic) Mellin symbols realizing a prescribed proper asymptotic type. The existence of such complete Mellin symbols is stated and proved in Theorem 2.42.

Added in proof. To keep this article of reasonable length, following the referee's advice, proofs of Theorems 2.6, 2.30, and 2.42 and Propositions 2.28 (b), 2.31, 2.32, 2.35, 2.36, 2.40, 2.44, 2.46, 2.47, 2.48, 2.49, and 2.52 are only sketchy or missing at all. They are available from the second author's homepage[1].

2.1 Fuchsian differential operators

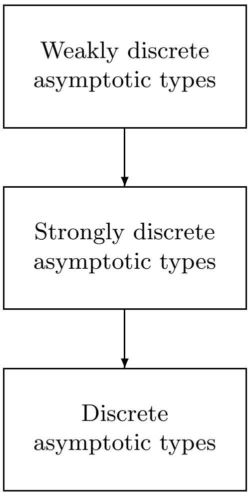
Let X be a compact C^∞ manifold with boundary, ∂X. Throughout, we fix a collar neighborhood $\mathcal{U}$ of ∂X and a diffeomorphism $\chi\colon \mathcal{U} \to [0,1) \times Y$, with Y being a closed C^∞ manifold diffeomorphic to ∂X. Hence, we work in a fixed splitting of coordinates (t,y) on $\mathcal{U}$, where $t \in [0,1)$ and $y \in Y$. Let (τ,η) be the covariables to (t,y). The compressed covariable $t\tau$ to t is denoted by $\tilde{\tau}$, i.e., $(\tilde{\tau},\eta)$ is the linear variable in the fiber of the compressed cotangent bundle $\tilde{T}^*X\big|_{\mathcal{U}}$. Finally, let $\dim X = n+1$.

Definition 2.1. A differential operator A with smooth coefficients of order μ on $X^\circ = X \setminus \partial X$ is called *Fuchsian* if

$$\chi_*\left(A\big|_{\mathcal{U}}\right) = t^{-\mu} \sum_{k=0}^{\mu} a_k(t,y,D_y)\big(-t\partial_t\big)^k, \tag{2.1}$$

where $a_k \in C^\infty([0,1); \mathrm{Diff}^{\mu-k}(Y))$ for $0 \le k \le \mu$. The class of all Fuchsian differential operators of order μ on X° is denoted by $\mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$.

[1] `http://www.ma.imperial.ac.uk/~ifw/asymptotics.html`



Singular exponents with multiplicities, (p_j, m_j), are prescribed, the coefficients $c_{jk}(y) \in C^\infty(Y)$ are arbitrary. The general form of asymptotics is observed, cf., e.g., KONDRAT'EV (1967), MELROSE (1993), SCHULZE (1998).

Singular exponents with multiplicities, (p_j, m_j), are prescribed, $c_{jk}(y) \in L_j \subset C^\infty(Y)$, where $\dim L_j < \infty$. The *production of asymptotics* is observed, cf. REMPEL–SCHULZE (1989), SCHULZE (1991).

Linear relation between the various coefficients $c_{jk}(y) \in L_j$, even for different j, are additionally allowed. Thus the *production/annihilation of asymptotics* is observed, cf. this article.

Figure 1: Schematic overview of asymptotic types

Henceforth, we shall suppress writing the restriction $\cdot\big|_{\mathcal{U}}$ and the operator push-forward χ_* in expressions like (2.1). For $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$, we denote by

$$\sigma^\mu_\psi(A)(t,y,\tau,\eta) = t^{-\mu} \sum_{k=0}^{\mu} \sigma^{\mu-k}_\psi(a_k(t))(y,\eta)(it\tau)^k$$

the principal symbol of A, by $\tilde{\sigma}^\mu_\psi(A)(t,y,\tilde{\tau},\eta)$ its compressed principal symbol related to $\sigma^\mu_\psi(A)(t,y,\tau,\eta)$ via

$$\sigma^\mu_\psi(A)(t,y,\tau,\eta) = t^{-\mu}\tilde{\sigma}^\mu_\psi(A)(t,y,t\tau,\eta)$$

in $(\tilde{T}^*X \setminus 0)\big|_{\mathcal{U}}$, and by $\sigma^\mu_M(A)(z)$ its *principal conormal symbol*,

$$\sigma^\mu_M(A)(z) = \sum_{k=0}^{\mu} a_k(0) z^k, \quad z \in \mathbb{C}.$$

Further, we introduce the *jth conormal symbol* $\sigma^{\mu-j}_M(A)(z)$ for $j = 1, 2, \dots$ by

$$\sigma^{\mu-j}_M(A)(z) = \sum_{k=0}^{\mu} \frac{1}{j!} \frac{\partial^j a_k}{\partial t^j}(0) z^k, \quad z \in \mathbb{C}.$$

Note that $\tilde{\sigma}^\mu_\psi(A)(t,y,\tilde{\tau},\eta)$ is smooth up to $t = 0$ and that $\sigma^{\mu-j}_M(z)$ for $j = 0, 1, 2, \dots$ is a holomorphic function in z taking values in $\mathrm{Diff}^\mu(Y)$. Moreover, if $A \in \mathrm{Diff}^{\mu}_{\mathrm{Fuchs}}(X)$, $B \in \mathrm{Diff}^{\nu}_{\mathrm{Fuchs}}(X)$, then $AB \in \mathrm{Diff}^{\mu+\nu}_{\mathrm{Fuchs}}(X)$,

$$\sigma_M^{\mu+\nu-l}(AB)(z) = \sum_{j+k=l} \sigma_M^{\mu-j}(A)(z+\nu-k)\sigma_M^{\nu-k}(B)(z) \tag{2.2}$$

for all $l = 0, 1, 2, \dots$ This formula is called the *Mellin translation product* (due to the shifts of $\nu - k$ in the argument of the first factors).

Definition 2.2. (a) The operator $A \in \mathrm{Diff}^{\mu}_{\mathrm{Fuchs}}(X)$ is called *elliptic* if A is an elliptic differential operator on X° and

$$\tilde{\sigma}^{\mu}_{\psi}(A)(t, y, \tilde{\tau}, \eta) \neq 0, \quad (t, y, \tilde{\tau}, \eta) \in (\tilde{T}^*X \setminus 0)\big|_{\mathcal{U}}. \tag{2.3}$$

(b) The operator $A \in \mathrm{Diff}^{\mu}_{\mathrm{Fuchs}}(X)$ is called *elliptic with respect to the weight* $\delta \in \mathbb{R}$ if A is elliptic in the sense of (a) and, in addition,

$$\sigma^{\mu}_{M}(A)(z)\colon H^s(Y) \to H^{s-\mu}(Y), \quad z \in \Gamma_{(n+1)/2-\delta}, \tag{2.4}$$

is invertible for some $s \in \mathbb{R}$ (and then for all $s \in \mathbb{R}$). Here, $\Gamma_\beta = \{z \in \mathbb{C};\ \mathrm{Re}\, z = \beta\}$ for $\beta \in \mathbb{R}$.

Under the assumption of interior ellipticity of A, (2.3) can be reformulated as

$$\sum_{k=0}^{\mu} \sigma^{\mu-k}_{\psi}(a_k(0))(y,\eta)\big(i\tilde{\tau}\big)^k \neq 0$$

for all $(0, y, \tilde{\tau}, \eta) \in (\tilde{T}^*X \setminus 0)\big|_{\partial\mathcal{U}}$. This relation implies that $\sigma^{\mu}_{M}(A)(z)\big|_{\Gamma_{(n+1)/2-\delta}}$ is parameter-dependent elliptic as an element in $L^{\mu}_{\mathrm{cl}}\big(Y; \Gamma_{(n+1)/2-\delta}\big)$, where the latter is the space of *classical pseudodifferential operators* on Y of order μ *with parameter* z varying in $\Gamma_{(n+1)/2-\delta}$, for

$$\sigma^{\mu}_{\psi}(\sigma^{\mu}_{M}(A))(y, z, \eta)\big|_{z=(n+1)/2-\delta-\tilde{\tau}} = \tilde{\sigma}^{\mu}_{\psi}(A)(0, y, \tilde{\tau}, \eta),$$

where $\sigma^{\mu}_{\psi}(\cdot)$ on the left-hand side denotes the parameter-dependent principal symbol. Thus, if (a) is fulfilled, then it follows that $\sigma^{\mu}_{M}(A)(z)$ in (2.4) is invertible for $z \in \Gamma_{(n+1)/2-\delta}$, $|z|$ large enough.

Lemma 2.3. *If $A \in \mathrm{Diff}^{\mu}_{\mathrm{Fuchs}}(X)$ is elliptic, then there exists a discrete set $\mathcal{D} \subset \mathbb{C}$ with $\mathcal{D} \cap \{z \in \mathbb{C};\ c_0 \leq \mathrm{Re}\, z \leq c_1\}$ is finite for all $-\infty < c_0 < c_1 < \infty$ such that* (2.4) *is invertible for all $z \in \mathbb{C} \setminus \mathcal{D}$. In particular, there is a discrete set $D \subset \mathbb{R}$ such that A is elliptic with respect to the weight δ for all $\delta \in \mathbb{R} \setminus D$; $D = \mathrm{Re}\, \mathcal{D}$.*

Proof. Since $\sigma^{\mu}_{M}(A)(z)\big|_{\Gamma_\beta} \in L^{\mu}(Y; \Gamma_\beta)$ is parameter-dependent elliptic for all $\beta \in \mathbb{R}$, for each $c > 0$ there is a $C > 0$ such that $\sigma^{\mu}_{M}(A)(z) \in L^{\mu}(Y)$ is invertible for all z with $|\mathrm{Re}\, z| \leq c$, $|\mathrm{Im}\, z| \geq C$. Then the assertion follows from results on the invertibility of holomorphic operator-valued functions. See Proposition 2.5, below, or Schulze [16, Theorem 2.4.20]. □

Next, we introduce the class of meromorphic functions arising in point-wise inverting parameter-dependent elliptic conormal symbols $\sigma_M^\mu(A)(z)$. The following definition is taken from Schulze [16, Definition 2.3.48]:

Definition 2.4. (a) $\mathcal{M}_{\mathcal{O}}^\mu(Y)$ for $\mu \in \mathbb{Z} \cup \{-\infty\}$ is the space of all holomorphic functions $f(z)$ on $\mathbb{C}$ taking values in $L_{\mathrm{cl}}^\mu(Y)$ such that $f(z)\big|_{z=\beta+i\tau} \in L_{\mathrm{cl}}^\mu(Y;\mathbb{R}_\tau)$ uniformly in $\beta \in [\beta_0, \beta_1]$ for all $-\infty < \beta_0 < \beta_1 < \infty$.
(b) $\mathcal{M}_{\mathrm{as}}^{-\infty}(Y)$ is the space of all meromorphic functions $f(z)$ on $\mathbb{C}$ taking values in $L^{-\infty}(Y)$ that satisfy the following conditions:
(i) The Laurent expansion around each pole $z = p$ of $f(z)$ has the form

$$f(z) = \frac{f_0}{(z-p)^\nu} + \frac{f_1}{(z-p)^{\nu-1}} + \cdots + \frac{f_{\nu-1}}{z-p} + \sum_{j\geq 0} f_{\nu+j}(z-p)^j, \tag{2.5}$$

where $f_0, f_1, \dots, f_{\nu-1} \in L^{-\infty}(Y)$ are finite-rank operators.
(ii) If the poles of $f(z)$ are numbered someway, $p_1, p_2, \dots$, then $|\operatorname{Re} p_j| \to \infty$ as $j \to \infty$ if the number of poles is infinite.
(iii) For any $\bigcup_j\{p_j\}$-excision function $\chi(z) \in C^\infty(\mathbb{C})$, i.e., $\chi(z) = 0$ if $\operatorname{dist}(z, \bigcup_j\{p_j\}) \leq 1/2$ and $\chi(z) = 1$ if $\operatorname{dist}(z, \bigcup_j\{p_j\}) \geq 1$, we have $\chi(z)f(z)\big|_{z=\beta+i\tau} \in L^{-\infty}(Y;\mathbb{R}_\tau)$ uniformly in $\beta \in [\beta_0, \beta_1]$ for all $-\infty < \beta_0 < \beta_1 < \infty$.
(c) Finally, we set $\mathcal{M}_{\mathrm{as}}^\mu(Y) = \mathcal{M}_{\mathcal{O}}^\mu(Y) + \mathcal{M}_{\mathrm{as}}^{-\infty}(Y)$ for $\mu \in \mathbb{Z}$. (Note that $\mathcal{M}_{\mathcal{O}}^\mu(Y) \cap \mathcal{M}_{\mathrm{as}}^{-\infty}(Y) = \mathcal{M}_{\mathcal{O}}^{-\infty}(Y)$.)
Functions $f(z)$ belonging to $\mathcal{M}_{\mathrm{as}}^\mu(Y)$ are called *Mellin symbols* of order μ.

$\bigcup_{\mu\in\mathbb{Z}} \mathcal{M}_{\mathrm{as}}^\mu(Y)$ is a filtered algebra under pointwise multiplication.
For $f \in \mathcal{M}_{\mathrm{as}}^\mu(Y)$ for $\mu \in \mathbb{Z}$ and $f(z) = f_0(z) + f_1(z)$, where $f_0 \in \mathcal{M}_{\mathcal{O}}^\mu(Y)$, $f_1 \in \mathcal{M}_{\mathrm{as}}^{-\infty}(Y)$, the parameter-dependent principal symbol $\sigma_\psi^\mu\big(f_0(z)\big|_{z=\beta+i\tau}\big)$ is independent of the choice of the decomposition of f and also independent of $\beta \in \mathbb{R}$. It is called the *principal symbol* of f. The Mellin symbol $f \in \mathcal{M}_{\mathrm{as}}^\mu(Y)$ is called *elliptic* if its principal symbol is everywhere invertible.
For the next result, see Schulze [16, Theorem 2.4.20]:

Proposition 2.5. *The Mellin symbol $f \in \mathcal{M}_{\mathrm{as}}^\mu(Y)$ for $\mu \in \mathbb{Z}$ is invertible in the filtered algebra $\bigcup_{\mu\in\mathbb{Z}} \mathcal{M}_{\mathrm{as}}^\mu(Y)$, i.e., there is a $g \in \mathcal{M}_{\mathrm{as}}^{-\mu}(Y)$ such that $(fg)(z) = (gf)(z) = 1$ on $\mathbb{C}$, if and only if f is elliptic.*

For $f \in \mathcal{M}_{\mathrm{as}}^\mu(Y)$, $p \in \mathbb{C}$, and $N \in \mathbb{N}$, we denote by $[f(z)]_p^N$ the Laurent series of $f(z)$ around $z = p$ truncated after the term containing $(z-p)^N$, i.e.,

$$[f(z)]_p^N = \frac{f_{-\nu}}{(z-p)^\nu} + \cdots + \frac{f_{-1}}{z-p} + f_\nu + f_1(z-p) + \cdots + f_N(z-p)^N. \tag{2.6}$$

Furthermore, $[f(z)]_p^* = [f(z)]_p^{-1}$ denotes the principal part of the Laurent series of $f(z)$ around $z = p$.
In various constructions, it is important to have examples of elliptic Mellin symbols $f \in \mathcal{M}_{\mathrm{as}}^\mu(Y)$ of controlled *singularity structure*:

Theorem 2.6. *Let* $\mu \in \mathbb{Z}$ *and* $\{p_j\}_{j=1,2,\dots} \subset \mathbb{C}$ *be a sequence obeying the property mentioned in* Definition 2.4 (b) (ii). *Let, for each* $j = 1, 2, \dots,$ *operators* $f^j_{-\nu_j}, \dots, f^j_{N_j}$ *in* $L^\mu_{\mathrm{cl}}(Y)$*, where* $\nu_j \geq 0$*,* $N_j + \nu_j \geq 0$*, be given such that*

- $f^j_{-\nu_j}, \dots, f^j_{\min\{N_j,0\}} \in L^{-\infty}(Y)$ *are finite-rank operators,*
- *there is an elliptic* $g \in \mathcal{M}^\mu_{\mathcal{O}}(Y)$ *such that, for all* j*,* $0 \leq k \leq N_j$*,*

$$f^j_k - \frac{1}{k!}\, g^{(k)}(p_j) \in L^{-\infty}(Y) \tag{2.7}$$

(*in particular,* $f^j_k \in L^{\mu-k}_{\mathrm{cl}}(Y)$ *for* $0 \leq k \leq N_j$ *and* $f^j_0 \in L^\mu_{\mathrm{cl}}(Y)$ *is elliptic of index zero*).

Then there is an elliptic Mellin symbol $f(z) \in \mathcal{M}^\mu_{\mathrm{as}}(Y)$ *such that, for all* j*,*

$$[f(z)]^{N_j}_{p_j} = \frac{f^j_{-\nu_j}}{(z-p_j)^{\nu_j}} + \dots + \frac{f^j_{-1}}{z-p_j} + f^j_0 + \dots + f^j_{N_j}(z-p_j)^{N_j}, \tag{2.8}$$

while $f(q) \in L^\mu_{\mathrm{cl}}(Y)$ *is* invertible for all $q \in \mathbb{C} \setminus \bigcup_{j=1,2,\dots}\{p_j\}$.

If $n = 0$, condition (2.7) is void. In case $n > 0$, however, this condition expresses several compatibility conditions among the $\sigma^{\mu-l}_\psi(f^j_k)$, where $j = 0, 1, 2, \dots,$ $0 \leq k \leq N_j$, and $l \geq k$, and also certain topological obstructions that must be fulfilled. For instance, for any $f \in \mathcal{M}^\mu_{\mathcal{O}}(Y)$,

$$\sigma^{\mu-j}_\psi(f(z))(y,\eta) = \sum_{k=0}^{j} \frac{(z-p)^k}{k!}\, \sigma^{\mu-j}_\psi(f^{(k)}(p))(y,\eta), \quad j = 0, 1, 2, \dots$$

in local coordinates (y, η) — showing, among others, that $\sigma^{\mu-j}_\psi(f(z))$ is polynomial of degree j with respect to $z \in \mathbb{C}$. The point is that we do not assume $g(q) \in L^\mu_{\mathrm{cl}}(Y)$ be invertible for $q \in \mathbb{C} \setminus \bigcup_{j=1,2,\dots}\{p_j\}$.

Proof of Theorem 2.6. This can be proved using the results of Witt [21]. In particular, the factorization result there gives directly the existence of $f(z)$ if the sequence $\{p_j\} \subset \mathbb{C}$ is void. ∎

Now, we are going to introduce the basic object of study — the *algebra of complete conormal symbols*. This algebra will enable us to introduce the refined notion of asymptotic type and to study the behavior of conormal asymptotics under the action of Fuchsian differential operators.

Definition 2.7. (a) For $\mu \in \mathbb{Z}$, the space $\mathrm{Symb}^\mu_M(Y)$ consists of all sequences $\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j}(z);\, j \in \mathbb{N}\} \subset \mathcal{M}^\mu_{\mathrm{as}}(Y)$.
(b) An element $\mathfrak{S}^\mu \in \mathrm{Symb}^\mu_M(Y)$ is called holomorphic if $\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j}(z);\, j \in \mathbb{N}\} \subset \mathcal{M}^\mu_{\mathcal{O}}(Y)$.

(c) $\bigcup_{\mu\in\mathbb{Z}} \mathrm{Symb}^\mu_M(Y)$ is a filtered algebra under the *Mellin translation product*, denoted by $\sharp_M$. Namely, for $\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j}(z);\, j \in \mathbb{N}\} \in \mathrm{Symb}^\mu_M(Y)$, $\mathfrak{T}^\nu = \{\mathfrak{t}^{\nu-k}(z);\, k \in \mathbb{N}\} \in \mathrm{Symb}^\nu_M(Y)$, we define $\mathfrak{U}^{\mu+\nu} = \mathfrak{S}^\mu \sharp_M \mathfrak{T}^\nu \in \mathrm{Symb}^{\mu+\nu}_M(Y)$, where $\mathfrak{U}^{\mu+\nu} = \{\mathfrak{u}^{\mu+\nu-l}(z);\, l \in \mathbb{N}\}$, by

$$\mathfrak{u}^{\mu+\nu-l}(z) = \sum_{j+k=l} \mathfrak{s}^{\mu-j}(z+\nu-k)\mathfrak{t}^{\nu-k}(z) \tag{2.9}$$

for $l = 0, 1, 2, \ldots$ See also (2.2).

From Proposition 2.5, we immediately get:

Lemma 2.8. *$\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j}(z);\, j \in \mathbb{N}\} \in \mathrm{Symb}^\mu_M(Y)$ is invertible in the filtered algebra $\bigcup_{\mu\in\mathbb{Z}} \mathrm{Symb}^\mu_M(Y)$ if and only if $\mathfrak{s}^\mu(z) \in \mathcal{M}^\mu_{\mathrm{as}}(Y)$ is elliptic.*

In the case of the preceding lemma, $\mathfrak{S}^\mu \in \mathrm{Symb}^\mu_M(Y)$ is called elliptic. A holomorphic elliptic $\mathfrak{S}^\mu \in \mathrm{Symb}^\mu_M(Y)$ is called elliptic with respect to the weight $\delta \in \mathbb{R}$ if the line $\Gamma_{(n+1)/2-\delta}$ is free of poles of $\mathfrak{s}^\mu(z)^{-1}$. Notice that a holomorphic elliptic $\mathfrak{S}^\mu \in \mathrm{Symb}^\mu_M(Y)$ is elliptic for all, but a discrete set of $\delta \in \mathbb{R}$. The inverse to $\mathfrak{S}^\mu$ with respect to the Mellin translation product is denoted by $(\mathfrak{S}^\mu)^{-1}$. The set of elliptic elements of $\mathrm{Symb}^\mu_M(Y)$ is denoted by $\mathrm{Ell}\,\mathrm{Symb}^\mu_M(Y)$.

There is a homomorphism of filtered algebras,

$$\bigcup_{\mu\in\mathbb{N}} \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X) \to \bigcup_{\mu\in\mathbb{Z}} \mathrm{Symb}^\mu_M(Y), \quad A \mapsto \left\{\sigma^{\mu-j}_M(A)(z);\, j \in \mathbb{N}\right\}.$$

By the remark preceding Lemma 2.3, $\left\{\sigma^{\mu-j}_M(A)(z);\, j \in \mathbb{N}\right\} \in \mathrm{Symb}^\mu_M(Y)$ is elliptic if $A \in \mathrm{Diff}_{\mathrm{Fuchs}}(X)$ is elliptic in the sense of Definition 2.2 (a).

2.2 Definition of asymptotic types

We now start to introduce discrete asymptotic types.

2.2.1 The spaces $\mathcal{E}^\delta(Y)$ and $\mathcal{E}_V(Y)$

Here, we construct the "coefficient" space $\mathcal{E}^\delta(Y) = \bigcup_{V\in\mathcal{C}^\delta} \mathcal{E}_V(Y)$ that admits the non-canonical isomorphism (2.13), below,

$$C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X) \xrightarrow{\cong} \mathcal{E}^\delta(Y),$$

where $C^{\infty,\delta}_{\mathrm{as}}(X)$ is the space of smooth functions on X° obeying *conormal asymptotic expansions* of the form (1.2) of conormal order at least δ, i.e., $\mathrm{Re}\, p_j < (n+1)/2 - \delta$ holds for all j (with the condition that the singular exponents p_j appear in conjugated pairs dropped), and $C^\infty_{\mathcal{O}}(X)$ is the subspace of all smooth functions on X° vanishing to infinite order at ∂X.

Definition 2.9. A *carrier V of asymptotics* for distributions of conormal order δ is a discrete subset of $\mathbb{C}$ contained in the half-space $\{z \in \mathbb{C};\ \operatorname{Re} z < (n+1)/2 - \delta\}$ such that, for all β_0, $\beta_1 \in \mathbb{R}$, $\beta_0 < \beta_1$, the intersection $V \cap \{z \in \mathbb{C};\ \beta_0 < \operatorname{Re} z < \beta_1\}$ is finite. The set of all these carriers is denoted by $\mathcal{C}^\delta$.

In particular, $V_p = p - \mathbb{N}$ for $p \in \mathbb{C}$ is such a carrier of asymptotics. Note that $V_p \in \mathcal{C}^\delta$ if and only if $\operatorname{Re} p < (n+1)/2 - \delta$. We set $T^\varrho V = \varrho + V \in \mathcal{C}^{-\varrho+\delta}$ for $\varrho \in \mathbb{R}$ and $V \in \mathcal{C}^\delta$. We further set $\mathcal{C} = \bigcup_{\delta\in\mathbb{R}} \mathcal{C}^\delta$.
Let $[C^\infty(Y)]^\infty = \bigcup_{m\in\mathbb{N}} [C^\infty(Y)]^m$ be the space of all *finite sequences* in $C^\infty(Y)$, where the sequences $(\phi_0, \dots, \phi_{m-1})$ and $(\underbrace{0, \dots, 0}_{h \text{ times}}, \phi_0, \dots, \phi_{m-1})$ for $h \in \mathbb{N}$ are identified. For $V \in \mathcal{C}^\delta$, we set $\mathcal{E}_V(Y) = \prod_{p\in V} [C^\infty(Y)]^\infty_p$, where $[C^\infty(Y)]^\infty_p$ is an isomorphic copy of $[C^\infty(Y)]^\infty$, and define $\mathcal{E}^\delta(Y)$ to be the space of all families $\Phi \in \mathcal{E}_V(Y)$ for some $V \in \mathcal{C}^\delta$ depending on Φ. Thereby, $\Phi \in \mathcal{E}_V(Y)$, $\Phi' \in \mathcal{E}_{V'}(Y)$ for possibly different V, $V' \in \mathcal{C}^\delta$ are identified if $\Phi(p) = \Phi'(p)$ for $p \in V \cap V'$, while $\Phi(p) = 0$ for $p \in V \setminus V'$, $\Phi'(p) = 0$ for $p \in V' \setminus V$. Under this identification,

$$\mathcal{E}^\delta(Y) = \bigcup_{V\in\mathcal{C}^\delta} \mathcal{E}_V(Y). \tag{2.10}$$

Moreover, $\mathcal{E}_V(Y) \cap \mathcal{E}_{V'}(Y) = \mathcal{E}_{V\cap V'}(Y)$.
On $[C^\infty(Y)]^\infty$, we define the *right shift operator* T by

$$(\phi_0, \dots, \phi_{m-2}, \phi_{m-1}) \mapsto (\phi_0, \dots, \phi_{m-2}).$$

On $\mathcal{E}^\delta(Y)$, the right shift operator T acts component-wise, i.e., $(T\Phi)(p) = T(\Phi(p))$ for $\Phi \in \mathcal{E}_V(Y)$ and all $p \in V$.

Remark 2.10. To designate different shift operators with the same symbol T, once T^ϱ for $\varrho \in \mathbb{R}$ for carriers of asymptotics, once T, T^2, etc. for vectors in $\mathcal{E}^\delta(Y)$ should not confuse the reader.

For $\Phi \in \mathcal{E}^\delta(Y)$, we define $\text{c-ord}(\Phi) = (n+1)/2 - \max\{\operatorname{Re} p;\ \Phi(p) \neq 0\}$. In particular, $\text{c-ord}(0) = \infty$. Note that $\text{c-ord}(\Phi) > \delta$ if $\Phi \in \mathcal{E}^\delta(Y)$. For $\Phi_i \in \mathcal{E}^\delta(Y)$, $\alpha_i \in \mathbb{C}$ for $i = 1, 2, \dots$ satisfying $\text{c-ord}(\Phi_i) \to \infty$ as $i \to \infty$, the sum

$$\Phi = \sum_{i=1}^\infty \alpha_i \Phi_i, \tag{2.11}$$

is defined in $\mathcal{E}^\delta(Y)$ in an obvious fashion: Let $\Phi_i \in \mathcal{E}_{V_i}(Y)$, where $V_i \in \mathcal{C}^{\delta_i}$, $\delta_i \geq \delta$, and $\delta_i \to \infty$ as $i \to \infty$. Then $V = \bigcup_i V_i \in \mathcal{C}^\delta$, and $\Phi \in \mathcal{E}_V(Y)$ is defined by $\Phi(p) = \sum_{i=1}^\infty \alpha_i \Phi_i(p)$ for $p \in V$, where, for each $p \in V$, the sum on the right-hand side is finite.

Lemma 2.11. *Let $\Phi_i \in \mathcal{E}^\delta(Y)$ for $i = 1, 2, \dots$, $\text{c-ord}(\Phi_i) \to \infty$ as $i \to \infty$. Then* (2.11) *holds if and only if*

$$\text{c-ord}\Big(\Phi - \sum_{i=1}^N \alpha_i \Phi_i\Big) \to \infty \ \text{ as } N \to \infty. \tag{2.12}$$

Note that (2.12) already implies that $\text{c-ord}(\alpha_i \Phi_i) \to \infty$ as $i \to \infty$.

Definition 2.12. Let Φ_i, $i = 1, 2, \dots$, be a sequence in $\mathcal{E}^\delta(Y)$ with the property that $\text{c-ord}(\Phi_i) \to \infty$ as $i \to \infty$. Then this sequence is called *linearly independent* if, for all $\alpha_i \in \mathbb{C}$,

$$\sum_{i=1}^{\infty} \alpha_i \Phi_i = 0$$

implies that $\alpha_i = 0$ for all i. A linearly independent sequence Φ_i for $i = 1, 2, \dots$ in J for a linear subspace $J \subseteq \mathcal{E}^\delta(Y)$ is called a *basis* for J if every vector $\Phi \in J$ can be represented in the form (2.11) with certain (then uniquely determined) coefficients $\alpha_i \in \mathbb{C}$.

Note that $\sum_{i=1}^{\infty} \alpha_i \Phi_i = 0$ in $\mathcal{E}^\delta(Y)$ if and only if $\text{c-ord}(\sum_{i=1}^{N} \alpha_i \Phi_i) \to \infty$ as $N \to \infty$ according to Lemma 2.11. We also obtain:

Lemma 2.13. *Let Φ_i, $i = 1, 2, \dots$, be a sequence in $\mathcal{E}^\delta(Y)$ such that* $\text{c-ord}(\Phi_i) \to \infty$ *as $i \to \infty$. Further, let $\{\delta_j\}_{j=1}^{\infty}$ be a strictly increasing sequence such that $\delta_j > \delta$ for all j and $\delta_j \to \infty$ as $j \to \infty$. Assume that the Φ_i are numbered in such a way that* $\text{c-ord}(\Phi_i) \le \delta_j$ *if and only if $1 \le i \le e_j$. Then the sequence Φ_i, $i = 1, 2, \dots$, is linearly independent provided that, for each $j = 1, 2, \dots$,*

$$\Phi_1, \dots, \Phi_{e_j} \text{ are linearly independent over the space } \mathcal{E}^{\delta_j}(Y).$$

We now introduce the notion of characteristic basis:

Definition 2.14. Let $J \subseteq \mathcal{E}^\delta(Y)$ be a linear subspace, $TJ \subseteq J$, and Φ_i for $i = 1, 2, \dots$ be a sequence in J. Then Φ_i, $i = 1, 2, \dots$, is called a *characteristic basis* of J if there are numbers $m_i \in \mathbb{N} \cup \{\infty\}$ such that $T^{m_i}\Phi_i = 0$ if $m_i < \infty$, while the sequence $\{T^k \Phi_i;\ i = 1, 2, \dots,\ 0 \le k < m_i\}$ forms a basis for J.

Remark 2.15. This notion generalizes a notion of Witt [18]: There, given a finite-dimensional linear space J and a nilpotent operator $T\colon J \to J$, the sequence $\Phi_1, \dots, \Phi_e$ in J has been called a *characteristic basis*, of characteristic $(m_1, \dots, m_e)$, if

$$\Phi_1, T\Phi_1, \dots, T^{m_1-1}\Phi_1, \dots, \Phi_e, T\Phi_e, \dots, T^{m_e-1}\Phi_e,$$

constitutes a Jordan basis of J. The numbers $m_1, \dots, m_e$ appear as the sizes of Jordan blocks; $\dim J = m_1 + \dots + m_e$. The tuple $(m_1, \dots, m_e)$ is also called the *characteristic* of J (with respect to T), e is called the *length* of its characteristic, and $\Phi_1, \dots, \Phi_e$ is sometimes said to be a an $(m_1, \dots, m_e)$-*characteristic basis* of J. The space $\{0\}$ has empty characteristic of length $e = 0$.

The question of the existence of a characteristic basis obeying one more special property is taken up in Proposition 2.20.
We also need following notion:

Definition 2.16. $\Phi \in \mathcal{E}^\delta(Y)$ is called a *special vector* if $\Phi \in \mathcal{E}^\delta_{V_p}(Y)$ for some $p \in \mathbb{C}$.

Thus, $\Phi \in \mathcal{E}_V(Y)$ is a special vector if there is a $p \in \mathbb{C}$, $\operatorname{Re} p < (n+1)/2-\delta$ such that $\Phi(p') = 0$ for all $p' \in V$, $p' \notin p-\mathbb{N}$. Obviously, if $\Phi \neq 0$, then p is uniquely determined by Φ, by the additional requirement that $\Phi(p) \neq 0$. We denote this complex number p by $\gamma(\Phi)$. In particular, $\text{c-ord}(\Phi) = (n+1)/2 - \operatorname{Re}\gamma(\Phi)$.

2.2.2 First properties of asymptotic types

In the sequel, we fix a splitting of coordinates $\mathcal{U} \to [0,1) \times Y$, $x \mapsto (t,y)$, near ∂X. Then we have the non-canonical isomorphism

$$C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X) \xrightarrow{\cong} \mathcal{E}^\delta(Y), \tag{2.13}$$

assigning to each *formal asymptotic expansion*

$$u(x) \sim \sum_{p\in V} \sum_{k+l=m_p-1} \frac{(-1)^k}{k!}\, t^{-p} \log^k t\, \phi^{(p)}_l(y) \quad \text{as } t \to +0 \tag{2.14}$$

for some $V \in \mathcal{C}^\delta$, $m_p \in \mathbb{N}$, the vector $\Phi \in \mathcal{E}_V(Y)$ given by

$$\Phi(p) = \begin{cases} \left(\phi^{(p)}_0, \phi^{(p)}_1, \dots, \phi^{(p)}_{m_p-1}\right) & \text{if } p \in V, \\ 0 & \text{otherwise,} \end{cases}$$

see also (2.30). "*Non-canonical*" in (2.13) means that the isomorphism depends explicitly on the chosen splitting of coordinates $\mathcal{U} \to [0,1) \times Y$, $x \mapsto (t,y)$, near ∂X. Coordinate invariance is discussed in Proposition 2.32.
Note the shift from m_p to $m_p - 1$ that for notational convenience has appeared in formula (2.14) compared to formula (1.2).

Definition 2.17. An *asymptotic type*, P, for distributions as $x \to \partial X$, of conormal order at least δ, *is represented* — in the given splitting of coordinates near ∂X — *by a linear subspace* $J \subset \mathcal{E}_V(Y)$ for some $V \in \mathcal{C}^\delta$ such that the following three conditions are met:
(a) $TJ \subseteq J$.
(b) $\dim J^{\delta+j} < \infty$ for all $j \in \mathbb{N}$, where $J^{\delta+j} = J/(J \cap \mathcal{E}^{\delta+j}(Y))$.
(c) There is a sequence $\{p_j\}_{j=1}^M \subset \mathbb{C}$, where $M \in \mathbb{N}\cup\{\infty\}$, $\operatorname{Re} p_j < (n+1)/2-\delta$, and $\operatorname{Re} p_j \to -\infty$ as $j \to \infty$ if $M = \infty$, such that $V \subseteq \bigcup_{j=1}^M V_{p_j}$ and

$$J = \bigoplus_{j=1}^M \left(J \cap \mathcal{E}_{V_{p_j}}(Y)\right). \tag{2.15}$$

The *empty asymptotic type*, $\mathcal{O}$, is represented by the trivial subspace $\{0\} \subset \mathcal{E}^\delta(Y)$. The set of all asymptotic types of conormal order δ is denoted by $\underline{\mathrm{As}}^\delta(Y)$.

Definition 2.18. Let $u \in C^{\infty,\delta}_{\mathrm{as}}(X)$ and $P \in \underline{\mathrm{As}}^{\delta}(Y)$ be represented by $J \subset \mathcal{E}_V(Y)$. Then u is said to have *asymptotics of type* P if there is a vector $\Phi \in J$ such that

$$u(x) \sim \sum_{p \in V} \sum_{k+l=m_p-1} \frac{(-1)^k}{k!} \log^k t\, \phi_l^{(p)}(y) \text{ as } t \to +0, \tag{2.16}$$

where $\Phi(p) = (\phi_0^{(p)}, \phi_1^{(p)}, \dots, \phi_{m_p-1}^{(p)})$ for $p \in V$. The space of all these u is denoted by $C^\infty_P(X)$.

Thus, by *representation of an asymptotic type* it is meant that P that — in the philosophy of asymptotic algebras, see Witt [20] — is the same as the linear subspace $C^\infty_P(X)/C^\infty_{\mathcal{O}}(X) \subset C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X)$, is mapped onto J by the isomorphism (2.13).
For $P \in \underline{\mathrm{As}}^{\delta}$ represented by $J \subset \mathcal{E}_V(Y)$, we introduce

$$\delta_P = \min\{\text{c-ord}(\Phi);\ \Phi \in J\}, \tag{2.17}$$

Notice that $\delta_P > \delta$ and $\delta_P = \infty$ if and only if $P = \mathcal{O}$.
Obviously, $\underline{\mathrm{As}}^{\delta}(Y) \subseteq \underline{\mathrm{As}}^{\delta'}(Y)$ if $\delta \geq \delta'$. We likewise set

$$\underline{\mathrm{As}}(Y) = \bigcup_{\delta \in \mathbb{R}} \underline{\mathrm{As}}^{\delta}(Y).$$

On asymptotic types $P \in \underline{\mathrm{As}}^{\delta}(Y)$, we have the shift operation T^ϱ for $\varrho \in \mathbb{R}$, namely $T^\varrho P$ is represented by the space

$$T^\varrho J = \big\{\Phi \in \mathcal{E}^{\varrho+\delta}_{T^\varrho V}(Y);\ \Phi(p) = \bar{\Phi}(p-\varrho),\ p \in \mathbb{C}, \text{ for some } \bar{\Phi} \in J\big\},$$

where $J \subset \mathcal{E}_V(Y)$ represents P.
Furthermore, for $J \subset \mathcal{E}_V(Y)$ as in Definition 2.17,

$$J_p = \{\Phi(p);\ \Phi \in J\} \subset [C^\infty(Y)]^\infty$$

for $p \in \mathbb{C}$ is the *localization of* J *at* p. Note that $TJ_p \subseteq J_p$ and $\dim J_p < \infty$; thus, J_p is a *local asymptotic type* in the sense of Witt [18].

We now investigate common properties of linear subspaces $J \subset \mathcal{E}_V(Y)$ satisfying (a) to (c) of Definition 2.17. Let $\Pi_j \colon J \to J^{\delta+j}$ be the canonical surjection. For $j' > j$, there is a natural surjective map $\Pi_{jj'} \colon J^{\delta+j'} \to J^{\delta+j}$ such that $\Pi_{jj''} = \Pi_{jj'}\Pi_{j'j''}$ for $j'' > j' > j$ and

$$\big(J, \Pi_j\big) = \operatorname*{proj\,lim}_{j \to \infty} \big(J^{\delta+j}, \Pi_{jj'}\big). \tag{2.18}$$

Note that $T\colon J^{\delta+j} \to J^{\delta+j}$ is nilpotent, where T denotes the map induced by $T\colon J \to J$. Furthermore, for $j' > j$, the diagram

$$\begin{array}{ccc} J^{\delta+j'} & \xrightarrow{\Pi_{jj'}} & J^{\delta+j} \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ J^{\delta+j'} & \xrightarrow{\Pi_{jj'}} & J^{\delta+j} \end{array} \tag{2.19}$$

commutes and the action of T on J is that one induced by (2.18), (2.19).

Proposition 2.19. *Let $J \subset \mathcal{E}_V(Y)$ be a linear subspace for some $V \in \mathcal{C}^\delta$. Then there is a sequence Φ_i for $i = 1, 2, \dots$ of special vectors with* c-ord$(\Phi_i) \to \infty$ *as $i \to \infty$ such that the vectors $T^k\Phi_i$ for $i = 1, 2, \dots,$ $k = 0, 1, 2 \dots$ span J if and only if J fulfills conditions* (a), (b), *and* (c).

In the situation just described, we write $J = \langle \Phi_1, \Phi_2, \dots \rangle$.

Proof. Let $J \subset \mathcal{E}_V(Y)$ fulfill conditions (a) to (c). Due to (c) we may assume that $V = V_p$ for some $p \in \mathbb{C}$. Suppose that the special vectors $\Phi_1, \dots, \Phi_e \in J$ have already been chosen (where $e = 0$ is possible). Then we choose the vector Φ_{e+1} among the special vectors $\Phi \in J$ which do not belong to $\langle \Phi_1, \dots, \Phi_e \rangle$ such that $\operatorname{Re}\gamma(\Phi_{e+1})$ is minimal. We claim that $J = \langle \Phi_1, \Phi_2, \dots \rangle$. In fact, c-ord$(\Phi_i) = (n+1)/2 - \operatorname{Re}\gamma(\Phi_i) \to \infty$ as $i \to \infty$ and, if Φ is a special vector in J, then $\Phi \in \langle \Phi_1, \dots, \Phi_e \rangle$, where e is such that $\operatorname{Re}\gamma(\Phi_e) \le \operatorname{Re}\gamma(\Phi)$, while $\operatorname{Re}\gamma(\Phi_{e+1}) > \operatorname{Re}\gamma(\Phi)$. Otherwise, Φ_{e+1} would not have been chosen in the $(e+1)$th step.
The other direction is obvious. □

For $j \ge 1$, let $(m_1^j, \dots, m_{e_j}^j)$ denote the characteristic of the space $J^{\delta+j}$, see Remark 2.15

Proposition 2.20. *Let $J \subset \mathcal{E}_V(Y)$ be a linear subspace and assume that the special vectors Φ_i for $i = 1, 2, \dots, e$, where $e \in \mathbb{N} \cup \{\infty\}$, as constructed in* Proposition 2.19, *form a characteristic basis of J. Then the following conditions are equivalent:*
(a) *For each j, $\Pi_j\Phi_1, \dots, \Pi_j\Phi_{e_j}^j$ is an $(m_1^j, \dots, m_{e_j}^j)$-characteristic basis of $J^{\delta+j}$;*
(b) *For each j, $T^{m_1^j-1}\Phi_1, \dots, T^{m_{e_j}-1}\Phi_{e_j}$ are linearly independent over the space $\mathcal{E}^{\delta+j}(Y)$, while $T^k\Phi_i \in \mathcal{E}^{\delta+j}(Y)$ if either $1 \le i \le e_j$, $k \ge m_i^j$ or $i > e_j$.*
In particular, if (a), (b) *are fulfilled, then, for any $j' > j$, $\Pi_{jj'}\Phi_1^{j'}, \dots, \Pi_{jj'}\Phi_{e_j}^{j'}$ is a characteristic basis of $J^{\delta+j}$, while $\Pi_{jj'}\Phi_{e_j+1}^{j'} = \cdots = \Pi_{jj'}\Phi_{e_j'}^{j'} = 0$. Here, $\Phi_i^{j'} = \Pi_{j'}\Phi_i$ for $1 \le i \le e_{j'}$.*

Proof. This is a consequence of Lemma 2.13 and Witt [18, Lemma 3.8]. □

Notice that, for a linear subspace $J \subset \mathcal{E}_V(Y)$ satisfying conditions (a) to (c) of Definition 2.17, a characteristic basis possessing the equivalent properties of Proposition 2.20 need not exist. We provide an example:

Example 2.21. Let the space $J = \langle \Phi_1, \Phi_2 \rangle \subset \mathcal{E}_{V_p}(Y)$ for some $p \in \mathbb{C}$, $\operatorname{Re} p < (n+1)/2 - \delta$, be spanned by two vectors Φ_1, Φ_2 in the sense of Proposition 2.19. We further assume that $\Phi_1(p) = (\psi_0, \star)$, $\Phi_1(p-1) = (\psi_1, \star, \star)$, $\Phi_2(p) = 0$, and $\Phi_2(p-1) = (\psi_1, \star)$, where ψ_0, $\psi_1 \in C^\infty(Y)$ are not identically zero and $\star$ stands for arbitrary entries, see Figure 2. Then, the asymptotic type represented by J is non-proper. In fact, assume that $\operatorname{Re} p \ge (n+1)/2 - \delta + 1$. Then

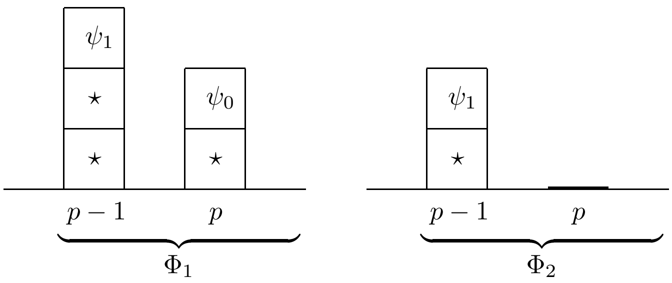


Figure 2: Example of a non-proper asymptotic type

$\Pi_2\Phi_1$, $T\Pi_2\Phi_1 - \Pi_2\Phi_2$ is a $(3,1)$-characteristic basis of $J^{\delta+2}$, and any other characteristic basis of $J^{\delta+2}$ is, up to a non-zero multiplicative constant, of the form

$$\begin{cases} \Pi_2\Phi_1 + \alpha_1 T\Pi_2\Phi_1 + \alpha_2 T^2\Pi_2\Phi_1 + \alpha_3\Pi_2\Phi_2, \\ \beta_1(T\Pi_2\Phi_1 - \Pi_2\Phi_2) + \beta_2 T^2\Pi_2\Phi_1, \end{cases} \tag{2.20}$$

where α_1, α_2, α_3, β_1, $\beta_2 \in \mathbb{C}$ and $\beta_1 \neq 0$. But then the conclusion in Proposition 2.20 is violated, since both vectors in (2.20) have non-zero image under the projection Π_{12}, while $\Pi_1\Phi_1$ is a (2)-characteristic basis of $J^{\delta+1}$.

Definition 2.22. An asymptotic type $P \in \underline{\mathrm{As}}^\delta(Y)$ represented by the linear subspace $J \subset \mathcal{E}_V(Y)$ is called *proper* if J admits a characteristic basis Φ_1, $\Phi_2, \ldots$ satisfying the equivalent conditions in Proposition 2.20. The set of all proper asymptotic types is denoted by $\underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y) \subsetneq \underline{\mathrm{As}}^\delta(Y)$.

For $\Phi \in \mathcal{E}^\delta(Y)$, $p \in \mathbb{C}$, and $\Phi(p) = (\phi_0^{(p)}, \phi_1^{(p)}, \ldots, \phi_{m_p-1}^{(p)})$ we shall use, for any $q \in \mathbb{C}$, the notation

$$\Phi(p)[z-q] = \frac{\phi_0^{(p)}}{(z-q)^{m_p}} + \frac{\phi_1^{(p)}}{(z-q)^{m_p-1}} + \cdots + \frac{\phi_{m_p-1}^{(p)}}{z-q} \in \mathcal{M}_q(C^\infty(Y)),$$

where $\mathcal{M}_q(C^\infty(Y))$ is the space of germs of *meromorphic functions* at $z = q$ taking values in $C^\infty(Y)$. Analogously, $\mathcal{A}_q(C^\infty(Y))$ is the space of germs of *holomorphic functions* at $z = p$ taking values in $C^\infty(Y)$.

Definition 2.23. For $\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j}(z);\, j \in \mathbb{N}\} \in \mathrm{Symb}^\mu_M(Y)$, the linear space $L^\delta_{\mathfrak{S}^\mu} \subseteq C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X)$ is represented by the space of $\Phi \in \mathcal{E}^\delta(Y)$ for which there are functions $\widetilde{\phi}^{(p)}(z) \in \mathcal{A}_p(C^\infty(Y))$ for $p \in \mathbb{C}$, $\operatorname{Re} p < (n+1)/2 - \delta$, such that

$$\sum_{j=0}^{[(n+1)/2-\delta+\mu-\operatorname{Re} q]^-} \mathfrak{s}^{\mu-j}(z-\mu+j)\Big(\Phi(q-\mu+j)[z-q] \\ + \widetilde{\phi}^{(q-\mu+j)}(z-\mu+j)\Big) \in \mathcal{A}_q(C^\infty(Y)) \tag{2.21}$$

for all $q \in \mathbb{C}$, $\operatorname{Re} q < (n+1)/2 - \delta + \mu$. Here, $[a]^-$ for $a \in \mathbb{R}$ is the largest integer strictly less than a, i.e., $[a]^- \in \mathbb{Z}$ and $[a]^- < a \le [a]^- + 1$.

Remark 2.24. (a) If $\Phi \in \mathcal{E}_V(Y)$ for $V \in \mathcal{C}^\delta$, then condition (2.21) is effective only if

$$q \in \bigcup_{j=0}^{[(n+1)/2-\delta+\mu-\operatorname{Re} q]^-} T^{\mu-j} V.$$

(b) If $\Phi \in \mathcal{E}^\delta(Y)$ belongs to the representing space of $L^\delta_{\mathfrak{S}^\mu}$, and if $u \in C^{\infty,\delta}_{\mathrm{as}}(X)$ possesses asymptotics given by the vector Φ according to (2.16), then there is a $v \in C^\infty_{\mathcal{O}}(X)$ such that

$$\sum_{j=0}^{\infty} \omega(c_j t) t^{-\mu+j} \operatorname{op}_M^{(n+1)/2-\delta}\left(\mathfrak{s}^{\mu-j}(z)\right) \tilde{\omega}(c_j t)\,(u+v) \in C^\infty_{\mathcal{O}}(X).$$

Here, the numbers $c_j > 0$ are chosen so that $c_j \to \infty$ as $j \to \infty$ sufficiently fast so that the infinite sum converges. For the notation $\operatorname{op}_M^{(n+1)/2-\delta}(\dots)$ see (2.35), below.

Definition 2.25. For $P \in \underline{\mathrm{As}}^\delta(Y)$ being represented by $J \subset \mathcal{E}_V(Y)$ and $\mathfrak{S}^\mu \in \operatorname{Symb}^\mu_M(Y)$, the *push-forward* $\mathcal{Q}^{\delta-\mu}(P;\mathfrak{S}^\mu)$ of P under $\mathfrak{S}^\mu$ is the asymptotic type in $\underline{\mathrm{As}}^{\delta-\mu}(Y)$ represented by the linear subspace $K \subset \mathcal{E}_{T^{-\mu}V}(Y)$ consisting of all vectors $\Psi \in \mathcal{E}_{T^{-\mu}V}(Y)$ such that there is a $\Phi \in J$ and there are functions $\widetilde{\phi}^{(p)}(z) \in \mathcal{A}_p(C^\infty(Y))$ for $p \in V$ such that

$$\Psi(q)[z-q] = \sum_{j=0}^{[(n+1)/2-\delta+\mu-\operatorname{Re} q]^-} \Big[\mathfrak{s}^{\mu-j}(z-\mu+j)\left(\Phi(q-\mu+j)[z-q] + \widetilde{\phi}^{(q-\mu+j)}(z-\mu+j)\right)\Big]^*_q, \quad (2.22)$$

holds for all $q \in T^\mu V$, see (2.6).

Remark 2.26. For a *holomorphic* $\mathfrak{S}^\mu \in \operatorname{Symb}^\mu_M(Y)$, one needs not to refer to the holomorphic functions $\widetilde{\phi}^{(p)}(z) \in \mathcal{A}_p(C^\infty(Y))$ for $p \in V$ in order to define the push-forward $\mathcal{Q}^{\delta-\mu}(P;\mathfrak{S}^\mu)$ in (2.22). We then also write $\mathcal{Q}(P;\mathfrak{S}^\mu)$ instead of $\mathcal{Q}^{\delta-\mu}(P;\mathfrak{S}^\mu)$.

Extending the notion of push-forward from asymptotic types to arbitrary linear subspaces of $C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X)$, the space $L^\delta_{\mathfrak{S}^\mu} \subseteq C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X)$ for $\mathfrak{S}^\mu \in \operatorname{Symb}^\mu_M(Y)$ appears as the largest subspace of $C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X)$ for which

$$\mathcal{Q}^{\delta-\mu}(L^\delta_{\mathfrak{S}^\mu};\mathfrak{S}^\mu) = \mathcal{Q}^{\delta-\mu}(\mathcal{O};\mathfrak{S}^\mu). \quad (2.23)$$

In this sense, it characterizes the *amount of asymptotics* of conormal order at least δ *annihilated by* $\mathfrak{S}^\mu \in \operatorname{Symb}^\mu_M(Y)$.

Definition 2.27. A *partial ordering* on $\underline{\mathrm{As}}^\delta(Y)$ is defined by $P \preccurlyeq P'$ for P, $P' \in \underline{\mathrm{As}}^\delta(Y)$ if and only if $J \subseteq J'$, where J, $J' \subset \mathcal{E}^\delta(Y)$ are the representing spaces for P and P', respectively.

Proposition 2.28. (a) *The p.o. set $(\underline{\mathrm{As}}^\delta(Y), \preccurlyeq)$ is a lattice in which each non-empty subset $\mathcal{S}$ admits a meet, $\bigwedge \mathcal{S}$, represented by $\bigcap_{P\in\mathcal{S}} J_P$, and each bounded subset $\mathcal{T}$ admits a join, $\bigvee \mathcal{T}$, represented by $\sum_{Q\in\mathcal{T}} J_Q$, where J_P and J_Q represent the asymptotic types P and Q, respectively. In particular, $\bigwedge \underline{\mathrm{As}}^\delta(Y) = \mathcal{O}$.*
(b) *For $P \in \underline{\mathrm{As}}^\delta(Y)$, $\mathfrak{S}^\mu \in \mathrm{Symb}^\mu_M(Y)$, we have $\mathcal{Q}^{\delta-\mu}(P;\mathfrak{S}^\mu) \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$.*

Proof. (a) is immediate from the definition of asymptotic type and (b) can be checked directly on the level of (2.22). $\square$

Remark 2.29. Each element $\mathfrak{S}^\mu \in \mathrm{Symb}^\mu_M(Y)$ induces a natural action $C^{\infty,\delta}_{\mathrm{as}}(X) \to C^{\infty,\delta}_{\mathrm{as}}(X)/C^\infty_{\mathcal{O}}(X)$. Its expression in the splitting of coordinates $\mathcal{U} \to [0,1) \times Y$, $x \mapsto (t,y)$, is given by (2.22).
In the language of Witt [20], this means that the quadruple $\big(\bigcup_{\mu\in\mathbb{Z}} \mathrm{Symb}^\mu_M(Y), C^{\infty,\delta}_{\mathrm{as}}(X), C^\infty_{\mathcal{O}}(X), \underline{\mathrm{As}}^\delta(Y)\big)$ is an *asymptotic algebra* that is even *reduced*; thus providing justification for the above choice of the notion of asymptotic type.

Theorem 2.30. *For a holomorphic $\mathfrak{S}^\mu \in \mathrm{Ell\,Symb}^\mu_M(Y)$, we have $L^\delta_{\mathfrak{S}^\mu} \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$.*

Proof. Let $\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j}(z);\, j \in \mathbb{N}\} \subset \mathcal{M}^\mu_{\mathcal{O}}(Y)$. Assume that, for some $p \in \mathbb{C}$, $\mathrm{Re}\, p < (n+1)/2 - \delta$, $\Phi_0 \in L_{\mathfrak{s}^\mu(z)}$ at $z = p$, with the obvious meaning, for this see Witt [18]. (Notice that $L_{\mathfrak{s}^\mu(z)}$ at $z = p$ is contained in the space $[C^\infty(Y)]^\infty$.) We then successively calculate the sequence Φ_0, Φ_1, $\Phi_2, \dots$ from the relations, at $z = p$,

$$\begin{aligned}\mathfrak{s}^\mu(z-j)\Phi_j[z-p] &+ \mathfrak{s}^{\mu-1}(z-j+1)\Phi_{j-1}[z-p] \\ &+ \dots + \mathfrak{s}^{\mu-j}(z)\Phi_0[z-p] \in \mathcal{A}_p(C^\infty(Y)),\ \ j = 0,1,2,\dots, \end{aligned} \tag{2.24}$$

see (2.22) and Remark 2.26. In each step, we find $\Phi_j \in [C^\infty(Y)]^\infty$ uniquely determined modulo $L_{\mathfrak{s}^\mu(z)}$ at $z = p - j$ such that (2.24) holds. We obtain the vector $\Phi \in \mathcal{E}_{V_p}(Y)$ define by $\Phi(p-j) = \Phi_j$ that belongs to the linear subspace $J \subset \mathcal{E}^\delta(Y)$ representing $L^\delta_{\mathfrak{S}^\mu}$.
Conversely, each vector in J is a sum like in (2.11) of vectors Φ obtained in that way. Thus, upon choosing in each space $L_{\mathfrak{s}^\mu(z)}$ at $z = p$ a characteristic basis and then, for each characteristic basis vector $\Phi_0 \in [C^\infty(Y)]^\infty$, exactly one vector $\Phi \in \mathcal{E}_{V_p}(Y)$ as just constructed, we obtain a characteristic basis of J in the sense of Definition 2.14 consisting completely of special vectors (since $L_{\mathfrak{s}^\mu(z)}$ at $z = p$ equals zero for all $p \in \mathbb{C}$, $\mathrm{Re}\, p < (n+1)/2 - \delta$, but a set of p belonging to $\mathcal{C}^\delta$). In particular, $J \subset \mathcal{E}_V(Y)$ for some $V \in \mathcal{C}^\delta$ and (a) to (c) of Definition 2.17 are satisfied. By its very construction, this characteristic basis fulfills condition (b) of Proposition 2.20. Therefore, the asymptotic type $L^\delta_{\mathfrak{S}^\mu}$ represented by J is proper. $\square$

In conclusion, we obtain:

Proposition 2.31. *Let* $\mathfrak{S}^\mu \in \operatorname{Ell}\operatorname{Symb}^\mu_M(Y)$. *Then:*
(a) $L^\delta_{\mathfrak{S}^\mu} = \mathcal{Q}^\delta(\mathcal{O}; (\mathfrak{S}^\mu)^{-1})$ *and* $L^{\delta-\mu}_{(\mathfrak{S}^\mu)^{-1}} = \mathcal{Q}^{\delta-\mu}(\mathcal{O}; \mathfrak{S}^\mu)$.
(b) *There is an order-preserving bijection*

$$\big\{P \in \underline{\mathrm{As}}^\delta(Y);\, P \succcurlyeq L^\delta_{\mathfrak{S}^\mu}\big\} \to \big\{Q \in \underline{\mathrm{As}}^{\delta-\mu}(Y);\, Q \succcurlyeq L^{\delta-\mu}_{(\mathfrak{S}^\mu)^{-1}}\big\}, \qquad (2.25)$$
$$P \mapsto \mathcal{Q}^{\delta-\mu}(P; \mathfrak{S}^\mu),$$

with the inverse given by $Q \mapsto \mathcal{Q}^\delta(Q; (\mathfrak{S}^\mu)^{-1})$.

Proof. Using Proposition 2.28 (b), the proof consists of a word-by-word repetition of the arguments given in the proof of Witt [18, Proposition 2.5]. □

In its consequence, Proposition 2.31 enables one to perform explicit calculations on asymptotic types.
We conclude this section with the following basic observation:

Proposition 2.32. *The notion of asymptotic type, as introduced above, is invariant under coordinates changes.*

Proof. Let $\kappa\colon X \to X$ be a C^∞ diffeomorphism and let $\kappa_*\colon C^\infty(X^\circ) \to C^\infty(X^\circ)$ be the corresponding push-forward on the level of functions, i.e., $(\kappa_* u)(x) = u(\kappa^{-1}(x))$ for $u \in C^\infty(X^\circ)$, where κ^{-1} denotes the inverse C^∞ diffeomorphism to κ. As is well-known, κ_* restricts to $\kappa_*\colon C^{\infty,\delta}_{\mathrm{as}}(X) \to C^{\infty,\delta}_{\mathrm{as}}(X)$ for any $\delta \in \mathbb{R}$, see, e.g., Schulze [15, Theorem 1.2.1.11].
We have to prove that, for each $P \in \underline{\mathrm{As}}^\delta(Y)$, there is a $\kappa_* P \in \underline{\mathrm{As}}^\delta(Y)$ so that the push-forward κ_* restricts further to a linear isomorphism $\kappa_*\colon C^\infty_P(X) \to C^\infty_{\kappa_* P}(X)$, i.e., we have to show that there is a $\kappa_* P \in \underline{\mathrm{As}}^\delta(Y)$ so that $\kappa_*(C^\infty_P(X)) = C^\infty_{\kappa_* P}(X)$. Using Proposition 2.19, we eventually have to prove that, for each $u \in C^{\infty,\delta}_{\mathrm{as}}(X)$ such that

$$u(x) \sim \sum_{j=0}^\infty \sum_{k+l=m_j-1} \frac{(-1)^k}{k!} \log^k t\, \phi^{(j)}_l(y) \ \text{ as } t \to +0, \qquad (2.26)$$

where $\Phi \in \mathcal{E}_{V_p}(Y)$ for a certain $p \in \mathbb{C}$, $\operatorname{Re} p < (n+1)/2 - \delta$, and $\Phi(p-j) = (\phi^{(j)}_0, \phi^{(j)}_1, \dots, \phi^{(j)}_{m_j-1})$ for all $j \in \mathbb{N}$, see (2.16), the push-forward $\kappa_* u$ is again of the form (2.26), with some other $\kappa_*\Phi \in \mathcal{E}_{V_p}(Y)$ in place of $\Phi \in \mathcal{E}_{V_p}(Y)$.
But this is immediate from a direct computation. □

2.2.3 Characteristics of proper asymptotic types

We introduce the notion of characteristic of a proper asymptotic type. This will be the main ingredient in the prove of Theorem 2.42.
Let $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$ be represented by $J \subset \mathcal{E}_V(Y)$ and let $\Phi_1, \Phi_2, \dots$ by a characteristic basis of J according to Definition 2.22. As before, let $(m^j_1, \dots, m^j_{e_j})$

be the characteristic of the space $J^{\delta+j}$. From Proposition 2.20, we conclude that $e_1 \leq e_2 \leq \ldots$ In the next lemma, we find a suitable "path through" the numbers m_i^j for $j \geq j_i$, where $j_i = \min\{j;\, e_j \geq i\}$, i.e., an appropriate re-ordering of the tuples $(m_1^j, \ldots, m_{e_j}^j)$.

Lemma 2.33. *The numbering within the tuples $(m_1^j, \ldots, m_{e_j}^j)$ can be chosen in such a way that, for each $j \geq 1$, there is a characteristic $(m_1^j, \ldots, m_{e_j}^j)$-basis $(\Phi_1^j, \ldots, \Phi_{e_j}^j)$ of $J^{\delta+j}$ such that, for all $j' > j$,*

$$\Pi_{jj'}\Phi_i^{j'} = \begin{cases} \Phi_i^j & \text{if } 1 \leq i \leq e_j, \\ 0 & \text{if } e_j + 1 \leq i \leq e_{j'} \end{cases}$$

holds.
Furthermore, the scheme

$$\begin{array}{r l} e_1 \text{ rows} & \left\{ \begin{array}{ccccc} m_1^1 & m_1^2 & m_1^3 & m_1^4 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_{e_1}^1 & m_{e_1}^2 & m_{e_1}^3 & m_{e_1}^4 & \cdots \end{array} \right. \\ e_2 - e_1 \text{ rows} & \left\{ \begin{array}{ccccc} \phantom{m_{e_1}^1} & m_{e_1+1}^2 & m_{e_1+1}^3 & m_{e_1+1}^4 & \cdots \\ & \cdots & \cdots & \cdots & \cdots \\ & m_{e_2}^2 & m_{e_2}^3 & m_{e_2}^4 & \cdots \end{array} \right. \\ e_3 - e_2 \text{ rows} & \left\{ \begin{array}{ccccc} \phantom{m_{e_1}^1} & \phantom{m_{e_1}^2} & m_{e_2+1}^3 & m_{e_2+1}^4 & \cdots \\ & & \cdots & \cdots & \cdots \\ & & m_{e_3}^3 & m_{e_3}^4 & \cdots \end{array} \right. \\ & \left\{ \begin{array}{ccccc} \phantom{m_{e_1}^1} & \phantom{m_{e_1}^2} & \phantom{m_{e_1}^3} & \vdots & \ddots, \end{array} \right. \end{array} \tag{2.27}$$

where in the jth column the characteristic of the space $J^{\delta+j}$ appears, is uniquely determined up to permutation of the kth and the k'th row, where $e_j+1 \leq k,\, k' \leq e_{j+1}$ for some j ($e_0 = 0$).

Proof. This is a reformulation of Proposition 2.20 in terms of the characteristics of the spaces $J^{\delta+j}$. Notice that one can recover the characteristic basis $\Phi_1, \Phi_2, \ldots$ of J, that was initially given, from the property that $\Pi_j\Phi_i = \Phi_i^j$ holds for all $1 \leq i \leq e_j$, while $\Pi_j\Phi_i = 0$ for $i > e_j$. $\square$

Performing the constructions of the foregoing lemma for each space $J \cap \mathcal{E}_{V_{p_j}}(Y)$ in (2.15) separately, one sees that the following notion is correctly defined:

Definition 2.34. Let $P \in \underline{\mathrm{As}}^{\delta}_{\mathrm{prop}}(Y)$ and $J \subset \mathcal{E}_V(Y)$ represent P. If $\Phi_1, \Phi_2, \ldots$ is a characteristic basis of J according to Definition 2.22 and if the tuples $(m_1^j, \ldots, m_{e_j}^j)$ are re-ordered according to Lemma 2.33, then the sequence

$$\operatorname{char} P = \left\{ \left(\gamma(\Phi_i) \;\middle|\; m_i^{j_i}, m_i^{j_i+1}, m_i^{j_i+2}, \ldots \right) \right\}_{i=1}^{e} \tag{2.28}$$

is called the *characteristic* of P.

The characteristic $\operatorname{char} P$ of an asymptotic type $P \in \underline{\mathrm{As}}^{\delta}_{\mathrm{prop}}(Y)$ is unique up to permutation of the kth and the k'th entry, where $e_j + 1 \le k,\ k' \le e_{j+1}$ for some j. So far, it is an invariant associated with the representing space J; so it still depends on the splitting of coordinates. However, we have:

Proposition 2.35. *The characteristic* $\operatorname{char} P$ *of an asymptotic type* $P \in \underline{\mathrm{As}}^{\delta}_{\mathrm{prop}}(Y)$ *is independent of the chosen splitting of coordinates* $\mathcal{U} \to [0,1) \times Y$, $x \mapsto (t, y)$, *near* ∂X.

Proof. Follow the proof of Proposition 2.32 to get the assertion. $\square$

Now, let $\left\{(p_i \,|\, m_i^{j_i}, m_i^{j_i+1}, \dots)\right\}_{i=1}^{e} \subset \mathbb{C} \times \mathbb{N}^{\mathbb{N}}$ be any given sequence, where we additionally assume that $\operatorname{Re} p_i < (n+1)/2 - \delta$ for all i, $\operatorname{Re} p_i \to -\infty$ as $i \to \infty$ when $e = \infty$, the p_i are ordered so that $\operatorname{Re} p_i \ge (n+1)/2 - \delta - j$ holds if and only if $i \le e_j$ for a certain (then uniquely determined) sequence $e_1 \le e_2 \le \dots$ satisfying $e = \sup_j e_j$, and

$$1 \le m_i^{j_i} \le m_i^{j_i+1} \le m_i^{j_i+2} \le \dots,$$

where $j_i = \min\{j;\, e_j \ge i\}$ as above.

Proposition 2.36. *Let the characteristic* $\left\{\left(p_i \,\middle|\, m_i^{j_i}, m_i^{j_i+1}, \dots\right)\right\}_{i=1}^{e}$ *satisfy the properties just mentioned. If* $n = 0$, *then we assume, in addition, that* $p_i \ne p_{i'}$ *for* $i \ne i'$ *and, for all* i, $k > 0$,

$$m_i^{j_i+k} - m_i^{j_i+k-1} = a > 0 \iff p_{i'} = p_i - k \text{ for some } i' \text{ and } m_{i'}^{j_{i'}} = a$$

(*where* $j_{i'} = j_i + k$). *Then there exists a holomorphic* $\mathfrak{S}^{\mu} \in \mathrm{Symb}^{\mu}_{M}(Y)$ *that is elliptic with respect to the weight* $\delta \in \mathbb{R}$ *such that* $L^{\delta}_{\mathfrak{S}^{\mu}} \in \underline{\mathrm{As}}^{\delta}_{\mathrm{prop}}(Y)$ *has exactly this characteristic.*

Proof. Multiplying $\mathfrak{S}^{\mu}$ by an elliptic element $\mathfrak{T}^{-\mu} = \{\mathfrak{t}^{-\mu}(z), 0, 0, \dots\}$ such that $\mathfrak{t}^{-\mu}(z) \in \mathcal{M}^{-\mu}_{\mathcal{O}}$ and $\mathfrak{t}^{-\mu}(z)^{-1} \in \mathcal{M}^{\mu}_{\mathcal{O}}$, we can assume $\mu = 0$.
If $n = 0$, then we choose an elliptic $\mathfrak{s}^{0}(z) \in \mathcal{M}^{0}_{\mathcal{O}}$ that has zeros precisely at $z = p_i$ of order $m_i^{j_i}$ for $i = 1, 2, \dots$ according to Theorem 2.6.
In case $\dim Y > 0$, let $\{\phi_i\}_{i=1}^{e}$ be an orthonormal set in $C^{\infty}(Y)$ with respect to a fixed C^{∞}-density $d\mu$ on Y. Let Π_i for $i = 1, \dots, e$ be the orthogonal projection in $L^2(Y, d\mu)$ onto the subspace spanned by ϕ_i. We then choose an elliptic $\mathfrak{s}^{\mu}(z) \in \mathcal{M}^{\mu}_{\mathcal{O}}(Y)$ such that, for every $p \in V_{p_i}$ and all i,

$$[\mathfrak{s}^{\mu}(z)]_p^{N_p} = \Bigl(1 - \sum_{p_{i'} - k = p} \Pi_{i'}\Bigr) + \sum_{p_{i'} - k = p} (z - p)^{m_{i'}^{j_{i'}+k}} \Pi_{i'}$$

where the sums are extended over all i', k such that $p_{i'} - k = p$, for some N_p sufficiently large, while $\mathfrak{s}^{\mu}(q) \in L^{\mu}_{\mathrm{cl}}(Y)$ is invertible for all $q \in \mathbb{C} \setminus V$, again according to Theorem 2.6.
In both cases, we set $\mathfrak{S}^{\mu} = \{\mathfrak{s}^{\mu-j}(z)\}_{j=0}^{\infty}$ with $\mathfrak{s}^{\mu-j}(z) \equiv 0$ for $j > 0$. Then $\mathfrak{S}^{\mu} \in \mathrm{Symb}^{\mu}_{M}(Y)$ is elliptic with respect to the weight δ, and the proper asymptotic type $L^{\delta}_{\mathfrak{S}^{\mu}}$ has characteristic $\left\{(p_i \,|\, m_i^{j_i}, m_i^{j_i+1}, \dots)\right\}_{i=1}^{e}$. $\square$

2.2.4 More properties of asymptotic types

Here, we study further properties of asymptotic types. First, asymptotic types are composed of elementary building blocks:

Proposition 2.37. (a) *An asymptotic type $P \in \underline{\mathrm{As}}^{\delta}(Y)$ is join-irreducible, i.e., $P \neq \mathcal{O}$ and $P = P_0 \vee P_1$ for $P_0, P_1 \in \underline{\mathrm{As}}^{\delta}(Y)$ implies $P = P_0$ or $P = P_1$, if and only if there is a $\Phi \in \mathcal{E}^{\delta}(Y)$, $\Phi \neq 0$, such that the representing space, J, for P, in the given splitting of coordinates near ∂X, has characteristic basis Φ, i.e., $J = \langle \Phi \rangle$. In particular, every join-irreducible asymptotic type is proper.*
(b) *The join-irreducible asymptotic types are join-dense in $\underline{\mathrm{As}}^{\delta}(Y)$.*

Proof. (a) Let $P \neq \mathcal{O}$. Assume that, for some $j \geq 1$, $J^{\delta+j}$ has characteristic of length larger 1. Then $J^{\delta+j} = K_0 + K_1$ for certain linear subspaces $K_i \subsetneq J^{\delta+j}$ satisfying $TK_i \subseteq K_i$, for $i = 0, 1$. Setting $J_i = \{\Phi \in J; \Pi_j \Phi \in K_i\}$, we get that $J = J_0 + J_1$, $J_i \subsetneq J$, and $TJ_i \subseteq J_i$ for $i = 0, 1$. Since this decomposition can be chosen compatible with (2.15), we obtain that a necessary condition for P to be join-irreducible is that each space $J^{\delta+j}$ for $j \geq 1$ has characteristic of length at most 1, i.e., $J = \langle \Phi \rangle$ for some $\Phi \neq 0$. Vice versa, if $J = \langle \Phi \rangle$ for some $\Phi \neq 0$, then P is join-irreducible, since the subspace $\langle T^k \Phi \rangle \subseteq J$ for $k \in \mathbb{N}$ are the only subspaces of J that are invariant under the action of T.
(b) This follows directly from Proposition 2.19. $\square$

Note that, by the foregoing proposition, also the proper asymptotic types are join-dense in $\underline{\mathrm{As}}^{\delta}(Y)$. We will utilize this fact in the definition of cone Sobolev spaces with asymptotics.
In constructing asymptotic types $P \in \underline{\mathrm{As}}^{\delta}(Y)$ obeying certain properties, one often encounters a situation in which P is successively constructed on strips $\{z \in \mathbb{C}; (n+1)/2 - \delta - \beta_h \leq \operatorname{Re} z < (n+1)/2 - \delta\}$ of finite width, where the sequence $\{\beta_h\}_{h=0}^{\infty} \subset \mathbb{R}_+$ is strictly increasing and $\beta_h \to \infty$ as $h \to \infty$. We will meet an example in Section 3.3.
To formulate the result, we need one more definition:

Definition 2.38. Let $P, P' \in \underline{\mathrm{As}}^{\delta}(Y)$ be represented by $J \subset \mathcal{E}_V(Y)$ and $J' \subset \mathcal{E}_V(Y)$, respectively. Then, for $\vartheta \geq 0$, the asymptotic types P and P' are said to be *equal up to the conormal order $\delta + \vartheta$* if $\Pi_\vartheta J = \Pi_\vartheta J'$, where $\Pi_\vartheta \colon J \to J/(J \cap \mathcal{E}^{\delta+\vartheta}(Y))$ is the canonical projection. Similarly, P and P' are said to be *equal up to the conormal order $\delta + \vartheta - 0$* if they are equal up to the conormal order $\delta + \vartheta - \epsilon$, for any $\epsilon > 0$. (Similarly for the order relation $\preccurlyeq$ instead of equality.)

Proposition 2.39. *Let $\{P_\iota\}_{\iota \in \mathcal{I}} \subset \underline{\mathrm{As}}^{\delta}(Y)$ be an increasing net of asymptotic types. Then the join $\bigvee_{\iota \in \mathcal{I}} P_\iota$ exists if and only if, for each $j \geq 1$, there is an $\iota_j \in \mathcal{I}$ such that $P_\iota = P_{\iota'}$ up to the conormal order $\delta + j$ for all $\iota, \iota' \geq \iota_j$.*

Proof. The condition is obviously sufficient.
Conversely, suppose that the join $\bigvee_{\iota \in \mathcal{I}} P_\iota$ exists. Let P_ι be represented by the subspace $J_\iota \subset \mathcal{E}_{V_\iota}(Y)$ for $V_\iota \in \mathcal{C}^{\delta}$. Since the join $\bigvee_{\iota \in \mathcal{I}} P_\iota$ exists, the carriers V_ι

can be chosen in such way that $\bigcup_{\iota\in\mathcal{I}} V_\iota \subseteq V$ for some $V \in \mathcal{C}^\delta$. Thus $J_\iota \subset \mathcal{E}_V(Y)$ for all ι. Now, for each $j \geq 1$, $\dim\bigl(\sum_{\iota\in\mathcal{I}} J_\iota^{\delta+j}\bigr) < \infty$, otherwise $\bigvee_{\iota\in\mathcal{I}} P_\iota$ does not exist. But since the net $\{J_\iota^{\delta+j}\}_{\iota\in\mathcal{I}}$ is increasing, this already implies that there is some $\iota_j \in \mathcal{I}$ such that $J_\iota^{\delta+j} = J_{\iota'}^{\delta+j}$ for ι, $\iota' \geq \iota_j$, i.e., $P_\iota = P_{\iota'}$ up to the conormal order $\delta + j$ for ι, $\iota' \geq \iota_j$. □

An equivalent condition is that the net $\{P_\iota\}_{\iota\in\mathcal{I}} \subset \underline{\mathrm{As}}^\delta(Y)$ of asymptotic types be bounded on each strip $\{z \in \mathbb{C}; (n+1)/2 - \delta - j \leq \operatorname{Re} z < (n+1)/2 - \delta\}$ of finite width.

2.3 Pseudodifferential theory

Here, we establish an analogue of Witt [18, Theorem 1.2]. We need:

Proposition 2.40. *Let P, $P_0 \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$, $Q \in \underline{\mathrm{As}}^{\delta-\mu}_{\mathrm{prop}}(Y)$ for $\mu \in \mathbb{R}$. Assume that $P \wedge P_0 = \mathcal{O}$. Then there is a holomorphic $\mathfrak{S}^\mu \in \operatorname{Ell}\operatorname{Symb}^\mu_M(Y)$ that is elliptic with respect to the weight δ such that $L^\delta_{\mathfrak{S}^\mu} = P_0$ and $\mathcal{Q}(P;\mathfrak{S}^\mu) = Q$ if and only if P and Q have the same characteristic shifted by μ, i.e., we have* $\operatorname{char} P = \operatorname{char} Q - \mu$ *(with the obvious meaning of* $\operatorname{char} Q - \mu$*).*

Proof. It is readily seen that $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$, $Q \in \underline{\mathrm{As}}^{\delta-\mu}_{\mathrm{prop}}(Y)$ have the same characteristic shifted by μ if there is a holomorphic $\mathfrak{S}^\mu \in \operatorname{Ell}\operatorname{Symb}^\mu_M(Y)$ such that $\mathcal{Q}(P;\mathfrak{S}^\mu) = Q$.

Suppose that $\operatorname{char} P = \operatorname{char} Q - \mu$. First, we deal with the case $P_0 = \mathcal{O}$. Let the asymptotic types P, Q be represented by $J \subset \mathcal{E}_V(Y)$ and $K \subset \mathcal{E}_{T^\mu V}(Y)$, respectively. Let $\{\Phi_i\}_{i=1}^e$ and $\{\Psi_i\}_{i=1}^e$ be characteristic bases of J and K corresponding to $\operatorname{char} P$ and $\operatorname{char} Q$, respectively.

We have to choose the sequence $\{\mathfrak{s}^{\mu-k}(z); k \in \mathbb{N}\} \subset \mathcal{M}^\mu_\mathcal{O}(Y)$. By Theorem 2.6, it suffices to construct the finite parts $[\mathfrak{s}^{\mu-k}(z)]_{p'}^{N_{p'k}}$ for $p' \in V$, $k \in \mathbb{N}$, and $N_{p'k}$ sufficiently large appropriately. Thereby, we can assume that $V = V_p$ for some $p \in \mathbb{C}$, $\operatorname{Re} p < (n+1)/2 - \delta$.

Let $e_1 \leq e_2 \leq \dots$, where $e = \sup_{j\in\mathbb{N}} e_j$, be such that $\gamma(\Phi_i) = \gamma(\Psi_i) - \mu = p - j$ for $e_{j-1} + 1 \leq i \leq e_j$ (and $e_0 = 0$). Then the finite parts $[\mathfrak{s}^{\mu-k}(z)]_{p-j}^{m^{j+k}}$ for all j, k must be chosen so that, for each $j \in \mathbb{N}$,

$$\begin{multline}\Phi_i(p-j)^{[\mathfrak{s}^\mu(z)]_{p-j}^{m^j}} + \Phi_i(p-j+1)^{[\mathfrak{s}^{\mu-1}(z)]_{p-j+1}^{m^j}} \\ + \dots + \Phi_i(p)^{[\mathfrak{s}^{\mu-j}(z)]_p^{m^j}} = \Psi_i(p+\mu-j) \quad (2.29)\end{multline}$$

for $1 \leq i \leq e_j$, where $m^j = \sup_{1\leq i\leq e_j} m_i^j$, and $\Phi_i(p-k) = 0$ if $e_k + 1 \leq i \leq e_j$. Here, $(m_1^j, \dots, m_{e_j}^j)$ is the characteristic of $J^{\delta+j}$ and, for $\Phi = (\phi_0, \dots, \phi_{m-1})$, $\Psi = (\psi_0, \dots, \psi_{m-1}) \in [C^\infty(Y)]^\infty$, and $\mathfrak{s}(z) \in \mathcal{M}^\mu_\mathcal{O}(Y)$, the relation

$$\Phi^{[\mathfrak{s}(z)]_p^m} = \Psi$$

stands for the linear system

$$\begin{aligned}\mathfrak{s}(p)\phi_0 &= \psi_0,\\ \mathfrak{s}(p)\phi_1 + \frac{\mathfrak{s}'(p)}{1!}\,\phi_0 &= \psi_1,\\ &\vdots\\ \mathfrak{s}(p)\phi_{m-1} + \frac{\mathfrak{s}'(p)}{1!}\,\phi_{m-2} + \cdots + \frac{\mathfrak{s}^{(m-1)}(p)}{(m-1)!}\,\phi_0 &= \psi_{m-1}.\end{aligned}$$

System (2.29) can successively be solved for $[\mathfrak{s}^\mu(z-k)]^{m^j}_{p-j+k}$ for $j = 0, 1, 2, \ldots$ and $0 \le k \le j$. In fact, this can be done by choosing $[\mathfrak{s}^{\mu-k}(z)]^{m^j}_{p-j+k}$ for $k > 0$ arbitrarily. In particular, we may choose $\mathfrak{s}^{\mu-k}(z) \equiv 0$ for $k > 0$.
The case $P_0 \neq \mathcal{O}$ can be reduced to the case $P_0 = \mathcal{O}$ as in the proof of Witt [18, Lemma 3.16], since the three rules from Witt [18, Lemma 2.3] applied there continues to hold in the present situation. □

Remark 2.41. (a) The proof of Proposition 2.40 shows that the holomorphic $\mathfrak{S}^\mu = \{\mathfrak{s}^{\mu-j};\, j \in \mathbb{N}\} \in \operatorname{Ell}\operatorname{Symb}^\mu_M(Y)$ satisfying $L^\delta_{\mathfrak{S}^\mu} = P_0$ and $\mathcal{Q}(P; \mathfrak{S}^\mu) = Q$ can always be chosen so that $\mathfrak{s}^{\mu-j}(z) \equiv 0$ for $j > 0$.
(b) Proposition 2.40 in connection with Theorem 2.30 also shows that $\underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$ consists precisely of those asymptotic types that are of the form $L^\delta_{\mathfrak{S}^\mu}$ for some holomorphic $\mathfrak{S}^\mu \in \operatorname{Ell}\operatorname{Symb}^\mu_M(Y)$ that is elliptic with respect to the weight δ. (Choose $P = Q = \mathcal{O}$ in Proposition 2.40.)
Now, we reach the final aim of this section:

Theorem 2.42. *Let $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$ and $Q \in \underline{\mathrm{As}}^{\delta-\mu}_{\mathrm{prop}}(Y)$. Then there exists a $\mathfrak{S}^\mu \in \operatorname{Symb}^\mu_M(Y)$ that is elliptic with respect to the weight δ such that $L^\delta_{\mathfrak{S}^\mu} = P$ and $L^{\delta-\mu}_{(\mathfrak{S}^\mu)^{-1}} = Q$ always when $\dim Y > 0$ and if and only if $P \wedge T^{-\mu}Q = \mathcal{O}$ when $\dim Y = 0$.*

Proof. The condition $P \wedge T^{-\mu}Q = \mathcal{O}$ is obviously necessary if $\dim Y = 0$.
In the general case, choose $P_1 \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$, $Q_1 \in \underline{\mathrm{As}}^{\delta-\mu}_{\mathrm{prop}}(Y)$ having the same characteristics as P and Q, respectively, such that $P_1 \wedge T^{-\mu}Q_1 = \mathcal{O}$. As in the proof of Witt [18, Theorem 1.2], it then suffices to construct holomorphic $\mathfrak{S}^0 \in \operatorname{Ell}\operatorname{Symb}^\mu_M(Y)$, $\mathfrak{T}^0 \in \operatorname{Ell}\operatorname{Symb}^{-\mu}_M(Y)$ that are elliptic with respect to the weight δ such that

$$L^\delta_{\mathfrak{S}^0} = P_1,\;\; \mathcal{Q}^\delta(Q_1; \mathfrak{S}^0) = Q,\;\; L^\delta_{\mathfrak{T}^0} = Q_1,\;\; \mathcal{Q}^\delta(P_1; \mathfrak{T}^0) = P.$$

This is achieved by using Proposition 2.40. □

2.4 Function spaces with asymptotics

The definition of cone Sobolev spaces with asymptotics is based on the Mellin transformation. See Schulze [15, Sections 1.2, 2.1] for this idea and also Remark 2.45. For more details on the Mellin transformation, see Jeanquartier [5].

2.4.1 Weighted cone Sobolev spaces

Let $Mu(z) = \tilde{u}(z) = \int_0^\infty t^{z-1}u(t)\,dt$, $z \in \mathbb{C}$, be the *Mellin transformation*, first defined for $u \in C_0^\infty(\mathbb{R}_+)$ and then extended to larger distribution classes. In particular, u will be allowed to be vector-valued. Recall the following properties of M:

$$\begin{aligned} M_{t\to z}\big\{(-t\partial_t - p)u\big\}(z) &= (z-p)\tilde{u}(z), \\ M_{t\to z}\big\{t^{-p}u\big\}(z) &= \tilde{u}(z-p), \qquad p \in \mathbb{C}, \end{aligned}$$

whenever both sides are defined, $M\colon L^2(\mathbb{R}_+) \to L^2(\Gamma_{1/2}; (2\pi i)^{-1}dz)$ is an isometry, and

$$M_{t\to z}\left\{\frac{(-1)^k}{k!}\, t^{-p}\log^k t\, \chi_{(0,1)}(t)\right\}(z) = \frac{1}{(z-p)^{k+1}}, \tag{2.30}$$

where $\chi_{(0,1)}$ is the characteristic function of the interval $(0,1)$. We infer that $h(z) = M_{t\to z}\left\{(-1)^k\omega(t)t^{-p}\log^k t/k!\right\}(z) \in \mathcal{M}_{\mathrm{as}}^{-\infty}$ is a meromorphic function of z having a pole precisely at $z = p$, and the principal part of the Laurent expansion around this pole is given by the right-hand side of (2.30), i.e., $[h(z)]_p^* = (z-p)^{-(k+1)}$. Here, $\omega(t)$ is a cut-off function near $t = 0$.

For s, $\delta \in \mathbb{R}$, let $\mathcal{H}^{s,\delta}(X)$ denote the space of $u \in H^s_{\mathrm{loc}}(X^\circ)$ such that $M_{t\to z}\{\omega u\}(z) \in L^2_{\mathrm{loc}}\big(\Gamma_{(n+1)/2-\delta}; H^s(Y)\big)$ and the expression

$$\|u\|_{\mathcal{H}^{s,\delta}(X)} = \left\{\frac{1}{2\pi i}\int_{\Gamma_{(n+1)/2-\delta}} \big\|R^s(z)M_{t\to z}\{\omega u\}(z)\big\|^2_{L^2(Y)}\right\}^{1/2} \tag{2.31}$$

is finite. Here, $R^s(z) \in L^s_{\mathrm{cl}}(Y;\Gamma_{(n+1)/2-\delta})$ is an *order-reducing family*, i.e., $R^s(z)$ is parameter-dependent elliptic and $R^s(z)\colon H^r(Y) \to H^{r-s}(Y)$ is an isomorphism for some $r \in \mathbb{R}$ (and then for all $r \in \mathbb{R}$) and all $z \in \Gamma_{(n+1)/2-\delta}$. For instance, if $f(z) \in \mathcal{M}^s_{\mathcal{O}}(Y)$ is elliptic and the line $\Gamma_{(n+1)/2-\delta}$ is free of poles of $f(z)^{-1}$, then $f(z)$ is such an order-reduction. We will employ this observation in the next section when defining cone Sobolev spaces with asymptotics.

2.4.2 Cone Sobolev spaces with asymptotics

Let s, $\delta \in \mathbb{R}$, $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$. By Theorem 2.42, there is an elliptic Mellin symbol $h^s_P(z) \in \mathcal{M}^s_{\mathcal{O}}(Y)$ such that the line $\Gamma_{(n+1)/2-\delta}$ is free of poles of $h^s_P(z)^{-1}$ and $L^\delta_{\mathfrak{S}^s} = P$ for $\mathfrak{S}^s = \big\{h^s_P(z), 0, 0, \dots\big\} \in \mathrm{Symb}^s_M(Y)$.

Definition 2.43. Let s, $\delta \in \mathbb{R}$, $\vartheta \geq 0$, and $P \in \underline{\mathrm{As}}^\delta(Y)$.
(a) For $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$, the space $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ consists of all functions $u \in \mathcal{H}^{s,\delta}(X)$ such that $M_{t\to z}\{\omega u\}(z)$, which is *a priori* holomorphic in $\big\{z \in \mathbb{C};\, \mathrm{Re}\, z > (n+1)/2-\delta\big\}$ taking values in $H^s(Y)$, possesses a meromorphic continuation to the half-space $\big\{z \in \mathbb{C};\, \mathrm{Re}\, z > (n+1)/2-\delta-\vartheta\big\}$, moreover,

$$h^s_P(z)M_{t\to z}\{\omega u\}(z) \in \mathcal{A}\big(\{z \in \mathbb{C}; \mathrm{Re}\, z > (n+1)/2-\delta-\vartheta\}; L^2(Y)\big)$$

and the expression

$$\sup_{\delta<\delta'<\delta+\vartheta}\left\{\frac{1}{2\pi i}\int_{\Gamma_{(n+1)/2-\delta'}}\left\|h_P^s(z)M_{t\to z}\{\omega u\}(z)\right\|^2_{L^2(Y)}\,dz\right\}^{1/2} \tag{2.32}$$

is finite.
(b) For a general $P \in \underline{\mathrm{As}}^\delta(Y)$, represented as the join $P = \bigvee_{\iota\in\mathcal{I}} P_\iota$ for a bounded family $\{P_\iota\}_{\iota\in\mathcal{I}} \subset \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$, we define $\mathcal{H}^{s,\delta}_{P,\vartheta}(X) = \sum_{\iota\in\mathcal{I}}\mathcal{H}^{s,\delta}_{P_\iota,\vartheta}(X)$.

It is readily seen that Definition 2.43 (a) is independent of the choice of the Mellin symbol $h_P^s(z)$. Moreover, under the condition that (2.32) is finite the limit

$$h_P^s(z)M_{t\to z}\{\omega u\}(z)\big|_{z=(n+1)/2-\delta'+i\tau} \to w(\tau) \quad \text{as } \delta' \to \delta+\vartheta-0$$

exists in $L^2(\mathbb{R}_\tau; L^2(Y))$. Thus, $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ is a Hilbert space with the norm

$$\|u\|_{\mathcal{H}^{s,\delta}_{P,\vartheta}(X)} = \left\{\|w\|^2_{L^2(\mathbb{R}_\tau;L^2(Y))} + \|u\|^2_{\mathcal{H}^{s,\delta}(X)}\right\}^{1/2}. \tag{2.33}$$

Definition 2.43 (b) is justified by Proposition 2.37 (b), since we obviously have $\mathcal{H}^{s,\delta}_{P,\vartheta}(X) = \mathcal{H}^{s,\delta+\vartheta}(X)$ for $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$ and $\delta_P > \delta+\vartheta$. Again, this definition is seen to be independent of the choice of the representing family $\{P_\iota\}_{\iota\in\mathcal{I}} \subset \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$, and it also yields a Hilbert space structure for $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$.

PROPOSITION 2.44. *Let s, $\delta \in \mathbb{R}$, $\vartheta \ge 0$, and $P \in \underline{\mathrm{As}}^\delta_{\mathrm{prop}}(Y)$. Further, let $\mathfrak{S}^s = \{\mathfrak{s}^{s-j}(z)\ j = 0,1,2,\dots\} \in \mathrm{Symb}^s_M(Y)$ be elliptic with respect to the weight δ and $L^\delta_{\mathfrak{S}^s} = P$, $L^{\delta-s}_{(\mathfrak{S}^s)^{-1}} = \mathcal{O}$. (Condition $L^{\delta-s}_{(\mathfrak{S}^s)^{-1}} = \mathcal{O}$ means that the Mellin symbols $\mathfrak{s}^{s-j}(z)$ are holomorphic when $\operatorname{Re} z > (n+1)/2-\delta$.) Then a function $u \in \mathcal{H}^{s,\delta}(X)$ belongs to the space $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ if and only if $M_{t\to z}\{\omega u\}(z)$ possesses a meromorphic continuation to the half-space $\{z \in \mathbb{C};\ \operatorname{Re} z > (n+1)/2-\delta-\vartheta\}$,*

$$\sum_{j=0}^{M}\mathfrak{s}^{s-j}(z-s+j)M_{t\to z}\{\omega u\}(z-s+j) \\ \in \mathcal{A}\big(\{z\in\mathbb{C}; \operatorname{Re} z > (n+1)/2-\delta+s-\vartheta\}; L^2(Y)\big),$$

and the expression

$$\sup_{\delta<\delta'<\delta+\vartheta}\left\{\frac{1}{2\pi i}\int_{\Gamma_{(n+1)/2-\delta'+s}} \left\|\sum_{j=0}^{M}\mathfrak{s}^{s-j}(z-s+j)M_{t\to z}\{\omega u\}(z-s+j)\right\|^2_{L^2(Y)}dz\right\}^{1/2}$$

is finite. Here, M is any integer larger than ϑ.

Proof. This is an application (of an adapted version) of Witt [18, Proposition 2.6]. Note that $\mathfrak{s}^{s-j}(z-s+j)M_{t\to z}\{\omega u\}(z-s+j) \in \mathcal{A}\big(\{z \in \mathbb{C}; \operatorname{Re} z > (n+1)/2-\delta+s-j\}; L^2(Y)\big)$ so that the condition is actually independent of the choice of the integer $M > \vartheta$. □

For s, $\delta \in \mathbb{R}$, $\vartheta > 0$, and $P \in \underline{\mathrm{As}}^{\delta}(Y)$, we will also employ the spaces

$$\mathcal{H}^{s,\delta}_{P,\vartheta-0}(X) = \bigcap_{\epsilon>0} \mathcal{H}^{s,\delta}_{P,\vartheta-\epsilon}(X). \tag{2.34}$$

These space $\mathcal{H}^{s,\delta}_{P,\vartheta-0}(X)$ are Fréchet-Hilbert spaces, i.e., Fréchet spaces whose topology is given by a countable family of Hilbert semi-norms. We will also use notations like

$$\mathcal{H}^{\infty,\delta}_{P,\vartheta}(X) = \bigcap_{s\in\mathbb{R}} \mathcal{H}^{s,\delta}_{P,\vartheta}(X), \qquad \mathcal{H}^{-\infty,\delta}_{P,\vartheta}(X) = \bigcup_{s\in\mathbb{R}} \mathcal{H}^{s,\delta}_{P,\vartheta}(X),$$
$$\mathcal{H}^{s,\delta}_{P,\vartheta+0}(X) = \bigcup_{\epsilon>0} \mathcal{H}^{s,\delta}_{P,\vartheta+\epsilon}(X), \qquad \text{etc.}$$

Remark 2.45. In case P is a *strongly discrete asymptotic type*, the spaces $\mathcal{H}^{s,\delta}_{P,\vartheta-0}(X)$ are the function spaces introduced by Schulze [15, Section 2.1.1]. There, the notation $\mathcal{H}^{s,\delta}_{P}(X)_\Delta$ with the half-open interval $\Delta = (-\vartheta, 0]$ has been used. The definition of the function spaces $\mathcal{H}^{s,\delta}_{P}(X)_\Delta$ refers to fixed splitting of coordinates near ∂X and is, in general, not coordinate invariant.

2.4.3 Functional-analytic properties

We list some properties of the function spaces $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$:

Proposition 2.46. *Let s, s', δ, $\delta' \in \mathbb{R}$, $\vartheta \geq 0$, $P \in \underline{\mathrm{As}}^{\delta}(Y)$, $P' \in \underline{\mathrm{As}}^{\delta'}(Y)$, and $\{P_\iota\}_{\iota\in\mathcal{I}} \subset \underline{\mathrm{As}}^{\delta}(Y)$ be a family of asymptotic types. Then*:
(a) $\mathcal{H}^{s,\delta}_{P,0}(X) = \mathcal{H}^{s,\delta}(X)$.
(b) $\mathcal{H}^{s,\delta}_{P,\vartheta}(X) = \mathcal{H}^{s,\delta-a}_{P,\vartheta+a}(X)$ *for any* $a > 0$.
(c) $\mathcal{H}^{s,\delta}_{\mathcal{O},\vartheta}(X) = \mathcal{H}^{s,\delta+\vartheta}(X)$.
(d) *We have*

$$\begin{aligned}\mathcal{H}^{s,\delta}_{P,\vartheta}(X) &= \mathcal{H}^{s,\delta}_{\mathcal{O},\vartheta}(X) \\ &\oplus \Bigg\{\omega(t) \sum_{\substack{p\in V,\\ \operatorname{Re} p>(n+1)/2-\delta-\vartheta}} \; \sum_{k+l=m_p-1} \frac{(-1)^k}{k!}\, t^{-p} \log^k t\, \phi^{(p)}_l(y); \\ &\qquad\qquad \Phi(p) = (\phi^{(p)}_0, \dots, \phi^{(p)}_{m_p-1}) \text{ for some } \Phi \in J\Bigg\},\end{aligned}$$

where $J \subset \mathcal{E}_V(Y)$ is the linear subspace representing the asymptotic type P, provided that $\operatorname{Re} p \neq (n+1)/2-\delta-\vartheta$ *holds for all* $p \in V$.

(e) *We have* $\mathcal{H}^{s,\delta}_{P,\vartheta}(X) \subseteq \mathcal{H}^{s',\delta'}_{P',\vartheta'}(X)$ *if and only if* $s \geq s'$, $\delta + \vartheta \geq \delta' + \vartheta'$, *and* $P \preccurlyeq P'$ *up to the conormal order* $\delta' + \vartheta'$.
(f) $\mathcal{H}^{s,\delta}_{\bigwedge_{\iota\in\mathcal{I}} P_\iota,\vartheta}(X) = \bigcap_{\iota\in\mathcal{I}} \mathcal{H}^{s,\delta}_{P_\iota,\vartheta}(X)$ *if the family* $\{P_\iota\}_{\iota\in\mathcal{I}}$ *is non-empty.*
(g) $\mathcal{H}^{s,\delta}_{\bigvee_{\iota\in\mathcal{I}} P_\iota,\vartheta}(X) = \sum_{\iota\in\mathcal{I}} \mathcal{H}^{s,\delta}_{P_\iota,\vartheta}(X)$ *if the family* $\{P_\iota\}_{\iota\in\mathcal{I}}$ *is bounded* (*where the sum sign stands for the non-direct sum of Hilbert spaces*);
(h) $C^\infty_P(X) = \bigcap_{s\in\mathbb{R},\,\vartheta\geq 0} \mathcal{H}^{s,\delta}_{P,\vartheta}(X)$.
(i) $C^\infty_P(X)$ *is dense in* $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$.

Proof. The proofs of (a) to (i) are straightforward. □

From (e) we get, in particular, $\mathcal{H}^{s,\delta}_{P,\vartheta}(X) = \mathcal{H}^{s',\delta'}_{P',\vartheta'}(X)$ if and only if $s = s'$, $\delta + \vartheta = \delta' + \vartheta'$, and $P = P'$ up to the conormal order $\delta + \vartheta$. (b) and also (c), in view of (a), are special cases.

Proposition 2.47. *For* $\delta \in \mathbb{R}$, $P \in \underline{\mathrm{As}}^\delta(Y)$, *and any* $a \in \mathbb{R}$, *the family* $\left\{\mathcal{H}^{s,\delta}_{P,s-a}(X);\ s \geq a\right\}$ *of Hilbert spaces forms an interpolation scale with respect to the complex interpolation method.*

Proof. This is immediate from the definition. □

Proposition 2.48. *The spaces* $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ *are invariant under coordinate changes, where this has to be understood in the sense of* Proposition 2.32.

Proof. Basically, this follows from the invariance of the spaces $C^\infty_P(X)$ under coordinate changes, where the latter is just a reformulation of the fact that the asymptotic types in $\underline{\mathrm{As}}^\delta(Y)$ are coordinate invariant. □

2.4.4 Mapping properties and elliptic regularity

We finally take the step from the algebra of complete conormal symbols to elliptic Fuchsian differential operators and their parametrices. These parametrices are cone pseudodifferential operators, where for the latter we refer to Schulze [16, Chapter 2]. While for general cone pseudodifferential operators, there might be a difference between the conormal asymptotics produced on the level of complete conormal symbols and operators, respectively — due to the appearance of so-called singular Green operators — for *Fuchsian differential operators* this does not happen.
In *cone pseudodifferential calculus*, one encounters operators of the form $\omega(t)t^{-\mu}\operatorname{op}_M^{(n+1)/2-\delta}(h)\,\tilde\omega(t)$, where $h(t,z) \in C^\infty(\overline{\mathbb{R}}_+; \mathcal{M}^\mu_{\mathrm{as}}(Y))$. Here,

$$\operatorname{op}_M^{(n+1)/2-\delta}(h(t,z))u = \frac{1}{2\pi i}\int_{\Gamma_{(n+1)/2-\delta}} t^{-z} h(t,z)\tilde u(z)\,dz \tag{2.35}$$

is a pseudodifferential operator, whose definition is based on the Mellin transformation instead of the Fourier transformation. The mapping properties of these operators in the spaces $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ are as follows:

Proposition 2.49. *Let $h(t,z) \in C^\infty(\overline{\mathbb{R}}_+; \mathcal{M}^\mu_{\mathrm{as}}(Y))$ and assume that the line $\Gamma_{(n+1)/2-\delta}$ is free of poles of $\partial^j h(0,z)/\partial t^j$ for all $j = 0, 1, 2, \ldots$ Then, for all $P \in \underline{\mathrm{As}}^\delta(Y)$, $s \in \mathbb{R}$, $\vartheta \geq 0$,*

$$\omega(t) t^{-\mu} \operatorname{op}_M^{(n+1)/2-\delta}(h)\, \tilde{\omega}(t) \colon \mathcal{H}^{s,\delta}_{P,\vartheta}(X) \to \mathcal{H}^{s-\mu,\delta-\mu}_{Q,\vartheta}(X),$$

where $\omega(t)$, $\tilde{\omega}(t)$ are cut-off functions, $\mathfrak{S}^\mu = \left\{ \frac{1}{j!} \frac{\partial^j h}{\partial t^j}(0,z);\, j = 0, 1, 2, \ldots \right\} \in \mathrm{Symb}^\mu_M(Y)$, and $Q = \mathcal{Q}^{\delta-\mu}(P, \mathfrak{S}^\mu) \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$.

Proof. The previous definitions are made to let this result hold. $\square$

Notation. Proposition 2.49 implies that, given a cone pseudodifferential operator A in Schulze's cone calculus $\mathcal{C}^\mu(X, (\delta, \delta-\mu, (-\infty, 0]))$, see Schulze [16, Chapter 2] again, for each $P \in \underline{\mathrm{As}}^\delta(Y)$, there is a $Q \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$ such that, for all $s \in \mathbb{R}$, $\vartheta \geq 0$,

$$A \colon \mathcal{H}^{s,\delta}_{P,\vartheta}(X) \to \mathcal{H}^{s-\mu,\delta-\mu}_{Q,\vartheta}(X). \tag{2.36}$$

Given $P \in \underline{\mathrm{As}}^\delta(Y)$, the minimal such asymptotic type $Q \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$, that exists by virtue of Proposition 2.28 (a) and Proposition 2.46 (f), is denoted by $\mathcal{Q}^{\delta-\mu}(P; A)$. If A is elliptic, given $Q \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$, the minimal asymptotic type $P \in \underline{\mathrm{As}}^\delta(Y)$ such that, for all $s \in \mathbb{R}$, $\vartheta \geq 0$, $u \in \mathcal{H}^{-\infty,\delta}(X)$, $Au \in \mathcal{H}^{s-\mu,\delta-\mu}_{Q,\vartheta}(X)$ implies $u \in \mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ is denoted by $\mathcal{P}^\delta(Q; A)$.
We shall employ this *push-forward* notation also if more than one operator A is involved, i.e., $\mathcal{Q}^{\delta-\mu}(P; A_1, \ldots, A_m)$ denotes the minimal asymptotic type Q for which $A_j \colon \mathcal{H}^{s,\delta}_{P,\vartheta}(X) \to \mathcal{H}^{s-\mu,\delta-\mu}_{Q,\vartheta}(X)$ for $1 \leq j \leq m$.

Theorem 2.50. *For $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$, $P \in \underline{\mathrm{As}}^\delta(Y)$, $Q \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$, we have $\mathcal{Q}^{\delta-\mu}(P; A) = \mathcal{Q}^{\delta-\mu}(P; \mathfrak{S}^\mu)$, where $\mathfrak{S}^\mu = \{\sigma^{\mu-j}_M(A)(z);\, j = 0, 1, \ldots\} \in \mathrm{Symb}^\mu_M(Y)$, as well as, in case A is elliptic, $\mathcal{P}^\delta(Q; A) = \mathcal{Q}^\delta(Q; (\mathfrak{S}^\mu)^{-1})$.*

Proof. In fact, $\mathcal{Q}^{\delta-\mu}(P; A) = \mathcal{Q}^{\delta-\mu}(P; \mathfrak{S}^\mu)$ follows from Proposition 2.49.
Furthermore, it is known that formal asymptotic solutions $u \in C^\infty_{\mathrm{as}}(X)$ to the equation $Au = f$ for $f \in C^\infty_R(X)$ and any $R \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$ can be constructed, see, e.g. Melrose [13, Lemma 5.13]. More precisely, it can be shown that there is a right parametrix B to A, $B \colon \mathcal{H}^{s-\mu,\delta-\mu}(X) \to \mathcal{H}^{s,\delta}(X)$ for all $s \in \mathbb{R}$, such that

$$AB = I + R, \quad R \colon \mathcal{H}^{-\infty,\delta-\mu}(X) \to C^\infty_{\mathcal{O}}(X),$$

i.e., R is smoothing over X° and flattening to infinite order near ∂X. In fact, $B \in \mathcal{C}^{-\mu}(X, (\delta-\mu, \delta, (-\infty, 0]))$ and, in particular, $B \in L^{-\mu}_{\mathrm{cl}}(X^\circ)$.
Now let $BA = I + R_0$. Obviously, R_0 is smoothing over X° such that $R_0 \colon \mathcal{H}^{s,\delta}(X) \to \mathcal{H}^{\infty,\delta-\mu}(X)$ for any $s \in \mathbb{R}$. Furthermore, $A(I + R_0) = ABA = (I + R)A$ so that

$$AR_0 = RA.$$

We conclude that $R_0\colon \mathcal{H}^{s,\delta}(X) \to C^\infty_{P_0}(X)$, where $P_0 = \mathcal{Q}^\delta(\mathcal{O};(\mathfrak{S}^\mu)^{-1})$. Hence, for $u \in \mathcal{H}^{-\infty,\delta}(X)$, $Au = f \in \mathcal{H}^{s-\mu,\delta-\mu}_{Q,\vartheta}(X)$, we get

$$u = Bf - R_0 u \in \mathcal{H}^{s,\delta}_{P,\vartheta}(X),$$

where $P = \mathcal{Q}^\delta(Q;(\mathfrak{S}^\mu)^{-1})$. Thus $\mathcal{P}^\delta(Q;A) = \mathcal{Q}^\delta(Q;(\mathfrak{S}^\mu)^{-1})$ as claimed. See also Witt [20, Remark after Proposition 5.5]. □

Notation. For $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$, $\mathcal{Q}^{\delta-\mu}(P;A)$ is even independent of $\delta \in \mathbb{R}$ in view of the holomorphy of the conormal symbols $\sigma_M^{\mu-j}(A)(z)$ for $j = 0,1,2,\dots$ In this case, we simply write $\mathcal{Q}(P;A) = \mathcal{Q}^{\delta-\mu}(P;A)$.

Proposition 2.51. *Let $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$ be elliptic. Then there is an order-preserving bijection*

$$\left\{P \in \underline{\mathrm{As}}^\delta(Y);\ P \succcurlyeq L^\delta_{\mathfrak{S}^\mu}\right\} \to \underline{\mathrm{As}}^{\delta-\mu}(Y), \quad P \mapsto \mathcal{Q}(P;A), \tag{2.37}$$

with its inverse given by $Q \mapsto \mathcal{P}^\delta(Q;A)$. In particular, $L^\delta_{\mathfrak{S}^\mu}$ is mapped to the empty asymptotic type, $\mathcal{O}$.

Proof. This is implied by Proposition 2.31 and Theorem 2.50. Note that $L^{\delta-\mu}_{(\mathfrak{S}^\mu)^{-1}} = \mathcal{O}$, since the $\sigma_M^{\mu-j}(A)(z)$ for $j = 0,1,2,\dots$ are holomorphic. □

Eventually, we have the following *locality principle*:

Proposition 2.52. *Let $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$ be elliptic, Q_0, $Q_1 \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$, and $P_0 = \mathcal{P}^\delta(Q_0;A)$, $P_1 = \mathcal{P}^\delta(Q_1;A)$. Then, for any $\vartheta > 0$, $P_0 = P_1$ up to the conormal order $\delta + \vartheta$ if $Q_0 = Q_1$ up to the conormal order $\delta - \mu + \vartheta$.*

Proof. This follows from $P_0 = \mathcal{Q}^\delta(Q_0;(\mathfrak{S}^{-\mu})^{-1})$, $P_1 = \mathcal{Q}^\delta(Q_1;(\mathfrak{S}^{-\mu})^{-1})$, where $\mathfrak{S}^\mu = \{\sigma_M^{\mu-j}(A)(z);\ j \in \mathbb{N}\} \in \mathrm{Ell\,Symb}^\mu_M(Y)$. □

Remark 2.53. Combined with Theorem 2.30, Theorem 2.50 shows that each solution $u \in C^{\infty,\delta}_{\mathrm{as}}(X)$ to the equation $Au = f \in C^\infty_{\mathcal{O}}(X)$, where $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$ is elliptic, can be written over finite weight intervals as a finite sum of functions of the form (2.16) modulo the corresponding flat class, where the Φ are taken from a characteristic basis of the linear subspace of $\mathcal{E}^\delta(Y)$ representing $\mathcal{P}^\delta(\mathcal{O};A)$. If $\Phi(p) = (\phi_0,\dots,\phi_{m-1})$ for such a vector Φ, where $p = \gamma(\Phi)$, then we say that A admits *an asymptotic series starting with the term $t^{-p}\log^{m-1} t\,\phi_0$*. Since this is then the most singular term (when $\gamma(\Phi)$ is highest possible), if it coefficient can be shown to vanish, then the whole series must vanish, up to the next appearance of a starting term for another asymptotic series.

3 Applications to semilinear equations

In this section, Theorem 1.1 is proved. To this end, multiplicatively closable and multiplicatively closed asymptotic types are investigated in Section 3.1. This allows the derivation of results concerning the action of nonlinear superposition operators on cone Sobolev spaces with asymptotics. In Section 3.2, the general scheme for establishing results of the type of Theorem 1.1 is established. This scheme is specified to multiplicatively closable asymptotic types in Section 3.3, then completing the proof of Theorem 1.1.

3.1 Multiplicatively closed asymptotic types

Here, we investigate multiplicative properties of asymptotic types and the behavior of cone Sobolev spaces $\mathcal{H}^{s,\delta}_{P,\vartheta}(X)$ under the action of nonlinear superposition.

Notation. In connection with pointwise multiplication, it is useful to employ the following notation:

$$H^s_{P,\vartheta}(X) = \begin{cases} \mathcal{H}^{s,\delta}_{P,\delta_P-\delta+\vartheta}(X) & \text{if } \vartheta \geq 0, \\ \mathcal{H}^{s,\delta_P+\vartheta}(X) & \text{otherwise,} \end{cases}$$

where $P \in \underline{\mathrm{As}}^\delta(Y)$, $P \neq \mathcal{O}$, and $\delta < \delta_P$ in the first line. (Proposition 2.46 (b) shows that this definition is independent of the choice of δ.) Thus, starting from δ_P, the conormal order is improved by ϑ upon allowing asymptotics of type P. Similarly for $H^s_{P,\vartheta-0}(X)$.
Furthermore, we write $\{\vartheta\}$ if we mean either ϑ or $\vartheta - 0$. For instance, $\{\vartheta\} \geq 0$ means $\vartheta \geq 0$ if $\{\vartheta\} = \vartheta$ and $\vartheta > 0$ if $\{\vartheta\} = \vartheta - 0$.

3.1.1 Multiplication of asymptotic types

The result admitting nonlinear superposition for function spaces with asymptotics is stated first:

Lemma 3.1. *Given $P \in \underline{\mathrm{As}}(Y)$, $Q \in \underline{\mathrm{As}}(Y)$, there is a minimal asymptotic type, $P \circ Q \in \underline{\mathrm{As}}(Y)$, such that*

$$C^\infty_P(X) \times C^\infty_Q(X) \to C^\infty_{P\circ Q}(X), \ \ (u, v) \mapsto uv. \tag{3.1}$$

Proof. Suppose that the asymptotic types P, Q are represented by subspaces $J \subset \mathcal{E}_V(Y)$ and $K \subset \mathcal{E}_W(Y)$, respectively, for suitable $V, W \in \mathcal{C}$. Then the asymptotic type $P \circ Q$ is carried by the set $V + W$, and it is represented by the linear subspace of $\mathcal{E}_{V+W}(Y)$ consisting of all $\Theta \in \mathcal{E}_{V+W}(Y)$ for which there are $\Phi \in J$, $\Psi \in K$ such that $\Theta(r) = \sum_{\substack{p+q=r,\\ p\in V,\, q\in W}} \Phi(p) \times \Psi(q)$ holds for all $r \in V + W$.

Here,

$$\Phi \times \Psi = \Bigg(\binom{m+n}{m} \phi_0\psi_0, \binom{m+n-1}{m} \phi_0\psi_1 + \binom{m+n-1}{m-1} \phi_1\psi_0, \\ \binom{m+n-2}{m} \phi_0\psi_2 + \binom{m+n-2}{m-1} \phi_1\psi_1 + \binom{m+n-2}{m-2} \phi_0\psi_2, \\ \dots, \binom{1}{1} \phi_{m-1}\psi_n + \binom{1}{0} \phi_m\psi_{n-1}, \binom{0}{0} \phi_m\psi_n \Bigg)$$

for $\Phi = (\phi_0, \phi_1, \dots, \phi_m)$, $\Psi = (\psi_0, \psi_1, \dots, \psi_n) \in [C^\infty(Y)]^\infty$. For this, see (2.16). Note that $T(\Phi \times \Psi) = (T\Phi) \times \Psi + \Phi \times (T\Psi)$ and, for $\Phi \in \mathcal{E}_{V_p}(Y)$, $\Psi \in \mathcal{E}_{V_q}(Y)$, we have $\Phi \times \Psi \in \mathcal{E}_{V_{p+q}}(Y)$ showing that the linear subspace of $\mathcal{E}_{V+W}(Y)$ described above actually represents an asymptotic type. □

The multiplication of asymptotic types possesses a unit, denoted by $\mathbf{1}$, that is represented by the space $\operatorname{span}\{(1)\} \subset \mathcal{E}_{\{0\}}(Y)$, where 1 is the function identically 1 on Y.

Definition 3.2. An asymptotic type $Q \in \underline{\mathrm{As}}(Y)$ is called *multiplicatively closed* if $Q \circ Q = Q$. An asymptotic type $Q \in \underline{\mathrm{As}}(Y)$ is called *multiplicatively closable* if it is dominated by a multiplicatively closed asymptotic type. In this case, the minimal multiplicatively closed asymptotic type dominating Q is called the *multiplicative closure* of Q and is denoted by $\widetilde{Q}$.

From the proof of Lemma 3.1,

$$\delta_{P \circ Q} \geq \delta_P + \delta_Q - (n+1)/2, \tag{3.2}$$

where equality holds if $P = Q$ or if $\dim Y = 0$. Especially, $\delta_Q = (n+1)/2$ if Q is multiplicatively closed and $\delta_Q \geq (n+1)/2$ if Q is multiplicatively closable. Furthermore, it is also seen $Q \succcurlyeq \mathbf{1}$ for any multiplicatively closed asymptotic type Q, see also Lemma 3.4 below.

3.1.2 The class $\underline{\mathrm{As}}^\sharp(Y)$ of multiplicatively closable asymptotic types

We study the class of asymptotic types that belong to bounded functions. It turns out that this class is intimately connected to the multiplication of asymptotic types.

Definition 3.3. (a) The class $\underline{\mathrm{As}}^b(Y)$ of *bounded asymptotic types* consists of all asymptotic types $Q \in \underline{\mathrm{As}}(Y)$ for which $\delta_Q \geq (n+1)/2$. Equivalently, a bounded asymptotic type Q is represented by a subspace $J \subset \mathcal{E}_V(Y)$ for some $V \in \mathcal{C}$, where $V \subset \{z \in \mathbb{C};\, \mathrm{Re}\, z \leq 0\}$.
(b) The class $\underline{\mathrm{As}}^\sharp(Y)$ consists of all bounded asymptotic types Q represented by a subspace $J \subset \mathcal{E}(Y)$ such that $J_0 \subseteq \operatorname{span}\{(1)\}$ and $J_p = \{0\}$ for $\mathrm{Re}\, p = 0$, $p \neq 0$.

Lemma 3.4. *For $Q \in \underline{\mathrm{As}}(Y)$, the following conditions are equivalent:*
(a) *Q is multiplicatively closable*;
(b) *the join $\bigvee_{k\geq 1} Q^k$ does exist, where $Q^k = \underbrace{Q \circ \cdots \circ Q}_{k \text{ times}}$ is the k-fold product*;
(c) *$Q \in \underline{\mathrm{As}}^\sharp(Y)$.*
In case (a) *to* (c) *are fulfilled, we have $\widetilde{Q} = \mathbf{1} \vee \bigvee_{k\geq 1} Q^k$.*

Proof. (a) and (b) are obviously equivalent. Moreover, (c) implies (b).
It remains to show that (a) also implies (c). If Q is multiplicatively closable, then $\widetilde{Q}$ exists and $\delta_{\widetilde{Q}} = (n+1)/2$. In particular, $\widetilde{Q} \in \underline{\mathrm{As}}^b(Y)$. Let $\widetilde{Q}$ be represented by $J \subset \mathcal{E}_V(Y)$ for some $V \in \mathcal{C}$, $V \subset \{z \in \mathbb{C};\, \mathrm{Re}\, z \leq 0\}$. Suppose that $\phi \in J_p$ for $p \in \mathbb{C}$, $\mathrm{Re}\, p = 0$, where $\phi \neq 0$. We immediately get $\phi^l \in J_{lp}$ for any $l \in \mathbb{N}$, $l \geq 1$. For $p \neq 0$, we obtain the contradiction $\{lp;\, l \in \mathbb{N}\} \subseteq V \in \mathcal{C}$. For $p = 0$ and ϕ not being constant, we obtain a contradiction to the fact that $\dim J_0 < \infty$. Thus, $\widetilde{Q} \in \underline{\mathrm{As}}^\sharp(Y)$ and, therefore, $Q \in \underline{\mathrm{As}}^\sharp(Y)$. □

Lemma 3.5. *For each $Q \in \underline{\mathrm{As}}(Y)$, there are asymptotic types $Q^b \in \underline{\mathrm{As}}^b(Y)$ and $Q^\sharp \in \underline{\mathrm{As}}^\sharp(Y)$ which are maximal among all asymptotic types possessing the property*

$$Q^b \preccurlyeq Q \text{ and } Q^\sharp \preccurlyeq Q, \tag{3.3}$$

respectively. In particular, $Q^\sharp \preccurlyeq Q^b$.

Proof. The proof is straightforward. □

3.1.3 Nonlinear superposition

We investigate expressions like $F(x, v(x))$, where $F(x,\nu) \in C^\infty_R(X \times \mathbb{R})$ for some $R \in \underline{\mathrm{As}}(Y)$ and $v \in H^s_{Q,\vartheta}(X) \cap L^\infty(X)$ with $s \geq 0$, $\vartheta > 0$, and $Q \in \underline{\mathrm{As}}^\sharp(Y)$. For later reference, we recall the following facts:

Proposition 3.6. (a) *For $s > (n+1)/2$, $0 \leq s' \leq s$, $\gamma, \delta \in \mathbb{R}$, pointwise multiplication induces a bilinear continuous map*

$$\mathcal{H}^{s,\gamma}(X) \times \mathcal{H}^{s',\delta}(X) \to \mathcal{H}^{s',\gamma+\delta-(n+1)/2}(X). \tag{3.4}$$

(b) *For $s, \delta \in \mathbb{R}$, $\mathcal{H}^{s,\delta}(X) \subset L^\infty(X)$ if and only if $s > (n+1)/2$, $\delta \geq (n+1)/2$.*
(c) *For $s \geq 0$, $\gamma, \delta \geq (n+1)/2$, pointwise multiplication induces a bilinear continuous map*

$$\big(\mathcal{H}^{s,\gamma}(X) \cap L^\infty(X)\big) \times \big(\mathcal{H}^{s,\delta}(X) \cap L^\infty(X)\big) \to \mathcal{H}^{s,\gamma+\delta-(n+1)/2}(X) \cap L^\infty(X).$$

(d) *For $s \geq 0$, $\delta \in \mathbb{R}$, $p \in \mathbb{C}$, $c(y) \in C^\infty(Y)$, the multiplication operator*

$$\omega(t)t^{-p}c(y)\colon \mathcal{H}^{s,\delta}(X) \to \mathcal{H}^{s,\delta-\mathrm{Re}\, p}(X),$$

where $\omega(t)$ is a cut-off function, is continuous.

(e) *For* $s \geq 0$, $v_1, \dots, v_K \in \big(1 + \mathcal{H}^{s,(n+1)/2}(X)\big) \cap L^\infty(X)$, *and* $F \in C^\infty(\mathbb{R}^K)$, *we have*

$$F(v_1, \dots, v_K) \in \big(1 + \mathcal{H}^{s,(n+1)/2}(X)\big) \cap L^\infty(X). \tag{3.5}$$

The map $\big(\big(1 + \mathcal{H}^{s,(n+1)/2}(X)\big) \cap L^\infty(X)\big)^K \to \big(1 + \mathcal{H}^{s,(n+1)/2}(X)\big) \cap L^\infty(X)$ *induced by* (3.5) *is continuous and sends bounded sets to bounded sets.*

Proof. A proof of (3.4) in case $s' = s$ has been supplied by WITT [19, Lemma 2.7] using a result of DAUGE [2, Theorem (AA.3)]. The other proofs are similar. □

Remark 3.7. Property (d) fails if logarithms appear and has to be replaced by

$$\omega(t) t^{-p} \log^k t\, c(y) \colon \mathcal{H}^{s,\delta}(X) \to \mathcal{H}^{s,\delta - \operatorname{Re} p - 0}(X)$$

is continuous when $k \in \mathbb{N}$, $k \geq 1$.

First, Lemma 3.1 is sharpened:

PROPOSITION 3.8. *For* $s > (n+1)/2$, $0 \leq s' \leq s$, $\vartheta > 0$, *and* P, $Q \in \underline{\mathrm{As}}(Y)$, *pointwise multiplication induces a bilinear continuous map*

$$H^s_{P,\vartheta-0}(X) \times H^{s'}_{Q,\vartheta-0}(X) \to H^{s'}_{P \circ Q,\vartheta-0}(X). \tag{3.6}$$

Proof. Let $u \in H^s_{P,\vartheta-0}(X)$, $v \in H^{s'}_{Q,\vartheta-0}(X)$. Then $u = u_0 + u_1$, $v = v_0 + v_1$, where

$$u_0 = \sum_{j=0}^{M} \sum_{k=0}^{m_j} \omega(t) t^{-p_j} \log^k t\, c_{jk}(y), \quad v_0 = \sum_{j'=0}^{N} \sum_{k'=0}^{n_{j'}} \omega(t) t^{-q_{j'}} \log^{k'} t\, d_{j'k'}(y), \tag{3.7}$$

$\omega(t)$ is a cut-off function, the sequences $\{(p_j, m_j, c_{jk})\}$, $\{(q_{j'}, n_{j'}, d_{j'k'})\}$ are given by the asymptotic types P and Q, respectively, according to Definition 2.18, and M, N are chosen so that $u_1 \in \mathcal{H}^{s,\delta_P+\vartheta-0}(X)$, $v_1 \in \mathcal{H}^{s',\delta_Q+\vartheta-0}(X)$. Since $u_0 \in \mathcal{H}^{\infty,\delta_P-0}(X)$, $v_0 \in \mathcal{H}^{\infty,\delta_Q-0}(X)$, we obtain

$$uv = u_0 v_0 + u_1 v_0 + u_0 v_1 + u_1 v_1,$$

where $u_1 v_0 + u_0 v_1 + u_1 v_1 \in \mathcal{H}^{s',\delta_{P\circ Q}+\vartheta-0}(X)$ by (3.4) and

$$u_0 v_0 = \sum_{j,j'=0}^{M,N} \sum_{k,k'=0}^{m_j,n_{j'}} \omega^2(t) t^{-(p_j+q_{j'})} \log^{k+k'} t\, c_{jk}(y) d_{j'k'}(y) \in H^\infty_{P\circ Q,\vartheta-0}(X),$$

for $\omega^2(t)$ is a cut-off function and the sequence

$$\Big\{\Big(r_{j''}, o_{j''}, \sum_{p_j+q_{j'}=r_{j''}} \sum_{k+k'=k''} c_{jk} d_{j'k'}\Big)\Big\},$$

where $o_{j''} = \max\{m_j + n_{j'};\, p_j + q_{j'} = r_{j''}\}$, is associated with an asymptotic type that equals $P \circ Q$ up to the conormal order $\delta_{P\circ Q} + \vartheta - 0$. This immediately gives $uv \in H^{s'}_{P\circ Q,\vartheta-0}(X)$. □

The significance of the class $\underline{\mathrm{As}}^b(Y)$ is uncovered by the next result:

Proposition 3.9. *For $s \geq 0$, $\delta \in \mathbb{R}$, $\delta + \{\vartheta\} \geq (n+1)/2$, and $Q \in \underline{\mathrm{As}}^\delta(Y)$,*

$$\mathcal{H}^{s,\delta}_{Q,\{\vartheta\}}(X) \cap L^\infty(X) = \mathcal{H}^{s,\delta}_{Q^b,\{\vartheta\}}(X) \cap L^\infty(X). \tag{3.8}$$

Proof. Let $u \in \mathcal{H}^{s,\delta}_{Q,\{\vartheta\}}(X) \cap L^\infty(X)$ and write

$$u(x) = \sum_{j=0}^{M} \sum_{k=0}^{m_j} \omega(t) t^{-p_j} \log^k t\, c_{jk}(y) + u_1(x),$$

where the sequences $\{(p_j, m_j, c_{jk})\}$ is given by the asymptotic type Q and M is chosen so that $u_1 \in \mathcal{H}^{s,(n+1)/2-0}(X)$. Since $u \in L^\infty(X) \subset \mathcal{H}^{0,(n+1)/2-0}(X)$, we get that $\sum_{j=0}^{M} \sum_{k=0}^{m_j} \omega(t) t^{-p_j} \log^k t\, c_{jk}(y) \in \mathcal{H}^{0,(n+1)/2-0}(X)$ which implies $c_{jk}(y) = 0$ for $\mathrm{Re}\, p_j > 0$. Thus $u \in \mathcal{H}^{s,\delta}_{Q^b,\vartheta}(X)$. □

Lemma 3.10. *For $s \geq 0$, $\vartheta > 0$, and P, $Q \in \underline{\mathrm{As}}^b(Y)$, pointwise multiplication induces a bilinear continuous map*

$$\big(H^s_{P,\vartheta-0}(X) \cap L^\infty(X)\big) \times \big(H^s_{Q,\vartheta-0}(X) \cap L^\infty(X)\big) \to H^s_{P\circ Q,\vartheta-0}(X) \cap L^\infty(X).$$

Proof. Represent $u = u_0 + u_1 \in H^s_{P,\vartheta-0}(X) \cap L^\infty(X)$, $v = v_0 + v_1 \in H^s_{Q,\vartheta-0}(X) \cap L^\infty(X)$ as in the proof of Proposition 3.8. Since u_0, $v_0 \in L^\infty(X)$ due to the assumption P, $Q \in \underline{\mathrm{As}}^b(Y)$, we get that $u_1 \in \mathcal{H}^{s,\delta_P+\vartheta-0}(X) \cap L^\infty(X)$, $v_1 \in \mathcal{H}^{s,\delta_Q+\vartheta-0}(X) \cap L^\infty(X)$ and, therefore, $u_1v_0 + u_0v_1 + u_1v_1 \in \mathcal{H}^{s,\delta_{P\circ Q}+\vartheta-0}(X) \cap L^\infty(X)$ in view of Proposition 3.6 (c). The assertion follows. □

A more precise statement is possible if P, $Q \in \underline{\mathrm{As}}^\sharp(Y)$:

Lemma 3.11. *For $s \geq 0$, $\vartheta \geq 0$, and P, $Q \in \underline{\mathrm{As}}^\sharp(Y)$ satisfying $P \succcurlyeq \mathbf{1}$, $Q \succcurlyeq \mathbf{1}$, pointwise multiplication induces a bilinear continuous map*

$$\big(H^s_{P,\vartheta}(X) \cap L^\infty(X)\big) \times \big(H^s_{Q,\vartheta}(X) \cap L^\infty(X)\big) \to H^s_{P\circ Q,\vartheta}(X) \cap L^\infty(X). \tag{3.9}$$

Especially, for $s \geq 0$, $\vartheta \geq 0$, and $Q \in \underline{\mathrm{As}}^\sharp(Y)$ being multiplicatively closed, $H^s_{Q,\vartheta}(X) \cap L^\infty(X)$ is an algebra under pointwise multiplication.

Proof. We may assume that $\vartheta > 0$. Write $u = u_0 + u_1 \in H^s_{P,\vartheta}(X) \cap L^\infty(X)$, $v = v_0 + v_1 \in H^s_{Q,\vartheta}(X) \cap L^\infty(X)$ as in the proof of Proposition 3.8, where $u_0 = u_{00} + u_{01}$, $v_0 = v_{00} + v_{01}$, $u_{00} = \omega(t)c_{00}$, and $v_{00} = \omega(t)d_{00}$ with c_{00}, d_{00} being constants and in the expressions for u_{01}, v_{01} only appear exponents with $\mathrm{Re}\, p_j < 0$ and $\mathrm{Re}\, q_{j'} < 0$, respectively. Then

$$u_1v_{01} + u_{01}v_1 + u_1v_1 \in \mathcal{H}^{s,(n+1)/2+\vartheta+0}(X),$$

$u_{00}v \in H^s_{Q,\vartheta}(X) \subseteq H^s_{P\circ Q,\vartheta}(X)$, $uv_{00} \in H^s_{P,\vartheta}(X) \subseteq H^s_{P\circ Q,\vartheta}(X)$, and

$$u_{01}v_{01} \in H^\infty_{P\circ Q,\vartheta+0}(X),$$

which proves the assertion. □

The fact which has actually been used in the last proof is that Proposition 3.6 (d) applies to the function $\omega(t)1$ ($p = 0$, $c(y) \equiv 1$). This is also used in part (b) of the next result:

Lemma 3.12. (a) *Let $s \geq 0$, $\vartheta > 0$, and $R, Q \in \underline{\mathrm{As}}(Y)$. Then pointwise multiplication induces a continuous map*

$$C_R^\infty(X) \times H^s_{Q,\vartheta-0}(X) \to H^s_{R\circ Q,\vartheta-0}(X). \tag{3.10}$$

(b) *If, in addition, $R \in \underline{\mathrm{As}}(Y)$ is so that the multiplicities of its highest singular values are one, i.e., $J_r \subseteq [C^\infty(Y)]^1$ for each $r \in V$, $\operatorname{Re} r = (n+1)/2-\delta_R$, where $J \subset \mathcal{E}_V(Y)$ represents R, then pointwise multiplication induces a continuous map*

$$C_R^\infty(X) \times H^s_{Q,\vartheta}(X) \to H^s_{R\circ Q,\vartheta}(X).$$

Proof. (a) is immediate from Proposition 3.8. To get (b), we argue as in the proof of Lemma 3.11. □

Proposition 3.13. *Let $s \geq 0$, $\vartheta \geq 0$, and $Q \in \underline{\mathrm{As}}^\sharp(Y)$ be multiplicatively closed. Then $v_1, \dots, v_K \in H^s_{Q,\vartheta}(X) \cap L^\infty(X)$ and $F \in C^\infty(\mathbb{R}^K)$ implies that*

$$F(v_1, \dots, v_K) \in H^s_{Q,\vartheta}(X) \cap L^\infty(X). \tag{3.11}$$

Proof. We are allowed to assume that $\vartheta > 0$. Then $v \in H^s_{Q,\vartheta}(X)$ implies that $v\big|_{\partial X}$ is a constant, where $v\big|_{\partial X}$ means the factor in front of t^0 in the asymptotic expansion (1.2) (with u replaced with v) of v as $t \to +0$. Let $\beta_J = v_J\big|_{\partial X}$ for $1 \leq J \leq K$ be these constants. Using Taylor's formula, we obtain

$$\begin{aligned} F(v_1, \dots, v_K) &= \sum_{|\alpha|<N} \frac{1}{\alpha!} (\partial^\alpha F)(\beta_1, \dots, \beta_K)(v_1 - \beta_1)^{\alpha_1} \dots (v_K - \beta_K)^{\alpha_K} \\ + N \sum_{|\alpha|=N} &\int_0^1 \frac{(1-\sigma)^{N-1}}{\alpha!} (\partial^\alpha F)(\beta_1 + \sigma(v_1 - \beta_1), \dots, \beta_K + \sigma(v_K - \beta_K))\, d\sigma \\ &\qquad \times (v_1 - \beta_1)^{\alpha_1} \dots (v_K - \beta_K)^{\alpha_K}. \end{aligned} \tag{3.12}$$

By Lemma 3.11, $(v_1 - \beta_1)^{\alpha_1} \dots (v_K - \beta_K)^{\alpha_K} \in H^s_{Q,\vartheta}(X) \cap L^\infty(X)$ for any $\alpha \in \mathbb{N}^K$, thus the first summand on the right-hand side of (3.12) belongs to $H^s_{Q,\vartheta}(X) \cap L^\infty(X)$. On the other hand, choosing N sufficiently large, we can arrange that $(v_1 - \beta_1)^{\alpha_1} \dots (v_K - \beta_K)^{\alpha_K} \in \mathcal{H}^{s,(n+1)/2+\vartheta}(X) \cap L^\infty(X)$ for $|\alpha| \geq N$, since $v_J - \beta_J \in \mathcal{H}^{s,(n+1)/2+0}(X) \cap L^\infty(X)$ for $1 \leq J \leq K$. By (3.5), $\big\{(\partial^\alpha F)(\beta_1 + \sigma(v_1 - \beta_1), \dots, \beta_K + \sigma(v_K - \beta_K))\, d\sigma;\ 0 \leq \sigma \leq 1\big\}$ is a bounded set in $\big(1 + \mathcal{H}^{s,(n+1)/2}(X)\big) \cap L^\infty(X)$ for any $\alpha \in \mathbb{N}^K$. This shows that the second summand on the right-hand side of (3.12) belongs to $\mathcal{H}^{s,(n+1)/2+\vartheta}(X) \cap L^\infty(X)$. □

Proposition 3.14. (a) *Let* $s \geq 0$, $\vartheta > 0$. *Further let* $Q \in \underline{\mathrm{As}}^{\sharp}(Y)$ *be multiplicatively closed and* $R \in \underline{\mathrm{As}}(Y)$. *Then* $v_1, \dots, v_K \in H^s_{Q,\vartheta-0}(X) \cap L^\infty(X)$ *and* $F \in C^\infty_R(X \times \mathbb{R}^K)$ *implies that*

$$F(x, v_1, \dots, v_K) \in H^s_{R\circ Q,\vartheta-0}(X). \tag{3.13}$$

(b) *If, in addition,* R *satisfies the assumption of* Lemma 3.12 (b), *then* $v_1, \dots, v_K \in H^s_{Q,\vartheta}(X) \cap L^\infty(X)$ *and* $F \in C^\infty_R(X \times \mathbb{R}^K)$ *implies that*

$$F(x, v_1, \dots, v_K) \in H^s_{R\circ Q,\vartheta}(X).$$

Proof. We prove (a), (b) is analogous. Since $C^\infty_R(X \times \mathbb{R}^K) = C^\infty_R(X) \hat{\otimes}_\pi C^\infty(\mathbb{R}^K)$, we can write

$$F(x, v) = \sum_{j=0}^{\infty} \alpha_j \, \varphi_j(x) \, F_j(v),$$

where $\{\alpha_j\}_{j=0}^\infty \in l^1$ and $\{\varphi_j\}_{j=0}^\infty \subset C^\infty_R(X)$ and $\{F_j\}_{j=0}^\infty \subset C^\infty(\mathbb{R}^K)$, respectively, are null sequences. By the preceeding proposition,

$$F_j(v_1, \dots, v_K) \to 0 \quad \text{as } j \to \infty \text{ in } H^s_{Q,\vartheta-0}(X).$$

By Lemma 3.12,

$$\varphi_j(x) F_j(v_1, \dots, v_K) \to 0 \quad \text{as } j \to \infty \text{ in } H^s_{R\circ Q,\vartheta-0}(X).$$

Thus

$$F(x, v_1, \dots, v_K) = \sum_{j=0}^{\infty} \alpha_j \, \varphi_j(x) F_j(v_1, \dots, v_K) \in H^s_{R\circ Q,\vartheta-0}(X),$$

where the sum on the right-hand side is absolutely convergent. ☐

3.2 The bootstrapping argument

We consider the equation

$$Au = \Pi(u), \tag{3.14}$$

where $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$ is an elliptic Fuchsian differential operator. Properties of the nonlinear operator $u \mapsto \Pi(u)$ are discussed below. The method proposed for deriving elliptic regularity for solutions to (3.14) amounts to *balancing* two asymptotic types — one for the left-hand and the other one for the right-hand side of (3.14).

We assume: There are asymptotic types $\bar{P} \in \underline{\mathrm{As}}^\delta(Y)$, $\bar{Q} \in \underline{\mathrm{As}}^{\delta-\mu}(Y)$, numbers a, b, s_0, $\vartheta_0 \in \mathbb{R}$ with

$$a < \mu, \quad b < \delta_{\bar{Q}} - \delta_{\bar{P}} + \mu, \quad s_0 \geq a^+, \quad \delta_{\bar{P}} + \{\vartheta_0\} \geq \delta,$$

and a subset $\mathcal{U} \subseteq H^{s_0}_{\bar{P},\{\vartheta_0\}}(X)$ such that the following conditions are met:
(A) *A is elliptic with respect to the conormal order δ and $\bar{P} \succcurlyeq \mathcal{P}^\delta(\bar{Q}; A)$, i.e., $u \in \mathcal{H}^{-\infty,\delta}(X)$, $Au \in C^\infty_{\bar{Q}}(X)$ implies $u \in C^\infty_{\bar{P}}(X)$;*
(B) *For $s \geq s_0$, $\vartheta \geq \vartheta_0$, we have*

$$\Pi \colon \mathcal{U} \cap H^s_{\bar{P},\{\vartheta\}}(X) \to H^{s-a}_{\bar{Q},\{\vartheta\}-b}(X).$$

Note that $\{\vartheta_0\} - b + \delta_{\bar{Q}} \geq \delta - \mu$.

Proposition 3.15. *Under the conditions* (A), (B), *each solution $u \in \mathcal{U} \subseteq H^{s_0}_{\bar{P},\{\vartheta_0\}}(X)$ to* (3.14) *belongs to the space $C^\infty_{\bar{P}}(X)$.*

Proof. We prove by induction on j that

$$u \in H^{s_0+j(\mu-a)}_{\bar{P},\{\vartheta_0\}+j(\mu-b+\delta_{\bar{Q}}-\delta_{\bar{P}})}(X) \tag{3.15}$$

for all $j \in \mathbb{N}$. Since $\mu - a > 0$, $\mu - b + \delta_{\bar{Q}} - \delta_{\bar{P}} > 0$, this implies $u \in C^\infty_{\bar{P}}(X)$. By assumption, (3.15) holds for $j = 0$. Now suppose that (3.15) for some j has already been proven. From (B) we conclude that $\Pi(u) \in H^{s_0+j(\mu-a)-a}_{\bar{Q},\{\vartheta_0\}+j(\mu-b+\delta_{\bar{Q}}-\delta_{\bar{P}})-b}(X)$. In view of (A), elliptic regularity gives $u \in H^{s_0+(j+1)(\mu-a)}_{\bar{P},\{\vartheta_0\}+(j+1)(\mu-b+\delta_{\bar{Q}}-\delta_{\bar{P}})}(X)$. □

Example 3.16. Here, we provide an example for a nonlinearity Π satisfying (B). Let $\Pi(u) = K_0(u)/K_1(u)$, where K_0, K_1 are polynomials of degree m_0 and m_1, respectively. Let $u \in H^s_{P,\vartheta-0}(X)$, where $s > (n+1)/2$, $\delta_P + \vartheta > (n+1)/2$, and $\vartheta > 0$. Further, we assume that the multiplicities of the highest singular values for P are simple and the coefficient functions for these singular values nowhere vanish on Y. Then we have $K_0(u) \in H^s_{P_0,\vartheta-0}(X)$, $K_1(u) \in H^s_{P_1,\vartheta-0}(X)$ for resulting asymptotic types P_0, P_1. In particular, P_0 is dominated by $\mathbf{1} \vee \bigvee_{k=1}^{m_0} P^k$ and P_1 is dominated by $\mathbf{1} \vee \bigvee_{k=1}^{m_1} P^k$. Furthermore, it is readily seen that $v \in H^s_{P_1,\vartheta-0}(X)$ and $v \neq 0$ everywhere on X° implies that $1/v \in H^s_{Q_1,\vartheta'-0}(X)$ for some resulting asymptotic type Q_1. Hence, we are allowed to set $\bar{P} = P$, $\bar{Q} = P_0 \circ Q_1$, and

$$\mathcal{U} = \left\{u \in H^s_{P,\vartheta-0}(X);\ K_1(u) \neq 0 \text{ everywhere on } X^\circ\right\}.$$

The condition $s > (n+1)/2$ can be replaced by $s \geq 0$. Then we additionally need $u \in L^\infty_{\mathrm{loc}}(X^\circ)$.

3.3 Proof of the main theorem

The main step consists in constructing asymptotic types $\bar{P}$, $\bar{Q}$ so that Proposition 3.15 applies. Thereby, upon choosing $\delta \in \mathbb{R}$ even smaller if necessary, we can assume that

$$\delta \leq \bar{\mu} + (n+1)/2$$

and that $A \in \mathrm{Diff}^\mu_{\mathrm{Fuchs}}(X)$ is elliptic with respect to the conormal order δ. Set $\Delta = \delta_R + (\mu - \bar{\mu}) - (n+1)/2$. By assumption (1.4), $\Delta > 0$.

3.3.1 Construction of asymptotic types P, Q

We construct by induction on h sequences $\{P_h\}_{h=0}^{\infty} \subset \underline{\mathrm{As}}^{\delta}(Y)$ and $\{Q_h\}_{h=0}^{\infty} \subset \underline{\mathrm{As}}^{\sharp}(Y)$ of asymptotic types as follows: Set $P_0 = \mathcal{P}^{\delta}(\mathcal{O}; A)$. Suppose that $P_0, \dots, P_h$ and $Q_0, \dots, Q_{h-1}$ for some h have already been constructed. Then

$$Q_h = \left(\mathcal{Q}(P_h; B_1, \dots, B_K)^{\sharp}\right)^{\sim}, \tag{3.16}$$

$$P_{h+1} = \mathcal{P}^{\delta}(R \circ Q_h; A). \tag{3.17}$$

Lemma 3.17. *For each $h \geq 0$,*

$$P_h = P_{h+1} \quad \text{up to the conormal order } \delta_R + \mu + h\Delta - 0, \tag{3.18}$$

$$Q_h = Q_{h+1} \quad \text{up to the conormal order } \delta_R + (\mu - \bar{\mu}) + h\Delta - 0. \tag{3.19}$$

In particular, the joins $P = \bigvee_{h=0}^{\infty} P_h$ and $Q = \bigvee_{h=0}^{\infty} Q_h$ exist.

Proof. We set $Q_{-1} = \mathcal{O}$ and proceed by induction on h. (3.19) holds for $h = -1$, since $Q_0 \in \underline{\mathrm{As}}^{\sharp}(Y)$ and, therefore, $Q_0 = \mathcal{O}$ up to the conormal order $(n+1)/2 - 0$.
Suppose that $Q_{h-1} = Q_h$ up to the conormal order $\delta_R + (\mu - \bar{\mu}) + (h-1)\Delta - 0$ for some $h \geq 0$ has already been proved. Then $R \circ Q_{h-1} = R \circ Q_h$ up to the conormal order $\delta_R + h\Delta - 0$ and $P_h = P_{h+1}$ up to the conormal order $\delta_R + \mu + h\Delta - 0$, since $P_h = \mathcal{P}^{\delta}(R \circ Q_h; A)$, $P_{h+1} = \mathcal{P}^{\delta}(R \circ Q_{h+1}; A)$.
Now suppose that $P_h = P_{h+1}$ up to the conormal order $\delta_R + \mu + h\Delta - 0$. We obtain $\mathcal{Q}(P_h; B_1, \dots, B_K) = \mathcal{Q}(P_{h+1}; B_1, \dots, B_K)$ up to the conormal order $\delta_R + (\mu - \bar{\mu}) + h\Delta - 0$ and, therefore, $Q_h = Q_{h+1}$ up to the conormal order $\delta_R + (\mu - \bar{\mu}) + h\Delta - 0$, since $Q_h = \left(\mathcal{Q}(P_h; B_1, \dots, B_K)^{\sharp}\right)^{\sim}$, $Q_{h+1} = \left(\mathcal{Q}(P_{h+1}; B_1, \dots, B_K)^{\sharp}\right)^{\sim}$.
This completes the inductive proof. $\square$

Lemma 3.18. *The asymptotic types $P = \bigvee_{h=0}^{\infty} P_h \in \underline{\mathrm{As}}^{\delta}(Y)$, $Q = \bigvee_{h=0}^{\infty} Q_h \in \underline{\mathrm{As}}^{\sharp}(Y)$ satisfy:*
(a) $\mathcal{Q}(P; B_1, \dots, B_K)^{b} = \mathcal{Q}(P; B_1, \dots, B_K)^{\sharp}$ *and* $Q = \left(\mathcal{Q}(P; B_1, \dots, B_K)^{\sharp}\right)^{\sim}$;
(b) $P = \mathcal{P}^{\delta}(R \circ Q; A)$;
(c) *Q is multiplicatively closed.*
Furthermore, P, Q are minimal among all asymptotic types in $\underline{\mathrm{As}}^{\delta}(Y)$ and $\underline{\mathrm{As}}^{\sharp}(Y)$, respectively, satisfying (a) *to* (c).

Proof. The assertions immediately follow from the description of the asymptotic types P_h, Q_h given in the previous lemma.
Only $\mathcal{Q}(P; B_1, \dots, B_K)^{b} = \mathcal{Q}(P; B_1, \dots, B_K)^{\sharp}$ needs an argument: But $P = P_0$ up to the conormal order $\delta_R + \mu - 0$, so we get $\mathcal{Q}(P; B_1, \dots, B_K) = \mathcal{Q}(P_0; B_1, \dots, B_K)$ up to the conormal order $\delta_R + (\mu - \bar{\mu}) - 0 = (n+1)/2 + \Delta - 0 > (n+1)/2$, and $\mathcal{Q}(P_0; B_1, \dots, B_K)^{b} = \mathcal{Q}(P_0; B_1, \dots, B_K)^{\sharp}$ is exactly the non-resonance condition (1.5). $\square$

Note that, by the non-resonance condition (1.5) and Proposition 3.9,

$$\begin{aligned} B_J u \in \mathcal{H}^{s-\bar{\mu},\delta-\bar{\mu}}_{\mathcal{Q}(P;B_1,\dots,B_K),\vartheta-0}(X) \cap L^\infty(X) \\ \subseteq \mathcal{H}^{s-\bar{\mu},\delta-\bar{\mu}}_{\mathcal{Q}(P;B_1,\dots,B_K)^\sharp,\vartheta-0}(X) \subseteq \mathcal{H}^{s-\bar{\mu},\delta-\bar{\mu}}_{Q,\vartheta-0}(X) \end{aligned} \tag{3.20}$$

if $u \in \mathcal{H}^{s,\delta}_{P,\vartheta-0}(X)$, $\delta - \bar{\mu} + \vartheta > (n+1)/2$, and $B_J u \in L^\infty(X)$.

3.3.2 End of the proof of Theorem 1.1

Since $B_J u \in L^\infty(X) \subset \mathcal{H}^{0,(n+1)/2-0}(X)$ for all $1 \le J \le K$, we have $F(x, B_1 u, \dots, B_K u) \in \mathcal{H}^{0,\delta_R-0}(X)$ and

$$u \in H^\mu_{P_0,\delta_R+\mu-\delta_P-0}(X) = H^\mu_{P,\delta_R+\mu-\delta_P-0}(X)$$

by elliptic regularity.
To conclude the proof of Theorem 1.1, we apply Proposition 3.15 with $\Pi u = F(x, B_1 u, \dots, B_K u)$, $\bar{P} = P$, $\bar{Q} = R \circ Q$, where $P \in \underline{\mathrm{As}}^\delta(Y)$, $Q \in \underline{\mathrm{As}}^\sharp(Y)$ have been constructed in Lemmas 3.17, 3.18, $s_0 = \mu$, $\{\vartheta_0\} = \delta_R + \mu - \delta_P - 0$, $a = \bar{\mu}$, $b = (n+1)/2 - \delta_P + \bar{\mu}$, and

$$\mathcal{U} = \bigl\{u \in H^\mu_{P,\delta_R+\mu-\delta_P-0}(X);\, B_J u \in L^\infty(X),\, 1 \le J \le K\bigr\}. \tag{3.21}$$

Then $a < \mu$, $b < \delta_{R\circ Q} - \delta_P + \mu$ for $\delta_{R\circ Q} = \delta_R$, $\Delta > 0$, and $\delta_P + \vartheta_0 = \delta_R + \mu > \bar{\mu} + (n+1)/2 \ge \delta$, i.e., $\delta_P + \{\vartheta_0\} \ge \delta$. Moreover, condition (A) is fulfilled.
To check condition (B), note that $u \in \mathcal{U} \cap H^s_{P,\vartheta-0}(X)$ for $s \ge \mu$, $\vartheta \ge \delta_R + \mu - \delta_P$ implies

$$F(x, B_1 u, \dots, B_K u) \in H^{s-\bar{\mu}}_{R\circ Q,\delta_P-\bar{\mu}-(n+1)/2+\vartheta-0}(X)$$

by (3.20) and Proposition 3.14.
Thus Proposition 3.15 applies to yield $u \in C^\infty_P(X)$.

Remark 3.19. From (3.21) it is seen that the asymptotic type $P \in \underline{\mathrm{As}}^\delta(Y)$ can be taken smaller, namely instead of $P = \mathcal{P}^\delta(R \circ Q; A)$ we can choose the asymptotic type

$$\bigvee \bigl\{P' \in \underline{\mathrm{As}}^\delta(Y);\, P' \preccurlyeq \mathcal{P}^\delta(R \circ Q; A),\, \mathcal{Q}(P'; B_1, \dots, B_K) \in \underline{\mathrm{As}}^\sharp(Y)\bigr\}.$$

In concrete problems, the resulting asymptotic type for u can be even smaller, e.g., due to nonlinear interaction caused by the *special structure* of the nonlinearity.

3.4 Example: The equation $\Delta u = Au^2 + B(x)u$ in three space dimensions

Let Ω be a bounded, smooth domain in $\mathbb{R}^3$ containing 0. We are going to study singular solutions to the equation

$$\Delta u = Au^2 + B(x)u \ \text{ on } \Omega \setminus \{0\}, \tag{3.22}$$

$$\gamma_0 u = c_0, \ \ u\big|_{\partial\Omega} = \phi, \tag{3.23}$$

where $\gamma_0 u = \lim_{x\to 0}|x|u(x)$, $A \in \mathbb{R}$, and $B \in C^\infty(\overline{\Omega})$ is real-valued. Since the quadratic polynomial $Au^2 + B(x)u$ rather than a general nonlinearity $F(x,u)$ enters, we may admit complex-valued solutions u to (3.22). In particular, $c_0 \in \mathbb{C}$.

Remark 3.20. By results in Véron [17], one expects the limit $\lim_{x\to 0}|x|u(x)$ exist for the solutions $u = u(x)$ to (3.22).

On $\Omega \setminus \{0\}$, we introduce polar coordinates $(t,y) \in \mathbb{R}_+ \times S^2$, $t = |x|$, $y = x/|x|$. We further introduce the function spaces

$$\mathcal{X}^2 = \left\{c_0 t^{-1} + c_{11}\log t + u_0(x);\, c_0, c_{11} \in \mathbb{C},\, u_0 \in H^2(\Omega)\right\},$$
$$\mathcal{Y}^0 = \left\{d_0 t^{-2} + v_0(x);\, d_0 \in \mathbb{C},\, v_0 \in L^2(\Omega)\right\},$$

the definition of which is suggested by formal asymptotic analysis. On the space $\mathcal{X}^2$, we have the trace operators γ_0, γ_1, γ_{11}, where $\gamma_{11}u = \lim_{t\to+0}\big(u(x) - (\gamma_0 u)t^{-1}\big)/\log t$, $\gamma_1 u = \lim_{t\to+0}\big(u(x) - (\gamma_0 u)t^{-1} - (\gamma_{11}u)\log t\big)$.

Proposition 3.21. *Suppose that $B(x) \geq 0$ for all $x \in \overline{\Omega}$. Then, for all $c_0 \in \mathbb{C}$, $\phi \in H^{3/2}(\partial\Omega)$ with $|c_0| + \|\phi\|_{H^{3/2}(\partial\Omega)}$ small enough, the boundary value problem* (3.22), (3.23) *admits a unique small solution $u \in \mathcal{X}^2$. This solution $u = u(x)$ obeys a complete conormal asymptotic expansion as $x \to 0$ that can successively be calculated. Especially,*

$$c_{11} = Ac_0^2, \tag{3.24}$$

where $c_{11} = \gamma_{11}u$.

Proof. Let us consider the nonlinear operator

$$\Psi\colon \mathcal{X}^2 \to \mathcal{Y}^0 \times \mathbb{C} \times H^{3/2}(\partial\Omega),\ \ u \mapsto \left(\Delta u - Au^2 - B(x)u, \gamma_0 u, u\big|_{\partial\Omega}\right).$$

It is readily seen that the linearization of Ψ about $u = 0$ is an isomorphism between the indicated spaces. Thus, the existence of a unique small solution $u \in \mathcal{X}^2$ to (3.22), (3.23) is implied by the inverse function theorem. (3.24) likewise follows.

Furthermore, writing this solution in the form $u(x) = c_0 t^{-1} + c_{11}\log t + u_0(x)$, where $u_0 \in H^2(\Omega)$, we get that u_0 fulfills the equation

$$\begin{aligned} c_{11}t^{-2} + \Delta u_0 &= A\left(c_0^2 t^{-2} + 2c_0 c_{11} t^{-1}\log t + c_{11}^2 \log^2 t\right) \\ &+ 2A\left(c_0 t^{-1} + c_{11}\log t\right)u_0 + Au_0^2 + B(x)\left(c_0 t^{-1} + c_{11}\log t\right) + B(x)u_0. \end{aligned} \tag{3.25}$$

This can be brought into the form (1.1) with $A = \Delta$,

$$\begin{aligned} F(x,\nu) = &\left(2Ac_0c_{11}t^{-1}\log t + B(x)c_0 t^{-1} + Ac_{11}^2\log^2 t + B(x)c_{11}\log t\right) \\ &+ \left(2Ac_0 t^{-1} + 2Ac_{11}\log t + B(x)\right)\nu + A\nu^2, \end{aligned}$$

since $\Delta = t^{-2}\left((-t\partial_t)^2 - (-t\partial_t) + \Delta_{S^2}\right) \in \mathrm{Diff}^2_{\mathrm{Fuchs}}(\Omega \setminus \{0\})$, where $0 \in \Omega$ is considered as conical point with cone base $S^2 = \{x \in \mathbb{R}^3;\, |x| = 1\}$, cf. Remark 1.3, and Δ_{S^2} being the Laplace-Beltrami operator on S^2. The conditions (1.4), (1.5) are obviously satisfied.

Thus, Theorem 1.1 applies to $u_0 \in H^2(\Omega) \subset L^\infty(\Omega)$ to yield that u_0 and, therefore, u obey a complete conormal asymptotic expansion. □

Remark 3.22. (a) Taking for P the asymptotic type in $\underline{\mathrm{As}}^0(S^2)$ that comes out of the calculation of the conormal asymptotic expansion for u, i.e., we have $u \in C_P^\infty(\Omega \setminus \{0\})$, and for Q the resulting asymptotic type in $\underline{\mathrm{As}}^{-2}(S^2)$ for the right-hand side of (3.22), we are in a situation in which Proposition 3.15 directly applies without having boundedness assumptions for u.
(b) Allowing more general functions $B \in C_R^\infty(\Omega \setminus \{0\})$ for some $R \in \underline{\mathrm{As}}^{-1/2}(S^2)$ (the conormal order $-1/2$ ensures that the term $Ac_0^2 t^{-2}$ dominates on the right-hand side of (3.25)) rather than $B \in C_{P_0}^\infty(\Omega \setminus \{0\})$, where P_0 is the asymptotic type for Taylor asymptotics, one can perform the same analysis as before upon replacing the space $H^2(\Omega)$ in the definition of $\mathcal{X}^2$ accordingly.

References

[1] P. M. Bleher, *Operators that depend meromorphically on a parameter*, Vestnik Moskov. Univ. Ser. I Mat. Meh. 24 (1969), 30–36, (In Russian).

[2] M. Dauge, *Elliptic boundary value problems on corner domains*, Lecture Notes in Math., vol. 1341, Springer, Berlin, 1988.

[3] B. A. Davey and H. A. Priestley, *Introduction to lattices and order*, Cambridge Math. Textbooks, Cambridge Univ. Press, Cambridge, 1990.

[4] I. C. Gohberg and E. I. Sigal, *An operator generalization of the logarithmic residue theorem and the theorem of Rouché*, Math. USSR Sbornik 13 (1971), no. 4, 603–625.

[5] P. Jeanquartier, *Transformation de Mellin et développements asymptotiques*, Enseign. Math. 25 (1979), no. 3-4, 285–308.

[6] D. Kapanadze, B.-W. Schulze, and I. Witt, *Coordinate invariance of the cone algebra with asymptotics*, Advances in Partial Differential Equations (Parabolicity, Volterra Calculus, and Conical Singularities) (S. Albeverio, M. Demuth, E. Schrohe, and B.-W. Schulze, eds.), Oper. Theory Adv. Appl., vol. 138, Birkhäuser, Basel, 2002, pp. 307–358.

[7] S. Kichenassamy, *The blow-up problem for exponential nonlinearities*, Comm. Partial Differential Equations 21 (1996), 125–162.

[8] ———, *Fuchsian equations in Sobolev spaces and blow-up*, J. Differential Equations 125 (1996), 299–327.

[9] V. A. Kondrat'ev, *Boundary problems for elliptic equations in domains with conical or angular points*, Trans. Mosc. Math. Soc. 16 (1967), 227–313.

[10] X. Liu, *A cone pseudodifferential calculus on the half-line with respect to conormal asymptotics of a given type*, Ph.D. thesis, Department of Mathematics, Wuhan University, Wuhan, 2000.

[11] X. Liu and I. Witt, *Nonlinear superposition in weighted Hölder-Zygmund spaces with asymptotics*, (in preparation).

[12] ———, *Pseudodifferential calculi on the half-line respecting prescribed asymptotic types*, (to appear in Integral Equations Operator Theory).

[13] R. B. Melrose, *The Atiyah-Patodi-Singer index theorem*, Res. Notes Math., A K Peters, Wellesley MA, Boston, 1993.

[14] S. Rempel and B.-W. Schulze, *Asymptotics for elliptic mixed boundary problems*, Math. Research, vol. 50, Akademie-Verlag, Berlin, 1989.

[15] B.-W. Schulze, *Pseudo-differential operators on manifolds with singularities*, Stud. Math. Appl., vol. 24, North-Holland, 1991.

[16] ———, *Boundary value problems and singular pseudo-differential operators*, Wiley Ser. Pure Appl. Math., J. Wiley, Chichester, 1998.

[17] L. Véron, *Singularities of solutions of second order quasilinear equations*, Pitman Research Notes Math. Series, vol. 353, Longman, Harlow, Essex, 1996.

[18] I. Witt, *Local asymptotic types*, (to appear in Manuscripta Math.).

[19] ———, *Non-linear hyperbolic equations in domains with conical points*, Math. Research, vol. 84, Akademie Verlag, Berlin, 1995.

[20] ———, *Asymptotic algebras*, Microlocal Analysis and Asymptotic Analysis of PDE, RIMS Kokyuroku, vol. 1211, Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 2001, pp. 21–33.

[21] ———, *On the factorization of meromorphic Mellin symbols*, Advances in Partial Differential Equations (Parabolicity, Volterra Calculus, and Conical Singularities) (S. Albeverio, M. Demuth, E. Schrohe, and B.-W. Schulze, eds.), Oper. Theory Adv. Appl., vol. 138, Birkhäuser, Basel, 2002, pp. 279–306.

Xiaochun Liu
School of Mathematics and Statistics
Wuhan University
Wuhan, 430072
P.R. China
liuxc72@public.wh.hb.cn

Ingo Witt
Department of Mathematics
Imperial College
180 Queen's Gate
London SW7 2AZ
United Kingdom
i.witt@imperial.ac.uk

Bounds for the Anticanonical Bundle of a Homogeneous Projective Rational Manifold

Dennis Snow

Received: February 26, 2004

Communicated by Thomas Peternell

Abstract. The following bounds for the anticanonical bundle $K_X^* = \det T_X$ of a complex homogeneous projective rational manifold X of dimension n are established:

$$3^n \leq \dim H^0(X, K_X^*) \leq \binom{2n+1}{n} \quad \text{and} \quad 2^n n! \leq \deg K_X^* \leq (n+1)^n$$

with equality in the lower bounds if and only if X is a flag manifold and equality in the upper bounds if and only if X is complex projective space. None of these bounds holds for general Fano manifolds.

2000 Mathematics Subject Classification: Primary 14M17; Secondary 14M15, 32M10

The homogeneous compact complex manifolds X that admit an equivariant embedding in projective space are precisely the quotients $X = G/P$ where G is a semisimple complex Lie group and P is a parabolic subgroup. Moreover, any such quotient is rational and has a very ample anticanonical bundle, $K_X^* = \det T_X$. In particular, X is a Fano manifold.
Various bounds have been established for the numerical invariants of K_X^* when X is a general Fano manifold, see [6, 8, 9, 10]. For example, there exists a constant $c(n)$ that depends only on $n = \dim X$ such that $\deg K_X^* \leq c(n)^n$. In this article we establish the following bounds when $X = G/P$:

$$3^n \leq \dim H^0(X, K_X^*) \leq \binom{2n+1}{n} \quad \text{and} \quad 2^n n! \leq \deg K_X^* \leq (n+1)^n$$

with equality in the lower bounds if and only if X is a flag manifold (i.e., P is a Borel subgroup of G), and equality in the upper bounds if and only if X is complex projective space, $\mathbf{P}^n$.

These bounds do not hold for general Fano manifolds. For example, let $X = \mathbf{P}(\mathcal{O}_{\mathbf{P}^{n-1}} \oplus \mathcal{O}_{\mathbf{P}^{n-1}}(n-1))$. Then X is a $\mathbf{P}^1$-bundle over $\mathbf{P}^{n-1}$, $\pi : X \to \mathbf{P}^{n-1}$, and $K_X^* = \pi^*\mathcal{O}_{\mathbf{P}^{n-1}}(1) \otimes \xi^2$ where ξ is the tautological line bundle on X whose restriction to any fiber $\mathbf{P}^1$ of π gives $\xi|_{\mathbf{P}^1} \cong \mathcal{O}_{\mathbf{P}^1}(1)$. It follows that X is a Fano manifold with $\dim H^0(X, K_X^*) = n + \binom{2n-1}{n-1} + \binom{3n-2}{n-1}$ and $\deg K_X^* = ((2n-1)^n - 1)/(n-1)$. An example where the lower bounds do not hold is given by $X = S \times (\mathbf{P}^1)^{n-2}$ where S is a del Pezzo surface.

In the homogeneous case there are well-known formulas from representation theory that can be used to calculate $\dim H^0(X, K_X^*)$ and $\deg K_X^*$ exactly. However, these formulas, which are products of rational numbers indexed by the roots of the group, do not easily lend themselves to comparison with expressions in $n = \dim X$. The point of this paper is to overcome this difficulty. The bounds are proved by first reducing to the case of simple Lie groups and then showing for each classical type that the known formulas can be broken up into subproducts of certain simple sequences. These subproducts are shown to satisfy inequalities that can be combined to yield the desired inequalities for the full product. The bounds for the exceptional types are verified through exhaustive calculations.

The above upper bounds can be trivially extended to any homogeneous compact complex manifold $X = G/H$. The sections of K_X^* define an equivariant map of X to projective space that coincides with the normalizer fibration $G/H \to G/N$, $N = N_G(H^0)$, [1, p.79]. Since the base $Y = G/N$ is a homogeneous projective rational manifold, the upper bounds hold for Y and hence for X.

For a homogeneous projective rational manifold X, the dimension of the holomorphic automorphism group, $\dim \mathrm{Aut}(X) = \dim H^0(X, T_X)$, never exceeds $n(n+2)$. In fact, this bound holds when X is any homogeneous compact Kähler manifold [5]. However, there are homogeneous compact complex manifolds for which $\dim \mathrm{Aut}(X)$ grows exponentially in n, see [12]. In [13], the above estimate for $\dim H^0(X, K_X^*)$ plays an important role in establishing the following bound for the non-Kähler case: $\dim \mathrm{Aut}(X) \le n^2 - 1 + \binom{2n-1}{n-1} \sim O(2^{2n-1}/\sqrt{(n-1)\pi})$.

1 Roots and Weights

In this section we introduce some notation and well-known facts about semisimple Lie groups [2, 7], and recall a formula for finding the weight μ_X associated to the line bundle K_X^* when $X = G/P$ [4, 11].

Let G be a semisimple complex Lie group and let T be a maximal torus of G. Let $\mathrm{Lie}(G)$ and $\mathrm{Lie}(T)$ be the corresponding Lie algebras. Let $\Phi \subset \mathrm{Lie}(T)^*$ denote the roots of G with respect to T and let $\{\alpha_1, \ldots, \alpha_\ell\}$ be a system of simple roots. Let Φ^+ denote the subset of positive roots—those that are positive integral combinations of the simple roots. For any root $\alpha \in \Phi$, let $e_\alpha \in \mathrm{Lie}(G)$ be the corresponding root vector: $[x, e_\alpha] = \alpha(x)e_\alpha$ for all $x \in \mathrm{Lie}(T)$.

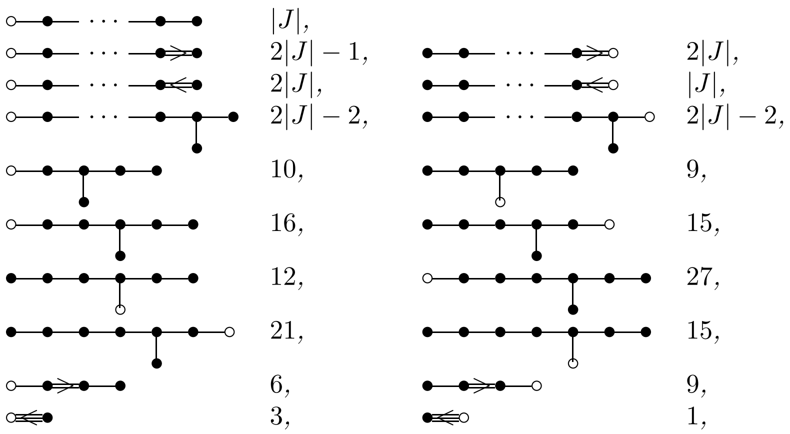
Let $\lambda_1, \ldots, \lambda_\ell$ be the fundamental dominant weights of G—those weights defined by $\langle \lambda_i, \alpha_j \rangle = 2(\lambda_i, \alpha_j)/(\alpha_j, \alpha_j) = \delta_{ij}$ where $(\ ,\ )$ denotes the Killing form.

Any weight $\mu \in \mathrm{Lie}(T)^*$ can be written $\mu = \sum_{i=1}^{\ell} \langle \mu, \alpha_i \rangle \lambda_i$.
A Borel subgroup is a maximal solvable subgroup of G, and any such subgroup is conjugate to the subgroup B generated by T and the root groups $\exp \mathbf{C} e_\alpha$, for all $\alpha \in -\Phi^+$. Let P be a parabolic subgroup of G, that is, a subgroup containing a Borel subgroup. We may assume that P contains B. Let $P = R \cdot S$ be a Levi decomposition of P where R is a maximal solvable normal subgroup of P and S is semisimple. We let Φ_P denote the subsystem of roots of S and let $\Phi_P^+ = \Phi_P \cap \Phi^+$. Let I denote the subset of indexes, $I \subset \{1, \ldots, \ell\}$, such that $\Phi_P^+ \cap \{\alpha_1, \ldots, \alpha_\ell\} = \{\alpha_i\}_{i \in I}$. The conjugacy class of P is uniquely determined by I and any such choice of indexes is associated to a parabolic subgroup of G. Let $X = G/P$, and define $\Phi_X^+ = \Phi^+ \setminus \Phi_P^+$. Since T_X is generated at the identity coset by the root vectors e_α for $\alpha \in \Phi_X^+$, the anticanonical bundle $K_X^* = \det T_X$, $n = \dim X$, is the homogeneous line bundle associated to the weight

$$\mu_X = \sum_{\alpha \in \Phi_X^+} \alpha$$

The weight μ_X is dominant: $\langle \mu_X, \alpha_i \rangle > 0$ for $i \notin I$, and $\langle \mu_X, \alpha_i \rangle = 0$ for $i \in I$. In particular, K_X^* is a very ample line bundle and μ_X is orthogonal to the roots Φ_P^+. If $P = B$, X is called a flag manifold.
We now recall a simple formula for calculating the coefficients $\langle \mu_X, \alpha_i \rangle$ of μ_X, see [11]: A set of indexes J is called connected if the subdiagram of the Dynkin diagram of G corresponding to the simple roots α_j, $j \in J$, is connected. An index i is said to be adjacent to J if $i \notin J$ and $J_0 \cup \{i\}$ is connected for some connected component J_0 of J. The set of indexes adjacent to J is denoted by ∂J. The number of elements in J is denoted $|J|$.

Definition 1 *Let J be a connected set of indexes. For $i \notin \partial J$ define $\nu_i(J) = 0$. For $i \in \partial J$ define $\nu_i(J)$ to be the number next to the appropriate diagram below. The black nodes correspond to J and the white node corresponds to i. Symmetry of Dynkin diagrams is tacitly assumed.*



For an arbitrary set of indexes I define $\nu_i(I) = \nu_i(I_1) + \cdots + \nu_i(I_p)$ where $I_1, \ldots, I_p$ are the connected components of I.

Proposition 1 ([11]) *Let $X = G/P$ where G is a semisimple complex Lie group and P is a parabolic subgroup defined by a set of indexes I. Let μ_X be the weight of the anticanonical bundle K_X^* of X. Then*

$$\mu_X = \sum_{i \notin I} (2 + \nu_i(I))\lambda_i$$

2 Estimating Products

We now prove some estimates for various products that appear in the proof of the main theorem.

Lemma 1 *For any non-negative integers s and t,*

$$\binom{2t+1}{t}\binom{2s+1}{s} \leq \binom{2(t+s)+1}{t+s} \tag{1}$$

$$\frac{(t+1)^t}{t!} \cdot \frac{(s+1)^s}{s!} \leq \frac{(t+s+1)^{t+s}}{(t+s)!} \tag{2}$$

with equality if and only if t or s is 0.

Proof. The inequalities are obviously equalities when s or t is 0. So we assume $t, s > 0$ and show strict inequalities hold for (1) and (2) by fixing s and applying induction on t. They are easily seen to hold for $t = 1$. Let $g(t) = \binom{2t+1}{t}$ (resp., $(t+1)^t/t!$), and let $f(t) = g(t+1)/g(t) = 4 - 2/(t+2)$ (resp., $[1 + 1/(t+1)]^{t+1}$), an increasing function of $t > 0$. By the induction hypothesis,

$$g(t+1)g(s) = f(t)g(t)g(s) < f(t)g(t+s) < f(t+s)g(t+s) = g(t+s+1)$$

□

Definition 2 *Let t and s be positive integers. A* simple sequence *(of length s) is a set S of rational numbers of the form*

$$S = S(t,s) = \left\{ \frac{3t+s-1+i}{t+i} \,\middle|\, 0 \leq i \leq s-1 \right\}$$

The shifted sequence *of $S(t,s)$ is*

$$S' = S'(t,s) = \left\{ \frac{3t+s-1+i}{t+i} - 1 \,\middle|\, 0 \leq i \leq s-1 \right\}$$

The products of the numbers in S and S' are denoted by

$$\Pi S = \prod_{i=0}^{s-1} \frac{3t+s-1+i}{t+i} = \binom{3t+2s-2}{s} \Big/ \binom{t+s-1}{s}$$

$$\Pi S' = \prod_{i=0}^{s-1} \frac{2t+s-1}{t+i} = \frac{(2t+s-1)^s (t-1)!}{(t+s-1)!}$$

Lemma 2 *Let $S(t,s)$ be a simple sequence and let $S'(t,s)$ be the shifted sequence of $S(t,s)$.*

1. *If the first and last elements of $S(t,s)$ are removed, the remaining set is the simple sequence $S(t+1, s-2)$.*

2. $\Pi S(t,s)$ *and* $\Pi S'(t,s)$ *are decreasing in t. In particular,*

$$3^s = \lim_{t\to\infty} \Pi S(t,s) \le \Pi S(t,s) \;\le\; \Pi S(1,s) = \binom{2s+1}{s}$$
$$2^s = \lim_{t\to\infty} \Pi S'(t,s) \le \Pi S'(t,s) \;\le\; \Pi S'(1,s) = \frac{(s+1)^s}{s!}$$

Proof. The first assertion is immediate from the definition. To prove the second assertion, let $f(t) = \Pi S(t,s)$ and let $m = [(s-1)/2]$ be the least integer $\le (s-1)/2$. Then, for $t > 0$,

$$\begin{aligned}\frac{d}{dt}\log f(t) &= \sum_{i=0}^{s-1} \frac{2i-(s-1)}{(3t+s-1+i)(t+i)} \\ &= \sum_{i=0}^{m} -\frac{s-1-2i}{(3t+s-1+i)(t+i)} + \frac{s-1-2i}{(3t+2s-2-i)(t+s-1-i)} \le 0\end{aligned}$$

and hence f is decreasing.
Now let $g(t) = \Pi S'(t,s)$ and define $h(t) = g(t+1)/g(t) = [1 + 2/(2t+s-1)]^s t/(t+s)$. Then

$$\frac{d}{dt}\log h(t) = \frac{s(s^2-1)}{t(t+s)((2t+s)^2-1)} \ge 0$$

so h is increasing and approaches 1 as $t \to \infty$. Therefore, g is decreasing. □

3 Bounds for K_X^*

Theorem 1 *Let X be a homogeneous projective rational manifold of dimension n. Then*

$$3^n \le \dim H^0(X, K_X^*) \le \binom{2n+1}{n} \quad \text{and} \quad 2^n n! \le \deg K_X^* \le (n+1)^n$$

with equality in the lower bounds if and only if X is a flag manifold and equality in the upper bounds if and only if $X = \mathbf{P}^n$.

Proof. Write $X = G/P$ where G is a semisimple Lie group and P is a parabolic subgroup. Let I be the subset of indexes that defines P, and let $I_1, \ldots, I_m$ be the connected components of I. Let μ_X be the weight of the anticanonical bundle as given in Proposition 1 so that $H^0(X, K_X^*)$ is the irreducible representation of G with highest weight $\mu_X = \sum_{i \notin I}(2 + \nu_i(I))\lambda_i$. Set $\delta = (1/2)\sum_{\alpha>0} \alpha = \lambda_1 + \cdots + \lambda_\ell$. By the Weyl dimension formula [7],

$$h = \dim H^0(X, K_X^*) = \prod_{\alpha \in \Phi_X^+} \frac{(\mu_X + \delta, \alpha)}{(\delta, \alpha)} \tag{3}$$

and the degree of K_X^* is given by [4]

$$d = \deg K_X^* = n! \prod_{\alpha \in \Phi_X^+} \frac{(\mu_X, \alpha)}{(\delta, \alpha)} \tag{4}$$

Let $G_1, \ldots, G_r$ be the simple factors of G. Then $X = X_1 \times \cdots \times X_r$ where $X_i = G_i/G_i \cap P$. Let $n = \dim X$, $n_i = \dim X_i$, $h_i = \dim H^0(X_i, K_{X_i}^*)$ and $d_i = \deg K_{X_i}^*$, $1 \le i \le r$. If $3^{n_i} \le h_i \le \binom{2n_i+1}{n_i}$ and $2^{n_i} n_i! \le d_i \le (n_i+1)^{n_i}$, $1 \le i \le r$, then the above formulas along with Lemma 1 imply $3^n \le h = h_1 \cdots h_r \le \prod_{i=1}^r \binom{2n_i+1}{n_i} \le \binom{2n+1}{n}$ and $2^n n! \le d = n!\,(d_1/n_1!) \cdots (d_r/n_r!) \le n! \prod_{i=1}^r (n_i+1)^{n_i}/n_i! \le (n+1)^n$, since $n = n_1 + \cdots + n_r$. We may therefore assume that G is simple.

The theorem can be verified by direct calculation for each of the exceptional simple Lie groups and their finite number of conjugacy classes of parabolic subgroups. While the details are too lengthy to include in this article, the results can be summarized as follows. The minimum of h is 3^n and is achieved only for Borel subgroups. The maximum of h is always strictly less than $\binom{2n+1}{n}$. In fact, the minimum of $\binom{2n+1}{n}/h$ over all parabolic subgroups for each type is greater than 3.11 for E_6, 9.96 for E_7, 758.2 for E_8, 3.24 for F_4, and 1.22 for G_2, and this minimum is achieved for the maximal parabolic subgroups defined by $I = \{2, \ldots, \ell\}$, or $\{1, 2, 3\}$ for F_4 (the simple roots are indexed from left to right in the diagrams shown in Definition 1).

The proof for the classical types A_ℓ, B_ℓ, C_ℓ and D_ℓ is accomplished by showing that the product (3) can be written as a product of simple sequences $S_1, \ldots, S_\sigma$. For, if we know that $h = \Pi S_1 \cdots \Pi S_\sigma$, it follows from (4) that $d = n!\,\Pi S_1' \cdots \Pi S_\sigma'$, and by Lemmas 1 and 2, we obtain $3^n \le h \le \prod_{i=1}^\sigma \binom{2|S_i|+1}{|S_i|} \le \binom{2n+1}{n}$ and $2^n n! \le d \le n! \prod_{i=1}^\sigma (|S_i|+1)^{|S_i|}/|S_i|! \le (n+1)^n$, since $n = |S_1| + \cdots + |S_\sigma|$. We now prove that such a decomposition of (3) is possible for each simple classical type.

Type A_ℓ: Let $\mu_X + \delta = m_1\lambda_1 + \cdots + m_\ell\lambda_\ell$, where $m_i = \langle \mu_X, \alpha_i \rangle + 1$, $1 \le i \le \ell$, and let $\epsilon_1, \ldots, \epsilon_{\ell+1}$ denote the standard orthonormal basis of $\mathbf{R}^{\ell+1}$. The simple roots for type A_ℓ are $\alpha_i = \epsilon_i - \epsilon_{i+1}$, $1 \le i \le \ell$ and the positive roots are $\alpha_i + \cdots + \alpha_{j-1} = \epsilon_i - \epsilon_j$, $1 \le i < j \le \ell+1$. The dimension formula (3) becomes $h = \prod a_{ij}$ where $a_{ij} = (m_i + \cdots + m_{j-1})/(j-i)$ and the product is taken over all indexes $i < j$ that are not both in same connected component of I.

According to Proposition 1, the coefficients $m_1, \dots, m_\ell$ are given by

$$m_i = \begin{cases} 1 & \text{if } i \in I \\ 3 + |I_\nu| & \text{if } i \in \partial I_\nu \text{ for some } \nu \\ 3 + |I_\nu| + |I_{\nu+1}| & \text{if } i \in \partial I_\nu \cap \partial I_{\nu+1} \text{ for some } \nu \\ 3 & \text{otherwise} \end{cases} \tag{5}$$

An example is given by the list of numbers at the top of Figure 1 (the indexes in I correspond to black nodes).
The numbers in the product $h = \prod a_{ij}$ can be arranged into rectangular arrays as follows. Let $i_1 < \cdots < i_k$ be an ordered list of those indexes i not in I and set $i_0 = 0$, $i_{k+1} = \ell + 1$. For $1 \le p \le q \le k$, define $R_{pq} = \{a_{ij} \,|\, i_{p-1} < i \le i_p,\ i_q < j \le i_{q+1}\}$, as illustrated in Figure 1.

Figure 1: Type A_ℓ decomposition

3 5 1 1 8 1 1 1 6

$\frac{3}{1}$	$\frac{8}{2}$	$\frac{9}{3}$	$\frac{10}{4}$	$\frac{18}{5}$	$\frac{19}{6}$	$\frac{20}{7}$	$\frac{21}{8}$	$\frac{27}{9}$
	$\frac{5}{1}$	$\frac{6}{2}$	$\frac{7}{3}$	$\frac{15}{4}$	$\frac{16}{5}$	$\frac{17}{6}$	$\frac{18}{7}$	$\frac{24}{8}$
				$\frac{10}{3}$	$\frac{11}{4}$	$\frac{12}{5}$	$\frac{13}{6}$	$\frac{19}{7}$
				$\frac{9}{2}$	$\frac{10}{3}$	$\frac{11}{4}$	$\frac{12}{5}$	$\frac{18}{6}$
				$\frac{8}{1}$	$\frac{9}{2}$	$\frac{10}{3}$	$\frac{11}{4}$	$\frac{17}{5}$
								$\frac{9}{4}$
								$\frac{8}{3}$
								$\frac{7}{2}$
								$\frac{6}{1}$

Then h is the product of the numbers in all the rectangular arrays R_{pq}, $1 \le p \le q \le k$. Each R_{pq} consists of rational numbers whose numerators and denominators both increase by 1 in each row and column, starting in the lower left corner, $a_{i_p(i_q+1)}$. From (5) it follows that $a_{i_p(i_q+1)}$ has the form $(3t+s-1)/t$ where $t = i_q - i_p + 1$ and s is the number of rows + columns $-1 = (i_p - i_{p-1}) + (i_{q+1} - i_q) - 1$. Therefore, R_{pq} may be decomposed into simple sequences, $R_{pq} = S_0 \cup \ldots \cup S_\sigma$ where $S_0 = \{a_{ij} \,|\, i = i_{p-1}+1 \text{ or } j = i_q+1\} = S(t,s)$ is the set of numbers in the left column and the top row of R_{pq}, and $S_i = S(t+i, s-2i)$ is obtained by removing the lower left and top right number from S_{i-1}, $1 \le i \le \sigma$, as illustrated in Figure 2.

Figure 2: Type A_ℓ rectangular array

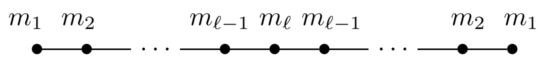
$\frac{10}{3}$	$\frac{11}{4}$	$\frac{12}{5}$	$\frac{13}{6}$
$\frac{9}{2}$	$\frac{10}{3}$	$\frac{11}{4}$	$\frac{12}{5}$
$\frac{8}{1}$	$\frac{9}{2}$	$\frac{10}{3}$	$\frac{11}{4}$

Repeating this procedure for each rectangular array R_{pq}, $1 \le p \le q \le k$, shows that for type A_ℓ the product (3) can be decomposed into a product of simple sequences.

Type B_ℓ: We now show the same type of decomposition is possible for type B_ℓ by embedding the appropriate numbers into a diagram for type $A_{2\ell-1}$. We again write $\mu_X + \delta = m_1\lambda_1 + \cdots + m_\ell\lambda_\ell$, $m_i = \langle \mu_X, \alpha_i \rangle + 1$, $1 \le i \le \ell$, and let $\epsilon_1, \ldots, \epsilon_\ell$ denote the standard orthonormal basis of $\mathbf{R}^\ell$. The simple roots are $\alpha_i = \epsilon_i - \epsilon_{i+1}$, $1 \le i \le \ell - 1$, and $\alpha_\ell = \epsilon_\ell$. The positive roots are $\alpha_i + \cdots + \alpha_{j-1} = \epsilon_i - \epsilon_j$, $\alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + \cdots + 2\alpha_\ell = \epsilon_i + \epsilon_j$, $1 \le i < j \le \ell$, and $\alpha_i + \cdots + \alpha_\ell = \epsilon_i$, $1 \le i \le \ell$. The dimension formula (3) becomes $h = \prod a_{ij} \times \prod b_{ij}$ where $a_{ij} = (m_i + \cdots + m_{j-1})/(j-i)$, $1 \le i < j \le \ell$, $b_{ij} = (m_i + \cdots + m_{j-1} + 2m_j + \cdots + 2m_{\ell-1} + m_\ell)/(2\ell - i - j + 1)$, $1 \le i \le j \le \ell$. To avoid trivial factors, these products should be taken over i, j not in the same connected component of I, although in the following arguments it is convenient to include all terms.

Define $\hat{I} = \{i \,|\, i \in I \text{ or } 2\ell - i \in I\}$. Then $\hat{I}$ defines a parabolic subgroup $\hat{P}$ of a simple group $\hat{G}$ of type $A_{2\ell-1}$. Let $\hat{X} = \hat{G}/\hat{P}$. By Proposition 1, the coefficients of $\mu_X + \delta$ appear as the first half of the coefficients of $\mu_{\hat{X}} + \delta$, see Figure 3.

Figure 3: Conversion of type B_ℓ to type $A_{2\ell-1}$

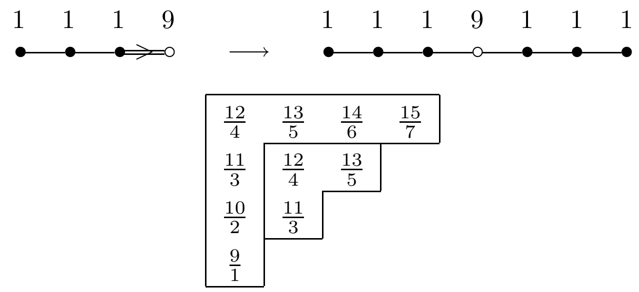


For a fixed i the product $h_i = \prod a_{ij} \times \prod b_{ij}$ can be arranged as

$$\frac{m_i}{1} \cdot \frac{m_i + m_{i+1}}{2} \cdots \frac{m_i + \cdots + m_\ell}{\ell - i + 1} \cdot \frac{s_i + m_{\ell-1}}{\ell - i + 2} \cdots \frac{s_i + m_{\ell-1} + \cdots + m_i}{2(\ell - i) + 1}$$

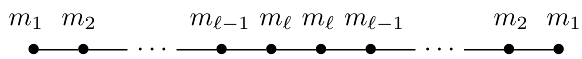
where $s_i = m_i + \cdots + m_\ell$. Therefore, the non-trivial terms in h correspond to the numbers in the upper left half of the rectangular arrays R_{pq} defined for type $A_{2\ell-1}$. These triangular arrays can clearly be broken up into simple sequences, see Figure 4, showing that h is a product of simple sequences.

Figure 4: Type B_ℓ decomposition



Type C_ℓ: The proof for this case is almost identical to type B_ℓ. The simple roots are $\alpha_i = \epsilon_i - \epsilon_{i+1}$, $1 \le i \le \ell-1$, and $\alpha_\ell = 2\epsilon_\ell$. The positive roots are $\alpha_i + \cdots + \alpha_{j-1} = \epsilon_i - \epsilon_j$, $\alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + \cdots + 2\alpha_{\ell-1} + \alpha_\ell = \epsilon_i + \epsilon_j$, $1 \le i < j \le \ell$, and $2\alpha_i + \cdots + 2\alpha_\ell + \alpha_\ell = 2\epsilon_i$, $1 \le i \le \ell$. The dimension formula (3) becomes $h = \prod a_{ij} \times \prod b_{ij}$ where $a_{ij} = (m_i + \cdots + m_{j-1})/(j-i)$, $1 \le i < j \le \ell$, $b_{ij} = (m_i + \cdots + m_{j-1} + 2m_j + \cdots + 2m_\ell)/(2\ell - i - j + 2)$, $1 \le i \le j \le \ell$.
Define $\hat{I} = \{i \,|\, i \in I \text{ or } 2\ell - i + 1 \in I\}$. Then $\hat{I}$ defines a parabolic subgroup $\hat{P}$ of a simple group $\hat{G}$ of type $A_{2\ell}$. Let $\hat{X} = \hat{G}/\hat{P}$. By Proposition 1, the coefficients of $\mu_X + \delta$ appear as the first half of the coefficients of $\mu_{\hat{X}} + \delta$, see Figure 5.

Figure 5: Conversion of type C_ℓ to type $A_{2\ell}$

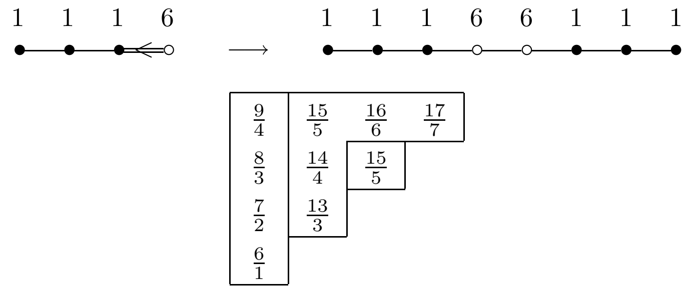


For a fixed i the product $h_i = \prod a_{ij} \times \prod b_{ij}$ can be arranged as

$$\frac{m_i}{1} \cdot \frac{m_i + m_{i+1}}{2} \cdots \frac{m_i + \cdots + m_\ell}{\ell - i + 1} \cdot \frac{s_i + m_\ell}{\ell - i + 2} \cdots \frac{s_i + m_\ell + \cdots + m_{i+1}}{2(\ell - i) + 1}$$

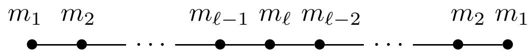
where $s_i = m_i + \cdots + m_\ell$. Therefore, the non-trivial terms in h correspond to the numbers in the upper left half (above the diagonal) of the rectangular arrays R_{pq} defined for type $A_{2\ell}$. These triangular arrays can be broken up into simple sequences as before, see Figure 6, showing that h is a product of simple sequences.
Type D_ℓ: The proof for this case must be handled somewhat differently than the previous two cases. The simple roots are $\alpha_i = \epsilon_i - \epsilon_{i+1}$, $1 \le i \le \ell - 1$, and $\alpha_\ell = \epsilon_{\ell-1} + \epsilon_\ell$. The positive roots are $\alpha_i + \cdots + \alpha_{j-1} = \epsilon_i - \epsilon_j$, $1 \le i < j \le \ell$,

Figure 6: Type C_ℓ decomposition



$\alpha_i + \cdots + \alpha_{j-1} + 2\alpha_j + \cdots + 2\alpha_{\ell-2} + \alpha_{\ell-1} + \alpha_\ell = \epsilon_i + \epsilon_j$, $1 \le i < j \le \ell - 1$, and $\alpha_i + \cdots + \alpha_{\ell-2} + \alpha_\ell = \epsilon_i + \epsilon_\ell$, $1 \le i \le \ell - 2$. The dimension formula (3) becomes $h = \prod a_{ij} \times \prod b_{ij} \times \prod c_i$ where $a_{ij} = (m_i + \cdots + m_{j-1})/(j-i)$, $1 \le i < j \le \ell$, $b_{ij} = (m_i + \cdots + m_{j-1} + 2m_j + \cdots + 2m_{\ell-2} + m_{\ell-1} + m_\ell)/(2\ell - i - j)$, $1 \le i < j \le \ell - 1$, $c_i = (m_i + \cdots + m_{\ell-2} + m_\ell)/(\ell - i)$, $1 \le i \le \ell - 2$, and $c_{\ell-1} = m_\ell$.
By symmetry of the Dynkin diagram, we may assume $m_{\ell-1} \le m_\ell$. We first assume $m_{\ell-1} = m_\ell$. Define $\hat{I} = \{i \,|\, i \in I \text{ or } 2\ell - i - 1 \in I \text{ (and } i > \ell)\}$. Then $\hat{I}$ defines a parabolic subgroup $\hat{P}$ of a simple group $\hat{G}$ of type $A_{2\ell-2}$. Let $\hat{X} = \hat{G}/\hat{P}$. By Proposition 1, the coefficients of $\mu_X + \delta$ appear as the first half of the coefficients of $\mu_{\hat{X}} + \delta$, see Figure 7.

Figure 7: Conversion of type D_ℓ to type $A_{2\ell-2}$

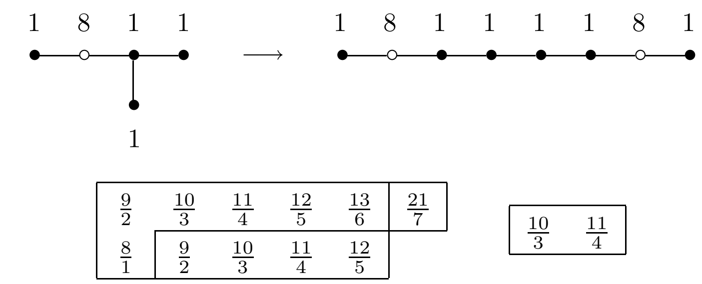


For a fixed i the product $h_i = \prod a_{ij} \times \prod b_{ij}$ can be arranged as

$$\frac{m_i}{1} \cdot \frac{m_i + m_{i+1}}{2} \cdots \frac{m_i + \cdots + m_\ell}{\ell - i + 1} \cdot \frac{s_i + m_{\ell-2}}{\ell - i + 2} \cdots \frac{s_i + m_{\ell-2} + \cdots + m_{i+1}}{2(\ell - i) - 1}$$

where $s_i = m_i + \cdots + m_\ell$. Therefore, the non-trivial terms in $\prod h_i$ correspond to the numbers in the upper left half (above the diagonal) of the rectangular arrays R_{pq} defined for type $A_{2\ell-2}$. These triangular arrays can be broken up into simple sequences as before, see Figure 8. The numbers in the remaining product, $\prod c_i$, are easily seen to form a product of simple sequences by Proposition 1. Therefore, the full product h is a product of simple sequences.
We now assume $m_{\ell-1} < m_\ell$. In this case, the product h is organized in a slightly different way. For fixed i, the previous product h_i is split into two

Figure 8: Type D_ℓ decomposition, $m_{\ell-1} = m_\ell$



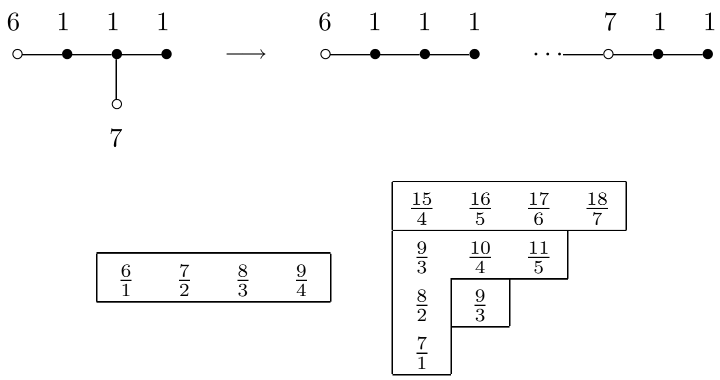
terms with c_i inserted at the beginning of the second term:

$$\frac{m_i}{1} \cdot \frac{m_i + m_{i+1}}{2} \cdots \frac{m_i + \cdots + m_{\ell-1}}{\ell - i}$$

$$\frac{m_i + \cdots + m_{\ell-2} + m_\ell}{\ell - i} \cdot \frac{s_i}{\ell - i + 1} \cdot \frac{s_i + m_{\ell-2}}{\ell - i + 2} \cdots \frac{s_i + m_{\ell-2} + \cdots + m_{i+1}}{2(\ell - i) - 1}$$

Therefore, the non-trivial terms in the product h come from two arrays, the first corresponding to the numbers in the rectangular arrays R_{pq} defined for type $A_{\ell-1}$ and the second corresponding to the numbers in the upper half of certain rectangular arrays R_{pq} defined for type $A_{2\ell-3}$. As before, these rectangular and triangular arrays can be broken up into simple sequences, see Figure 9, and hence the product h is a product of simple sequences.

Figure 9: Type D_ℓ decomposition, $m_{\ell-1} < m_\ell$



It remains to show that equality is obtained only in the designated cases. From Lemmas 1 and 2 is is clear that if $h = 3^n$ then all the simple sequences making up h must have length one and each consists of the number 3. Consequently, $m_i = 3$ for $1 \le i \le \ell$, so that $\mu_X = 2\delta$, and, by Proposition 1, X is a flag manifold. Likewise, if $h = \binom{2n+1}{n}$, then h must be the product of just one simple sequence, $h = S(1,n)$. By Proposition 1, this situation occurs either in type A_n when $m_1 = n+2$ and $m_i = 1$, $2 \le i \le n$ (or $m_n = n+2$ and $m_i = 1$, $1 \le i \le n-1$), or in type C_ℓ when $n = 2\ell - 1$, $m_1 = n+2 = 2\ell+1$, and $m_i = 1$, $2 \le i \le \ell$. In both of these cases the underlying manifold is projective space, $\mathbf{P}^n$. If the degree is $d = 2^n n!$ (resp. $(n+1)^n$), then from (3) and (4), $h = 3^n$ (resp. $\binom{2n+1}{n}$), and the same argument applies. $\square$

References

[1] Akhiezer, D., *Lie Group Actions in Complex Analysis*, Vieweg, Braunschweig, 1995.

[2] Borel, A., *Linear Algebraic Groups*, 2nd ed., Benjamin, New York, 1991.

[3] Borel, A., Hirzebruch, F., *Characteristic classes and homogeneous spaces, part I*, Amer. J. Math. 80 (1958), 459–538.

[4] Borel, A., Hirzebruch, F., *Characteristic classes and homogeneous spaces, part II*, Amer. J. Math. 81 (1959), 315–381.

[5] Borel, A., Remmert, R., *Über kompakte homogene Kählersche Mannigfaltigkeiten*, Math. Ann. 145 (1962), 429–439.

[6] Clemens, H., Ran, Z., *A new method in Fano geometry*, Internat. Math. Res. Notices 2000, no.10, 527–549.

[7] Fulton, W., Harris, J., *Representation Theory*, Springer, Berlin, Heidelberg, New York, 1991.

[8] Iskovskikh, V., Prokhorov, Y., *Fano Varieties*, Algebraic Geometry V, Encyclopaedia of Mathematical Sciences, Vol. 47, Springer, Berlin, Heidelberg, New York, 1999.

[9] Kollár, J., *Rational Curves on Algebraic Varieties*, Springer, Berlin, Heidelberg, New York, 1996.

[10] Kollár, J., Miyaoka, Y., Mori, S., *Rational Connectedness and Boundedness of Fano Manifolds*, J. Diff. Geom. 36 (1992), 765–779.

[11] Snow, D., *The nef value of homogeneous line bundles and related vanishing theorems*, Forum Math. 7 (1995), 385–392.

[12] Snow, D., Winkelmann, J., *Compact complex homogeneous manifolds with large automorphism groups*, Invent. Math. 134 (1998), 139–144.

[13] Snow, D., *A bound for the dimension of the automorphism group of a homogeneous compact complex manifold*, to appear Proc. Amer. Math. Soc.

Dennis Snow
Department of Mathematics
University of Notre Dame
Notre Dame, Indiana 46556
snow.1@nd.edu

Moduli of Prym Curves

Edoardo Ballico, Cinzia Casagrande,
and Claudio Fontanari[1]

Received: May 22, 2003
Revised: March 22, 2004

Communicated by Thomas Peternell

Abstract. Here we focus on the compactification of the moduli space of curves of genus g together with an unramified double cover, constructed by Arnaud Beauville in order to compactify the Prym mapping. We present an alternative description of it, inspired by the moduli space of spin curves of Maurizio Cornalba, and we discuss in detail its main features, both from a geometrical and a combinatorial point of view.

2000 Mathematics Subject Classification: 14H10, 14H30.
Keywords and Phrases: moduli of curves, étale and admissible double covers, spin and Prym curves.

Introduction

Let X be a smooth curve of genus g. As it is well known (see for instance [Bea96], p. 104, or [ACGH85], Appendix B, § 2, 13.), a square root of $\mathcal{O}_X$ corresponds to an unramified double cover of X.
A compactification $\overline{R}_g$ of the moduli space of curves of genus g together with an unramified double cover was constructed by Arnaud Beauville ([Bea77], Section 6; see also [DS81], Theorem 1.1) by means of admissible double covers of stable curves. This moduli space was introduced as a tool to compactify the mapping which associates to a curve plus a 2-sheeted cover the corresponding Prym variety; however, we believe that it is interesting also in its own and worthy of a closer inspection.
Here we explore some of the geometrical and combinatorial properties of $\overline{R}_g$. In order to do that, we present a description of this scheme which is different from the original one and is inspired by the construction performed by Maurizio

Cornalba in [Cor89] of the moduli space of spin curves $\overline{S}_g$. This is a natural compactification over $\overline{\mathcal{M}}_g$ of the space of pairs (X, ζ), where $\zeta \in \operatorname{Pic} X$ is a square root of the canonical bundle K_X.

In Section 1 we define a *Prym curve* to be just the analogue of a spin curve; Cornalba's approach can be easily adapted to our context and allows us to put a structure of projective variety on the set $\overline{Pr}_g$ of isomorphism classes of Prym curves of genus g. This variety has two irreducible components $\overline{Pr}_g^-$ and $\overline{Pr}_g^+$, where $\overline{Pr}_g^- \simeq \overline{\mathcal{M}}_g$ contains "trivial" Prym curves; moreover, by comparing Prym curves and admissible double covers, we give an explicit isomorphism between $\overline{Pr}_g^+$ and $\overline{R}_g$ over $\overline{\mathcal{M}}_g$.

Next, in Section 2 we reproduce the arguments in [Fon02] in order to show that $\overline{Pr}_g$ is endowed with a natural injective morphism into the compactification of the universal Picard variety constructed by Lucia Caporaso in [Cap94], just like $\overline{S}_g$.

Finally, in Section 3 we turn to the combinatorics of $\overline{Pr}_g$. Applying the same approach used in [CC03] for spin curves, we study the ramification of the morphism $\overline{Pr}_g \to \overline{\mathcal{M}}_g$ over the boundary. We describe the numerical properties of the scheme-theoretical fiber Pr_Z over a point $[Z] \in \overline{\mathcal{M}}_g$, which turn out to depend only on the dual graph Γ_Z of Z. From this combinatorial description, it follows that the morphisms $\overline{Pr}_g \to \overline{\mathcal{M}}_g$ and $\overline{S}_g \to \overline{\mathcal{M}}_g$ ramify in a different way.

The moduli space $\overline{R}_g$ of admissible double covers has been studied also by Mira Bernstein in [Ber99], where $\overline{R}_g$ is shown to be of general type for $g = 17, 19, 21, 23$ ([Ber99], Corollary 3.1.7) (for $g \geq 24$ it is obvious, since $\overline{\mathcal{M}}_g$ is).

We work over the field $\mathbb{C}$ of complex numbers.

We wish to thank Lucia Caporaso for many fruitful conversations. We are also grateful to the anonymous referee for pointing out a gap in a previous version of this paper.

1 Prym curves and admissible double covers

1.1. Defining the objects. Let X be a Deligne-Mumford semistable curve and E an irreducible component of X. One says that E is *exceptional* if it is smooth, rational, and meets the other components in exactly two points. Moreover, one says that X is *quasistable* if any two distinct exceptional components of X are disjoint. The stable model of X is the stable curve Z obtained from X by contracting each exceptional component to a point. In the sequel, $\widetilde{X}$ will denote the subcurve $\overline{X \smallsetminus \cup_i E_i}$ obtained from X by removing all exceptional components.

We fix an integer $g \geq 2$.

Definition 1. A *Prym curve* of genus g is the datum of (X, η, β) where X is a quasistable curve of genus g, $\eta \in \operatorname{Pic} X$, and $\beta \colon \eta^{\otimes 2} \to \mathcal{O}_X$ is a homomorphism of invertible sheaves satisfying the following conditions:

(i) η has total degree 0 on X and degree 1 on every exceptional component of X;

(ii) β is non zero at a general point of every non-exceptional component of X.

We say that X is the support of the Prym curve (X, η, β).
An *isomorphism* between two Prym curves (X, η, β) and (X', η', β') is an isomorphism $\sigma\colon X \to X'$ such that there exists an isomorphism $\tau\colon \sigma^*(\eta') \to \eta$ which makes the following diagram commute [2]

$$\begin{tikzcd} \sigma^*(\eta')^{\otimes 2} \arrow[r, "\tau^{\otimes 2}"] \arrow[d, "\sigma^*(\beta')"'] & \eta^{\otimes 2} \arrow[d, "\beta"] \\ \sigma^*(\mathcal{O}_{X'}) \arrow[r, "\sim"] & \mathcal{O}_X. \end{tikzcd}$$

Let (X, η, β) be a Prym curve and let $E_1, \ldots, E_r$ be the exceptional components of X. From the definition it follows that β vanishes identically on all E_i and induces an isomorphism

$$\eta^{\otimes 2}|_{\widetilde{X}} \xrightarrow{\sim} \mathcal{O}_{\widetilde{X}}(-q_1^1 - q_1^2 - \cdots - q_r^1 - q_r^2),$$

where $\{q_i^1, q_i^2\} = \widetilde{X} \cap E_i$ for $i = 1, \ldots, r$. In particular, when X is smooth, η is just a point of order two in the Picard group of X. The number of such points, as it is well-known, is exactly 2^{2g}.
We denote by $\operatorname{Aut}(X, \eta, \beta)$ the group of automorphisms of the Prym curve (X, η, β). As in [Cor89], p. 565, one can show that $\operatorname{Aut}(X, \eta, \beta)$ is finite.
We say that an isomorphism between two Prym curves (X, η, β) and (X, η', β') having the same support is *inessential* if it induces the identity on the stable model of X. We denote by $\operatorname{Aut}_0(X, \eta, \beta) \subseteq \operatorname{Aut}(X, \eta, \beta)$ the subgroup of inessential automorphisms. We have the following

Lemma 2 ([Cor89], Lemma 2.1). *There exists an inessential isomorphism between two Prym curves (X, η, β) and (X, η', β') if and only if*

$$\eta|_{\widetilde{X}} \simeq \eta'|_{\widetilde{X}}.$$

So the set of isomorphism classes of Prym curves supported on X is in bijection with the set of square roots of $\mathcal{O}_{\widetilde{X}}(-q_1^1 - q_1^2 - \cdots - q_r^1 - q_r^2)$ in $\operatorname{Pic} \widetilde{X}$, modulo the action of the group of automorphisms of $\widetilde{X}$ fixing $q_1^1, q_1^2, \ldots, q_r^1, q_r^2$.
A *family of Prym curves* is a flat family of quasistable curves $f\colon \mathcal{X} \to S$ with an invertible sheaf $\boldsymbol{\eta}$ on $\mathcal{X}$ and a homomorphism

$$\boldsymbol{\beta}\colon \boldsymbol{\eta}^{\otimes 2} \longrightarrow \mathcal{O}_{\mathcal{X}}$$

[2] Observe that we are adopting the convention that the datum of τ is not included in the definition of isomorphism, as in [Cor89]. This is different from the convention in [Cor91]; see [Cor91], end of section 2, for a discussion about this.

such that the restriction of these data to any fiber of f gives rise to a Prym curve. An isomorphism between two families of Prym curves $(\mathcal{X} \to S, \boldsymbol{\eta}, \boldsymbol{\beta})$ and $(\mathcal{X}' \to S, \boldsymbol{\eta}', \boldsymbol{\beta}')$ over S is an isomorphism $\sigma\colon \mathcal{X} \to \mathcal{X}'$ over S such that there exists an isomorphism $\tau\colon \sigma^*(\boldsymbol{\eta}') \to \boldsymbol{\eta}$ compatible with the canonical isomorphism between $\sigma^*(\mathcal{O}_{\mathcal{X}'})$ and $\mathcal{O}_{\mathcal{X}}$.

We define the moduli functor associated to Prym curves in the obvious way: $\overline{\mathcal{P}\mathrm{r}}_g$ is the contravariant functor from the category of schemes to the one of sets, which to every scheme S associates the set $\overline{\mathcal{P}\mathrm{r}}_g(S)$ of isomorphism classes of families of Prym curves of genus g over S.

1.2. The universal deformation. Fix a Prym curve (X, η, β), call Z the stable model of X and denote by $E_1, \dots, E_r$ the exceptional components of X. Let $\mathcal{Z}' \to B'$ be the universal deformation of Z, where B' is the unit policylinder in $\mathbb{C}^{3g-3}$ with coordinates $t_1, \dots, t_{3g-3}$ such that $\{t_i = 0\} \subset B'$ is the locus where the node corresponding to E_i persists for $i = 1, \dots, r$. Let B be another unit policylinder in $\mathbb{C}^{3g-3}$ with coordinates $\tau_1, \dots, \tau_{3g-3}$, and consider the map $B \to B'$ given by $t_i = \tau_i^2$ for $1 \le i \le r$ and $t_i = \tau_i$ for $i > r$. Call $\mathcal{Z}$ the pull-back of $\mathcal{Z}'$ to B. For $i \in \{1, \dots, r\}$ the family $\mathcal{Z}_{|\{\tau_i = 0\}} \to \{\tau_i = 0\} \subset B$ has a section V_i, corresponding to the locus of the ith node. Let $\mathcal{X} \to \mathcal{Z}$ be the blow-up of $V_1, \dots, V_r$ and call $\mathcal{E}_1, \dots, \mathcal{E}_r$ the exceptional divisors.

$$\begin{tikzcd}
\mathcal{X} \arrow[r] \arrow[dr] & \mathcal{Z} \arrow[r] \arrow[d] & \mathcal{Z}' \arrow[d] \\
 & B \arrow[r] & B'
\end{tikzcd}$$

The variety $\mathcal{X}$ is smooth and $\mathcal{X} \to B$ is a family of quasistable curves, with X as central fiber. Up to an inessential automorphism, we can assume that $\eta^{\otimes 2} \simeq \mathcal{O}_{\mathcal{X}}(-\sum_i \mathcal{E}_i)_{|X}$ and that this isomorphism is induced by β. By shrinking B if necessary, we can extend η to $\boldsymbol{\eta} \in \operatorname{Pic} \mathcal{X}$ such that $\boldsymbol{\eta}^{\otimes 2} \simeq \mathcal{O}_{\mathcal{X}}(-\sum_i \mathcal{E}_i)$. Denote by $\boldsymbol{\beta}$ the composition of this isomorphism with the natural inclusion $\mathcal{O}_{\mathcal{X}}(-\sum_i \mathcal{E}_i) \hookrightarrow \mathcal{O}_{\mathcal{X}}$. Then $(\mathcal{X} \to B, \boldsymbol{\eta}, \boldsymbol{\beta})$ is a family of Prym curves, and there is a morphism $\psi\colon X \to \mathcal{X}$ which induces an isomorphism of Prym curves between (X, η, β) and the fiber of the family over $b_0 = (0, \dots, 0) \in B$. This family provides a *universal deformation* of (X, η, β):

Theorem 3. *Let $(\mathcal{X}' \to T, \boldsymbol{\eta}', \boldsymbol{\beta}')$ be a family of Prym curves and let $\varphi\colon X \to \mathcal{X}'$ be a morphism which induces an isomorphism of Prym curves between (X, η, β) and the fiber of the family over $t_0 \in T$.*

Then, possibly after shrinking T, there exists a unique morphism $\gamma\colon T \to B$ satisfying the following conditions:

(i) $\gamma(t_0) = b_0$;

(ii) there is a cartesian diagram

$$\begin{tikzcd} \mathcal{X}' \arrow[r, "\delta"] \arrow[d] & \mathcal{X} \arrow[d] \\ T \arrow[r, "\gamma"] & B; \end{tikzcd}$$

(iii) $\boldsymbol{\eta}' \simeq \delta^*(\boldsymbol{\eta})$ *and* $\boldsymbol{\beta}' = \delta^*(\boldsymbol{\beta})$*;*

(iv) $\delta \circ \varphi = \psi$*.*

Since the proof of [Cor89], Proposition 4.6 applies verbatim to our case, we omit the proof of Theorem 3.

1.3. The moduli scheme. Let $\overline{Pr}_g$ be the set of isomorphism classes of Prym curves of genus g. We define a natural structure of analytic variety on $\overline{Pr}_g$ following [Cor89], § 5.
Consider a Prym curve (X, η, β) and its universal deformation $(\mathcal{X} \to B, \boldsymbol{\eta}, \boldsymbol{\beta})$ constructed in 1.2. By the universality, the group $\mathrm{Aut}(X, \eta, \beta)$ acts on B and on $\mathcal{X}$. This action has the following crucial property:

Lemma 4 ([Cor89], Lemma 5.1). *Let $b_1, b_2 \in B$ and let $(X_{b_1}, \eta_{b_1}, \beta_{b_1})$ and $(X_{b_2}, \eta_{b_2}, \beta_{b_2})$ be the fibers of the universal family over b_1 and b_2 respectively. Then there exists $\sigma \in \mathrm{Aut}(X, \eta, \beta)$ such that $\sigma(b_1) = b_2$[3] if and only if the Prym curves $(X_{b_1}, \eta_{b_1}, \beta_{b_1})$ and $(X_{b_2}, \eta_{b_2}, \beta_{b_2})$ are isomorphic.*

Lemma 4 implies that the natural (set-theoretical) map $B \to \overline{Pr}_g$, associating to $b \in B$ the isomorphism class of the fiber over b, descends to a well-defined, injective map

$$J\colon B/\mathrm{Aut}(X, \eta, \beta) \longrightarrow \overline{Pr}_g.$$

This allows to define a complex structure on the subset $\mathrm{Im}\, J \subseteq \overline{Pr}_g$. Since $\overline{Pr}_g$ is covered by these subsets, in order to get a complex structure on $\overline{Pr}_g$ we just have to check that the complex structures are compatible on the overlaps.
This compatibility will follow from the following remark, which is an immediate consequence of the construction of the universal family in 1.2:

- *the family of Prym curves $(\mathcal{X} \to B, \boldsymbol{\eta}, \boldsymbol{\beta})$ is a universal deformation for any of its fibers.*

In fact, assume that there are two Prym curves (X_1, η_1, β_1) and (X_2, η_2, β_2) such that the images of the associated maps J_1, J_2 intersect. Choose a Prym curve (X_3, η_3, β_3) corresponding to a point in $\mathrm{Im}\, J_1 \cap \mathrm{Im}\, J_2$. Let B_i $(i = 1, 2, 3)$ be the basis of the universal deformation of (X_i, η_i, β_i). Then by the remark above, for $i = 1, 2$ there are natural open immersions $h_i\colon B_3 \hookrightarrow B_i$, equivariant with respect to the actions of the automorphism groups. Hence h_i induces an open immersion $\overline{h}_i\colon B_3/\mathrm{Aut}(X_3, \eta_3, \beta_3) \hookrightarrow B_i/\mathrm{Aut}(X_i, \eta_i, \beta_i)$, and $J_3 = J_i \circ \overline{h}_i$.

[3] Where we still denote by σ the automorphism of B induced by σ.

Observe now that the morphisms

$$B/\mathrm{Aut}(X,\eta,\beta) \longrightarrow B'/\mathrm{Aut}(Z)$$

glue together and yield a morphism $p\colon \overline{Pr}_g \to \overline{\mathcal{M}}_g$. Clearly p is finite, as a morphism between analytic varieties (see [Ray71]). Hence $\overline{Pr}_g$ is projective, because $\overline{\mathcal{M}}_g$ is. The variety $\overline{Pr}_g$ has finite quotient singularities; in particular, it is normal.

The degree of p is 2^{2g}. The fiber over a smooth curve Z is just the set of points of order two in its Picard group, modulo the action of $\mathrm{Aut}(Z)$ if non trivial. When Z is a stable curve, the set-theoretical fiber over $[Z]$ consists of isomorphism classes of Prym curves (X,η,β) such that the stable model of X is Z. In section 3 we will describe precisely the scheme-theoretical fiber over $[Z]$, following [Cor89] and [CC03]. We will show that p is étale over $\overline{\mathcal{M}}_g^0 \smallsetminus D_{irr}$, where $\overline{\mathcal{M}}_g^0$ is the locus of stable curves with trivial automorphism group and D_{irr} is the boundary component whose general member is an irreducible curve with one node.

Finally, $\overline{Pr}_g$ is a coarse moduli space for the functor $\overline{\mathcal{P}\mathrm{r}}_g$. For any family of Prym curves over a scheme T, the associated moduli morphism $T \to \overline{Pr}_g$ is locally defined by Theorem 3.

Let $\overline{Pr}_g^-$ be the closed subvariety of $\overline{Pr}_g$ consisting of classes of Prym curves (X,η,β) where $\eta \simeq \mathcal{O}_X$. Observe that when η is trivial, the curve X is stable. So $\overline{Pr}_g^-$ is the image of the obvious section of $p\colon \overline{Pr}_g \to \overline{\mathcal{M}}_g$, and it is an irreducible (and connected) component of $\overline{Pr}_g$, isomorphic to $\overline{\mathcal{M}}_g$.

Let $\overline{Pr}_g^+$ be the complement of $\overline{Pr}_g^-$ in $\overline{Pr}_g$, and denote by Pr_g^+ its open subset consisting of classes of Prym curves supported on smooth curves. Then Pr_g^+ parametrizes connected unramified double covers of smooth curves of genus g; it is well-known that this moduli space is irreducible, being a finite quotient of the moduli space of smooth curves of genus g with a level 2 structure, which is irreducible by [DM69]. So $\overline{Pr}_g^+$ is an irreducible component of $\overline{Pr}_g$.

1.4. Admissible double covers. Consider a pair (C,i) where C is a stable curve of genus $2g-1$ and i is an involution of C such that:

- the set I of fixed points of i is contained in $\operatorname{Sing} C$;
- for any fixed node, i does not exchange the two branches of the curve.

Then the quotient $Z := C/i$ is a stable curve of genus g, and $\pi\colon C \to Z$ is a finite morphism of degree 2, étale over $Z \smallsetminus \pi(I)$. This is called an *admissible double cover*. Remark that π is not a cover in the usual sense, since it is not flat at I.

The moduli space $\overline{R}_g$ of admissible double covers of stable curves of genus g is constructed [Bea77], Section 6 (see also [DS81, ABH01]), as the moduli space for pairs (C,i) as above.

An isomorphism of two admissible covers $\pi_1\colon C_1 \to Z_1$ and $\pi_2\colon C_2 \to Z_2$ is an isomorphism $\varphi\colon Z_1 \xrightarrow{\sim} Z_2$ such that there exists[4] an isomorphism $\widetilde{\varphi}\colon C_1 \xrightarrow{\sim} C_2$ with $\pi_2 \circ \widetilde{\varphi} = \varphi \circ \pi_1$.

We denote by $\mathrm{Aut}(C \to Z)$ the automorphism group of the admissible cover $C \to Z$, so $\mathrm{Aut}(C \to Z) \subseteq \mathrm{Aut}(Z)$. All elements of $\mathrm{Aut}(C \to Z)$ are induced by automorphisms of C, different from i, and that commute with i.

Let (C, i) be as above; we describe its universal deformation. Let $\mathcal{C}' \to W'$ be a universal deformation of C. By the universality, there are compatible involutions i' of W' and $\boldsymbol{i}'$ of $\mathcal{C}'$, extending the action of i on the central fiber. Let $W \subset W'$ be the locus fixed by i', $\mathcal{C} \to W$ the induced family and $\boldsymbol{i}$ the restriction of $\boldsymbol{i}'$ to $\mathcal{C}$. Then $(\mathcal{C}, \boldsymbol{i}) \to W$ is a universal deformation of (C, i) and the corresponding family of admissible double covers is $\mathcal{C} \to \mathcal{Q} := \mathcal{C}/\boldsymbol{i} \to W$.

We are going to show that $\overline{R}_g$ is isomorphic over $\overline{\mathcal{M}}_g$ to the irreducible component $\overline{Pr}_g^+$ of $\overline{Pr}_g$.

First of all we define a map Φ from the set of non trivial Prym curves of genus g to the set of admissible double covers of stable curves of genus g.

Let $\xi = (X, \eta, \beta)$ be a Prym curve with $\eta \not\simeq \mathcal{O}_X$; then $\Phi(\xi)$ will be an admissible double cover of the stable model Z of X, constructed as follows. The homomorphism β induces an isomorphism

$$\eta_{|\widetilde{X}}^{\otimes(-2)} \simeq \mathcal{O}_{\widetilde{X}}(q_1^1 + q_1^2 + \cdots + q_r^1 + q_r^2).$$

This determines a double cover $\widetilde{\pi}\colon \widetilde{C} \to \widetilde{X}$, ramified over $q_1^1, q_1^2, \ldots, q_r^1, q_r^2$, which are smooth points of $\widetilde{X}$. Now call C_ξ the stable curve obtained identifying $\widetilde{\pi}^{-1}(q_i^1)$ with $\widetilde{\pi}^{-1}(q_i^2)$ for all $i = 1, \ldots, r$. Then the induced map $C_\xi \to Z$ is the admissible double cover $\Phi(\xi)$.

Now consider two Prym curves ξ_1 and ξ_2 supported respectively on X_1 and X_2. Suppose that $\sigma\colon X_1 \xrightarrow{\sim} X_2$ induces an isomorphism between ξ_1 and ξ_2. Let $\overline{\sigma}\colon Z_1 \xrightarrow{\sim} Z_2$ be the induced isomorphism between the stable models. Then it is easy to see that $\overline{\sigma}$ is an isomorphism between the admissible covers $\Phi(\xi_1)$ and $\Phi(\xi_2)$. Moreover, any isomorphism between $\Phi(\xi_1)$ and $\Phi(\xi_2)$ is obtained in this way. Hence we have an exact sequence of automorphism groups:

$$1 \to \mathrm{Aut}_0(\xi) \to \mathrm{Aut}(\xi) \to \mathrm{Aut}(C_\xi \to Z) \to 1. \tag{1}$$

We show that Φ is surjective. Let $C \to Z$ be an admissible double cover, $I \subset C$ the set of fixed points of the involution and $J \subset Z$ their images. Let $\widetilde{C} \to C$ and $\nu\colon \widetilde{X} \to Z$ be the normalizations of C at I and of Z at J respectively. Then i extends to an involution on $\widetilde{C}$, whose quotient is $\widetilde{X}$, namely: $\widetilde{C} \to \widetilde{X}$ is a double cover, ramified over $q_1^1, q_1^2, \ldots, q_r^1, q_r^2$, where $r = |J|$ and $\nu(q_i^1) = \nu(q_i^2) \in J$ for $i = 1, \ldots, r$. Let $L \in \mathrm{Pic}\,\widetilde{X}$ be the associated line bundle, satisfying $L^{\otimes 2} \simeq \mathcal{O}_{\widetilde{X}}(q_1^1 + q_1^2 + \cdots + q_r^1 + q_r^2)$. Finally let X be the quasistable curve obtained by attaching to $\widetilde{X}$ r rational components $E_1, \ldots, E_r$ such that

[4]Given φ, there are exactly two choices for $\widetilde{\varphi}$; if $\widetilde{\varphi}$ is one, the other is $\widetilde{\varphi} \circ i_1$.

$E_i \cap \widetilde{X} = \{q_i^1, q_i^2\}$. Choose $\eta \in \operatorname{Pic} X$ having degree 1 on all E_i and such that $\eta|_{\widetilde{X}} = L^{\otimes(-1)}$. Let $\beta\colon \eta^{\otimes 2} \to \mathcal{O}_X$ be a homomorphism which agrees with $\eta|_{\widetilde{X}} \simeq \mathcal{O}_{\widetilde{X}}(-q_1^1 - q_1^2 - \cdots - q_r^1 - q_r^2) \hookrightarrow \mathcal{O}_X$ on $\widetilde{X}$. Then $\xi = (X, \eta, \beta)$ is a Prym curve with $\eta \not\simeq \mathcal{O}_X$, and $C \to Z$ is $\Phi(\xi)$. For different choices of η, the corresponding Prym curves differ by an inessential isomorphism.

Proposition 5. *The map Φ just defined induces an isomorphism*

$$\widehat{\Phi}\colon \overline{Pr}_g^+ \longrightarrow \overline{R}_g$$

over $\overline{\mathcal{M}}_g$.

Proof. By what precedes, Φ induces a bijection $\widehat{\Phi}\colon \overline{Pr}_g^+ \to \overline{R}_g$. The statement will follow if we prove that $\widehat{\Phi}$ is a local isomorphism at every point of $\overline{Pr}_g^+$.
Fix a point $\xi = (X, \eta, \beta) \in \overline{Pr}_g^+$ and consider its universal deformation $(\mathcal{X} \to B, \boldsymbol{\eta}, \boldsymbol{\beta})$ constructed in 1.2. Keeping the notations of 1.2, the line bundle $\boldsymbol{\eta}^{\otimes(-1)}$ determines a double cover $\overline{\mathcal{P}} \to \mathcal{X}$, ramified over $\mathcal{E}_1, \dots, \mathcal{E}_r$. The divisor $\mathcal{E}_i$ is a $\mathbb{P}^1$-bundle over $V_i \subset B$, and the restriction of its normal bundle to a non trivial fiber F is $(\mathcal{N}_{\mathcal{E}_i/\mathcal{X}})_{|F} \simeq \mathcal{O}_{\mathbb{P}^1}(-2)$. The inverse image $\overline{\mathcal{E}}_i$ of $\mathcal{E}_i$ in $\overline{\mathcal{P}}$ is again a $\mathbb{P}^1$-bundle over $V_i \subset B$, but now the restriction of its normal bundle to a non trivial fiber $\overline{F}$ is $(\mathcal{N}_{\overline{\mathcal{E}}_i/\overline{\mathcal{P}}})_{|\overline{F}} \simeq \mathcal{O}_{\mathbb{P}^1}(-1)$. Let $\overline{\mathcal{P}} \to \mathcal{P}$ be the blow-down of $\overline{\mathcal{E}}_1, \dots, \overline{\mathcal{E}}_r$. We get a diagram

$$\begin{tikzcd}
\overline{\mathcal{P}} \arrow[r] \arrow[d] & \mathcal{P} \arrow[d] & \\
\mathcal{X} \arrow[r] & \mathcal{Z} \arrow[r] & B
\end{tikzcd}$$

where $\mathcal{P} \to \mathcal{Z} \to B$ is a family of admissible double covers whose central fiber is $C_\xi \to Z$. Therefore, up to shrinking B, there exists a morphism $B \to W$ such that $\mathcal{P} \to \mathcal{Z} \to B$ is obtained by pull-back from the universal deformation $\mathcal{C} \to \mathcal{Q} \to W$ of $C_\xi \to Z$. Now notice that $\mathcal{Q} \to W$ is a family of stable curves of genus g, with Z as central fiber: so (again up to shrinking) it must be a pull-back of the universal deformation $\mathcal{Z}' \to B'$. In the end we get a diagram:

$$\begin{tikzcd}
\overline{\mathcal{P}} \arrow[r] \arrow[d] & \mathcal{P} \arrow[r] \arrow[d] & \mathcal{C} \arrow[d] & & \text{(genus } 2g-1) \\
\mathcal{X} \arrow[r] \arrow[dr] & \mathcal{Z} \arrow[r] \arrow[d] & \mathcal{Q} \arrow[r] \arrow[d] & \mathcal{Z}' \arrow[d] & \text{(genus } g) \\
& B \arrow[r, "\varphi"] \arrow[loop below, "\mathrm{Aut}(\xi)"] & W \arrow[r, "\psi"] \arrow[loop below, "\mathrm{Aut}(C_\xi \to Z)"] & B' \arrow[loop below, "\mathrm{Aut}(Z)"] & \text{(bases)}
\end{tikzcd}$$

We can assume that φ and ψ are surjective. Observe that both maps are equivariant with respect to the actions of the automorphism groups indicated in the diagram.
Clearly φ is just the restriction of Φ to the set of Prym curves parametrized by B.
Now by (1), $\varphi(b_1) = \varphi(b_2)$ if and only if there exists an inessential isomorphism between $(X_{b_1}, \eta_{b_1}, \beta_{b_1})$ and $(X_{b_2}, \eta_{b_2}, \beta_{b_2})$. Hence φ induces an equivariant isomorphism $\widehat{\varphi}$:

$$\begin{tikzcd} B/\mathrm{Aut}_0(\xi) \arrow[r, "\widehat{\varphi}"] \arrow[loop below, "\mathrm{Aut}(\xi)/\mathrm{Aut}_0(\xi)"] & W \arrow[r] \arrow[loop below, "\mathrm{Aut}(C_\xi \to Z)"] & B' \arrow[loop below, "\mathrm{Aut}(Z)"] \end{tikzcd}$$

and finally if we mod out by all the automorphism groups, we get

$$\begin{tikzcd} B/\mathrm{Aut}(\xi) \arrow[r, "\sim"] \arrow[d, hook] & W/\mathrm{Aut}(C_\xi \to Z) \arrow[r] \arrow[d, hook] & B'/\mathrm{Aut}Z \arrow[d, hook] \\ \overline{Pr}_g^{\,+} \arrow[r, "\widehat{\Phi}"] & \overline{R}_g & \overline{\mathcal{M}}_g. \end{tikzcd}$$

This shows that $\widehat{\Phi}$ is a local isomorphism in ξ. ■

2 Embedding $\overline{Pr}_g$ in the compactified Picard variety

Let $g \geq 3$. For every integer d, there is a *universal Picard variety*

$$P_{d,g} \longrightarrow \mathcal{M}_g^0$$

whose fiber $J^d(X)$ over a point X of $\mathcal{M}_g^0$ parametrizes line bundles on X of degree d, modulo isomorphism. Denote by Pr_g^0 the inverse image of $\mathcal{M}_g^0$ under the finite morphism $\overline{Pr}_g \to \overline{\mathcal{M}}_g$; then we have a commutative diagram

$$\begin{tikzcd} Pr_g^0 \arrow[r, hook] \arrow[dr] & P_{0,g} \arrow[d] \\ & \mathcal{M}_g^0. \end{tikzcd}$$

Assume $d \geq 20(g-1)$; this is not a real restriction, since for all $t \in \mathbb{Z}_{\geq 0}$ there is a natural isomorphism $P_{d,g} \cong P_{d+t(2g-2),g}$. Then $P_{d,g}$ has a natural compactification $\overline{P}_{d,g}$, endowed with a natural morphism $\phi_d \colon \overline{P}_{d,g} \to \overline{\mathcal{M}}_g$, such that $\phi_d^{-1}(\mathcal{M}_g^0) = P_{d,g}$. It was constructed in [Cap94] as a GIT quotient

$$\pi_d \colon H_d \longrightarrow H_d /\!/ G = \overline{P}_{d,g},$$

where $G = SL(d-g+1)$ and

$$H_d := \{h \in \mathrm{Hilb}_{d-g}^{dx-g+1} \mid h \text{ is } G\text{-semistable}$$
$$\text{and the corresponding curve is connected } \}$$

(the action of G is linearized by a suitable embedding of $\mathrm{Hilb}_{d-g}^{dx-g+1}$ in a Grassmannian).
Fix now and in the sequel an integer $t \geq 10$ and define

$$\begin{aligned} K_{2t(g-1)} := \{h \in \mathrm{Hilb}_{2t(g-1)-g}^{2t(g-1)x-g+1} | & \text{ there is a Prym curve } (X,\eta,\beta) \text{ and} \\ & \text{an embedding } h_t \colon X \to \mathbb{P}^{2t(g-1)-g} \text{ induced by } \eta \otimes \omega_X^{\otimes t}, \\ & \text{such that } h \text{ is the Hilbert point of } h_t(X)\}. \end{aligned}$$

Our result is the following.

Theorem 6. *The set $K_{2t(g-1)}$ is contained in $H_{2t(g-1)}$; consider its projection*

$$\Pi_t := \pi_{2t(g-1)}(K_{2t(g-1)}) \subset \overline{P}_{2t(g-1),g}.$$

There is a natural injective morphism

$$f_t \colon \overline{Pr}_g \longrightarrow \overline{P}_{2t(g-1),g}$$

whose image is Π_t.

In particular, the Theorem implies that Π_t is a closed subvariety of $\overline{P}_{2t(g-1),g}$.
The proof of Theorem 6 will be achieved in several steps and will take the rest of this section. The argument is the one used in [Fon02] to show the existence of an injective morphism $\overline{S_g} \to \overline{P}_{(2t+1)(g-1),g}$ of the moduli space of spin curves in the corresponding compactified Picard variety.
One can define (see [Cap94], §8.1) the contravariant functor $\overline{\mathcal{P}}_{d,g}$ from the category of schemes to the one of sets, which to every scheme S associates the set $\overline{\mathcal{P}}_{d,g}(S)$ of equivalence classes of polarized families of quasistable curves of genus g

$$f \colon (\mathcal{X},\mathcal{L}) \longrightarrow S$$

such that $\mathcal{L}$ is a relatively very ample line bundle of degree d whose multidegree satisfies the following Basic Inequality on each fiber.

Definition 7. Let $X = \bigcup_{i=1}^n X_i$ be a projective, nodal, connected curve of arithmetic genus g, where the X_i's are the irreducible components of X. We say that the multidegree $(d_1,\ldots,d_n)$ satisfies the *Basic Inequality* if for every complete subcurve Y of X of arithmetic genus g_Y we have

$$m_Y \leq d_Y \leq m_Y + k_Y$$

where

$$d_Y = \sum_{X_i \subseteq Y} d_i, \quad k_Y = |Y \cap \overline{X \smallsetminus Y}| \quad \text{and} \quad m_Y = \frac{d}{g-1}\left(g_Y - 1 + \frac{k_Y}{2}\right) - \frac{k_Y}{2}$$

(see [Cap94] p. 611 and p. 614).

Two families over S, $(\mathcal{X}, \mathcal{L})$ and $(\mathcal{X}', \mathcal{L}')$ are equivalent if there exists an S-isomorphism $\sigma\colon \mathcal{X} \to \mathcal{X}'$ and a line bundle M on S such that $\sigma^*\mathcal{L}' \cong \mathcal{L} \otimes f^*M$. By [Cap94], Proposition 8.1, there is a morphism of functors:

$$\overline{\mathcal{P}}_{d,g} \longrightarrow \mathit{Hom}(\,\cdot\,, \overline{P}_{d,g}) \tag{2}$$

and $\overline{P}_{d,g}$ coarsely represents $\overline{\mathcal{P}}_{d,g}$ if and only if

$$(d - g + 1, 2g - 2) = 1. \tag{3}$$

Proposition 8. *For every integer $t \geq 10$ there is a natural morphism:*

$$f_t\colon \overline{Pr}_g \longrightarrow \overline{P}_{2t(g-1),g}.$$

Proof. First of all, notice that in this case (3) does not hold, so the points of $\overline{P}_{2t(g-1),g}$ are *not* in one-to-one correspondence with equivalence classes of very ample line bundles of degree $2t(g-1)$ on quasistable curves, satisfying the Basic Inequality (see [Cap94], p. 654). However, we claim that the thesis can be deduced from the existence of a morphism of functors:

$$F_t\colon \overline{\mathcal{P}\mathrm{r}}_g \longrightarrow \overline{\mathcal{P}}_{2t(g-1),g}. \tag{4}$$

Indeed, since $\overline{Pr}_g$ coarsely represents $\overline{\mathcal{P}\mathrm{r}}_g$, any morphism of functors $\overline{\mathcal{P}\mathrm{r}}_g \to \mathit{Hom}(\,\cdot\,, T)$ induces a morphism of schemes $\overline{Pr}_g \to T$, so the claim follows from (2). Now, a morphism of functors as (4) is the datum for any scheme S of a set-theoretical map

$$F_t(S)\colon \overline{\mathcal{P}\mathrm{r}}_g(S) \longrightarrow \overline{\mathcal{P}}_{2t(g-1),g}(S),$$

satisfying obvious compatibility conditions. Let us define $F_t(S)$ in the following way:

$$(f\colon \mathcal{X} \to S, \boldsymbol{\eta}, \boldsymbol{\beta}) \;\mapsto\; (f\colon (\mathcal{X}, \boldsymbol{\eta} \otimes \omega_f^{\otimes t}) \to S).$$

In order to prove that $F_t(S)$ is well-defined, the only non-trivial matter is to check that the multidegree of $\boldsymbol{\eta} \otimes \omega_f^{\otimes t}$ satisfies the Basic Inequality on each fiber, so the thesis follows from the next Lemma. ■

Lemma 9. *Let (X, η, β) be a Prym curve. If Y is a complete subcurve of X and d_Y is the degree of $(\eta \otimes \omega_X^{\otimes t})|_Y$, then $m_Y \leq d_Y \leq m_Y + k_Y$ in the notation of the Basic Inequality. Moreover, if $d_Y = m_Y$ then $\widetilde{k}_Y := |\widetilde{Y} \cap \overline{\widetilde{X} \smallsetminus \widetilde{Y}}| = 0$.*

Proof. In the present case, the Basic Inequality simplifies as follows:

$$-\frac{k_Y}{2} \leq e_Y \leq \frac{k_Y}{2},$$

where $e_Y := \deg \eta|_Y$. By the definition of a Prym curve, the degree e_Y depends only on the exceptional components of X intersecting Y.
For any exceptional component E of X with $E \subseteq \overline{X \setminus Y}$, let $m := |E \cap Y|$. The contribution of E to k_Y is m, while its contribution to e_Y is $-\frac{m}{2}$.
Next, for any exceptional component E of X with $E \subseteq Y$, let $l := |E \cap \overline{X \setminus Y}|$. The contribution of E to k_Y is l, while its contribution to e_Y is $1 - \frac{2-l}{2} = \frac{l}{2}$.
Summing up, we see that the Basic Inequality holds. Finally, if $\widetilde{k}_Y \neq 0$, then there exists a non-exceptional component of X intersecting Y. Such a component contributes at least 1 to k_Y, but it does not affect e_Y; hence $-\frac{k_Y}{2} < e_Y$ and the proof is over. ∎

By applying [Cap94], Proposition 6.1, from the first part of Lemma 9 we deduce

$$K_{2t(g-1)} \subset H_{2t(g-1)}.$$

Moreover, the second part of the same Lemma provides a crucial information on Hilbert points corresponding to Prym curves.

Lemma 10. *If $h \in K_{2t(g-1)}$, then the orbit of h is closed in the semistable locus.*

Proof. Let (X, η, β) be a Prym curve such that h is the Hilbert point of an embedding $h_t \colon X \to \mathbb{P}^{2t(g-1)-g}$ induced by $\eta \otimes \omega_X^{\otimes t}$. Just recall the first part of [Cap94], Lemma 6.1, which says that the orbit of h is closed in the semistable locus if and only if $\widetilde{k}_Y = 0$ for every subcurve Y of X such that $d_Y = m_Y$, so the thesis is a direct consequence of Lemma 9. ∎

Proof of Theorem 6. It is easy to check that $f_t(\overline{Pr}_g) = \Pi_t$. Indeed, if $(X, \eta, \beta) \in \overline{Pr}_g$, then any choice of a base for $H^0(X, \eta \otimes \omega_X^{\otimes t})$ gives an embedding $h_t \colon X \to \mathbb{P}^{2t(g-1)-g}$ and $f_t(X, \eta, \beta) = \pi_{2t(g-1)}(h)$, where $h \in K_{2t(g-1)}$ is the Hilbert point of $h_t(X)$. Conversely, if $\pi_{2t(g-1)}(h) \in \Pi_t$, then there is a Prym curve (X, η, β) and an embedding $h_t \colon X \to \mathbb{P}^{2t(g-1)-g}$ such that h is the Hilbert point of $h_t(X)$ and $f_t(X, \eta, \beta) = \pi_{2t(g-1)}(h)$.
Next we claim that f_t is injective. Indeed, let (X, η, β) and (X', η', β') be two Prym curves and assume that $f_t(X, \eta, \beta) = f_t(X', \eta', \beta')$. Choose bases for $H^0(X, \eta \otimes \omega_X^{\otimes t})$ and $H^0(X', \eta' \otimes \omega_{X'}^{\otimes t})$ and embed X and X' in $\mathbb{P}^{2t(g-1)-g}$. If h and h' are the corresponding Hilbert points, then $\pi_{2t(g-1)}(h) = \pi_{2t(g-1)}(h')$ and the Fundamental Theorem of GIT implies that $\overline{O_G(h)}$ and $\overline{O_G(h')}$ intersect in the semistable locus. It follows from Lemma 10 that $O_G(h) \cap O_G(h') \neq \emptyset$, so $O_G(h) = O_G(h')$ and there is an isomorphism $\sigma \colon (X, \eta, \beta) \to (X', \eta', \beta')$. ∎

Observe that Theorem 6 and Lemma 10 imply that $K_{2t(g-1)}$ is a constructible set in $H_{2t(g-1)}$.

3 Fiberwise description

Let Z be a stable curve of genus g. We recall that the dual graph Γ_Z of Z is the graph whose vertices are the irreducible components of Z and whose edges are the nodes of Z. The first Betti number of Γ_Z is $b_1(\Gamma_Z) = \delta - \gamma + 1 = g - g^\nu$, where δ is the number of nodes of Z, γ the number of its irreducible components and g^ν the genus of its normalization.
We denote by Pr_Z the scheme parametrizing Prym curves (X, η, β) such that the stable model of X is Z, modulo inessential isomorphisms, and by S_Z the analogue for spin curves. Since by Lemma 2 the homomorphism β is not relevant in determining the inessential isomorphism class of (X, η, β), in this section we will omit it and just write (X, η).
When $\mathrm{Aut}(Z) = \{\mathrm{Id}_Z\}$, Pr_Z is the scheme-theoretical fiber over $[Z]$ of the morphism $p\colon \overline{Pr}_g \to \overline{\mathcal{M}}_g$. Recall that p is finite of degree 2^{2g}, and étale over $\mathcal{M}_g^0$.
For any 0-dimensional scheme P we denote by $L(P)$ the set of integers occurring as multiplicities of components of P.
In this section we describe the numerical properties of Pr_Z, namely the number of irreducible components and their multiplicities, showing that they depend only on the dual graph Γ_Z of Z. Using this, we give some properties of $L(Pr_Z)$, and show that in some cases the set of multiplicities $L(Pr_Z)$ gives informations on Z. In particular, we show that the morphism $\overline{Pr}_g \to \overline{\mathcal{M}}_g$ is étale over $\overline{\mathcal{M}}_g^0 \smallsetminus D_{irr}$.
We use the techniques and results of [CC03], where the same questions about the numerics of S_Z are studied (see also [CS03], § 3). Quite surprisingly, the schemes P_Z and S_Z are not isomorphic in general.
Finally we will show with an example that, differently from the case of spin curves, the set of multiplicities $L(Pr_Z)$ appearing in Pr_Z does not always identify curves having two smooth components.
Let X be a quasistable curve having Z as stable model and consider the set

$$\Delta_X := \{z \in \mathrm{Sing}\, Z \mid z \text{ is } \textit{not} \text{ the image of an exceptional component of } X\}.$$

Given Z, the quasistable curve X is determined by Δ_X, or equivalently by

$$\Delta_X^c := \mathrm{Sing}\, Z \smallsetminus \Delta_X = \{\text{images in } Z \text{ of the exceptional components of } X\}.$$

Remark that any subset of $\mathrm{Sing}\, Z$ can be seen as a subgraph of the dual graph Γ_Z of Z.
We recall that the *valency* of a vertex of a graph is the number of edges ending in that vertex and a graph Γ is *eulerian* if it has all even valencies. Thus Γ_Z is eulerian if and only if for any irreducible component C of Z, $|C \cap \overline{Z \smallsetminus C}|$ is even. The set $\mathcal{C}_\Gamma$ of all eulerian subgraphs of Γ is called the cycle space of Γ. There is a natural identification of $\mathcal{C}_\Gamma$ with $H_1(\Gamma, \mathbb{Z}_2)$, so $|\mathcal{C}_\Gamma| = 2^{b_1(\Gamma)}$ (see [CC03]). Reasoning as in [CC03], Section 1.3, we can show the following:

Proposition 11. *Let X be a quasistable curve having Z as stable model. The curve X is the support of a Prym curve if and only if Δ_X^c is eulerian. If so, there are $2^{2g^\nu+b_1(\Delta_X)}$ different choices for $\eta \in \operatorname{Pic} X$ such that $(X,\eta) \in Pr_Z$. For each such η, the point (X,η) has multiplicity $2^{b_1(\Gamma_Z)-b_1(\Delta_X)}$ in Pr_Z.*

Hence the number of irreducible components of Pr_Z is

$$2^{2g^\nu} \cdot \sum_{\Sigma \in \mathcal{C}_{\Gamma_Z}} 2^{b_1(\Sigma^c)},$$

and its set of multiplicities is given by

$$L(Pr_Z) = \{2^{b_1(\Gamma_Z)-b_1(\Delta)} \mid \Delta^c \in \mathcal{C}_{\Gamma_Z}\}.$$

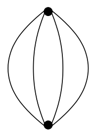
Remark that since $|\mathcal{C}_{\Gamma_Z}| = 2^{b_1(\Gamma_Z)}$, we can check immediately from the proposition that the length of Pr_Z is

$$\sum_{\Sigma \in \mathcal{C}_{\Gamma_Z}} (2^{2g^\nu+b_1(\Sigma^c)} \cdot 2^{b_1(\Gamma_Z)-b_1(\Sigma^c)}) = 2^{b_1(\Gamma_Z)} \cdot 2^{2g^\nu+b_1(\Gamma_Z)} = 2^{2g}.$$

As a consequence of Proposition 11, we see that

- *a point (X,η) in Pr_Z is non reduced if and only if X is non stable.*

Example (curves having two smooth components). Let $Z = C_1 \cup C_2$, C_i smooth irreducible, $|C_1 \cap C_2| = \delta \geq 2$.



Γ_Z

Let X be a quasistable curve having Z as stable model and let Δ_X be the corresponding subset of $\operatorname{Sing} Z$. The subgraph Δ_X^c is eulerian if and only if $|\Delta_X^c|$ is even. Therefore X is support of a Prym curve if and only if it has an even number $2r$ of exceptional components. If so, for each choice of $\eta \in \operatorname{Pic} X$ such that $(X,\eta) \in Pr_Z$, this point will have multiplicity $2^{b_1(\Gamma_Z)-b_1(\Delta_X)}$. We have $b_1(\Gamma_Z) = \delta - 1$ and $|\Delta_X| = \delta - 2r$, so

$$b_1(\Delta_X) = \begin{cases} \delta - 2r - 1 & \text{if } 2r \leq \delta - 2, \\ 0 & \text{if } \delta - 1 \leq 2r \leq \delta \end{cases}$$

and we get

$$L(Pr_Z) = \{2^{2r} \mid 0 \leq r \leq \frac{1}{2}\delta - 1\} \cup \{2^{\delta-1}\}.$$

Proposition 12 (combinatorial properties of $L(Pr_Z)$). *The following properties hold:*

(1) $1 \in L(Pr_Z)$*;*

(2) $\max L(Pr_Z) = 2^{b_1(\Gamma_Z)}$*;*

(3) $2^g \in L(Pr_Z)$ *if and only if* Z *has only rational components;*

(4) Pr_Z *is reduced if and only if* Z *is of compact type;*

(5) if Γ_Z *is an eulerian graph, then* $L(Pr_Z) = L(S_Z)$*.*

Proof. (1) Choosing $\Delta_X = \Gamma_Z$, we get $X = Z$; since the empty set is trivially in $\mathcal{C}_{\Gamma_Z}$, there always exists $\eta \in \operatorname{Pic} Z$ such that $(Z, \eta) \in Pr_Z$. This η is a square root of $\mathcal{O}_Z$; there are $2^{2g^\nu + b_1(\Gamma_Z)}$ choices for it, and it will appear with multiplicity 1 in Pr_Z. So $1 \in L(Pr_Z)$.
(2) From Proposition 11 we get $\max L(Pr_Z) \leq 2^{b_1(\Gamma_Z)}$. Set $M = \max\{b_1(\Sigma) \,|\, \Sigma \in \mathcal{C}_{\Gamma_Z}\}$ and let $\Sigma_0 \in \mathcal{C}_{\Gamma_Z}$ be such that $b_1(\Sigma_0) = M$. By Proposition 11, we know that $2^{b_1(\Gamma_Z) - b_1(\Sigma_0^c)} \in L(Pr_Z)$. We claim that $b_1(\Sigma_0^c) = 0$. Indeed, if not, Σ_0^c contains a subgraph σ with $b_1(\sigma) = 1$ and having all valencies equal to 2. Then $\Sigma_0 \cup \sigma \in \mathcal{C}_{\Gamma_Z}$ and $b_1(\Sigma_0 \cup \sigma) > M$, a contradiction. Hence we have points of multiplicity $2^{b_1(\Gamma_Z)}$ in Pr_Z, so $\max L(Pr_Z) = 2^{b_1(\Gamma_Z)}$.
Property (3) is immediate from (2), since $b_1(\Gamma_Z) = g$ if and only $g^\nu = 0$.
Also property (4) is immediate from (2), because $L(Pr_Z) = \{1\}$ if and only if $b_1(\Gamma_Z) = 0$.
(5) Assume that Γ_Z is eulerian. Then $\Delta_X^c \in \mathcal{C}_{\Gamma_Z}$ if and only if $\Delta_X \in \mathcal{C}_{\Gamma_Z}$, so we have
$$L(Pr_Z) = L(S_Z) = \{2^{b_1(\Gamma_Z) - b_1(\Delta_X)} \,|\, \Delta_X \in \mathcal{C}_{\Gamma_Z}\}$$
(see [CC03] for the description of $L(S_Z)$). ■

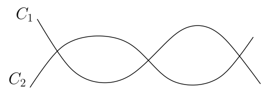
Property (4) implies the following

Corollary 13. *The morphism* $p\colon \overline{Pr}_g \to \overline{\mathcal{M}}_g$ *is étale over* $\overline{\mathcal{M}}_g^0 \smallsetminus D_{irr}$*.*

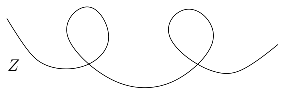
Consider now property (1) of Proposition 12. It shows, in particular, that in general Pr_Z and S_Z are not isomorphic and do not have the same set of multiplicities: indeed, for spin curves, it can very well happen that $1 \notin L(S_Z)$ (see example after Corollary 14).
The following shows that in some cases, $L(Pr_Z)$ gives informations on Z.

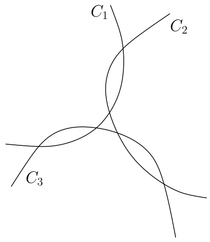
Corollary 14. *Let* Z *be a stable curve and* $\nu\colon Z^\nu \to Z$ *its normalization. Assume that for every irreducible component* C *of* Z*, the number* $|\nu^{-1}(C \cap \operatorname{Sing} Z)|$ *is even and at least* 4*.*

(i) If $2^{b_1(\Gamma_Z)-2} \notin L(Pr_Z)$*, then* $Z = C_1 \cup C_2$*, with* C_1 *and* C_2 *smooth and irreducible.*



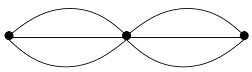
(ii) If $2^{b_1(\Gamma_Z)-3} \notin L(Pr_Z)$*, then either* Z *is irreducible with two nodes, or* $Z = C_1 \cup C_2 \cup C_3$*, with* C_i *smooth irreducible and* $|C_i \cap C_j| = 2$ *for* $1 \leq i < j \leq 3$.



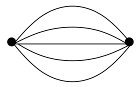


Proof. By hypothesis Γ_Z is eulerian, so property (5) says that $L(Pr_Z) = L(S_Z)$. Then (i) follow immediately from [CC03], Theorem 11. Let us show (ii). If $b_1(\Gamma_Z) \geq 4$, by property (2) we can apply [CC03], Theorem 13; then Z has three smooth components meeting each other in two points. Assume $b_1(\Gamma_Z) \leq 3$ and let δ, γ be the number of nodes and of irreducible components of Z. Since all vertices of Γ_Z have valency at least 4, we have $\delta \geq 2\gamma$, so $\gamma \leq b_1(\Gamma_Z) - 1 \leq 2$. Then by an easy check we see that the only possibility which satisfies all the hypotheses is $\gamma = 1$ and $\delta = 2$. ∎

In [CC03] it is shown (Theorem 11) that $L(S_Z)$ allows to recover curves having two smooth components. Instead, when the number of nodes is odd, it is no more true that these curves are characterized by $L(Pr_Z)$. For instance, consider the graphs:



Γ_1



Γ_2

It is easy to see that if Z_1, Z_2 are stable curves with $\Gamma_{Z_i} = \Gamma_i$ for $i = 1, 2$, we have $L(P_{Z_1}) = L(P_{Z_2}) = \{1, 4, 16\}$, while $L(S_{Z_1}) = \{4, 8, 16\}$ and $L(S_{Z_2}) = \{2, 8, 16\}$.

References

[ABH01] Valery Alexeev, Christina Birkenhake, and Klaus Hulek. Degenerations of Prym varieties. *J. Reine Angew. Math.*, 553:73–116, 2001.

[ACGH85] Enrico Arbarello, Maurizio Cornalba, Phillip Griffiths, and Joseph Harris. *Geometry of Algebraic Curves, I*, volume 267 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1985.

[Bea77] Arnaud Beauville. Prym varieties and the Schottky problem. *Invent. Math.*, 41:149–196, 1977.

[Bea96] Arnaud Beauville. *Complex Algebraic Surfaces.* Cambridge University Press, second edition, 1996.

[Ber99] Mira Bernstein. *Moduli of curves with level structure.* PhD thesis, Harvard University, 1999. Available at the web page `http://www.math.berkeley.edu/~mira`.

[Cap94] Lucia Caporaso. A compactification of the universal Picard variety over the moduli space of stable curves. *J. Amer. Math. Soc.*, 7(3):589–660, 1994.

[CC03] Lucia Caporaso and Cinzia Casagrande. Combinatorial properties of stable spin curves. *Comm. Algebra*, 31(8):3653–3672, 2003. Special Issue in Honor of Steven L. Kleiman.

[Cor89] Maurizio Cornalba. Moduli of curves and theta-characteristics. In *Lectures on Riemann Surfaces: Proceedings of the College on Riemann Surfaces, International Centre for Theoretical Physics, Trieste, 1987*, pages 560–589. World Scientific, 1989.

[Cor91] Maurizio Cornalba. A remark on the Picard group of spin moduli space. *Atti dell'Accademia Nazionale dei Lincei, Matematica e Applicazioni*, IX(2):211–217, 1991.

[CS03] Lucia Caporaso and Edoardo Sernesi. Characterizing curves by their odd theta-characteristics. *J. Reine Angew. Math.*, 562:101–135, 2003.

[DM69] Pierre Deligne and David Mumford. The irreducibility of the space of curves of given genus. *Publ. Math. I. H. E. S.*, 36:75–109, 1969.

[DS81] Ron Donagi and Roy Campbell Smith. The structure of the Prym map. *Acta Math.*, 146:25–102, 1981.

[Fon02] Claudio Fontanari. On the geometry of the compactification of the universal Picard variety. Preprint math.AG/0202168, 2002.

[Ray71] Michel Raynaud. Exposé XII. Géométrie algébrique et géométrie analytique. In *Revêtements Étales et Groupe Fondamental. Séminaire de Géométrie Algébrique du Bois Marie 1960–1961 (SGA1). Dirigé par Alexandre Grothendieck*, volume 224 of *Lecture Notes in Mathematics*, pages 311–343. Springer-Verlag, 1971.

Edoardo Ballico
Università degli Studi di Trento
Dipartimento di Matematica
Via Sommarive 14
38050 Povo (Trento), Italy
ballico@science.unitn.it

Claudio Fontanari
Università degli Studi di Trento
Dipartimento di Matematica
Via Sommarive 14
38050 Povo (Trento), Italy
fontanar@science.unitn.it

Cinzia Casagrande
Università degli Studi Roma Tre
Dipartimento di Matematica
Largo S. L. Murialdo 1
00146 Roma, Italy
casagran@mat.uniroma3.it

Nodal Domain Theorems à la Courant

A. Ancona, B. Helffer, and T. Hoffmann-Ostenhof*

Received: April 23, 2004
Revised: July 4, 2004

Communicated by Heinz Siedentop

Abstract. Let $H(\Omega_0) = -\Delta + V$ be a Schrödinger operator on a bounded domain $\Omega_0 \subset \mathbb{R}^d$ ($d \geq 2$) with Dirichlet boundary condition. Suppose that Ω_ℓ ($\ell \in \{1, \dots, k\}$) are some pairwise disjoint subsets of Ω_0 and that $H(\Omega_\ell)$ are the corresponding Schrödinger operators again with Dirichlet boundary condition. We investigate the relations between the spectrum of $H(\Omega_0)$ and the spectra of the $H(\Omega_\ell)$. In particular, we derive some inequalities for the associated spectral counting functions which can be interpreted as generalizations of Courant's nodal theorem. For the case where equality is achieved we prove converse results. In particular, we use potential theoretic methods to relate the Ω_ℓ to the nodal domains of some eigenfunction of $H(\Omega_0)$.

2000 Mathematics Subject Classification: 35B05

1 Introduction

Consider a Schrödinger operator

$$H = -\Delta + V \tag{1.1}$$

on a bounded domain $\Omega_0 \subset \mathbb{R}^d$ with Dirichlet boundary condition. Further we assume that V is real valued and satisfies $V \in L^\infty(\Omega_0)$. (We could relax this condition and extend our results to the case $V \in L^\beta(\Omega_0)$ for some $\beta > d/2$ using [11].)
The operator H is selfadjoint if viewed as the Friedrichs extension of the quadratic form of H with form domain $W_0^{1,2}(\Omega_0)$ and form core $C_0^\infty(\Omega_0)$ and

*Supported by Ministerium für Bildung, Wissenschaft und Kunst der Republik Österreich

we denote it by $H(\Omega_0)$. Further $H(\Omega_0)$ has compact resolvent. So the spectrum of $H(\Omega_0)$, $\sigma\big(H(\Omega_0)\big)$, can be described by an increasing sequence of eigenvalues

$$\lambda_1 < \lambda_2 \le \lambda_3 \le \cdots \le \lambda_j \le \lambda_{j+1} \le \dots \tag{1.2}$$

tending to $+\infty$, such that the associated eigenfunctions u_j form an orthonormal basis of $L^2(\Omega_0)$. We can assume that these eigenfunctions u_j are real valued and by elliptic regularity, [9] (Corollary 8.36), u_j belongs to $C^{1,\alpha}(\Omega_0)$ for every $\alpha < 1$. Moreover λ_1 is simple and the corresponding eigenfunction u_1 can be chosen to satisfy, see e.g. [17],

$$u_1(x) > 0\ ,\ \text{ for all } x \in \Omega_0\ . \tag{1.3}$$

For a bounded domain D we let $H(D)$ be the corresponding selfadjoint operator, with Dirichlet boundary condition on ∂D. Its lowest eigenvalue will be denoted by $\lambda(D)$.
We denote the zero set of an eigenfunction u by

$$N(u) = \overline{\{x \in \Omega_0 \mid u(x) = 0\}}. \tag{1.4}$$

The nodal domains of u, which are by definition the connected components of $\Omega_0 \setminus N(u)$, will be denoted by $D_j, j = 1, \dots, \mu(u)$, where $\mu(u)$ denotes the number of nodal domains of u.
Suppose that Ω_ℓ ($\ell = 1, 2, \dots, k$) are k open pairwise disjoint subsets of Ω_0. In this paper we shall investigate relations between the spectrum of $H(\Omega_0)$ and the spectra of the $H(\Omega_\ell)$. Roughly speaking, we shall derive an inequality between the counting function of $H(\Omega_0)$ and those of the $H(\Omega_\ell)$. This inequality can be interpreted as a generalization of Courant's classical nodal domain theorem. For the case where equality is achieved this will lead to a partial characterization of the Ω_ℓ which will turn out to be related to the nodal domains of one of the eigenfunctions of $H(\Omega_0)$.
These results will be given in sections 2 and 3. From these results some natural questions of potential theoretic nature arise which will be analyzed and answered in section 7.
The proofs of the results stated in sections 2 and 3 are given in sections 4 and 5. In section 6 some illustrative explicit examples are given.

2 Main results

We start with a result which will turn out to be a generalization of Courant's nodal theorem. We consider again (1.1) on a bounded domain Ω_0 and the corresponding eigenfunctions and eigenvalues. We first introduce

$$\overline{n}(\lambda, \Omega_0) = \#\{j \mid \lambda_j(\Omega_0) \le \lambda\}, \tag{2.1}$$

where $\lambda_j(\Omega_0)$ is the j-th eigenvalue of $H(\Omega_0)$.
We also define

$$\underline{n}(\lambda, \Omega_0) = \#\{j \mid \lambda_j(\Omega_0) < \lambda\}, \tag{2.2}$$

and

$$n(\lambda,\Omega_0)=\begin{cases}\underline{n}(\lambda,\Omega_0) & \text{if } \lambda\notin\sigma\big(H(\Omega_0)\big)\\ \underline{n}(\lambda,\Omega_0)+1 & \text{if } \lambda\in\sigma\big(H(\Omega_0)\big).\end{cases} \tag{2.3}$$

So we always have :

$$\underline{n}(\lambda,\Omega_0)\le n(\lambda,\Omega_0)\le \overline{n}(\lambda,\Omega_0), \tag{2.4}$$

with equality when λ is not an eigenvalue. Note that $\overline{n}(\lambda,\Omega_0)-\underline{n}(\lambda,\Omega_0)$ is the multiplicity of λ when λ is an eigenvalue of $H(\Omega_0)$, i.e. the dimension of the eigenspace associated to λ. We shall consider a family of k open sets Ω_ℓ ($\ell=1,\dots,k$) contained in Ω_0 and the corresponding Dirichlet realizations $H(\Omega_\ell)$. For each $H(\Omega_\ell)$ the corresponding eigenvalues counted with multiplicity are denoted by $(\lambda_j^\ell)_{j\in\mathbb{N}\setminus\{0\}}$ (with $\lambda_j^\ell\le\lambda_{j+1}^\ell$). When counting the eigenvalues less than some given λ , we shall for simplicity write

$$n_\ell=n_\ell(\lambda)=n(\lambda,\Omega_\ell), \tag{2.5}$$

and analogously for the quantities with over-, respectively, underbars.

Theorem 2.1
Suppose $\Omega_0\subset\mathbb{R}^d$ is a bounded domain and that $\lambda\in\sigma\big(H(\Omega_0)\big)$. Suppose that the sets Ω_ℓ ($\ell=1,\dots,k$) are pairwise disjoint open subsets of Ω_0. Then

$$\sum_{\ell=1}^k \overline{n}_\ell\le n_0+\min_{\ell\ge 0}\left(\overline{n}_\ell-n_\ell\right) . \tag{2.6}$$

A direct weaker consequence of (2.6) is the more standard

Corollary 2.2
Under the assumptions of Theorem 2.1, we have

$$\sum_{\ell=1}^k \overline{n}_\ell\le \overline{n}_0 . \tag{2.7}$$

This corollary is actually present in the proofs of the asymptotics of the counting function (see for example the Dirichlet-Neumann bracketing in Lieb-Simon [14]).

Remark 2.3
Inequality (2.6) is also true if $\lambda\notin\sigma\big(H(\Omega_0)\big)$. The statement becomes

$$\sum_{\ell=1}^k \overline{n}_\ell\le n_0 ,$$

and is proved essentially in the same way.

Remark 2.4
The assumption that Ω_0 is connected is necessary. Indeed, suppose Ω_1 and Ω_2 are connected and assume that $\Omega_0 = \Omega_1 \cup \Omega_2$ with $\Omega_1 \cap \Omega_2 = \emptyset$ and that $\lambda = \lambda_1(\Omega_1) = \lambda_1(\Omega_2)$. Then $\lambda_1(\Omega_0) = \lambda_2(\Omega_0)$ and we deduce $n(\lambda, \Omega_0) = 1$. If we no longer assume the connectedness of Ω_0, we in general just have Corollary 2.2.

Finally we show that Courant's nodal theorem is an easy corollary of Theorem 2.1.

Corollary 2.5 : Courant's nodal theorem
If Ω_0 is connected and if u is an eigenfunction of $H(\Omega_0)$ associated to some eigenvalue λ, then

$$\mu(u) \leq n(\lambda, \Omega_0) .$$

Proof.
We now simply apply Theorem 2.1 by taking $\Omega_1, \dots, \Omega_{\mu(u)}$ as the nodal domains associated to u. We just have to use (1.3) for each Ω_ℓ, $\ell = 1, \dots, \mu(u)$, which gives $\overline{n}_\ell = n_\ell = 1$. □

Remark 2.6
Courant's nodal theorem is one of the basic results in spectral theory of Schrödinger-type operators. It is the natural generalization of Sturm's oscillation theorem for second order ODE's. For recent investigations see for instance [1] and [4].

3 Converse results.

In this section we consider some results that are converse to Theorem 2.1.

Theorem 3.1
Suppose that the Ω_ℓ, $1 \leq \ell \leq k$, are pairwise disjoint open subsets of Ω_0. If $\lambda \in \sigma\big(H(\Omega_0)\big)$ and

$$\sum_{\ell=1}^{k} \overline{n}_\ell \geq n_0 , \tag{3.1}$$

then $\lambda \in \sigma(H(\Omega_\ell))$ for each Ω_ℓ. If $U_\ell(\lambda)$ denotes the eigenspace of $H(\Omega_\ell)$ associated to the eigenvalue λ, then there is an eigenfunction u of $H(\Omega_0)$ with eigenvalue λ such that

$$u = \sum_{\ell=1}^{k} \varphi_\ell \text{ in } W_0^{1,2}(\Omega_0) , \tag{3.2}$$

where each φ_ℓ belongs to $U_\ell(\lambda) \setminus \{0\}$ and is identified with its extension by 0 outside Ω_ℓ.

Remark 3.2
One can naturally think that formula (3.2) *has immediate consequences on the family* Ω_ℓ*, which should for example have some covering property. The question is a bit more subtle because we do not a priori want to assume strong regularity properties for the boundaries of the* Ω_ℓ*. We shall discuss this point in detail in the last section.*

Another consequence of equalities in Theorems 2.1 or 3.1 is given by the following result.

Theorem 3.3
Suppose that, for some bounded domain Ω_0 *in* $\mathbb{R}^d$*, some* $\lambda \in \sigma(H(\Omega_0))$ *and some family of pairwise disjoint open sets* $\Omega_\ell \subset \Omega_0$*,* $0 < \ell \leq k$*, we have*

$$\sum_{\ell=1}^{k} \overline{n}_\ell = n_0 + \min_{\ell \geq 0} \left(\overline{n}_\ell - n_\ell\right) . \tag{3.3}$$

Then, for any subset $L \subset \{1, 2, \ldots, k\}$ *such that* $\Omega_L^* = \mathrm{Int}\left(\cup_{\ell \in L} \overline{\Omega_\ell}\right) \setminus \partial\Omega_0$ *is connected, we have*

$$\sum_{\ell \in L} \overline{n}_\ell = n(\lambda, \Omega_L^*) + \min\left(\min_{\ell \in L}\left(\overline{n}_\ell - n_\ell\right),\, \overline{n}(\lambda, \Omega_L^*) - n(\lambda, \Omega_L^*)\right) . \tag{3.4}$$

A simpler variant is the following :

Theorem 3.4
Suppose (3.1) *holds and that* Ω_L^* *is defined as above. Then we have the inequality :*

$$\sum_{\ell \in L} \overline{n}_\ell \geq n(\lambda, \Omega_L^*) . \tag{3.5}$$

On the sharpness of Courant's nodal theorem
It is well known that Courant's nodal theorem is sharp only for finitely many k's [15].
Let Ω_0 be connected. We will say that an eigenfunction u associated to an eigenvalue λ of $H(\Omega_0)$ is Courant-sharp if $\mu(u) = n(\lambda, \Omega_0)$. Theorem 3.3 now implies :

Corollary 3.5
i) Let u *be a* Courant-sharp *eigenfunction of* $H(\Omega_0)$ *with* $\mu(u) = k$*. Let* $\{D_i\}_{i=1,\ldots,k}$ *be the family of the nodal domains associated to* u*, let* L *be a subset of* $\{1, \ldots, k\}$ *with* $\#L = \ell$ *and let* $\Omega_L^* = \mathrm{Int}\,(\overline{\cup_{i \in L} D_i}) \setminus \partial\Omega_0$*. Then*

$$\lambda_\ell(\Omega_L^*) = \lambda_k , \tag{3.6}$$

where $\lambda_j(\Omega_L^*)$ *are the eigenvalues of* $H(\Omega_L^*)$*.*
ii) Moreover, if Ω_L^* *is connected, and if* $\ell < k$*, then* $u\big|_{\Omega_L^*}$ *is* Courant-sharp *and* $\lambda_\ell(\Omega_L^*)$ *is simple.*

4 Basic tools

Let us first recall some basic tools (see e.g. [17]) which were already vital for the proof of Courant's classical result.

4.1 Variational characterization

Let us first recall the variational characterization of eigenvalues.

Proposition 4.1
Let Ω be a bounded open set in $\mathbb{R}^d$ and let $V \in L^\infty(\Omega)$ be real-valued. Suppose $\lambda \in \sigma\big(H(\Omega)\big)$ and let $\mathcal{U}_\pm = span\ \langle u_1, \dots, u_{k_\pm}\rangle$ where

$$k_- = \underline{n}(\lambda, \Omega)\ ,\ k_+ = \overline{n}(\lambda, \Omega)\ , \tag{4.1}$$

and $(u_j)_{j\geq 1}$ is as before an orthonormal basis of eigenfunctions of $H(\Omega)$ associated to $(\lambda_j)_{j\geq 1}$. Then

$$\lambda = \inf_{\varphi \perp \mathcal{U}_-,\, \varphi \in W_0^{1,2}(\Omega)} \frac{\langle \varphi,\ H(\Omega)\varphi\rangle}{\|\varphi\|^2} \tag{4.2}$$

and

$$\lambda < \lambda_{\overline{n}(\lambda,\, \Omega)+1} = \inf_{\varphi \perp \mathcal{U}_+,\, \varphi\, \in W_0^{1,2}(\Omega)} \frac{\langle \varphi,\ H(\Omega)\varphi\rangle}{\|\varphi\|^2}. \tag{4.3}$$

If equality is achieved in (4.2) for some $\varphi \not\equiv 0$, then φ is an eigenfunction in the eigenspace of λ.

Note that (4.2) and (4.3) are actually the same statement. We just stated them separately for later reference. Note that we have not assumed that Ω is connected.

4.2 Unique continuation

Next we restate a weak form of the unique continuation property:

Theorem 4.2
Let Ω be an open set in $\mathbb{R}^d$ and let $V \in L^\infty_{loc}(\Omega)$ be real-valued. Then any distributional solution in Ω to $(-\Delta + V)u = \lambda u$ which vanishes on an open subset ω of Ω is identically zero in the connected component of Ω containing ω.

There are stronger results of this type under weaker assumptions on the potential, see [11].

4.3 A consequence of Harnack's inequality

The standard Harnack's inequality (see e.g. Theorem 8.20 in [9]), together with the unique continuation theorem leads to the following theorem :

Theorem 4.3
If Ω is a bounded domain in $\mathbb{R}^d$ and u is an eigenfunction of $H(\Omega)$, then for any x in $N(u) \cap \Omega$ and any ball $B(x,r)$ ($r > 0$), there exist $y_\pm \in B(x,r) \cap \Omega$ such that $\pm u(y_\pm) > 0$.

5 Proof of the main theorems

5.1 Proof of Theorem 2.1

Assume first for contradiction that

$$\sum_{\ell \geq 1} \overline{n}_\ell > n_0 + \min_{\ell \geq 0} \left(\overline{n}_\ell - n_\ell \right) , \tag{5.1}$$

and recall that we assume that $\lambda \in \sigma\big(H(\Omega_0)\big)$. Pick some ℓ_0 such that

$$\overline{n}_{\ell_0} - n_{\ell_0} = \min_{\ell \geq 0} \left(\overline{n}_\ell - n_\ell \right) .$$

Suppose first that $\ell_0 \geq 1$.
We can rewrite (5.1) to obtain

$$\sum_{\ell \neq \ell_0,\, \ell \geq 1} \overline{n}_\ell + n_{\ell_0} > n_0 . \tag{5.2}$$

Let $\varphi_i^{\ell_0}$, $i = 1, \ldots, \underline{n}(\lambda, \Omega_{\ell_0})$, denote the first $\underline{n}_{\ell_0}$ eigenfunctions of $H(\Omega_{\ell_0})$. The corresponding eigenvalues are strictly smaller than λ. The functions $\varphi_i^{\ell_0}$ and the remaining $\sum_{\ell \neq \ell_0} \overline{n}_\ell$ eigenfunctions associated to the other $H(\Omega_\ell)$ span a space of dimension at least n_0. We can pick a linear combination $\Phi \not\equiv 0$ of these functions which is orthogonal to the $\underline{n}_0$ eigenfunctions of $H(\Omega_0)$. By assumption

$$\frac{\langle \Phi, H(\Omega_0)\Phi \rangle}{\|\Phi\|^2} \leq \lambda, \tag{5.3}$$

hence Φ must by the variational principle be an eigenfunction and there must be equality in (5.3).
There are two possibilities: either some $\varphi_i^{\ell_0}, i < n_{\ell_0}$ contributes to the linear combination which makes up Φ or not. In the first case this means that the left hand side of (5.3) is strictly smaller than λ, contradicting the variational characterization of λ. In the other case we obtain a contradiction to unique continuation, since then $\Phi \equiv 0$ in Ω_{ℓ_0} and hence Φ vanishes identically in all of Ω_0.

Consider now the case when $\ell_0 = 0$.
We have to show that the assumption

$$\sum_{\ell \geq 1} \overline{n}_\ell > \overline{n}_0 \ , \tag{5.4}$$

leads to a contradiction. To this end it suffices to apply (4.3). Indeed, we can find a linear combination Φ of the eigenfunctions φ_j^ℓ, $j \leq \overline{n}_\ell$, corresponding to the different $H(\Omega_\ell)$ such that $\Phi \perp \mathcal{U}_+$, $\Phi \not\equiv 0$, but Φ satisfies

$$\frac{\langle \Phi,\ H(\Omega_0)\ \Phi \rangle}{\|\Phi\|^2} \leq \lambda = \lambda_{\overline{n}_0} \ ,$$

and this contradicts (4.3). This proves (2.6).

5.2 Proof of Theorem 3.1

The inequality (3.1) implies that we can find a non zero $u \perp \mathcal{U}_-$ in the span of the eigenfunctions φ_j^ℓ, $j = 1, \ldots \overline{n}_\ell$, of the different $H(\Omega_\ell)$. Again by the variational characterization, (4.2) and (5.3) hold and hence u must be an eigenfunction. □

5.3 Proof of Theorem 3.3

We assume (3.3). Without loss we might assume that we have labeled the Ω_ℓ such that $L = \{1, \ldots, K\}$, with $K \leq k$. Let $n_* = n(\lambda, \Omega_L^*)$. We apply Theorem 2.1 to the family Ω_ℓ ($\ell \in L$) and replace Ω_0 by Ω_L^* and obtain :

$$\sum_{1 \leq \ell \leq K} \overline{n}_\ell \leq n_* + \min \left(\overline{n}_* - n_*,\ \min_{1 \leq \ell \leq K} (\overline{n}_\ell - n_\ell) \right) . \tag{5.5}$$

We assume for contradiction that

$$\sum_{1 \leq \ell \leq K} \overline{n}_\ell < n_* + \min \left(\overline{n}_* - n_*,\ \min_{1 \leq \ell \leq K} (\overline{n}_\ell - n_\ell) \right) . \tag{5.6}$$

This implies

$$\sum_{1 \leq \ell \leq K} \overline{n}_\ell < \overline{n}_* \ , \tag{5.7}$$

and

$$\sum_{1 \leq \ell \leq K} \overline{n}_\ell < n_* + \min_{1 \leq \ell \leq K} (\overline{n}_\ell - n_\ell) \ . \tag{5.8}$$

Theorem 2.1, applied to the family Ω_L^*, Ω_ℓ ($\ell > K$), implies that

$$\overline{n}_* + \sum_{K < \ell \leq k} \overline{n}_\ell \leq n_0 + \min \left(\overline{n}_0 - n_0,\ \min_{K < \ell \leq k} (\overline{n}_\ell - n_\ell) \right) , \tag{5.9}$$

and

$$n_* + \sum_{K<\ell\leq k} \overline{n}_\ell \leq n_0 \; . \tag{5.10}$$

By adding (5.7) and (5.9), we get :

$$\sum_{1\leq\ell\leq k} \overline{n}_\ell < n_0 + \min\big(\overline{n}_0 - n_0,\ \min_{K<\ell\leq k}(\overline{n}_\ell - n_\ell)\big) \; . \tag{5.11}$$

By adding (5.8) and (5.10), we obtain

$$\sum_{1\leq\ell\leq k} \overline{n}_\ell < n_0 + \min_{1\leq\ell\leq K}(\overline{n}_\ell - n_\ell) \; . \tag{5.12}$$

The combination of (5.11) and (5.12) is in contradiction with (3.3).

5.4 Proof of Theorem 3.4

For the case that (3.1) holds, (3.5) can be shown similarly. (3.1) reads

$$\sum_{1\leq\ell\leq k} \overline{n}_\ell \geq n_0 \; .$$

We assume for contradiction that

$$\sum_{1\leq\ell\leq K} \overline{n}_\ell < n_* \; , \tag{5.13}$$

where n_* is defined as above. The addition of (5.10) and (5.13) leads to a contradiction. □

6 Illustrative examples

6.1 Examples for a rectangle

We illustrate Theorem 2.1 by the analysis of various examples in rectangles. Pick a rectangle $\Omega_0 = (0, 2\pi) \times (0, \pi)$ and take $\Omega_1 = (0, \pi) \times (0, \pi)$ and consequently $\Omega_2 = (\pi, 2\pi) \times (0, \pi)$. The eigenvalues corresponding to Ω_0 for $-\Delta$ with Dirichlet boundary condition are given by

$$\sigma\big(H(\Omega_0)\big) = \left\{ \lambda \in \mathbb{R} \,\middle|\, \lambda = m^2/4 + n^2,\ (m,n) \in \mathbb{Z}^2,\ m,n > 0 \right\} , \tag{6.1}$$

while those for Ω_1, and hence for Ω_2 which can be obtained by a translation of Ω_1, are given by

$$\sigma\big(H(\Omega_1)\big) = \sigma\big(H(\Omega_2)\big) = \left\{ \lambda \in \mathbb{R} \,\middle|\, \lambda = m^2 + n^2,\ (m,n) \in \mathbb{Z}^2, m,n > 0 \right\}. \tag{6.2}$$

Denote the eigenvalues associated to Ω_0 by $\{\lambda_i\}$ and those to Ω_1 by $\{\nu_i\}$. We easily check that $\lambda_5 = \lambda_6 = \nu_2 = \nu_3 = 5$, $\lambda_{11} = \lambda_{12} = \nu_5 = \nu_6 = 10$ so that Theorem 2.1 is sharp for these cases.
One could ask whether there are arbitrarily high eigenvalues cases for which we have equality in (2.6). This is not the case, as can be seen from the following standard number theoretical considerations. We have (see [18] and for more recent contributions [16] and [2]) the following asymptotic estimate for the number of lattice points in an ellipse. Let $a, b > 0$, then

$$A(\lambda) := \#\left\{(m,n) \in \mathbb{Z}^2 \;\middle|\; am^2 + bn^2 \leq \lambda\right\} \tag{6.3}$$

has the following asymptotics as λ tends to infinity:

$$A(\lambda) = \frac{\pi}{\sqrt{ab}}\lambda + \mathcal{O}(\lambda^{1/3}). \tag{6.4}$$

We have not to consider $A(\lambda)$ but rather

$$A^+ = \#\left\{(m,n) \in \mathbb{Z}^2, m,n > 0 \;\middle|\; am^2 + bn^2 \leq \lambda\right\}. \tag{6.5}$$

Hence we get

$$\begin{aligned} A(\lambda) = 4A^+(\lambda) + 2\#&\left\{m \in \mathbb{N},\ m > 0 \;\middle|\; m \leq \left[(\lambda/a)^{1/2}\right]\right\} \\ &+2\#\left\{n \in \mathbb{N},\ n > 0 \;\middle|\; n \leq \left[(\lambda/b)^{1/2}\right]\right\} + 1\,. \end{aligned} \tag{6.6}$$

If we apply this to A^+ with $a = 1/4, b = 1$ (in this case denoted by A_0^+) and to A^+ with $a = 1, b = 1$ (in this case denoted by A_1^+), we get asymptotically

$$A_0^+(\lambda) - 2A_1^+(\lambda) = \frac{1}{2}\sqrt{\lambda} + o\,(\sqrt{\lambda})\,. \tag{6.7}$$

Note that

$$\overline{n}_i(\lambda) = A_i^+(\lambda),\ i = 0, 1\,.$$

In order to control $n_i(\lambda)$, we observe that, for any $\epsilon > 0$:

$$\overline{n}_i(\lambda - \epsilon) \leq n_i(\lambda) \leq \overline{n}_i(\lambda)\,.$$

This implies

$$\overline{n}_i(\lambda) - n_i(\lambda) = \mathcal{O}(\lambda^{\frac{1}{3}})\,. \tag{6.8}$$

The asymptotic formula (6.4) implies

$$\overline{n}_i(\lambda) - n_i(\lambda) = o(\sqrt{\lambda})\,, \tag{6.9}$$

and this shows that (2.6) is never sharp for large λ.

6.2 About Corollary 3.5

One can ask whether there is a converse to Corollary 3.5 in the following sense. Suppose we have an eigenfunction u with k nodal domains and eigenvalue λ. For each pair of neighboring nodal domains of u, say, D_i and D_j, let $\Omega_{i,j} = \mathrm{Int}\,(\overline{D_i \cup D_j})$ and suppose that $\lambda = \lambda_2(\Omega_{i,j})$. Does this imply that $\lambda = \lambda_k$? The answer to the question is negative, as the following easy example shows :

Consider the rectangle $Q = (0,a) \times (0,1) \subset \mathbb{R}^2$ and consider $H_0(Q)$. We can work out the eigenvalues explicitly as

$$\{\pi^2(\frac{m^2}{a^2} + n^2)\}, \ \text{ for } m, n \in \mathbb{N} \setminus 0, \tag{6.10}$$

with corresponding eigenfunctions $(x,y) \mapsto \sin(\pi m \frac{x}{a})(\sin \pi n y)$. If

$$a^2 \in (\frac{9}{4}, \frac{8}{3}) \ , \tag{6.11}$$

then

$$\lambda_3(Q) = \pi^2(\frac{1}{a^2} + 4) < \lambda_4(Q) = \pi^2(\frac{9}{a^2} + 1) \ ,$$

and the zeroset of u_4 is given by $\{(x,y) \in Q \mid x = a/3, \ x = 2a/3\}$. For u_4 we have $\Omega_{1,2} = Q \cap \{0 < x < 2a/3\}$. If $2a/3 > 1$ (which is the case under assumption (6.11)), then $\lambda_2(\Omega_{1,2}) = \lambda_4(Q)$. We have consequently an example with $k = 3$.

7 Converse theorems in the case of regular open sets

7.1 Preliminaries

As a consequence of Theorem 3.1 and using (1.3), we get that each nodal domain $D_{k\ell}$ of φ_ℓ is included in a nodal domain D_{j0} of u. Using a result of Gesztesy and Zhao ([8], Theorem 1), this implies also that the capacity (see next subsection) of $D_{j0} \setminus D_{k\ell}$ (hence the Lebesgue-measure) is 0.

We now would like to show that under some extra condition the nodal domains of u are those of the φ_ℓ. This is easy when it is assumed that the boundaries of the Ω_ℓ are $C^{1,\alpha}$. However, this regularity assumption is rather strong. A natural weaker regularity condition involving the notion of capacity will be given in this section.

7.2 Capacity

There are various equivalent definitions of polar sets and capacity (see e.g. [5], [7], [10], [13]). If U is a bounded open subset of $\mathbb{R}^d$, we denote by $\|.\|_{W_0^{1,2}(U)}$ the Hilbert norm on $W_0^{1,2}(U)$:

$$\|u\|_{W_0^{1,2}(U)} := (\int_U |\nabla u|^2 \, dx)^{\frac{1}{2}} \ .$$

The capacity in U of $A \subset U$ is defined[†] as

$$\begin{aligned}\mathrm{Cap}_U(A) \quad &:= \inf\{\|s\|^2_{W_0^{1,2}(U)}\,;\, s \in W_0^{1,2}(U) \\ &\qquad\qquad \text{and } s \geq 1 \text{ a.e. in some neighborhood of } A\}\,.\end{aligned}$$

It is easily checked that if K is compact and $K \subset U \cap V$, where V is also open and bounded in $\mathbb{R}^d$, then there is a $c = c(K, U, V)$ such that $\mathrm{Cap}_U(A) \leq c\ \mathrm{Cap}_V(A)$ for $A \subset K$. So $\mathrm{Cap}_U(A) = 0$ for some bounded open $U \supset A$ iff for each $a \in A$ there exists an $r > 0$ and a bounded domain V such that $V \supset B(a, r)$ and $\mathrm{Cap}_V(B(a, r) \cap A) = 0$. In this case we may simply write $\mathrm{Cap}(A) = 0$ without referring to U.

7.3 Converse theorem

We are now able to formulate our definition of a regular point.

Definition 7.1
Let D be an open set in $\mathbb{R}^d$. We shall say that a point $x \in \partial D$ is (capacity)-regular (for D) if, for any $r > 0$, the capacity of $B(x, r) \cap \complement D$ is strictly positive.

Theorem 7.2
Under the assumptions of Theorem 3.1, any point $x \in \partial\Omega_\ell \cap \Omega_0$ which is (capacity)-regular with respect to Ω_ℓ (for some ℓ) is in the nodal set of u.

This theorem admits the following corollary :

Corollary 7.3
Under the assumptions of Theorem 3.1 and if, for all ℓ, every point in $(\partial\Omega_\ell) \cap \Omega_0$ is (capacity)-regular for Ω_ℓ, then the family of the nodal domains of u coincides with the union over ℓ of the family of the nodal domains of the φ_ℓ, where u and φ_ℓ are introduced in (3.2).

Proof of corollary
It is clear that any nodal domain of φ_ℓ is contained in a unique nodal domain of u.
Conversely, let D be a nodal domain of u and let $\ell \in \{1, \ldots, k\}$. Then, by combining the assumption on $\partial\Omega_\ell$, Proposition 7.4 and (3.2), we obtain the property :

$$\partial\Omega_\ell \cap D = \emptyset\,.$$

Now, D being connected, either $\Omega_\ell \cap D = \emptyset$ or $D \subset \Omega_\ell$. Moreover the second case should occur for at least one ℓ, say $\ell = \ell_0$. Coming back to the definition of a nodal set and (3.2), we observe that D is necessarily contained in a nodal domain $D_j^{\ell_0}$ of φ_{ℓ_0}.
Combining the two parts of the proof gives that any nodal set of u is a nodal set of φ_ℓ and vice-versa.

[†]For $d \geq 3$ the restriction that U is bounded can be removed and one may take $U = \mathbb{R}^d$.

7.4 Proof of Theorem 7.2

The proof is a consequence of (3.2) and of the following proposition :

Proposition 7.4
Let $D, \Omega \subset \mathbb{R}^d$ be open sets such that $D \subset \Omega$, and let $x_0 \in \partial D \cap \Omega$. Assume that, for some given $r_0 > 0$ such that $B(x_0, r_0) \subset \Omega$, there exists $u \in W_0^{1,2}(D)$ and $v \in C^0(B(x_0, r_0))$ such that :

$$u_{|D\cap B(x_0,r_0)} = v_{|D\cap B(x_0,r_0)} \ \text{ a.e. in } D \cap B(x_0, r_0)\ .$$

Then if $v(x_0) \neq 0$, there exists a ball $B(x_0, r_1)$ ($r_1 > 0$), such that $B(x_0, r_1)\setminus D$ is polar, that is, of capacity 0.

Remark 7.5
Using some standard potential theoretic arguments, Proposition 7.4 can be deduced from Théorème 5.1 in [6] which characterizes, in the case where $d \geq 3$, those $u \in W^{1,2}(\Omega)$ that belong to $W_0^{1,2}(\Omega)$. The proof below should be more elementary in character.

Remark 7.6
Given an open subset $D \subset \mathbb{R}^d$ and a ball $B = B(x, r)$, $x \in \partial D$, the difference set $B \setminus D$ is polar if and only if $B \cap \partial D$ is polar. This follows from the fact that a polar subset of B does not disconnect B [3].

Remark 7.7
If D is a nodal domain of an eigenfunction u of $H(\Omega)$, then any point of $\partial D \cap \Omega$ is capacity-regular for D. This is an immediate consequence of Theorem 4.3 (it also follows from the preceding remark). Indeed, if x is in $\partial D \cap \Omega$, then for any $r > 0$, one can find a ball $B(y, r')$ in $\complement D \cap B(x, r)$.

To prove Proposition 7.4 we require some well-known facts stated in the next three lemmas.

Lemma 7.8
Let U be a bounded convex domain in $\mathbb{R}^d$ and let $B(a, \rho)$, $\rho > 0$ be a ball such that $\overline{B}(a, \rho) \subset U$. There exists a positive constant $c = c(a, \rho, U)$ such that, for every $f \in W^{1,2}(U)$ vanishing a.e. in $B(a, \rho)$,

$$\|f\|_{W^{1,2}(U)} \leq c\, \|\nabla f\|_{L^2(U)}\ .$$

Proof of Lemma 7.8

We can assume without loss of generality that $a = 0$ and let $U' = U \setminus B(0, \rho)$. Fix R so large that $U \subset B(0, R)$. By approximating f by smooth functions (e.g. regularize the function $x \mapsto f((1-\delta)x)$ for $\delta > 0$ and small to get $f_1 \in C^\infty(\overline{U})$), we may restrict to functions $f \in C^\infty(U)$ vanishing in $B(0, \rho)$. Then, since

$$|f(x)|^2 = |\int_0^1 x \cdot \nabla f(sx)\, ds|^2 \leq R^2 \int_{\frac{\rho}{|x|}}^1 |\nabla f(sx)|^2\, ds \text{ for } x \in U' ,$$

we have

$$\begin{aligned} \int_{U'} |f(x)|^2\, dx &\leq R^2 \iint_{x \in U',\, \frac{\rho}{|x|} \leq s \leq 1} |\nabla f(sx)|^2\, dx\, ds \\ &\leq R^2 \iint_{z \in sU',\, \rho \leq |z|,\, s \leq 1} |\nabla f(z)|^2\, dz\, \frac{ds}{s} \qquad (7.1) \\ &\leq \frac{R^3}{\rho} \int_{U'} |\nabla f(x)|^2\, dx, \end{aligned}$$

and the lemma follows.

Lemma 7.9

Let U be a domain in $\mathbb{R}^d$. For every real-valued $f \in W^{1,2}(U)$ the function $g = f_+$ is also in $W^{1,2}(U)$, with $\|g\|_{W^{1,2}(U)} \leq \|f\|_{W^{1,2}(U)}$. Moreover the map $f \mapsto g$ from $W^{1,2}(U)$ into itself is continuous (in the norm topology).

Remark 7.10

Since $\inf\{f_n, 1\} = 1 - (1 - f_n)_+$, *it follows from the lemma that* $\inf\{f_n, 1\} \to \inf\{f, 1\}$ *in* $W^{1,2}(U)$ *whenever* $f_n \to f$ *in* $W^{1,2}(U)$.

Proof of Lemma 7.9

For the first two facts we refer to [12] or [13], where it is moreover shown that the weak partial derivatives $\partial_j f_+$ and $\partial_j f$ satisfy

$$\partial_j f_+ = 1_{\{f>0\}}\, \partial_j f = 1_{\{f \geq 0\}}\, \partial_j f \quad \text{a.e. in} \quad U.$$

Therefore, for any $\delta > 0$, we have :

$$\begin{aligned} &\|\nabla [f_n]_+ - \nabla f_+\|_{L^2} \\ &\qquad = \|1_{\{f_n>0\}} \nabla f_n - 1_{\{f>0\}} \nabla f\|_{L^2} \\ &\qquad \leq \|1_{\{f_n>0\}} (\nabla f_n - \nabla f)\|_{L^2} + \|(1_{\{f>0\}} - 1_{\{f_n>0\}}) \nabla f\|_{L^2} \qquad (7.2) \\ &\qquad \leq \|\nabla f_n - \nabla f\|_{L^2} + \|(1_{\{f>0; f_n \leq 0\}} + 1_{\{f \leq 0; f_n > 0\}}) \nabla f\|_{L^2} \\ &\qquad \leq \|\nabla f_n - \nabla f\|_{L^2} + \|1_{\{0 \leq |f| \leq \delta\}} \nabla f\|_{L^2} + 2\|1_{\{|f_n - f| \geq \delta\}} \nabla f\|_{L^2}. \end{aligned}$$

Given $\varepsilon > 0$, fix $\delta > 0$ so that $\|1_{\{0 \leq |f| \leq \delta\}} \nabla f\|_{L^2} \leq \varepsilon$ (recall that $\nabla f = 0$ a.e. in $\{f = 0\}$). Since $\nabla f \in L^2(U)$ and $\|1_{\{|f - f_n| \geq \delta\}}\|_{L^1} \leq \frac{\|f_n - f\|_{L^2}^2}{\delta^2}$, it follows that $\lim_{n\to\infty} \|(1_{\{|f-f_n|\geq\delta\}})\nabla f\|_{L^2} = 0$. Therefore $\limsup_{n\to\infty} \|\nabla[f_n]_+ - \nabla f_+\|_{L^2} \leq \varepsilon$, which proves that $[f_n]_+ \to f_+$ in $W^{1,2}(U)$, if $f_n \to f$ in $W^{1,2}(U)$.

Lemma 7.11

Let ω be open in $\mathbb{R}^d$ and let $\{f_n\}$ be a sequence of functions continuous in ω such that $f_n \in W^{1,2}(\omega)$ for each $n \geq 1$ and $\lim_{n\to\infty} \|f_n\|_{W^{1,2}(\omega)} = 0$.
Then the set $F = \{x \in \omega\,;\, \liminf_{n\to\infty} |f_n(x)| > 0\}$ is polar.

Proof of Lemma 7.11

It suffices to show that $\mathrm{cap}_\omega(F \cap K) = 0$ for any compact subset K of ω. Let $\varphi \in C_0^\infty(\mathbb{R}^d)$ be such that $0 \leq \varphi \leq 1$ in $\mathbb{R}^d$, $\varphi = 1$ in K and $\mathrm{supp}(\varphi) \subset \omega$. Then $g_n = f_n\varphi \to 0$ in $W_0^{1,2}(\omega)$ and $g_n = f_n$ in K.
Set $F_\nu = \{x \in \omega\,;\, |g_n(x)| \geq 2^{-\nu}$ for all $n \geq \nu\}$. By the definition of the capacity, we have $\mathrm{Cap}_\omega(F_\nu) \leq 2^{2\nu}\|\nabla g_n\|_{L^2}^2$ for all $n \geq \nu$ and $\mathrm{cap}(F_\nu) = 0$. Therefore $\mathrm{cap}_\omega(\bigcup_{\nu\geq 1} F_\nu) = 0$ and $\mathrm{cap}_\omega(F \bigcap K) = 0$, since $F \bigcap K \subset \bigcup_{\nu\geq 1} F_\nu$.

Proposition 7.12

Let U be a non-empty open subset of the ball $B = B(a,r)$ in $\mathbb{R}^d$. Suppose there exist a function f continuous in U and a sequence $\{f_n\}$ of functions continuous in B such that
(i) $f \geq 1$ in U and $f \in W^{1,2}(U)$,
(ii) $f_n = 0$ in a neighborhood of $B \setminus U$ and $f_n \in W^{1,2}(U)$ for each $n \geq 1$,
(iii) $\lim_{n\to\infty} \|f - f_n\|_{W^{1,2}(U)} = 0$.
Then the set $F := B \setminus U$ is polar.

Proof of Proposition 7.12

Replacing f by $\inf\{f, 1\}$ and f_n by $\inf\{f_n, 1\}$, we see‡ from Lemma 7.9 that we may assume that $f = 1$ in U. So

$$\lim_{n\to\infty} \|\nabla f_n\|_{L^2(U)} = 0 \text{ and } \lim_{n\to\infty} \|1 - f_n\|_{L^2(U)} = 0\,.$$

Fix a ball $\overline{B}(z_0, 2\rho) \subset U$, $\rho > 0$, and a cut-off function $\alpha \in C^\infty(\mathbb{R}^d)$ such that $\alpha = 1$ in $B(z_0, \rho)$, $\alpha = 0$ in $\mathbb{R}^d \setminus B(z_0, 2\rho)$. Set $g = 1 - \alpha$, $g_n = (1-\alpha)f_n$. Then g, g_n belong to $W^{1,2}(B)$, $\nabla g = \nabla g_n = 0$ a.e. in F and

$$\lim_{n\to\infty} \|\nabla(g - g_n)\|_{L^2(B)} = \lim_{n\to\infty} \|\nabla(g - g_n)\|_{L^2(U)} = 0.$$

So, by Lemma 7.8, $\lim_{n\to\infty} \|g - g_n\|_{W^{1,2}(B)} = 0$. But $g - g_n \geq 1$ in F and it follows from Lemma 7.11 that F is polar.

‡The weak convergence $\inf\{f_n, 1\} \xrightarrow{w} \inf\{f, 1\}$ suffices here. It allows the approximation of $1 = \inf\{f, 1\}$ in the norm topology in $W^{1,2}(U)$ by finite convex combination of the $\inf\{f_n, 1\}$. So we are again left with the case when $f = 1$ in U.

Proof of Proposition 7.4
Without loss of generality, we can assume that $v(x_0) > 0$. Choose $r_1 > 0$ so small that $v \geq c_0 := \frac{1}{2}v(x_0)$ in $B(x_0, r_1)$. Since $u \in W_0^{1,2}(D)$, there is a sequence $\{u_n\}$ in $C_0^\infty(\mathbb{R}^d)$ such that $\mathrm{supp}(u_n) \subset D$ and $u_n \to u$ in $W^{1,2}(\mathbb{R}^d)$. Applying Proposition 7.12 to the ball $B(x_0, r_1)$ and the functions $f = c_0^{-1} u_{|B(x_0,r_1)}$, $f_n = c_0^{-1} u_{n|B(x_0,r_1)}$, we see that $B(x_0, r_1) \setminus D$ is polar.

Acknowledgements.
It is a pleasure to thank Fritz Gesztesy and Vladimir Maz'ya for useful discussions. We thank also the referee for his careful reading. The two last authors were partly supported by the European Science Foundation Programme *Spectral Theory and Partial Differential Equations* (SPECT) and the EU IHP network *Postdoctoral Training Program in Mathematical Analysis of Large Quantum Systems* HPRN-CT-2002-00277.

References

[1] G. Alessandrini. On Courant's nodal domain theorem. *Forum Math.*, 10, p. 521-532 (1998).

[2] P. Bérard. On the number of lattice points in some domains. *Comm. Partial Differential Equations*, Vol 3 (4) p. 335-348 (1978).

[3] M. Brelot. *Éléments de la théorie classique du Potentiel.* Les cours de Sorbonne. Centre de Documentation Universitaire, Paris 1965.

[4] E.B. Davies, G.M.L. Gladwell, J. Leydold, and P.F. Stadler. Discrete nodal domain theorems. *Linear Algebra Appl.* 336, p. 51-60 (2001).

[5] J. Deny. Méthodes hilbertiennes en théorie du potentiel. *Potential Theory (C.I.M.E., I Ciclo, Stresa, 1969)*, p. 121-201. Edizioni Cremonese, Roma 1970.

[6] J. Deny, J. L. Lions. Les espaces du type de Beppo Levi. *Ann. Inst. Fourier, Grenoble* 5, p. 305-370 (1955).

[7] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes.* de Gruyter Studies in Mathematics, de Gruyter & Co., Berlin 1994.

[8] F. Gesztesy, Z. Zhao. Domain perturbations, Brownian motion, capacities, and groundstate of Dirichlet Schrödinger operators. *Math. Z.* 215, p. 143-150 (1994).

[9] D. Gilbarg, N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order.* Springer 1983.

[10] L.L. Helms. *Introduction to potential theory.* Wiley-Interscience, New York-London-Sydney, 1969.

[11] D. Jerison, C.E. Kenig. Unique continuation and absence of positive eigenvalues for Schrödinger operators (with an appendix by E.M. Stein). *Annals of Math.* (2) 121, n^{o}3, p. 463-494 (1985).

[12] D. Kinderlehrer, G. Stampacchia. *An introduction to variational inequalities and their applications.* Pure and Applied Mathematics 88. Academic Press. New York-London, 1980.

[13] E.H. Lieb, M. Loss. *Analysis.* Graduate Studies in Mathematics, 14. American Mathematical Society, Providence, RI, 1997.

[14] E.H. Lieb, B. Simon. The Thomas-Fermi theory of atoms, molecules and solids, *Adv. in Math.* 23, p. 22-116 (1977).

[15] A. Pleijel. Remarks on Courant's nodal line theorem. *Comm. Pure. Appl. Math.* 9, p. 543-550 (1956).

[16] B. Randol. A lattice-point problem. *Trans. AMS* 121, p. 257-268 (1966).

[17] M. Reed, B. Simon. *Methods of modern mathematical physics IV: Analysis of operators.* Academic Press, 1978.

[18] J. Van der Corput. Zahlentheoretische Abschätzungen mit Anwendungen auf Gitterpunkteprobleme. *Math. Z.* 17, p. 250-259 (1923).

A. Ancona
Département de Mathématiques
Bat. 425
Université Paris-Sud
91 405 Orsay Cedex, France
Alano.Ancona@math.u-psud.fr

B. Helffer
Département de Mathématiques
Bat. 425
Université Paris-Sud
91 405 Orsay Cedex, France
Bernard.Helffer@math.u-psud.fr

T. Hoffmann-Ostenhof
Institut für Theoretische Chemie
Universität Wien
Währinger Strasse 17
A-1090 Wien, Austria
and
International Erwin Schrödinger
Institute for Mathematical Physics
Boltzmanngasse 9, A-1090
Wien, Austria
thoffman@esi.ac.at

Additive Structure of Multiplicative Subgroups of Fields and Galois Theory

Dedicated to Professor Tsit-Yuen Lam
in friendship and admiration.

Louis Mahé, Ján Mináč[1], and Tara L. Smith[2]

Received: January 9, 2004

Communicated by Ulf Rehmann

Abstract. One of the fundamental questions in current field theory, related to Grothendieck's conjecture of birational anabelian geometry, is the investigation of the precise relationship between the Galois theory of fields and the structure of the fields themselves. In this paper we initiate the classification of additive properties of multiplicative subgroups of fields containing all squares, using pro-2-Galois groups of nilpotency class at most 2, and of exponent at most 4. This work extends some powerful methods and techniques from formally real fields to general fields of characteristic not 2.

2000 Mathematics Subject Classification: Primary 11E81; Secondary 12D15

[1]The two first authors were supported by the Volkswagen-Stiftung (RIP-Program at Oberwolfach), research supported in part by the Natural Sciences and Engineering Research Council of Canada and by the special Dean of Science Fund at the University of Western Ontario, supported by the Mathematical Sciences Research Institute, Berkeley

[2]Research supported in part by the Taft Memorial Fund of the University of Cincinnati

§1. Introduction

Let F be a field of characteristic not 2 and T be a multiplicative subgroup of $\dot{F} = F \setminus \{0\}$ containing the squares. By the additive structure of T, we mean a description of the T-cosets forming $T + aT$. The purpose of this article is to relate the additive structure of such a group T, to some Galois pro-2-group H associated with T. In the case when T is a usual ordering, the group H is a group of order 2. In the general case, H is a pro-2-group of nilpotency class at most 2, and of exponent at most 4. Therefore the structure of H is relatively simple, and this is one of the attractive features of this investigation.
One of our main motivations is to extend Artin-Schreier theory to this general situation. In classical Artin-Schreier theory as modified by Becker, one studies euclidean closures and their relationship with Galois theory [ArSc1, ArSc2, Be]. Recall that such a closure is a maximal 2-extension of an ordered field to which the given ordering extends. (See [Be].)
It came as a surprise to us that for a good number of isomorphism types of groups H as above, we could provide a complete algebraic characterization of the multiplicative subgroups of $\dot{F}/\dot{F}^2$ associated with H, entirely analogous to the classical algebraic description of orderings of fields. We thus obtain a fascinating direct link between Galois theory and additive properties of multiplicative subgroups of fields.
We obtain in particular a Galois-theoretic characterization of rigidity conditions (Proposition 3.4 and Proposition 3.5) using "small" Galois groups, and a full classification of rigid groups T (§7). We also know how to make closures (as defined below) with respect to these rigid "orderings" (§8).
In §9 we refine the notion of H-orderings of fields. We show that under natural conditions, we can control the behaviour of the additive structure of these orderings under quadratic extensions. It is worthwhile to point out that each finite Galois 2-extension can be obtained by successive quadratic extensions. Therefore, it is sufficient to investigate quadratic extensions.
We have in §2 a nice illustration of what a W-group can or cannot be. Since the W-group of the field F, together with its level, determines the Witt ring $W(F)$, it is clear that every result about the W-group of F and its subgroups will provide information on $W(F)$.
This fits together with one of the main ideas behind this work (see §10): obtaining new Local-Global Principles for quadratic forms, with respect to these new "orderings." This will be the subject of a subsequent article.

We now enter into more detail, fix some notation, and present a more technical outline of the structure of the paper.

Notation 1.1. All fields in this paper are assumed to be of characteristic not 2, with any exceptions clearly pointed out. Occasionally we denote a field extension K/F as $F \longrightarrow K$. The compositum of two fields K and L contained in a larger field is denoted as KL. Recall that the *level* of a field F is the smallest natural number $n > 0$ such that -1 is a sum of n squares in F or

∞ if no such n exists. Given a field F, we denote by $F(\sqrt{\dot{F}})$ the compositum of all quadratic extensions of F, and by $F^{(3)}$ the compositum of all quadratic extensions of $F(\sqrt{\dot{F}})$ which are Galois over F. (The field $F(\sqrt{\dot{F}})$ was denoted by $F^{(2)}$ in previous papers (e.g. [MiSm2]), and this explains the notation $F^{(3)}$.) The W-group of the field F is then defined as $\mathcal{G}_F = \mathrm{Gal}(F^{(3)}/F)$. This W-group is the Galois-theoretic analogue of the Witt ring, in that if two fields have isomorphic Witt rings, then their W-groups are also isomorphic. Conversely, if two fields have isomorphic W-groups, then their Witt rings are also isomorphic, provided that the fields have the same level when the quadratic form $\langle 1, 1\rangle$ is universal over one of the fields. (See [MiSp2, Theorem 3.8].)
We denote by $\Phi(\mathcal{G}_F)$ the Frattini subgroup of $\mathcal{G}_F$. The Frattini subgroup is by definition the intersection of the maximal proper subgroups H of $\mathcal{G}_F$. (This means that H is a maximal subgroup of $\mathcal{G}_F$ among the family of all closed subgroups of $\mathcal{G}_F$ not equal to $\mathcal{G}_F$. It is a basic fact in the theory of pro-2-groups that each such subgroup of $\mathcal{G}_F$ is a closed subgroup of $\mathcal{G}_F$ of index two.) Notice that $\mathrm{Gal}(F^{(3)}/F(\sqrt{\dot{F}})) = \Phi(\mathcal{G}_F)$. In the case of a pro-2-group G, the Frattini subgroup is exactly the closure of the group generated by squares. Observe that for each closed subgroup H of $\mathcal{G}_F$ we have $\Phi(H) \subseteq \Phi(\mathcal{G}_F) \cap H$. We say that a closed subgroup $H \subseteq \mathcal{G}_F$ satisfying $\Phi(H) = H \cap \Phi(\mathcal{G}_F)$ is an *essential* subgroup of $\mathcal{G}_F$. Two essential subgroups H_1, H_2 are *equivalent* if $H_1\Phi(\mathcal{G}_F) = H_2\Phi(\mathcal{G}_F)$. In general, for a closed subgroup H of $\mathcal{G}_F$, we have $H = \mathcal{E} \times \prod_i(\mathbb{Z}/2\mathbb{Z})_i$ where $\mathcal{E}$ is essential: $\Phi(H) = \Phi(\mathcal{E})$ and $\Phi(\mathcal{G}_F) \cap H \cong \Phi(\mathcal{E}) \times \prod_i(\mathbb{Z}/2\mathbb{Z})_i$. The equivalence class of $\mathcal{E}$ is that of H, and equivalent essential subgroups are always isomorphic. (See [CrSm, Theorem 2.1]. The proof is carried out in the case when H is finite, and the routine technical details necessary for extending the proof for an infinite H have been omitted.)

We recall that a subset $S = \{\sigma_i, i \in I\}$ of a pro-p-group G is called a set of generators of G if G is the smallest closed subgroup containing S, and for each open subgroup U of G, all but finitely many elements of S are in U. It is well-known that each pro-p-group G contains a set of generators. A set of generators S of G is called minimal if no proper subset of S generates S. (See [Koc, 4.1].)
We now give the field-theoretic interpretation of the notion of an essential subgroup of $\mathcal{G}_F$. Let H be any closed subgroup of $\mathcal{G}_F$ and let L be the fixed field of H. Let N and M be the fixed fields of $\Phi(H)$ and $\Phi(\mathcal{G}_F) \cap H$ respectively. Because $\Phi(H) \subseteq \Phi(\mathcal{G}_F) \cap H$, we see that $M \subseteq N$ and equality holds for one of the inclusions if it holds for the other. Finally observe that M is the compositum of $F(\sqrt{\dot{F}})$ and L, and that N is the compositum of all quadratic extensions of L contained in $F^{(3)}$. Summarizing the discussion above we obtain:

Proposition 1.2. *Let H be a closed subgroup of $\mathcal{G}_F$ and L be the fixed field of H. Then H is an essential subgroup of $\mathcal{G}_F$ if and only if the maximal multiquadratic extension of L contained in $F^{(3)}$ is equal to the compositum of L and $F(\sqrt{\dot{F}})$.*

Kummer theory and Burnside's Basis Theorem allow us to prove the following:

PROPOSITION 1.3. *For H a closed subgroup of $\mathcal{G}_F$, the assignment*

$$H \mapsto u(H) = P_H := \{a \in \dot{F} \mid (\sqrt{a})^\sigma = \sqrt{a}, \quad \forall \sigma \in H\}$$

induces a $1-1$ correspondence between equivalence classes of essential subgroups of $\mathcal{G}_F$ and multiplicative subgroups of $\dot{F}/\dot{F}^2$.

Proof. Recall from Kummer theory that $\mathrm{Gal}(F(\sqrt{\dot{F}})/F)$ is the Pontrjagin dual of the discrete group $\dot{F}/\dot{F}^2$ under the pairing $(g,[f]) = g(\sqrt{f})/\sqrt{f}$ of $\mathrm{Gal}(F(\sqrt{\dot{F}})/F)$ with $\dot{F}/\dot{F}^2$, with values in $\mathbb{Z}/2\mathbb{Z} \cong \{\pm 1\}$. (See [ArTa, Chapter 6].)
Assume that H_1 and H_2 are two essential subgroups of $\mathcal{G}_F$ such that $P_{H_1} = P_{H_2} =: P$. This means $\frac{H_1\Phi(\mathcal{G}_F)}{\Phi(\mathcal{G}_F)} = \frac{H_2\Phi(\mathcal{G}_F)}{\Phi(\mathcal{G}_F)}$ because they are both the annihilator of P under the pairing above. (See [Mo, Chapter 5].) Therefore $H_1\Phi(\mathcal{G}_F) = H_2\Phi(\mathcal{G}_F)$. Hence u is injective on equivalent classes of essential subgroups.
In order to prove that u is surjective, consider any subgroup P of $\dot{F}$ containing $\dot{F}^2$. Let $\{[a_i], i \in I\} \subset \dot{F}/P$ be an $\mathbb{F}_2$-basis of $\dot{F}/P$ and $\{\bar{\sigma}_i, i \in I\}$ be elements of $\mathcal{G}_F/\Phi(\mathcal{G}_F)$ such that $\bar{\sigma}_i(\sqrt{a_i})/\sqrt{a_i} = -1, \bar{\sigma}_i(\sqrt{a_j}) = \sqrt{a_j}$ for $j \neq i$ and $\bar{\sigma}_i(\sqrt{p}) = \sqrt{p}$ for all $p \in P$.
From [Koc, 4.4] we see that there exists a subset $S = \{\sigma_i | i \in I\}$ of $\mathcal{G}_F$ such that the image of each σ_i in $\mathcal{G}_F/\Phi(\mathcal{G}_F)$ is $\bar{\sigma}_i$ and for each open subgroup U of $\mathcal{G}_F$ all but finitely many elements of S are in U. Set H to be the smallest closed subgroup of $\mathcal{G}_F$ containing S. Because $H/\Phi(H) = \langle \bar{\sigma}_i | i \in I\rangle :=$ the smallest closed subgroup of $\mathcal{G}_F/\Phi(\mathcal{G}_F)$ generated by $\{\bar{\sigma}_i | i \in I\}$, and $P = P_H$ we see that H is an essential subgroup of $\mathcal{G}_F$ such that $u(H) = P$. $\square$

The motivation for this study of essential subgroups grew out of the observation in [MiSp1] that for $H \cong \mathbb{Z}/2\mathbb{Z}$, if $P_H \neq \dot{F}/\dot{F}^2$ (i.e. if $H \cap \Phi(\mathcal{G}_F) = \{1\}$), then P_H is in fact the positive cone of some ordering on F. The reader is referred to [L2] for further details on orderings and connections to quadratic forms. Some convenient references for basic facts on quadratic forms are [L1] and [Sc].
Since the presence or absence of $\mathbb{Z}/2\mathbb{Z}$ as an essential subgroup of $\mathcal{G}_F$ determines the orderings or lack thereof on F, one wonders whether other subgroups of $\mathcal{G}_F$ also yield interesting information about F. We make the following definition.

DEFINITION 1.4.
(1) Let $\mathcal{C}$ denote the category of pro-2-groups of exponent at most 4, for which squares and commutators are central. (Observe that since each commutator is a product of (three) squares, it is sufficient to assume that all squares are central.) All W-groups are in category $\mathcal{C}$. In particular $\Phi(\mathcal{G}_F)$ is in the center of $\mathcal{G}_F$, for any $\mathcal{G}_F$. See [MiSm2] for further details. Note that $\mathcal{C}$ is a full subcategory of the category of pro-2-groups. This allows us to freely use all of the properties of pro-2-groups.

(2) Let H be a pro-2-group. An embedding $\varphi\colon H \longrightarrow \mathcal{G}_F$ is an *essential embedding* if $\varphi(H)$ is an essential subgroup of $\mathcal{G}_F$. Note that if H embeds in $\mathcal{G}_F$, then H has to be in category $\mathcal{C}$.

(3) An *H-ordering* on F is a set $P_{\varphi(H)}$ where φ is an essential embedding of H in $\mathcal{G}_F$.

(4) Let (F,T) be a field with an H-ordering T. We say that (L,S) *extends* (F,T) if L is an extension field of F in the maximal Galois 2-extension $F(2)$ of F, S is a subgroup of $\dot{L}$ containing $\dot{L}^2$, $T = S \cap \dot{F}$, and the induced injection $\dot{F}/T \longrightarrow \dot{L}/S$ is an isomorphism. We also say (L,S) is a *T-extension* of F. (We will see in Propositions 4.1 and 4.2 that maximal T-extensions always exist, and that a maximal such extension (L,S) in $F(2)$ has $S = \dot{L}^2$.) An extension (L,S) of (F,T) is said to be an *H-extension* if S is an H-ordering of L.

(5) An extension (L,S) of (F,T) is called an *H-closure* if it is a maximal T-extension which is also an H-extension. Note this implies $S = \dot{L}^2$ and $\mathcal{G}_L \cong H$. Observe that maximal H-extensions (K,S) need not satisfy $S = \dot{K}^2$.

We set the following notation: C_n denotes the cyclic group of order n, D denotes the dihedral group of order 8, Q denotes the quaternion group of order 8.

If G_1 and G_2 are in $\mathcal{C}$, we denote by $G_1 * G_2$ the free product (i.e. the coproduct) of the two groups in category $\mathcal{C}$. Then G_1 and G_2 are canonically embedded in $G_1 * G_2$ and the latter can be thought of as $(G_1 \times [G_1, G_2]) \rtimes G_2$ with the obvious action of G_2 on the inner factor. (See [MiSm2].) For example, $D \cong C_2 * C_2$.

Let $a \in \dot{F} \backslash \dot{F}^2$. By a C_4^a-extension of a field F, we mean a cyclic Galois extension K of F of degree 4, with $F(\sqrt{a})$ as its unique quadratic intermediate extension. Let $a, b \in \dot{F}$ be linearly independent modulo $\dot{F}^2$. By a $D^{a,b}$-extension of F we mean a dihedral Galois extension L of F of degree 8, containing $F(\sqrt{a}, \sqrt{b})$, for which $\mathrm{Gal}(L/F(\sqrt{ab})) \cong C_4$. Observe that any C_4-extension is a C_4^a-extension for an $a \in F$, and that any D-extension is a $D^{a,b}$-extension for a suitable $a, b \in \dot{F}$.

The following result is not hard to prove, and is a special case of more general results in [Fr]. (See also [L1, Exercise VII.8].)

Proposition 1.5. *There exists a C_4^a-extension of F if and only if $a \in \dot{F} \setminus F^2$ and the quaternion algebra $\left(\frac{a,a}{F}\right)$ is split. There exists a $D^{a,b}$-extension of F if and only if $a, b \in \dot{F}$ are independent modulo squares and the quaternion algebra $\left(\frac{a,b}{F}\right)$ is split.*

This proposition is one of the main tools we use to link the Galois-theoretic properties of an essential subgroup H of $\mathcal{G}_F$ to the algebraic properties of an H-ordering. Since we will need to refer to such extensions often in the sequel, we sketch the subfield lattice of a $D^{a,b}$-extension L/F.

$$\begin{tikzcd}[arrows=dash]
 & & L \arrow[dll] \arrow[dl] \arrow[d] \arrow[dr] \arrow[drr] & & \\
K_1 \arrow[dr] & K_2 \arrow[d] & F(\sqrt{a},\sqrt{b}) \arrow[dl] \arrow[d] \arrow[dr] & K_3 \arrow[d] & K_4 \arrow[dl] \\
 & F(\sqrt{a}) \arrow[dr] & F(\sqrt{ab}) \arrow[d] & F(\sqrt{b}) \arrow[dl] & \\
 & & F & &
\end{tikzcd}$$

The paper is organized as follows.
In §2, we determine centralizers of involutions in W-groups. These results imply in particular that the only abelian groups which can appear as essential nontrivial subgroups of a W-group are C_2 and $(C_4)^I$ where I is some nonempty set. We also determine the possible nonabelian subgroups generated by two elements. In Theorem 2.7 we provide a strong restriction on possible finite subgroups of a W-group. Some of these results are important in determining the cohomology rings of W-groups.
In §3 we show how properties of an H-ordering T, such as stability under addition or rigidity, may be described in a Galois-theoretic way. The definition and first properties of extensions and closures are given in §4. We illustrate with Proposition 4.4 that even in a very geometric situation, we cannot expect that every H-ordering T admits a closure. In Proposition 4.5, that is a corollary of [Cr2, Theorem 5.5], we also point out that this leads to a negative answer to a strong version of the question asked in [Ma]: there are fields F having no field extension $F \longrightarrow K$ with $W_{red}(K) \cong W(K)$, such that the induced map $W_{red}(F) \longrightarrow W_{red}(K)$ is an isomorphism. Later in §8 we are able to provide a similar example of a field F with a subgroup T of $\dot{F}$ such that the associated Witt ring $W_T(F)$ is isomorphic to $W(\mathbb{Q}_p), p \equiv 1(4)$ but again there is no field extension $F \longrightarrow K$ inducing the isomorphism $W_T(K) \cong W(K)$. This example is interesting because $\mid \dot{F}/T \mid$ is finite. (For details see Example 8.14, Proposition 8.15, and Remark 8.16.)
In §5 and §6 we study the case of essential subgroups H generated by 1 or 2 elements, and show that they admit closures.
In §7 we give a complete Galois-theoretic, as well as an algebraic classification of rigid orderings, and in §8 we show that they admit closures, provided that in the case of $C(I)$, the associated valuation is not dyadic. (See Theorem 8.15 and Example 8.14.) In Example 6.4 we see that the link between the additive structure of an H-ordering and the Galois-theoretic properties of H is not as tight as we might have expected. This leads us to investigate this question more thoroughly in §9. Actually, with a few natural extra requirements on the

Galois groups we are considering, this can be fixed. We are then able to obtain a perfect identification between the two aspects.
As we have already said, application of this theory to local-global principles for quadratic forms will constitute the core of a subsequent paper. In the conclusion we illustrate by an easy example, what we intend to do in this direction.
The authors would like to acknowledge Professors A. Adem, J.-L. Colliot-Thélène, T. Craven, B. Jacob, D. Karagueuzian, J. Koenigsmann, T.-Y. Lam, D. Leep and H. W. Lenstra, Jr. for valuable discussions concerning the results in this paper; and also the hospitality of the Mathematical Sciences Research Institute at Berkeley, the Department of Mathematics at the University of California at Berkeley, and the Mathematisches Forschungsinstitut at Oberwolfach, which the authors were privileged to visit during the preparation of this paper. We also wish to thank the anonymous referee for valuable comments and also for suggestions for polishing the exposition.

§2. Groups not appearing as subgroups of W-groups

In this section we show that no essential subgroup of $\mathcal{G}_F$ can have C_2 as a direct factor (except in the trivial case where the subgroup is C_2), nor can Q appear as a subgroup of $\mathcal{G}_F$. These two facts will then be used to show that the four nonabelian groups $C_2 * C_2 = D, C_2 * C_4, C_4 \rtimes C_4$ and $C_4 * C_4$, together with the abelian group $C_4 \times C_4$, comprise all of the possible two-generator essential subgroups of W-groups. Thus we have a good picture of the minimal realizable and unrealizable subgroups. We further show that every finite subgroup of a W-group is in fact an "S-group" as defined in [Jo]. (We shall call such groups "split groups" here.) The fact that Q is not a subgroup of $\mathcal{G}_F$ is actually a consequence of this last result.
Since we are working in category $\mathcal{C}$ in the presentations of groups by generators and relations, we write only those relations which do not follow from the fact that our groups are in $\mathcal{C}$.

Lemma 2.1. [Mi], [CrSm] *The groups $C_2 \times C_2$ and $C_4 \times C_2$ cannot be realized as essential subgroups of $\mathcal{G}_F$ for any field F.*

Proof. Assume $H = \langle \sigma, \tau \mid \sigma^2 = \tau^2 = [\sigma,\tau] = 1\rangle \subseteq \mathcal{G}_F$ or $H = \langle \sigma, \tau \mid \sigma^2 = [\sigma,\tau] = 1\rangle$, and assume $\sigma, \tau, \sigma\tau \notin \Phi(\mathcal{G}_F)$. Then from [MiSp1] we know that $P_{\langle\sigma\rangle}$ is a C_2-ordering which is a usual ordering. In particular $-1 \notin P_{\langle\sigma\rangle}$ and $\sigma(\sqrt{-1}) = -\sqrt{-1}$.
Now choose an element $b \in \dot{F}\backslash\dot{F}^2$ for which $\sqrt{b}^{\sigma} = \sqrt{b}$ and $\sqrt{b}^{\tau} = -\sqrt{b}$. Such an element b exists since $\sigma, \tau, \sigma\tau \notin \Phi(\mathcal{G}_F)$. Consider the image $\langle\bar\sigma, \bar\tau\rangle$ of H inside the Galois group G of a $D^{b,-b}$-extension K of F. (Because $(\sqrt{-1})^{\sigma} = -\sqrt{-1}$ we see that $-b$ is not a square in F, and we can conclude that the elements b and $-b$ are linearly independent when they are considered as elements in $\dot{F}/\dot{F}^2$.) The fixed field K_σ of $\bar\sigma$ cannot contain $\sqrt{-b}$, so it must be one of the two subfields of index 2 in K not containing $\sqrt{-b}$. On the other hand, the fixed field K_τ of τ cannot contain $\sqrt{b}$, so considering the subfield lattice, we see that

$K_\sigma \cap K_\tau = F$. Then the image of H in G generates G, which means σ and τ cannot commute. This is a contradiction, so H cannot exist as an essential subgroup of $\mathcal{G}_F$. $\square$

From the lemma above we immediately obtain the following result, which is used in [AKMi] to investigate those fields F for which the cohomology ring $H^*(\mathcal{G}_F)$ is Cohen-Macaulay.

Corollary 2.2. *Let σ be any involution in $\mathcal{G}_F \backslash \Phi(\mathcal{G}_F)$ and set $E_\sigma = \Phi(\mathcal{G}_F) \times \langle\sigma\rangle$. Then the centralizer $Z(E_\sigma)$ of E_σ in $\mathcal{G}_F$ is E_σ itself.*

Proof. If $\tau \in Z(E_\sigma) \setminus E_\sigma$ then $[\tau, \sigma] = 1$ and $\langle\tau, \sigma\rangle = C_2 \times C_2$ or $C_4 \times C_2$, where $\langle\tau, \sigma\rangle$ is an essential subgroup of $\mathcal{G}_F$. From Lemma 2.1, this is a contradiction, and we see $\tau \in E_\sigma$ as desired. $\square$

Corollary 2.3. *No essential subgroup of $\mathcal{G}_F$ can have C_2 as a direct factor (except in the trivial case where the subgroup is C_2).*

Proof. Since $\Phi(H \times C_2) = \Phi(H)$, if $H \times C_2$ is a subgroup of $\mathcal{G}_F$ with $\Phi(H \times C_2) = (H \times C_2) \cap \Phi(\mathcal{G}_F)$, then the C_2-factor is not in $\Phi(\mathcal{G}_F)$. Take any single element $\sigma \in H \backslash \Phi(H)$. Then $\langle\sigma\rangle \times C_2 \cong C_2 \times C_2$ or $C_4 \times C_2$, which cannot be an essential subgroup. Therefore neither can $H \times C_2$. $\square$

Proposition 2.4. *The quaternion group Q cannot appear as a subgroup of $\mathcal{G}_F$.*

Proof. Suppose $Q = \langle\sigma, \tau | \sigma^2 = \tau^2 = [\sigma, \tau]\rangle \subseteq \mathcal{G}_F$. If $-1 \in F^2$, then $F = F^2 + F^2$ and since $\mathcal{G}_F$ is not trivial, we have $F \neq F^2$. Therefore there exists an element $a \in \dot{F} \setminus F^2$ and for any such a we have a C_4^a-extension L/F. Since Q does not admit C_4 as a quotient, the images $\bar{\sigma}, \bar{\tau}$ of σ, τ in $\mathrm{Gal}(L/F)$ have order ≤ 2 and they fix the only subfield $F(\sqrt{a})$ of codimension 2 in L. Then σ, τ act as the identity on the compositum $F(\sqrt{\dot{F}})$ of these fields and hence are in $\Phi(F)$. Since they do not commute, this is impossible and we must have $-1 \notin F^2$.
Now suppose $-1 \in P_{\langle\sigma\rangle}$. Since $\sigma \notin \Phi(F)$, there exists $a \in \dot{F}$ such that $a \notin P_{\langle\sigma\rangle}$ and hence $-a \notin P_\sigma$. Then a and $-a$ are linearly independent modulo $\dot{F}^2$ and there exists a $D^{a,-a}$-extension L/F. Again, since Q has no C_4 quotient, the image $\bar{\sigma}$ of σ in $\mathrm{Gal}(L/F)$ has order ≤ 2 and must fix a codimension 2 subfield of L. Therefore $\bar{\sigma}$ must fix $\sqrt{a}$ or $\sqrt{-a}$, and this is a contradiction with $a, -a \notin P_{\langle\sigma\rangle}$. Hence we see that $-1 \notin P_{\langle\sigma\rangle}$.
Because σ and τ are linearly independent modulo $\Phi(\mathcal{G}_F)$, there exists an element $b \in \dot{F} \setminus \dot{F}^2$ such that $\sqrt{b}^\sigma = \sqrt{b}$ and $\sqrt{b}^\tau = -\sqrt{b}$. Then b and $-b$ are linearly independent modulo $\dot{F}^2$, and there exists a $D^{b,-b}$-extension K/F. Because D is not a homomorphic image of Q, the image of Q is a proper subgroup of $\mathrm{Gal}(K/F)$. On the other hand, because both σ and τ act nontrivially and in a different way on $F(\sqrt{b}, \sqrt{-b})/F$ we see that their images $\bar{\sigma}$ and $\bar{\tau}$ in $\mathrm{Gal}(K/F)$ generate the entire Galois group $\mathrm{Gal}(K/F)$, which is a contradiction! $\square$

Theorem 2.5. *The only groups generated by two elements which can arise as essential subgroups of $\mathcal{G}_F$ are the five groups $C_2 * C_2$, $C_2 * C_4$, $C_4 * C_4$, $C_4 \times C_4$, and $C_4 \rtimes C_4$.*

Proof. Let H be generated by x, y. We have an exact sequence

$$1 \to \Phi(H) \to H \to C_2 \times C_2 \to 1,$$

where $\Phi(H) \cong (C_2)^k$ is generated by $x^2, y^2, [x, y]$, so $k \leq 3$. Then $|H| = 2^{k+2}$, so $|H| \leq 32$, and $|H| = 32$ if and only if $|\Phi(H)| = 8$, if and only if $H \cong C_4 * C_4$. Otherwise $|H| = 8$ or 16, and there are only a few groups to consider. If $|H| = 8$, necessarily $H \cong C_2 * C_2$, as all other groups of order 8 and exponent at most 4 either have C_2 as a direct factor or are isomorphic to Q.
There are fourteen groups of order 16; among these, five are abelian, and by Lemma 2.1 only $C_4 \times C_4$ among these can be an essential subgroup of $\mathcal{G}_F$. Among the nine nonabelian groups, two have C_2 as a direct factor, and four more have exponent 8. The remaining three are the groups $C_2 * C_4$, $C_4 \rtimes C_4$, and DC, the central product of D and C_4 amalgamating the unique central subgroup of order 2 in each group. This group, however, has Q as a subgroup (see [LaSm]), so cannot be an essential subgroup of $\mathcal{G}_F$. $\square$

That the group Q cannot appear as a subgroup of any W-group is a special case of a more general description of the kinds of groups which can appear as essential subgroups of W-groups. All finite subgroups must in fact be "split groups", which we define next. These are the same as "S-groups" as defined in [Jo]. The quaternion group Q is not such a group.

Definition 2.6. Let G be a nontrivial finite group and $X = \{x_1, x_2, \ldots, x_n\}$ be an ordered minimal set of generators for G. We say that G satisfies the *split condition with respect to* X if $\langle x_1 \rangle \cap [G, G]\langle x_2, \ldots, x_n \rangle = \{1\}$. The group G is called a *split group* if it has a minimal generating set with respect to which it satisfies the split condition. We also take the trivial group to be a split group.

We refer to G above as split because if G satisfies the split condition with respect to X then G can be written as a semidirect product $G = ([G, G]\langle x_2, \ldots, x_n \rangle) \rtimes \langle x_1 \rangle$.

Theorem 2.7. *Let $\mathcal{G}_F$ be a W-group, and let G be any finite subgroup of $\mathcal{G}_F$. Then G is a split group.*

Proof. Each finite subgroup H of $\mathcal{G}_F$ can be written as $H = G \times \prod_1^m C_2$ for some $m \in \mathbb{N} \cup \{0\}$, where G is an essential subgroup of $\mathcal{G}_F$ [CrSm]. Thus it is enough to prove the theorem for G a finite essential subgroup of $\mathcal{G}_F$.
Then let G be such a group and let P_G be the associated G-ordering. Let $\dot{F}/P_G = \langle a_1 P_G, \ldots, a_n P_G \rangle$ so that the cosets $a_i P_G$ give a minimal generating set for $\dot{F}/P_G$. Further set $\{\sigma_1, \ldots, \sigma_n\}$ to be a minimal generating set for G such that $\sigma_i(\sqrt{a_j}) = (-1)^{\delta_{ij}}\sqrt{a_j}$ where δ_{ij} is the Kronecker delta. (This is possible because G is an essential subgroup of $\mathcal{G}_F$, so that a minimal set of

generators for G can be extended to a minimal (topological) generating set of $\mathcal{G}_F$.)

Assume first that we can choose the representatives a_i in such a way that $a_1t_1 + a_1t_2 = f^2 \in \dot{F}^2$ for some $t_1, t_2 \in P_G$. (Note that this is equivalent to saying that $a_1 \in P_G + P_G$.) In this instance, there are two cases to consider.

First, suppose that t_1, t_2 are congruent mod $\dot{F}^2$. Then there exists $g \in \dot{F}$ such that $a_1t_1 + a_1t_1g^2 = f^2$, and so $a_1t_1f^2 = (a_1t_1)^2 + (a_1t_1g)^2$, and a_1t_1 is a sum of two squares in F which is not itself a square. Thus we have a $C_4^{a_1t_1}$-extension L of F. We claim that G satisfies the split condition with respect to $\{\sigma_1, \ldots, \sigma_n\}$. Checking this condition is equivalent to showing $\sigma_1^2 \notin [G,G]\langle\sigma_2, \ldots, \sigma_n\rangle$. Suppose it is not true. Then we have an identity $\sigma_1^2 \prod_{1\le i<j\le n}[\sigma_i, \sigma_j]^{\epsilon_{ij}} \prod_{k=2}^n \sigma_k^{2\epsilon_k} = 1$ in G, where $\epsilon_{ij}, \epsilon_k \in \{0,1\}$. Restricting to L we see that $\sigma_1^2|_L = 1$. This cannot be the case as σ_1 does not fix $\sqrt{a_1t_1}$. Thus in this case G is a split group.

Next suppose that $t_1\dot{F}^2 \neq t_2\dot{F}^2$. In this case we can find a $D^{a_1t_1, a_1t_2}$-extension L/F. Assuming again that G does not satisfy the split condition with respect to $\{\sigma_1, \ldots, \sigma_n\}$, we again have an identity $\sigma_1^2 \prod_{1\le i<j\le n}[\sigma_i, \sigma_j]^{\epsilon_{ij}} \prod_{k=2}^n \sigma_k^{2\epsilon_k} = 1$ in G, where $\epsilon_{ij}, \epsilon_k \in \{0,1\}$. Since each of the $\sigma_i, i = 2, \ldots, n$ acts trivially on $F(\sqrt{a_1t_1}, \sqrt{a_1t_2})$, we see that each $\sigma_i, i > 1$ is central when restricted to L. Thus again $\sigma_1^2|_L = 1$. But $\sigma_1|_L$ generates $\mathrm{Gal}(L/F(\sqrt{a_1t_1} \cdot \sqrt{a_1t_2})) \cong C_4$. Hence G is a split group.

Finally, assume that we cannot choose $a_1 \in P_G + P_G$. Then necessarily $P_G + P_G \subseteq P_G \cup \{0\}$. If $-1 \in P_G$, then $P_G = \dot{F}$ and $G = \{1\}$ which is a split group. Otherwise P_G is a preordering in F, and we may write $P_G = \cap_{i=1}^n P_i$ where each P_i is an ordering, and each $P_i = \{\, f \in \dot{F} \mid \sqrt{f}^{\sigma_i} = \sqrt{f}\,\}$. Then $\{\sigma_1, \ldots, \sigma_n\}$ is a minimal generating set for G. Furthermore, each $\sigma_i^2 = 1$. (See [MiSp1] for details. The definition of a preordering in a field F can be found in [L2, Chapter 1], together with the basic properties of preordered rings.) Thus again we see that G is a split group. $\square$

Corollary 2.8. *Each nontrivial finite subgroup G of a W-group $\mathcal{G}_F$ can be obtained inductively from copies of C_2 and C_4 by taking semidirect products at each step. Thus we have $G = G_n \supseteq G_{n-1} \supseteq \cdots \supseteq G_1 \supseteq G_0$ where $G_0 \in \{C_2, C_4\}$, and $G_i = G_{i-1} \rtimes C_2$ or $G_i = G_{i-1} \rtimes C_4$ for each $i = 1, \ldots, n$.*

Proof. We proceed by induction on the number of generators of G. The statement clearly holds for any group G generated by a single element. Let G be any (nontrivial) finite subgroup of the W-group $\mathcal{G}_F$. Then we can write $G = H \times \prod_1^m C_2$ where H is essential, and G, if not equal to H, is clearly built up as described from H, where the action in the semidirect product is trivial. We can choose a minimal set of generators $\{\sigma_1, \ldots, \sigma_n\}$ for H such that H satisfies the split condition with respect to these generators. Clearly $N := [H,H]\langle\sigma_2, \ldots, \sigma_n\rangle$ is a normal subgroup of H, and $H \cong N \rtimes \langle\sigma_1\rangle$, where $\langle\sigma_1\rangle \cong C_2$ or C_4. Since $N \cong \langle\sigma_2, \ldots \sigma_n\rangle \times \prod_1^k C_2$ (for some positive integer k), we finish by induction. $\square$

Example 2.9. Consider the W-group $\mathcal{G}_2$ of the 2-adic numbers $\mathbb{Q}_2$. It has the presentation $\langle \sigma, \tau, \rho \mid \sigma^2[\tau, \rho]\rangle$ in the category $\mathcal{C}$ of groups of exponent at most four with squares and commutators central. (See [MiSp2, Example 4.4].) A basis for $\dot{\mathbb{Q}}_2/\dot{\mathbb{Q}}_2^2$ is given by $\{[-1],[2],[5]\}$, and σ may be chosen to fix $\sqrt{2}$ and $\sqrt{5}$ but not $\sqrt{-1}$, τ to fix $\sqrt{-1}$ and $\sqrt{5}$ but not $\sqrt{2}$, and ρ to fix $\sqrt{-1}$ and $\sqrt{2}$ but not $\sqrt{5}$. Then $\mathcal{G}_2$ can be constructed inductively from copies of C_4 and C_2 using semidirect products as follows:

$$\begin{aligned}
G_0 &= \langle \rho \rangle \cong C_4\\
G_1 &= G_0 \times \langle [\sigma,\rho]\rangle \cong G_0 \times C_2\\
G_2 &= G_1 \rtimes \langle \sigma \rangle \cong G_1 \rtimes C_4\\
G_3 &= G_2 \times \langle [\sigma,\tau]\rangle \cong G_0 \times C_2\\
\mathcal{G}_2 &= G_3 \rtimes \langle \tau \rangle \cong G_3 \rtimes C_4
\end{aligned}$$

Thus $\mathcal{G}_2 \cong \{[(C_4 \times C_2) \rtimes C_4] \times C_2\} \rtimes C_4$.

Corollary 2.8 is an interesting generalization of the known structure of W-groups associated with Witt rings of finite elementary type. (See [Ma: pages 122 and 123].) In fact, all W-groups associated with Witt rings of finite elementary type can easily be seen to be built up from cyclic groups of order 2 or 4, using only semidirect products. First one checks that the groups associated with basic indecomposable groups are such groups. Then the group ring construction for Witt rings corresponds directly to taking a semidirect product with a cyclic group of order 4, while the direct product construction for Witt rings corresponds to taking a free product of W-groups in the appropriate category. But this in turn just involves taking a direct product with an appropriate number of copies of C_2 (representing the necessary commutators) and then taking a semidirect product with the generators of one of the initial W-groups. See [MiSm2] for details.

Corollary 2.8 is quite useful for the investigation of cohomology rings of W-groups. This is important in light of the recent proof of the Milnor Conjecture by Voevodsky [Vo]. In particular, Voevodsky's result shows that the cohomology rings of absolute Galois groups with $\mathbb{F}_2$-coefficients carry no more information about the base field than Milnor's K-theory mod 2. On the other hand, the cohomology rings of W-groups carry substantial additional information. (See [AKMi].)

Using [Jo: Cor, p. 370] and Theorem 2.7 above, we immediately obtain the following.

Corollary 2.10. *Let G be any nontrivial finite subgroup of a W-group $\mathcal{G}_F$. Then the cohomology ring $H^*(G,\mathbb{F}_2)$ contains nonnilpotent elements of degree 2, and hence of every even degree.*

§3. Galois groups and additive structures (1)

In this section we give a simple Galois-theoretic characterization of two important additive properties of H-orderings: stability under addition and rigidity.

This generalizes the results on rigidity and on the realizability of certain Galois groups obtained in [MiSm1].
For the rest of this paper, unless otherwise mentioned, or if clearly some non-essential subgroups are also considered, subgroups of $\mathcal{G}_F$ will always be *essential.* Nevertheless for the sake of the convenience of the reader we occasionally recall that the considered subgroups are essential. Throughout this paper we write $T + aT = \{t_1 + at_2 \mid t_1, t_2 \in T \cup \{0\}, t_1 + at_2 \neq 0\}$, so T and aT are always subsets of $T + aT$, and $T + aT \supseteq \dot{F}^2$. (Here T is any subgroup of $\dot{F}$ containing all squares in $\dot{F}$.)

PROPOSITION 3.1. *Let H be an essential subgroup of $\mathcal{G}_F$, and T its associated H-ordering. Then H has C_4 as a quotient if and only if $T + T \neq T$.*

Proof. First assume there exists $a \in T+T$ which is not in T. Let K be the fixed field of H in $F^{(3)}$. We construct a C_4^a-extension F_1 of $F_0 = F(\sqrt{T}) = K \cap F^{(2)}$ inside $F^{(3)}$. Then $L = KF_1$ is a C_4^a-extension of K in $F^{(3)}$, showing H has C_4 as a quotient. We may write $a = t_1 + t_2$, so $a^2 - at_1 = at_2$. Let $y = a - \sqrt{a}\sqrt{t_1} \in F_0(\sqrt{a})$, so $N_{F_0(\sqrt{a})/F_0}(y) = [a] \in \dot{F}_0/\dot{F}_0^2$. Then $F_1 = F_0(\sqrt{a}, \sqrt{y})$ is a C_4^a-extension of F_0. Since $yy^\sigma = y^2$ or $at_2 \in (\dot{F}_0(\sqrt{a}))^2$ for all $\sigma \in \mathrm{Gal}(F_0(\sqrt{a})/F)$, we see F_1 is Galois over F, and hence is contained in $F^{(3)}$.
Conversely, assume $T + T = T$. If $-1 \in T$, then $T = \dot{F}$ and $H = \{1\}$. If $-1 \notin T$, then T is a preordering, so T is an intersection of orderings, and there is an essential subgroup H_1 of $\mathcal{G}_F$ isomorphic with H and $K \subset \Phi(\mathcal{G}_F)$ such that $H_1 \times K$ is generated by involutions. This follows from the fact that each preordering is an intersection of C_2-orderings ([L2, Theorem 1.6]), a characterization of C_2-orderings in [MiSp1] and Proposition 1.3. Thus H_1 and consequently H as well, cannot have C_4 as a quotient. $\square$

Remark. If H has a C_4-quotient, then there exists a C_4^a-extension of F_0 where we may take a to be in F. However, it is not necessarily the case that $a \in T+T$. That is, the quaternion algebra $\left(\frac{a,a}{F(\sqrt{T})}\right)$ is split, so a can be represented as the sum of two squares in $F(\sqrt{T})$, but not necessarily as the sum of two elements in T. This can be seen in Example 6.4.
The following definition generalizes the notion of the rigidity of a field, and introduces the notion of the level of T. (See [Wa, page 1349].)

DEFINITION 3.2. Let T be a subgroup of $\dot{F}/\dot{F}^2$. We say that T has *level s* if -1 is a sum of s elements of T, and not a sum of $s-1$ elements of T. We say that this level is infinite if -1 is not such a sum for any natural number s. We say that the field F is *T-rigid*, or equivalently that T is rigid, if for every $a \notin T \cup -T$, we have $T + aT \subseteq T \cup aT$.

We have the following easy-to-prove but important property of rigid H-orderings:

PROPOSITION 3.3. *Let T be a rigid H-ordering on F. Then*
(1) The level of T is $1, 2$ or infinite.

(2) If the level of T is 2, then $T+T=T\cup -T$.

Proof. Let T be an H-ordering of finite level $s>1$ and let us write $-1=a+a_s$ with $a=a_1+\ldots+a_{s-1}$ and $a_i\in T$ for $i=1,\ldots,s$. If $a\in T\cup -T$ then since $a\notin -T$ we see $a\in T$ and s must be 2. Thus we may assume $a\notin T\cup -T$. If T is rigid, then $-1=a+a_s\in T+aT=T\cup aT$. This is a contradiction, proving (1).

Assume the level of T is 2. Then $-1\in T+T$ and $T\cup -T\subseteq T+T$. Suppose there is $a\in (T+T)\setminus(T\cup -T)$ and let us write $a=s+t, s,t\in T$. Then of course $-a\notin T\cup -T$ and we have $-t=s-a\in T+(-a)T=T\cup -aT$ by rigidity. But $-t\notin T$ because the level is 2, and $-t\notin -aT$ because $a\notin T$. This is again a contradiction, proving (2). □

Proposition 3.4. *Let H be an essential subgroup of $\mathcal{G}_F$, and let T be an H-ordering. Assume $-1\in T$. The following are equivalent.*

(1) *F is T-rigid.*
(2) *D is not a quotient of H.*
(3) *H is abelian.*

Proof. We will show (2) $\implies$ (1) $\implies$ (3) $\implies$ (2). For the first implication, we show the contrapositive. Thus assume that F is not T-rigid. Let K be the fixed field of H, and let $F_0=K\cap F(\sqrt{\dot F})=F(\{\sqrt{t}:t\in T\})$. We will construct a D-extension F_1 of F_0 inside $F^{(3)}$, and linearly disjoint with K. Then $L=KF_1$ will be a D-extension of K in $F^{(3)}$, showing that H has D as a quotient. Since F is not T-rigid and $-1\in T$, there exist $a,b\in\dot F\backslash T$ such that $b=t_1-at_2$, where $t_1,t_2\in T$ but $b\notin T\cup aT$. Let $y=\sqrt{t_1}+\sqrt{a}\sqrt{t_2}\in F_0(\sqrt{a})$, and let $F_1=F_0(\sqrt{a},\sqrt{b},\sqrt{y})$. Notice that $yy^\sigma\in\{\pm y^2,\pm b\}\subseteq F_0(\sqrt{a},\sqrt{b})^2$ for all $\sigma\in \mathrm{Gal}(F_0(\sqrt{a},\sqrt{b})/F)$, so F_1/F is Galois, and $F_1\subseteq F^{(3)}$. Then the usual argument (see [Sp] or [Ki, Theorem 5]) shows $\mathrm{Gal}(F_1/F_0)\cong D$. Also F_1 is linearly disjoint with K, as no proper quadratic extension of F_0 is in K.

Now assume F is T-rigid. To see that H is abelian, it is sufficient to show that for all $\sigma,\tau\in H$, the restrictions of σ,τ to any D-extension L of F commute. (This is because $F^{(3)}$ is the compositum of all quadratic, C_4- and D-extensions of F. (See [MiSp2, Corollary 2.18].) Thus if σ,τ commute on all D-extensions, they commute in $\mathcal{G}_F$.) Let $D^{a,b}$ be some dihedral quotient of $\mathcal{G}_F$, and let L be the corresponding extension of F. Denote as $\bar\sigma,\bar\tau$ the images of σ and τ in $D^{a,b}$ and suppose $[\bar\sigma,\bar\tau]\neq 1$. Then σ,τ must each move at least one of $\sqrt{a},\sqrt{b}$, and they cannot both act in the same way on these square roots. That implies $a,b,ab\notin T$. But $(\frac{a,b}{F})$ splits, so $b\in F^2-aF^2\subseteq T-aT=T+aT=T\cup aT$ by (1). Since $b\notin T$, we have $b\in aT$, which contradicts the fact that $ab\notin T$. Thus $[\sigma,\tau]=1$.

The final implication is trivial. □

It is worth observing that if $4\le|\dot F/T|$ and if H is abelian then $-1\in T$. Indeed if $4\le|\dot F/T|$ and $-1\notin T$, there exists $[a]\in\dot F/T$ such that $[a],[-a]$ are linearly independent in $\dot F/T$. Then there exist elements $\sigma,\tau\in H$ such that their restrictions to $F(\sqrt{a}, \sqrt{-a})$ generate $\mathrm{Gal}(F(\sqrt{a}, \sqrt{-a}))/F$. Subsequently images of σ, τ generate $\mathrm{Gal}(L/F)$ for any $D^{a,-a}$ extension L/F. Thus H is not abelian. In the next proposition we freely use the fact that if H_1 is an essential part of the subgroup H_0 of $\mathcal{G}_F$, then H_1 admits a quotient D if and only if H_0 admits a quotient D.

Proposition 3.5. *Let H be an essential subgroup of $\mathcal{G}_F$, and let T be an H-ordering. Assume $-1 \notin T$. Let K be the fixed field of H, and let H_0 be the subgroup of H which is the Galois group of $F^{(3)}/K(\sqrt{-1})$. The following are equivalent.*

(1) *F is $(T \cup -T)$-rigid.*
(2) *D is not a quotient of H_0.*
(3) *H_0 is abelian.*
(4) *Every D-extension of K in $F^{(3)}$ contains $K(\sqrt{-1})$.*

Proof. Let $S = T \cup -T$ and let H_1 be an essential part of H_0. Then S is an H_1-ordering, and the equivalence of the first three statements follows from the preceding proposition. If there exists a D-extension L of K not containing $K(\sqrt{-1})$, then $L(\sqrt{-1})$ will be a D-extension of $K(\sqrt{-1})$, and H_0 will have D as a quotient. This shows (2) $\implies$ (4). In order to show that (4) $\implies$ (3), assume there exist $\sigma, \tau \in H_0$ which do not commute. Then there exists some $D^{a,b}$-extension M of F such that $\mathrm{Gal}(M/F) = \langle \bar{\sigma}, \bar{\tau} \rangle$, where we denote by $\bar{\sigma}$ and $\bar{\tau}$ the images of σ and τ in $\mathrm{Gal}(M/F)$. Then σ and τ each move at least one of $\sqrt{a}, \sqrt{b}$ and cannot act in the same way on each. Thus $a, b, ab \notin S$. This gives a D-extension MK of K, which does not contain $\sqrt{-1}$. $\square$

§4. Maximal extensions, closures and examples

Given any C_2-ordering P on a field F, one can find a real closure L of F with respect to that ordering, inside a fixed algebraic closure $\bar{F}$. This means L is real closed and $P = \dot{L}^2 \cap F$, and then $\mathrm{Gal}(\bar{F}/L) \cong C_2$. Notice that for our purposes nothing is lost by considering a real closure of F inside a euclidean closure $F(2)$ rather than inside the algebraic closure $\bar{F}$. (See [Be].) We then obtain a C_2-closure $(L, \dot{L}^2)$ of (F, P), and this observation actually motivated the definition of H-closure given in Definition 1.4. The following two propositions show that for any subgroup T of $\dot{F}$, containing $\dot{F}^2$, maximal T-extensions always exist and have a nice property.

Proposition 4.1. *Let T be a subgroup of $\dot{F}/\dot{F}^2$. Then (F, T) possesses a maximal T-extension.*

Proof. Let $\mathcal{S}$ be the set of T-extensions (L, S) of (F, T) inside $F(2)$, and let us order $\mathcal{S}$ by $(L_1, S_1) \leq (L_2, S_2)$ if $L_1 \subset L_2$ and $S_2 \cap L_1 = S_1$. Then $\mathcal{S}$ is nonempty, since $(F, T) \in \mathcal{S}$. Now consider a totally ordered family (F_j, T_j) in $\mathcal{S}$. Let $K = \cup F_j$, $S = \cup T_j$. Then (K, S) is an upper bound for the family (F_j, T_j) in $\mathcal{S}$. Then by Zorn's Lemma $\mathcal{S}$ contains a maximal element, which is a maximal T-extension of (F, T). $\square$

Proposition 4.2. *Let (K, S) be a maximal T-extension of (F, T). Then $S = \dot{K}^2$.*

Proof. Assume $S \neq \dot{K}^2$ and choose $c \in S \setminus \dot{K}^2$. Let $L = K(\sqrt{c})$. Then $\dot{K}/S \cong \dot{K}\dot{L}^2/S\dot{L}^2$ is an $\mathbb{F}_2$-vector subspace and hence a summand of $\dot{L}/S\dot{L}^2$. Pick any projection φ of $\dot{L}/S\dot{L}^2$ onto $\dot{K}\dot{L}^2/S\dot{L}^2$. Set S' as the inverse image of $\ker\varphi$ in $\dot{L}$. Then the natural inclusions $\dot{F} \longrightarrow \dot{K}$ and $\dot{K} \longrightarrow \dot{L}$ induce the isomorphisms $\dot{F}/T \cong \dot{K}/S \cong \dot{L}/S'$, contradicting the maximality of (K, S). Thus we conclude that $S = \dot{K}^2$. □

Corollary 4.3. *An H-ordered field (F, T) is an H-closure if and only if $T = \dot{F}^2$.*

Proof. If (F, T) is an H-closure, then it is also a maximal T-extension, and $T = \dot{F}^2$ by the preceding proposition. Conversely, suppose $T = \dot{F}^2$. Let $L \supset F$ be any proper extension of F in $F(2)$. Then L contains a quadratic extension of F, so $\dot{L}^2 \cap F \supsetneq \dot{F}^2$ and L cannot extend (F, T). This shows that (F, T) is its own maximal T-extension, and as it is an H-ordering, it is an H-closure. □

Thus, if we want to show the existence of an H-closure for an H-ordered field (F, T), we have to show that there exists a maximal T-extension $(K, \dot{K}^2)$ for an H-ordered field, which is itself H-ordered, i.e. for which $\mathcal{G}_K \cong H$.
The following proposition indicates that even very simple preorderings may have a surprising behaviour in the context of a T- or H-extension. The proof of this proposition is no less interesting than the proposition itself, as it relies upon visual geometrical arguments involving topology of the plane.

Proposition 4.4. *Let $F = \mathbb{R}(X, Y)$ and let T be the set of nonzero sums of squares in F. If H is an essential subgroup of $\mathcal{G}_F$ such that $T = P_H$, then the H-ordered field (F, T) does not admit an H-closure.*

Proof. Suppose that we are in the situation described in our proposition. Then $H \neq \{1\}$ and by Proposition 3.1 the group H does not admit a C_4 quotient. Thus again by Proposition 3.1, if $(\dot{K}, \dot{K}^2)$ is an H-closure of (F, T), then $\dot{K}^2$ is a preordering in K. Choose $s \in T \setminus \dot{F}^2$, fix an embedding of $L = F(\sqrt{s}) = F[Z]/(Z^2 - s)$ in K and set $P = L \cap \dot{K}^2$. The intermediate extension (L, P) between (F, T) and $(K, \dot{K}^2)$ is a T-extension of (F, T) and $P = L \cap \dot{F}^2$ is a preordering of L.
Call z the class of Z in L. For every element $h \in \dot{L}$ there is a $g \in \dot{F}$ such that $gh \in P$. In particular, there is $f \in \dot{F}$ such that $zf \in P$. Call $\hat{P}$ the set of orderings of L that contain P, and denote by N the norm of L down to F. The embedding $F \longrightarrow L$ induces a map $\pi \colon X(L) \longrightarrow X(F)$ between the corresponding spaces of orderings, defined by $\alpha \mapsto \alpha \cap F$.
Let us show first that π induces an injection from $\hat{P}$ to $X(F)$. Let α_1, α_2 be two orderings of L containing P such that $\alpha = \alpha_1 \cap F = \alpha_2 \cap F$. Then the element $f \in F$ introduced above has a given sign $\epsilon = \pm 1$ at α, and thus has this same sign at α_1 and α_2. Since $zf \in P \subset \alpha_1 \cap \alpha_2$, z also has the same sign

at α_1 and α_2. But this cannot be, since the product of these signs is the sign of $N(z) = -s$ at α, which is negative.

Now, π also induces a surjection from $\hat{P}$ onto $X(F)$, and this is a bit deeper. Briefly, it goes as follows. Suppose α is an ordering of F such that none of the extensions α_1, α_2 to L contain P. Then we can find $u \in L$ such that $u \in P$ and $-u \in \alpha_1 \cap \alpha_2$. Denote by $D_E(w_1, \dots, w_n)$ the set of orderings of a field E containing the given elements $w_1, \dots, w_n \in E$. It is an open set for the Harrison topology on $X(E)$. Considering α_1, α_2 as points in $X(L)$ and $\hat{P}$ as a subset of $X(L)$, we may write $\alpha_1, \alpha_2 \in D_L(-u)$ and $D_L(-u) \cap \hat{P} = \emptyset$. In other words, $D_L(-u)$ separate $\{\alpha_1, \alpha_2\}$ from $\hat{P}$. Now, one may check easily that there exists an open nonempty set V in $X(F)$ such that $\pi^{-1}(V) \subset D_L(-u)$. Due to the fact that F is the function field of an algebraic variety over a real closed field, we know that every open set of $X(F)$, and in particular V contains a nonempty set $D_F(v)$ for some $v \in F$. Since $D_L(v) \cap \hat{P} = \emptyset$, $-v$ must be in any ordering containing P, and thus must be in P. Hence $-V \in T$, and V, are in any ordering of F. Since $D_F(v) \neq \emptyset$, this is a contradiction which proves the surjectivity of π on $\hat{P}$.

Since π is surjective on $\hat{P}$, we have $\pi(D_L(w) \cap \hat{P}) = D_F(w)$ for $w \in F$, and since $zf \in P$, $\pi(D_L(wz) \cap \hat{P}) = D_F(wf)$. Coming back to $h = a + bz \in L$ with $a, b \in F$, it is known (and easy to see) that $D_L(h) = D_L(N(h), a) \cup D_L(-N(h), bz)$. Since π is injective on $\hat{P}$, it preserves intersection (and of course unions) and thus $\pi(D_L(h) \cap \hat{P}) = D_F(N(h), a) \cup D_F(-N(h), bf)$. On the other hand, for $g \in F$ such that $gh \in P$, we have $\pi(D_L(h) \cap \hat{P}) = D_F(g)$. What we have proved so far is that for any $h = a + bz \in L$, $D_F(N(h), a) \cup D_F(-N(h), bf)$ is a "principal" set $D_F(g)$ in $X(F)$.

Let us show that this is impossible in general. Take $s = 1 + X^2$ and $h = Y + c + bz \in L$ with $c, b \in \mathbb{R}, b > 0$. Assume that the corresponding set $D_F(N(h), Y + c) \cup D_F(-N(h), f)$ is the principal set $D_F(g)$ for a given square-free polynomial $g \in F$. (This can always be achieved.) Note that $N(h) = 0$ is the equation $(Y + c)^2 = b^2(1 + X^2)$ of a hyperbola $\mathcal{H}$ in $\mathbb{R}^2$. Set $A := \{(X, Y) \in \mathbb{R}^2 \mid N(h) > 0, \quad Y + c > 0\}$ (respectively $B := \{(X, Y) \in \mathbb{R}^2 \mid N(h) > 0, \quad Y + c < 0\}$) the open region of the plane above (respectively below) the upper (respectively lower) branch of $\mathcal{H}$. Denote by $\tilde{A}, \tilde{B}$ the subsets defined in $X(F)$ by the same inequalities as for A, B. By assumption, we know that $g > 0$ on $\tilde{A} \cap X(F) = D_F(N(h), Y + c)$ and $g < 0$ on $\tilde{B} \cap X(F) = D_F(N(h), -(Y + c))$. This implies that $g \geq 0$ on A and $g \leq 0$ on B (see [BCR], §7.6) and that A and B are separated by a branch (i.e. a 1-dimensional irreducible connected component) of $g = 0$. Moreover, no branch of $g = 0$ can go inside $A \cup B$, or else g would change sign on A or B. (This is due to the fact that g is square free, and thus every branch is a sign-changing branch.) Set $C := \mathbb{R}^2 \setminus A \cup B$. Then $\tilde{C} \cap X(F) = D_F(-N(h))$. Since $D_F(g, -N(h)) = D_F(bf, -N(h)) = D_F(f, -N(h))$, we know that f and g have the same sign on C, up to a 0-dimensional set. Thus $f = 0$ must also have a sign-changing branch contained in C, and since f may be chosen square free, any branch of $f = 0$ having a

nonempty intersection with the interior of C must be contained in C.
Suppose this is true at the same time for $h = h_1 = Y+z$ and $h = h_2 = Y+4+2z$. Then

(1) no branch of $f = 0$ is allowed to cross a branch of the hyperbolas $\mathcal{H}_i$, $i = 1, 2$, and
(2) there is a branch of $f = 0$ splitting the plane into two connected components, each of them containing one branch of $\mathcal{H}_i$.

As the upper branch of $\mathcal{H}_2$ crosses the two branches of $\mathcal{H}_1$, this is impossible. This provides a contradiction to the existence of an H-closure for T, finishing the proof of Proposition 4.4. $\square$

Associated to the group $\dot{F}/T$ of the preceding proposition is the "abstract Witt ring" of T-forms (see [Ma]), which is actually the reduced Witt ring $W_{red}(F)$. (See also [L2, Chapter 1] for the definition of $W_{red}(F)$.) From Proposition 4.4 we can show there is no extension $F \longrightarrow K$ with $W_{red}(K) \cong W(K)$ such that the induced homomorphism $W_{red}(F) \longrightarrow W_{red}(K)$ is an isomorphism. Note that $W_{red}(F)$ might actually be isomorphic to $W(K)$ for some field K not related to F, as shown in Example 8.14. This is why we can view this result as a weak version of the "unrealizability" of $W_{red}(F)$ as a "true" Witt ring. (See [Ma], as well as [Cr2].)
Actually T. Craven kindly called our attention to [Cr2, Theorem 5.5], which can be applied to obtain the following more general result.

Proposition 4.5 (Craven). *Let $F = L(X)$ where L is a formally real field, which is not a pythagorean field. Then for each pythagorean field extension K/F, the natural homomorphism $W_{red}(F) \longrightarrow W_{red}(K) = W(K)$ induced by the inclusion map $F \longrightarrow K$ is not an isomorphism.*

Proof. Assume that K is a pythagorean field extension of $F = L(X)$, where L is a formally real field which is not pythagorean, and suppose that the field extension $F \longrightarrow K$ induces an isomorphism $W_{red}(F) \longrightarrow W_{red}(K)$.
Because L is not a pythagorean field, there exists an element $l = l_1^2 + l_2^2, l_1, l_2 \in L$ such that $l \notin \dot{L}^2$. Because K is a pythagorean field, there exists an element $k \in \dot{K}$ such that $k^2 = l$. Hence the polynomial $f(X) = X^2 - l$ has a root in K. Then from [Cr2, Theorem 5.5(b)], we see that $f(X)$ has exactly one root in every real closure of L. Of course this is not true, as each real closure of L must contain both roots of $f(X)$. Hence we have arrived at a contradiction, completing the proof. $\square$

Of course we may take $L = \mathbb{R}(Y)$ and get the result for $\mathbb{R}(X,Y)$.
In the other direction we present a case below, where (F,T) admits a maximal preordered T-extension $(\dot{K}, \dot{K}^2)$. We recall that a preordering T in F is SAP (Strong Approximation Property) if and only if for each pair of elements $a, b \in \dot{F}$ there exists an element $c \in \dot{F}$ such that $D_F(a,b) \cap \hat{T} = D_F(c) \cap \hat{T}$. (Here as above, $\hat{T}$ is the set of all orderings $\alpha \in X(F)$ such that $T \subset \alpha$.) If T is SAP and R is a preordering of F containing T, then R is SAP as well. (See

[L2, Theorem 17.12 and Corollary 16.8].) Note that this definition implies that every finite union of basic open sets in $X(F)$ is a "principal" set $D_F(c)$.

Proposition 4.6. *Let F be a formally real field, and let T be a SAP preordering in F. Then (F,T) admits a maximal preordered T-extension $(K,\dot{K}^2)$, which is again SAP.*

Proof. Let F be a formally real field and let T be a SAP preordering in F. Using Zorn's lemma we see that there exists a T-extension (L,S) of (F,T) which is maximal among the preordered T-extensions. We claim that S is a SAP preordering in L. In order to show this, pick any elements $a,b \in \dot{L}$. Because (L,S) is a T-extension of (F,T) we see that there exist elements $a',b' \in \dot{F}$ such that $aa', bb' \in S$. Because T is SAP there exists an element $c \in \dot{F}$ such that $D_F(a',b') \cap \hat{T} = D_F(c) \cap \hat{T}$. Let $\alpha \in D_L(c) \cap \hat{S}$ and $\beta = \alpha \cap F$, then $\beta \in \hat{T}$ and $c \in \beta$. Thus $a',b' \in \beta \subset \alpha$ and $\alpha \in D_L(a',b') \cap \hat{S}$. Conversely, if $\alpha \in D_L(a',b') \cap \hat{S}$, then $\beta \in D_F(a',b') \cap \hat{T} = D_F(c) \cap \hat{T}$ and $\alpha \in D_L(c) \cap \hat{S}$. Since it is clear that $D_L(a,b) \cap \hat{S} = D_L(a',b') \cap \hat{S}$, we have shown that $D_L(a,b) \cap \hat{S} = D_L(c) \cap \hat{S}$ and that S is SAP.

Now, we just have to prove that $S = \dot{L}^2$. Suppose it is not true. Then there exists an element $s \in S \setminus \dot{L}^2$ and we can set $E = L(\sqrt{s}) = L[Z]/(Z^2 - s)$. Let α be an ordering of L containing S. We know there are two orderings α_1, α_2 on E extending α and giving opposite signs to z. Denote by α_1 the ordering containing z.

Define P as $\bigcap_{S \subseteq \alpha} \alpha_1$, then $P \cap L = \bigcap_{S \subseteq \alpha} (\alpha_1 \cap L) = \bigcap_{S \subseteq \alpha} (\alpha) = S$ and we have proved that $\dot{L}/S \longrightarrow \dot{E}/P$ is one-to-one. Take $h = a + bz \in E$ with $a,b \in L$. Because S is SAP we know there exists $g \in L$ such that $[D_L(N(h),a) \cup D_L(-N(h),b)] \cap \hat{S} = D_L(g) \cap \hat{S}$.

Let us show $gh \in P$. Suppose $S \subset \alpha$, then $g \in \alpha_1 \iff g \in \alpha \iff [N(f), a \in \alpha]$ or $[-N(f), b \in \alpha] \iff h \in \alpha_1$. Thus $gh \in \bigcap_{S \subseteq \alpha} \alpha_1 = P$ and $\dot{L}/S \longrightarrow \dot{E}/P$ is onto.

But then (E,P) is a strict preordered T-extension of (L,S), contradicting the maximality of (L,S). This proves $S = \dot{L}^2$ and finishes the proof of the proposition. □

According to [ELP], we say that a field F satisfies the property H_4 if each totally indefinite quadratic form of dimension four represents zero in F. When a formally real field F satisfies H_4, the nonzero sums of squares in F form a SAP preordering. By [ELP, Theorem F], every field F such that $F(\sqrt{-1})$ is C_1 (i. e. "quasi-algebraically closed") satisfies H_4.

Therefore the preceding proposition will apply in particular to any formally real field of transcendence degree 1 over a real closed field. But in this case one can even prove the following addition to Proposition 4.6.

Proposition 4.7. *Let F be a formally real field which satisfies H_4, and let T be the set of nonzero sums of squares in F. Let $T = P_H$ for some essential subgroup H of $\mathcal{G}_F$. Then (F,T) admits an H-closure $(L,\dot{L}^2)$.*

Because we shall not use this result in this paper, and because our proof is quite long, we shall omit its details.

§5. Cyclic subgroups of W-groups

In this section we consider the subgroups H of $\mathcal{G}_F$ which are the easiest to understand in terms of their associated H-orderings, namely the two cyclic groups C_2 and C_4. As mentioned earlier, C_2 in many ways is the motivating example for this entire theory, and we cite here the results previously given in [MiSp1] for this group, as a means of illustrating the results we are attempting to generalize in this paper. As any single element of $\mathcal{G}_F$ necessarily generates a cyclic subgroup of order 2 or 4, those which generate subgroups of order 4 are precisely those not associated with usual orderings on the field F. These are the so-called half-orders of F, as investigated in [K1]; this concept was first introduced by Sperner [S] in 1949, in a geometrical context.

Definition 5.1. A *nonsimple involution* of $\mathcal{G}_F$ is an element $\sigma \in \mathcal{G}_F$ such that $\sigma^2 = 1$ and $\sigma \notin \Phi(\mathcal{G}_F)$. In other words, a nonsimple involution is an element of $\mathcal{G}_F$ which generates an essential subgroup of order 2.

Theorem 5.2. [MiSp1] *The field F is formally real if and only if $\mathcal{G}_F$ contains a nonsimple involution. There is a one-one correspondence between orderings on F and nontrivial cosets of $\Phi(\mathcal{G}_F)$ which have an involution as a coset representative.*

We have the well-known characterization of those subgroups of $\dot{F}$ that are orderings, which we include here for the sake of completeness.

Proposition 5.3. *A subgroup S of $\dot{F}$ containing $\dot{F}^2$ is a C_2-ordering of F if and only if the following conditions hold.*

(1) $|\dot{F}/S| = 2$ *and*
(2) $1 + s \in S\ \forall s \in S$.

We can now characterize those subgroups S of $\dot{F}$ which are C_4-orderings. They are precisely those subgroups of index 2 which fail to be orderings. We also see that C_4-ordered fields always admit a closure.

Proposition 5.4. *A subgroup S of $\dot{F}$ containing $\dot{F}^2$ is a C_4-ordering of F if and only if the following conditions hold.*

(1) $|\dot{F}/S| = 2$ *and*
(2) $\exists s \in S$ *such that* $1 + s \notin S$.

Proof. We know S is a C_4-ordering of F if and only if there exists $\sigma \in \mathcal{G}_F$ such that $S = \{a \in \dot{F} | \sqrt{a}^{\sigma} = \sqrt{a}\}$ where $\sigma^2 \neq 1$. Now any subgroup of index 2 in $\dot{F}$ is of the form $\{a \in \dot{F} | \sqrt{a}^{\sigma} = \sqrt{a}\}$ for some $\sigma \in \mathcal{G}_F$, so we need only guarantee that S is not an ordering, which condition (2) does. $\square$

REMARK 5.5.

(1) Note that it is easy to see that condition (2) above can be replaced by (2') $S+S=\dot{F}$. (Recall that here we use our definition of the sum $S+S$ as described in the beginning of §3. If instead we set the sum $S\oplus S$ as $\{s_1+s_2|s_1,s_2\in S\}$ then we can replace (2) by the condition $\dot{F}\subset S\oplus S$ provided that F contains more than five elements. (See [K1, Remark after Def. 1.1].)

(2) There are actually two kinds of C_4-orderings, distinguished by whether or not they contain -1. If S is a C_4-ordering such that $-1\in S$, we say that S has level 1. The prototype is given by $\dot{\mathbb{F}}_p^2$ when $p\equiv 1 \bmod 4$. If $-1\notin S$, then necessarily $-1\in S+S$, and we say that S has level 2. The model is $\dot{\mathbb{F}}_p^2$ when $p\equiv -1 \bmod 4$. It is clear that every C_4-extension preserves the level.

PROPOSITION 5.6. *Let $(K,\dot{K}^2)$ be a maximal T-extension of a C_4-ordered field (F,T). Then*

(1) *K is characterized by the condition of being maximal in $F(2)$ among fields $L\supseteq F$ such that $\sqrt{a}\notin L$ $\forall a\in\dot{F}\backslash T$.*

(2) *$\mathcal{G}_K\cong C_4$.*

(3) *$\mathrm{Gal}(K(2)/K)\cong\mathbb{Z}_2$, the group of 2-adic integers.*

In particular, every maximal T-extension of a C_4-ordered field (F,T) is a C_4-closure, and thus C_4-closures always exist.

Proof. Let $(K,\dot{K}^2)$ be a maximal T-extension of the C_4-ordered field (F,T). Since $\dot{K}^2\cap F=T$, we see that for any $a\in\dot{F}\backslash T$, we have $\sqrt{a}\notin K$, while for any $a\in T$, we have $\sqrt{a}\in K$. Now if $L\supsetneq K$ in $F(2)$, then $L\supseteq K(\sqrt{a})$ for some $a\in\dot{K}\backslash\dot{K}^2$. Since the cosets of $\dot{K}^2$ in $\dot{K}$ correspond naturally to the cosets of T in $\dot{F}$, we see that L contains $\sqrt{a'}$ for some $a'\in\dot{F}\setminus T$, and thus K is maximal among such extensions of F in $F(2)$. Conversely, suppose K is maximal in $F(2)$ among fields $L\supseteq F$ such that $\sqrt{a}\notin L$ $\forall a\in F\backslash T$. Then we see that $\dot{K}^2\cap F=T$. We need to see that $|\dot{K}/\dot{K}^2|=2$. Suppose it is not true. Fix $a\in\dot{F}\backslash T$, so that $a\notin\dot{K}^2$, and suppose there exists some $b\in\dot{K}$ such that a,b are linearly independent in $\dot{K}/\dot{K}^2$. Then certainly $b\notin aT$, and setting $L=K(\sqrt{b})$ contradicts the maximality of K. Thus we have that $(K,\dot{K}^2)$ is a maximal T-extension for (F,T), and this proves (1).
Now observe that $\mathcal{G}_K$ is generated by one generator, since $|\dot{K}/\dot{K}^2|=2$, so $\mathcal{G}_K\cong C_2$ or C_4. It cannot be C_2, or else T would be an ordering on F. Thus $\mathcal{G}_K\cong C_4$. Finally, $\mathrm{Gal}(K(2)/K)$ is cyclic and cannot be finite, since it is not C_2 (see [Be]). Thus $\mathrm{Gal}(K(2)/K)\cong\mathbb{Z}_2$. □

§6. SUBGROUPS OF W-GROUPS GENERATED BY TWO ELEMENTS

As we saw in Theorem 2.5, a group generated by two elements appearing as a subgroup of $\mathcal{G}_F$ may only be one in the list $C_2*C_4, C_4*C_4, C_2*C_2, C_4\times C_4, C_4\rtimes C_4$. The last two are particular cases of the groups studied in § 7 and

§ 8, and we will focus in this section on the first three. The third one is better known as the dihedral group D.
We will give an algebraic characterization for the orderings associated with these groups and show that it is always possible to make closures. Portions of the proofs rely on the characterizations of $C_4 \times C_4$- and $C_4 \rtimes C_4$-orderings obtained in § 7; but since the results in § 7 do not rely on those in § 6, we freely use these results where needed.

Lemma 6.1. *Let T be a subgroup of $\dot{F}$ such that $\dot{F}^2 \subseteq T$ and $|\dot{F}/T| = 4$. If $-1 \notin T$, then F is $(T \cup -T)$-rigid.*

Proof. Let $\dot{F}/T = \{1, -1, a, -a\}$. Then $(T \cup -T) + a(T \cup -T) \subseteq (T \cup -T) \cup a(T \cup -T) = \dot{F}$. $\square$

Proposition 6.2. *A subgroup T of $\dot{F}$ is a $C_2 * C_4$-ordering if and only if $\dot{F}^2 \subseteq T$, $|\dot{F}/T| = 4$, and the following two conditions hold.*

(1) *$T + T \neq T$, and*
(2) *$-1 \notin \sum T$, where $\sum T$ denotes the set of all finite sums of elements of T.*

Proof. The conditions $\dot{F}^2 \subseteq T$ and $|\dot{F}/T| = 4$ are necessary and sufficient for T to be a G-ordering for some essential subgroup $G \subseteq \mathcal{G}_F$ generated by two elements σ, τ, independent mod $\Phi(\mathcal{G}_F)$. We next show the necessity of conditions (1) and (2). Let $G \cong C_2 * C_4$ be a subgroup of $\mathcal{G}_F$, where $T = P_G$. We assume G is generated by two noncommuting (hence independent mod $\Phi(\mathcal{G}_F)$) elements σ, τ such that $\sigma^2 = 1, \tau^4 = 1$. If $T + T = T$, then by Proposition 6.14, T would be a D-ordering (this is independent of previous results). Since it is not, we see that (1) holds. Also $-T \nsubseteq \sum T$, since $\sum T \subseteq P_\sigma$, which is an ordering because σ is an involution. Thus P_σ cannot contain $-T$ and condition (2) holds.
We now show the sufficiency of the conditions. Since T is a G-ordering for some essential subgroup with two generators, it must be isomorphic to one of the five groups listed in Theorem 2.5. Since $-1 \notin T$ by (2), it cannot be $C_4 \times C_4$ by Proposition 7.2 in the next section. Also (1) shows that G cannot be isomorphic to $D \cong C_2 * C_2$ by Proposition 6.14, and (2) shows that G cannot be isomorphic to $C_4 \rtimes C_4$ by Proposition 7.6. Finally, from (1) and (2) we can see that $\sum T$ is an ordering on F, since it is clearly a proper subgroup of $\dot{F}$, which properly contains T, so must be of index 2 in $\dot{F}$; it does not contain -1, and it is closed under addition. Then $\sum T = T \cup aT$ for some $a \notin T$, and G is generated by elements σ, τ where the intersection of the fixed field of σ with $F^{(2)}$ is $K(\sqrt{a})$, and the intersection of the fixed field of τ with $F^{(2)}$ is $K(\sqrt{-1})$. Then $P_\sigma = \sum T$ is an ordering, so σ is an involution. This shows G cannot be isomorphic to $C_4 * C_4$. Thus the only remaining possibility is $G \cong C_2 * C_4$. $\square$

Proposition 6.3. *A subgroup T of $\dot{F}$ is a $C_4 * C_4$-ordering if and only if $\dot{F}^2 \subseteq T$, $|\dot{F}/T| = 4$, and one of the following two conditions hold.*

(1) *$-1 \in T$ and F is not T-rigid, or*

(2) $-1 \notin T$, $-1 \in \sum T$, *but* $T + T \neq T \cup -T$.

Proof. If $-1 \in T$, the only possible subgroups H of $\mathcal{G}_F$ with two generators for which T can be an H-ordering are $C_4 \times C_4$ and $C_4 * C_4$. The other three are eliminated by Propositions 6.14, 6.2, and 7.6. Also, if $-1 \in T$, then F is T-rigid if and only if T is a $C_4 \times C_4$-ordering by Proposition 7.2. This leaves $C_4 * C_4$ as the only possibility.
If $-1 \notin T$, there are three possibilities to consider: $-1 \notin \sum T$, $T+T = T \cup -T$, or $-1 \in \sum T$ but $T+T \neq T \cup -T$. The first case occurs if and only if T is either a D-ordering (by Proposition 6.14) or a $C_2 * C_4$-ordering (by Proposition 6.2). The second case occurs if and only if T is a $C_4 \rtimes C_4$-ordering by Proposition 7.6 and Lemma 6.1. Thus, the third case must occur if and only if T is a $C_4 * C_4$-ordering as claimed. □

The following example constructs a $C_4 * C_4$-ordering of $\mathbb{Q}_2$. It is illustrative, in that it shows how even in a relatively "small" setting, the additive structure of T can behave quite differently from the additive structure of $\dot{F}(\sqrt{T})^2$. In particular, it shows that $\langle 1, 1\rangle$ may represent elements in $F(\sqrt{T})$ which are not in $T + T$. In this example, $T + T$ is not multiplicatively closed, but of course the form $\langle 1, 1\rangle$, being a Pfister form, is multiplicative in $F(\sqrt{T})$.

EXAMPLE 6.4. In $F = \mathbb{Q}_2$ consider the subgroup $T = \dot{F}^2 \cup 5\dot{F}^2$ of $\dot{F}$. Using the notation for $\mathcal{G}_2$ as in Example 2.9, we see that the corresponding subgroup of $\mathcal{G}_2$ is $H = \langle \sigma, \tau\rangle \cong C_4 * C_4$. This is a W-group associated with the Witt ring $\mathbb{Z}/4\mathbb{Z} \times_M \mathbb{Z}/4\mathbb{Z}$, where the product "$\times_M$" is taken in the category of Witt rings (see [Ma] and [MiSm2]). The fixed field of H is $K = \mathbb{Q}_2(\sqrt{5})$. The form $\langle 1, 1\rangle$ represents -1 over K, and this can be shown as follows. It is well known and easy to show that for any quadratic field extension $F \longrightarrow K = F(\sqrt{a})$, one has $(K^2 + K^2) \cap \dot{F} = (F^2 + F^2)(F^2 + aF^2)$. If $F = \mathbb{Q}_2$ and $a = 5$, we have $30 = 5 \times 6 \in (K^2 + aK^2)$ and $2 \in (K^2 + K^2)$. Then $15 \in K^2 + K^2$, and since 15 is congruent to -1 mod 16: it is a negative square in $\mathbb{Q}_2$. This shows that $-1 \in K^2 + K^2$.
However, when one considers which elements of $\dot{F}/\dot{F}^2$ are in $T + T$, one finds only the six classes represented by $1, 2, 5, 10, -2, -10$. In particular, $-1 \notin T+T$, and $T + T$ is not multiplicatively closed (so forms mod T-equivalence do not behave as quadratic forms over a field behave). Nonetheless, it is easy to see that $-1 \in T + T + T$, so that $T + T \neq T \cup -T$, but $-1 \in \sum T$, consistent with the proposition above.

In §9 we introduce natural conditions for a subgroup H of $\mathcal{G}_F$ in order to keep track of the additive properties of $\dot{F}/T$ under 2-extensions. We shall see in §9 that the group $H \subset \mathcal{G}_F$ above does not possess one of the key properties we require.

THEOREM 6.5. *A* $(C_2 * C_4)$*-ordered field* (F, T) *admits a closure.*

Proof. Let $\mathcal{S}$ be the set of extensions (L, S) of (F, T) inside $F(2)$ satisfying the additional condition that $-1 \notin \sum S$. As in the proof of Proposition 4.1, we see that $\mathcal{S}$ has a maximal element (K, T_0) with $\dot{K}/T_0 \cong \dot{F}/T, T = T_0 \cap F$, and $-1 \notin \sum T_0$. Then (K, T_0) is a $(C_2 * C_4)$-ordered field. To see this we need only show that conditions (1) and (2) of Proposition 6.2 hold, and condition (2) is given by construction of (K, T_0). Condition (1) holds since if $T_0 + T_0 = T_0$, then $T + T \subseteq (T_0 + T_0) \cap F = T_0 \cap F = T$, contradicting the fact that T is a $C_2 * C_4$-ordering on F.

To conclude, we must show $T_0 = \dot{K}^2$. Notice $\sum T_0$ is an ordering on K, so K is formally real. We may write $\dot{K}/T_0 = \{\pm T_0, \pm aT_0\}$, where $a \in T + T$. If $T_0 \neq \dot{K}^2$, we can choose $c \in T - \dot{K}^2$, and consider $L = K(\sqrt{c})$. Since $-c \notin \sum T_0$, $\sum T_0$ extends to an ordering S_0 on L. Then $S_0 \cup -S_0 = \dot{L}$ and $a \in S_0$. Let S be a subgroup of S_0 containing T_0 and maximal with respect to excluding a. Then $\dot{L}/S = \{\pm S, \pm aS\} \cong \dot{K}/T_0 \cong \dot{F}/T$. Also $S \cap K \supseteq T_0$ by construction, and if there exists $b \in S \cap K, b \notin T_0$, then $b \in aT_0 \cup -T_0 \cup -aT_0$, which implies either $a \in S$ or $-1 \in S$, which leads to a contradiction in either case. Thus $S \cap K = T_0$, and (L, S) is an extension contradicting the maximality of (K, T_0). We conclude $T_0 = \dot{K}^2$. $\square$

Theorem 6.6. *A $(C_4 * C_4)$-ordered field (F, T) admits a $(C_4 * C_4)$-closure $(K, \dot{K}^2)$.*

Proof. Let $(K, \dot{K}^2)$ be a maximal T-extension for (F, T). First assume $-1 \in T$. We must show K is not a rigid field. Let $\{1, a, b, ab\}$ be a set of representatives for $\dot{F}/T$ that lifts to a set of representatives for $\dot{K}/\dot{K}^2$. Since F is not T-rigid, we may, without loss of generality, assume $b \in T + aT$. Then $T + aT \subseteq \dot{K}^2 + a\dot{K}^2$, but $b \notin \dot{K}^2 \cup a\dot{K}^2$, so K is not rigid, and $\dot{K}^2$ is a $(C_4 * C_4)$-ordering on K.

Now assume $-1 \notin T = F \cap \dot{K}^2$. Then $-1 \notin \dot{K}^2$, and $-1 \in \sum T \subseteq \sum \dot{K}^2$. Letting $\{1, -1, a, -a\}$ be a set of representatives for $\dot{F}/T$, this again lifts to a set of representatives for $\dot{K}/\dot{K}^2$. Since $T + T \neq T \cup -T$, but clearly also $T + T \neq T$, we may assume $a \in T + T$, so $a \in \dot{K}^2 + \dot{K}^2$ as well. This shows $\dot{K}^2$ is a $(C_4 * C_4)$-ordering on K. $\square$

Remark 6.7. We have defined in Definition 3.2 the level of an H-ordering. It is then easy to see that the level of a $(C_4 * C_4)$-ordering T is at most 4. The level of the closure K (which is the "usual" level) is less than or equal to the level of T. The level of a $(C_4 * C_4)$-closure is either 1 or 2, as any field of finite level with at most four square classes has level at most 2. The level of T is 1 if and only if the level of K is 1, but in the other cases the level may actually decrease: Example 6.4 shows that T has level 3 and that its closure has level 2.

Now we turn our attention to D-orderings. We showed in § 2 that $C_2 \times C_2$ cannot be an essential subgroup of $\mathcal{G}_F$, so if H is an essential subgroup of $\mathcal{G}_F$ generated by two elements of order 2, necessarily $H \cong D$. Recall that according to [Br], a 2-*element fan* in F is a set of two distinct orderings P_1, P_2 on F, and it can be identified with the preordering $T = P_1 \cap P_2$.

Lemma 6.8. *The dihedral group D is a subgroup of $\mathcal{G}_F$ if and only if there is a 2-element fan in F. In this case, $T \subseteq \dot{F}$ is a D-ordering if and only if T is a 2-element fan in F.*

Proof. Let $H = \langle \sigma, \tau | \sigma^2 = \tau^2 = [\sigma, \tau]^2 = 1 \rangle \cong D$ be a subgroup of $\mathcal{G}_F$. Then P_σ and P_τ are positive cones of two distinct orderings on F, and $P_H = P_\sigma \cap P_\tau$. Conversely, if P_1, P_2 are positive cones corresponding to distinct orderings on F, then there exist nontrivial involutions $\sigma, \tau \in \mathcal{G}_F$, in distinct cosets of $\Phi(\mathcal{G}_F)$, such that $P_1 = P_\sigma$ and $P_2 = P_\tau$. Then $H = \langle \sigma, \tau \rangle$ is an essential subgroup of $\mathcal{G}_F$, and $H \cong D$. $\square$

In [BEK], a field F with two orderings P_1, P_2 is defined to be maximal with respect to P_1, P_2 if for any algebraic extension K of F, at least one of the two orderings cannot be extended to K. Since we prefer to work inside $F(2)$, we modify this as follows.

Definition 6.9. A field F with two orderings P_1, P_2 is *maximal with respect to* P_1, P_2 if for any 2-extension K of F, at least one of the orderings does not extend to K.

Proposition 6.10. *(F, P_1, P_2) is maximal if and only if (F, T_F) is a D-ordered field, where $T_F = P_1 \cap P_2$, and there exists no proper D-ordered extension field $(L, T_L) \subseteq F(2)$ with $T_L \cap F = T_F$.*

Proof. Suppose that the field (F, P_1, P_2) is maximal. Let σ_1, σ_2 be involutions in $\mathcal{G}_F$ such that $P_i = \{a \in \dot{F} | \sqrt{a}^{\sigma_i} = \sqrt{a}\}, i = 1, 2$. Then the subgroup $\langle \sigma_1, \sigma_2 \rangle \subseteq \mathcal{G}_F$ is isomorphic to D, as we have seen, and (F, T_F) is a D-ordered field as claimed.
Now suppose that L is a D-ordered field containing F inside $F(2)$, such that $T_L \cap F = T_F$. Then $\mathcal{G}_L$ contains a subgroup isomorphic to D, which we can take to be generated by two involutions τ_1, τ_2 such that $T_L = Q_1 \cap Q_2$, where $Q_i = \{a \in \dot{L} | \sqrt{a}^{\tau_i} = \sqrt{a}\}, i = 1, 2$ are distinct orderings of L. Now $Q_i \cap F \supseteq T_L \cap F = T_F$, so $Q_i \cap F$ is an ordering of F which contains $T_F, i = 1, 2$. Thus $\{Q_1 \cap F, Q_2 \cap F\} = \{P_1, P_2\}$. Then by maximality of (F, P_1, P_2), we see $L = F$.
Conversely, suppose that F is a D-ordered field contained in no proper D-ordered extension field as described. Then F has at least two distinct orderings P_1 and P_2 corresponding to the two involutions generating the subgroup D of $\mathcal{G}_F$, and since there is no proper D-ordered extension field, we see that it is not possible for both orderings to extend to any extension of F. Thus (F, P_1, P_2) is maximal, as claimed. $\square$

By Zorn's Lemma we immediately see the following.

Proposition 6.11. [BEK, Prop.3] *Given a field F with two orderings P_1, P_2, there always exists an algebraic extension $\tilde{F}$ of F which is maximal with respect to $\tilde{P}_1, \tilde{P}_2$, where $\tilde{P}_1, \tilde{P}_2$ are extensions of P_1, P_2 to $\tilde{F}$.*

Theorem 6.12. *A field (F, P_1, P_2) is maximal if and only if*

(1) *there exist exactly two orderings on F and*
(2) *F is pythagorean, i.e. any sum of squares is a square in F.*

Proof. [BEK] Suppose three different orderings P_1, P_2, P_3 are possible in F. Let $x \in \dot{F}$ be such that x is positive with respect to the first two orderings, and negative with respect to P_3. Then $\sqrt{x} \notin F$, so $F(\sqrt{x})$ is a proper algebraic extension of F, and since x is positive with respect to P_1 and P_2, they extend to $F(\sqrt{x})$, and (F, P_1, P_2) cannot be maximal. Similarly, if α, β are elements of F such that $\sqrt{\alpha^2 + \beta^2} \notin F$, then P_1, P_2 can be extended to the proper extension $F(\sqrt{\alpha^2 + \beta^2})$ of F, again contradicting maximality. Thus conditions (1) and (2) are necessary.
Conversely, one can show that any field F satisfying conditions (1) and (2) has $\dot{F}/\dot{F}^2 = \{1, -1, a, -a\}$ for some $a \in \dot{F}$. Now let F be such a field and let P_1, P_2 be the two unique orderings in F, so that a is positive with respect to P_1 and negative with respect to P_2. Suppose (F, P_1, P_2) were not maximal, and let $K = F(\sqrt{b})$ be a proper quadratic extension of F such that both P_1 and P_2 extend to K. Since K is an ordered proper extension of F, $b \neq 1, -1 \in \dot{F}/\dot{F}^2$, so $b = a$ or $-a$. Then either $\sqrt{a} \in K$ or $\sqrt{-a} \in K$, so that not both P_1 and P_2 extend to K. This is a contradiction. $\square$

Corollary 6.13. *The D-ordered field (F, T) is a maximal D-ordered field if and only if $\mathcal{G}_F \cong D$. Thus any D-ordered field admits a D-closure.*

Proof. By the preceding theorem, if F is maximal, it has exactly two orderings, so $\mathcal{G}_F$ has exactly two involutions which are independent mod $\Phi(\mathcal{G}_F)$. Also F is pythagorean, so by [MiSp1] $\mathcal{G}_F$ is generated by involutions. Thus $\mathcal{G}_F$ is generated by two elements of order 2, and since $\mathcal{G}_F$ is necessarily an essential subgroup of itself, we see that $\mathcal{G}_F \cong D$.
Conversely, if $\mathcal{G}_F \cong D$, then F is a D-ordered field, and since orderings on F correspond to independent involutions in $\mathcal{G}_F$, we see that F has precisely two distinct orderings. Also, since $\mathcal{G}_F$ is generated by these involutions, we see that F is pythagorean. Thus, by the preceding theorem, F is a maximal D-ordered field. Then we see that for any D-ordered field (L, P_H), a maximal D-ordered extension $(F, \dot{F}^2)$ containing (L, P_H) will be a closure for (L, P_H). $\square$

Proposition 6.14. *A subgroup S of $\dot{F}$ containing $\dot{F}^2$ is a D-ordering of F if and only if $|\dot{F}/S| = 4$ and $1 + s \in S$ whenever $s \in S$.*

Proof. All that is necessary for S to be a D-ordering of F is that it be a 2-element fan in F. In other words, S must be a preordering of index 4 in F. A subgroup S of $\dot{F}$ is such a preordering if and only if the conditions in the statement of the proposition are met. $\square$

§7. Classification of rigid orderings

This section will provide a full Galois-theoretic and algebraic characterization of all possible rigid orderings. We start with the following definition.

Definition 7.1. Let I be a possibly empty index set. We call G a $C(I)$-group if G is isomorphic to $(C_4)^I \times C_4$, an $S(I)$-group if G is isomorphic to $(C_4)^I \rtimes C_4$, and a $D(I)$-group if G is isomorphic to $(C_4)^I \rtimes C_2$, the semidirect product being defined with the nontrivial action of C_4 or C_2 on each inner factor in the last two cases, when I is nonempty. A G-ordering on F is called a $C(I)$-(respectively $S(I)$-, $D(I)$-) ordering if G is a $C(I)$- (respectively $S(I)$-, $D(I)$-) group. When $I = \emptyset$ the $C(I)$- and $S(I)$-orderings are the C_4-orderings, and the $D(I)$-orderings are the C_2-orderings, that is the usual orderings. Observe that $C(\emptyset)$- and $S(\emptyset)$-orderings both correspond to the same group C_4. The difference between them is that a $C(\emptyset)$-ordering has level 1, while an $S(\emptyset)$-ordering has level 2. (See Remark 5.5 for comparison.) When $|I| = 1$, we obtain the groups generated by two elements which are respectively $(C_4) \times C_4$, $(C_4) \rtimes C_4$ and D.

In this section we will characterize $C(I)$-orderings, $S(I)$-orderings and $D(I)$-orderings in terms of their algebraic properties as subgroups of $\dot{F}$. We will see in particular that they are all rigid, and that they constitute the whole class of rigid orderings. The group $\coprod_{i\in I} G_i$ will denote the direct sum of the groups $G_i, i \in I$ and in $I \cup \{x\}$ the letter x is added to denote the new index.

Proposition 7.2. *A subgroup T of $\dot{F}$ containing $\dot{F}^2$ is a $C(I)$-ordering if and only if the following three conditions hold.*

(1) $-1 \in T$,
(2) *F is T-rigid, and*
(3) $\dot{F}/T \cong \coprod_{i\in I\cup\{x\}} (C_2)_i$.

In other words, the $C(I)$-orderings are exactly the rigid orderings of level 1.

Proof. If $I = \emptyset$, the result follows from Proposition 5.4 and Remark 5.5, so we shall assume $I \neq \emptyset$. We begin by showing that the three conditions above are necessary. Let $G \cong C(I)$ and let T be a G-ordering. Suppose $-1 \notin T$. Let $\{\sigma_i, i \in I; \sigma_x\}$ generate G. Then $T = \cap_{i\in I\cup\{x\}} P_{\sigma_i}$ and $|\dot{F}/T| \geq 4$. Thus there are at least four classes mod T, which we can represent as $1, -1, a, -a$ for some $a \in \dot{F}$, and there exists a $D^{a,-a}$-extension L of F. Hence there exist elements $\sigma, \tau \in G$ such that $a \in P_\sigma \backslash P_\tau$ and $-a \in P_\tau \backslash P_\sigma$. It then follows that the restriction of $\sigma\tau$ to L has order 4, so that $\sigma|_L, \tau|_L$ generate $\mathrm{Gal}(L/F) \cong D$, and hence cannot commute. Yet $\sigma, \tau \in G$, which is an abelian group. This is a contradiction, so $-1 \in T$, and (1) holds.
Since $-1 \in T$, we have $T \cup -T = T$. Suppose we have a nonrigid element $c \in \dot{F}\backslash T$, so that we have $t_1, t_2 \in T$ with $t_1 + ct_2 \notin T \cup cT$. Then $b = 1 + ct_2/t_1 \notin T \cup cT$. Let $a = -ct_2/t_1 \notin T$. Then $a + b = 1$, so $(\frac{a,b}{F})$ splits. Since $b \notin T \cup cT = T \cup aT$, a and b are independent mod T and thus mod $\dot{F}^2$. Hence we have a $D^{a,b}$-extension L of F, and by the same argument as above, we find $\sigma, \tau \in G$ which do not commute, leading to a contradiction. Thus F is T-rigid and (2) holds. Finally, by Kummer theory we know that $\dot{F}/T$ is isomorphic to the dual $(G/\Phi(G))^* \cong \coprod_{i\in I\cup\{x\}} (C_2)_i$, giving (3).
We now show that the three conditions are sufficient for T to be a $C(I)$-ordering. By (3) we see that $T = \cap_{i\in I\cup\{x\}} P_i$ where P_i is the kernel of the

projection $\dot{F} \to \dot{F}/T \cong \coprod_{i \in I \cup \{x\}} (C_2)_i \to (C_2)_i$. Further, for each P_i we have a $\sigma_i \in \mathcal{G}_F$ such that $P_i = P_{\sigma_i}$. Let G be the closed subgroup of $\mathcal{G}_F$ generated by $\{\sigma_i | i \in I \cup \{x\}\}$. Then $G \subseteq \{\sigma | P_\sigma \supseteq T\}$ because every element of G must fix every $\sqrt{a}$ left fixed by the σ_i. So we also have $T = \cap_{\sigma \in G} P_\sigma$, and T is a G-ordering. It remains to show that G is a $C(I)$-group.
Since $-1 \in T \subseteq P_{\sigma_i}$, none of the P_{σ_i} can be usual orderings on F, so each σ_i must have exponent 4 in G. Since $-1 \in T$ and F is T-rigid, we see by Proposition 3.4 that G is abelian. Then G is a compact abelian group of exponent 4, and $(G/\Phi(G))^* \cong \coprod_{i \in I \cup \{x\}} (C_2)_i$ is a discrete group of exponent 2. Then $((G/\Phi(G))^*)^* \cong G/\Phi(G) \cong \prod_{i \in I \cup \{x\}} (C_2)_i$, and $G \cong \prod_{i \in I \cup \{x\}} (C_4)_i$, so G is a $C(I)$-group as claimed. $\square$

In order to characterize the subgroups of $\dot{F}$ which are $S(I)$-orderings, we will first prove three lemmas. Let G be an $S(I)$-group. It will be helpful to fix the following notation: write $G = G_1 \rtimes G_2$ where $G_1 \cong \prod_{i \in I} (C_4)_i$ and $G_2 \cong C_4$. Let τ be a generator of G_2 and $P_\tau = \{a \in \dot{F} | \sqrt{a}^\tau = \sqrt{a}\}$.

Lemma 7.3. *Let T be a G-ordering. Then T has index* 2 *in* P_{G_1}.

Proof. If $P_{G_1} \subseteq P_\tau$, we would have $T = P_{G_1} \cap P_\tau = P_{G_1} = P_G$. But by Kummer theory and the Burnside Basis Theorem, that would imply $G = G_1$. Thus $P_{G_1} \nsubseteq P_\tau$, $T \subsetneq P_{G_1}$, and $|P_{G_1}/T| \geq 2$. On the other hand, since $T = P_{G_1} \cap P_\tau$, we have $|P_{G_1}/T| \leq 2$, and so $|P_{G_1}/T| = 2$. $\square$

Lemma 7.4. *For any group homomorphism $\theta : G \to C_4 = \langle \sigma \rangle$, we have $\theta(G_1) \subseteq \langle \sigma^2 \rangle$.*

Proof. If $a \in G_1$, writing multiplicatively, we have

$$\theta(a^{-1}) = \theta(\tau a \tau^{-1}) = \theta(\tau)\theta(a)\theta(\tau)^{-1} = \theta(a),$$

so $\theta(a)^2 = 1$. $\square$

Lemma 7.5. *We have $T + T \subseteq P_{G_1}$.*

Proof. Let $a \in T + T, a \notin T$, and consider the following three cases.
Case 1: $a = x^2 + y^2$. Then there exists a C_4^a-extension L of F, and we have a map $\theta : G \to \mathrm{Gal}(L/F) \cong C_4$, and by Lemma 7.4 $\theta(G_1)$ has order at most 2. Thus $\theta(G_1)$ fixes $\sqrt{a}$ and $a \in P_{G_1}$.
Case 2: $a = x^2 + t, t \in T \backslash \dot{F}^2$. We have $a^2 = ax^2 + at$, and a, at are independent modulo $\dot{F}^2$. Thus there exists a $D^{a,at}$-extension L of F, and $\mathrm{Gal}(L/F(\sqrt{t})) \cong C_4$. Since $t \in T$, we have $\sqrt{t}^\sigma = \sqrt{t}$ for $\sigma \in G$, which means we have a homomorphism $\theta : G \to \mathrm{Gal}(L/F(\sqrt{t})) \cong C_4$. Again applying Lemma 7.4, $\theta(G_1)$ has order at most 2, so G_1 must fix $\sqrt{a}$ and $a \in P_{G_1}$.
Case 3: $a = s + t$, $s, t \in T \backslash \dot{F}^2$. We can write $as^{-1} = 1 + ts^{-1}$, and then we are in one of the previous two cases. Hence $as^{-1} \in P_{G_1}$, and it follows that $a \in P_{G_1}$. $\square$

Proposition 7.6. *A subgroup T of $\dot{F}$ containing $\dot{F}^2$ is an $S(I)$-ordering if and only if the following four conditions hold.*

(1) $-1 \notin T$,
(2) F *is* $(T \cup -T)$*-rigid,*
(3) $T + T = T \cup -T$, *and*
(4) $\dot{F}/T \cong \coprod_{i \in I \cup \{x\}} (C_2)_i$.

Proof. When $I = \emptyset$ the result follows from Proposition 5.4 and Remark 5.5. Thus we may assume that $I \neq \emptyset$. We begin by showing the conditions above are necessary. Condition (4) follows from Kummer theory. Condition (1) follows from Lemma 7.5 above, for if $-1 \in T$, we would have $\dot{F} \subseteq \dot{F}^2 - \dot{F}^2 \subseteq T - T = T + T \subseteq P_{G_1}$, but as $|I| \geq 1$, we cannot have P_{G_1} being all of $\dot{F}$.
To show the necessity of condition (3), first observe that $-1 \in P_{G_1}$, $-1 \notin T$, and $|P_{G_1}/T| = 2$, so $P_{G_1} = T \cup -T$, and thus $T + T \subseteq T \cup -T$. To show equality, we need to show that some element of $-T$ is in $T + T$. In this case, that amounts to showing that T is not additively closed. Suppose that T were additively closed. Then T would be a preordering, so contained in some ordering P_σ for some $\sigma \in \mathcal{G}_F$. Further, σ is an involution not contained in $\Phi(\mathcal{G}_F)$, and $\sigma \in G = G_1 \rtimes G_2$. In particular, σ is not a square in G, and $\sigma \neq \tau$. Thus $\sigma = \sigma_1\tau$ for some $\sigma_1 \in G_1$ and

$$\sigma^2 = \sigma_1\tau\sigma_1\tau = \sigma_1\tau\sigma_1\tau^{-1}\tau^2 = \sigma_1\sigma_1^{-1}\tau^2 = \tau^2 \neq 1.$$

Thus σ is not an involution, which is a contradiction, and so $-1 \in T + T$. Finally, since F is P_{G_1}-rigid and $T \cup -T = P_{G_1}$, we see that (2) holds.
Now we must show that conditions (1) - (4) are sufficient for T to be an $S(I)$-ordering. Letting $S = T \cup -T$, we see that S satisfies the condition for being a G_1-ordering, with $G_1 \cong \prod_{i \in I} (C_4)_i$, as given in Proposition 7.2. Let Q be a subgroup of index 2 in $\dot{F}$ such that $T = S \cap Q$, and let $\tau \in \mathcal{G}_F$ such that $Q = P_\tau$. Let G be the subgroup of $\mathcal{G}_F$ generated by G_1 and τ. We need to see that $G = G_1 \rtimes G_2$ where G_2 is the subgroup of $\mathcal{G}_F$ generated by τ. Specifically, we need to show that $G_1 \cap G_2 = \{1\}$ and that $[\sigma, \tau]\sigma^2 = 1 \; \forall \sigma \in G_1$.
Since G_1 fixes $\sqrt{-1}$ and τ does not, we cannot have τ or τ^{-1} in G_1. Suppose $\tau^2 \in G_1$. Then it has order 2 in G_1 and hence must be a square. Let $\sigma \in G_1$ such that $\sigma^2 = \tau^2$. Since $P_\sigma \neq P_\tau$, there exists $a \in P_\tau \backslash P_\sigma$, and neither a nor $-a$ can be a square, since neither is in P_σ. Since also $-1 \notin \dot{F}^2$, we have a $D^{a,-a}$-extension L of F, and $\sigma|_L$ has order 4 in $\mathrm{Gal}(L/F)$. However, since τ fixes $\sqrt{a}$, $\tau|_L \in \mathrm{Gal}(L/F(\sqrt{a})) \cong C_2 \times C_2$, and so $\sigma^2 \neq \tau^2$, contradicting the assumption. Thus $G_1 \cap G_2 = \{1\}$.
To prove $[\sigma, \tau]\sigma^2 = 1 \; \forall \sigma \in G_1$, it is sufficient to show that this condition holds for the restriction of σ, τ to each C_4- and D-extension of F. Suppose L is a C_4^a-extension of F. Then a is a sum of two squares, so $a \in T + T = T \cup -T = P_{G_1}$ and $[\sigma, \tau]\sigma^2|_L = \sigma^2|_L$. Since $\sigma \in G_1$, $\sigma \in \mathrm{Gal}(L/F(\sqrt{a}))$ and $\sigma^2|_L = 1$.
Now suppose L is a $D^{a,b}$-extension of F. We may assume $\sigma \notin Z(\mathrm{Gal}(L/F))$ (the centralizer), since otherwise clearly $[\sigma, \tau]\sigma^2|_L = 1$. Without loss of generality,

we may assume $\sqrt{a}^\sigma = -\sqrt{a}$. Then $a \notin T \cup -T$, and since $1 = ax^2 + by^2$, we have $b \in T - aT$, and by rigidity, $b \in T \cup -aT \cup -T \cup aT$. However, if b were in $-T$ or aT, then we would obtain $a \in T + T = T \cup -T$, a contradiction. Thus $b \in T \cup -aT$.

If $b \in T$, then σ and τ both fix $\sqrt{b}$ and both have order 2. If τ does not fix $\sqrt{a}$, then σ, τ act the same on $\sqrt{a}$ and $\sqrt{b}$ and hence commute. If τ fixes $\sqrt{a}$ then $\tau \in Z(\mathrm{Gal}(L/F))$ so in either case $[\sigma, \tau]\sigma^2 = \sigma^2 = 1$.

If $b \in -aT$, then σ fixes neither $\sqrt{a}$ nor $\sqrt{-a}$, so has order 4. Since τ acts differently on $\sqrt{a}$ and $\sqrt{b}$, it must fix one of them and be of order 2, and the same holds for $\sigma\tau$. Then $[\sigma, \tau]\sigma^2 = \sigma\tau\sigma^{-1}\tau^{-1}\sigma^2 = \tau^{-1}\sigma^{-2}\tau^{-1}\sigma^2 = 1$ since $\sigma^2 \in Z(\mathrm{Gal}(L/F))$. □

We have another convenient formulation of Proposition 7.6 as follows:

Corollary 7.7. *A subgroup T of $\dot F$ containing $\dot F^2$ is an $S(I)$-ordering if and only if the following three conditions hold.*
(a) T has level 2,
(b) F is T-rigid, and
(c) $\dot F/T \cong \coprod_{i \in I \cup \{x\}} (C_2)_i$.
In other words the $S(I)$-orderings are exactly the rigid orderings of level 2.

Proof. If $I = \emptyset$, the result follows from Definition 7.1, so we shall assume that $I \neq \emptyset$. Assume that T satisfies (1), (2) and (3) of Proposition 7.6. We show it is rigid. Let $a \in \dot F \setminus (T \cup -T)$. Then $T + aT \subset (T \cup -T) + a(T \cup -T) = T \cup -T \cup aT \cup -aT$. Take $s + at \in T + aT$ and suppose it is not in $T \cup aT$. Then it is in $-T \cup -aT$. If $s + at = -u \in -T$ then $-a = t(u+s) \in T + T = T \cup -T$, a contradiction. If $s + at = -au \in -aT$ then $-a = s/(u+t) \in T + T = T \cup -T$, a contradiction. Thus T is rigid.

By Proposition 3.3, a rigid ordering of finite level greater than 1 is exactly a rigid ordering of level 2. This proves (a) and (b).

Conversely, if T satisfies (a) and (b), then it satisfies (1) and (3) by Proposition 3.3. Let us show we also have (2). Let $a \in \dot F \setminus \pm(T \cup -T) = T \cup -T$. Then $(T \cup -T) + a(T \cup -T) = \pm(T + aT) \cup \pm(T - aT) \subseteq \pm(T \cup aT) \cup \pm(T \cup -aT) = (T \cup -T) \cup a(T \cup -T)$. Since we always have $S \cup aS \subseteq S + aS$ for any subgroup S, we see that F is $T \cup -T$-rigid. □

Example 7.8. It is well-known that if $K \longrightarrow L$ is a field extension and if T is a usual ordering of L, then $S = K \cap T$ is a usual ordering of K. This need not hold for $C(\emptyset)$-orderings nor for $S(\emptyset)$-orderings. Consider for example $L = K(\sqrt{\dot K})$ and assume that L is equipped with some $C_\emptyset$-ordering T. Since $\dot L^2 \cap K = \dot K$ and $\dot L^2 \subseteq T$, we also have $T \cap K = \dot K$: the $C_\emptyset$-ordering T "vanishes" under the restriction. This happens in particular if K is the finite field $\mathbb{F}_q$ with an odd number q of elements. With $L = \mathbb{F}_{q^2}$, $\dot L^2$ is a $C_\emptyset$-ordering. Observe that this cannot happen when T is an $S(\emptyset)$-ordering in an extension L of K: since -1 is not in T, it cannot be in $S = T \cap K$, and S cannot be the trivial index 1 subgroup. But $S(\emptyset)$-orderings are subject to another pathology of their own: it may happen that the restriction of an $S(\emptyset)$-ordering is a C_2-ordering. (Observe that this cannot happen with $C(\emptyset)$-orderings.) Take for example $K = \mathbb{Q}, L = K(\sqrt{10})$, and denote by N the norm map from L down to K. Let α be the ordering of L containing $\sqrt{10}$. Let v be the discrete rank 1 valuation on $\mathbb{Q}$ associated to the prime 3. Define $T := \{h \in \dot{L} \mid (-1)^{v(N(h))}h \in \alpha\}$. Then $-1 \notin T$ and T is a subgroup containing $\dot{K}^2$, of index 2 in $\dot{K}$ (if $x \notin T, -x \in T$). It is not a usual ordering, since $-4 - \sqrt{10}$ is negative at the two orderings of L but belongs to T, as its norm 6 has an odd 3-valuation. Thus it must be an $S(\emptyset)$-ordering. Since $N(f) = f^2$ has an even valuation when $f \in K$, we see that $S := T \cap K$ is the usual ordering of $\mathbb{Q}$.

The proof of the next proposition is nearly identical with the proof of Proposition 7.6. Therefore in the proof below, we merely indicate the key points of the proof. For the definition of a fan preordering, see [L2, Section 5].

Proposition 7.9. *A subgroup T of $\dot{F}$ containing $\dot{F}^2$ is a $D(I)$-ordering if and only if the following three conditions hold.*

(1) $-1 \notin T$,
(2) $T + T = T$,
(3) *F is T-rigid, and*
(4) $\dot{F}/T \cong \coprod_{i \in I \cup \{x\}} (C_2)_i$.

In particular a subgroup T is a $D(I)$-ordering for some index set I, if and only if it is a fan, and this happens if and only if T is a rigid ordering of infinite level.

Proof. Assume that T is a $P_{D(I)}$-ordering. Then $D(I) = G_1 \rtimes C_2$ where $G_1 = \prod_i (C_4)_i$ and $C_2 = \langle \tau \rangle$. Further, all elements in τG_1 are involutions not in $\Phi(D(I))$. Therefore we see that T is the intersection of the orderings $P_{\langle \gamma \rangle}, \gamma \in \tau G_1$. Hence T is a preordering and conditions (1) and (2) follow. Condition (4) follows from Kummer theory. By Proposition 7.2, $-1 \in P_{G_1}$, hence $P_{G_1} = T \cup -T$ and F is P_{G_1}-rigid. Since T is a preordering, this implies condition (3). Conversely, if H is an essential subgroup of $\mathcal{G}_F$ such that $T = P_H$ and T satisfies conditions (1), (2), (3), and (4), one can write H as a topological group generated by G_1 and τ where $P_{G_1} = T \cup -T$ and $P_{\langle \tau \rangle}$ is a C_2-ordering of F. Using Proposition 7.2 we see that $G_1 = \prod_i (C_4)_i$ and using the restrictions of the elements $\sigma^2[\sigma, \tau], \sigma \in G_1$, on C_4^a and $D^{a,b}$ extensions, we check that $\sigma^2[\sigma, \tau] = 1$ for all $\sigma \in G_1$. This forces $H \cong G_1 \rtimes \langle \tau \rangle$ with action $\tau^{-1}\sigma\tau = \sigma^{-1}$ for each $\sigma \in G_1$. Hence $H \cong D(I)$ as required.
It is known that conditions (1), (2), and (3) characterize fans [L2, Theorem 5.5], and by Proposition 3.3 we see that they are rigid orderings of infinite level. □

To conclude the section we may summarize the results with the following

Theorem 7.10. *Rigid orderings are exactly $C(I)-, S(I)-$ or $D(I)$-orderings for some (possibly empty) index set I.*

Proof. This is a straightforward application of Proposition 3.3, Proposition 7.2, Corollary 7.7 and Proposition 7.9. □

§8. Construction of closures for rigid orderings

In this section we employ valuation-theoretic techniques to construct closures for
$C(I)$-, $S(I)$- and $D(I)$-orderings. From the preceding section, we know that both $C(I)$- and $S(I)$-orderings are T-rigid. Then for such an ordering we will be able to use results of Arason, Elman, Jacob [AEJ], Efrat [Ef] and Ware [Wa] to associate a valuation to T. For $D(I)$-orderings, it is the "Fan Trivialization Theorem" of Bröcker [Br, Theorem 2.7] that will be used. Since it is well known (see [Ri]) that for each algebraic extension K/F we can extend any valuation v on F to a valuation w on K, we can then use this to extend $S(I)$- or $D(I)$-orderings, and in most cases also $C(I)$-orderings, from F to $F(\sqrt{t})$, $t \in T$. This will allow us to prove the existence of $S(I)$- and $D(I)$-closures, and in most cases also $C(I)$-closures.

For the reader's convenience we define here some of the valuation-theoretic notation we will be using below. For more detailed information, we refer the reader to [End] and [Ri] as well as [AEJ], [Wa] and [Br].

Let $v : F \to \Gamma \cup \{\infty\}$ be a valuation on the field F, where Γ is some linearly ordered abelian group. Then we set A_v to be the valuation subring of F, M_v to be the unique maximal ideal of A_v (consisting of those elements $f \in F$ such that $v(f) > 0$), and U_v to be the group of invertible elements of A_v. We say T is *compatible* with v (or A_v) if $1 + M_v \subseteq T$. We denote the residue field A_v/M_v by F_v, and we set $\pi_v : A_v \to F_v$ to denote the canonical epimorphism from A_v onto F_v.

The strategy of the proof is as follows: It is easy to reduce the problem of constructing H-closures to the problem of extending a given H-ordering T of a field F to an H-ordering T' of any quadratic extension $L = F(\sqrt{t}), t \in T$, such that $T' \cap F = T$. (Here $H \cong C(I), S(I)$, or $D(I)$.) In order to extend T in this manner, we first find a suitable T-compatible valuation v on F and then extend v to a valuation w on L. We then extend the induced ordering $\bar{T}$ of the residue field F_v to $\hat{T}$ on the residue field L_w of L with respect to the valuation w. Finally we lift the ordering $\hat{T}$ from the residue field L_w to an ordering $\tilde{T}$ on L, and then show that $\tilde{T}$ is the desired extension of T from F to L.

Suppose first that we are given some $S(I)$-ordering T of F. In this case, T is "not exceptional" in the sense of [AEJ, Definition 2.15]. Thus we can apply [AEJ, Theorem 2.16] to obtain the following.

Proposition 8.1. *Let T be any $S(I)$-ordering of F. Then there exists a T-compatible nondyadic valuation v of F such that $U_vT = T \cup -T$. The set $\bar{T} := \pi_v(T \cap U_v)$ is an $S(\emptyset)$-ordering of F_v.*

Proof. By [AEJ, Theorem 2.16], we have a T-compatible valuation v such that $U_vT = T \cup -T$. The last statement of the proposition follows from this. Indeed we have
$$\frac{U_v}{U_v \cap T} \cong \frac{U_vT}{T} \cong \frac{T \cup -T}{T}.$$

Since $-1 \notin T$ we see that $F_v = \bar{T} \cup -\bar{T}$ and $-1 \notin \bar{T}$. Therefore $\bar{T}$ has index 2 in $\dot{F}_v$.
Since T is an $S(I)$-ordering on F, we see that there exist elements $t_1, t_2, t_3 \in T$ such that $t_1 + t_2 + t_3 = 0$. Dividing through by that element t_i whose value $v(t_i)$ is minimal among the three elements considered (say t_1), we may assume we have

$$-1 = t_2 + t_3, v(t_2), v(t_3) \geq 0$$

Passing to the residue field we obtain $\bar{t_1} + \bar{t_2} = -\bar{1}$ in F_v. Since $-1 \notin \bar{T}$ we see that $\bar{t_i} \neq 0, i = 2, 3$. Thus $-1 \in \bar{T} + \bar{T} \setminus \bar{T}$, and $\bar{T}$ is a $S(\emptyset)$-ordering of F_v, as claimed.
Observe also that $-1 \notin \bar{T}$ implies $-1 \neq 1$ and $\operatorname{char} F_v \neq 2$. Thus v is nondyadic. □

Next suppose we have a $C(I)$-ordering T of F. Then we may apply [Ef, Propositions 2.1 and 2.3 and Theorem 4.1], to yield the following result.

Proposition 8.2. *Let T be any $C(I)$-ordering of F. Then there exists a T-compatible valuation ring A_v of F such that $[U_vT : T] \leq 2$ and $\dim_{\mathbb{F}_2} \Gamma/2\Gamma \geq |I|$, where Γ is the associated value group. The set $\bar{T} := \pi_v(T \cap U_v)$ is either $\dot{F}_v$ itself or a $C(\emptyset)$-ordering of F_v.*

Proof. Observe again that the last statement claiming that $\bar{T} := \pi_v(T \cap U_v)$ is either $\dot{F}_v$ itself or a $C(\emptyset)$-ordering, and also the statement $\dim_{\mathbb{F}_2} \Gamma/2\Gamma \geq |I|$, are consequences of the first part of the proposition. We have $\frac{U_vT}{T} \cong \frac{U_v}{U_v \cap T}$, so $[U_v : U_v \cap T] \leq 2$; hence $U_v = U_vT$ or $[U_v : U_v \cap T] = 2$. In the latter case, we see that $\bar{T} = \pi_v(T \cap U_v)$ is a $C(\emptyset)$-ordering as $-\bar{1} \in \bar{T}$. Also observe that we have $|I| + 1 = \dim_{\mathbb{F}_2} \frac{\dot{F}}{T} = \dim_{\mathbb{F}_2} \frac{\dot{F}}{U_vT} + \dim_{\mathbb{F}_2} \frac{U_vT}{T}$. From the hypothesis $[U_vT : T] \leq 2$ we see that $\dim_{\mathbb{F}_2} \frac{U_vT}{T} \leq 1$. Hence $\dim_{\mathbb{F}_2} \frac{\dot{F}}{U_vT} \geq |I|$. Therefore $\dim_{\mathbb{F}_2} \Gamma_v \geq \dim_{\mathbb{F}_2} \frac{\dot{F}}{U_vT} \geq |I|$ as claimed. □

Proposition 8.3. (Fan Trivialization Theorem [Br, Theorem 2.7]) *Let T be any $D(I)$-ordering of F. Then there exists a T-compatible valuation ring A_v of F such that the set $\bar{T} := \pi_v(T \cap U_v)$ is either an ordering of F_v or a D-ordering of F_v. (When $\bar{T}$ is an ordering, T is called a valuation fan.) Moreover, the valuation v may be chosen such that $v(T)$ contains no convex subgroups of $v(F)$.*

Now suppose that we have an $S(I)$-ordering (respectively $C(I)$-, $D(I)$-ordering) T together with a T-compatible valuation v on F. Assume $t \in T$, and let $K = F(\sqrt{t})$. Our goal is to find an $S(I)$-ordering (respectively $C(I)$-, $D(I)$-ordering) T' of K such that $T' \cap F = T$ and $\dot{F}/T \cong \dot{K}/T'$ is the isomorphism of multiplicative groups induced by the inclusion $F \hookrightarrow K$. Note that if $T' \cap F = T$, then the map $\dot{F}/T \to \dot{K}/T'$ is injective, so we need only worry about surjectivity. Then recall the well-known Krull's Theorem ([Ri, Theorem 5]):

Theorem 8.4. (Krull) *Let F be a field and $\tilde{F}$ any overfield of F. Any valuation v in F can be extended to a valuation $\tilde{v}$ in $\tilde{F}$.*

Thus there exists a valuation w on K which extends v. We now make the following convenient reduction.

Lemma 8.5. *Assume that $T_1 \subseteq T_2$ are respectively $S(I_1)$- and $S(I_2)$-orderings of F, and let $t \in T_1 \setminus \dot{F}^2$. Let $K = F(\sqrt{t})$. Suppose T_1' is an extension of T_1 to an $S(I_1)$-ordering of K. Then $T_2' := T_1'T_2$ is an $S(I_2)$-ordering of K extending T_2.*

Proof. We first show that $T_2' \cap F = T_2$. By definition, $T_2 \subseteq T_2' \cap F$, and if $f \in T_2' \cap F$ then there exists $t_1' \in T_1', t_2 \in T_2$ such that $f = t_1't_2$. This implies $t_1' \in F \cap T_1' = T_1 \subseteq T_2$, and $f \in T_2$. Thus $T_2' \cap F = T_2$.
Let $\varphi_2 : \dot{F}/T_2 \to \dot{K}/T_2'$ denote the natural homomorphism induced by the inclusion map $F \hookrightarrow K$. Because $T_2' \cap F = T_2$ we see that φ_2 is injective. Consider the following diagram:

$$\begin{array}{ccc} \dot{F}/T_1 & \xrightarrow{\varphi_1} & \dot{K}/T_1' \\ \downarrow & & \downarrow \\ \dot{F}/T_2 & \xrightarrow{\varphi_2} & \dot{K}/T_2' \end{array}$$

Since $\varphi_1 : \dot{F}/T_1 \to \dot{K}/T_1'$ is bijective and $T_1' \subseteq T_2'$, we see that φ_2 is also surjective.
Finally we shall show that T_2' is an $S(I_2)$-ordering by checking that conditions (a),(b),(c) of Corollary 7.7 hold. Since $T_2' \cap F = T_2$, we see that $-1 \notin T_2'$. As $-1 \in T_1' + T_1' \subseteq T_2' + T_2'$, we see that T_2' satisfies condition (a).
Suppose $s = u + av \in K$ with $u, v \in T_2'$ and $a \notin (T_2' \cup -T_2')$. By definition of T_2', u, v can be written $u = u_1'u_2, v = v_1'v_2$ with $u_1', v_1' \in (T_1' \cup -T_1')$, $u_2, v_2 \in T_2$. Then $su_2^{-1} = u_1' + (av_2u_2^{-1})v_1'$. Because $av_2u_2^{-1} \notin (T_1' \cup -T_1')$, the T_1'-rigidity of K implies $su_2^{-1} \in T_1' \cup (av_2u_2^{-1})T_1'$, and thus $s \in T_2' \cup aT_2'$, giving condition (b).
Finally, to check condition (c), observe that $\dot{K}/T_2' \cong \dot{F}/T_2 \cong \coprod_{i \in I_2 \cup \{x\}} (C_2)_i$. Thus T_2' is an $S(I_2)$-ordering which extends T_2. $\square$

Lemma 8.6. *Assume that $T_1 \subseteq T_2$ are respectively $C(I_1)$- and $C(I_2)$-orderings of F, and let $t \in T_1 \setminus \dot{F}^2$. Let $K = F(\sqrt{t})$. Suppose T_1' is an extension of T_1 to a $C(I_1)$-ordering of K. Then $T_2' := T_1'T_2$ is a $C(I_2)$-ordering of K extending T_2.*

Proof. The proof is identical to that of Lemma 8.5, except that one must now check that $-1 \in T_2'$. Since $T_2' \cap F = T_2$, we see $-1 \in T_2'$. $\square$

Lemma 8.7. *Assume that $T_1 \subseteq T_2$ are respectively $D(I_1)$- and $D(I_2)$-orderings of F, and let $t \in T_1 \setminus \dot{F}^2$. Let $K = F(\sqrt{t})$. Suppose T_1' is an extension of T_1 to a $D(I_1)$-ordering of K. Then $T_2' := T_1'T_2$ is a $D(I_2)$-ordering of K extending T_2.*

Proof. Again the proof takes the same arguments as in the proof of Lemma 8.5 to show that T_2' extends T_2, that $-1 \notin T_2'$ and that K is T_2'-rigid. Let us prove

$T_2' + T_2' = T_2'$. Consider $u, v \in T_2'$ and write them as above, $u = u_1' u_2, v = v_1' v_2$, with $u_1', v_1' \in T_1'$ and $u_2, v_2 \in T_2$. Then $u + v = u_2(u_1' + (v_2 u_2^{-1})v_1')$. We know that $-1 \notin T_2'$, and this implies that $v_2 u_2^{-1} \notin -T_1'$. If $v_2 u_2^{-1} \in T_1'$, then $(u+v)u_2^{-1} \in T_1' + T_1' = T_1'$ and $u + v \in T_2'$. The remaining possibility is $v_2 u_2^{-1} \notin T_1' \cup -T_1'$, and by T_1'-rigidity of K, we have $(u+v)u_2^{-1} \in T_1' \cup (v_2 u_2^{-1})T_1'$ and $u + v \in T_2'$. Hence condition (2) holds. □

We consider the following situation. Assume that $v : F \to \Gamma_v \cup \{\infty\}$ is a valuation on the field F, with valuation ring A_v and maximal ideal M_v. Let $F_v = A_v/M_v$ be the residue field, and denote by π_v the canonical homomorphism of A_v onto its quotient ring F_v.

Lemma 8.8. *Assume that v is a valuation on the field F and that T_0 is an $S(I_0)$-ordering of $\dot{F}_v$ for some (possibly empty) set I_0. Set $T_1 = \pi_v^{-1}(T_0)$. Then the group $T = T_1 \dot{F}^2$ is an $S(I)$-ordering of F with $|I| = dim_{\mathbb{F}_2}(\frac{\dot{F}}{T \cup -T})$.*

Proof. We need to check that the conditions in Corollary 7.7 hold for T. First, suppose that $-1 \in T$. Then $-1 = t_0 f^2$ for some $t_0 \in T_1, f \in \dot{F}$. Hence $f^2 = (-t_0)^{-1} \in -T_1 \subseteq U_v$, and so $f \in U_v$ as well. Passing to the residue field F_v and knowing $\dot{F_v}^2 \subseteq T_0$ we see $-1 = \bar{t}_0 \bar{f}^2 \in T_0$, which is a contradiction. Thus we must have $-1 \notin T$. Since $-1 \in T_0 + T_0$, we have $-1 + m \in T_1 + T_1$ for some m in the maximal ideal of the valuation, and $-1 + m \in -T_1 \subset T$. This shows that the level of T is 2.
To see that F is T-rigid, let $a \in \dot{F} \setminus (T \cup -T), t_1, t_2 \in T$, and consider $b := t_1 + t_2 a$. We consider various possibilities for $v(t_1)$ relative to $v(t_2 a)$. First suppose that $v(t_1) = v(t_2 a)$. Then $b = t_1(1 + t_1^{-1} t_2 a)$, with $u := t_1^{-1} t_2 a \in U_v$. Since $a \notin T \cup -T$, we see that $\pi_v(u) = \bar{u} \notin T_0 \cup -T_0$. (Otherwise $u \in \pi_v^{-1}(T_0) = T_1 \subseteq T$ or $u \in -\pi_v^{-1}(T_0) = -T_1 \subseteq -T$ and hence $a \in T \cup -T$, a contradiction.) Since we are assuming F_v is T_0-rigid, we see that $1 + \bar{u} \in T_0 \cup \bar{u} T_0$. Hence $1 + u \in \pi_v^{-1}(T_0 \cup \bar{u} T_0) = T_1 \cup u T_1$. Thus, rewriting $u = t_1^{-1} t_2 a$ and multiplying through by t_1, we see

$$b = t_1 + t_2 a \in T_1 \cup a T_1 \subseteq T \cup aT$$

as required. Now assume that $v(t_1) \neq v(t_2 a)$. If $v(t_1) < v(t_2 a)$, then again let $b = t_1(1+u)$, where $u = t_1^{-1} t_2 a$. Now, however, $v(u) > 0$, so $1 + u \in 1 + M_v \subseteq T_1 = \pi_v^{-1}(T_0)$, and thus $b \in T$. If $v(t_1) > v(t_2 a)$, set $b = a t_2(1 + t_1 t_2^{-1} a^{-1})$. We see $v(t_1 t_2^{-1} a^{-1}) > 0$, and therefore $b \in aT$. In each case $b = t_1 + a t_2 \in T \cup aT$ as desired.
It remains to see that $\dot{F}/T \cong \coprod_{i \in I \cup \{x\}} (C_2)_i$. This condition follows from the fact that $\dot{F}/T$ is an $\mathbb{F}_2$-vector space and that $dim_{\mathbb{F}_2} \dot{F}/T$ is $1 + |I|$. □

We have the analogue to Lemma 8.8 for the case of $C(I)$-orderings.

Lemma 8.9. *Assume that v is a valuation on the field F such that $[\Gamma_v : 2\Gamma_v] \geq 2$. Let T_0 be $\dot{F}_v$ or a $C(I_0)$-ordering of F_v for some (possibly empty) set I_0.*

Set $T_1 = \pi_v^{-1}(T_0)$. Then the group $T = T_1\dot{F}^2$ is a $C(I)$-ordering of F with $|I| = dim_{\mathbb{F}_2}(\frac{\dot{F}}{T}) - 1$.

Proof. We must check that the conditions of Proposition 7.2 hold for T. Clearly if $-1 \in T_0$, then $-1 \in T_1 \subseteq T$. To see that F is T-rigid, one applies the same argument as in Lemma 8.8. As in the case for $S(I)$-orderings, $\dot{F}/T$ is clearly an $\mathbb{F}_2$-vector space. Since $[\Gamma_v\colon 2\Gamma_v] \geq 2$, its dimension is strictly positive and thus may be written $dim_{\mathbb{F}_2}(\dot{F}/T) = 1 + |I|$. □

Again, we also have the analogue to Lemma 8.8 for the case of $D(I)$-orderings.

Lemma 8.10. ([Br]) *Assume that v is a valuation on the field F. Let T_0 be a fan of $\dot{F}_v$. Set $T_1 = \pi_v^{-1}(T_0)$. Then the group $T = T_1\dot{F}^2$ is a fan (i.e. a $D(I)$-ordering) of F.*

We now formulate the key results in this section.

Theorem 8.11. *Let T be any $S(I)$-ordering of F and let $L = F(\sqrt{t}), t \in T$. Then there exists an $S(I)$-ordering T' on L such that (L, T') is an $S(I)$-extension of (F, T).*

Proof. From Proposition 8.1, we see that there exists a nondyadic T-compatible valuation ring A_v in F such that $U_vT = T \cup -T$ and that $\bar{T} := \pi_v(U_v \cap T)$ is an $S(\emptyset)$-ordering of F_v. As $\pi_v^{-1}(\bar{T}) = (U_v \cap T)(1+M_v)$ and because $(1+M_v) \subseteq T$, one has $T_1 := \pi_v^{-1}(\bar{T})\dot{F}^2 \subseteq T$. By Lemma 8.8, we see that T_1 is an $S(J)$-ordering in F for a suitable set J.
Let w be any valuation of L which extends v. Let L_w denote its residue field, and Γ_v, Γ_w denote the valuation groups of v and w. We may assume $\Gamma_v \subseteq \Gamma_w$, and we set $e = [\Gamma_w : \Gamma_v]$, the ramification degree of w with respect to v, and $f = [L_w : F_v]$, the residue class degree of w with respect to v. It is well known that $ef \leq [L : F] = 2$ and in particular we have $f = [L_w : F_v] \leq 2$. More precisely, one has $L_w = F_v(\sqrt{\pi_v(u_0)})$ with $u_0 = 1$ if $f = 1$, and $u_0/t \in \dot{F}^2$ if $f = 2$. By Proposition 5.6 and Remark 5.5, C_4-orderings are known to admit C_4-closures of the same level, and as $\pi_v(u_0) \in \bar{T}$, the $S(\emptyset)$-ordering $\bar{T}$ admits an $S(\emptyset)$-extension $\tilde{T}$ to $F_v(\sqrt{\pi_v(u_0)}) = L_w$. Calling $T_2 = \pi_w^{-1}(\tilde{T})L^2$, Lemma 8.8 implies that T_2 is an $S(K)$-ordering of L for a suitable set K.
Let us first show that $T_1 = T_2 \cap F$. By definition of T_1, an element $s \in T_1$ has the same square class as an element $u \in U_v$ such that $\pi_v(u) \in \bar{T} \subseteq \tilde{T}$. This implies that $\pi_w(u) \in \tilde{T}$, and thus u and s are in T_2. This shows $T_1 \subseteq T_2 \cap F$. For the reverse inclusion, we state the following claim:

Claim. *With notation as above, one has $\dot{L} = U_w\dot{F} \cup \sqrt{t}U_w\dot{F}$.*

Proof. We know that $e \leq 2$. If $e = 1$, then $\dot{L} = \dot{F}U_w$ and we are done. If $e = 2$, then $f = 1$ and we may show that $w(\sqrt{t}) \notin \Gamma_v$. Otherwise $\sqrt{t} = xu$ with $x \in F$ and $u \in U_w$, and denoting by σ the nontrivial element of the Galois group $\mathrm{Gal}(L/F)$, we know that $\frac{\sigma(\sqrt{t})}{\sqrt{t}} = -1$ and thus $\pi_w(\frac{\sigma(\sqrt{t})}{\sqrt{t}}) = \pi_w(\frac{\sigma(u)}{u}) = -1$. Since $f = 1$, $L_w = F_v$, and so $\pi_w(\frac{\sigma(u)}{u})$ must also be 1. Since the valuation

v is not dyadic, this would be a contradiction. Thus we see that since $\Gamma_w \cong \dot{L}/U_w, \Gamma_v \cong \dot{F}/U_v$, and $[\Gamma_w : \Gamma_v] = 2$, the factor group $\dot{L}/U_w\dot{F}$ is $\{1,\sqrt{t}\}$, and we can write $\dot{L} = U_w\dot{F} \cup \sqrt{t}U_w\dot{F}$. $\square$

We now finish the proof of the theorem. If $\alpha \in T_2 \cap F$, we may write $\alpha = u\lambda^2$ with $u \in \pi_w^{-1}(\tilde{T}), \lambda \in \dot{L}$, and writing $\lambda = \sqrt{t}^{\eta} u_1 g$ with $u_1 \in U_w, g \in \dot{F}, \eta = 0$ or 1, this yields $\alpha = uu_1^2 t^\eta g^2$. Since $t^\eta g^2 \in T_1$, we may assume $\alpha = uu_1^2$. Then $\pi_v(\alpha) = \pi_w(\alpha) \in \tilde{T} \cap F_v = \bar{T}$ and $\alpha \in T_1$. This proves $T_1 = T_2 \cap F$.

We define a new subgroup T_2' of $\dot{L}$ as follows.

(1) If $\sqrt{t} \in (T_2 \cup -T_2)$, set $T_2' = T_2$.
(2) If $\sqrt{t} \notin (T_2 \cup -T_2)$ and $[\Gamma_w : \Gamma_v] = 1$, again set $T_2' = T_2$.
(3) If $\sqrt{t} \notin (T_2 \cup -T_2)$ and $[\Gamma_w : \Gamma_v] = 2$, set $T_2' = T_2 \cup \sqrt{t}T_2$.

Then again $T_1 = T_2' \cap F$, the only thing to prove being that in the third case, $\sqrt{t}T_2 \cap F \subseteq T_1$. But if $\alpha \in \sqrt{t}T_2 \cap F$ we have $\alpha = \sqrt{t}ug^2$ with $u \in U_w, g \in \dot{F}$ and this implies $w(\sqrt{t}) \in \Gamma_v$, contradicting $[\Gamma_w : \Gamma_v] = 2$. This shows that $\sqrt{t}T_2 \cap F = \emptyset$ in the third case.

Since T_2 is an $S(K)$-ordering, it is easy to check that conditions (1)-(3) of Proposition 7.6 hold for T_2' and to see that T_2' is also an $S(K')$-ordering for a suitable set K'.

We want to show that the injection $\dot{F}/T_1 \longrightarrow \dot{L}/T_2'$ is also surjective, which reduces to showing that $\dot{L} = T_2'\dot{F}$. We already know $\dot{L} = U_w\dot{F} \cup \sqrt{t}U_w\dot{F}$, and by Lemma 8.1, $U_w \subseteq T_2 \cup -T_2$. This gives us $U_w\dot{F} \subseteq T_2\dot{F} \subseteq T_2'\dot{F}$. In cases (1) and (3), one has $\sqrt{t} \in T_2' \cup -T_2'$, and so $\dot{L} \subseteq T_2'\dot{F}$. In case (2), there exists $x_0 \in \dot{F}$ such that $\sqrt{t}x_0 \in U_w \subseteq T_2\dot{F}$. So $\sqrt{t} \in T_2\dot{F}$, finishing the proof that $\dot{F}/T_1 \longrightarrow \dot{L}/T_2'$ is an isomorphism.

We have proved so far that (L, T_2') is an $S(J)$-extension of (F, T_1), and that T_1 is contained in the $S(I)$-ordering T. We may then apply Lemma 8.5 to show that (L, T_1T_2') is an $S(I)$-extension of (F, T), and the theorem is proved. $\square$

Corollary 8.12. *An $S(I)$-ordered field (F, T) admits an $S(I)$-closure.*

Proof. Let $\mathcal{S}$ be the set of extensions (L, S) of (F, T) inside $F(2)$ such that S is an $S(I)$-ordering on L. Then by a Zorn's Lemma argument $\mathcal{S}$ has a maximal element (K, T_0) with $\dot{K}/T_0 \cong \dot{F}/T, T = T_0 \cap F$, and T_0 is an $S(I)$-ordering on K. We are done by Corollary 4.3 if we can show $T_0 = \dot{K}^2$. If not, choose $t \in T_0 \backslash \dot{K}^2$. Then by Theorem 8.11 we can extend T_0 to an $S(I)$-ordering on $K(\sqrt{t})$, contradicting the maximality of (K, T_0). $\square$

Corollary 8.12 can be reformulated in the language of Galois theory as in the following corollary, which tells us that a certain family of subgroups of $G_F := \mathrm{Gal}(F(2)/F)$ occurs whenever G_F contains certain subquotients of G_F. Observe that in Corollary 8.13 we do not specify the action of the outer factor $\mathbb{Z}_2$ on the normal subgroup $(\mathbb{Z}_2)^I$ as this action depends upon a subtler analysis of the roots of unity belonging to the fields under consideration.

Corollary 8.13. *Let F be a field of characteristic $\neq 2$. Suppose that we have a tower of field extensions $F \subset N_1 \subset N_2 \subset N_1^{(3)} \subset F(2)$, such that*

$\mathrm{Gal}(N_1^{(3)}/N_2) \cong (C_4)^I \rtimes C_4$ *for* I *some nonempty set. Then* $G_F = \mathrm{Gal}(F(2)/F)$ *contains the closed subgroup* $(\mathbb{Z}_2)^I \rtimes \mathbb{Z}_2$.

Proof. Let $F \subset N_1 \subset N_2 \subset N_1^{(3)} \subset F(2)$ be a tower of field extensions, where $N_1^{(3)}/N_2$ is a Galois extension and $\mathrm{Gal}(\frac{N_1^{(3)}}{N_2}) \cong (C_4)^I \rtimes C_4$ for I some nonempty set. Set $T = \{t \in \dot{N}_1 \mid (\sqrt{t})^\sigma = \sqrt{t}$ for each $\sigma \in \mathrm{Gal}(N_1^{(3)}/N_2)\}$. From Definition 7.1 we see that T is an $S(I)$-ordering of N_1. From Corollary 8.12 it follows that there exists a field extension N of N_1 such that $\dot{N}^2$ is an $S(I)$-ordering of N and $\dot{N}^2 \cap N_1 = T$. Then Proposition 8.1 implies the existence of an $\dot{N}^2$-compatible valuation ring A_v of N such that $U_v\dot{N}^2 = \dot{N}^2 \cup -\dot{N}^2$.
It is well known that an $\dot{N}^2$-compatible valuation v on N is 2-henselian. Moreover N is a rigid field (and is $S(I)$-closed). In Proposition 8.1 we observed that v is a nondyadic valuation (i.e., char $F_v \neq 2$) and in this case it follows from basic valuation theory (see e.g. [End, §20]) that we have a split short exact sequence

$$1 \longrightarrow I_v \longrightarrow G_N(2) \longrightarrow G_{N_v}(2) \longrightarrow 1,$$

where I_v is the inertia subgroup of $G_N(2) := \mathrm{Gal}(N(2)/N) = \mathrm{Gal}(F(2)/N)$ and N_v is the residue field of v. Moreover it is well known that I_v is an abelian group. (See e.g. [EnKo].)
Because $\dot{N}^2$ is an $S(I)$-ordering of N we see that $s(N) = 2$. In particular N is not a formally real field, and so $G_N(2)$ is a torsion-free group. (See [Be].) Therefore using Pontrjagin's duality and the well-known structure of abelian divisible groups, we see that $I_v \cong (\mathbb{Z}_2)^J$ for some set J. (See e.g. [RZ, §4.3, Theorem 4.3.3].)
Because $\dot{N}^2$ is compatible with v and

$$\frac{U_v}{U_v \cap \dot{N}^2} \quad \cong \quad \frac{\dot{N}^2 \cup -\dot{N}^2}{\dot{N}^2},$$

we see that $\mid \dot{N}_v/\dot{N}_v^2 \mid = 2$. Hence $G_{N_v}(2) \cong \mathbb{Z}_2$. Since $\dot{N}^2$ is an $S(I)$-ordering of N, it follows that the cardinality of I is the same as the cardinality of J. Hence $I_v \cong (\mathbb{Z}_2)^I$. Since the Galois group $G_N(2) = I_v \rtimes \mathbb{Z}_2$ is a closed subgroup of G_F, the proof is completed. □

In the case of $C(I)$-orderings, we cannot always find a closure. The problem arises from the fact that the valuation whose existence is guaranteed by Proposition 8.2 may be dyadic, and thus the appropriate modification of Theorem 8.11 will not go through. For $S(I)$- and $D(I)$-orderings we do not have this problem, as the valuation in question will be nondyadic. Example 8.14 below constructs a $C(1)$-ordered field which we show in Proposition 8.15 does not admit a $C(1)$-closure.

Example 8.14. Recall that a field K of characteristic 2 is called *perfect* if $K^2 = K$. S. MacLane has shown that for any field K of characteristic 2, there exists a field F of characteristic 0 with a valuation $v : F \to \mathbb{Z} \cup \{\infty\}$ such that

$F_v \cong K$ ([Mac, Theorem 2]. For some more general theorems on valued fields with prescribed residue fields, see [Ri, Chapter I]). Then let F be such a field where $F_v = K$ is a field of characteristic 2 which is not perfect. Let T_0 be a multiplicative subgroup of $\dot{K}$ of index 2 in $\dot{K}$ such that $\dot{K}^2 \subsetneq T_0 \subsetneq \dot{K}$. Let $T = \dot{F}^2\pi_v^{-1}(T_0)$, a subgroup of $\dot{F}$. Here π_v is the residue map $U_v \longrightarrow \dot{K}$. Then $|\, \dot{F}/T \,| = 4$, and one can choose as representatives of the factor group $\dot{F}/T$ the elements $1, u, \rho, \rho u$ where $v(\rho) = 1, u \in U_v$, and $\pi_v(u) \notin T_0$.
We claim that F is T-rigid. Since any element in ρT or in $\rho u T$ lies outside of $U_v T$, we see that all elements of $\rho T \cup \rho u T$ are T-rigid. (See [AEJ, Proposition 1.5.]) Consider an element $\alpha = t_1 + t_2 u \in T + uT$, with $t_1, t_2 \in \dot{F}$. Then $\alpha = t_2(t_1 t_2^{-1} + u)$, so it is enough to show $t_1 t_2^{-1} + u \in T \cup uT$. Thus we may restrict our attention to elements which can be written as $tf^2 + u$, where $t \in \pi_v^{-1}(T_0), f \in \dot{F}$. If $v(f) = 0$, then $tf^2 + u \in U_v \subseteq T \cup uT$. If $v(f) > 0$, then $tf^2 + u = u(1 + tf^2u^{-1}) \in uT$. Finally, if $v(f) < 0$, then $tf^2 + u = tf^2(1 + uf^{-2}t^{-1}) \in T$. Thus F is T-rigid.
Since $-1 \in T_0$, we have $-1 \in T$, and T is a $C(1)$-ordering of F. Observe that $T \neq \dot{F}^2$ and (F, T) is not $C(1)$-closed.

Proposition 8.15. *The $C(1)$-ordered field (F, T) does not admit a $C(1)$-closure.*

Proof. Recall that a valuation ν on a field L is said to be *T-coarse* if $\nu(T)$ contains no nontrivial convex subgroups of the valuation group Γ_ν of ν. Suppose that $F \subsetneq N \subsetneq F(2), \dot{N}^2 \cap F = T$, and $\dot{N}^2$ is a $C(1)$-ordering of N. Then applying [AEJ, Corollary 2.1.7] or [Wa, Theorem 2.16], we see that there exists a $\dot{N}^2$-compatible valuation w on N such that $[U_w\dot{N}^2 : \dot{N}^2] \leq 2$. This means that $|\, U_w/U_w \cap \dot{N}^2 \,| \leq 2$. We may further choose w to be the unique finest N^2-coarse N^2-compatible valuation on N (see [AEJ, Theorem 3.8]). Consider $z :=$ the restriction of the valuation w to F. First observe that z is a T-compatible valuation on F. Indeed, from $M_w \cap F = M_z$ we get $(1 + M_w) \cap F = 1 + M_z$. Thus we have
$$1 + M_z = (1 + M_w) \cap F \subseteq \dot{N}^2 \cap F = T.$$

Let Δ be the maximal convex subgroup of Γ_z contained in $z(T)$. Then set y to be the composite valuation
$$y : \dot{F} \xrightarrow{z} \Gamma_z \xrightarrow{\rho} \Gamma_z/\Delta,$$
where the last map $\rho : \Gamma_z \to \Gamma_z/\Delta$ is the natural projection. Then, following the notation of [AEJ, Definition 2.2], the valuation ring $A_y = O_F(U_zT, T)$, and $y(T)$ contains no nontrivial convex subgroups of the value group $\Gamma_y = \Gamma_z/\Delta$ ([AEJ, Lemma 3.1 and Proposition 3.2]), so y is T-coarse. Observe that y is also T-compatible. However, since $\Gamma_v = \mathbb{Z}$ and $v(T) = 2\mathbb{Z} \neq \mathbb{Z}$, the valuation v is also T-coarse. Hence, by [AEJ, Corollary 3.7], we see that the valuations v and y are comparable. Since A_v is a maximal proper subring of F (because $\Gamma_v = \mathbb{Z}$), we see that $A_v \supseteq A_y \supseteq A_z$. However, since $M_z \supseteq M_y \supseteq M_v$ and $2 \in M_v$, we

see that both valuations y and z are dyadic. Since $F_z \subseteq F_w$, it follows that w is also dyadic. But from [AEJ, Theorem 3.8 and Lemma 4.4], it follows that w cannot be a dyadic valuation. Indeed, $[D_N\langle 1, -n^2\rangle \dot{N}^2 : \dot{N}^2] = 4 > 2$ for all $n \in \dot{N}$. Thus we have a contradiction, and there can be no $C(1)$-closure of (F, T). $\square$

Remark 8.16. Example 8.14 is analogous to Proposition 4.5. What makes this example striking when compared to Proposition 4.5 is that here we have $\mid \dot{F}/T \mid = 4 < \infty$, but in Proposition 4.5 $\mid \dot{F}/T \mid = \infty$. Although this example is a relatively simple consequence of the work in [AEJ], it seems to be the first example where the Witt ring of a field with finitely many square classes is realizable as a "Witt ring of T-forms over some field F", but it is not realizable as an actual Witt ring of any field extension K of F. We make this last comment more precise.

First observe that, analogous to the definition of reduced Witt rings of fields, one may define $W_T(F)$ for any subgroup T of $\dot{F}$ which contains all nonzero squares in F. One possible definition is as follows: (See also [La2, Corollary 1.27] and [Sc, Chapter 2, § 9].)

Let $\mathbb{Z}[\dot{F}/T]$ be the group ring of $\dot{F}/T$ with coefficients in $\mathbb{Z}$. Let J be the ideal of $\mathbb{Z}[\dot{F}/T]$ generated by

(1) $[T] + [-T]$,
(2) $[aT] + [bT] - [(a+b)T] - [ab(a+b)T], (a, b, a+b \in \dot{F})$,
(3) $[aT][bT] - [abT], (a, b \in \dot{F})$.

Then we set $W_T(F) = \mathbb{Z}[\dot{F}/T] \diagup J$.

A systematic study of $W_T(F)$ for H-orderings T of F is very desirable, but it is not pursued in this particular paper. Here we just point out that if T is any $C(1)$-ordering of F then $W_T(F) \cong W(\mathbb{Q}_p)$, where p is any prime such that $p \equiv 1 \pmod 4$, and $\mathbb{Q}_p$ is the field of p-adic numbers.

Since T is a $C(1)$-ordering in $\dot{F}$ and $\dot{\mathbb{Q}}_p^2$ is a $C(1)$-ordering in $\mathbb{Q}_p$ (see Proposition 7.2 and [L1, Chapter 6]), we see that there exists a group homomorphism $\varphi: \dot{F}/T \longrightarrow \dot{\mathbb{Q}}_p/\dot{\mathbb{Q}}_p^2$ such that φ takes any relation in the form (1), (2) or (3) above again to a relation of the same type. Using the same argument for φ^{-1} rather than φ, we see that φ indeed induces an isomorphism $\tilde{\varphi}: W_T(F) \cong W(\mathbb{Q}_p)$.

Similar to Proposition 4.12, we have the following proposition.

Proposition 8.17. *Let (F, T) be the field F with $C(1)$-ordering T constructed in Example 8.14 above. Then there is no field extension K/F with $C(1)$-ordering $\dot{K}^2$ which is a T-extension of (F, T). (Equivalently, $W_T(F)$ cannot be realized as $W(K)$ for any field extension K of F.)*

Proof. Suppose to the contrary that there exists a field extension K/F such that $\dot{K}^2$ is a $C(1)$-ordering of K and $(\dot{K}, \dot{K}^2)$ is a T-extension of (F, T). Assume that both K and a quadratic closure $F(2)$ of F are contained in some

common overfield so that we can consider the field $L = K \cap F(2)$. The natural isomorphism $\psi\colon \dot F/T \longrightarrow \dot K/\dot K^2$ factors through $\theta\colon \dot F/T \longrightarrow \dot L/(\dot K^2 \cap L)$. Because ψ is injective, so is θ. Observe that θ is also surjective. Indeed since ψ is surjective, we see that for each $l \in \dot L$ there exists an element $f \in \dot F$ such that $lf^{-1} \in \dot K^2 \cap L$. Thus we see that $(L, \dot K^2 \cap L)$ is a T-extension of (F, T).
We claim that $(\dot L, \dot K^2 \cap L)$ is a $C(1)$-closure of $(F, \dot T)$. Observe that $\dot K^2 \cap L = \dot L^2$. Indeed if $k^2 \in L, k \in \dot K$ then $k \in \dot K \cap F(2) = \dot L$. Since $\dot L^2 \subset \dot K^2 \cap L$ is obvious, we see that $\dot K^2 \cap L = \dot L^2$. In order to conclude the proof, it is enough to show that $\dot L^2$ is a $C(1)$-ordering in $\dot L$. Because $\sqrt{-1} \in \dot K$ we see that $\sqrt{-1} \in \dot L$ as well, and $-1 \in \dot L^2$. From the isomorphism $\theta\colon \dot F/T \longrightarrow \dot L/\dot L^2$ we see that $\dot L/\dot L^2 = C_2 \oplus C_2$. By Proposition 7.2 it remains only to show that L is $\dot L^2$-rigid. Consider an element $a \in \dot L \diagdown \dot L^2$. For any $l \in \dot L$ we have $l^2 + a \in \dot K^2 \cup a\dot K^2$ because $\dot K$ is $\dot K^2$-rigid and $\dot L^2 = \dot K^2 \cap L$. Hence $l^2 + a \in (\dot K^2 \cap L) \cup (a\dot K^2 \cap L)$. Finally since $\dot K^2 \cap L = \dot L^2$ and $a\dot K^2 \cap L = a\dot L^2$ we see that $\dot L$ is $\dot L^2$-rigid. □

Theorem 8.18. *A $C(I)$-ordered field (F,T) possessing a nondyadic T-compatible valuation ring A_v as in Proposition 8.2 admits a $C(I)$-closure.*

Proof. The proof is essentially the same as the proof of Theorem 8.11 and Corollary 8.12, and we will follow the same plan and the same notation. Applying Proposition 8.2, we find a valuation v on F such that $\bar T := \pi_v(U_v \cap T)$ is either $\dot F_v$ or a C-ordering. By assumption here this valuation is nondyadic. By Lemma 8.9, T_1 is a $C(J)$-ordering contained in T. Taking any valuation w on $L = F(\sqrt{t})$ extending v, we extend $\bar T$ to $\tilde T$ in L_w. We obtain, by Lemma 8.9, a $C(K)$-ordering T_2 in L. We enlarge it to a $C(K')$-ordering T_2', according to the three cases (1), (2), (3), replacing $T_2 \cup -T_2$ by T_2. The only serious change is in proving that $\dot L = T_2'\dot F$. For this it is enough to show that $U_w \subseteq T_2\dot F$, which can be done as follows. If the index $[U_vT : T] = [\dot F_v : \bar T]$ is 1, then $[\dot L_w : \tilde T] = [U_wT_2 : T_2] = 1$ and $U_w \subseteq T_2$. If this index is 2, there exists $a \in U_v$ such that $U_w \subseteq T_2 \cup aT_2 \subseteq T_2\dot F$. This shows that (L, T_2') is a $C(J)$-extension of (F, T_1), and we apply Lemma 8.5 to show that (L, T_1T_2') is a $C(I)$-extension of (F,T). We finish by applying the same argument as in Corollary 8.12. □

The following observation about valuations when F contains a real-closed field was pointed out to us by J.-L. Colliot-Thélène. It contains a convenient condition for a valuation v to be nondyadic, and thus it is related to Theorem 8.18.

Example 8.19. *Let v be a valuation with value group Γ, and denote by U_v the units of the valuation ring. Suppose there exists an integer $n > 1$ such that any n-divisible subgroup of Γ is trivial. Assume that F contains a real-closed field R. Then R is contained in U_v, and in particular the valuation is nondyadic.*

Proof. Assume F contains a real-closed field R. If $a \in R$ is positive, for the given integer n there exists $b \in R$ such that $a = b^n$, and thus $v(a) = nv(b)$. Thus $v(a)$, being divisible by any power of n, must be 0, and the nonzero elements of R must be units. This implies that the residue field F_v contains an isomorphic copy of R, and the valuation v cannot be dyadic. □

Theorem 8.20. *A $D(I)$-ordered field (F,T) admits a $D(I)$-closure.*

Proof. We have already proved that D-orderings admit closures, and thus we may assume that $|I| > 1$. It has also already been shown in [Sch] that valuation fans admit closures. Here is a more general situation and a different proof, that consists again in transpositions of the proofs of Theorem 8.11 and Corollary 8.12. As in Theorem 8.11, if $t \in T$ and $L = F(\sqrt{t})$, applying Proposition 8.3, we find a valuation v on F such that $\bar{T} := \pi_v(U_v \cap T)$ is either an ordering or a D-ordering. By Lemma 8.10, T_1 is a $D(J)$-ordering contained in T. Taking any valuation w on $L = F(\sqrt{t})$ extending v, we extend $\bar{T}$ to $\tilde{T}$ in L_w. By Lemma 8.10 we obtain a $D(K)$-ordering T_2 in L. We enlarge it to a $D(K')$-ordering T_2', according to the three cases (1), (2), (3), replacing $T_2 \cup -T_2$ by T_2. As in the case for $C(I)$-ordered fields, the only serious change is in proving that $\dot{L} = T_2'\dot{F}$, and the proof is identical to that for $C(I)$-ordered fields. □

§9. Galois groups and additive structures (2)

Throughout this paper, we have considered a number of subgroups H of $\mathcal{G}_F$ which behave pretty well, in that we have a certain control over the additive structure of the associated orderings, and we are able to make closures. Actually some of these groups H have an additional property which helped us in a subtle but important way. Let us introduce the following definition and notation.

Definition and Notation 9.1.
(1) We say that an essential subgroup H of $\mathcal{G}_F$ is *lifted* if we can write $\mathcal{G}_F = G \rtimes H$ for some normal subgroup G of $\mathcal{G}_F$. This means that H is not only a subgroup of $\mathcal{G}_F$, but also a quotient $\mathcal{G}_F \longrightarrow H$ such that $H \longrightarrow \mathcal{G}_F \longrightarrow H$ is the identity map. The H-ordering P_H is called a *lifted* ordering. (The name *lifted* was chosen because such an H corresponds, as a quotient of $\mathcal{G}_F$, to a Galois extension of F inside $F^{(3)}$, of group H, which can be lifted as a Galois subextension of $F^{(3)}$ of same group H.)
(2) If we want to realize some subgroup H of $\mathcal{G}_F$ as a $\mathcal{G}_K$ for some field K, we certainly need to use an H which satisfies known properties of W-groups. In particular, if $H \neq \{1\}, C_2$, then by Corollary 2.18 of [MiSp2], we see that H can be embedded in a suitable product $\prod_I D \times \prod_J C_4$, where each factor is a quotient of H. According to the use in universal algebra, see e.g. [Gr, p. 123], we refer to H as the *subdirect product* of $\prod_I D \times \prod_J C_4$. (Also we say that H as above satisfies the *subdirect product* condition.)
(3) We say that an essential subgroup H of $\mathcal{G}_F$ is a *fair* subgroup if it is lifted and if it is either $\{1\}$ or C_2 or a subdirect product of some $\prod_I D \times \prod_J C_4$. The H-ordering P_H will be called a *fair* ordering if H is a fair subgroup of $\mathcal{G}_F$.

Remark 9.2. We observed in Example 6.4 that the subgroup $H = \langle \sigma, \tau \rangle \cong C_4 * C_4$ in $\mathcal{G}_{\mathbb{Q}_2}$ has associated H-ordering $T = \dot{F}^2 \cup 5\dot{F}^2, F = \mathbb{Q}_2$, such that $T + T$ is not multiplicatively closed. We now use the description of $\mathcal{G}_F = \mathbb{G}_2$

as in Example 2.9, to show that H is <u>not</u> a lifted subgroup of $\mathcal{G}_F$. Suppose instead that H is a lifted subgroup of $\mathcal{G}_F$. Then there exists a subgroup G of $\mathcal{G}_F$ such that $\mathcal{G}_F = G \rtimes H$. Then G must contain some element of the form $\alpha = \rho h \phi$ where ρ is an element of $\mathcal{G}_F$ such that ρ, σ, τ generate $\mathcal{G}_F$ and $\sigma^2[\rho,\tau] = 1, h \in H$ and ϕ is some element in $\Phi(\mathcal{G}_F)$. Because G is a normal subgroup of $\mathcal{G}_F$ we see that $\alpha \in G$ implies $\alpha^{-1}(\tau^{-1}\alpha\tau) = [\alpha,\tau] \in G$ as well. Hence $[\alpha,\tau] = [\rho h\phi,\tau] = [\rho,\tau][h,\tau] \in G$. On the other hand $[\rho,\tau][h,\tau] = \sigma^2[h,\tau] \in H$. Because $\mathcal{G}_F = G \rtimes H$ we see $G \cap H = \{1\}$ and thus $\sigma^2[h,\tau] = 1$. This equality is impossible as H is a free group in category $\mathcal{C}$. Therefore H is not lifted.

Observe that it is sometimes fairly easy to establish the "fairness" of a given subgroup. For example if $H = \langle\sigma\rangle$ is an essential subgroup of $\mathcal{G}_F$ of order 2, then for $f \notin P_H$ the restriction $H \longrightarrow \mathrm{Gal}(F(\sqrt{f})/F)$ induces an isomorphism. Since the subdirect product condition is empty, H is fair. We can also readily check the following:

Proposition 9.3. *Let $\varphi: D(I) \longrightarrow \mathcal{G}_F$ be an essential embedding. Then $\varphi(D(I))$ is a lifted subgroup of $\mathcal{G}_F$. As the subdirect product condition is also trivially satisfied, it is a fair subgroup of $\mathcal{G}_F$.*

Proof. Consider a $D(I)$-ordering T of F for some $|\, I \,| \geq 1$. Pick a basis for $\dot{F}/T$ of the form $\{[-1]\} \cup \{[a_i], i \in I\}$. (As usual $[f]$ means the class represented by f in the factor group $\dot{F}/T$.) Set $K/F = F(\sqrt{-1}, \sqrt[4]{a_i}: i \in I)$. Then $\mathrm{Gal}(K/F) \cong (\prod_I C_4) \rtimes C_2$, where we can choose generators $\bar{\tau}_i$, $i \in I$ for factors in the inner product and $\bar{\sigma}$ for the outer factor such that $\bar{\sigma}(\sqrt{-1}) = -\sqrt{-1}$, $\bar{\sigma}\,(\sqrt[4]{a_i}) = \sqrt[4]{a_i}$, $\bar{\tau}_i\,(\sqrt{-1}) = \sqrt{-1}$ and $\bar{\tau}_i\,(\sqrt[4]{a_i}) = \sqrt{-1}\,\sqrt[4]{a_i}$, $\bar{\tau}_i\,(\sqrt[4]{a_j}) = \sqrt[4]{a_j}$ for $j \neq i$. Moreover the action of $\bar{\sigma}$ on $\prod_I C_4$ is described as $\bar{\sigma}^{-1}\ \bar{\tau}_i\ \bar{\sigma} = \bar{\tau}_i^3$ for each $i \in I$. (Or equivalently $\bar{\sigma}^{-1}\ \bar{\tau}\ \bar{\sigma} = \bar{\tau}^{-1}$ for each $\bar{\tau} \in \prod_I C_4$.)
Pick any elements σ, τ_i, $i \in I \in H := \varphi(D(I))$ such that their homomorphic image from H to $\mathrm{Gal}(K/F)$ are elements $\bar{\sigma}$, $\bar{\tau}_i$, $i \in I$. This is possible as H surjects on $\mathrm{Gal}(K/F)$. Then the essential subgroup H of $\mathcal{G}_F$ is generated by the minimal set of generators $\{\sigma, \tau_i, i \in I\}$. Moreover the natural restriction map $r: H \longrightarrow \mathrm{Gal}(K/F)$ is an isomorphism, as r takes the generators of H to the generators of $\mathrm{Gal}(K/F)$ and both sets of generators satisfy the same relations. $\square$

Now we consider C_4-orderings and determine when they are fair orderings. Observe that a C_4-ordering is automatically fair provided it is lifted, so it is enough to decide when a C_4-ordering T is lifted.

Proposition 9.4. *Let T be a C_4-ordering of F. Then T is lifted if and only if there exists an element $f \in (F^2 + F^2) \setminus (T \cup \{0\})$.*

Proof. Suppose that T is a C_4-ordering of $F, T = P_H$ for $H \cong C_4$, and H is essentially embedded in $\mathcal{G}_F$. Suppose also that $f \in (F^2+F^2) \setminus (T \cup \{0\})$. Then since $f \notin T$ and $T \supset \dot{F}^2$, we see that $f \notin F^2$ and a C_4^f-extension K of F exists. Because $f \in \dot{F} \setminus T$, an element $h \in H$ exists such that $h(\sqrt{f}) = -\sqrt{f}$.

Then the image of h in $\mathrm{Gal}(K/F)$ under the natural homomorphism $H \longrightarrow \mathrm{Gal}(K/F)$ is a generator of $\mathrm{Gal}(K/F)$. Therefore the homomorphism is in fact an isomorphism, and H is lifted as asserted. Assume now that $H \cong C_4$ is a lifted subgroup of $\mathcal{G}_F$. Then a surjective homomorphism $\varphi\colon \mathcal{G}_F \longrightarrow C_4$ exists, which induces an isomorphism $\psi\colon H \longrightarrow C_4$. Let K be the fixed field of the kernel of φ. Then K/F is a Galois extension on $\mathrm{Gal}(K/F) \cong C_4$. Let $F(\sqrt{f})$ be a unique quadratic extension of F contained in K. Also let $T = P_H$. Then H acts nontrivially on $\sqrt{f}$ and $f \in (F^2+F^2)\setminus\{0\}$. Hence $f \in (F^2+F^2)\setminus(T\cup\{0\})$ as claimed. □

Example 9.5. The following simple example shows that we cannot drop the condition $\exists\, f \in (F^2 + F^2) \setminus (T \cup \{0\})$ from the proposition above, and that unfair C_4-orderings exist in nature. Consider again $F = \mathbb{Q}_2$ and set $T = (F^2 + F^2) \setminus \{0\}$. Then T is a subgroup of $\dot{F}$ of index 2. Because $\mathbb{Q}_2$ is not a formally real field, $\mathbb{Q}_2$ does not admit any usual ordering, and T is a C_4-ordering of F. However T contains all sums of two squares, and therefore T is not lifted.

On the bright side, we wish to point out that for each C_4-ordering there exists a quadratic extension of the base field, and an extension of the original C_4-ordering on this quadratic extension where this extended ordering become a fair ordering. In other words an unfair ordering may become fair in some algebraic extension. More precisely we have the following proposition, in which we use Definition 1.4(4) of an H-extension

Proposition 9.6. *Let T be a C_4-ordering in F. If T is not fair, there exists $t \in T$ and a C_4-extension $(F(\sqrt{t}), V)$ of (F, T) such that V is a fair ordering in $F(\sqrt{t})$.*

Proof. Suppose that T is a C_4-ordering in F. Then by Proposition 5.4 there must exist an element $t \in T$ such that $1 + t \notin T$. If T is not a fair ordering, we know from the characterization of fair orderings in Proposition 9.4 that $t \notin \dot{F}^2$. Hence $K = F(\sqrt{t})$ is a quadratic extension of F and $[K\colon F] = 2$. From the proof of Proposition 4.2, we know that there exists some subgroup V in K such that $\mid \dot{K}/V \mid = 2$ and $V \cap \dot{F} = T$. Then V is a C_4-ordering of K, and V is fair as $1 + (\sqrt{t})^2 \notin V$. □

In this section we merely give a few examples of fair orderings and are not pursuing a systematic check of which orderings considered in this paper are fair and which will become fair after extension to a suitable 2-extension of the base field. The development of a theory of fair orderings of fields is planned for a subsequent paper.

We complete our family of examples of orderings by considering $H = \mathcal{F}(I)$, where I is some nonempty index set and $\mathcal{F}(I)$ is the free pro-2-group in the category $\mathcal{C}$, on a minimal set $\{\sigma_i \mid i \in I\}$ of generators I. (We assume as usual that each open subgroup V of $\mathcal{F}(I)$ contains all but finitely many $\sigma_i, i \in I$. See [Koc, Chapter 4].)

Proposition 9.7. *Let K/F be a Galois extension such that* $\mathrm{Gal}(K/F) \cong \mathcal{F}(I) = \langle \sigma_i | i \in I\rangle$ *where* $\{\sigma_i, i \in I\}$ *is a family of minimal generators of the free pro-2-group* $\mathcal{F}(I)$ *in our category* $\mathcal{C}$. *Then there exists a fair* $\mathcal{F}(I)$*-ordering in* F.

Proof. We first embed the group $\mathcal{F}(I)$ essentially in $\mathcal{G}_F$. Since $F^{(3)}$ is the maximal Galois subextension of a quadratic closure F_q of F such that $\mathrm{Gal}(F^{(3)}/F)$ belongs to the category $\mathcal{C}$, and since $\mathcal{F}(I)$ also belongs to $\mathcal{C}$, we see that $K \subset F^{(3)}$. Therefore there exists a surjective natural homomorphism $\pi: \mathcal{G}_F \longrightarrow \mathrm{Gal}(K/F)$.
It is well known that there exists a continuous map $s: \mathrm{Gal}(K/F) \longrightarrow \mathcal{G}_F$ such that $\pi \circ s$ is the identity map on $\mathrm{Gal}(K/F)$ (See [Koc, 1.3]). (Here we use only the fact that both groups $\mathrm{Gal}(K/F)$ and $\mathcal{G}_F$ are profinite groups.) Set $s(\sigma_i) = \omega_i$ for each $i \in I$. Then for each open subgroup V of $\mathcal{G}_F$ the set $s^{-1}(V)$ is an open subset of $\mathrm{Gal}(K/F)$, and because open subgroups of $\mathrm{Gal}(K/F)$ form a basis for the topology of $\mathrm{Gal}(K/F)$ we see that all but finitely many $\sigma_i, i \in I$, are in $\sigma^{-1}(V)$. Hence all but finitely many ω_i are in V.
Because $\mathcal{F}(I)$ is a free object of $\mathcal{C}$ on the set of generators $(\sigma_i), i \in I$ we see that there exists a continuous homomorphism $p: \mathrm{Gal}(K/F) \longrightarrow \mathcal{G}_F$ such that $p(\sigma_i) = \omega_i$ for each $i \in I$. Set $H = p(\mathrm{Gal}(K/F))$. Then we have $\pi \circ p = 1$ and $\mathcal{G}_F \cong \ker \pi \rtimes H$. Moreover, π restricted to H induces an isomorphism $\varphi: H \longrightarrow \mathrm{Gal}(K/F)$. Observe that $\varphi(\omega_i) = \sigma_i$ for each $i \in I$. Because σ_i mod $\phi(\mathrm{Gal}(K/F))$ are topologically independent, we see that ω_i must be topologically independent mod $\phi(\mathcal{G}_F)$. This means that $\{\omega_i, i \in I\}$ generates the essential subgroup H of $\mathcal{G}_F$.
One can check that $\mathcal{F}(I)$ is a subdirect product of its dihedral and C_4 quotients directly from the structure of $\mathcal{F}(I)$, but it is also possible simply to observe that $\mathcal{F}(I)$ is the W-group of a suitable field A and all W-groups have this property. That each $\mathcal{F}(I)$ is the W-group of a suitable field A follows from the fact that for each index set $I \neq \phi$ we can find a field A such that the Galois group of its quadratic closure is a free pro-2-group (see e.g., [GM, page 98]), and therefore its W-group is $\mathcal{F}(I)$. □

The following corollary applies, for example, in the case of $F = \mathbb{Q}_p(t)$.

Corollary 9.8. *Let F be the quotient field of a complete local integral domain properly contained in F. Let $\mathcal{F}(I)$ be any free object of category $\mathcal{C}$ on generators I, where I is a nonempty finite set. Then F admits a fair $\mathcal{F}(I)$-ordering.*

Proof. From Proposition 9.7 we see that it is sufficient to show that each group $\mathcal{F}(I), I$ finite and nonempty, occurs as a Galois group over F. Harbater's well-known result [Har, p. 186] says that each finite group is realizable over F. (For a very nice and rather elementary proof of this result see [HaVöl, Theorem 4.4].) □

Let us fix the following notation.

Notation 9.9. Let $i\colon F_1 \longrightarrow F_2$ be a quadratic extension and let $i^\star\colon \mathcal{G}_{F_2} \longrightarrow \mathcal{G}_{F_1}$ be the associated restriction map. (See e.g. [MiSm3] for the existence of this map.) Let H_2 be a subgroup of $\mathcal{G}_{F_2}$ and let $H_1 = i^\star(H_2)$. Assume H_1 is essential in $\mathcal{G}_{F_1}$. Observe that this property is not automatically satisfied since the image of an essential group under the restriction map $i^\star$ need not be essential. (See Remark 7.8 for an example exhibiting such a case.) When this is the case, we say that the extension $(F_1, H_1) \longrightarrow (F_2, H_2)$ is essential. Put $T_1 = P_{H_1}, T_2 = P_{H_2}$. Then it follows that $T_1 = T_2 \cap F_1$.

If we are working with fair groups H as above, then we can show that for an essential quadratic extension $(F_1, H_1) \longrightarrow (F_2, H_2)$, the additive structure of the associated orderings is preserved if and only if $i^\star$ induces an isomorphism between H_2 and H_1.

Theorem 9.10. *Assume the hypotheses in Notation 9.9 hold and that H_1, H_2 are fair subgroups of $\mathcal{G}_{F_1}, \mathcal{G}_{F_2}$ respectively. Then the restriction $i^\star$ induces an isomorphism between H_2 and H_1 if and only if $\dot F_1/T_1 \cong \dot F_2/T_2$ and for each $a \in F_1, T_1 + aT_1 = (T_2 + aT_2) \cap F_1$.*

Since the proof is a bit long and since the two directions are not using the same assumptions on H_1, H_2, we split the theorem in two parts, Proposition 9.11 and Proposition 9.12

Proposition 9.11. *Assume that H_1 is lifted. Following Notation 9.9, if the restriction $i^\star$ induces an isomorphism between H_2 and H_1, then $\dot F_1/T_1 \cong \dot F_2/T_2$ and for each $a \in F_1$, $\quad T_1 + aT_1 = (T_2 + aT_2) \cap F_1$.*

Proof. We know that $\dot F_i/T_i$ is the Pontrjagin dual of $H_i/\Phi(H_i)$ for $i = 1, 2$. Thus the natural isomorphism $H_2 \longrightarrow H_1$ yields an isomorphism $\dot F_1/T_1 \cong \dot F_2/T_2$. In order to show that for each $a \in F_1$ we have $T_1 + aT_1 = (T_2 + aT_2) \cap F_1$, it is enough to show that for every $b, c \in \dot F_1 \setminus T_1$, if there exists $s_2, t_2 \in T_2$ such that $bs_2 + ct_2 = 1$, then there exists $s_1, t_1 \in T_1$ such that $bs_1 + ct_1 = 1$. Indeed, assume that the latter condition involving $b, c \in \dot F_1 \setminus T_1$ is valid. Consider any $a \in \dot F_1$ and any relation $u_2 + av_2 = d$, where $u_2, v_2 \in T_2 \cup \{0\}$ and $d \in \dot F_1$. We want to show that there exist elements $u_1, v_1 \in T_1 \cup \{0\}$ such that $u_1 + av_1 = d$. If $u_2 = 0$ then $v_2 \in \dot F_1 \cap T_2 = T_1$, and we are done. If $v_2 = 0$ then $u_2 = d \in \dot F_1 \cap T_2 = T_1$, and again we are done. Then assume $u_2, v_2 \in T_2$. If $-a \in T_1$, let us write $d = s^2 - t^2$ for some elements $s, t \in \dot F_1$. We then have $d = s^2 + a(-at^2/a^2) \in T_1 + aT_1$. Hence we may assume that $-a \notin T_1$. Finally we also assume that $d \notin T_1$. From the equation $u_2 + av_2 = d$ we obtain $u_2 = d - av_2$, and since $u_2, v_2 \in T_2$ we can rewrite this equation as $1 = ds_2 - at_2$ where $d, -a \in \dot F_1 \setminus T_1$. Using our hypothesis we see that there exist elements $s_1, t_1 \in T_1$ such that $1 = ds_1 - at_1$. Hence $d \in T_1 + aT_1$ as required.
Now take $b, c \in \dot F_1 \setminus T_1$ and assume that $bs_2 + ct_2 = 1$ for some $s_2, t_2 \in T_2$. Then the quaternion algebra $\left(\frac{bs_2, ct_2}{F_2}\right)$ splits. We consider the following cases.
(1) Suppose bs_2, ct_2 are linearly independent in $\dot F_2/T_2$. Then they are also independent modulo $\dot F_2^2$, and by Proposition 1.5 we have a dihedral extension

L_2/F_2 such that $F_2(\sqrt{bs_2}, \sqrt{ct_2}) \subset L_2$ and $\mathrm{Gal}(L_2/F_2(\sqrt{bcs_2t_2})) \cong C_4$. In particular we have an exact sequence

$$1 \longrightarrow C_2 \longrightarrow \mathrm{Gal}(L_2/F_2) \cong D \longrightarrow \mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2) \cong C_2 \times C_2 \longrightarrow 1.$$

Let θ denote the restriction map from H_2 to $\mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$. We show it is surjective. Denote by u_1, u_2 the two generators of $\mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$ defined by $u_1(\sqrt{bs_2})/\sqrt{bs_2} = -1, u_1(\sqrt{ct_2})/\sqrt{ct_2} = 1, u_2(\sqrt{bs_2})/\sqrt{bs_2} = 1, u_2(\sqrt{ct_2})/\sqrt{ct_2} = -1$. We may look at u_1, u_2 as linear functions on the $\mathbb{F}_2$-vector subspace of $\dot{F}_2/\dot{F}_2^2$ spanned by bs_2, ct_2, which are assumed to be independent, and since $bT_2 \cap cT_2 = \emptyset$, we may extend them to linear functions v_1, v_2 defined on the subspace generated by bT_2, cT_2, by putting $v_i(x) = u_i(b)$ if $x \in bT_2$ and $v_i(x) = u_i(c)$ if $x \in cT_2$. Then v_i may be viewed as a function on the $\mathbb{F}_2$-vector subspace generated by the cosets bT_2, cT_2 in $\dot{F}_2/T_2$. Again, these functions v_i's may be extended to w_i defined on the whole vector space $\dot{F}_2/T_2$. By duality, one has $(\dot{F}_2/T_2)^\star \cong H_2/\Phi(H_2)$, and the w_i's yield to elements in $H_2/\Phi(H_2)$ which may be lifted as elements $h_1, h_2 \in H_2$. Since the duality is precisely given by the pairing $H_2/\Phi(H_2) \times \dot{F}_2/T_2 \longrightarrow \{\pm 1\}$ defined by $(h, f) \mapsto h(\sqrt{f})/\sqrt{f}$, it is immediate that h_i goes to u_i under the restriction map $\theta\colon H_2 \longrightarrow \mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$. This shows the surjectivity of θ. Since θ factors through $\psi\colon H_2 \longrightarrow \mathrm{Gal}(L_2/F_2) \cong D$ and since the kernel of $\mathrm{Gal}(L_2/F_2) \longrightarrow \mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$ is the Frattini subgroup of $\mathrm{Gal}(L_2/F_2)$, we see that ψ is also surjective. This means that D may be viewed as a quotient of H_2 and that we have inclusion maps ${F_2^{(3)}}^{H_2} \longrightarrow L'_2 \longrightarrow F_2^{(3)}$ such that $\mathrm{Gal}(L'_2/{F_2^{(3)}}^{H_2}) \cong D$. Since $i^\star(H_2) = H_1$, applying $i^\star$ to this diagram gives us another diagram ${F_1^{(3)}}^{H_1} \longrightarrow L'_1 \longrightarrow F_1^{(3)}$ with $\mathrm{Gal}(L'_1/{F_1^{(3)}}^{H_1}) \cong D$. Since H_1 is lifted, we know that there exists an H_1-extension K/F_1 inside $F_1^{(3)}$ containing a D-extension L_1/F_1. This extension is a $D^{u,v}$-extension for suitable $u, v \in F_1$ by Proposition 1.5. We claim that we have $u = bs_1, v = ct_1$ for suitable $s_1, t_1 \in T_1$. Consider the surjective homomorphism

$$\theta\colon H_2 \longrightarrow \mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$$

defined above. Then θ factors through the surjective homomorphism $\psi\colon H_2 \longrightarrow \mathrm{Gal}(L_2/F_2) \cong D$. Using the isomorphism $\beta\colon H_2 \longrightarrow H_1$ induced by $i^\star$ and our construction of L_1/F_1, we see that the homomorphism $\psi\colon H_2 \longrightarrow \mathrm{Gal}(L_2/F_2)$ is compatible, via identification of H_2 with H_1 using $i^\star$, with the restriction homomorphism $\tilde{\psi}\colon H_1 \longrightarrow \mathrm{Gal}(L_1/F_1)$. Passing to the quotients $\mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$ and $\mathrm{Gal}(F_1(\sqrt{u}, \sqrt{v})/F_1)$ of $\mathrm{Gal}(L_2/F_2)$ and $\mathrm{Gal}(L_1/F_1)$ respectively, we see that we can identify the homomorphism $\theta\colon H_2 \longrightarrow \mathrm{Gal}(F_2(\sqrt{bs_2}, \sqrt{ct_2})/F_2)$ with the restriction homomorphism $\tilde{\theta}\colon H_1 \longrightarrow \mathrm{Gal}(F_1(\sqrt{u}, \sqrt{v})/F_1)$ via the isomorphism $i^\star\colon H_2 \longrightarrow H_1$. Finally from the natural isomorphism $\dot{F}_1/T_1 \cong \dot{F}_2/T_2$ we may assume that $u = bs_1$ and $v = ct_1$ for suitable elements $s_1, t_1 \in T_1$. By Proposition 1.5, this implies

that the quaternion algebra $\left(\frac{bs_1,ct_1}{F_1}\right)$ splits, and that there exist $\tilde{s_1}, \tilde{t_1} \in T_1$ such that $b\tilde{s_1} + c\tilde{t_1} = 1$.
Suppose now that bs_2, ct_2 are linearly dependent in $\dot{F}_2/T_2$. Then b and c are equal modulo T_2 and we may assume $b = c$. There are still two more cases to consider.
(2) Suppose we have $cs_2 + ct_2 = 1$ with $s_2 = t_2 \bmod \dot{F}_2^2$. By Proposition 1.5, there exists a $C_4^{cs_2}$-extension L_2/F_2 with $F_2(\sqrt{cs_2}) \subset L_2$. Using arguments similar to those in (1), we show that the restriction $\psi: H_2 \longrightarrow \mathrm{Gal}(L_2/F_2)$ is onto, and we find $s_1 \in T_1$ such that $\left(\frac{cs_1,cs_1}{F_1}\right)$ splits. This implies that there exist $\tilde{s_1}, \tilde{t_1} \in T_1$ such that $c\tilde{s_1} + c\tilde{t_1} = 1$.
(3) Suppose we have $cs_2 + ct_2 = 1$ with $s_2 \neq t_2 \bmod \dot{F}_2^2$. As in (1) we find L_2 with $\mathrm{Gal}(L_2/F_2) \cong D$ and we have a tower of fields $F_2 \longrightarrow F_2(\sqrt{s_2t_2}) \longrightarrow F_2(\sqrt{cs_2}, \sqrt{ct_2}) \longrightarrow L_2$. Since H_2 fixes $F_2(\sqrt{s_2t_2})$, the restriction map $\psi: H_2 \longrightarrow \mathrm{Gal}(L_2/F_2)$ induces a surjective homomorphism $\psi': H_2 \longrightarrow \mathrm{Gal}(L_2/F_2(\sqrt{s_2t_2})) \cong C_4$. We finish with arguments as in (2) and replacing F_2 by $F_2(\sqrt{s_2t_2})$, we find $\tilde{s_1}, \tilde{t_1} \in T_1$ such that $c\tilde{s_1} + c\tilde{t_1} = 1$. □

We now prove the result in the other direction.

Proposition 9.12. *Let H_1, H_2 be as in Notation 9.9 and assume they are fair subgroups. If the inclusion $i: F_1 \longrightarrow F_2$ induces an isomorphism $\dot{F}_1/T_1 \longrightarrow \dot{F}_2/T_2$ and if $(T_2 + aT_2) \cap \dot{F} = T_1 + aT_1$ for any $a \in F_1$, then $i^\star$ induces an isomorphism between H_2 and H_1.*

Proof. If $H_2 = \{1\}$ then $H_1 = \{1\}$ as well. If $H_2 = C_2$ then $i^\star(H_2) \neq \{1\}$ because T_2 is a usual ordering in $\dot{F}_2$, and it cannot contain $\dot{F}_1$. However if H_1 were $\{1\}$ then $T_1 = \dot{F}_1$. Therefore $i^\star$ induces an isomorphism between H_2 and H_1.
For the rest of our proof we assume that $H_2 \neq \{1\}, C_2$. Call $\beta: H_2 \longrightarrow H_1$ the restriction of $i^\star$ to H_2. Because $i^\star$ is a group homomorphism from $\mathcal{G}_{F_2}$ into $\mathcal{G}_{F_1}$, we have $i^\star(\Phi(\mathcal{G}_{F_2})) \subset \Phi(\mathcal{G}_{F_1})$. Also we have $\beta(\Phi(H_2)) \subset \Phi(H_1)$. Then the map β induces $\hat{\beta}: H_2/\Phi(H_2) \longrightarrow H_1/\Phi(H_1)$, which is an isomorphism because its dual map $\dot{F}_1/T_1 \longrightarrow \dot{F}_2/T_2$ is an isomorphism. By definition β is onto. We want to show that β is injective. From the fact that $\hat{\beta}$ is an isomorphism, we see that $\ker\beta \subseteq \Phi(H_2)$. Take a fixed set of minimal (topological) generators $(\sigma_i)_{i\in I}$ for H_2. Then $\gamma \in \Phi(H_2)$ has a unique description, up to a permutation, as $\gamma = \prod_{i\in I}\sigma_i^2 \times \prod_{(u,v)\in K}[\sigma_u, \sigma_v]$ for some possibly infinite sets I, K.
To complete the proof we use the following lemma.

Lemma 9.13. *Assume that H_1, H_2, T_1, T_2 are as in Proposition 9.12, and let δ be σ_i^2 or $[\sigma_u, \sigma_v]$. Suppose that we have a surjective map $\varphi: H_2 \longrightarrow G$ where $G = D$ or C_4. Then there exists a group $\tilde{G}$ which is again either D or C_4 and a homomorphism $\psi: H_1 \longrightarrow \tilde{G}$ such that $\psi(\beta(\delta)) \neq 1 \in \tilde{G}$ if and only if $\varphi(\delta) \neq 1 \in G$. Moreover $\tilde{G}$ and the homomorphism ψ depend only on G and on the fields F_1 and F_2, but not on δ.*

Proof. (1) Assume first that $G = C_4$. Since H_2 is lifted, there exist an H_2-extension K_2/F_2 and a C_4^u-extension L_2 of F_2 with $F_2 \longrightarrow F_2(\sqrt{u}) \longrightarrow L_2 \longrightarrow K_2$. Since $\dot{F}_1/T_1 \cong \dot{F}_2/T_2$, there exist $a \in \dot{F}_1, s_2 \in T_2$ such that $u = as_2$. Let $\delta = \sigma^2$, which is the only case to be considered when $G = C_4$. Then $\varphi(\sigma^2) \neq 1 \in \mathrm{Gal}(L_2/F_2)$ if and only if $\varphi(\sigma)$ has order 4. Thus $\varphi(\sigma^2) \neq 1$ if and only if $\varphi(\sigma)$ generates $\mathrm{Gal}(L_2/F_2)$. This happens precisely when $\varphi(\sigma)(\sqrt{as_2}) = -\sqrt{as_2}$. Since H_2, and thus $\varphi(\sigma)$, fixes $\sqrt{s_2}$, this is equivalent to $\varphi(\sigma)(\sqrt{a}) = -\sqrt{a}$. On the other hand, we know by Proposition 1.5 that the quaternion algebra $\left(\frac{as_2, as_2}{F_2}\right)$ splits, and this implies the existence of $s_2', t_2' \in T_2$ such that $as_2' + at_2' = 1$. From the assumption on the additive structure, this implies the existence of $s_1, t_1 \in T_1$ such that $as_1 + at_1 = 1$. Two cases are to be considered.
(1.1) If $s_1 = t_1 \bmod \dot{F}_1^2$, then there is a $C_4^{as_1}$-extension L_1 of F_1 with $F_1 \longrightarrow F_1(\sqrt{as_1}) \longrightarrow L_1$. Denoting by $\psi\colon H_1 \longrightarrow \mathrm{Gal}(L_1/F_1)$ the restriction, because H_1 fixes $\sqrt{T_1}$ we have $\psi(\beta(\sigma))(\sqrt{as_1})/\sqrt{as_1} = \psi(\beta(\sigma))(\sqrt{a})/\sqrt{a} = \varphi(\sigma)(\sqrt{a})/\sqrt{a} = -1$, showing $\psi(\beta(\delta)) \neq 1 \in C_4 = \tilde{G}$.
(1.2) If $s_1 \neq t_1 \bmod \dot{F}_1^2$, there is a D^{as_1, at_1}-extension L_1 of F_1 with $F_1 \longrightarrow F_1(\sqrt{s_1t_1}) \longrightarrow L_1$. Here $L_1/F_1(\sqrt{s_1t_1})$ is a C_4-extension. Since $\beta(\sigma) \in H_1$ fixes $F_1(\sqrt{s_1t_1})$, $\psi(\beta(\sigma))$ is in the Galois group of the latter extension, which is again a C_4-extension. We then use the same argument as in (1.1) to conclude that $\psi(\beta(\delta)) \neq 1 \in \tilde{G} = C_4$.
(2) Assume $G = D$. Again there is an H_2-extension $K_2\, of F_2$ and a D^{as_2, bs_2}-extension L_2 of F_2 with $F_2 \longrightarrow F_2(\sqrt{abs_2t_2}) \longrightarrow L_2 \longrightarrow K_2$. Since φ is surjective, there is an element $\tau \in H_2$ such that $\tau(\sqrt{abs_2t_2})/\sqrt{abs_2t_2} = -1$, or else $\varphi(H_2)$ would fix $F_2(\sqrt{abs_2t_2})$ and would be contained in a proper subgroup of $\mathrm{Gal}(L_2/F_2) \cong D$. This implies $ab \notin T_2$. Since there exist $s_2', t_2' \in T_2$ such that $as_2' + bt_2' = 1$, we also have, by the assumption on the additive structures, $as_1 + bt_1 = 1$ for some $s_1, t_1 \in T_1$. Since $ab \notin T_1$, we see that as_1, bt_1 are independent modulo $\dot{F}_1^2$, and there is a D^{as_1, bt_1}-extension L_1 of F_1 with $F_1 \longrightarrow F_1(\sqrt{abs_1t_1}) \longrightarrow L_1$. Denote by $\psi\colon H_1 \longrightarrow \mathrm{Gal}(L_1/F_1) \cong D$ the restriction map.
(2.1) Suppose $\delta = \sigma^2$ and $\varphi(\delta) \neq 1$. Then $\varphi(\sigma)$ has order 4 and must fix the quadratic extension $F_2(\sqrt{abs_2t_2})$. Then it belongs to $\mathrm{Gal}(L_2/F_2(\sqrt{abs_2t_2})) \cong C_4$. With the same arguments as in (1), we show that $\psi(\beta(\delta)) \neq 1$.
(2.2) Suppose $\delta = [\sigma_u, \sigma_v]$ and $\varphi(\delta) \neq 1$. Then none of $\varphi(\sigma_u), \varphi(\sigma_v)$ is in $\Phi(D)$ (i.e. they do not fix the biquadratic extension $F_2(\sqrt{as_2}, \sqrt{bt_2})$), and they act differently on this biquadratic extension. Since $\varphi(\sigma_u)$ (respectively $\varphi(\sigma_v)$) acts the same way on elements in $\sqrt{\dot{F}}$ as $\psi(\beta(\sigma_u))$ (respectively $\psi(\beta(\sigma_v))$, we see that $\psi(\beta(\delta)) \neq 1 \in G$.
To conclude the proof, we point out that in all cases above, we first associated $\tilde{G}$ with the given homomorphism $\varphi\colon H_2 \longrightarrow G$ and only then checked that $\varphi(\delta) \neq 1 \in G$ is equivalent to $\psi(\beta(\delta)) \neq 1 \in \tilde{G}$. □

We can now finish the proof of Proposition 9.12. Suppose $\gamma \neq 1 \in \Phi(H_2)$. Since H_2 satisfies the subdirect product condition, there exists a surjective map $\varphi: H_2 \longrightarrow G$ with $G \cong D$ or C_4 and with $\varphi(\gamma) \neq 1 \in G$. Recall that the minimal set of generators $(\sigma_i)_{i\in I}$ may be chosen in such a way that for any open set U of H_2 there are at most finitely many σ_i's outside U. (See for example [Koc, Chapter 4].) Since $\ker\varphi$ is open, we may thus assume, when working with a given φ, that $\gamma = \gamma_0 \times \gamma_1$, with $\gamma_0 = \prod_{i\in I_0} \sigma_i^2 \times \prod_{(u,v)\in K_0}[\sigma_u, \sigma_v]$, $\gamma_1 = \prod_{i\in I_1} \sigma_i^2 \times \prod_{(u,v)\in K_1}[\sigma_u, \sigma_v]$, with the following properties. The sets I_0, K_0 are finite. Any individual factor $\sigma_i^2, [\sigma_u, \sigma_v]$ of γ_0 is not in $\ker\varphi$, while any individual factor of γ_1 is in $\ker\varphi$. We may assume that $\gamma = \gamma_0$, and in particular we have only a finite number n of terms δ_i's with $\delta_i = \sigma_i^2$ or $[\sigma_u, \sigma_v]$. The Frattini group $\Phi(G) \cong C_2$ may be written $\{1, \epsilon\}$, and each $\varphi(\delta_i)$ must be ϵ, since it is not 1 by assumption. Since $\varphi(\gamma) = \epsilon^n \neq 1$, n must be odd. By Lemma 9.13, we know that there exists a group $\tilde{G}$ which is again D or C_4 and a homomorphism $\psi: H_1 \longrightarrow \tilde{G}$, such that $\varphi(\delta_i) = \epsilon \neq 1$ is equivalent to $\psi(\beta(\delta_i)) = \epsilon \neq 1$. Because n is odd, this shows that $\psi(\beta(\gamma)) \neq 1$, and therefore $\beta(\gamma) \neq 1$. This shows the injectivity of β and finishes the proof of Proposition 9.12. □

§10. Concluding Remarks

In this article we have considered all $C(I)$- and $S(I)$-orderings. These groups correspond to W-groups for p-adic fields, for odd primes p. In particular, the W-group $\mathcal{G}_p$ of $\mathbb{Q}_p$ is $C_4 \times C_4$ for $p \equiv 1(4)$ and is $C_4 \rtimes C_4$ for $p \equiv 3(4)$. It is then natural to look for a characterization of $\mathcal{G}_2$-orderings, i.e. those orderings corresponding to subgroups isomorphic to the W-group of $\mathbb{Q}_2$. This is currently under investigation [MiSm4].

For the field $\mathbb{Q}$, there is a unique C_2-ordering, which is the unique ordering on $\mathbb{Q}$. In addition there is a one-to-one correspondence between $C_4 \times C_4$-orderings on $\mathbb{Q}$ and primes $p \equiv 1(4)$, and a one-to-one correspondence between $C_4 \rtimes C_4$-orderings on $\mathbb{Q}$ and primes $p \equiv 3(4)$. In each case the correspondence is given by $T_p = \dot{\mathbb{Q}}_p^2 \cap \mathbb{Q}$. It is not hard to see that each such intersection gives rise to an H-ordering of the appropriate type. To see that every such ordering may be obtained in this way, one shows that each such ordering corresponds to a certain valuation on $\mathbb{Q}$, and the valuations on $\mathbb{Q}$ are well-known to be classified by the primes. (See e.g. [End, Theorem 1.16].)

This observation then lends itself to an alternative perspective on the Hasse-Minkowski Theorem, which states that a quadratic form defined over $\mathbb{Q}$ is isotropic over $\mathbb{Q}$ if and only if it is isotropic over each $\mathbb{Q}_p$, including $\mathbb{Q}_\infty$, the real numbers. Using Hilbert's reciprocity law, one can prove that a ternary quadratic form is isotropic over $\mathbb{Q}$ if and only if it is isotropic over all but one of these fields. Thus we see that a ternary quadratic form over $\mathbb{Q}$ is isotropic if and only if it is isotropic with respect to all C_2-, $(C_4 \times C_4)$-, and $(C_4 \rtimes C_4)$-orderings on $\mathbb{Q}$.

We point out that the case of a ternary quadratic form over $\mathbb{Q}$, together with the clever use of Dirichlet's theorem on the existence of an infinite number of primes in an arithmetic progression, where first term and increment are

relatively prime, are the main ingredients of a proof of the full Hasse-Minkowski theorem over $\mathbb{Q}$. For a very nice exposition of the Hasse-Minkowski theorem over $\mathbb{Q}$, see [BS]. See also [L1, Chapter 6, Exercise 22].
It is easy however, to find a quaternary quadratic form φ over $\mathbb{Q}$ such that φ is isotropic over all $\mathbb{Q}_p$, p is an odd prime, and $\mathbb{Q}_\infty = \mathbb{R}$ but φ is anisotropic over $\mathbb{Q}_2$. For example $\varphi = X_1^2 + X_2^2 - 7X_3^2 - 31X_4^2$ and $\psi = X_1^2 + X_2^2 + X_3^2 - 7X_4^2$ are such forms.
In a subsequent paper we will present several applications of this theory to different kinds of local-global principles for quadratic forms. In order to get a sense of what can be done in this direction, we show below an example of a simple situation in which our theory applies.
Consider a field F. Recall that a $C(\emptyset)$-ordering T on F is an index 2 multiplicative subgroup of $\dot F/\dot F^2$ containing -1. Additively speaking, it is a hyperplane containing -1 in the $\mathbb{F}_2$-vector space $\dot F/\dot F^2$. If $f \in \dot F \setminus (\dot F^2 \cup -\dot F^2)$ and if V is any subspace of $\dot F/\dot F^2$ such that $\dot F/\dot F^2 = \operatorname{Span}\{f, -1\} \oplus V$, then $T := \operatorname{Span}\{-1\} + V$ is a $C(\emptyset)$-ordering not containing f. Then the next lemma follows immediately.

LEMMA 10.1. *Let $C_0(F)$ denote the set of $C(\emptyset)$-orderings of F. Then $C_0(F) = \emptyset$ if and only if $\dot F = \dot F^2 \cup -\dot F^2$, and in general,*

$$\bigcap_{T \in C_0(F)} T = \dot F^2 \cup -\dot F^2.$$

To every $C(\emptyset)$-ordering T we associate a fixed closure F_T of F in the quadratic closure of F. Denote by $\langle\langle a_1, \dots, a_n\rangle\rangle$ the Pfister form $\langle 1, -a_1\rangle \otimes \dots \otimes \langle 1, -a_n\rangle$. (For the basic theory of Pfister forms see e.g. [L1, Chapter 10] or [Sc, Chapter 4]. Observe that both Lam and Scharlau denote by $\langle\langle a_1, \dots, a_n\rangle\rangle$ the Pfister form $\langle 1, a_1\rangle \otimes \dots \otimes \langle 1, a_n\rangle$.) Then we have the following.

PROPOSITION 10.2. *Assume $C_0(F) \neq \emptyset$ and let $\varphi : W(F) \longrightarrow \prod_{T \in C_0(F)} W(F_T)$ denote the map induced by the inclusions $F \longrightarrow F_T$. Then $\operatorname{Ker}\varphi = I^2F + 2W(F)$ where IF denotes the fundamental ideal of $W(F)$.*

Proof. For $T \in C_0(F)$ we have $\dot F/T = \{\bar 1, \bar f\}$ for a certain $f \in \dot F$, and it is easy to see that $W(F_T) \cong C_2[\epsilon]/\epsilon^2$ and that the isomorphism, which we call π, is defined by $\pi(\langle\bar 1\rangle) = 1, \pi(\langle\bar f\rangle) = 1 + \epsilon$. If $a, b \in \dot F$ then the possibilities for $\bar a, \bar b$ are (1) $\bar a = 1$ or $\bar b = 1$, or (2) $\bar a = \bar b = \bar f$. In any case the image in $W(F_T)$ of the 2-fold Pfister form $\langle\langle a, b\rangle\rangle$ is in $2W(F_T) = 0$, and we have shown the inclusion $I^2F + 2W(F) \subseteq \operatorname{Ker}\varphi$.
Take $q \in \operatorname{Ker}\varphi$. Then $q \in IF$, because any odd-dimensional form is nonzero in $W(F_T)$. But it is known ([Pf, p. 122, Kor. to Satz 13]) that any element q of IF may be written $q = \langle\langle u\rangle\rangle + q_1$, with $q_1 \in I^2F$. Since $q \in \operatorname{Ker}\varphi$, and $I^2F \subset \operatorname{Ker}\varphi$, we deduce $\langle\langle u\rangle\rangle \in \operatorname{Ker}\varphi$. The latter is equivalent to $u \in T$ for every T, meaning $u \in \dot F^2 \cup -\dot F^2$, or in other words $\langle\langle u\rangle\rangle = 0$ or 2 in $W(F)$. $\square$

Recall that a field F is said to have virtual cohomological dimension n, denoted $\mathrm{vcd}(F) = n$, if $H^d(\mathrm{Gal}(F(2))/F(\sqrt{-1})), \mu_2) = 0$ for $d > n$, and $H^n(\mathrm{Gal}(F(2))/F\ (\sqrt{-1})), \mu_2) \neq 0$. (If we also considered the case of $\mathbb{F}_p, p$ an odd prime, as coefficients of the cohomology groups of absolute Galois groups, it would be more appropriate to say that F as above has a virtual 2-cohomological dimension equal to n.) If $\mathrm{vcd}(F) \leq 1$, then I^2F is torsion-free. To see this, observe first that $\mathrm{vcd}(F) \leq 1$ implies each binary quadratic form over $F(\sqrt{-1})$ is universal. Then use [L1, Chapter 11, Theorem 1.8 and Exercise 20] to conclude that I^2F is torsion free. An example of a formally real field F with $vcd(F) = 1$ is $F = \mathbb{R}(X)$. We have the following local-global principle:

Theorem 10.3. *Let F be a field with $\mathrm{vcd}(F) \leq 1$. Let $D_0(F)$ (resp. $C_0(F)$, $S_0(F)$) denote the set of usual orderings $X(F)$ (resp. $C(\emptyset)$-orderings, $S(\emptyset)$-orderings) of F. Then*

$$\Lambda\colon W(F) \longrightarrow \prod_{T \in D_0(F) \cup C_0(F) \cup S_0(F)} W(F_T)$$

is injective. If F is formally real, we may drop $S_0(F)$. (If not, we may drop $D_0(F)$.)

Proof. It is clear that a form $q \in \operatorname{Ker}\Lambda$ is in IF, and thus can be written $q = \langle\langle a\rangle\rangle + q_2$ with $q_2 \in I^2F$. By Pfister's Local-Global Principle [L1, Chapter 8, §4], q is torsion and it is therefore the case for $\langle\langle a\rangle\rangle$ and q_2. (It is trivial when $D_0(F) = \emptyset$, and if not, we use the fact that the signature $\hat{q}$ of q is 0 and that $\hat{q_2} \equiv 0(\mathrm{mod} 4)$.)
Since I^2F is torsion-free, one has $q_2 = 0$, and $q = \langle\langle a\rangle\rangle$. Since q vanishes on $C_0(F)$, by Proposition 10.2 we have $a \in \dot{F}^2 \cup -\dot{F}^2$. (If $C_0(F) = \emptyset$, this condition is trivially satisfied.) If the level $s(F)$ is 1, our proof is completed. Otherwise $D_0(F) \cup S_0(F) \neq \emptyset$, which shows that $q \neq \langle\langle -1\rangle\rangle$. Thus $q = \langle\langle 1\rangle\rangle = 0$. □

Remark 10.4. In this case we even have a strong Hasse principle, that is a local-global principle for detecting whether a quadratic form is anisotropic rather than just hyperbolic. Indeed, the fact that each ternary form over $F(\sqrt{-1})$ is isotropic and [ELP, Theorem F] give us the strong Hasse principle for forms of rank greater than or equal to 3. Then the use of $C_0(F), S_0(F)$ and $D_0(F)$ provides the result for rank 2 forms.

Finally let us point out that our results are closely related to some ideas in birational anabelian Grothendieck geometry. In particular there is a close connection between ideas explored in this paper and the work of Bogomolov, Tschinkel and Pop ([Bo], [BoT], [Po1], and [Po2]; see also Koenigsmann's thesis [K1] and paper [K2]). Roughly speaking, they establish that for certain fields K the isomorphy type of K, modulo purely inseparable extensions of K, is functorially encoded in the maximal pro-p-quotient of the absolute Galois group $\tilde{G} := \mathrm{Gal}(\bar{K}/K), \mathrm{char}\, K \neq p$. In fact Bogomolov in [Bo] and also Pop in lectures at MSRI in the fall of 1999, considered smaller Galois groups than the

Galois group defined above, namely the maximal pro-p-quotient of the group $\tilde{G}/[[\tilde{G},\tilde{G}],\tilde{G}]$. In this paper we consider $p = 2$, because of the connections with quadratic forms. It is expected however that a substantial part of our results can be extended to any prime p provided that the base field F contains a primitive pth root of unity. We allow F to be any field with $\operatorname{char} F \neq 2$, and we are concerned with even smaller Galois groups than were considered by Bogomolov and Pop. Of course in this more general setting we cannot obtain as precise results as Bogomolov and Pop, but we do get some interesting information about the additive properties of multiplicative subgroups of fields. It would be very interesting to investigate further relationships between our work and the quoted work of Bogomolov, Pop and Tschinkel.

References

[AKMi] A. Adem, D. Karagueuzian and J. Mináč, *On the cohomology of Galois groups determined by Witt rings*, Adv. Math. **148** (1999), 105–160.

[ArTa] E. Artin and J. Tate, *Class field theory*, Benjamin, 1967.

[AEJ] J. Arason, R. Elman and B. Jacob, *Rigid elements, valuations, and realization of Witt rings*, J. Algebra **110** (1987), 449–467.

[ArSc1] E. Artin and O. Schreier, *Algebraische Konstruktion reeller Körper*, Abh. Math. Sem. Univ. Hamburg **5** (1927), 85–99.

[ArSc2] E. Artin and O. Schreier, *Eine Kennzeichnung der reell abgeschlossenen Körper*, Abh. Math. Sem. Univ. Hamburg **5** (1927), 225–231.

[Be] E. Becker, *Euklidische Körper und euklidische Hüllen von Körpern*, J. reine angew. Math. **268/269** (1974), 41–52.

[BCR] J. Bochnak, M. Coste, M. -F. Roy, *Real Algebraic Geometry*, Ergeb. Math., vol. 36, Springer-Verlag, Berlin, Heidelberg, 1998.

[Bo] F. A. Bogomolov, *On two conjectures in birational algebraic geometry*, Algebraic Geometry and Analytic Geometry (A. Fujiki et. al., eds.), ICM-90 Satellige Conference Proceedings, Springer-Verlag, Tokyo, 1991.

[BoT] F. A. Bogomolov and Y. Tschinkel, *Commuting elements of Galois groups of function fields*, Motives, polylogarithms and Hodge theory, Part I (Irvine, CA, 1998), Int. Press Lect. Ser. **3, I,** (2002), Int. Press, Somerville, MA, 75–120.

[BS] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Academic Press, Inc., 1966.

[BEK] Bredikhin, Ershov, and Kal'nei, *Fields with two linear orderings*, Translated from Mat. Zametki Vol. 7 (1970), 525–536, Math Notes, Institute of Mathematics, Siberian Branch, Academy of Sciences of hte USSR (1970), 319–325.

[Br] L. Bröcker, *Zur Theorie der quadratischen Formen über formal reellen Körpern*, Math. Ann. **210** (1974), 233–256.

[Cr1] T. Craven, *Characterizing reduced Witt rings of fields*, J. Algebra **53** (1978), 68–77.

[Cr2] ———, *Fields maximal with respect to a set of orderings*, J. Algebra **115** (1988), 200–218.

[CrSm] T. Craven and T. Smith, *Formally real fields from a Galois theoretic perspective*, J. Pure Appl. Alg. **145** (2000), 19–36.

[Ef] I. Efrat, *Construction of valuations from K - theory*, Mathematical Research Letters **6** (1999), 335–343.

[ELP] R. Elman, T.-Y. Lam and A. Prestel, *On some Hasse Principles over formally real fields*, Math. Z. **134** (1973), 291–301.

[End] O. Endler, *Valuation theory*, Springer-Verlag, New York, 1972.

[EnKo] A. J. Engler and J. Koenigsmann, *Abelian subgroups of pro-p Galois groups*, Trans. A.M.S. **350** (1998), 2473–2485.

[Fr] A. Fröhlich, *Orthogonal representations of Galois groups, Stiefel-Whitney classes and Hasse-Witt invariants*, J. Reine Angew. Math. **360** (1985), 84–123.

[GM] W. Gao and J. Mináč, *Milnor conjecture and Galois theory I*, Fields Institute Communications **16** (1997), 95–110.

[Gr] G. Grätzer, *Universal Algebra*, Second Edition, Springer-Verlag, 1979.

[Har] D. Harbater, *Galois coverings of the arithmetic line*, Number Theory, New York 1984-1985, Lecture Notes in Mathematics **1240** (1987), Springer-Verlag, New York, 165–195.

[HaVöl] D. Haran and H. Völklein, *Galois groups over complete valued fields*, Israel Journal of Mathematics **93** (1996), 9–27.

[Jo] D. L. Johnson, *Non-nilpotent elements of cohomology rings*, Math. Zeit. **112** (1969), 364–374.

[Ki] I. Kiming, *Explicit classifications of some 2-extensions of a field of characteristic different from* 2, Can. J. of Math. **42** (1990), 825–855.

[Koc] H. Koch, *Galoische Theorie der p-Erweiterungen*, Springer-Verlag, New York, 1970.

[K1] J. Koenigsmann, *Half-ordered fields*, Dissertation Thesis, Universität Konstanz, Konstanz, 1993.

[K2] ———, *From p-rigid elements to valuations (with a Galois characterization of p-adic fields)*, with appendix by F. Pop, J. Reine Angew. Math. **465** (1995), 165–182.

[L1] T. Y. Lam, *The algebraic theory of quadratic forms*, Benjamin/Cummings Publishing Co., Reading, Mass., 1980.

[L2] ———, *Orderings, valuations and quadratic forms*, Conference Board of the Mathematical Sciences No. 52, Amer. Math. Soc., Providence, RI, 1983.

[LaSm] T. Y. Lam and T. Smith, *On the Clifford-Littlewood-Eckmann groups: A new look at periodicity mod 8*, Rocky Mtn. J. Math. **19** (1989), 749–786.

[Mac] S. MacLane, *Subfields and automorphism groups of p-adic fields*, Ann. of Math. **40** (1939), 423–442.

[Ma] M. Marshall, *Abstract Witt rings*, Queen's Papers in Pure and Appl. Math., Vol. 57, Queen's University, Kingston, Ontario, 1980.

[Me] A. Merkurjev, *On the norm residue symbol of degree two*, English translation: Soviet Math. Dokladi **24** (1981), 546-551, Dokladi Akad. Nauk. SSSR **261** (1981), 542-547.

[Mi] J. Mináč, *Elementary 2-abelian subgroups of W-groups,*, unpublished manuscript (1987).

[MiSm1] J. Mináč and T. Smith, *A characterization of C-fields via Galois groups*, J. Algebra **137** (1991), 1–11.

[MiSm2] ———, *Decomposition of Witt rings and Galois groups*, Canad. J. Math **47** (1995), 1274–1289.

[MiSm3] ———, *W-Groups under quadratic extensions of fields*, Canad. J. Math. **52** (2000), 833–848.

[MiSm4] ———, *Formally dyadic fields*, in preparation.

[MiSp1] J. Mináč and M. Spira, *Formally real fields, pythagorean fields, C-fields and W-groups*, Math. Z. **205** (1990), 519–530.

[MiSp2] ———, *Witt rings and Galois groups*, Ann. of Math. **144** (1996), 35–60.

[Mo] S. Morris, *Pontryagin duality and the structure of locally compact abelian groups*, London Mathematical Society Lecture Note Series, vol. 29, Cambridge University Press, Cambridge, 1980.

[Pf] A. Pfister, *Quadratische Formen in beliebigen Körpern*, Invent. Math. **1** (1966), 116–132.

[Po1] F. Pop, *On Grothendieck's conjecture of birational anabelian geometry*, Ann. Math. **139** (1994), 145–182.

[Po2] ———, *Glimpses of Grothendieck's anabelian geometry*, Geometric Galois Action 1: Around Grothendieck's Esquisse d'un Programme (L. Schneps and P. Lochak, eds.), Lecture Notes in Mathematics, vol. 242, Springer-Verlag, Berlin, 1997, pp. 113–126.

[Ri] P. Ribenboim, *Théorie des Valuations*, Université de Montreal, Les Presses de l'Université Montreal, 1965.

[RZ] L. Ribes and P. Zalesskii, *Profinite Groups*, Springer-Verlag, Berlin, 2000.

[Sc] W. Scharlau, *Quadratic and Hermitian Forms*, Springer-Verlag, 1985.

[Sch] N. Schwartz, *Signatures and Real Closures of Fields*, Publications de l'Université Paris VII **32** (1990), 65–78.

[S] E. Sperner, *Die Ordnungsfunktionen einer Geometrie*, Math. Ann. **121** (1949), 107–130.

[Sp] M. Spira, *Witt rings and Galois groups*, Ph.D. Thesis, University of California, Berkeley, California, 1987.

[Vo] V. Voevodsky, Motivic cohomology with $\mathbb{Z}/2$-coefficients, Publ. Math. Inst. Hautes Études Sci. **98** (2003), 59–104.

[Wa] R. Ware, *Valuation rings and rigid elements in fields*, Canad. J. Math. **33** (1981), 1338–1355.

Louis Mahé
IRMAR, Campus de Beaulieu
F-35042 Rennes-Cedex
France
louis.mahe@univ-rennes1.fr

Ján Mináč
Dept. of Mathematics
Univ. of Western Ontario
London, Ontario N6A 5B7
Canada
minac@uwo.ca

Tara L. Smith
Dept. of Mathematical Sciences
Univ. of Cincinnati
Cincinnati Ohio 45221-0025
U.S.A.
tsmith@math.uc.edu

On the Values of Equivariant Zeta Functions of Curves over Finite Fields

David Burns

Received: November 12, 2003

Communicated by Stephen Lichtenbaum

Abstract. Let K/k be a finite Galois extension of global function fields of characteristic p. Let C_K denote the smooth projective curve that has function field K and set $G := \mathrm{Gal}(K/k)$. We conjecture a formula for the leading term in the Taylor expansion at zero of the G-equivariant truncated Artin L-functions of K/k in terms of the Weil-étale cohomology of $\mathbb{G}_m$ on the corresponding open subschemes of C_K. We then prove the ℓ-primary component of this conjecture for all primes ℓ for which either $\ell \neq p$ or the relative algebraic K-group $K_0(\mathbb{Z}_\ell[G], \mathbb{Q}_\ell)$ is torsion-free. In the remainder of the manuscript we show that this result has the following consequences for K/k: if $p \nmid |G|$, then refined versions of all of Chinburg's 'Ω-Conjectures' in Galois module theory are valid; if the torsion subgroup of $K^\times$ is a cohomologically-trivial G-module, then Gross's conjectural 'refined class number formula' is valid; if K/k satisfies a certain natural class-field theoretical condition, then Tate's recent refinement of Gross's conjecture is valid.

2000 Mathematics Subject Classification: Primary 11G40; Secondary 11R65 19A31 19B28

1. Introduction

Let K/k be a finite Galois extension of global function fields of characteristic p. Let C_K be the unique geometrically irreducible smooth projective curve which has function field equal to K and set $G := \mathrm{Gal}(K/k)$. For each finite non-empty set S of places of k that contains all places which ramify in K/k, we write $\mathcal{O}_{K,S}$ for the subring of K consisting of those elements that are integral at all places of K which do not lie above an element of S and we set $U_{K,S} := \mathrm{Spec}(\mathcal{O}_{K,S})$. With R denoting either $\mathbb{Z}$ or $\mathbb{Z}_\ell$ for some prime ℓ and E an extension field of the field of fractions of R, we write $K_0(R[G], E)$ for the relative algebraic K-group defined by Swan in [46].

In §2 we formulate a conjectural equality $\mathrm{C}(K/k)$ between an element of $K_0(\mathbb{Z}[G],\mathbb{R})$ constructed from the leading term in the Taylor expansion at $s=0$ of the G-equivariant Artin L-function of $U_{K,S}$ and the refined Euler characteristic of a pair comprising the Weil-étale cohomology of $\mathbb{G}_m$ on $U_{K,S}$ (considered as an object of an appropriate derived category) and a natural logarithmic regulator mapping. This conjecture is motivated both by the general approach described by Lichtenbaum in [40, §8] and also by analogy to a special case of the equivariant refinement of the Tamagawa Number Conjecture of Bloch and Kato (which was formulated by Flach and the present author in [13]). The equality $\mathrm{C}(K/k)$ can be naturally reinterpreted as a conjectural equality in $K_0(\mathbb{Z}[G],\mathbb{Q})$ involving the leading term at $t=1$ of the G-equivariant Zeta-function of $U_{K,S}$ and in §3 we shall prove the validity, resp. the validity modulo torsion, of the projection of the latter conjectural equality to $K_0(\mathbb{Z}_\ell[G],\mathbb{Q}_\ell)$ for all primes $\ell\neq p$, resp. for $\ell=p$ (this is Theorem 3.1). If $\ell\neq p$, then our proof combines Grothendieck's formula for the Zeta-function in terms of the action of frobenius on ℓ-adic cohomology together with a non-commutative generalisation of a purely algebraic observation of Kato in [35] (this result may itself be of some independent interest) and an explicit computation of certain Bockstein homomorphisms in ℓ-adic cohomology. In the case that $\ell=p$ we are able to deduce our result from Bae's verification of the 'Strong-Stark Conjecture' [3] which in turn relies upon results of Milne [43] concerning relations between Zeta-functions and p-adic cohomology.

In the remainder of the manuscript we show that $\mathrm{C}(K/k)$ provides a universal approach to the study of several well known conjectures. A key ingredient in all of our results in this direction is a previous observation of Flach and the present author which allows an interpretation in terms of Weil-étale cohomology of the canonical extension classes defined using class field theory by Tate in [49].

In §4 we consider connections between $\mathrm{C}(K/k)$ and the central conjectures of classical Galois module theory. To be specific, we prove that $\mathrm{C}(K/k)$ implies the validity for K/k of a strong refinement of the '$\Omega(3)$-Conjecture' formulated by Chinburg in [18, §4.2]. Taken in conjunction with Theorem 3.1 this result allows us to deduce that if $K_0(\mathbb{Z}_p[G],\mathbb{Q}_p)$ is torsion-free, resp. $p\nmid|G|$, then the $\Omega(3)$-Conjecture, resp. the $\Omega(1)$-, $\Omega(2)$- and $\Omega(3)$-Conjectures, formulated by Chinburg in loc. cit., are valid for K/k. This is a strong refinement of previous results in this area.

We assume henceforth that G is abelian. In this case Gross has conjectured a 'refined class number formula' which expresses an explicit congruence relation between the values at $s=0$ of the Dirichlet L-functions associated to K/k [31]. This conjecture has attracted much attention and indeed Tate has recently formulated a strong refinement in the case that G is cyclic [51]. However, whilst the conjecture of Gross has already been verified in several interesting cases [31, 1, 47, 37, 39], much of this evidence is obtained either by careful analysis of special cases or by induction on $|G|$ and, as yet, no coherent overview of or systematic approach to these conjectures of Gross and Tate has emerged. In contrast, in §5 we shall use the general approach of algebraic height pairings developed by Nekovář in [44] to interpret the integral regulator mapping of Gross as a Bockstein homomorphism in Weil-étale cohomology, and we shall also apply this interpretation to prove that if the torsion subgroup μ_K of $K^\times$ is a cohomologically-trivial G-module (a condition that is automatically satisfied if, for example, $|\mu_K|$ is coprime to $|G|$), then $\mathrm{C}(K/k)$ implies the validity of Gross's conjecture for K/k. Under a certain natural class-field theoretical assumption on K/k we shall also show (in §6) that $\mathrm{C}(K/k)$ implies the validity of Tate's refinement of Gross's conjecture. When combined with Theorem 3.1 (and earlier results of Tan concerning p-extensions) these observations allow us to deduce the validity of Gross's conjecture for all extensions K/k for which μ_K is a cohomologically-trivial G-module and also to prove the validity of Tate's refinement of Gross's conjecture for a large family of extensions.

A further development of the approach used here should allow the removal of any hypothesis on μ_K (indeed, in special cases this is already achieved in the present manuscript). However, even at this stage, our results constitute a strong improvement of previous results in this area and also provide a philosophical underpinning to the conjectures of Gross and Tate that was not hitherto apparent. Indeed, the approach presented here leads to the formulation of natural analogues of these conjectures concerning the values of (higher order) derivatives of L-functions that vanish at $s = 0$. These developments have in turn led to a proof of Tate's conjecture under the hypothesis only that $|\mu_K|$ is coprime to $|G|$ and have also provided new insight into Gross's 'refined p-adic abelian Stark conjecture' as well as several other conjectures due, for example, to Rubin, to Darmon, to Popescu and to Tan. For more details of this aspect of the theory the reader is referred to [10, 34].

Acknowledgements. The author is very grateful to J. Tate and B. H. Gross for their encouragement concerning this project and for their hospitality during his visits to the Universities of Texas at Austin and Harvard respectively. In addition, he is most grateful to M. Kurihara for his hospitality during the author's visit to the Tokyo Metropolitan University, where a portion of this project was completed. The author is also grateful to J. Nekovář for a number of very helpful discussions.

2. The leading term conjecture

2.1. Relative Algebraic K-Theory. In this subsection we quickly recall certain useful constructions in algebraic K-theory.

If Λ is any ring, then all modules are to be understood as left modules. We write $\zeta(\Lambda)$ for the centre of Λ, $K_1(\Lambda)$ for the Whitehead group of Λ and $K_0(\Lambda)$ for the Grothendieck group of the category of finitely generated projective Λ-modules. We also write $\mathcal{D}(\Lambda)$ for the derived category of complexes of Λ-modules with only finitely many non-zero cohomology groups, and we let $\mathcal{D}^{\mathrm{fpd}}(\Lambda)$, resp. $\mathcal{D}^{\mathrm{perf}}(\Lambda)$,denote the full triangulated subcategory of $\mathcal{D}(\Lambda)$

consisting of those complexes that are quasi-isomorphic to a bounded complex of projective Λ-modules, resp. to a bounded complex of finitely generated projective Λ-modules.

We let R denote either $\mathbb{Z}$ or $\mathbb{Z}_\ell$ for some prime ℓ, E and F denote extension fields of the field of fractions of R and we fix a finite group G. For finitely generated $E[G]$-modules V and W we write $\mathrm{Is}_{E[G]}(V,W)$ for the set of $E[G]$-module isomorphisms from V to W. The relative algebraic K-group $K_0(R[G],E)$ is an abelian group with generators (X,ϕ,Y), where X,Y are finitely generated projective $R[G]$-modules and ϕ is an element of $\mathrm{Is}_{E[G]}(X\otimes_R E, Y\otimes_R E)$. For the defining relations we refer to [46, p. 215]. We systematically use the following facts: there is a long exact sequence of relative K-theory (cf. [46, Th. 15.5])

$$K_1(R[G]) \to K_1(E[G]) \xrightarrow{\partial^1_{R[G],E}} K_0(R[G],E) \xrightarrow{\partial^0_{R[G],E}} K_0(R[G]) \to K_0(E[G]);$$

if $E \subseteq F$, then there is a natural injective 'inclusion' homomorphism $K_0(R[G],E) \subseteq K_0(R[G],F)$; for each rational prime ℓ the assignment $(X,\phi,Y) \mapsto (X\otimes_{\mathbb{Z}}\mathbb{Z}_\ell, \phi\otimes_{\mathbb{Q}}\mathbb{Q}_\ell, Y\otimes_{\mathbb{Z}}\mathbb{Z}_\ell)$ induces a homomorphism

$$\rho_\ell : K_0(\mathbb{Z}[G],\mathbb{Q}) \to K_0(\mathbb{Z}_\ell[G],\mathbb{Q}_\ell)$$

and the product of these homomorphisms over all primes ℓ induces an isomorphism (cf. the discussion following [26, (49.12)])

$$\prod_\ell \rho_\ell : K_0(\mathbb{Z}[G],\mathbb{Q}) \cong \bigoplus_\ell K_0(\mathbb{Z}_\ell[G],\mathbb{Q}_\ell). \tag{1}$$

Let A be a finite dimensional central simple F-algebra, F' an extension of F which splits A and e an indecomposable idempotent of $A\otimes_F F'$. If V is any finitely generated A-module and $\phi \in \mathrm{End}_A(V)$, then the 'reduced determinant' of ϕ is defined by setting $\mathrm{detred}_A(\phi) := \det_{F'}(\phi\otimes_F \mathrm{id}_{F'}|e(V\otimes_F F'))$. This is an element of F which is independent of the choices of F' and e. This construction extends to finite-dimensional semi-simple F-algebras in the obvious way. In particular, the group $K_1(E[G])$ is generated by symbols of the form $[\phi]$ with $\phi \in \mathrm{Aut}_{E[G]}(V)$ and the assignment $[\phi] \mapsto \mathrm{detred}_{E[G]}(\phi)$ induces a well-defined injective 'reduced norm' homomorphism $\mathrm{nr}_{E[G]} : K_1(E[G]) \to \zeta(E[G])^\times$ [26, §45A]. For each ℓ the map $\mathrm{nr}_{\mathbb{Q}_\ell[G]}$ is bijective and so there exists a unique homomorphism $\delta_\ell : \zeta(\mathbb{Q}_\ell[G])^\times \to K_0(\mathbb{Z}_\ell[G],\mathbb{Q}_\ell)$ with $\partial^1_{\mathbb{Z}_\ell[G],\mathbb{Q}_\ell} = \delta_\ell \circ \mathrm{nr}_{\mathbb{Q}_\ell[G]}$. (When we need to be more precise we write $\delta_{G,\ell}$ rather than δ_ℓ.) The map $\mathrm{nr}_{\mathbb{R}[G]}$ is not in general surjective, but nevertheless there exists a canonical 'extended boundary' homomorphism $\delta : \zeta(\mathbb{R}[G])^\times \to K_0(\mathbb{Z}[G],\mathbb{R})$ which satisfies $\partial^1_{\mathbb{Z}[G],\mathbb{R}} = \delta\circ\mathrm{nr}_{\mathbb{R}[G]}$ and is such that $\zeta(\mathbb{Q}[G])^\times$ is the full pre-image of $K_0(\mathbb{Z}[G],\mathbb{Q})$ under δ (cf. [13, Lem. 9]).

The map $\mathrm{nr}_{E[G]}$ induces an equivalence relation '$\sim$' on each set $\mathrm{Is}_{E[G]}(V,W)$ in the following way: $\phi \sim \phi'$ if $\mathrm{nr}_{E[G]}([\phi'\circ\phi^{-1}]) = 1$. In the sequel we shall often not distinguish between an element of $\mathrm{Is}_{E[G]}(V,W)$ and its associated equivalence class in $\mathrm{Is}_{E[G]}(V,W)/\sim$.

For each $\mathbb{Z}$-graded module $C^\cdot$ we write C^{all}, C^- and C^+ for the direct sum of C^i as i ranges over all, all odd and all even integers respectively.

An 'E-trivialisation' of an object $C^{\cdot}$ of $\mathcal{D}^{\mathrm{perf}}(R[G])$ is an element τ of $\mathrm{Is}_{E[G]}(H^+(C^{\cdot})\otimes_R E, H^-(C^{\cdot})\otimes_R E)/\sim$. In [9] it is shown that a variant of the classical construction of Whitehead torsion allows one to associate to each such pair $(C^{\cdot},\tau)$ a canonical 'refined Euler characteristic' element $\chi_{R[G],E}(C^{\cdot},\tau)$ which belongs to $K_0(R[G],E)$ and has image under $\partial^0_{R[G],E}$ equal to the Euler characteristic of $C^{\cdot}$ in $K_0(R[G])$. Further details of this construction are recalled in the Appendix.

In the sequel we shall use the following notation and conventions. We abbreviate 'cohomologically-trivial' to 'c-t', '$\chi_{\mathbb{Z}[G],\mathbb{R}}$' to '$\chi$', '$\chi_{\mathbb{Z}[G],\mathbb{Q}}$' to '$\chi_{\mathbb{Q}}$' and '$\chi_{\mathbb{Z}_\ell[G],\mathbb{Q}_\ell}$' to '$\chi_\ell$' (or to '$\chi_{G,\ell}$' when we need to be more precise); if X is any scheme of finite type over the finite field $\mathbb{F}_p$ of cardinality p and $\mathcal{F}$ is any étale (pro-) sheaf, resp. Weil-étale sheaf, on X, then we abbreviate $R\Gamma(X_{\text{ét}},\mathcal{F})$, resp. $R\Gamma(X_{\mathrm{Weil-ét}},\mathcal{F})$ to $R\Gamma(X,\mathcal{F})$, resp. $R\Gamma_{\mathcal{W}}(X,\mathcal{F})$, and we also use similar abbreviations on cohomology; for any commutative ring Λ we write $x\mapsto x^{\#}$ for the Λ-linear involution of $\zeta(\Lambda[G])$ that is induced by setting $g^{\#}:=g^{-1}$ for each $g\in G$; for any group H and any H-module M we write M^H, resp. M_H, for the maximal submodule, resp. quotient, of M upon which H acts trivially; for any abelian group A we let A_{tors} denote its torsion subgroup; unless explicitly indicated otherwise, all tensor products and exterior powers are to be considered as taken in the category of abelian groups.

2.2. Formulation of the conjecture. We assume henceforth that S is a finite non-empty set of places of k containing all places which ramify in K/k. We let $\mathrm{Irr}_{\mathbb{C}}(G)$ denote the set of irreducible finite dimensional complex characters of G. For each $\chi\in\mathrm{Irr}_{\mathbb{C}}(G)$ we write $L_S(\chi,s)$ for the associated S-truncated Artin L-function and $L_S^*(\chi,0)$ for the leading term in the Taylor expansion of $L_S(\chi,s)$ at $s=0$. Recalling that $\zeta(\mathbb{C}[G])$ identifies with $\prod_{\mathrm{Irr}_{\mathbb{C}}(G)}\mathbb{C}$, we define a $\zeta(\mathbb{C}[G])$-valued meromorphic function of a complex variable s by setting

$$\theta_{K/k,S}(s):=(L_S(\chi,s))_{\chi\in\mathrm{Irr}_{\mathbb{C}}(G)}.$$

The leading term $\theta^*_{K/k,S}(0)$ in the Taylor expansion of $\theta_{K/k,S}(s)$ at $s=0$ is equal to $(L_S^*(\chi,0))_{\chi\in\mathrm{Irr}(G)}$ and hence belongs to $\zeta(\mathbb{R}[G])^\times$. In this subsection we follow the philosophy introduced by Lichtenbaum in [40] to formulate a conjectural description of $\delta(\theta^*_{K/k,S}(0)^{\#})$ in terms of Weil-étale cohomology.

For any intermediate field F of K/k we write $Y_{F,S}$ for the free abelian group on the set of places $S(F)$ of F which lie above those in S and $X_{F,S}$ for the kernel of the homomorphism $Y_{F,S}\to\mathbb{Z}$ that sends each element of $S(F)$ to 1. We write $\mathcal{O}_{F,S}$ for the ring of $S(F)$-integers in F and $\mathcal{O}^\times_{F,S}$ for its unit group. We also set $U_{F,S}:=\mathrm{Spec}(\mathcal{O}_{F,S})$ and $A_{F,S}:=\mathrm{Pic}(\mathcal{O}_{F,S})$.

Lemma 1.

i) *Let $j: U_{K,S}\to C_K$ denote the natural open immersion. Then there exists a canonical isomorphism in $\mathcal{D}(\mathbb{Z}[G])$ of the form*

$$R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)\cong R\,\mathrm{Hom}_{\mathbb{Z}}(R\Gamma_{\mathcal{W}}(C_K,j_!\mathbb{Z}),\mathbb{Z}[-2]).$$

ii) *There exists a natural distinguished triangle in* $\mathcal{D}(\mathbb{Z}[G])$ *of the form*

$$X_{K,S}\otimes\mathbb{Q}[-2]\to R\Gamma(U_{K,S},\mathbb{G}_m)\to R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)\to X_{K,S}\otimes\mathbb{Q}[-1]$$

where the map induced on cohomology (in degree 2*) by the first morphism is the composite of the projection* $X_{K,S}\otimes\mathbb{Q}\to X_{K,S}\otimes\mathbb{Q}/\mathbb{Z}$ *and the canonical identification* $X_{K,S}\otimes\mathbb{Q}/\mathbb{Z}\cong H^2(U_{K,S},\mathbb{G}_m)$*.*

iii) $R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$ *is an object of* $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ *that is acyclic outside degrees* 0 *and* 1*. One has a canonical identification* $H^0_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)=\mathcal{O}^\times_{K,S}$ *and a natural exact sequence of* G*-modules*

$$0\to A_{K,S}\to H^1_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)\to X_{K,S}\to 0.$$

iv) *If* J *is any normal subgroup of* G*, then there exists a natural isomorphism in* $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G/J])$ *of the form*

$$R\Gamma_{\mathcal{W}}(U_{K^J,S},\mathbb{G}_m)\cong R\operatorname{Hom}_{\mathbb{Z}[J]}(\mathbb{Z},R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)).$$

With respect to the descriptions of cohomology given in iii) (for both K *and* K^J*) the displayed isomorphism induces the natural identification* $\mathcal{O}^\times_{K^J,S}=(\mathcal{O}^\times_{K,S})^J$ *and also identifies* $X_{K^J,S}$ *with a submodule of* $X_{K,S}$ *by means of the map that sends each place* v *of* $S(K^J)$ *to* $\sum_{j\in J}j(w)$ *where* w *is any place of* K *lying above* v*.*

Proof. Claim i) is proved by the argument of [40, proof of Th. 6.5].
The existence of the distinguished triangle in claim ii) can be proved by comparing the spectral sequences of [40, Prop. 2.3(f)] or by using the approach of Geisser in [30, Th. 6.1].
The descriptions of the groups $H^i_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$ given in claim iii) are proved by Lichtenbaum in [40, Th. 7.1c)]. They follow from the long exact sequence of cohomology associated to the triangle in claim ii), the canonical identifications $H^0(U_{K,S},\mathbb{G}_m)\cong\mathcal{O}^\times_{K,S}, H^1(U_{K,S},\mathbb{G}_m)\cong\operatorname{Pic}(\mathcal{O}_{K,S})$ and $H^2(U_{K,S},\mathbb{G}_m)\cong X_{K,S}\otimes\mathbb{Q}/\mathbb{Z}$ and the fact that $H^i(U_{K,S},\mathbb{G}_m)=0$ if $i>2$. Since each cohomology group of $R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$ is finitely generated, a standard argument of homological algebra shows that this complex belongs to $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ if and only if it belongs to $\mathcal{D}^{\mathrm{fpd}}(\mathbb{Z}[G])$ (cf. [11, proof of Prop. 1.20, Steps 3 and 4]). On the other hand, any G-module that is c-t has finite projective dimension as a $\mathbb{Z}[G]$-module and so it suffices to show that $R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$ is isomorphic to a bounded complex of G-modules which are each c-t. Now the G-module $X_{K,S}\otimes\mathbb{Q}$ is c-t and so the distinguished triangle of claim ii) implies that we need only prove that $R\Gamma(U_{K,S},\mathbb{G}_m)$ is isomorphic to a bounded complex of G-modules which are each c-t. But this is true because the natural morphism $\pi: U_{K,S}\to U_{k,S}$ is étale and $\mathbb{G}_m=\pi^*\mathbb{G}_m$ on $(U_{K,S})_{\text{ét}}$ (cf. [11, proof of Prop. 1.20, Steps 1 and 2]).
Claim iv) follows from the triangle of claim ii) and the description of cohomology given in iii) (for both K and K^J) together with an explicit computation of the maps induced on cohomology by the natural isomorphism $R\Gamma(U_{K^J,S},\mathbb{G}_m)\cong R\operatorname{Hom}_{\mathbb{Z}[J]}(\mathbb{Z},R\Gamma(U_{K,S},\mathbb{G}_m))$ in $\mathcal{D}(\mathbb{Z}[G/J])$ (for more details as to the latter see, for example, the proof of [12, Lem. 12]). □

For each place w of K we let $|\cdot|_w$ denote the absolute value of w normalised as in [50, Chap. 0, 0.2]. We write $\mathrm{R}_{K,S}$ for the $\mathbb{R}[G]$-equivariant isomorphism $\mathcal{O}_{K,S}^\times \otimes \mathbb{R} \longrightarrow X_{K,S} \otimes \mathbb{R}$ which at each $u \in \mathcal{O}_{K,S}^\times$ satisfies

$$\mathrm{R}_{K,S}(u) = -\sum_{w \in S(K)} \log |\, u \,|_w \cdot w. \tag{2}$$

We also denote by $\mathrm{R}_{K,S}$ the $\mathbb{R}$-trivialisation of $R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$ that is induced by $\mathrm{R}_{K,S}$ and the descriptions of Lemma 1iii).
We can now state the central conjecture of this paper.

CONJECTURE $\mathrm{C}(K/k)$: *In $K_0(\mathbb{Z}[G], \mathbb{R})$ one has an equality*

$$\delta(\theta_{K/k,S}^*(0)^\#) = \chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S}).$$

Remark 1. Lemma 1i) shows that $\mathrm{C}(K/k)$ can be naturally rephrased in terms of $R\Gamma_{\mathcal{W}}(C_K, j_!\mathbb{Z})$. We have chosen to work in terms of $\mathbb{G}_m$ rather than $j_!\mathbb{Z}$ for the purposes of explicit computations that we make in subsequent sections (see also Remark 3 in this regard).

Remark 2. If G is abelian, then the equality of $\mathrm{C}(K/k)$ is equivalent to a formula for the $\mathbb{Z}[G]$-submodule of $\mathbb{R}[G]$ which is generated by $\theta_{K/k,S}^*(0)^\#$ in terms of the $\mathbb{Z}[G]$-equivariant graded determinant of $R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$ (see Remark A1 in the Appendix). By using this observation in conjunction with Remark 1 it can be shown that $\mathrm{C}(k/k)$ is equivalent to a special case of the conjecture formulated by Lichtenbaum in [40, Conj. 8.1e)].

Remark 3. Let $j : U_{K,S} \to C_K$ denote the natural open immersion. Then the Poincaré Duality Theorem of [42, Chap. II, Th. 3.1] gives rise to a commutative diagram in $\mathcal{D}(\mathbb{Z}[G])$ of the form

$$\begin{tikzcd}
X_{K,S} \otimes \mathbb{Q}[-2] \arrow[r] \arrow[d, equal] & R\Gamma(U_{K,S}, \mathbb{G}_m) \arrow[r] \arrow[d] & R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \\
X_{K,S} \otimes \mathbb{Q}[-2] \arrow[r] & \mathrm{Hom}_{\mathbb{Z}}(R\Gamma(C_K, j_!\mathbb{Z}), \mathbb{Q}/\mathbb{Z}[-3]) \arrow[d] & \\
 & (\hat{\mathcal{O}}_{K,S}^\times / \mathcal{O}_{K,S}^\times)[0] &
\end{tikzcd}$$

where the top row is as in Lemma 1ii), $\hat{\mathcal{O}}_{K,S}^\times$ denotes the profinite completion of $\mathcal{O}_{K,S}^\times$ and the second column is a distinguished triangle. This diagram implies that $R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$ is a precise analogue of the complex Ψ_S that occurs in [12, Rem. following Prop. 3.1] and [7, Prop. 2.1.1]. For this reason, $\mathrm{C}(K/k)$ is an analogue of the conjectural vanishing of the element $T\Omega(K/k, 0)$ defined in [7, Th. 2.1.2], where K/k is a Galois extension of number fields of group G, and also coincides in the abelian case with the function field case of [8, Conj. 2.1]. We recall that the vanishing of the element $T\Omega(K/k, 0)$ is equivalent to the validity of the 'Lifted Root Number Conjecture' of Gruenberg, Ritter and

Weiss [33] (see [7, Th. 2.3.3] for a proof of this fact) and also to the validity of the 'Equivariant Tamagawa Number Conjecture' of [13, Conj. 4(iv)] as applied to the pair $(h^0(\mathrm{Spec}\ K), \mathbb{Z}[G])$ where $h^0(\mathrm{Spec}\, K)$ is considered as a motive that is defined over k and has coefficients $\mathbb{Q}[G]$ (see [7, Th. 2.4.1] or [14, §3] for different proofs of this fact). We further recall that [13, Conj. 4(iv)] is itself a natural equivariant version of the seminal conjecture of Bloch and Kato from [6], and that if G is abelian, then it refines the 'Generalized Iwasawa Main Conjecture' formulated by Kato in [35, §3.2] (cf. [14, §2] in this regard). Finally we recall that strong evidence in favour of [13, Conj. 4(iv)] has recently been obtained in [15, 16].

By a change of variable we now remove all of the transcendental terms which occur in $\mathrm{C}(K/k)$ and then decompose the conjecture according to (1).
To do this we set $t := p^{-s}$ and then define a $\zeta(\mathbb{C}[G])$-valued function of the complex variable t by means of the equality $Z_{K/k,S}(t) := \theta_{K/k,S}(s)$. For each place $w \in S(K)$ we write val_w and $k(w)$ for its valuation and residue field and let $\deg(w)$ denote the degree of the field extension $k(w)/\mathbb{F}_p$. We write $\mathrm{D}_{K,S} : \mathcal{O}_{K,S}^\times \to X_{K,S}$ for the homomorphism which at each $u \in \mathcal{O}_{K,S}^\times$ satisfies

$$\mathrm{D}_{K,S}(u) = \sum_{w\in S(K)} \mathrm{val}_w(u)\deg(w)\cdot w.$$

Lemma 2. *Let $e : \mathrm{Spec}(\zeta(\mathbb{R}[G])) \to \mathbb{Z}$ denote the algebraic order of $Z_{K/k,S}(t)$ at $t = 1$ (which we regard as an element of $\mathbb{Z}^{\pi_0(\mathrm{Spec}(\zeta(\mathbb{R}[G])))}$ in the natural way). Then the element*

$$Z^*_{K/k,S}(1) := \lim_{t\to 1}(1-t)^{-e} Z_{K/k,S}(t)$$

belongs to $\zeta(\mathbb{Q}[G])^\times$ and $\mathrm{C}(K/k)$ is valid if and only if in $K_0(\mathbb{Z}[G],\mathbb{Q})$ one has

$$\delta(Z^*_{K/k,S}(1)^\#) = \chi_{\mathbb{Q}}(R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m), \mathrm{D}_{K,S}\otimes\mathbb{Q}). \tag{3}$$

Proof. The algebraic order of $\theta_{K/k,S}(s)^\#$ at $s = 0$ is equal to e. In addition, by an explicit computation one verifies that

$$\begin{aligned}\theta^*_{K/k,S}(0)^\# &= \lim_{s\to 0} s^{-e}\theta_{K/k,S}(s)^\# \\ &= (\log(p))^e \cdot Z^*_{K/k,S}(1)^\#.\end{aligned}$$

When combined with the known validity of Stark's Conjecture for K/k [50, p. 111], this equality proves that $Z^*_{K/k,S}(1)$ belongs to $\zeta(\mathbb{Q}[G])^\times$. Also, since $\chi_{\mathbb{Q}}(R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m), \mathrm{D}_{K,S}\otimes\mathbb{Q})$ is equal to $\chi(R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m), \mathrm{D}_{K,S}\otimes\mathbb{R})$ in $K_0(\mathbb{Z}[G],\mathbb{R})$, the above equality shows that $\mathrm{C}(K/k)$ is equivalent to (3) provided that in $K_0(\mathbb{Z}[G],\mathbb{R})$ one has

$$\chi(R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m), \mathrm{R}_{K,S}) = \chi(R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m), \mathrm{D}_{K,S}\otimes\mathbb{R}) + \delta((\log(p))^e).$$

The validity of this equality follows directly from [7, Prop. 1.2.1(ii)] in conjunction with the equality $\mathrm{R}_{K,S}(u) = \log(p)\cdot \mathrm{D}_{K,S}(u)$ for each $u \in \mathcal{O}_{K,S}^\times$ and the fact that the reduced rank (as defined in [13, §2.6]) of the $\mathbb{R}[G]$-module $X_{K,S} \otimes \mathbb{R}$ is equal to e [50, Chap. I, Prop. 3.4]. □

From Lemma 2 and the bijectivity of the map (1) it is clear that $\mathrm{C}(K/k)$ is valid if and only if, for each prime ℓ, the following conjecture is valid.

Conjecture $\mathrm{C}_\ell(K/k)$: *The image of (3) under ρ_ℓ is valid.*

Remark 4. There are several useful functorial properties of $\mathrm{C}(K/k)$ that can be proved directly or by combining Remark 3 with the relevant arguments from either [7] or [13, §4.4-5]. For example, in this way it can be shown that the validity of $\mathrm{C}(K/k)$ is independent of the choice of S (cf. [7, Th. 2.1.2]). In addition, if ℓ is any prime and H is any subgroup of G, then it can be shown that the validity of the image of the equality of $\mathrm{C}_\ell(K/k)$ under the natural restriction map $K_0(\mathbb{Z}_\ell[G], \mathbb{Q}_\ell) \to K_0(\mathbb{Z}_\ell[H], \mathbb{Q}_\ell)$ is equivalent to the validity of $\mathrm{C}_\ell(K/K^H)$ (cf. [7, Prop. 2.1.4(i)]). In a similar way, if J is any normal subgroup of G, then Lemma 1iv) implies that the validity of the image of the equality of $\mathrm{C}_\ell(K/k)$ under the natural coinflation map $K_0(\mathbb{Z}_\ell[G], \mathbb{Q}_\ell) \to K_0(\mathbb{Z}_\ell[G/J], \mathbb{Q}_\ell)$ is equivalent to the validity of $\mathrm{C}_\ell(K^J/k)$ (cf. [7, Prop. 2.1.4(ii)]).

3. Evidence

In this section we shall provide the following evidence in support of $\mathrm{C}(K/k)$.

Theorem 3.1. *Let K/k be a finite Galois extension of global function fields of characteristic p and set $G := \mathrm{Gal}(K/k)$.*

i) *If $\ell \neq p$, then $\mathrm{C}_\ell(K/k)$ is valid.*
ii) *$\mathrm{C}_p(K/k)$ is valid modulo the torsion subgroup of $K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)$.*

Corollary 1. *$\mathrm{C}(K/k)$ is valid modulo the torsion subgroup of $K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)$.*

Proof. Clear. □

Remark 5. The group $K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)$ is torsion-free if $p \nmid |G|$ (cf. [13, proof of Lem. 11c)]) and also if $p = 2$ and G is either of order 2 or is dihedral of order congruent to 2 modulo 4 [5, Lem. 8.2].

3.1. The descent formalism. In this subsection we prepare for the proof of Theorem 3.1i) by proving a purely algebraic result. This provides a natural generalisation of several results that have already been used elsewhere (cf. Remark 6) and so the material of this subsection may well itself be of some independent interest.

We fix an arbitrary rational prime ℓ and for each $\mathbb{Z}_\ell$-module M we set $M_{\mathbb{Q}_\ell} := M \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$. We say that an endomorphism ψ of a finitely generated $\mathbb{Z}_\ell[G]$-module M is 'semi-simple at 0' if the natural composite homomorphism

$$\ker(\psi) \xrightarrow{\subseteq} M \twoheadrightarrow \mathrm{cok}(\psi) \tag{4}$$

has both finite kernel and finite cokernel. We note that this condition is satisfied if and only if there exists a $\mathbb{Q}_\ell[G][\psi]$-equivariant direct complement to the submodule $\ker(\psi)_{\mathbb{Q}_\ell}$ of $M_{\mathbb{Q}_\ell}$.

Let t be an indeterminate. Then for any element f of $\zeta(\mathbb{Q}_\ell[G])[[t]]$ we write $e_f : \mathrm{Spec}(\zeta(\mathbb{Q}_\ell[G])) \to \mathbb{Z}$ for the algebraic order of $f(t)$ at $t = 1$. We identify e_f with an element of $\mathbb{Z}^{\pi_0(\mathrm{Spec}(\zeta(\mathbb{Q}_\ell[G])))}$ in the natural way and then set

$$f^*(1) := \lim_{t\to 1}(1-t)^{-e_f} f(t) \in \zeta(\mathbb{Q}_\ell[G])^\times.$$

In particular, if θ is any endomorphism of a finitely generated $\mathbb{Z}_\ell[G]$-module M for which $1-\theta$ is semi-simple at 0 and

$$f(t) = \mathrm{detred}_{\mathbb{Q}_\ell[G]}(1-\theta\cdot t : M_{\mathbb{Q}_\ell}),$$

then we set

$$\begin{aligned}\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta : M_{\mathbb{Q}_\ell}) :=& f^*(1)\\ =& \mathrm{detred}_{\mathbb{Q}_\ell[G]}(1-\theta : D)\end{aligned}$$

where D is any choice of a $\mathbb{Q}_\ell[G][\theta]$-equivariant direct complement to the submodule $\ker(1-\theta)_{\mathbb{Q}_\ell}$ of $M_{\mathbb{Q}_\ell}$.

We now suppose given a bounded complex of finitely generated projective $\mathbb{Z}_\ell[G]$-modules $P^\cdot$ and a $\mathbb{Z}_\ell[G]$-equivariant endomorphism θ of $P^\cdot$ which is such that the induced endomorphism $H^i(1-\theta)$ of $H^i(P^\cdot)$ is semi-simple at 0 in each degree i.

We let $\mathrm{C}(\theta)^\cdot$ denote the -1-shift of the mapping cone of the endomorphism of $P^\cdot$ induced by $1-\theta$. Then from the long exact sequence of cohomology that is associated to the distinguished triangle

$$P^\cdot \xrightarrow{1-\theta} P^\cdot \to C(\theta)^\cdot[1] \to P^\cdot[1]$$

one obtains in each degree i a short exact sequence

$$0 \to \mathrm{cok}(H^{i-1}(1-\theta)) \to H^i(\mathrm{C}(\theta)^\cdot) \to \ker(H^i(1-\theta)) \to 0.$$

Upon combining these sequences with the isomorphisms

$$\ker(H^i(1-\theta))_{\mathbb{Q}_\ell} \xrightarrow{\sim} \mathrm{cok}(H^i(1-\theta))_{\mathbb{Q}_\ell}$$

induced by (4) (with $\psi = H^i(1-\theta)$ and $M = H^i(P^\cdot)$) one obtains a well-defined $\mathbb{Q}_\ell$-trivialisation τ_θ of $\mathrm{C}(\theta)^\cdot$.

Proposition 3.1. *Let $P^\cdot$ be a bounded complex of finitely generated projective $\mathbb{Z}_\ell[G]$-modules and θ a $\mathbb{Z}_\ell[G]$-equivariant endomorphism of $P^\cdot$ for which $H^i(1-\theta)$ is semi-simple at 0 in each degree i. Then in $K_0(\mathbb{Z}_\ell[G],\mathbb{Q}_\ell)$ one has*

$$\chi_\ell(\mathrm{C}(\theta)^\cdot,\tau_\theta) = \sum_{i\in\mathbb{Z}}(-1)^i\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta : H^i(P^\cdot)_{\mathbb{Q}_\ell})).$$

Proof. We shall argue by induction on the quantity

$$|P^\cdot| := \max\{i : P^i \neq 0\} - \min\{j : P^j \neq 0\}.$$

We first assume that $|P^\cdot| = 0$ so that $P^\cdot$ has only one non-zero term. To be specific, we assume that $P^\cdot = P^n[-n]$ (so that $H^n(P^\cdot) = P^n$). In this case $C(\theta)^\cdot$

is equal to the complex $P^n \xrightarrow{1-\theta^n} P^n$, where the first term is placed in degree n. In addition, upon choosing a $\mathbb{Q}_\ell[G][\theta^n]$-equivariant direct complement D to $\ker(1-\theta^n)_{\mathbb{Q}_\ell}$ in $P^n_{\mathbb{Q}_\ell}$, and using (4) to identify $H^n(C(\theta)^\cdot)_{\mathbb{Q}_\ell} = \ker(1-\theta^n)_{\mathbb{Q}_\ell}$ with $H^{n+1}(C(\theta)^\cdot)_{\mathbb{Q}_\ell} = \mathrm{cok}(1-\theta^n)_{\mathbb{Q}_\ell}$, the trivialisation τ_θ is induced by the identity map on cohomology. Hence, from Lemma A1, one has

$$\begin{aligned}\chi_\ell(C(\theta)^\cdot,\tau_\theta) =&(-1)^n\partial^1_{\mathbb{Z}_\ell[G],\mathbb{Q}_\ell}([\mathrm{id}_{\ker(1-\theta^n)_{\mathbb{Q}_\ell}} \oplus (1-\theta^n)\mid_D])\\ =&(-1)^n\partial^1_{\mathbb{Z}_\ell[G],\mathbb{Q}_\ell}([1-\theta^n\mid_D])\\ =&(-1)^n\delta_\ell(\mathrm{detred}_{\mathbb{Q}_\ell[G]}(1-\theta^n:D))\\ =&(-1)^n\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta:H^n(P^n[-n])_{\mathbb{Q}_\ell})),\end{aligned}$$

as required.

We now assume that $|P^\cdot| = n$ and, to fix notation, that $\min\{j : P^j \neq 0\} = 0$. We also assume that the claimed formula is true for any pair of the form $(Q^\cdot,\phi)$ where $Q^\cdot$ is a bounded complex of finitely generated projective $\mathbb{Z}_\ell[G]$-modules for which $|Q^\cdot| \le n-1$ and ϕ is a $\mathbb{Z}_\ell[G]$-equivariant endomorphism of $Q^\cdot$ for which $H^i(1-\phi)$ is semi-simple at 0 in each degree i. For any complex $C^\cdot$ and any integer i we write $B^i(C^\cdot), Z^i(C^\cdot)$ and $d^i(C^\cdot)$ for the boundaries, cycles and differential in degree i. If necessary, we use the argument of [25, Lem. 7.10] to change θ by a homotopy in order to ensure that, in each degree i, the restriction of $1-\theta^i$ to $B^i(P^\cdot)$ induces an automorphism of $B^i(P^\cdot)_{\mathbb{Q}_\ell}$. We shall make frequent use of this assumption (without explicit comment) in the remainder of this argument.

We henceforth let $Q^\cdot$ denote the naive truncation in degree $n-1$ of $P^\cdot$ (so $Q^i = P^i$ if $i \le n-1$ and $Q^n = 0$). Then one has a tautological short exact sequence of complexes $0 \to P^n[-n] \to P^\cdot \to Q^\cdot \to 0$. From the associated long exact cohomology sequence we deduce that $H^i(Q^\cdot) = H^i(P^\cdot)$ if $i < n-1$ and that there are commutative diagrams of short exact sequences of the form

$$\begin{array}{ccccccccc} 0 & \longrightarrow & H^{n-1}(P^\cdot) & \longrightarrow & H^{n-1}(Q^\cdot) & \longrightarrow & B^n & \longrightarrow & 0\\ & & \big\downarrow{\scriptstyle H^{n-1}(1-\theta)} & & \big\downarrow{\scriptstyle H^{n-1}(1-\phi)} & & \big\downarrow{\scriptstyle 1-\theta^n} & & \\ 0 & \longrightarrow & H^{n-1}(P^\cdot) & \longrightarrow & H^{n-1}(Q^\cdot) & \longrightarrow & B^n & \longrightarrow & 0\end{array}$$

$$\begin{array}{ccccccccc} 0 & \longrightarrow & B^n & \longrightarrow & H^n(P^n[-n]) & \longrightarrow & H^n(P^\cdot) & \longrightarrow & 0\\ & & \big\downarrow{\scriptstyle 1-\theta^n} & & \big\downarrow{\scriptstyle H^n(1-\theta^n[-n])} & & \big\downarrow{\scriptstyle H^n(1-\theta)} & & \\ 0 & \longrightarrow & B^n & \longrightarrow & H^n(P^n[-n]) & \longrightarrow & H^n(P^\cdot) & \longrightarrow & 0.\end{array}$$

We write ϕ, resp. $\theta^n[-n]$, for the endomorphism of $Q^\cdot$, resp. $P^n[-n]$, which is induced by θ. Then the above diagrams imply that $\ker(H^i(1-\phi))_{\mathbb{Q}_\ell} = \ker(H^i(1-\theta))_{\mathbb{Q}_\ell}$ and $\mathrm{cok}(H^i(1-\phi))_{\mathbb{Q}_\ell} = \mathrm{cok}(H^i(1-\theta))_{\mathbb{Q}_\ell}$ for all $i < n$ and also that $\ker(H^n(1-\theta^n[-n]))_{\mathbb{Q}_\ell} = \ker(H^n(1-\theta))_{\mathbb{Q}_\ell}$ and $\mathrm{cok}(H^n(1-\theta^n[-n]))_{\mathbb{Q}_\ell} = \mathrm{cok}(H^n(1-\theta))_{\mathbb{Q}_\ell}$. This implies that $1-\phi$ and $1-\theta^n[-n]$ induce endomorphisms of $H^i(Q^\cdot)$ and $H^i(P^n[-n])$ respectively which are each semi-simple at 0 in all degrees i.

We set $C := \mathrm{Cone}(1-\theta^n[-n])[-1]$, $D := C(\theta)^\cdot = \mathrm{Cone}(1-\theta)[-1]$ and $E := \mathrm{Cone}(1-\phi)[-1]$ so that there is a natural short exact sequence of complexes

$$0 \to C \xrightarrow{\alpha} D \xrightarrow{\beta} E \to 0. \tag{5}$$

Now, since $|Q^\cdot| < n$, our inductive hypothesis implies that

$$\begin{aligned}\chi_\ell(E,\tau_\phi) =& \sum_{i=0}^{n-1}(-1)^i\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\phi : H^i(Q^\cdot)_{\mathbb{Q}_\ell}))\\ =&(-1)^{n-1}\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\phi : H^{n-1}(Q^\cdot)_{\mathbb{Q}_\ell}))\\ &+\sum_{i=0}^{n-2}(-1)^i\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta : H^i(P^\cdot)_{\mathbb{Q}_\ell})).\end{aligned}$$

In addition, since $|P^n[-n]| = 0$, our earlier argument proves that

$$\chi_\ell(C,\tau_{\theta^n[-n]}) = (-1)^n\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta^n[-n] : H^n(P^n[-n])_{\mathbb{Q}_\ell}).$$

From the commutative diagrams displayed above, one also has

$$\begin{aligned}&\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\phi : H^{n-1}(Q^\cdot)_{\mathbb{Q}_\ell}) =\\ &\qquad \mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta^n : B^n_{\mathbb{Q}_\ell})\cdot \mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta : H^{n-1}(P^\cdot)_{\mathbb{Q}_\ell})\end{aligned}$$

and

$$\begin{aligned}&\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta^n[-n] : H^n(P^n[-n])_{\mathbb{Q}_\ell}) =\\ &\qquad \mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta^n : B^n_{\mathbb{Q}_\ell})\cdot \mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta : H^n(P^\cdot)_{\mathbb{Q}_\ell}).\end{aligned}$$

Upon combining the last four displayed formulas we obtain an equality

$$\chi_\ell(C,\tau_{\theta^n[-n]}) + \chi_\ell(E,\tau_\phi) = \sum_{i\in\mathbb{Z}}(-1)^i\delta_\ell(\mathrm{detred}^*_{\mathbb{Q}_\ell[G]}(1-\theta : H^i(P^\cdot)_{\mathbb{Q}_\ell})),$$

and so the claimed result will follow if we can show that

$$\chi_\ell(D,\tau_\theta) = \chi_\ell(C,\tau_{\theta^n[-n]}) + \chi_\ell(E,\tau_\phi). \tag{6}$$

Before discussing the proof of this equality we introduce some convenient notation: for any $\mathbb{Z}_\ell$-module A we set $\overline{A} := A\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$, and we use similar abbreviations for both complexes and morphisms of $\mathbb{Z}_\ell$-modules. For any complex A we also set $H^+_A := H^+(A)$ and $H^-_A := H^-(A)$.
The key to proving (6) is the observation (which is itself straightforward to verify directly) that one can choose elements κ_1, κ_2 and κ_3 of $\tau_{\theta^n[-n]}, \tau_\theta$ and τ_ϕ respectively which together lie in a commutative diagram of short exact sequences of the form

$$\begin{tikzcd} 0 \arrow[r] & H^+_{\overline{C}} \arrow[r, "H^+(\overline{\alpha})"] \arrow[d, "\kappa_1"] & H^+_{\overline{D}} \arrow[r, "H^+(\overline{\beta})"] \arrow[d, "\kappa_2"] & H^+_{\overline{E}} \arrow[r] \arrow[d, "\kappa_3"] & 0 \\ 0 \arrow[r] & H^-_{\overline{C}} \arrow[r, "H^-(\overline{\alpha})"] & H^-_{\overline{D}} \arrow[r, "H^-(\overline{\beta})"] & H^-_{\overline{E}} \arrow[r] & 0. \end{tikzcd} \tag{7}$$

Indeed, the equality (6) follows directly upon combining such a diagram with the exact sequence (5) and the result of [9, Th. 2.8]. However, for the convenience of the reader, we also now indicate a more direct argument.
After taking account of the construction of $\chi_\ell(\cdot,\cdot)$ given in the Appendix (the notation of which we now assume) and the definitions of $\tau_{\theta^n[-n]}(C), \tau_\theta(D)$ and $\tau_\phi(E)$ it is enough to prove the existence of a commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \overline{C}^+ & \xrightarrow{\overline{\alpha}^+} & \overline{D}^+ & \xrightarrow{\overline{\beta}^+} & \overline{E}^+ & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & B_{\overline{C}} \oplus H^+_{\overline{C}} & \xrightarrow{(\alpha', H^+(\overline{\alpha}))} & B_{\overline{D}} \oplus H^+_{\overline{D}} & \xrightarrow{(\beta', H^+(\overline{\beta}))} & B_{\overline{E}} \oplus H^+_{\overline{E}} & \longrightarrow & 0 \\
 & & \Big\downarrow{\scriptstyle (\mathrm{id},\kappa_1)} & & \Big\downarrow{\scriptstyle (\mathrm{id},\kappa_2)} & & \Big\downarrow{\scriptstyle (\mathrm{id},\kappa_3)} & & \\
0 & \longrightarrow & B_{\overline{C}} \oplus H^-_{\overline{C}} & \xrightarrow{(\alpha', H^-(\overline{\alpha}))} & B_{\overline{D}} \oplus H^-_{\overline{D}} & \xrightarrow{(\beta', H^-(\overline{\beta}))} & B_{\overline{E}} \oplus H^-_{\overline{E}} & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \overline{C}^- & \xrightarrow{\overline{\alpha}^-} & \overline{D}^- & \xrightarrow{\overline{\beta}^-} & \overline{E}^- & \longrightarrow & 0,
\end{array}$$

where $B_{\overline{C}}$ denotes $B^{\mathrm{all}}(\overline{C})$, and similarly for $B_{\overline{D}}$ and $B_{\overline{E}}$, $\alpha' : B_{\overline{C}} \to B_{\overline{D}}$ and $\beta' : B_{\overline{D}} \to B_{\overline{E}}$ are the natural homomorphisms that are induced by α and β respectively, κ_1, κ_2 and κ_3 are as in (7) and all unlabelled vertical maps are the isomorphisms induced by a choice of sections to each of the natural homomorphisms $\overline{C}^i \to B^{i+1}(\overline{C}), Z^i(\overline{C}) \to H^i(\overline{C}), \overline{D}^i \to B^{i+1}(\overline{D}), Z^i(\overline{D}) \to H^i(\overline{D}), \overline{E}^i \to B^{i+1}(\overline{E})$ and $Z^i(\overline{E}) \to H^i(\overline{E})$. Indeed, if such a diagram exists, then the composite vertical isomorphisms belong to $\tau_{\theta^n[-n]}(C), \tau_\theta(D)$ and $\tau_\phi(E)$ respectively, and so the commutativity of the diagram formed by the first and fourth rows combines with the exactness of the sequences $0 \to C^+ \xrightarrow{\alpha^+} D^+ \xrightarrow{\beta^+} E^+ \to 0$ and $0 \to C^- \xrightarrow{\alpha^-} D^- \xrightarrow{\beta^-} E^- \to 0$ and the defining relations of $K_0(\mathbb{Z}_\ell[G], \mathbb{Q}_\ell)$ [46, p. 415] to imply the required equality (6). Thus, upon noting that the rows of this diagram are all exact (the second and third as a consequence of the exactness of the rows of (7)), it is enough to prove that sections of the above form can be chosen in such a way that the top and bottom two squares of the diagram commute, and this in turn can be proved by a straightforward and explicit exercise using the following facts. After choosing $\mathbb{Q}_\ell[G]$-equivariant direct sum decompositions $\overline{P^{n-1}} = \overline{\mathrm{im}(1-\theta^{n-1})} \oplus S^{n-1}$ and $\overline{P^n} = \overline{\ker(1-\theta^n)} \oplus S^n$, one obtains a direct sum decomposition $\overline{D}^n = \overline{P^{n-1}} \oplus \overline{P^n} = B^n(\overline{D}) \oplus S^{n-1,*} \oplus (0, \overline{\ker(1-\theta^n)}) \oplus (0, S^n)$ where $S^{n-1,*}$ denotes the set of elements (π, π') where π runs over S^{n-1} and π' denotes the unique element of $B^n(\overline{P^\cdot})$ which is such that $(\pi, \pi') \in Z^n(\overline{D})$; one has $Z^n(\overline{D}) = B^n(\overline{D}) \oplus S^{n-1,*} \oplus (0, \overline{\ker(1-\theta^n)})$; the natural projection maps induce isomorphisms $B^n(\overline{D}) \cong B^n(\overline{E}), Z^{n-1}(\overline{D}) \cong Z^{n-1}(\overline{E})$ and $B^{n-1}(\overline{D}) \cong B^{n-1}(\overline{E})$. $\square$

Remark 6. There are two special cases in which the formula of Proposition 3.1 has already been proved: if $C(\theta)^{\cdot}_{\mathbb{Q}_\ell}$ is acyclic, then $H^i(1-\theta)$ is automatically semi-simple at 0 in each degree i and the given formula has been proved by Greither and the present author in [16, proof of Prop. 4.1]; if G is abelian, then Proposition 3.1 can be reinterpreted in terms of graded determinants and in this case the given formula has been proved to within a sign ambiguity by Kato in [35, Lem. 3.5.8]. (This sign ambiguity arises because Kato uses ungraded determinants - for more details in this regard see [loc cit., Rem. 3.2.3(3) and 3.2.6(3),(5)] and [13, Rem. 9].)

3.2. Zeta functions of varieties. In this subsection we fix a prime ℓ that is distinct from p. We also fix an algebraic closure $\mathbb{F}_p^c$ of $\mathbb{F}_p$, we set $\Gamma := \mathrm{Gal}(\mathbb{F}_p^c/\mathbb{F}_p)$ and we write σ for the (arithmetic) Frobenius element in Γ. For any scheme X over $\mathbb{F}_p$ we write X^c for the associated scheme $\mathbb{F}_p^c \times_{\mathbb{F}_p} X$ over $\mathbb{F}_p^c$.

We let J be a finite group, X and Y separated schemes of finite type over $\mathbb{F}_p$ and $\pi : X \to Y$ an étale morphism that is Galois of group J. For each ℓ-adic sheaf $\mathcal{G}$ on $Y_{\text{ét}}$ we follow the approach of Deligne [27, Rem. 2.12] to define a J-equivariant Zeta function by setting

$$\begin{aligned} &Z_J(Y, \pi_*\pi^*\mathcal{G} \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell, t) := \\ &\qquad \prod_y \mathrm{detred}_{\mathbb{Q}_\ell[J]}(1 - f_y^{-1} \cdot t^{\deg(y)} \mid (\pi_*\pi^*\mathcal{G} \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell)_{\overline{y}})^{-1} \in \zeta(\mathbb{Q}_\ell[J])[[t]], \end{aligned}$$

where y runs over the set of closed points of Y, f_y denotes the arithmetic Frobenius of y, $\deg(y)$ the degree of y and subscript $\overline{y}$ denotes taking stalk at a geometric point over y.

We now combine the algebraic approach of the previous subsection with a well known result of Grothendieck from [32] to describe, for each integer r, the image of the leading term $Z_J^*(Y, \pi_*\pi^*\mathbb{Z}_\ell(r) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell, 1)^\#$ under the homomorphism $\delta_{J,\ell} : \zeta(\mathbb{Q}_\ell[J])^\times \to K_0(\mathbb{Z}_\ell[J], \mathbb{Q}_\ell)$.

To this end we observe that π_* is exact and hence that, for each sheaf $\mathcal{G}$ as above, there is a natural isomorphism $R\Gamma(Y, \pi_*\pi^*\mathcal{G}) \cong R\Gamma(X, \pi^*\mathcal{G})$ in $\mathcal{D}(\mathbb{Z}_\ell[J])$. This implies that if $\mathcal{G}$ is any étale (pro-)sheaf of finitely generated $\mathbb{Z}_\ell$-modules on Y and we set $\mathcal{F} := \pi^*\mathcal{G}$, then the complexes $R\Gamma(X, \mathcal{F})$ and $R\Gamma(X^c, \mathcal{F})$ both belong to $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[J])$ (cf. [29, Th. 5.1]). We may therefore fix a bounded complex of finitely generated projective $\mathbb{Z}_\ell[J]$-modules $C^\cdot$ for which there exists an isomorphism $\alpha : C^\cdot \xrightarrow{\sim} R\Gamma(X^c, \mathcal{F})$ in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[J])$ and a $\mathbb{Z}_\ell[J]$-endomorphism θ of $C^\cdot$ that induces the action of σ on $R\Gamma(X^c, \mathcal{F})$ (the existence of such a θ follows from [41, Chap. VI, Lem. 8.17] - but note that the map ψ in loc. cit. need not, in general, be a quasi-isomorphism). In this way we obtain a commutative diagram in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[J])$ of the form

$$\begin{array}{ccccccc} & & C^{\cdot} & \xrightarrow{1-\theta} & C^{\cdot} & & \\ (8) & & \Big\downarrow{\scriptstyle\alpha} & & \Big\downarrow{\scriptstyle\alpha} & & \\ R\Gamma(X,\mathcal{F}) & \longrightarrow & R\Gamma(X^c,\mathcal{F}) & \xrightarrow{1-\sigma} & R\Gamma(X^c,\mathcal{F}) & \longrightarrow & R\Gamma(X,\mathcal{F})[1], \end{array}$$

where the lower row denotes the natural distinguished triangle. Taken in conjunction with the Octahedral axiom, this diagram implies the existence of an isomorphism $\alpha' : C(\theta)^{\cdot} \xrightarrow{\sim} R\Gamma(X,\mathcal{F})$ in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[J])$. Further, the hypothesis that the composite (4) with $\psi = H^i(1-\theta)$ and $M = H^i(C^{\cdot})$ has both finite kernel and finite cokernel is equivalent to the hypothesis that σ acts 'semi-simply' on the space $H^i(X^c,\mathcal{F})\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ and is therefore expected to be true under some very general conditions [35, Rem. 3.5.4]. In this context, and in terms of the notation of Lemma A2, we write $\tau_{X,\mathcal{F},\sigma}$ for the $\mathbb{Q}_\ell$-trivialisation of $R\Gamma(X,\mathcal{F})$ which is equal to $(\tau_\theta)_{\alpha'}$ where τ_θ is the $\mathbb{Q}_\ell$-trivialisation of $C(\theta)^{\cdot}$ that is defined just prior to Proposition 3.1 (with $P^{\cdot} = C^{\cdot}$).

Remark 7. The trivialisation $\tau_{X,\mathcal{F},\sigma}$ defined above has an alternative description. To explain this we let $\mathcal{C}(\mathcal{F})^{\cdot}$ denote the complex

$$H^0(X,\mathcal{F}) \xrightarrow{\kappa} H^1(X,\mathcal{F}) \xrightarrow{\kappa} H^2(X,\mathcal{F}) \xrightarrow{\kappa} \cdots$$

where $H^0(X,\mathcal{F})$ occurs in degree 0 and κ denotes cup-product with the element of $H^1(X,\mathbb{Z}_\ell)$ obtained by pulling back the element ϕ_p of $H^1(\mathrm{Spec}(\mathbb{F}_p),\mathbb{Z}_\ell) = \mathrm{Hom}_{\mathrm{cont}}(\Gamma,\mathbb{Z}_\ell)$ which sends σ to 1. Then the complex $\mathcal{C}(\mathcal{F})^{\cdot}\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ is acyclic if and only if σ acts semi-simply on each space $H^i(X^c,\mathcal{F})\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ [35, Lem. 3.5.3]. Further, in each degree i the homomorphism $H^i(X,\mathcal{F}) \xrightarrow{\kappa} H^{i+1}(X,\mathcal{F})$ is equal to the 'Bockstein homomorphism'

$$\beta^i_{X,\mathcal{F},\sigma} : H^i(X,\mathcal{F}) \to H^{i+1}(X,\mathcal{F})$$

that is obtained as the composite

$$H^i(X,\mathcal{F}) \to H^i(X^c,\mathcal{F})^{\Gamma} \to H^i(X^c,\mathcal{F})_{\Gamma} \to H^{i+1}(X,\mathcal{F})$$

where the first and third maps are induced by the long exact sequence of cohomology associated to the lower row of (8) and the second map is as in (4). Indeed, this equality is a consequence of the description of κ on the level of complexes that is given by Rapaport and Zink in [45, 1.2] (cf. [43, Prop. 6.5] and [35, §3.5.2] in this regard). These equalities imply in turn that $\tau_{X,\mathcal{F},\sigma}$ coincides with the $\mathbb{Q}_\ell$-trivialisation of $R\Gamma(X,\mathcal{F})$ that is induced by the acyclicity of $\mathcal{C}(\mathcal{F})^{\cdot}\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ together with the assignment $\tau \mapsto \tau(\mathcal{C}(\mathcal{F})^{\cdot}\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell)$ which is described just prior to Lemma A1.

We now state the main result of this subsection.

Theorem 3.2. *Let $\pi : X \to Y$ be a finite étale morphism of separated schemes of dimension d over $\mathbb{F}_p$. If π is Galois of group J and r is any integer for* *which σ acts semi-simply on $H^i(X^c, \mathbb{Z}_\ell(r)))\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ in all degrees i, then in $K_0(\mathbb{Z}_\ell[J], \mathbb{Q}_\ell)$ one has*

$$\delta_{J,\ell}(Z_J^*(Y, \pi_*\pi^*\mathbb{Z}_\ell(r)\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell, 1)^{\#}) = -\chi_{J,\ell}(R\Gamma(X, \mathbb{Z}_\ell(d-r)), \tau_{X,\mathbb{Z}_\ell(d-r),\sigma}).$$

Proof. We set $r' := d-r$ and make a choice of morphisms θ and α as in diagram (8) with $\mathcal{F} = \mathbb{Z}_\ell(r')$. Upon applying Lemma A2 to the induced isomorphism $\alpha' : C(\theta)^{\cdot} \xrightarrow{\sim} R\Gamma(X, \mathbb{Z}_\ell(r'))$ and then Proposition 3.1 with $P^{\cdot} = C^{\cdot}$ and $G = J$, we find that

$$\begin{aligned}
&\chi_{J,\ell}(R\Gamma(X, \mathbb{Z}_\ell(r')), \tau_{X,\mathbb{Z}_\ell(r'),\sigma})\\
=&\,\chi_{J,\ell}(C(\theta)^{\cdot}, \tau_\theta)\\
=&\sum_{i\in\mathbb{Z}}(-1)^i\delta_{J,\ell}(\mathrm{detred}^*_{\mathbb{Q}_\ell[J]}(1-\theta : H^i(C^{\cdot})_{\mathbb{Q}_\ell}))\\
=&\sum_{i\in\mathbb{Z}}(-1)^i\delta_{J,\ell}(\mathrm{detred}^*_{\mathbb{Q}_\ell[J]}(1-\sigma : H^i(X^c, \mathbb{Q}_\ell(r')))).
\end{aligned}$$

For each integer i we set $V^i := H^i_c(Y^c, \pi_*\pi^*\mathbb{Z}_\ell(r)\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell) \cong H^i_c(X^c, \mathbb{Q}_\ell(r))$, where subscript '$c$' denotes cohomology with compact support. Then, by Poincaré Duality (cf. [41, Chap. VI, Cor. 11.2]), in each degree i one has an isomorphism of $\mathbb{Q}_\ell[J]$-modules $H^i(X^c, \mathbb{Q}_\ell(r')) \cong \mathrm{Hom}_{\mathbb{Q}_\ell}(V^{2d-i}, \mathbb{Q}_\ell)$. This isomorphism respects the action of Frobenius in the sense that the action of σ on $H^i(X^c, \mathbb{Q}_\ell(r'))$ corresponds to the inverse of the action of σ that is induced on $\mathrm{Hom}_{\mathbb{Q}_\ell}(V^{2d-i}, \mathbb{Q}_\ell)$ by its natural action on V^{2d-i} (since the linear duality functor is contravariant). Hence one has

$$\begin{aligned}
&\mathrm{detred}_{\mathbb{Q}_\ell[J]}(1-\sigma\cdot t : H^i(X^c, \mathbb{Q}_\ell(r')))\\
=&\,\mathrm{detred}_{\mathbb{Q}_\ell[J]}(1-\sigma^{-1}\cdot t : \mathrm{Hom}_{\mathbb{Q}_\ell}(V^{2d-i}, \mathbb{Q}_\ell))\\
=&\,\mathrm{detred}_{\mathbb{Q}_\ell[J]}(1-\sigma^{-1}\cdot t : V^{2d-i})^{\#},
\end{aligned}$$

where the involution $x \mapsto x^{\#}$ acts coefficient-wise on elements of $\zeta(\mathbb{Q}_\ell[J])[[t]]$ and the second equality is valid because J acts contragrediently on $\mathrm{Hom}_{\mathbb{Q}_\ell}(V^{2d-i}, \mathbb{Q}_\ell)$ (cf. [7, (2.0.5)]). From the above formula one therefore has

$$\begin{aligned}
&\chi_{J,\ell}(R\Gamma(X, \mathbb{Z}_\ell(r')), \tau_{X,\mathbb{Z}_\ell(r'),\sigma})\\
=&\,\delta_{J,\ell}(\prod_{i\in\mathbb{Z}}(\mathrm{detred}^*_{\mathbb{Q}_\ell[J]}(1-\sigma^{-1} : V^i))^{\#,(-1)^i})\\
=&-\delta_{J,\ell}((\prod_{i\in\mathbb{Z}}\mathrm{detred}^*_{\mathbb{Q}_\ell[J]}(1-\sigma^{-1} : V^i)^{(-1)^{i+1}})^{\#}).
\end{aligned}$$

To complete the proof it is thus sufficient to observe that, by Grothendieck [32], one has an equality of functions of the complex variable t

$$\prod_{i\in\mathbb{Z}}\mathrm{detred}_{\mathbb{Q}_\ell[J]}(1-\sigma^{-1}\cdot t : V^i)^{(-1)^{i+1}} = Z_J(Y, \pi_*\pi^*\mathbb{Z}_\ell(r)\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell, t).$$

Indeed, the exposition of [41, Chap. VI, proof of Th. 13.3] proves just such an equality with $\mathbb{Q}_\ell[J]$ replaced by an arbitrary finite degree field extension Ω of $\mathbb{Q}_\ell$ and $\pi_*\pi^*\mathbb{Z}_\ell(r)\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ by any constructible sheaf of vector spaces over Ω, and the last displayed equality can be verified by reduction to such cases since both sides are defined via Galois descent (cf. [27, Rem. 2.12]). $\square$

3.3. The case $\ell \neq p$. In this subsection we deduce Theorem 3.1i) from a special case of Theorem 3.2.

To this end we first reinterpret $\mathrm{C}_\ell(K/k)$ in the style of Theorem 3.2. We note that the isomorphism ι_ℓ constructed in the following result is as predicted by [30, Conj. 7.2] (with $X = U_{K,S}$ and $n = 1$).

Lemma 3. *There exists a natural isomorphism in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[G])$ of the form*

$$\iota_\ell : R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)\otimes\mathbb{Z}_\ell \xrightarrow{\sim} R\Gamma(U_{K,S},\mathbb{Z}_\ell(1))[1]\,.$$

Set $\mathrm{D}_{K,S,\ell} := H^1(\iota_\ell)\circ(\mathrm{D}_{K,S}\otimes\mathbb{Z}_\ell)\circ H^0(\iota_\ell)^{-1}$. Then the inverse of $\mathrm{D}_{K,S,\ell}\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell$ induces a $\mathbb{Q}_\ell$-trivialisation of $R\Gamma(U_{K,S},\mathbb{Z}_\ell(1))$ and $\mathrm{C}_\ell(K/k)$ is valid if and only if in $K_0(\mathbb{Z}_\ell[G],\mathbb{Q}_\ell)$ one has

$$\delta_\ell(Z^*_{K/k,S}(1)^{\#}) = -\chi_\ell(R\Gamma(U_{K,S},\mathbb{Z}_\ell(1)),(-\mathrm{D}_{K,S,\ell}\otimes_{\mathbb{Z}_\ell}\mathbb{Q}_\ell)^{-1}). \tag{9}$$

Proof. Following Lemma 1iii) we fix a bounded complex of finitely generated projective $\mathbb{Z}[G]$-modules $P^\cdot$ that is isomorphic in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ to $R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$. Since $R\Gamma(U_{K,S},\mathbb{Z}_\ell(1))$ is an object of $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[G])$ we may also fix a bounded complex of finitely generated projective $\mathbb{Z}_\ell[G]$-modules $Q^\cdot$ that is isomorphic in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[G])$ to $R\Gamma(U_{K,S},\mathbb{Z}_\ell(1))$.

For each natural number n we consider the following diagram

$$\begin{array}{ccccccc}
X_{K,S}\otimes\mathbb{Q}[-2] & \longrightarrow & R\Gamma(U_{K,S},\mathbb{G}_m) & \longrightarrow & P^\cdot & \longrightarrow & X_{K,S}\otimes\mathbb{Q}[-1]\\
\ell^n\big\downarrow & & \ell^n\big\downarrow & & \ell^n\big\downarrow & & \\
X_{K,S}\otimes\mathbb{Q}[-2] & \longrightarrow & R\Gamma(U_{K,S},\mathbb{G}_m) & \longrightarrow & P^\cdot & \longrightarrow & X_{K,S}\otimes\mathbb{Q}[-1]\\
\big\downarrow & & \big\downarrow & & \big\downarrow & & \\
0 & & Q^\cdot/\ell^n[1] & & P^\cdot/\ell^n. & &
\end{array}$$

The first two rows of this diagram are the distinguished triangles that are induced by Lemma 1ii) and the isomorphism $P^\cdot \cong R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$. In addition, all columns of the diagram are distinguished triangles: the first is obviously so, the second is the triangle which is induced by the exact sequence of étale sheaves $1\to\mu_{\ell^n}\to\mathbb{G}_m\xrightarrow{\ell^n}\mathbb{G}_m\to 1$, the exact sequence of étale pro-sheaves $0\to\mathbb{Z}_\ell(1)\xrightarrow{\ell^n}\mathbb{Z}_\ell(1)\to\mu_{\ell^n}\to 1$, the isomorphism $Q^\cdot\cong R\Gamma(U_{K,S},\mathbb{Z}_\ell(1))$ and the exact sequence of modules $0\to Q^i\xrightarrow{\ell^n}Q^i\to Q^i/\ell^n\to 0$ in each degree i, and the third column is the distinguished triangle which is induced by the exact sequence of modules $0\to P^i\xrightarrow{\ell^n}P^i\to P^i/\ell^n\to 0$ in each degree i. Since the diagram commutes in $\mathcal{D}(\mathbb{Z}[G])$ and all rows and columns are distinguished triangles, one can deduce the existence of an isomorphism $\alpha_n : Q^\cdot/\ell^n[1] \cong P^\cdot/\ell^n$ in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}/\ell^n[G])$. Further, as n varies, the isomorphisms α_n may be chosen to be compatible with the natural transition morphisms (cf. [12, the proof of Prop. 3.3]). The inverse limit of such a compatible system of isomorphisms $\{\alpha_n\}_n$ then gives an isomorphism in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[G])$ of the form $R\Gamma(U_{K,S}, \mathbb{Z}_\ell(1))[1] \cong Q^\cdot[1] \cong \varprojlim_n Q^\cdot/\ell^n[1] \cong \varprojlim_n P^\cdot/\ell^n \cong P^\cdot \otimes \mathbb{Z}_\ell \cong R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \otimes \mathbb{Z}_\ell$, as required.

Taken in conjunction with Lemma A2 the quasi-isomorphism ι_ℓ implies that

$$\begin{aligned}
&\rho_\ell(\chi_{\mathbb{Q}}(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{D}_{K,S} \otimes \mathbb{Q}))\\
=&\chi_\ell(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \otimes \mathbb{Z}_\ell, \mathrm{D}_{K,S} \otimes \mathbb{Q}_\ell)\\
=&\chi_\ell(R\Gamma(U_{K,S}, \mathbb{Z}_\ell(1))[1], \mathrm{D}_{K,S,\ell} \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell)\\
=&-\chi_\ell(R\Gamma(U_{K,S}, \mathbb{Z}_\ell(1)), (-\mathrm{D}_{K,S,\ell} \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell)^{-1}),
\end{aligned}$$

where the last equality follows from [9, Th. 2.1(3)]. To prove the final assertion of the lemma we need therefore only observe that $\rho_\ell(\delta(Z^*_{K/k,S}(1)^\#)) = \delta_\ell(Z^*_{K/k,S}(1)^\#)$. Indeed, this equality follows from the fact that on $\zeta(\mathbb{Q}[G])^\times$ one has $\rho_\ell \circ \delta = \delta_\ell \circ i_\ell$ where i_ℓ denotes the natural inclusion $\zeta(\mathbb{Q}[G])^\times \to \zeta(\mathbb{Q}_\ell[G])^\times$. $\square$

To prove $\mathrm{C}_\ell(K/k)$ we need only show that (9) coincides with the formula of Theorem 3.2 in the case $X = U_{K,S}$, $Y = U_{k,S}$ (so that $d = 1$), $\pi : U_{K,S} \to U_{k,S}$ is the natural morphism of spectra, $J = G$ and $r = 0$.

We first compare the left hand sides of the respective formulas. If y is any closed point of $U_{k,S}$, then, after fixing a $\overline{y}$ point x of $U_{K,S}$ and writing G_x for the decomposition subgroup of x in G, the stalk of $\pi_*\pi^*\mathbb{Z}_\ell(0) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$ at $\overline{y}$ identifies as a (left) $G \times G_x$-module with $\mathbb{Q}_\ell[G]$ where elements of the form $(g, \mathrm{id}) \in G \times G_x$ act via left multiplication by g and elements of the form $(\mathrm{id}, g_x) \in G \times G_x$ act via right multiplication by g_x^{-1} (in this regard compare the discussion of [7, beginning of §2]). By using this identification one computes that $Z_G(U_{k,S}, \pi_*\pi^*\mathbb{Z}_\ell(0) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell, t)$ has the same Euler factor at y as does $Z_{K/k,S}(t)$. Since this is true for all closed points y it follows that there is an equality of functions of the complex variable t

$$Z_{K/k,S}(t) = Z_G(U_{k,S}, \pi_*\pi^*\mathbb{Z}_\ell(0) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell, t).$$

This implies that the left hand side of (9) is equal to the left hand side of the relevant special case of the formula in Theorem 3.2. Hence our proof of (9) will be complete if we can verify the relevant semi-simplicity hypothesis (in order to apply Theorem 3.2) and then prove that the trivialisation $\tau_{U_{K,S},\mathbb{Z}_\ell(1),\sigma}$ is induced by the isomorphism $(-D_{K,S,\ell} \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell)^{-1}$. Our proof is therefore completed by combining the description of $\tau_{U_{K,S},\mathbb{Z}_\ell(1),\sigma}$ in Remark 7 together with the following result.

Lemma 4. i) *σ acts semi-simply on $H^i(U^c_{K,S}, \mathbb{Z}_\ell(1)) \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$ in all degrees i.*
ii) $\beta^1_{U_{K,S},\mathbb{Z}_\ell(1),\sigma} = -\mathrm{D}_{K,S,\ell}$.

Proof. Lemma 3 combines with Lemma 1iii) to imply that $R\Gamma(U_{K,S}, \mathbb{Z}_\ell(1))$ is acyclic outside degrees 1 and 2. Remark 7 therefore implies claim i) is equivalent to asserting that the map $\beta^1_{U_{K,S},\mathbb{Z}_\ell(1),\sigma} \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$ is bijective and this is an immediate consequence of the explicit description given in claim ii).

We now fix an arbitrary place v in S and write $c_v : Y_{K,S} \otimes \mathbb{Z}_\ell \to \bigoplus_{w|v} \mathbb{Z}_\ell$ for the homomorphism induced by projecting each element of $Y_{K,S}$ to its respective coefficient at each place w of K above v. Then claim ii) will follow if we show that the composite homomorphism

$$\begin{aligned}(10)\quad \mathcal{O}^\times_{K,S} \otimes \mathbb{Z}_\ell \xrightarrow{H^0(\iota_\ell)} H^1(U_{K,S}, \mathbb{Z}_\ell(1)) \xrightarrow{\beta^1_{U_{K,S},\mathbb{Z}_\ell(1),\sigma}} H^2(U_{K,S}, \mathbb{Z}_\ell(1)) \\ \xrightarrow{H^1(\iota_\ell)^{-1}} X_{K,S} \otimes \mathbb{Z}_\ell \xrightarrow{\subseteq} Y_{K,S} \otimes \mathbb{Z}_\ell \xrightarrow{c_v} \bigoplus_{w|v} \mathbb{Z}_\ell\end{aligned}$$

is equal to $(-\deg(w) \cdot \mathrm{val}_w(-))_{w|v}$. To prove this we set $S' := S \setminus \{v\}$, let Z denote the complement of $U_{K,S}$ in $U_{K,S'}$ and write $j : U_{K,S} \to U_{K,S'}$, resp. $i : Z \to U_{K,S'}$, for the natural open, resp. closed, immersion. Then there exists a natural morphism of étale sheaves $j_*\mathbb{G}_m \to i_*i^*\mathbb{Z}$ on $U_{K,S'}$ that is induced by taking valuations. In turn this gives rise to a morphism $R\Gamma(U_{K,S}, \mathbb{G}_m) \to R\Gamma(Z, \mathbb{Z})$ in $\mathcal{D}(\mathbb{Z}[G])$ and hence, for each non-negative integer n, to a morphism $R\Gamma(U_{K,S}, \mu_{\ell^n}) \to R\Gamma(Z, \mathbb{Z}/\ell^n)[-1]$ in $\mathcal{D}(\mathbb{Z}/\ell^n[G])$. These morphisms are compatible with the natural transition maps as n varies and therefore induce, upon passage to the inverse limit, a morphism in $\mathcal{D}(\mathbb{Z}_\ell[G])$ of the form $\lambda : R\Gamma(U_{K,S}, \mathbb{Z}_\ell(1)) \to R\Gamma(Z, \mathbb{Z}_\ell)[-1]$.

Now $H^0(Z, \mathbb{Z}_\ell) = \bigoplus_{w|v} \mathbb{Z}_\ell$ and each w-component of $H^1(\lambda) \circ H^0(\iota_\ell)$ is induced by the respective valuation map val_w. In addition, if we identify $H^1(Z, \mathbb{Z}_\ell) = \bigoplus_{w|v} \mathrm{Hom}_{\mathrm{cont}}(\mathrm{Gal}(\mathbb{F}^c_p/k(w)), \mathbb{Z}_\ell)$ with $\bigoplus_{w|v} \mathbb{Z}_\ell$ by evaluating each homomorphism at the topological generator $\sigma^{\deg(w)}$ of $\mathrm{Gal}(\mathbb{F}^c_p/k(w))$, then $H^2(\lambda) \circ H^1(\iota_\ell)$ is induced by projection of an element of $X_{K,S}$ to its respective coefficients at each place w above v.

Upon replacing $U_{K,S}$ and Z by $U^c_{K,S}$ and Z^c one obtains in a similar manner a morphism $\lambda^c : R\Gamma(U^c_{K,S}, \mathbb{Z}_\ell(1)) \to R\Gamma(Z^c, \mathbb{Z}_\ell)[-1]$ in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}_\ell[G])$ that induces a morphism of distinguished triangles of the form

$$\begin{array}{ccccc} R\Gamma(U_{K,S}, \mathbb{Z}_\ell(1)) & \longrightarrow & R\Gamma(U^c_{K,S}, \mathbb{Z}_\ell(1)) & \xrightarrow{1-\sigma} & R\Gamma(U^c_{K,S}, \mathbb{Z}_\ell(1)) \\ {\scriptstyle\lambda}\downarrow & & {\scriptstyle\lambda^c}\downarrow & & {\scriptstyle\lambda^c}\downarrow \\ R\Gamma(Z, \mathbb{Z}_\ell)[-1] & \longrightarrow & R\Gamma(Z^c, \mathbb{Z}_\ell)[-1] & \xrightarrow{1-\sigma} & R\Gamma(Z^c, \mathbb{Z}_\ell)[-1]. \end{array}$$

After passing to cohomology this diagram induces a commutative diagram

$$\begin{array}{ccc} H^1(U_{K,S}, \mathbb{Z}_\ell(1)) & \xrightarrow{\beta^1_{U_{K,S},\mathbb{Z}_\ell(1),\sigma}} & H^2(U_{K,S}, \mathbb{Z}_\ell(1)) \\ {\scriptstyle H^1(\lambda)}\downarrow & & {\scriptstyle H^2(\lambda)}\downarrow \\ H^0(Z, \mathbb{Z}_\ell) & \xrightarrow{-\beta^0_{Z,\mathbb{Z}_\ell,\sigma}} & H^1(Z, \mathbb{Z}_\ell), \end{array}$$

where the minus sign in the lower row occurs because of the -1-shift in the lower row of the previous diagram. Now the pull-back to $H^1(Z,\mathbb{Z}_\ell)$ of ϕ_p is the element $(\phi_w)_{w|v}$ where $\phi_w(\sigma^{\deg(w)}) = \deg(w)$ for each w dividing v. After identifying both $H^0(Z,\mathbb{Z}_\ell)$ and $H^1(Z,\mathbb{Z}_\ell)$ with $\bigoplus_{w|v}\mathbb{Z}_\ell$ in the manner prescribed above, the description of Remark 7 (with $X = Z$ and $\mathcal{F} = \mathbb{Z}_\ell$) therefore implies that $\beta^0_{Z,\mathbb{Z}_\ell,\sigma}$ is given by component-wise multiplication with the element $(\deg(w))_{w|v}$. Upon combining the commutativity of this diagram with the explicit descriptions of $H^1(\lambda)$ and $H^2(\lambda)$ given above, it follows that the composite homomorphism (10) is indeed equal to $(-\deg(w)\cdot \mathrm{val}_w(-))_{w|v}$, as required. □

3.4. The case $\ell = p$. In this subsection we prove Theorem 3.1ii).
For each subgroup H of G we let $\rho_H^{G,*}$ denote the natural restriction of scalars homomorphism $K_0(\mathbb{Z}_p[G],\mathbb{Q}_p) \to K_0(\mathbb{Z}_p[H],\mathbb{Q}_p)$. For each abelian group H and each subgroup J of H we also let $q^H_{H/J,*}$ denote the natural coinflation homomorphism $K_0(\mathbb{Z}_p[H],\mathbb{Q}_p) \to K_0(\mathbb{Z}_p[H/J],\mathbb{Q}_p)$. Then one has

$$K_0(\mathbb{Z}_p[G],\mathbb{Q}_p)_{\mathrm{tors}} = \bigcap \ker(q^H_{H/J,*}\circ \rho_H^{G,*})$$

where the intersection runs over all cyclic subgroups H of G and over all subgroups J of H which are such that $p \nmid |H/J|$ [9, Th. 4.1].
Taken in conjunction with the functorial properties of $\mathrm{C}_p(K/k)$ under change of group (Remark 4), the above displayed equality implies that $\mathrm{C}_p(K/k)$ is valid modulo $K_0(\mathbb{Z}_p[G],\mathbb{Q}_p)_{\mathrm{tors}}$ if and only if $\mathrm{C}_p(F/E)$ is valid for each cyclic extension F/E with $k \subseteq E \subseteq F \subseteq K$ and $p \nmid [F:E]$. But, for each such extension F/E, the argument of [7, Lem. 2.2.7] shows that $\mathrm{C}_p(F/E)$ is implied by the Strong-Stark Conjecture for F/E, as formulated by Chinburg (cf. [3, §3.1]). The required result therefore follows directly from Bae's proof of the Strong-Stark Conjecture in this case [3, Th. 3.5.4].
This completes our proof of Theorem 3.1.

4. The conjectures of Chinburg

4.1. Canonical 2-extensions. In the sequel we shall say that two complexes of G-modules $C^\cdot$ and $D^\cdot$ are 'equivalent' if $H^i(C) = H^i(D)$ in each degree i and there exists an isomorphism in $\mathcal{D}(\mathbb{Z}[G])$ from $C^\cdot$ to $D^\cdot$ which induces the identity map in all degrees of cohomology.
If now $C^\cdot$ is any complex of G-modules which is acyclic outside degrees 0 and 1, then $C^\cdot$ is naturally isomorphic in $\mathcal{D}(\mathbb{Z}[G])$ to its double truncation $\tau_{\ge 0}\tau_{\le 1}C^\cdot$. In addition, the tautological exact sequence

$$0 \to H^0(C^\cdot) \to (\tau_{\ge 0}\tau_{\le 1}C^\cdot)^0 \to (\tau_{\ge 0}\tau_{\le 1}C^\cdot)^1 \to H^1(C^\cdot) \to 0$$

determines a unique Yoneda extension class $e(C^\cdot) \in \mathrm{Ext}^2_G(H^1(C^\cdot),H^0(C^\cdot))$.

Lemma 5. *Let $C^\cdot$ and $D^\cdot$ be any complexes of G-modules which are acyclic outside degrees 0 and 1 and are also such that $H^i(C^\cdot) = H^i(D^\cdot)$ for $i = 0,1$. Then $C^\cdot$ and $D^\cdot$ are equivalent if and only if one has $e(C^\cdot) = e(D^\cdot)$.*

Proof. An easy consequence of the definition of equivalence of Yoneda extensions. □

This result implies that $R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$ corresponds to a unique element $c_{\mathcal{W},S}(K/k)$ of $\mathrm{Ext}^2_G(H^1_{\mathcal{W}}(U_{K,S},\mathbb{G}_m),H^0_{\mathcal{W}}(U_{K,S},\mathbb{G}_m))$. In this subsection we relate $c_{\mathcal{W},S}(K/k)$ to the canonical extension class which is defined in terms of class field theory by Tate in [49].
To make such a connection we assume that the G-module $A_{K,S}$ is c-t. In this case the displayed short exact sequence in Lemma 1iii) splits (since $\mathrm{Ext}^1_G(X_{K,S},A_{K,S})=0$) and also $\mathrm{Ext}^2_G(A_{K,S},\mathcal{O}^\times_{K,S})=0$ and so there exists a natural isomorphism

$$\iota_S:\mathrm{Ext}^2_G(X_{K,S},\mathcal{O}^\times_{K,S})\xrightarrow{\sim}\mathrm{Ext}^2_G(H^1_{\mathcal{W}}(U_{K,S},\mathbb{G}_m),H^0_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)).$$

We choose a finite set of places W of k which do not belong to S, are each totally split in K/k and are such that $A_{K,S}$ is generated by the classes of places in $W(K)$. We set $S':=S\cup W$ (so that $A_{K,S'}$ is trivial) and we observe that there are natural exact sequences of G-modules of the form

$$0\to X_{K,S}\xrightarrow{\subseteq}X_{K,S'}\to Y_{K,W}\to 0$$

$$0\to\mathcal{O}^\times_{K,S}\xrightarrow{\subseteq}\mathcal{O}^\times_{K,S'}\to Y_{K,W}\to A_{K,S}\to 0.$$

Since $Y_{K,W}$ is a free $\mathbb{Z}[G]$-module these sequences combine to induce an isomorphism of extension groups

$$\iota_{S',S}:\mathrm{Ext}^2_G(X_{K,S'},\mathcal{O}^\times_{K,S'})\xrightarrow{\sim}\mathrm{Ext}^2_G(X_{K,S},\mathcal{O}^\times_{K,S}).$$

In the sequel we shall identify Yoneda-Ext-groups with derived functor Ext-groups by means of a projective resolution of the first variable (this convention differs from that used in [12] - see in particular [loc. cit., Lem. 3]). We also write $c_{S'}(K/k)$ for the canonical element of $\mathrm{Ext}^2_G(X_{K,S'},\mathcal{O}^\times_{K,S'})$ which is defined in [49].

Proposition 4.1. *If the G-module $A_{K,S}$ is c-t, then one has $c_{\mathcal{W},S}(K/k)=\iota_S\circ\iota_{S',S}(-c_{S'}(K/k))$.*

Proof. For each $w\in S'(K)$ we set $V_w:=\mathrm{Spec}(K_w)$. We also let j' denote the natural open immersion $U_{K,S'}\to C_K$ and we consider the following diagram in $\mathcal{D}(\mathbb{Z}[G])$

$$\begin{array}{ccccc}
X_{K,S'}\otimes\mathbb{Q}[-2] & \longrightarrow & R\Gamma(U_{K,S'},\mathbb{G}_m) & \longrightarrow & R\Gamma_{\mathcal{W}}(U_{K,S'},\mathbb{G}_m)\\
\downarrow & & \downarrow & & \\
Y_{K,S'}\otimes\mathbb{Q}[-2] & \xrightarrow{\alpha} & \oplus_{w\in S'(K)}R\Gamma(V_w,\mathbb{G}_m) & & \\
\downarrow & & \downarrow & & \\
\mathbb{Q}[-2] & & R\Gamma(C_K,j'_!\mathbb{G}_m)[1]. & &
\end{array}$$

The top row of this diagram is the distinguished triangle from Lemma 1ii) (with S replaced by S'), the first column is the distinguished triangle induced by the tautological exact sequence $0 \to X_{K,S'} \xrightarrow{\subseteq} Y_{K,S'} \to \mathbb{Z} \to 0$ and the second column is the distinguished triangle from [42, Chap. II, Prop. 2.3]. Further, under the isomorphism

$$\mathrm{Hom}_{\mathcal{D}(\mathbb{Z}[G])}(Y_{K,S'} \otimes \mathbb{Q}[-2], \oplus_{w\in S'(K)} R\Gamma(V_w, \mathbb{G}_m)) \cong \mathrm{Hom}_G(Y_{K,S'} \otimes \mathbb{Q}, \oplus_{w\in S'(K)} H^2(V_w, \mathbb{G}_m))$$

that is induced by [12, Lem. 7(b)], the morphism α corresponds to the composite of the projection $Y_{K,S'} \otimes \mathbb{Q} \to Y_{K,S'} \otimes \mathbb{Q}/\mathbb{Z}$ and the natural identification $Y_{K,S'} \otimes \mathbb{Q}/\mathbb{Z} \cong \oplus_{w\in S'(K)} H^2(V_w, \mathbb{G}_m)$.

It is straightforward to show that the square in the above diagram commutes (for example, by using [12, Lem. 7(b)] to reduce to cohomological considerations). By comparing this diagram to the diagrams (85) and (88) from loc. cit., and then using the Octahedral axiom, one may therefore conclude that $R\Gamma_{\mathcal{W}}(U_{K,S'}, \mathbb{G}_m)$ is equivalent to the complex $\Psi_{S'}$ which is defined in [12, Prop. 3.1]. From the proof of [12, Prop. 3.5] we may thus deduce that $c_{\mathcal{W},S'}(K/k) = -c_{S'}(K/k)$. (We remark that whilst the results of [12] are phrased solely in terms of number fields, all of the constructions and arguments of loc. cit. extend directly to the case of global function fields. In addition, we obtain $-c_{S'}(K/k)$ rather than $c_{S'}(K/k)$ in the present context because we have changed conventions regarding Yoneda-Ext-groups.)

To conclude that $c_{\mathcal{W},S}(K/k) = \iota_S \circ \iota_{S',S}(-c_{S'}(K/k))$ it suffices to prove that there exists a morphism $R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \to R\Gamma_{\mathcal{W}}(U_{K,S'}, \mathbb{G}_m)$ in $\mathcal{D}(\mathbb{Z}[G])$ which induces upon cohomology the natural maps $\mathcal{O}_{K,S}^\times \xrightarrow{\subseteq} \mathcal{O}_{K,S'}^\times$ and $H^1_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \twoheadrightarrow X_{K,S} \xrightarrow{\subseteq} X_{K,S'}$. But, following [40, the proof of Th. 7.1], the existence of such a morphism can be seen to be a consequence of the morphism of étale sheaves $\mathbb{G}_m \to j_*\mathbb{G}_m$ on $U_{K,S}$ where $j : U_{K,S'} \to U_{K,S}$ denotes the natural open immersion. □

4.2. Galois module theory. In this subsection we relate $\mathrm{C}(K/k)$ to the conjectures formulated by Chinburg in [18, §4.2]. We recall that the conjectures of loc. cit. are natural function field analogues of the central conjectures of Galois module theory which had earlier been formulated by Chinburg in [19, 21]. We write $\Omega(K/k,1), \Omega(K/k,2)$ and $\Omega(K/k,3)$ for the Galois structure invariants defined by Chinburg in [18, the end of §4.1] and $W_{K/k}$ for the so-called 'Cassou-Noguès-Fröhlich Root Number Class' (cf. [loc. cit., Rem. 4.18]).

Conjecture $\mathrm{Ch}(K/k)$ (Chinburg, [18, §4.2, Conj. 3]): *In* $K_0(\mathbb{Z}[G])$ *one has*

i) $\Omega(K/k,1) = 0$,
ii) $\Omega(K/k,2) = W_{K/k}$,
iii) $\Omega(K/k,3) = W_{K/k}$.

We now state the main results of this section.

Theorem 4.1. *The image under* $\partial^0_{\mathbb{Z}[G],\mathbb{R}}$ *of the equality of* $\mathrm{C}(K/k)$ *is equivalent to the equality of* $\mathrm{Ch}(K/k)$iii).

Proof. Following Remark 4, we may consider $\mathrm{C}(K/k)$ with respect to a set S which is large enough to ensure that $A_{K,S}$ is trivial, and in this case Proposition 4.1 (with $S = S'$) implies that $c_{\mathcal{W},S} = \iota_S(-c_S(K/k))$.
Let now $C^\cdot$ and $D^\cdot$ be any objects of $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ which are acyclic outside degrees 0 and 1 and are such that $H^i(C^\cdot) = H^i(D^\cdot)$ for $i = 0, 1$. It is easily shown that if $e(C^\cdot) = -e(D^\cdot)$, then $C^\cdot$ and $D^\cdot$ have the same Euler characteristic in $K_0(\mathbb{Z}[G])$. This observation combines with the equality $c_{\mathcal{W},S} = \iota_S(-c_S(K/k))$ and the very definition of $\Omega(K/k, 3)$ to imply that the latter element can be computed as the Euler characteristic of $R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$ in $K_0(\mathbb{Z}[G])$. It therefore follows that $\Omega(K/k, 3) = \partial^0_{\mathbb{Z}[G],\mathbb{R}}(\chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S}))$.
On the other hand, the same argument as used to prove [7, Lem. 2.3.7] shows that $\partial^0_{\mathbb{Z}[G],\mathbb{R}}(\delta(\theta^*_{K/k,S}(0)^{\#})) = W_{K/k}$. The claimed result is now clear. □

Corollary 2. i) $\mathrm{Ch}(K/k)$iii) *is valid modulo* $\partial^0_{\mathbb{Z}[G],\mathbb{R}}(K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)_{\mathrm{tors}})$.
ii) *If* $p \nmid |G|$, *then* $\mathrm{Ch}(K/k)$ *is valid.*

Proof. Claim i) follows directly from Theorem 4.1 and Corollary 1.
We now assume that $p \nmid |G|$. In this case $K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)$ is torsion-free [13, proof of Lem. 11c)] and hence claim i) implies $\Omega(K/k, 3) = W_{K/k}$. In addition, K/k is tamely ramified and so $\mathrm{Ch}(K/k)$ii) has been proved by Chinburg. Indeed, the equality $\Omega(K/k, 2) = W_{K/k}$ follows directly upon combining [18, §4.2, Th. 4] with [23, Cor. 4.10]. Finally, we observe that the validity of $\mathrm{Ch}(K/k)$i) now follows immediately from the fact that $\Omega(K/k, 1) = \Omega(K/k, 2) - \Omega(K/k, 3)$ [18, §4.1, Th. 2 and the remarks which follow it]. □

Remark 8. The image of $K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)_{\mathrm{tors}}$ under $\partial^0_{\mathbb{Z}[G],\mathbb{R}}$ is equal to the group $D^p(\mathbb{Z}[G])$ that arises in [24, Th. 6.13]. We recall that the arguments of Chinburg in loc. cit., and of Bae in [3] (the results of which provided the key ingredient in our proof of Theorem 3.1ii) in §3.4), rely crucially upon results of Milne and Illusie concerning p-adic cohomology. In particular, in both cases the occurrence of the term $D^p(\mathbb{Z}[G])$ reflects difficulties involved in formulating and proving suitable equivariant refinements of the results of [43].

5. The conjecture of Gross

In this section we assume unless explicitly stated otherwise that G is abelian. We set $G^* := \mathrm{Hom}(G, \mathbb{C}^\times)$ and for each $\chi \in G^*$ we let e_χ denote the associated idempotent $|G|^{-1}\sum_{g\in G}\chi(g)g^{-1}$ of $\mathbb{C}[G]$. In terms of this notation one has $\theta_{K/k,S}(s) = \sum_{\chi\in G^*} e_\chi L_S(\chi, s)$.
We let I_G denote the kernel of the homomorphism $\epsilon : \mathbb{Z}[G] \to \mathbb{Z}$ which sends each element of G to 1.

5.1. Statement of the conjecture. We set $n := |S| - 1$ and let $|n|$, resp. $|n|^*$, denote the set of integers j which satisfy $1 \le j \le n$, resp. $0 \le j \le n$. We henceforth label (and thereby order) the places in S as $\{v_i : i \in |n|^*\}$. For each $j \in |n|^*$ we fix a place w_i of K which restricts to v_i on k. For any place v of k which is unramified in K/k we write σ_v for its frobenius automorphism in G and $\mathrm{N}v$ for the cardinality of the associated residue field. We also fix a finite non-empty set T of places of k which is disjoint from S and then set $\Delta_T := \prod_{v\in T}(1-\sigma_v\cdot \mathrm{N}v) \in \mathbb{Q}[G]^\times$ and

$$\theta_{K/k,S,T}(s) = \Delta_T \cdot \theta_{K/k,S}(s).$$

This $\mathbb{C}[G]$-valued function is holomorphic at $s=0$ and, by using results of Weil, Gross has shown that $\theta_{K/k,S,T}(0)$ belongs to $\mathbb{Z}[G]$ [31, Prop. 3.7].

For any intermediate field F of K/k and any place w of K we let w' denote the restriction of w to F and then write $f_{K/F,w}$ for the homomorphism $F^\times \to G$ which is obtained as the composite of the natural inclusion $F^\times \to F^\times_{w'}$, the reciprocity map $F^\times_{w'} \to \mathrm{Gal}(K_w/F_{w'})$ and the natural injection $\mathrm{Gal}(K_w/F_{w'}) \to G$. We also write $\mathcal{O}^\times_{F,S,T}$ for the subgroup of $\mathcal{O}^\times_{F,S}$ consisting of those $S(F)$-units which are congruent to 1 modulo all places in $T(F)$. It is known that each such group $\mathcal{O}^\times_{F,S,T}$ is torsion-free. In particular, after choosing an ordered $\mathbb{Z}$-basis $\{u_j : j\in|n|\}$ of $\mathcal{O}^\times_{k,S,T}$, we may define an element of $\mathbb{Z}[G]$ by setting

$$\mathrm{Reg}_{G,S,T} := \det((f_{K/k,w_i}(u_j)-1)_{1\le i,j\le n}).$$

At the same time we also define a rational integer $m_{k,S,T}$ by means of the following equality in $\wedge^n X_{k,S}\otimes\mathbb{R}$

$$(\lim_{s\to 0} s^{-n}\theta_{k/k,S,T}(s))\cdot \wedge_{j\in|n|}(v_j-v_0) = m_{k,S,T}\cdot\lambda_{k,S}(\wedge_{j\in|n|}u_j), \tag{11}$$

where $\lambda_{k,S}$ denotes the isomorphism

$$\wedge^n\mathcal{O}^\times_{k,S,T}\otimes\mathbb{R} \to \wedge^n X_{k,S}\otimes\mathbb{R}$$

induced by the n-th exterior power of the map $-\mathrm{R}_{k,S}$ as defined in (2) (cf. [31, (1.7)]).

Conjecture $\mathrm{Gr}(K/k)$ (Gross, [31, Conj. 4.1]): *One has*

$$\theta_{K/k,S,T}(0) \equiv m_{k,S,T}\cdot \mathrm{Reg}_{G,S,T} \pmod{I_G^{n+1}}.$$

Remark 9. The term $m_{k,S,T}\cdot\mathrm{Reg}_{G,S,T}$ belongs to I_G^n and is, when considered modulo I_G^{n+1}, independent of the chosen ordering of S and of the precise choice of the places $\{w_i : i\in|n|^*\}$ and of the ordered basis $\{u_j : j\in|n|\}$.

5.2. Statement of the main results. At the present time, the best results concerning $\mathrm{Gr}(K/k)$ are due to Tan and to Lee. Specifically, it is known that $\mathrm{Gr}(K/k)$ is valid if either $|G|$ is a power of p [47] or if $|G|$ is coprime to both $|\mu_K|$ and the order of the group of divisors of degree 0 of the curve C_k [39]. However, these results are proved either by reduction to special cases or by induction on $|G|$ and so do not provide an insight into why $\mathrm{Gr}(K/k)$ should be true in general. In contrast, in this section we shall show that Gross's integral regulator mapping $\mathcal{O}^\times_{k,S} \to X_{k,S}\otimes G$ [31, (2.1)] arises as a natural Bockstein homomorphism in Weil-étale cohomology and we shall use this observation to prove the following result.

Theorem 5.1. *If the G-module μ_K is c-t, then* $\mathrm{C}(K/k)$ *implies* $\mathrm{Gr}(K/k)$.

Corollary 3. *If the G-module μ_K is c-t, then* $\mathrm{Gr}(K/k)$ *is valid.*

Proof of Corollary 3. It is easily seen to be enough to prove $\mathrm{Gr}(K/k)$ in the case that $|G|$ is a prime power. The aforementioned result of Tan therefore allows us to assume that $p \nmid |G|$ (so that $K_0(\mathbb{Z}_p[G], \mathbb{Q}_p)$ is torsion-free). But since the G-module μ_K is assumed to be c-t, in this case the validity of $\mathrm{Gr}(K/k)$ follows directly from Theorem 5.1 and Corollary 1. □

Remark 10. The G-module μ_K is c-t if and only if for each prime divisor ℓ of $|G|$ one has either $\ell \nmid |\mu_K|$ or $\ell \nmid [K : k(\mu_K^{(\ell)})]$ where $\mu_K^{(\ell)}$ is the maximal subgroup of μ_K of ℓ-power order. It seems likely that a further development of the method we use to prove Theorem 5.1 will allow the removal of any such hypothesis on μ_K. Indeed, in certain special cases this is already achieved in the present manuscript (cf. Corollary 5).

The proof of Theorem 5.1 will be the subject of the next three subsections.

5.3. The computation of $\chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S})$. In this subsection we assume that the G-module μ_K is c-t, but we do not assume that G is abelian. We set $\mathrm{Tr}_G := \sum_{g\in G} g \in \mathbb{Z}[G]$. For any abelian group A we write $\overline{A}$ in place of A/A_{tors} and for any extension field E of $\mathbb{Q}$ we set $A_E := A \otimes E$. For any homomorphism of abelian groups $\phi : A \to A'$ we also let ϕ_E denote the induced homomorphism $\phi \otimes \mathrm{id}_E : A_E \to A'_E$.
In the following result we let $\mathrm{Cone}(\alpha)$ denote the 'mapping cone' of a particular morphism α in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ - our application of this construction can be made rigorous by the same observation as used in [15, Rem. 5.2].

Lemma 6. *There exists an endomorphism ϕ of a finitely generated free $\mathbb{Z}[G]$-module F which satisfies both of the following conditions.*
Let $F^\cdot$ denote the complex $F \xrightarrow{\phi} F$, where the first term is placed in degree 0.

i) *There exists a distinguished triangle in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ of the form*
$$F^\cdot \xrightarrow{\beta} \mathrm{Cone}(\alpha) \to Q[0] \to F^\cdot[1]$$
where α is the morphism $\mu_K[0] \to R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$ in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ that is induced by the inclusion $\mu_K \subset \mathcal{O}_{K,S}^\times$ and Q is a finite G-module of order coprime to $|G|$.

ii) *The endomorphism ϕ^G of F^G is semi-simple at* 0. *Indeed, there exists an integer d with $d \geq n$ and an ordered $\mathbb{Z}[G]$-basis $\{b_i : 1 \leq i \leq d\}$ of F which satisfies both of the following conditions.*

   a) *The $\mathbb{Z}[G]$-module F_1 which is generated by $\{b_i : i \in |n|\}$ satisfies $F_1^G = \ker(\phi^G)$ and, for each $i \in |n|$, the element $\mathrm{Tr}_G(b_i)$ is a pre-image of $v_i - v_0$ under the composite map*

$$F_1^G \subseteq F^G \twoheadrightarrow \mathrm{cok}(\phi^G) \to H^1(R\operatorname{Hom}_{\mathbb{Z}[G]}(\mathbb{Z}, R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)))$$
$$\cong H^1(R\Gamma_{\mathcal{W}}(U_{k,S}, \mathbb{G}_m)) \twoheadrightarrow X_{k,S},$$

where the third, fourth and fifth maps are induced by $H^1(\beta)$, the isomorphism of Lemma 1iv) (with $J = G$) and the short exact sequence of Lemma 1iii) (with $K = k$) respectively.

b) *The $\mathbb{Z}[G]$-module F_2 which is generated by $\{b_i : n < i \leq d\}$ is such that $\phi^G(F_2^G) \subseteq F_2^G$.*

Proof. We set $C^\cdot := \mathrm{Cone}(\alpha)$. Then, since μ_K is c-t, Lemma 1iii) implies that $C^\cdot$ is an object of $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ which is acyclic outside degrees 0 and 1 and that $H^0(C^\cdot) = \overline{\mathcal{O}_{K,S}^\times}$ and $H^1(C^\cdot) = H^1_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$. It is therefore clear that $C^\cdot$ is equivalent to a complex $\hat{F}^\cdot$ of the form $P \xrightarrow{\psi} F$ where F, resp. P, is a finitely generated free $\mathbb{Z}[G]$-module, resp. a finitely generated $\mathbb{Z}[G]$-module which is both c-t and $\mathbb{Z}$-free, and P is placed in degree 0. Now any such $\mathbb{Z}[G]$-module P is projective [2, Th. 8]. In addition, since the $\mathbb{Q}[G]$-modules $H^0(C^\cdot)_\mathbb{Q}$ and $H^1(C^\cdot)_\mathbb{Q}$ are isomorphic, Wedderburn's Theorem implies that the $\mathbb{Q}[G]$-modules $P_\mathbb{Q}$ and $F_\mathbb{Q}$ are also isomorphic. From Swan's Theorem [26, Th. (32.1)] we may therefore deduce that, for each prime q, the $\mathbb{Z}_q[G]$-modules $P \otimes \mathbb{Z}_q$ and $F \otimes \mathbb{Z}_q$ are isomorphic. We may thus apply Roiter's Lemma [26, (31.6)] to deduce the existence of a $\mathbb{Z}[G]$-submodule P' of P for which the quotient P/P' is finite and of order coprime to $|G|$ and one has an isomorphism of $\mathbb{Z}[G]$-modules $\iota : F \xrightarrow{\sim} P'$. We set $\lambda := \psi \circ \iota \in \mathrm{End}_{\mathbb{Z}[G]}(F)$.

The $\mathbb{Z}$-module $\mathrm{im}(\lambda^G)$ is free and so the exact sequence $0 \to \ker(\lambda^G) \xrightarrow{\subseteq} F^G \xrightarrow{\lambda^G} \mathrm{im}(\lambda^G) \to 0$ splits. Hence we may choose a submodule D of F^G which λ^G maps isomorphically to $\mathrm{im}(\lambda^G)$. We next let T denote the pre-image under the tautological surjection $F^G \to \mathrm{cok}(\lambda^G)$ of the subgroup $\mathrm{cok}(\lambda^G)_{\mathrm{tors}}$. Then the exact sequence $0 \to T \to F^G \to \overline{\mathrm{cok}(\lambda^G)} \to 0$ is also split and so we may choose a submodule D' of F^G which is mapped isomorphically to $\overline{\mathrm{cok}(\lambda^G)}$ under the natural surjection. Now D' and $\ker(\lambda^G)$ have the same $\mathbb{Z}$-rank since $D'_\mathbb{Q} \cong \mathrm{cok}(\lambda^G)_\mathbb{Q} \cong \mathrm{cok}(\lambda)^G_\mathbb{Q} \cong \ker(\lambda)^G_\mathbb{Q} \cong \ker(\lambda^G)_\mathbb{Q}$. The direct sum decompositions $\ker(\lambda^G) \oplus D = F^G = T \oplus D'$ therefore imply that there exists an automorphism ψ' of F^G such that both $\psi'(T) = D$ and $\psi'(D') = \ker(\lambda^G)$. It is then easily checked that $\psi' \circ \lambda^G(D) \subseteq D$ and that $\ker(\psi' \circ \lambda^G) = \ker(\lambda^G)$ is mapped bijectively to $\overline{\mathrm{cok}(\psi' \circ \lambda^G)}$ under the composite of the tautological surjections $F^G \to \mathrm{cok}(\psi' \circ \lambda^G)$ and $\mathrm{cok}(\psi' \circ \lambda^G) \to \overline{\mathrm{cok}(\psi' \circ \lambda^G)}$.

Since F is a free $\mathbb{Z}[G]$-module we may choose an element $\tilde{\psi}$ of $\mathrm{Aut}_{\mathbb{Z}[G]}(F)$ such that $\tilde{\psi}^G = \psi'$. We now set $\phi := \tilde{\psi} \circ \lambda \in \mathrm{End}_{\mathbb{Z}[G]}(F)$ and we let β denote the morphism in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ which corresponds to the morphism from the complex $F^\cdot$ (as described in the statement of the Lemma) to $\hat{F}^\cdot$ that is induced by ι in degree 0 and is equal to $\tilde{\psi}^{-1}$ in degree 1. It is then easily checked that this gives rise to a distinguished triangle of the form stated in i) in which $Q := P/P'$.

Now $\phi^G = \psi' \circ \lambda^G$ and so the above remarks imply both that $\phi^G(D) \subseteq D$ and that the natural map $\ker(\phi^G) \to \overline{\mathrm{cok}(\phi^G)}$ is bijective. We next observe that the decomposition $F^G = \ker(\phi^G) \oplus D$ can be lifted to a direct sum decomposition $F = F_1 \oplus F_2$ in which both F_1 and F_2 are free $\mathbb{Z}[G]$-modules (of ranks n and

$d-n$ respectively), $F_1^G = \ker(\phi^G)$ and $F_2^G = D$. We let κ denote the composite homomorphism described in claim ii)a). Our earlier observations imply that κ is bijective, and so $\{\kappa^{-1}(v_i - v_0) : i \in |n|\}$ is a $\mathbb{Z}$-basis of $F_1^G = \mathrm{Tr}_G(F_1)$. It is then easily shown that there exists a $\mathbb{Z}[G]$-basis $\{b_i : i \in |n|\}$ of F_1 such that $\mathrm{Tr}_G(b_i) = \kappa^{-1}(v_i - v_0)$ for each $i \in |n|$. To complete the proof of claim ii) we simply let $\{b_i : n < i \leq d\}$ denote any choice of (ordered) $\mathbb{Z}[G]$-basis of F_2. $\square$

If M is any finite G-module which is c-t, then $M[0]$ is an object of $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ and we set $\chi(M) := \chi(M[0], \mathrm{id}_0) \in K_0(\mathbb{Z}[G], \mathbb{Q})$ where id_0 denotes the identity map on the zero space. We also set $\mathrm{R}^{\beta}_{K,S} := H^1(\beta)_{\mathbb{R}}^{-1} \circ \mathrm{R}_{K,S} \circ H^0(\beta)_{\mathbb{R}}$. Then upon applying Lemma A2 firstly to the distinguished triangle

$$\mu_K[0] \xrightarrow{\alpha} R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \to \mathrm{Cone}(\alpha) \to \mu_K[1]$$

and then to the distinguished triangle in Lemma 6i) we obtain equalities

$$\begin{aligned} \chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S}) &= \chi(\mathrm{Cone}(\alpha), \mathrm{R}_{K,S}) + \chi(\mu_K) \\ &= \chi(F^{\cdot}, \mathrm{R}^{\beta}_{K,S}) + \chi(Q) + \chi(\mu_K) \\ &= \delta(\mathrm{detred}_{\mathbb{R}[G]}(\langle \mathrm{R}^{\beta}_{K,S}, \phi\rangle_{\iota_1,\iota_2})) + \chi(Q) + \chi(\mu_K), \end{aligned} \tag{12}$$

where ι_1 and ι_2 are any choices of $\mathbb{R}[G]$-equivariant sections to the tautological surjections $F_{\mathbb{R}} \to \mathrm{im}(\phi)_{\mathbb{R}}$ and $F_{\mathbb{R}} \to \mathrm{cok}(\phi)_{\mathbb{R}}$ and the last equality follows from Lemma A1.

5.4. The connection to $\mathrm{Gr}(K/k)$. In this subsection we assume that G is abelian and identify $K_0(\mathbb{Z}[G], \mathbb{R})$ with the multiplicative group of invertible $\mathbb{Z}[G]$-lattices in $\mathbb{R}[G]$ (see Remark A1). In particular, we note that if M is any finite G-module which is c-t, then its (initial) Fitting ideal $\mathrm{Fitt}_{\mathbb{Z}[G]}(M)$ is an invertible ideal of $\mathbb{Z}[G]$ and under the stated identification one has $\chi(M) = \mathrm{Fitt}_{\mathbb{Z}[G]}(M)^{-1}$ in $\mathbb{R}[G]$.

Now $\theta^*_{K/k,S,T}(0)^{\#} = \Delta_T^{\#} \cdot \theta^*_{K/k,S}(0)^{\#}$ and $\Delta_T^{\#} \in \mathrm{Ann}_{\mathbb{Z}[G]}(\mu_K) = \mathrm{Fitt}_{\mathbb{Z}[G]}(\mu_K)$. Hence, in this case, (12) implies that the validity of $\mathrm{C}(K/k)$ is equivalent to the existence of an element x_T of $\mathbb{Q}[G]^{\times}$ which satisfies both

$$\theta^*_{K/k,S,T}(0)^{\#} = x_T \cdot \det_{\mathbb{R}[G]}(\langle \mathrm{R}^{\beta}_{K,S}, \phi\rangle_{\iota_1,\iota_2}) \in \mathbb{R}[G]^{\times} \tag{13}$$

and

$$\mathbb{Z}[G] \cdot x_T = \Delta_T^{\#} \cdot \mathrm{Fitt}_{\mathbb{Z}[G]}(\mu_K)^{-1}\, \mathrm{Fitt}_{\mathbb{Z}[G]}(Q)^{-1} \subseteq \mathrm{Fitt}_{\mathbb{Z}[G]}(Q)^{-1}. \tag{14}$$

We let $G^*_{(0)}$ denote the set of characters $\chi \in G^*$ at which $L_S(\chi, 0) \neq 0$, and we set $e_0 := \sum_{\chi \in G^*_{(0)}} e_\chi$. Then the criterion of [50, Chap. I, Prop. 3.4] implies that $e_0 \in \mathbb{Q}[G]$, that $e_0 \cdot \ker(\phi)_{\mathbb{Q}} = 0$ and hence $e_0\mathrm{det}_{\mathbb{R}[G]}(\langle \mathrm{R}^{\beta}_{K,S}, \phi\rangle_{\iota_1,\iota_2}) = e_0\mathrm{det}_{\mathbb{Z}[G]}(\phi)$, and also that for any $\chi \in G^* \setminus G^*_{(0)}$ one has $e_\chi \cdot \ker(\phi)_{\mathbb{C}} \neq 0$ and so $e_0\mathrm{det}_{\mathbb{Z}[G]}(\phi) = \mathrm{det}_{\mathbb{Z}[G]}(\phi)$. Since $\theta_{K/k,S,T}(0)^{\#} = e_0\theta^*_{K/k,S,T}(0)^{\#}$ we therefore deduce from (13) that

$$\theta_{K/k,S,T}(0)^{\#} = x_T\mathrm{det}_{\mathbb{Z}[G]}(\phi).$$

Now $|Q|$ is coprime to $|G|$ and I_G^n/I_G^{n+1} is annihilated by a power of $|G|$, and so (14) implies that x_T acts naturally on I_G^n/I_G^{n+1}. In addition, Lemma 6ii) implies that the matrix of ϕ with respect to the ordered $\mathbb{Z}[G]$-basis $\{b_i : 1 \le i \le d\}$ of F is a block matrix of the form

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \tag{15}$$

where $A := (A_{ij})_{1\le i,j\le n} \in \mathrm{M}_n(I_G)$, $D \in M_{d-n}(\mathbb{Z}[G])$ and all entries of both B and C belong to I_G. Since $\det(A) \in I_G^n$ one has

$$\det(A)^{\#} \equiv (-1)^n \det(A) \pmod{I_G^{n+1}}$$

and so the above matrix representation combines with the previous displayed equality to imply that

$$\theta_{K/k,S,T}(0) \equiv (-1)^n \epsilon(x_T)\epsilon(\det(D)) \cdot \det(A) \pmod{I_G^{n+1}}. \tag{16}$$

To compute the term $(-1)^n\epsilon(x_T)\epsilon(\det(D))$ we first multiply (13) by Tr_G and obtain an equality

$$\lim_{s\to 0} s^{-n}\theta_{k/k,S,T}(s) = \epsilon(x_T)\mathrm{det}_{\mathbb{R}}(\langle \mathrm{R}^{\beta}_{K,S}, \phi\rangle_{\iota_1,\iota_2})^G.$$

For convenience we fix the sections ι_1 and ι_2 so that ι_1^G is equal to the inverse of the automorphism of $F^G_{2,\mathbb{R}}$ induced by ϕ^G and ι_2^G is the inverse of the composite map $F^G_{1,\mathbb{R}} \subseteq F^G_{\mathbb{R}} \twoheadrightarrow \mathrm{cok}(\phi^G)_{\mathbb{R}}$. Then $(\langle \mathrm{R}^{\beta}_{K,S}, \phi\rangle_{\iota_1,\iota_2})^G = \psi_1 \oplus \psi_2$ where ψ_2 is equal to the restriction of ϕ to $F^G_{2,\mathbb{R}}$ and ψ_1 is the automorphism of $F^G_{1,\mathbb{R}}$ that is obtained as the composite

$$F^G_{1,\mathbb{R}} = \ker(\phi^G)_{\mathbb{R}} \xrightarrow{H^0(\beta)^G_{\mathbb{R}}} (\mathcal{O}^{\times}_{K,S})^G_{\mathbb{R}} \xrightarrow{\mathrm{R}_{K,S}} X^G_{K,S,\mathbb{R}} \xrightarrow{\sigma} X_{k,S,\mathbb{R}} \to F^G_{1,\mathbb{R}}$$

where σ is the bijection induced by the injection $X_{k,S} \to X_{K,S}$ described in Lemma 1iv) (with $J = G$), and the final arrow denotes the inverse of the isomorphism induced by the displayed map in Lemma 6ii)a). Now, with respect to the ordered $\mathbb{Z}$-basis $\{\mathrm{Tr}_G(b_i) : n < i \le d\}$ of F_2^G, each component of the matrix of ψ_2 is the image under ϵ of the corresponding component of D and so

$$\lim_{s\to 0} s^{-n}\theta_{k/k,S,T}(s) = \epsilon(x_T) \cdot \mathrm{det}_{\mathbb{R}}(\psi_1) \cdot \epsilon(\det(D)).$$

On the other hand, the commutative diagram

$$\begin{tikzcd} (\mathcal{O}^{\times}_{K,S})^G_{\mathbb{R}} \arrow[r, "\mathrm{R}_{K,S}"] & X^G_{K,S,\mathbb{R}} \arrow[d, "\sigma"] \\ (\mathcal{O}^{\times}_{k,S})_{\mathbb{R}} \arrow[u, "\subseteq"] \arrow[r, "\mathrm{R}_{k,S}"] & X_{k,S,\mathbb{R}} \end{tikzcd}$$

(cf. [50, Chap. I, §6.5]) combines with the above description of ψ_1 to imply that $\mathrm{det}_{\mathbb{R}}(\psi_1)$ is equal to the determinant of the map $\wedge^n(H^0(\beta)(F_1^G))_{\mathbb{R}} \to (\wedge^n X_{k,S})_{\mathbb{R}}$ induced by $\wedge^n_{\mathbb{R}}\mathrm{R}_{k,S} = (-1)^n\lambda_{k,S}$, as computed with respect to the $\mathbb{R}$-bases $\wedge_{i\in|n|}H^0(\beta)(\mathrm{Tr}_G(b_i))$ and $\wedge_{i\in|n|}(v_i - v_0)$. Hence, if we fix an ordered $\mathbb{Z}$-basis $\{d_i : i \in |n|\}$ of $\overline{\mathcal{O}^{\times}_{k,S}}$, regard $\mathcal{O}^{\times}_{k,S,T}$ as a subgroup of $\overline{\mathcal{O}^{\times}_{k,S}}$ in the natural way, and define elements $a := (a_{ij})_{1\le i,j\le n}$ and $b := (b_{ij})_{1\le i,j\le n}$ of $\mathrm{M}_n(\mathbb{Z})$ by the equalities $u_i = \sum_{j\in|n|} a_{ij}d_j$ and $\mathrm{Tr}_G(H^0(\beta)(b_i)) = \sum_{j\in|n|} b_{ij}d_j$ for each $i \in |n|$, then the last displayed formula implies that

$$\begin{aligned}
&(\lim_{s\to 0} s^{-n}\theta_{k/k,S,T}(s))\cdot \wedge_{j\in|n|}(v_j - v_0)\\
&= \epsilon(x_T)\epsilon(\det(D))\det{}_{\mathbb{R}}(\psi_1)\cdot \wedge_{j\in|n|}(v_j - v_0)\\
&= (-1)^n\epsilon(x_T)\epsilon(\det(D))\cdot \lambda_{k,S}(\wedge_{j\in|n|}H^0(\beta)(\mathrm{Tr}_G(b_j)))\\
&= (-1)^n\epsilon(x_T)\epsilon(\det(D))\det(b)\det(a)^{-1}\cdot \lambda_{k,S}(\wedge_{j\in|n|}u_j).
\end{aligned}$$

Comparing this equality with (11) implies that

$$(-1)^n\epsilon(x_T)\epsilon(\det(D))\det(b)\det(a)^{-1} = m_{k,S,T}$$

and hence that

$$(-1)^n\epsilon(x_T)\epsilon(\det(D)) = m_{k,S,T}\det(a)\det(b)^{-1}.$$

In turn, upon substituting this equality into (16) we obtain a congruence

$$\theta_{K/k,S,T}(0) \equiv m_{k,S,T}\det(a)\det(b)^{-1}\cdot\det(A) \pmod{I_G^{n+1}}. \tag{17}$$

5.5. Bockstein homomorphisms. In this subsection we complete our proof of Theorem 5.1 by showing that the factor $\det(a)\det(b)^{-1}\cdot\det(A)$ which occurs in (17) is equal to $\mathrm{Reg}_{G,S,T}$. The key to our proof of this equality will be the observation that the 'regulator map' $\mathcal{O}_{k,S}^\times \to X_{k,S}\otimes G$ introduced by Gross in [31, (2.1)] arises as a natural Bockstein homomorphism in Weil-étale cohomology (this is Lemma 8). The material in this subsection is strongly influenced by the general philosophy of algebraic height pairings that is developed by Nekovář in [44, §11].

At the outset we let Γ be any finite abelian group and $C^\cdot$ any object of $\mathcal{D}^{\mathrm{fpd}}(\mathbb{Z}[\Gamma])$. Then, upon tensoring $C^\cdot$ with the tautological exact sequence $0 \to I_\Gamma \to \mathbb{Z}[\Gamma] \to \mathbb{Z} \to 0$ we obtain a distinguished triangle in $\mathcal{D}(\mathbb{Z})$ of the form

$$C^\cdot \to C^\cdot_\Gamma \to C \otimes^{\mathbb{L}}_{\mathbb{Z}[\Gamma]} I_\Gamma[1] \to C^\cdot[1],$$

where $C^\cdot_\Gamma := C^\cdot\otimes^{\mathbb{L}}_{\mathbb{Z}[\Gamma]}\mathbb{Z}$. In addition, if $C^\cdot$ is acyclic outside degrees 0 and 1, then there are natural identifications $H^0(C^\cdot_\Gamma) \cong H^0(C^\cdot)^\Gamma$ (induced by the action of Tr_Γ), $H^1(C^\cdot_\Gamma) \cong H^1(C^\cdot)_\Gamma$ and $H^1(C^\cdot \otimes^{\mathbb{L}}_{\mathbb{Z}[\Gamma]} I_\Gamma) \cong H^1(C^\cdot)\otimes_{\mathbb{Z}[\Gamma]} I_\Gamma$. In this case the canonical identification $I_\Gamma/I_\Gamma^2 \cong \Gamma$ therefore combines with the cohomology sequence of the above triangle to induce a 'Bockstein homomorphism'

$$\begin{aligned}
\beta_{C^\cdot,\Gamma} :\ H^0(C^\cdot)^\Gamma \to H^1(C^\cdot \otimes^{\mathbb{L}}_{\mathbb{Z}[\Gamma]} I_\Gamma) &\cong H^1(C^\cdot)\otimes_{\mathbb{Z}[\Gamma]} I_\Gamma\\
&\to \overline{H^1(C^\cdot)_\Gamma}\otimes_{\mathbb{Z}[\Gamma]} (I_\Gamma/I_\Gamma^2) \cong \overline{H^1(C^\cdot_\Gamma)}\otimes\Gamma
\end{aligned}$$

and also an associated pairing

$$\rho_{C^\cdot,\Gamma} : H^0(C^\cdot)^\Gamma \times \mathrm{Hom}_{\mathbb{Z}}(\overline{H^1(C^\cdot_\Gamma)}, \mathbb{Z}) \to I_\Gamma/I_\Gamma^2.$$

In the remainder of this subsection we shall use these constructions in the cases that $\Gamma = G$ and $C^\cdot$ is equal to both $F^\cdot$ (as described in Lemma 6) and

$R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$, and also in the case that Γ is equal to a given decomposition subgroup of G and $C^\cdot$ is a local analogue of $R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$. In the course of so doing we shall always use the $\mathbb{Z}$-basis $\{v_i - v_0 : i \in |n|\}$ to identify $X_{k,S}$ with $\mathrm{Hom}_{\mathbb{Z}}(X_{k,S},\mathbb{Z})$.

Before stating our first result we observe that the action of Tr_G (in each degree) induces an isomorphism in $\mathcal{D}(\mathbb{Z})$ between $F^\cdot_G = F^\cdot \otimes_{\mathbb{Z}[G]} \mathbb{Z}$ and the complex $F^G \xrightarrow{\phi^G} F^G$ in which the first term is placed in degree 0. We shall use this isomorphism to identify $\overline{H^1(F^\cdot_G)}$ with $X_{k,S}$ by means of the map $F^G \to X_{k,S}$ described in Lemma 6ii)a).

Lemma 7. *With respect to the ordered $\mathbb{Z}$-bases $\{\mathrm{Tr}_G(b_i) : i \in |n|\}$ and $\{v_i - v_0 : i \in |n|\}$ of $H^0(F^\cdot)^G$ and $\mathrm{Hom}_{\mathbb{Z}}(\overline{H^1(F^\cdot_G)},\mathbb{Z})$ respectively, the matrix of $\rho_{F^\cdot,G}$ is equal to A (mod $M_n(I_G^2)$).*

Proof. The homomorphism $\beta_{F^\cdot,G}$ can be computed as the composite of the connecting homomorphism in the following commutative diagram

$$\begin{tikzcd}
 & & & H^0(F^\cdot)^G \arrow[d] & \\
0 \arrow[r] & F \otimes_{\mathbb{Z}[G]} I_G \arrow[r, "\subseteq"] \arrow[d, "\phi\otimes_{\mathbb{Z}[G]}\mathrm{id}"] & F \arrow[r, "\cdot\,\mathrm{Tr}_G"] \arrow[d, "\phi"] & F^G \arrow[r] \arrow[d, "\phi^G"] & 0 \\
0 \arrow[r] & F \otimes_{\mathbb{Z}[G]} I_G \arrow[r, "\subseteq"] \arrow[d] & F \arrow[r, "\cdot\,\mathrm{Tr}_G"] & F^G \arrow[r] & 0 \\
 & H^1(F^\cdot)\otimes_{\mathbb{Z}[G]} I_G & & &
\end{tikzcd}$$

with the natural surjection $H^1(F^\cdot)\otimes_{\mathbb{Z}[G]} I_G \to \overline{H^1(F^\cdot_G)} \otimes I_G/I_G^2$. Upon computing the above connecting homomorphism by using the matrix representation of ϕ given in (15), and observing Lemma 6ii) implies that the tautological surjection $F^G \to \overline{\mathrm{cok}(\phi^G)} \cong \overline{H^1(F^\cdot_G)}$ factors through the projection $F^G \to F_1^G$, one finds that the required composite sends each element $\mathrm{Tr}_G(b_i)$ to $\sum_{j\in|n|}(v_j - v_0)\otimes A_{ij}$ (mod I_G^2). This implies the stated result. □

The construction of the pairing $\rho_{C^\cdot,G}$ is natural in $C^\cdot$ in the following sense: if $\mu : C^\cdot \to D^\cdot$ is any morphism in $\mathcal{D}^{\mathrm{fpd}}(\mathbb{Z}[G])$ which induces a bijection $\overline{H^1(\mu_G)}$ from $\overline{H^1(C^\cdot_G)}$ to $\overline{H^1(D^\cdot_G)}$, then there is a commutative diagram

$$\begin{tikzcd}
H^0(C^\cdot)^G \times \mathrm{Hom}_{\mathbb{Z}}(\overline{H^1(C^\cdot_G)},\mathbb{Z}) \arrow[r, "\rho_{C^\cdot,G}"] \arrow[d, "(H^0(\mu)^G,\mathrm{Hom}_{\mathbb{Z}}(\overline{H^1(\mu_G)},\mathbb{Z})^{-1})"'] & I_G/I_G^2 \arrow[d, equal] \\
H^0(D^\cdot)^G \times \mathrm{Hom}_{\mathbb{Z}}(\overline{H^1(D^\cdot_G)},\mathbb{Z}) \arrow[r, "\rho_{D^\cdot,G}"] & I_G/I_G^2.
\end{tikzcd}$$

When taken in conjunction with the computation of Lemma 7 and the fact that multiplication by $\det(b)$ is invertible on I_G^n/I_G^{n+1} (since $|Q^G|$ is coprime to

$|G|$), this observation implies that the term $\det(a)\det(b)^{-1}\cdot\det(A) \pmod{I_G^{n+1}}$ in (17) is equal to the discriminant of the restriction of $\rho_{R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m),G}$ to $\mathcal{O}_{k,S,T}^\times \times X_{k,S}$ as computed with respect to the ordered $\mathbb{Z}$-bases $\{u_i : i \in |n|\}$ and $\{v_i - v_0 : i \in |n|\}$.

To prove Theorem 5.1 we therefore need only show that the homomorphism $\beta_{R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m),G}$ coincides with the regulator mapping $\mathcal{O}_{k,S}^\times \to X_{k,S}\otimes G$ defined by Gross in [31, (2.1)]. In turn, this is achieved by the following result (which, we observe, does not assume that the G-module μ_K is c-t).

Lemma 8. *Set $C^\cdot := R\Gamma_{\mathcal{W}}(U_{K,S},\mathbb{G}_m)$. Then for each $u \in \mathcal{O}_{k,S,T}^\times$ one has $\beta_{C^\cdot,G}(u) = \sum_{i\in|n|}(v_i - v_0)\otimes f_{K/k,w_i}(u)$.*

Proof. We fix an index $i \in |n|$ and set $v := v_i$, $w := w_i$ and $D := \mathrm{Gal}(K_w/k_v)$. We let $\beta_{C^\cdot,G,v}$ denote the composite of $\beta_{C^\cdot,G}$ with the inclusion $X_{k,S}\otimes G \subset Y_{k,S}\otimes G$ and the homomorphism $Y_{k,S}\otimes G \to G$ which is induced by mapping each element of $Y_{k,S}$ to its coefficient at v. Then we need to show that $\beta_{C^\cdot,G,v} = f_{K/k,w}$.

We set $V_w := \mathrm{Spec}(K_w)$. Then the result of [12, Lem. 7(b)] combines with the fact that $H^1(V_w,\mathbb{G}_m) = 0$ to imply that there exists a unique morphism α_w from $\mathbb{Q}[-2]$ to $R\Gamma(V_w,\mathbb{G}_m)$ in $\mathcal{D}(\mathbb{Z}[D])$ for which $H^2(\alpha_w)$ is equal to the composite of the natural projection $\mathbb{Q}\to\mathbb{Q}/\mathbb{Z}$ and the canonical identification $\mathbb{Q}/\mathbb{Z} \cong H^2(V_w,\mathbb{G}_m)$. We set $C_w^\cdot := \mathrm{Cone}(\alpha_w)$ (cf. the remark just prior to Lemma 6). Then, by an argument similar to that used in the proof of Lemma 1iii), one shows that $C_w^\cdot$ is an object of $\mathcal{D}^{\mathrm{fpd}}(\mathbb{Z}[D])$ which is acyclic outside degrees 0 and 1 and is such that $H^0(C_w^\cdot)$ and $H^1(C_w^\cdot)$ identify canonically with $K_w^\times$ and $\mathbb{Z}$ respectively. Further, in the notation of §4.1, the result of [12, Prop. 3.5(a)] implies that the associated Yoneda extension class $e(C_w^\cdot)$ is equal to the element $-e_w$ of $\mathrm{Ext}_D^2(\mathbb{Z},K_w^\times)\cong H^2(D,K_w^\times)$ where $\mathrm{inv}_{k_v}(e_w) = \frac{1}{|D|}$ (recall that, following the approach of §4.1, we are here using a different convention regarding Yoneda-Ext-groups than that used in [12], and hence $e(C_w^\cdot)$ is equal to $-e_w$ rather than e_w.)

The natural localisation morphism $R\Gamma(U_{K,S},\mathbb{G}_m)\to\mathbb{Z}[G]\otimes_{\mathbb{Z}[D]}R\Gamma(V_w,\mathbb{G}_m)$ in $\mathcal{D}(\mathbb{Z}[G])$ induces a morphism $C^\cdot\to\mathbb{Z}[G]\otimes_{\mathbb{Z}[D]}C_w^\cdot$ and by consideration of this morphism one finds that $\beta_{C^\cdot,G,v}$ is equal to the composite of the embedding $\mathcal{O}_{k,S}^\times\to k_v^\times$, the homomorphism $\beta_{C_w^\cdot,D}$ and the natural injection $D\subseteq G$. It is therefore enough for us to prove that $\beta_{C_w^\cdot,D}$ is equal to the reciprocity map $\mathrm{rec}_w : k_v^\times\to D$ of the extension K_w/k_v.

To this end we first recall that rec_w is defined to be the map induced by the inverse of the isomorphism $D\cong\hat{H}^0(D,K_w^\times)$ which results from the canonical identifications $D\cong I_D/I_D^2 = \hat{H}^{-1}(D,I_D)$, the isomorphism $\hat{H}^{-1}(D,I_D)\cong\hat{H}^{-2}(D,\mathbb{Z})$ which is induced by the connecting homomorphism associated to the tautological exact sequence $0\to I_D\to\mathbb{Z}[D]\to\mathbb{Z}\to 0$ and the isomorphism $\hat{H}^{-2}(D,\mathbb{Z})\cong\hat{H}^0(D,K_w^\times)$ which is given by cup-product with e_w.

To proceed we choose an extension of D-modules

$$0\to K_w^\times\xrightarrow{\iota}A\xrightarrow{\psi}\mathbb{Z}[D]\xrightarrow{\epsilon}\mathbb{Z}\to 0$$

of Yoneda extension class $-e_w$. Then $C_w^{\cdot}$ is equivalent to the complex $A^{\cdot}$ which is given by $A \xrightarrow{\psi} \mathbb{Z}[D]$, where the modules are placed in degrees 0 and 1 and the cohomology is identified with $K_w^\times$ and $\mathbb{Z}$ by means of the given maps. Taken in conjunction with the description of rec_w in the preceding paragraph and the compatibility of cup-products with connecting homomorphisms in Tate cohomology (cf. [2, Th. 3 and Th. 4(iii),(iv)]), this observation implies that rec_w is induced by the canonical isomorphism $D \cong I_D/I_D^2 = (I_D)_D$ together with the inverse of the connecting homomorphism in the following commutative diagram

$$\begin{tikzcd}
 & & & (I_D)_D \arrow[d, equal] & \\
 & (K_w^\times)_D \arrow[r, "\iota_D"] \arrow[d, "\mathrm{Tr}_w"] & A_D \arrow[r, "\psi_D"] \arrow[d, "\mathrm{Tr}_w"] & (I_D)_D \arrow[r] \arrow[d, "\mathrm{Tr}_w"] & 0 \\
0 \arrow[r] & (K_w^\times)^D \arrow[r, "\iota^D"] \arrow[d] & A^D \arrow[r, "\psi^D"] & (I_D)^D = 0 & \\
 & \hat{H}^0(D, K_w^\times) & & &
\end{tikzcd}$$

where $\mathrm{Tr}_w := \sum_{d\in D} d \in \mathbb{Z}[D]$. On the other hand, the fact that $C_w^{\cdot}$ is equivalent to $A^{\cdot}$ combines with the definition of $\beta_{C_w^{\cdot},D}$ to imply that the latter homomorphism can be computed as the composite of the natural identification $D \cong (I_D)_D$ and the connecting homomorphism in the following commutative diagram

$$\begin{tikzcd}
 & & & k_v^\times \arrow[d, "\iota^D"] & \\
 & A \otimes_{\mathbb{Z}[D]} I_D \arrow[r] \arrow[d, "\psi\otimes_{\mathbb{Z}[G]}\mathrm{id}"] & A \arrow[r, "\mathrm{Tr}_w"] \arrow[d, "\psi"] & A^D \arrow[r] \arrow[d, "\psi^D"] & 0 \\
0 \arrow[r] & I_D \arrow[r, "\subset"] \arrow[d] & \mathbb{Z}[D] \arrow[r, "\mathrm{Tr}_w"] & \mathbb{Z}\cdot \mathrm{Tr}_w \arrow[r] & 0 \\
 & (I_D)_D. & & &
\end{tikzcd}$$

We remark that the upper row of this diagram is exact since the D-module A is c-t. Our proof now concludes by means of an explicit diagram chase showing that the connecting homomorphism in the second of these diagrams induces the inverse of the connecting homomorphism in the first diagram. $\square$

6. The conjecture of Tate

In this section we provide evidence for Tate's refinement of $\mathrm{Gr}(K/k)$. To do so we continue to use the notation of §5.1. In addition, we fix a prime number ℓ and assume henceforth that G has order ℓ^m with $m \geq 1$. For each index j in $|n|^*$ we let G_j denote the decomposition subgroup of w_j in G and we define an integer m_j by the equality $|G_j| = \ell^{m-m_j}$.

6.1. Statement of the conjecture. In this subsection we assume S to be ordered so that $m_0 \leq m_1 \leq \cdots \leq m_n$.

Conjecture $\mathrm{Ta}(K/k, S, T)$ (Tate, [51]): *If G is cyclic of order ℓ^m, $m_0 = 0$ and $m_n = m - 1$, then one has*

$$\theta_{K/k,S,T}(0) \equiv m_{k,S,T} \cdot \mathrm{Reg}_{G,S,T} \pmod{I_G^{(\sum_{i=0}^{n-1} \ell^{m_i})+1}}.$$

For a further discussion of this conjecture see, for example, [38, §4].

6.2. Statement of the main results. We recall from §4.1 that if the G-module $A_{K,S} := \mathrm{Pic}(\mathcal{O}_{K,S})$ is c-t, then one can define a canonical element $c_S(K/k) := \iota_{S',S}(c_{S'}(K/k))$ of $\mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S})$, where S' is any set as described in §4.1 (and $c_S(K/k)$ is indeed independent of the choice of S').
For each index j in $|n|$ we write I_j for the kernel of the natural projection map $\mathbb{Z}[G] \to \mathbb{Z}[G/G_j]$. We consider the following hypothesis on K/k.

Hypothesis (S,T): There exist finite non-empty sets S and T of places of k which satisfy each of the following conditions:

i) S contains all places which ramify in K/k,
ii) the G-module $A_{K,S}$ is c-t,
iii) $G_0 = G$, $n > 0$ and G_j is cyclic for each $j \in |n|$,
iv) T is disjoint from S and $c_S(K/k)$ lies in the image of the map

$$\mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S,T}) \to \mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S})$$

induced by the inclusion $\mathcal{O}^\times_{K,S,T} \subset \mathcal{O}^\times_{K,S}$.

Remark 11. If K/k is cyclic, then there always exists a set of places S which satisfies conditions i), ii) and iii) above. In general however, for a given field K there are restrictions on the abstract structure of the decomposition group G_0 and therefore (under condition iii)) also on G. Nevertheless, the validity of Hypothesis (S,T) does not itself imply, for example, that G is abelian. If $\ell \nmid |\mu_k|$, then (since $|G|$ is a power of ℓ) one has $\ell \nmid |\mu_K|$ and so [50, Chap. IV, Lem. 1.1] implies that there exists a set T which is disjoint from S and satisfies $\ell \nmid [\mathcal{O}^\times_{K,S} : \mathcal{O}^\times_{K,S,T}]$ and hence also condition iv). In fact, condition iv) can be shown to be satisfied under reasonably general conditions even if $\ell \mid |\mu_k|$ (cf. [17, Lem. 2]).

The following result will be proved in §6.4.

Theorem 6.1. *If S and T are as in Hypothesis (S,T) and G is abelian, then* $\mathrm{C}(K/k)$ *implies that*

$$\theta_{K/k,S,T}(0) \equiv m_{k,S,T} \cdot \mathrm{Reg}_{G,S,T} \pmod{I_G \cdot \prod_{j \in |n|} I_j}.$$

Corollary 4. *Assume the notation and hypotheses of* $\mathrm{Ta}(K/k, S, T)$*. If the G-module* $\mathrm{Pic}(\mathcal{O}_{K,S})$ *is c-t and $c_S(K/k)$ lies in the image of the map*

$$\mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S,T}) \to \mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S})$$

induced by the inclusion $\mathcal{O}^\times_{K,S,T} \subset \mathcal{O}^\times_{K,S}$, then $\mathrm{C}(K/k)$ *implies that*

$$\theta_{K/k,S,T}(0) \equiv m_{k,S,T} \cdot \mathrm{Reg}_{G,S,T} \pmod{I_G \cdot \prod_{j \in |n|} I_j}.$$

In particular, in this case $\mathrm{Ta}(K/k, S, T)$ *is valid.*

Proof. Since, by assumption, $m_0 = 0$ the sets S and T satisfy all parts of Hypothesis (S,T). The first assertion thus follows directly from Theorem 6.1. To prove the second assertion we recall that if $\ell = p$, then $\mathrm{Ta}(K/k, S, T)$ has been proved by Tan [48]. We may therefore assume that $\ell \neq p$ so that $\mathrm{C}(K/k)$ is valid by Corollary 1. It thus suffices to deduce the validity of $\mathrm{Ta}(K/k, S, T)$ from the stated congruence for $\theta_{K/k,S,T}(0)$ and this is true because $\prod_{j\in|n|} I_j \subseteq I_G^{\sum_{i=0}^{n-1} \ell^{m_i}}$. Indeed, since $m_n = m - 1$, the required inclusion follows directly from the criterion of [8, Lem. 5.11]. □

The next result improves upon Corollary 3 and also the main result of Lee in [37].

Corollary 5. *If G has prime exponent, then* $\mathrm{Gr}(K/k)$ *is valid.*

Proof. In this case, the functorial properties of $\theta_{K/k,S,T}(0)$ and $\mathrm{Reg}_{G,S,T}$ under change of K/k combine with results on the structure of I_G/I_G^{n+1} to show that it is enough to prove $\mathrm{Gr}(L/k)$ for each sub-extension L/k of K/k which is of prime degree. The theorem of Tan [47] also allows us to assume that $[L : k]$ is a prime number different from p, and in this case the required congruence can be proved by combining the result of Corollary 4 (with $K = L$) together with arguments of Gross from [31, §6]. The precise details of this argument are presented in joint work of the author with Lee [17]. □

6.3. $\chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S})$ revisited. In this subsection we prepare for the proof of Theorem 6.1 by using Hypothesis (S,T) to refine the computation of $\chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S})$ given in §5.3. We do not assume here that G is abelian or that the G-module μ_K is c-t.

At the outset we fix sets S and T as in Hypothesis (S,T). Since S is fixed we abbreviate $\mathcal{O}^\times_{K,S,T}, X_{K,S}$ and $\mathcal{O}^\times_{K,S}$ to $\mathcal{O}^\times_{K,T}, X_K$ and $\mathcal{O}^\times_K$ respectively. We also set $A_K := \mathrm{Pic}(\mathcal{O}_{K,S})$ and write $A_{K,T}$ for the quotient of the group of fractional

ideals of $\mathcal{O}_{K,S}$ that are prime to T by the subgroup of principal ideals with a generator congruent to 1 modulo all places in $T(K)$.
For each $j \in |n|$ we fix a generator g_j of G_j and a set of representatives $S(j)$ of the orbits of G_j on the set of places of K lying above $\{v_i : i \in |n|\}$. We assume that $S(j)$ contains w_i for each $i \in |n|$. For each place w in $S(j)$ we define δ_{jw} to be 1 if $w = w_j$ and to be 0 otherwise. For each $j \in |n|$ we also set $\mathrm{Tr}_j := \sum_{g \in G_j} g \in \mathbb{Z}[G]$ and $K_j := K^{G_j}$.
If d is any strictly positive integer, then in the sequel we shall use the canonical basis of $\mathbb{R}[G]^d$ to identify the groups $\mathrm{GL}_d(\mathbb{R}[G])$ and $\mathrm{Aut}_{\mathbb{R}[G]}(\mathbb{R}[G]^d)$.

Proposition 6.1. *Let S and T be as in Hypothesis (S,T). Assume also that G_j is not trivial for any $j \in |n|$. Then, for each $j \in |n|$ there exists an element ϵ_j of $\mathcal{O}^\times_{K_j,T}$ which satisfies all of the following conditions.*

i) *For each $w \in S(j)$ one has $f_{K/K_j,w}(\epsilon_j) = g_j^{\delta_{jw}}$.*
ii) *For each pair of integers i, j in $|n|$ let y_{ji} denote the (unique) element of $\mathbb{R}[G] \cdot \mathrm{Tr}_i$ which satisfies*
$$\frac{1}{|G_j|}\mathrm{R}_{K,S}(\epsilon_j) = \sum_{i \in |n|} y_{ji}(w_i - w_0).$$
Then the matrix $M_T := (\delta_{ij}(g_i - 1) + y_{ij})_{1 \le i,j \le n}$ belongs to $\mathrm{GL}_n(\mathbb{R}[G])$.
iii) *The G-module $\mathcal{E}$ that is generated by the set $\{\epsilon_j : j \in |n|\}$ has finite index in $\mathcal{O}^\times_{K,T}$. The G-modules $\mathcal{O}^\times_{K,S}/\mathcal{E}$ and $A_{K,T}$ are both c-t and in $K_0(\mathbb{Z}[G], \mathbb{R})$ one has*

$$\chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S}) = \chi(\mathcal{O}^\times_{K,T}/\mathcal{E}) - \chi(A_{K,T}) + \delta(\mathrm{detred}_{\mathbb{R}[G]}(M_T)) - \delta(\mathrm{detred}_{\mathbb{R}[G]}(\Delta_T^{\#})).$$

To prove this result we let $\hat{\Psi}^{\cdot}$ denote any complex of G-modules of the form $\hat{\Psi}^0 \xrightarrow{d} \hat{\Psi}^1$ where $\hat{\Psi}^0$ occurs in degree 0 and $e(\hat{\Psi}^{\cdot}) = c_{\mathcal{W},S}(K/k)$ in the notation of §4.1. We write Ψ^1 for the pullback of the natural surjection $\hat{\Psi}^1 \to H^1_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m)$ and a choice of section γ to the surjection $H^1_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \to X_K$ provided by Lemma 1iii) (such a section always exists under Hypothesis (S,T)ii)). In this way we obtain a complex $\Psi^{\cdot}$ of the form $\hat{\Psi}^0 \xrightarrow{\hat{d}} \Psi^1$ which satisfies $e(\Psi^{\cdot}) = \iota_S^{-1}(c_{\mathcal{W},S}(K/k)) \in \mathrm{Ext}^2_G(X_K, \mathcal{O}^\times_K)$ and lies in a distinguished triangle in $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ of the form

$$\Psi^{\cdot} \xrightarrow{\alpha} R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m) \to A_K[-1] \to \Psi^{\cdot}[1],$$

where $H^0(\alpha)$ is the identity map and $H^1(\alpha) = \gamma$. Upon applying Lemma A2 to this triangle we obtain an equality

$$\chi(R\Gamma_{\mathcal{W}}(U_{K,S}, \mathbb{G}_m), \mathrm{R}_{K,S}) = \chi(\Psi^{\cdot}, \mathrm{R}_{K,S}) - \chi(A_K). \tag{18}$$

To compute $\chi(\Psi^{\cdot}, \mathrm{R}_{K,S})$ we shall first be more explicit about the computation of the group $\mathrm{Ext}^2_G(X_K, \mathcal{O}^\times_K)$. For each $j \in |n|$ one has an exact sequence

$$0 \to \mathbb{Z}[G] \cdot \mathrm{Tr}_j \xrightarrow{\subseteq} \mathbb{Z}[G] \xrightarrow{d_j} \mathbb{Z}[G] \xrightarrow{\theta_j} \mathbb{Z}[G](w_j - w_0) \to 0$$

where $d_j(x) = (g_j - 1)x$ and $\theta_j(x) = x(w_j - w_0)$ for each $x \in \mathbb{Z}[G]$. By taking the direct sum of these sequences over j in $|n|$ we obtain a resolution of $X_K = \bigoplus_{j\in|n|} \mathbb{Z}[G](w_j - w_0)$ of the form $0 \to \Sigma_K \xrightarrow{\subseteq} F \xrightarrow{d} F \xrightarrow{\theta} X_K \to 0$ in which $\Sigma_K := \bigoplus_{j\in|n|} \mathbb{Z}[G] \cdot \mathrm{Tr}_j, F = \bigoplus_{j\in|n|} \mathbb{Z}[G], d = \bigoplus_{j\in|n|} d_j$ and $\theta = \bigoplus_{j\in|n|} \theta_j$. When computing $\mathrm{Ext}^2_G(X_K, \mathcal{O}_K^\times)$ with respect to this resolution, we may choose an injective G-homomorphism $\phi : \Sigma_K \to \mathcal{O}_K^\times$ which represents $\iota_S^{-1}(c_{\mathcal{W},S}(K/k))$ [4, Lem. 2.4]. In addition, from Proposition 4.1, one has $\iota_S^{-1}(c_{\mathcal{W},S}(K/k)) = -c_S(K/k)$ and so Hypothesis (S,T)iv) allows us to assume that ϕ factors through a homomorphism $\phi_T : \Sigma_K \to \mathcal{O}_{K,T}^\times$. In this case one has $\phi_T = \bigoplus_{i\in|n|} \phi_j$ with $\phi_j \in \mathrm{Hom}_{\mathbb{Z}[G]}(\mathbb{Z}[G] \cdot \mathrm{Tr}_j, \mathcal{O}_{K,T}^\times)$ and so we set $\epsilon_j := \phi_j(\mathrm{Tr}_j) \in (\mathcal{O}_{K,T}^\times)^{G_j} = \mathcal{O}_{K_j,T}^\times$. Then, since ϕ represents $-c_S(K/k)$, the descriptions of [22, Prop. 3.2.1, Prop. 4.5.2] imply that condition i) is satisfied. (Note that the result of [22, Prop. 4.5.2] should state that $\psi_w(t_v(c))$ is equal to g^{-1} rather than g. Indeed, the compatibility of cup-product with connecting homomorphisms in Tate cohomology implies that (in the notation of the proof given in loc. cit.) cup-product with β is equal to the *negative* of the composite of the connecting homomorphisms $H^{-2}(G_v, \mathbb{Z}) \to H^{-1}(G_v, I_\Gamma a)$ and $H^{-1}(G_v, I_\Gamma a) \to H^0(G_v, \mathbb{Z}c)$ described there. See also the proof of [20, Cor. 2.1] in this regard.)

We next let $\hat{F}$ denote the push-out of ϕ and the inclusion map $\Sigma_K \xrightarrow{\subseteq} F$, and we write $F^\cdot$ and $\hat{F}^\cdot$ for the complexes $F \xrightarrow{d} F$ and $\hat{F} \xrightarrow{\hat{d}} F$ where (in both cases) the modules are placed in degrees 0 and 1 and $\hat{d}$ denotes the morphism induced by d. Then Lemma 5 combines with our choice of ϕ to imply that the complexes $\hat{F}^\cdot$ and $\Psi^\cdot$ are equivalent and hence there exists a distinguished triangle in $\mathcal{D}(\mathbb{Z}[G])$ of the form

$$F^\cdot \xrightarrow{\beta} \Psi^\cdot \to \mathrm{cok}(\phi)[0] \to F^\cdot[1]$$

in which $H^0(\beta) = \phi$ and $H^1(\beta)$ is the identity map. Note that since both $F^\cdot$ and $\Psi^\cdot$ belong to $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$ this triangle implies that $\mathrm{cok}(\phi)[0]$ (and hence also the triangle itself) belongs to $\mathcal{D}^{\mathrm{perf}}(\mathbb{Z}[G])$. In particular, it follows that the G-module $\mathrm{cok}(\phi)$ is both finite (since ϕ is injective) and c-t. In addition, we may apply Lemma A2 to the triangle to deduce that

$$\chi(\Psi^\cdot, \mathrm{R}_{K,S}) = \chi(F^\cdot, \mathrm{R}_{K,S} \circ \phi) + \chi(\mathrm{cok}(\phi)). \tag{19}$$

To compute $\chi(F^\cdot, \mathrm{R}_{K,S} \circ \phi)$ we observe that the differential of $F^\cdot$ is semisimple at 0, when considered as an endomorphism of F. Indeed, the submodule $D := \bigoplus_{j\in|n|} \mathbb{Q}[G]\cdot(g_j - 1)$ is a $\mathbb{Q}[G][d]$-equivariant direct complement to $\Sigma_K \otimes \mathbb{Q}$ in $F \otimes \mathbb{Q}$. We may therefore apply Lemma A1 with $P = F, R = \mathbb{Z}, E = \mathbb{R}, \phi = d, \lambda = \mathrm{R}_{K,S} \circ \phi$ and with ι_1, ι_2 equal to the sections which are induced by D. In this context, the definition of the elements y_{ij} in the statement of claim ii) implies that the restriction of the automorphism $\langle \lambda, \phi \rangle_{\iota_1,\iota_2}$ which occurs in Lemma A1 to the direct summand $\Sigma_K \otimes \mathbb{R}$, resp. $D \otimes_{\mathbb{Q}} \mathbb{R}$, of $F \otimes \mathbb{R}$ is the map which sends each Tr_j to $\sum_{i\in|n|} y_{ji} \mathrm{Tr}_i$, resp. is the map which is induced by multiplication by (g_j-1) on each summand $\mathbb{R}[G]\cdot(g_j-1)$. It follows that, with respect to a suitable ordered $\mathbb{R}[G]$-basis of $F\otimes\mathbb{R}$, the matrix of $\langle\lambda,\phi\rangle_{\iota_1,\iota_2}$ is equal to M_T and hence that M_T is invertible, as required by claim ii). In addition, in this case Lemma A1 implies that $\chi(F^{\cdot},\mathrm{R}_{K,S}\circ\phi)=\delta(\mathrm{detred}_{\mathbb{R}[G]}(M_T))$. Our proof of Proposition 6.1 is thus completed by combining this equality with (18), (19) and the following two results.

Lemma 9. *If the G-module $\mathrm{cok}(\phi_T)=\mathcal{O}_{K,T}^{\times}/\mathcal{E}$ is c-t, then so also are $\mathrm{cok}(\phi)=\mathcal{O}_K^{\times}/\mathcal{E}$, A_K and $A_{K,T}$, and in $K_0(\mathbb{Z}[G],\mathbb{R})$ one has*

$$\chi(\mathrm{cok}(\phi))-\chi(A_K)=\chi(\mathrm{cok}(\phi_T))-\chi(A_{K,T})-\delta(\mathrm{detred}_{\mathbb{R}[G]}(\Delta_T^{\#})).$$

Proof. We use the natural exact sequence of finite G-modules

$$0\to\mathrm{cok}(\phi_T)\xrightarrow{\subseteq}\mathrm{cok}(\phi)\to\bigoplus_{v\in T}\mathbb{F}_{(v)}^{\times}\to A_{K,T}\to A_K\to 0 \tag{20}$$

where $\mathbb{F}_{(v)}$ denotes the direct sum of the residue fields $\mathbb{F}_w$ of each place w of K which lies above v [31, (1.5)]. Let G_w denote the decomposition group of w in G. Then, if η is any generator of the cyclic group $\mathbb{F}_w^{\times}$, there exists a G_w-equivariant surjection $\mathbb{Z}[G_w]\to\mathbb{F}_w^{\times}$ which sends 1 to η. In this way one obtains an exact sequence $0\to\mathbb{Z}[G]\overset{1-\sigma_v^{-1}\mathrm{N}v}{\to}\mathbb{Z}[G]\to\mathbb{F}_{(v)}^{\times}\to 0$ of G-modules. These sequences combine to imply that the G-module $\bigoplus_{v\in T}\mathbb{F}_{(v)}^{\times}$ is c-t and moreover that $\chi(\bigoplus_{v\in T}\mathbb{F}_{(v)}^{\times})=\sum_{v\in T}\chi(\mathbb{F}_{(v)}^{\times})=-\sum_{v\in T}\delta(\mathrm{detred}_{\mathbb{R}[G]}(1-\sigma_v^{-1}\mathrm{N}v))=-\delta(\mathrm{detred}_{\mathbb{R}[G]}(\Delta_T^{\#}))$.
At this stage we know that all of the modules which occur in (20) are c-t, except possibly for $A_{K,T}$. The exactness of this sequence therefore implies that $A_{K,T}$ is also c-t. Finally, the claimed equality follows upon decomposing (20) into short exact sequences and then using Lemma A2 (repeatedly). $\square$

Lemma 10. *The G-module $\mathrm{cok}(\phi_T)$ is c-t. Indeed, one has $\ell\nmid|\mathrm{cok}(\phi_T)|$.*

Proof. It suffices to prove that $\ell\nmid|\mathrm{cok}(\phi_T)^G|$. Now $\mathcal{E}\cong\Sigma_K$ so $H^1(G,\mathcal{E})\cong H^1(G,\Sigma_K)=0$. This implies $\mathrm{cok}(\phi_T)^G=(\mathcal{O}_{K,T}^{\times}/\mathcal{E})^G\cong\mathcal{O}_{k,T}^{\times}/\mathcal{E}^G$ and also that $\mathcal{E}^G$ is generated by $\{\mathrm{N}_j(\epsilon_j):j\in|n|\}$ where, for each $j\in|n|$, we write N_j for the field theoretic norm map $K_j^{\times}\to k^{\times}$.
We fix an ordered $\mathbb{Z}$-basis $\{u_i:i\in|n|\}$ of $U_{k,T}$ and define an element $b:=(b_{ij})$ of $\mathrm{M}_n(\mathbb{Z})$ by the equalities $\mathrm{N}_i(\epsilon_i)=\prod_{j=1}^n u_j^{b_{ij}}$ for each i,j in $|n|$. Then $|\mathcal{O}_{k,T}^{\times}/\mathcal{E}^G|=\pm\det(b)$ and so we must show that $\ell\nmid\det(b)$. To prove this we choose for each i,j in $|n|$, an integer a_{ij} such that $f_{K/k,w_i}(u_j)=g_i^{a_{ji}}$, we set $a:=(a_{ij})\in\mathrm{M}_n(\mathbb{Z})$ and we show that $b\cdot a\equiv I_n\ (\mathrm{mod}\ \ell\cdot\mathrm{M}_n(\mathbb{Z}))$.
For each intermediate field F of K/k we write J_F for the idele group of F and $f_F:J_F\to\mathrm{Gal}(K/F)$ for the global reciprocity map. For each $j\in|n|$ we write $f_{F,j}:F^{\times}\to\mathrm{Gal}(K/F)$ for the composite of f_F and the natural inclusion of $F^{\times}$ into $\prod_s F_s^{\times}\subset J_F$ where the product is taken over the set of places s of F which lie above v_j. We note that if $F=k$, then $f_{F,j}=f_{K/k,w_j}$.

For each pair of elements i, j of $|n|$ we set $S(ij) := \{w \in S(i) : w \mid v_j\}$. Then property i) in the statement of the Proposition implies $f_{K_i,j}(\epsilon_i) = \prod_{w\in S(ij)} f_{K/K_i,w}(\epsilon_i) = \prod_{w\in S(ij)} g_i^{\delta_{iw}} = g_i^{\delta_{ij}}$. After taking account of the functorial behaviour of Artin maps, this implies $g_i^{\delta_{ij}} = f_{K/k,w_j}(\mathrm{N}_i(\epsilon_i)) = \prod_{s\in|n|} f_{K/k,w_j}(u_s)^{b_{is}} = \prod_{s\in|n|} g_j^{b_{is}a_{sj}} = g_j^{\sum_{s\in|n|} b_{is}a_{sj}}$ and hence, since by assumption no element g_j is trivial, that $\sum_{s\in|n|} b_{is}a_{sj} \equiv \delta_{ij} \pmod{\ell}$. It follows that $b \cdot a \equiv I_n \pmod{\ell \cdot M_n(\mathbb{Z})}$, as required. □

6.4. The proof of Theorem 6.1. In this subsection we use Proposition 6.1 to prove Theorem 6.1. We assume throughout that G is abelian. Our argument is similar to that used in §5.4 and so we continue to use the notation $G^*_{(0)}$ and e_0 introduced in that subsection.

At the outset we observe that if G_j is trivial for any $j \in |n|$, then $\theta_{K/k,S,T}(0)$ and $\mathrm{Reg}_{G,S,T}$ are both equal to 0 and so the congruence of Theorem 6.1 is valid trivially. In the sequel we shall therefore assume that G_j is not trivial for any $j \in |n|$, as is required by Proposition 6.1.

Now, since G is abelian, Proposition 6.1iii) shows that $\mathrm{C}(K/k)$ implies the existence of an element x_T of $\mathbb{Q}[G]^\times$ which satisfies both

$$\theta^*_{K/k,S,T}(0)^\# = x_T \cdot \det(M_T) \in \mathbb{R}[G]^\times \tag{21}$$

and

$$\mathbb{Z}[G] \cdot x_T = \mathrm{Fitt}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,T}/\mathcal{E})^{-1}\, \mathrm{Fitt}_{\mathbb{Z}[G]}(A_{K,T}) \subseteq \mathrm{Fitt}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,T}/\mathcal{E})^{-1}. \tag{22}$$

For all i, j in $|n|$ one has $e_0 \cdot \mathrm{Tr}_j = 0$ so that $e_0 y_{ij} = 0$ and hence $(M_T e_0)_{ij} = \delta_{ij}(g_i - 1)e_0$. Also, for each $\chi \in G^* \setminus G^*_{(0)}$ there exists $j \in |n|$ such that $\chi(g_j) = 1$ and so $\prod_{i\in|n|}(g_i - 1)(1 - e_0) = 0$. It follows that

$$\begin{aligned} e_0\det(M_T) &= \det(M_T e_0) \\ &= \prod_{i\in|n|} (g_i - 1)e_0 \\ &= \prod_{i\in|n|} (g_i - 1). \end{aligned}$$

This combines with (21) to imply that $\theta_{K/k,S,T}(0)^\# = x_T \cdot e_0\det(M_T) = x_T \prod_{i\in|n|}(g_i - 1)$. In addition, (22) combines with Lemma 10 to imply that $x_T \in \mathbb{Z}_\ell[G]$ and hence one has

$$\theta_{K/k,S,T}(0)^\# \equiv \epsilon(x_T) \prod_{i\in|n|} (g_i - 1) \pmod{I_G \cdot \prod_{i\in|n|} I_i}.$$

Since $\mathrm{Reg}_{G,S,T} \in \prod_{i\in|n|} I_i$ one also has

$$(\mathrm{Reg}_{G,S,T})^\# \equiv (-1)^n \mathrm{Reg}_{G,S,T} \pmod{I_G \cdot \prod_{i\in|n|} I_i}.$$

To deduce the congruence of Theorem 6.1 from the previous displayed congruence we therefore need only show that

$$\epsilon(x_T)\prod_{i\in|n|}(g_i-1)\equiv(-1)^n m_{k,S,T}\cdot\mathrm{Reg}_{G,S,T}\pmod{I_G\cdot\prod_{i\in|n|}I_i}$$

where $m_{k,S,T}$ is as defined in (11). In addition, with the matrix b as defined in the proof of Lemma 10, one has

$$\begin{aligned}\prod_{i\in|n|}(g_i-1) &\equiv \det((f_{K/k,w_j}(\mathrm{N}_i(\epsilon_i))-1)_{1\le i,j\le n})\\ &\equiv \det(b)\cdot\det((f_{K/k,w_j}(u_i)-1)_{1\le i,j\le n})\\ &\equiv \det(b)\cdot\mathrm{Reg}_{G,S,T}\pmod{I_G\cdot\prod_{j\in|n|}I_j},\end{aligned}$$

and so it suffices to show that $\epsilon(x_T)\cdot\det(b)=(-1)^n m_{k,S,T}$. But, just as in the deduction of (17) from (13), this can be proved by first multiplying (21) by Tr_G and then comparing the resulting equality to (11).
This completes our proof of Theorem 6.1.

Appendix

We recall some relevant properties of the refined Euler characteristic construction discussed in §2.1 (the notation of which we continue to use). For further details we refer the reader to [9] (or to [7, §1] for a fuller review than that given here).

We let R denote either $\mathbb{Z}$ or $\mathbb{Z}_\ell$ for some prime ℓ and E an extension of the field of fractions of R. For any $R[G]$-module M, resp. homomorphism of R-modules ϕ, we set $M_E := M\otimes_R E$, resp. $\phi_E := \phi\otimes_R \mathrm{id}_E$.

Let $P^\cdot$ be a bounded complex of finitely generated projective $R[G]$-modules. For each integer i we let B^i, resp. Z^i, denote the submodules of coboundaries, resp. cocycles, of $P^\cdot_E$ in degree i. After choosing $E[G]$-equivariant splittings of the tautological exact sequences $0\to Z^i\to P^i_E\to B^{i+1}\to 0$ and $0\to B^i\to Z^i\to H^i(P^\cdot_E)\to 0$ one obtains non-canonical isomorphisms

$$\begin{aligned}P^+_E &\cong B^{\mathrm{all}}\oplus H^+(P^\cdot)_E\\ P^-_E &\cong B^{\mathrm{all}}\oplus H^-(P^\cdot)_E.\end{aligned}$$

By using the identity map on B^{all} one can therefore extend each element ϕ of $\mathrm{Is}_{E[G]}(H^+(P^\cdot)_E,H^-(P^\cdot)_E)$ to give an element $\phi(P^\cdot_E)$ of $\mathrm{Is}_{E[G]}(P^+_E,P^-_E)$. This construction clearly depends upon the above choice of splittings but nevertheless induces a well-defined map from $\mathrm{Is}_{E[G]}(H^+(P^\cdot)_E,H^-(P^\cdot)_E)/\sim$ to $\mathrm{Is}_{E[G]}(P^+_E,P^-_E)/\sim$ which is independent of all such choices. We denote this map by $\tau\mapsto\tau(P^\cdot_E)$ and we obtain a well-defined element of $K_0(R[G],E)$ by setting $\chi_{R[G],E}(P^\cdot,\tau):=(P^+,\phi,P^-)$ for any (and therefore every) $\phi\in\tau(P^\cdot_E)$.
In the following result we record this construction in a special case.

Lemma A1. *Let P be a finitely generated projective $R[G]$-module, ϕ an $R[G]$-endomorphism of P and $\lambda:\ker(\phi)_E\to\mathrm{cok}(\phi)_E$ an $E[G]$-isomorphism. Choose*

$E[G]$-equivariant sections ι_1 and ι_2 to the tautological surjections $P_E \to \mathrm{im}(\phi)_E$ and $P_E \to \mathrm{cok}(\phi)_E$, and let $\langle\lambda,\phi\rangle_{\iota_1,\iota_2}$ denote the automorphism of P_E which is equal to $\iota_2 \circ \lambda$ on $\ker(\phi)_E$ and to ϕ_E on $\iota_1(\mathrm{im}(\phi)_E)$. If $P^\cdot$ denotes the complex $P \xrightarrow{\phi} P$, where the first term is placed in degree 0, then in $K_0(R[G],E)$ one has $\chi_{R[G],E}(P^\cdot,\lambda) = \partial^1_{R[G],E}([\langle\lambda,\phi\rangle_{\iota_1,\iota_2}])$.

For each $i \in \{1,2,3\}$ let $P_i^\cdot$ be a bounded complex of finitely generated projective $R[G]$-modules. We assume that there exists a distinguished triangle in $\mathcal{D}^{\mathrm{perf}}(R[G])$ of the form

$$P_1^\cdot \xrightarrow{\alpha} P_2^\cdot \to P_3^\cdot \to P_1^\cdot[1]$$

and that $P_{3,E}^\cdot$ is acyclic (so that $H^i(\alpha)_E$ is bijective in each degree i). For any E-trivialisation τ of $P_1^\cdot$ we let τ_α denote the unique E-trivialisation of $P_2^\cdot$ that contains $H^-(\alpha)_E \circ \phi \circ H^+(\alpha)_E^{-1}$ for any (and therefore every) $\phi \in \tau$. The following result is a special case of [9, Th. 2.8].

Lemma A2. *If $P_{3,E}^\cdot$ is acyclic, then for any E-trivialisation τ of $P_1^\cdot$ one has*

$$\chi_{R[G],E}(P_2^\cdot,\tau_\alpha) = \chi_{R[G],E}(P_1^\cdot,\tau) + \chi_{R[G],E}(P_3^\cdot,\mathrm{id}_0),$$

where id_0 *denotes the identity map on the zero space.*

Note that if $P_3^\cdot$ is acyclic, then $\chi_{R[G],E}(P_3^\cdot,\mathrm{id}_0) = 0$ and so Lemma A2 implies $\chi_{R[G],E}(\cdot,\cdot)$ is well-defined on pairs of the form (X,τ) where X is an object of $\mathcal{D}^{\mathrm{perf}}(R[G])$ and τ an element of $\mathrm{Is}_{E[G]}(H^+(X)_E, H^-(X)_E)/\sim$.

Remark A1. The element $\chi_{R[G],E}(X,\tau)$ of $K_0(R[G],E)$ constructed above can be naturally reinterpreted as an isomorphism class of objects in a fibre product category involving a suitable category of virtual objects as introduced by Deligne in [28]. (Indeed, this more conceptual approach has important technical advantages and is used systematically in [13]). As a result, if G is abelian, then $\chi_{R[G],E}(X,\tau)$ can also be described by using the graded determinant functor of Grothendieck, Knudsen and Mumford (that is described in [36]). In fact, if G is abelian, then $K_0(R[G],E)$ identifies naturally with the multiplicative group of invertible $R[G]$-lattices in $E[G]$ (cf. [4, Lem. 2.6]), the reduced norm map $\mathrm{nr}_{E[G]} : K_1(E[G]) \to E[G]^\times$ is bijective and, with respect to the stated identification, for each $x \in E[G]^\times$ one has

$$\partial^1_{R[G],E}(\mathrm{nr}^{-1}_{E[G]}(x)) = R[G]\cdot x \subset E[G].$$

This shows in particular that, if K/k is abelian, then the equality of C(K/k) is equivalent to a formula for the sublattice $\mathbb{Z}[G]\cdot\theta^*_{K/k,S}(0)^\#$ of $\mathbb{R}[G]$.

References

[1] N. Aoki, Gross' conjecture on the special values of abelian L-functions at $s = 0$, Comm. Math. Univ. Sacti Pauli 40 (1991) 101-124.

[2] M. F. Atiyah, C. T. C. Wall, Cohomology of Groups, In: 'Algebraic Number Theory', J. W. S. Cassels, A. Fröhlich (eds.), Academic Press, London, 1967.

[3] S. Bae, On the conjectures of Lichtenbaum and Chinburg over function fields, Math. Ann. 285 (1989) 417-445.

[4] W. Bley, D. Burns, Equivariant Tamagawa Numbers, Fitting Ideals and Iwasawa Theory, Compositio Math. 126 (2001) 213-247.

[5] W. Bley, D. Burns, Equivariant epsilon constants, discriminants and étale cohomology, Proc. London Math. Soc. 87 (2003) 545-590.

[6] S. Bloch, K. Kato, L-functions and Tamagawa numbers of motives, In: 'The Grothendieck Festschrift' vol. 1, Progress in Math. 86, 333-400, Birkhäuser, Boston, 1990.

[7] D. Burns, Equivariant Tamagawa Numbers and Galois module theory, Compositio Math. 129 (2001) 203-237.

[8] D. Burns, Equivariant Tamagawa Numbers and refined abelian Stark conjectures, J. Math. Soc. Univ. Tokyo. 10 (2003) 225-259.

[9] D. Burns, Equivariant Whitehead torsion and refined Euler characteristics, CRM Proceedings and Lecture Notes 36 (2004) 35-59.

[10] D. Burns, Congruences between derivatives of abelian L-functions at $s = 0$, manuscript submitted for publication.

[11] D. Burns, M. Flach, Motivic L-functions and Galois module structures, Math. Ann. 305 (1996) 65-102.

[12] D. Burns, M. Flach, On Galois structure invariants associated to Tate motives, Amer. J. Math. 120 (1998) 1343-1397.

[13] D. Burns, M. Flach, Tamagawa numbers for motives with (non-commutative) coefficients, Documenta Math. 6 (2001) 501-570.

[14] D. Burns, M. Flach, Tamagawa numbers for motives with (noncommutative) coefficients, II, Amer. J. Math. 125 (2003) 475-512.

[15] D. Burns, C. Greither, On the Equivariant Tamagawa Number Conjecture for Tate motives, Invent. math. 153 (2003) 303-359.

[16] D. Burns, C. Greither, Equivariant Weierstrass Preparation and values of L-functions at negative integers, Documenta Math., Extra volume: Kazuya Kato's Fiftieth Birthday, (2003) 157-185.

[17] D. Burns, J. Lee, On the refined class number formula of Gross, J. Number Theory 107 (2004) 282-286.

[18] Ph. Cassou-Noguès, T. Chinburg, A. Fröhlich, M. J. Taylor, L-functions and Galois modules, (notes by D. Burns and N. P. Byott, In: 'L-functions and Arithmetic', J. Coates, M. J. Taylor (eds.), London Math. Soc. Lecture Note Series 153, 75-139, Cam. Univ. Press 1991.

[19] T. Chinburg, On the Galois structure of algebraic integers and S-units, Invent. math. 74 (1983) 321-349.

[20] T. Chinburg, The Galois structure of S-units, Sém. Th. Nombres Bordeaux 1982-1983, exposé 40.

[21] T. Chinburg, Exact sequences and Galois module structure, Annals of Math. 121 (1985) 351-376.

[22] T. Chinburg, The analytic theory of multiplicative Galois structure, Memoirs of the Amer. Math. Soc., 77 (1989).

[23] T. Chinburg, Galois module structure of de Rham cohomology, J. Th. Nombres Bordeaux 4 (1991) 1-18.

[24] T. Chinburg, Galois structure of de Rham cohomology of tame covers of schemes, Annals of Math. 139 (1994) 443-490.

[25] T. Chinburg, M. Kolster, G. Pappas, and V. Snaith, Galois Structure of K-Groups of Rings of Integers, K-theory 14 (1998) 319-369.

[26] C. W. Curtis, I. Reiner, Methods of Representation Theory, Vol. I and II, John Wiley and Sons, New York, 1987.

[27] P. Deligne, Valeurs de fonctions L et périods d'intégrales, Proc. Sym. Pure Math. 33 (2), (1979) 313-346.

[28] P. Deligne, Le déterminant de la cohomologie, Contemp. Math. 67 (1987) 313-346.

[29] M. Flach, Euler characteristics in relative K-groups, Bull. London Math. Soc. 32 (2000) 272-284.

[30] T. Geisser, Weil-Étale cohomology over finite fields, to appear in Math. Ann.

[31] B. H. Gross, On the value of abelian L-functions at $s = 0$, J. Fac. Sci. Univ. Tokyo, Sect. IA, Math., 35 (1988) 177-197.

[32] A. Grothendieck, Formule de Lefschetz et rationalité des fonctions L, in Sém. Bourbaki vol. 1965-1966, Benjamin (1966), exposé 306.

[33] K. W. Gruenberg, J. Ritter and A. Weiss, A Local Approach to Chinburg's Root Number Conjecture, Proc. London Math. Soc. 79 (1999) 47-80.

[34] A. Hayward, A class number formula for higher derivatives of abelian L-functions, Compos. Math. 140 (2004) 99-129.

[35] K. Kato, Lectures on the approach to Iwasawa theory of Hasse-Weil L-functions via B_{dR}, Part I, In: Arithmetical Algebraic Geometry (ed. E. Ballico), Lecture Notes in Math. 1553, 50-163, Springer, New York, 1993.

[36] F. Knudsen, D. Mumford, The projectivity of the moduli space of stable curves I: Preliminaries on 'det' and 'Div', Math. Scand. 39 (1976) 19-55.

[37] J. Lee, On Gross's Refined Class Number Formula for Elementary Abelian Extensions, J. Math. Sci. Univ. Tokyo 4 (1997) 373-383.

[38] J. Lee, Stickelberger elements for cyclic extensions and the order of vanishing of abelian L-functions at $s = 0$, Compositio Math. 138 (2003) 157-163.

[39] J. Lee, On the refined class number formula for global function fields, to appear in Math. Res. Letters.

[40] S. Lichtenbaum, The Weil Étale Topology, preprint 2001.

[41] J. S. Milne, Étale Cohomology, Princeton Math. Series 17, Princeton University Press, 1980.

[42] J. S. Milne, Arithmetic Duality Theorems, Perspectives in Math. 1, Academic Press, 1986.

[43] J. S. Milne, Values of Zeta functions of varieties over finite fields, Amer. J. Math. 108 (1986) 297-360.

[44] J. Nekovář, Selmer Complexes, to appear in Astérisque.

[45] M. Rapoport, Th. Zink, Über die lokale Zetafunktion von Shimuravarietäten. Monodromiefiltration und verschwindende Zyklen in ungleicher Charakteristik, Invent. math. 68 (1982) 21-101.

[46] R. G. Swan, Algebraic K-theory, Lecture Note in Math. 76, Springer, 1968.

[47] K.-S. Tan, On the special values of abelian L-functions, J. Math. Sci. Univ. Tokyo 1 (1994) 305-319.

[48] K.-S. Tan, A note on the Stickelberger elements for cyclic p-extensions over global function fields of characteristic p, Math. Res. Letters 11 (2004) 273-278.

[49] J. Tate, The cohomology groups of tori in finite Galois extensions of number fields, Nagoya Math. J. 27 (1966) 709-719.

[50] J. Tate, Les Conjectures de Stark sur les Fonctions L d'Artin en $s = 0$ (notes par D. Bernardi et N. Schappacher), Progress in Math., 47, Birkhäuser, Boston, 1984.

[51] J. Tate, Letter to J. Lee, 22 July 1997.

David Burns
Dept. of Mathematics
King's College London
London WC2R 2LS
United Kingdom
david.burns@kcl.ac.uk

Tight Embeddings of Simply Connected 4-Manifolds

Wolfgang Kühnel

Received: January 8, 2004
Revised: November 10, 2004

Communicated by Günter M. Ziegler

Dedicated to Thomas F. Banchoff on the occasion of his 65^{th} birthday[1]

Abstract. The classification of compact and simply connected PL 4-manifolds states that the homeomorphism classes coincide with the homotopy classes, and that these are classified by the intersection form. We show here that "most" of these classes with an indefinite intersection form can be represented by a tight polyhedral embedding into some Euclidean space. It remains open which of the PL structures can be realized in such a way.

2000 Mathematics Subject Classification: 52B70, 53C42, 57Q35
Keywords and Phrases: intersection form, tight polyhedra, polyhedral handles, tight surgery

Introduction and Result

An embedding $M \to \mathbb{E}^N$ of a compact manifold into Euclidean space is called *tight*, if for any open half space $E_+ \subset \mathbb{E}^N$ the induced homomorphism

$$H_*(M \cap E_+) \longrightarrow H_*(M)$$

is injective where H_* denotes an appropriate homology theory with coefficients in a certain field. In the smooth case (and, with certain modifications, also in the polyhedral case) this is equivalent to the condition that almost all height functions on M are perfect functions, i.e., have the minimum number of critical points which coincides with the sum of the Betti numbers. For a survey on tightness see [14] or [3].

[1] The main result of this paper was first presented at the TFB conference on this occasion at Providence, RI, Nov. 1, 2003.

For compact 2-manifolds without boundary this is equivalent to the Two-piece property (TPP) which states that the intersection of M with any (open or closed) halfspace is connected. Smooth tight surfaces were investigated by N.H.Kuiper [13] and others, the study of tight polyhedral surfaces was initiated by T.F.Banchoff [1]. One of the results is that any given closed surface admits a tight polyhedral embedding into some Euclidean space. For obtaining this, it is sufficient to start with the three cases of the sphere, the real projective plane and the Klein bottle [2] and then to attach handles tightly. For tight polyhedral immersions into 3-space the situation is the following: Any given closed surface (except for the real projective plane and the Klein bottle) admits a tight polyhedral immersion into 3-space. The crucial and most difficult case $\chi = -1$ had been open for many years and was solved only recently by D.Cervone [5]. Smooth tight immersions into 3-space exist for all surfaces except for the real projective plane, the Klein bottle and the surface with $\chi = -1$. The latter is again the most crucial case and was solved by F.Haab [7]. There is a smooth tight embedding $\mathbb{R}P^2 \to \mathbb{E}^4$ as a suitable linear projection of the Veronese surface. The cases of the Klein bottle and $\chi = -1$ seem still to be open. One approach might be to attach a handle tightly to the Veronese surface in 4-space (or a slightly distorted version of it) but that has so far turned out to be unmanageable. In 5-space the only smooth tight surface is the classical Veronese surface itself by a theorem of N.H.Kuiper.

In the case of compact 3-manifolds not too much seems to be known at all. Smooth tight examples include the Veronese embedding $\mathbb{R}P^3 \to \mathbb{E}^9$, connected sums of handles $S^1 \times S^2$ and cartesian products of a circle with tight surfaces as well as tubes around embedded tight surfaces in 4-space. The more restrictive class of taut 3-manifolds was classified in [17]. In particular it includes an embedding of the twisted product $S^1 \times_h S^2$ as a "complexified 2-sphere" and the quaternion space as Cartan's isoparametric hypersurface in S^4. There are a number of constructions for tight polyhedral embeddings of 3-manifolds, compare [9]. However, we are far from being able to cover major parts of the class of all 3-manifolds. It seems that we do not know a tight embedding of any Lens space (except for $\mathbb{R}P^3$) and it seems also that we do not even know a tight polyhedral embedding of $\mathbb{R}P^3$. For any given tight polyhedral 3-manifold it is easy to attach handles tightly but that procedure does not help too much if the other building blocks are missing. Unfortunately there is no simple combinatorial condition which implies the tightness. Instead one has to check all the homology classes in all the open halfspaces, just by applying the definition above. This is better in the case of simply connected 4-manifolds.

For compact and simply connected 4-manifolds without boundary the tightness is equivalent to the requirement that $M \cap E_+$ is always connected and simply connected. The only smooth tight immersions of simply connected 4-manifolds which are known are spheres as convex hypersurfaces in 5-space, the Veronese-type embedding of $\mathbb{C}P^2$ into 8-space [13] and certain embeddings of arbitrary connected sums of copies of $S^2 \times S^2$ in 5-space [8]. G.Thorbergsson [18] found topological obstructions to the existence of smooth tight immersions in terms of the intersection form and Stiefel-Whitney classes. This leads to restrictions for the existence of smooth tight immersions of connected sums of copies of $\mathbb{C}P^2$ and $-\mathbb{C}P^2$. In particular, it turned out that the $K3$ surface does not admit any smooth tight immersion. The obstruction is that it does not admit a splitting as a connected sum of two smooth manifolds even though the intersection form splits as a connected sum.

This is much different in the polyhedral case because the same type of topological obstruction is not there. The polyhedral tight embedding $\mathbb{C}P^2 \to \mathbb{E}^8$ [10] leads to tight embeddings $\mathbb{C}P^2\#k(-\mathbb{C}P^2) \to \mathbb{E}^8$ for any number k, see [9, Sect.6C], and a tight embedding of the $K3$ surface into 15-space was recently found in [4]. We use them as building blocks and show in our Theorem 7 below how these – together with attaching 2-handles of type $S^2 \times S^2$ – lead to polyhedral tight embeddings of any given topological type of a simply connected PL 4-manifold, subject to a certain extra assumption on the intersection form. Our proof relies on the following results from the classification of 4-manifolds. For an outline of them see [16, Sect.5].

Definition The *intersection form* Q of a compact 4-manifold M is the symmetric bilinear form $Q\colon H_2(M;\mathbb{Z}) \times H_2(M;\mathbb{Z}) \to \mathbb{Z}$ which is dual to the cup product defined on the cohomology $H^2(M;\mathbb{Z})$. It satisfies the equation $Q(M_1\#M_2) \cong Q(M_1) \oplus Q(M_2)$. If we represent the intersection form in a basis over the integers then the corresponding matrix is invertible and hence *unimodular*, i.e., it has determinant ± 1. The *rank* of Q is the rank of $H_2(M;\mathbb{Z})$ as a $\mathbb{Z}$-module, also known as the second Betti number, the *signature* is the number of negative eigenvalues minus the number of positive eigenvalues. A quadratic form is called *odd* if some diagonal entry in the representing integer matrix is odd, otherwise it is called *even*. It is known from algebra [15] that an indefinite quadratic form over the integers is uniquely classified by its rank, its signature and by its type (even or odd).

Theorem 1 (S.S.Cairns 1940)
The equivalence classes of smooth 4-manifolds and PL 4-manifolds are in $(1-1)$-correspondence. More precisely, every smooth 4-manifold induces precisely one PL manifold (up to PL-homeomorphism) and, vice versa, every PL 4-manifold admits exactly one smoothing (up to diffeomorphism).

Theorem 2 (V.A.Rohlin 1952)
The signature of any simply connected smooth or PL 4-manifold with an even intersection form is an integer multiple of 16.

Theorem 3 (S.Donaldson 1983)
If the intersection form of a simply connected PL 4-manifold is definite then it is diagonalizable over the integers and, in particular, odd.

Theorem 4 (J.Milnor 1958)
The homotopy classes of simply connected 4*-manifolds are uniquely classified by their intersection forms.*

The topological classification turned out to be much harder, and it took almost 25 more years until this problem was solved by M.Freedman. The smooth (or PL) classification appears still to be open.

Theorem 5 (M.Freedman 1982)
The homeomorphism classes of simply connected PL 4*-manifolds are uniquely classified by their intersection forms. More precisely: Two such* PL *manifolds* $M, \widetilde{M}$ *are homeomorphic (not necessarily* PL homeomorphic*) if and only if their intersection forms* $Q, \widetilde{Q}$ *are equivalent over the integers.*

There is an algebraic classification of indefinite unimodular quadratic forms as follows:

Theorem 6 (see [16, Sect.5])

1. *Any indefinite, odd and unimodular quadratic form over the integers is equivalent to* $l(+1) \oplus k(-1)$.
2. *Any indefinite, even and unimodular quadratic form over the integers is equivalent to* $n(\mp E_8) \oplus m\binom{0\ 1}{1\ 0}$.

The rank is $k + l$ or $8n + 2m$, respectively, the signature is $k - l$ or $\pm 8n$, respectively. Conversely, rank and signature of the quadratic form determine these numbers k, l, m, n uniquely. Here E_8 denotes the following unimodular and positive definite matrix:

$$E_8 = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}$$

Corollary
Let $K3$ *denote the* $K3$ *surface with its intersection form* $(-E_8) \oplus (-E_8) \oplus 3\binom{0\ 1}{1\ 0}$. *Then the manifolds*

$$l(\mathbb{C}P^2)\#k(-\mathbb{C}P^2) \text{ with } k, l \geq 0$$

and

$$n(\pm K3)\#m(S^2 \times S^2) \text{ with } m, n \geq 0$$

cover all homotopy classes (and, in fact, homeomorphism classes) of simply connected PL 4*-manifolds with intersection forms*

$$l(+1) \oplus k(-1) \quad \text{or} \quad 2\nu(\mp E_8) \oplus \mu \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where $k, l \geq 0$ and $\mu \geq 3\nu \geq 0$, respectively.

Remark For the intersection form of the $K3$ surface compare [15]. Theorem 3 together with the $\frac{11}{8}$-conjecture [6] implies that no other quadratic form can occur as the intersection form of any simply connected PL 4-manifold. In more detail this conjecture states that for an even intersection form Q the rank of Q is always at least $\frac{11}{8}$ times the absolute value of the signature of Q. It is easily seen that for the form $Q = 2\nu(\mp E_8) \oplus \mu\binom{0\ 1}{1\ 0}$ we have

$$\frac{\operatorname{rank}(Q)}{|\operatorname{sign}(Q)|} = \frac{16\nu + 2\mu}{16\nu} \geq \frac{11}{8} \quad \text{if and only if} \quad \mu \geq 3\nu.$$

Our main result is the following Theorem 7 which provides a construction of polyhedral tight embeddings for a large class of simply connected 4-manifolds. This follows the pattern in the case of 2-manifolds which was mentioned at the very beginning above: Start with certain building blocks and then attach handles tightly.

Theorem 7 *Let M be a simply connected* PL 4*-manifold with an indefinite intersection form Q. Assume further that* $\operatorname{rank}(Q) \geq \frac{11}{8}|\operatorname{sign}(Q)| + 44$ *in case that Q is even with* $|\operatorname{sign}(Q)| \geq 32$*. Then there exists a tight polyhedral embedding $\widetilde{M} \to \mathbb{E}^N$ for some N such that M and $\widetilde{M}$ are homeomorphic.*

Since this result relies on the classification in terms of the intersection form, we cannot obtain by this method that M and $\widetilde{M}$ are PL homeomorphic. However, by a theorem of C.T.C.Wall 1964 there is always a number $k \geq 0$ such that the manifolds $M\#k(S^2 \times S^2)$ and $\widetilde{M}\#k(S^2 \times S^2)$ in Theorem 7 are PL homeomorphic. So in some sense in "most" of the cases we can not only prescribe the topological type but also the PL type. Compare Remark 2 at the end of the paper. However, there are an infinite number of undecided cases left. In particular we do not have any example of a tight polyhedral realization of a manifold homeomorphic to $K3\#K3\#\cdots\#K3$. Such examples could remove the number 44 from the extra assumption in Theorem 7 which then would just transform into the hypothesis of the $\frac{11}{8}$-conjecture. For the case of a positive definite intersection form it would be sufficient – by Theorem 3 – to find a tight polyhedral embedding of $k(\mathbb{C}P^2)$ for arbitrary $k \geq 2$. However, such an example (for any $k \geq 2$) is still missing.

The building blocks and connected sums of them

First of all, there are tight triangulations of $\mathbb{C}P^2$ and of the $K3$ surface. This means, there is a triangulation of $\mathbb{C}P^2$ (with 9 vertices, see [10], [11]) and one of the $K3$ surface (with 16 vertices, see [4]) such that *any* simplexwise linear embedding into *any* Euclidean space is tight. In particular, we can regard $\mathbb{C}P^2_9$ as a tightly embedded subcomplex of the 8-simplex $\triangle^8$ and $(K3)_{16}$ as a tightly embedded subcomplex of the 15-simplex $\triangle^{15}$. In each case the manifold contains the complete 2-dimensional skeleton of the ambient 8-simplex or 15-simplex, respectively. This implies that the intersection with any open half-space is connected and simply connected. Compare [9] for general properties of tight triangulations and [12] for a list of known examples.

By truncating each of these subcomplexes at a vertex and by glueing in another copy of the same kind, one gets tight embeddings

$$\mathbb{C}P^2\#(-\mathbb{C}P^2) \to \mathbb{E}^8 \quad \text{and} \quad K3\#(-K3) \to \mathbb{E}^{15},$$

each with signature zero. This construction is quite similar to the original version [2] of Banchoff's tight Klein bottle in 5-space as a geometric connected sum $\mathbb{R}P^2\#\mathbb{R}P^2$. The process of truncating and glueing in additional copies of the same combinatorial type can be repeated arbitrarily often, as shown in [9, Sect.6C]. This implies that we can realize any quadratic form of type

$$(+1) \oplus k(-1), \quad k \geq 1$$

or

$$2(-E_8) \oplus 2n(E_8) \oplus 3(n+1)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cong 2(n-1)E_8 \oplus (3n+19)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad n \geq 1$$

by a tightly and polyhedrally embedded simply connected 4-manifold. In the latter case we have the equations rank $= 16(n-1) + 6n + 38 = 22(n+1)$ and $|\text{sign}| = 16(n-1)$, so in particular rank $\geq \frac{11}{8}|\text{sign}| + 44$.

In order to cover the other cases in Theorem 7, we have to attach handles, thus realizing the sum of a previously given quadratic form Q and copies of $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Attaching handles tightly

There is an obvious procedure to attach a handle to a tight polyhedral surface in 3-space: Pick two faces opposite to one another (not necessarily in parallel planes), cut out a certain triangle in each of them, and glue in a polyhedral cylinder (as the boundary of a triangular prism), see Figure 1. It is, however, much less obvious how one can attach a 2-handle or a 3-handle tightly to a given polyhedron. One needs to fill in something within the convex hull of its boundary without hitting the rest of the manifold.

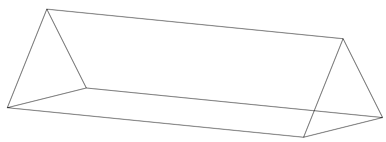


Figure 1: Attaching a 1-handle tightly

In general the procedure of attaching a k-handle of type $S^k \times S^{n-k}$ to an n-manifold is equivalent to cutting out a submanifold of type $S^{k-1} \times B^{n-k+1}$ inside a topological ball (e.g., a coordinate chart) and replacing it by $B^k \times S^{n-k}$. This is the classical *surgery* which we have to carry out in a polyhedral setting. The case $k = 1$ corresponds to attaching an ordinary 1-handle, like a bridge between two parts of the manifold. The case $k = 2, n = 4$ is crucial in the proof of our Theorem 7. For our purpose we have to realize this surgery within the class of tight polyhedral submanifolds. Therefore, we have to describe this process of *tight surgery* geometrically in the ambient space. It will always be carried out in some Euclidean $(n + 1)$-space if the manifold is n-dimensional

Definition A *simple polyhedral sphere* Σ^{k-1} is a triangulation of the sphere S^{k-1} with $k + 1$ vertices. This is nothing but the boundary complex of a k-dimensional simplex. A *short link* of a certain $(n-k)$-simplex in a triangulated n-manifold is a link which is combinatorially equivalent to a simple polyhedral sphere Σ^{k-1}. Notice that the link of a codimension–1 face is always short, the link of a codimension–2 face is short if and only if it has exactly 3 vertices and 3 edges. In the sequel let Δ^k denote a certain k-dimensional simplex in the simplicial complex which is considered, and let $\triangle^k$ denote an abstract k-simplex which is not necessarily in the complex.

Lemma *Assume that $M^n \subset \mathbb{E}^N$ is a simplicial submanifold containing a simplex Δ^{n-k} with a short link Σ^{k-1} such that the $n + 2$ vertices of the star of Δ^{n-k} are in general position. Then there is a polyhedral "solid torus" of type $S^{k-1} \times B^{n-k+1}$ within the open star of Δ^{n-k} which is a tight submanifold-with-boundary in the affine subspace $\mathbb{E}^{n+1}$ of $\mathbb{E}^N$ which is spanned by the $n+2$ vertices of the star of Δ^{n-k}. Moreover, it can be arranged that the convex hull of the short link does not hit M except for its boundary. Therefore, we can choose the tight solid torus in such a way that its convex hull does not hit M either except for the solid torus itself.*

Proof. The procedure of attaching a handle will be carried out inside the open star of Δ^{n-k} without using any of the original vertices. Since the tightness is affinely invariant, we can assume that the $n + 2$ vertices of the star of Δ^{n-k} form a regular simplex in $(n+1)$-space. In the classical case $k = 1$ we take the two barycenters of the two n-faces meeting at Δ^{n-1}. These form a 0-sphere. Then the procedure of attaching a handle tightly is suggested by Figure 1.

If $k = 2$ we take the three barycenters of the three $(n-1)$-faces meeting at Δ^{n-2}. Any two of them can be joined by a straight line segment inside of one of the three n-faces of M. The union of these three line segments is a simple polyhedral 1-sphere $\partial\triangle^2$ in M (but not as a subcomplex) such that its convex hull does not hit M except along exactly those three line segments. Then we construct a tight thickening of this polyhedral 1-sphere in the n-manifold as the union of three prisms of type $\triangle^1 \times \triangle^{n-1}$ such that they fit mutually together in $(n+1)$-space as an embedded solid handle $\partial \triangle^2 \times\triangle^{n-1}$.

In the general case for arbitrary k we proceed similarly: Take the $k+1$ barycenters of all the $(n-k+1)$-faces meeting at Δ^{n-k}. These span a regular k-simplex $\triangle^k$ in an $(n+1)$-dimensional Euclidean space. Its boundary is a simple polyhedral $(k-1)$-sphere inside the star of Δ^{n-k} in M (but not as a subcomplex). Then take a similar n-dimensional thickening of that simple sphere inside M and inside the same $(n+1)$-space. Then again replace the interior of a "solid torus" of type $\partial \triangle^k \times\triangle^{n-k+1}$ by the exterior of type $\triangle^k \times \partial\triangle^{n-k+1}$.

In order to describe this procedure in more detail we use the unique projective transformation $\Phi\colon \mathrm{star}(\Delta^{n-k})\backslash\Delta^{n-k} \to \mathbb{E}^{n+1}$ which sends to infinity the hyperplane which contains Δ^{n-k} and the k-plane parallel to the opposite k-simplex in the star of Δ^{n-k}. Then the rest of the open star becomes an orthogonal cartesian product of the link of Δ^{n-k} with an open part of some Euclidean $\mathbb{E}^{n-k+1}$. Furthermore Φ maps the union of the $k+1$ open n-faces meeting at Δ^{n-k} to an open part of the orthogonal cartesian product $\partial \triangle^k \times\mathbb{E}^{n-k+1}$ in $\mathbb{E}^k \times \mathbb{E}^{n-k+1} = \mathbb{E}^{n+1}$. Hence the polyhedral thickening of $\partial\triangle^k$ can be defined as the cartesian product $\partial\triangle^k \times\triangle^{n-k+1}$ where $\triangle^{n-k+1}$ denotes a small simplex in $(n-k+1)$-space. This is tightly embedded since it is a product of two tight subsets. The boundary is the product $\partial \triangle^k \times\partial\triangle^{n-k+1}$ which is also a tight polyhedral embedding of $S^{k-1} \times S^{n-k}$.

By applying Φ^{-1} we obtain the tight solid torus in the actual open star of Δ^{n-k}. Note that projective transformations preserve tightness. For the surgery we cut out the interior of $\partial \triangle^k \times\triangle^{n-k+1}$ and replace it by the interior of $\triangle^k \times \partial\triangle^{n-k+1}$. A picture for $n = 3$ is shown in Figure 2. Note, however, that this is a 3-dimensional projection of a 3-dimensional solid torus in 4-space. It is not a solid torus in 3-space. □

Corollary *Given a tight triangulation of a* PL *n-manifold M where some $(n-k)$-simplex ($k \leq n/2$) has a short link, we can tightly attach arbitrarily many handles of type $S^k \times S^{n-k}$. Hence for any m we obtain a manifold of* PL *type $M\#m(S^k \times S^{n-k})$ tightly embedded into Euclidean space.*

Proof: We carry out the construction of the lemma above. It is quite clear that we can repeat it arbitrarily often within one star since these solid tori can be chosen arbitrarily thin and disjoint. It is not essential to use the exact barycenters in the construction. The tightness of the solid torus implies that

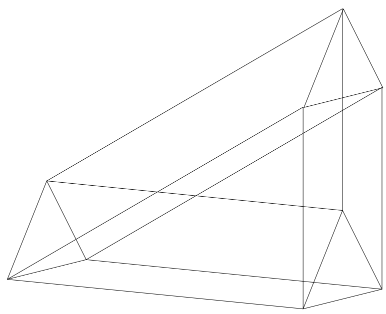


Figure 2: Attaching a 2-handle tightly; the case $k = 2, n = 3$

M minus the interior will still be tight. The same holds after the surgery. In the case of $n = 4$ and $k = 2$ which is most important for Theorem 7. We can easily see that the intersection with any halfspace is still connected and simply connected after the surgery if it was before. The tightness in the other cases follows similarly by considering the homology cycles created by the surgery. In any case the original S^{k-1} for starting the surgery is nullhomotopic in M, so that the topology after one step is that of $M\#(S^k \times S^{n-k})$. $\square$

Proof of Theorem 7:
In case 1 we consider a simply connected 4-manifold M with an odd intersection form which is equivalent to $l(+1) \oplus k(-1)$. By assumption it is indefinite, so we can assume that the signature is nonnegative and thus $k \geq l \geq 1$. This quadratic form is also equivalent to

$$(+1) \oplus (k - l + 1)(-1) \oplus (l - 1)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We can realize this by starting with the tight $\mathbb{C}P^2_9$ in 8-space, by truncating vertices and by glueing in $k - l + 1$ combinatorially equivalent copies of $-\mathbb{C}P^2_9$ with an open vertex star removed (see [9, Sect.6C]), and finally by attaching $l - 1$ handles of type $S^2 \times S^2$. Here it is crucial that $\mathbb{C}P^2_9$ does contain triangles with a short link, e.g., the triangle $\Delta^2 = \langle 1, 2, 3\rangle$ in the labeling of [10]. Hence our lemma above is applicable.

In case 2 we consider an even intersection form with signature 0, 16 or $16m \geq 32$. If the signature is zero we just take the standard ladder construction of tight connected sums of $S^2 \times S^2$, as described in [3, Ex.2.6.4]. The case of the 4-sphere itself is realized by the boundary of any convex polyhedron. If the signature is 16 we start with the tight $K3$ surface in 15-space and attach handles of type $S^2 \times S^2$ tightly. Here it is crucial that this triangulation contains a triangle with a short link. In the labeling of Figure 1 in [4] this is the triangle $\Delta^2 = \big\langle \binom{0}{0}, \binom{1}{0}, \binom{x}{0} \big\rangle$.

If the signature is $16m \geq 32$ we first build a tight $(-K3)\#(K3)\#m(K3)$ by the truncation process from [9, Sect.6C] and then attach handles of type $S^2 \times S^2$ tightly. Here the extra assumption

$$\mathrm{rank}(Q) \geq \frac{11}{8}\mathrm{sign}(Q) + 44$$

comes in since the signature of $(-K3)\#(K3)\#m(K3)$ is $16m$ whereas the rank is $22(m+2)$, so it is our assumption which implies that we have a nonnegative number of handles to attach. In any case the resulting tightly embedded 4-manifold has the same intersection form as M and is, therefore, homeomorphic to M by Theorem 5. □

Remarks:

1. The cases of 4-manifolds which are not covered by Theorem 7 are $k(\mathbb{C}P^2)$ where $k \geq 2$ and $m(K3)\#n\binom{0\ 1}{1\ 0}$ where $m \geq 2$ and $n < 22$. Examples of that kind would imply that – modulo the validity of the $\frac{11}{8}$-conjecture – *every* homotopy (or homeomorphism) class of simply connected PL 4-manifolds would be realizable by a tight embedding into some Euclidean space .

2. It seems that no example is known of any pair $M, \widetilde{M}$ of PL manifolds which are homeomorphic to one another but not PL homeomorphic and where each admits a tight PL embedding. One might expect that the "standard structure" is preferred for tight polyhedral embeddings if there is any. This is true at least for the sphere and for any homology sphere: The image of any tight polyhedral embedding of a homology k-sphere is the boundary of a convex polyhedron in $(k+1)$-space, for a simple proof see [9, Cor.3.6].

3. The same construction of attaching handles can be applied to other classes of manifolds. In the case of simply connected 5-manifolds we have tight connected sums of $S^2 \times S^3$ on the one hand and also a tight 13-vertex triangulation of $SU(3)/SO(3)$ on the other, see [12, p.170]. Since the tetrahedron $\Delta^3 = \langle 0, 1, 4, 6\rangle$ in the latter one has a short 1-dimensional link, it is possible to attach 2-handles of type $S^2 \times S^3$ tightly.

4. The construction above of attaching handles does not raise the essential codimension of the embedding. In fact, reaching or estimating the maximum codimension is a different interesting problem. Here a conjecture states that for any simply connected 4-manifold M a tight polyhedral embedding into $\mathbb{E}^N$ (not in any hyperplane) can exist only if the Heawood type inequality

$$\binom{N-3}{3} \leq 10\beta_2(M)$$

is satisfied where β_2 denotes the second Betti number (similarly for $(k-1)$-connected $2k$-manifolds), see [9, Sect.4]. Equality is attained for the tight triangulations of $\mathbb{C}P^2$ and the $K3$ surface, perhaps also in other cases. By standard arguments this conjecture would follow if the following generalized van

Kampen–Flores theorem is true: *Assume that a simply connected* 4*-manifold* M *admits a polyhedral embedding of the complete* 2*-skeleton of the* N*-dimensional simplex. Then the inequality* $\binom{N-3}{3} \leq 10\beta_2(M)$ *holds.* The classical van Kampen–Flores theorem is nothing but the case of the 4-sphere.

References

[1] T. F. Banchoff, Tightly embedded 2-dimensional polyhedral manifolds, *Amer. J. Math.* 87, 462 - 472 (1965)

[2] T. F. Banchoff, Tight polyhedral Klein bottles, projective planes, and Möbius bands, *Math. Ann.* 207, 233 - 243 (1974)

[3] T. F. Banchoff and W. Kühnel, Tight submanifolds, smooth and polyhedral, *Tight and Taut Submanifolds* (T. E. Cecil and S.-s. Chern, eds.), MSRI Publ. 32, pp. 51 - 118, Cambridge Univ. Press 1997

[4] M. Casella and W. Kühnel, A triangulated $K3$ surface with the minimum number of vertices, *Topology* 40, 753 - 772 (2001)

[5] D. P. Cervone, A tight polyhedral immersion in three-space of the real projective plane with one handle, *Pacific J. Math.* 196, 113 - 122 (2000)

[6] M. Furuta, Monopole equation and the $\frac{11}{8}$-conjecture, *Math. Res. Lett.* 8, 279 - 291 (2001)

[7] F. Haab, Immersioins tendues de surfaces dans $\mathbb{E}^n$, *Comment. Math. Helv.* 67, 182 - 202 (1992)

[8] J. J. Hebda Some new tight embeddings which cannot be made taut, *Geom. Dedicata* 17 49 - 60 (1984)

[9] W. Kühnel, *Tight Polyhedral Submanifolds and Tight Triangulations*, Lecture Notes in Mathematics 1612, Springer, Berlin - Heidelberg - New York 1995

[10] W. Kühnel and T. F. Banchoff, The 9-vertex complex projective plane, *The Math. Intelligencer* 5:3, 11 - 22 (1983)

[11] W. Kühnel and G. Lassmann, The unique 3-neighborly 4-manifold with few vertices, *J. Combin. Th. (A)* 35, 173 - 184 (1983)

[12] W. Kühnel and F. H. Lutz, A census of tight triangulations, *Period. Math. Hung.* 39, 161 - 183 (1999)

[13] N. H. Kuiper, Minimal total absolute curvature for immersions, *Invent. Math.* 10, 209 - 238 (1970)

[14] N. H. Kuiper, Geometry in total absolute curvature theory, *Perspectives in Mathematics*, Anniversary of Oberwolfach 1984 (W. Jäger et al., eds.), 377 - 392, Birkhäuser, Basel - Boston - Stuttgart 1984

[15] J. Milnor, On simply connected 4-manifolds, 1958 Symposium internacional de topología algebraica, 122 - 128, Universidad Nacional Autónoma de México and UNESCO, Mexico City

[16] N. Saveliev, *Lectures on the Topology of* 3*-Manifolds*, de Gruyter, Berlin - New York 1999

[17] U. Pinkall and G. Thorbergsson, Taut 3-manifolds, *Topology* 28, 389 - 401 (1989)

[18] G. Thorbergsson, Tight immersions of highly connected manifolds, *Comment. Math. Helv.* 61, 102 - 121 (1986)

Wolfgang Kühnel
Fachbereich Mathematik
Universität Stuttgart
70550 Stuttgart, Germany
kuehnel@mathematik.uni-stuttgart.de

Irrégularité et Conducteur de Swan p-adiques

Adriano Marmora

Received: July 6, 2004

Communicated by Takeshi Saito

Abstract. Let V be a de-Rham representation of the Galois group of a local field of mixed characteristic $(0,p)$. We relate the Swan conductor of the associated Weil-Deligne representation to the irregularity of the corresponding p-adic differential equation.

2000 Mathematics Subject Classification: 11F80,11F85,11S15,12H25
Keywords and Phrases: p-adic representation, de-Rham representation, Swan conductor, p-adic differential equation, p-adic irregularity.

1 Introduction

Soient K un corps de valuation discrète complet de caractéristique 0, de corps résiduel k parfait de caractéristique $p > 0$, et $\overline{K}$ une clôture algébrique de K. On note G_K le groupe de Galois de $\overline{K}/K$ et I_K le sous-groupe d'inertie. Fontaine a défini une hiérarchie sur les représentations p-adiques de G_K(i.e. les $\mathbb{Q}_p$-espaces vectoriels de dimension finie munis d'une action continue de G_K) : représentations de de Rham $\supset$ rep. semi-stables $\supset$ rep. cristallines. Le théorème de monodromie p-adique affirme que toute représentation de de Rham est potentiellement semi-stable, i.e. sa restriction à un sous-groupe ouvert de G_K est semi-stable. Le but de cet article est l'étude d'invariants numériques qui mesurent le défaut de semi-stabilité de représentations potentiellement semi-stables. Une telle représentation est entièrement décrite par son module de Weil-Deligne $\mathrm{D}_{\mathrm{pst}}(V)$. Celui-ci est muni d'une action de G_K dont la restriction à l'inertie se factorise par un quotient fini. Fontaine [10] définit les conducteurs de Swan et d'Artin de V, notés respectivement $\mathrm{sw}(V)$ et $\mathrm{ar}(V)$, comme étant les conducteurs de Swan et d'Artin de $\mathrm{D}_{\mathrm{pst}}(V)$. Dans un travail récent [3], Berger associe à toute représentation de de Rham V une équation différentielle p-adique $\mathrm{N}_{\mathrm{dR}}(V)$ munie d'une structure de Frobenius. À une équation différentielle p-adique M munie d'une structure de Frobenius, Christol et Mebkhout [7] associent un invariant entier $\mathrm{irr}(M)$, l'irrégularité de M.

Pour tout $n \in \mathbb{N}$, soient μ_n le groupe des racines p^n-ièmes de l'unité dans $\overline{K}$ et $K_n = K(\mu_n)$. Pour une représentation p-adique V de G_K, on note V_n sa restriction au sous-groupe $\mathrm{Gal}(\overline{K}/K_n)$. Le résultat principal de cet article est le suivant.

Théorème 1.1. *Pour toute représentation de de Rham V de G_K, on a*

$$\mathrm{irr}(\mathrm{N}_{\mathrm{dR}}(V)) = \lim_{n \to +\infty} \mathrm{sw}(V_n).$$

Le théorème 1.1 est l'analogue en caractéristique zéro d'un théorème de Tsuzuki en caractéristique $p > 0$. Soit E un corps de valuation discrète complet, de caractéristique p et de corps résiduel parfait. Dans [22], Tsuzuki montre que la catégorie des représentations p-adiques de $\mathrm{G}_E = \mathrm{Gal}(E^{\mathrm{sep}}/E)$ dont l'inertie agit par un quotient fini (monodromie finie), est équivalente à la catégorie des φ-∇-modules étales sur un corps valué $\mathcal{E}^{\dagger}(E)$ de caractéristique 0, d'anneau d'entiers hensélien et de corps résiduel E (voir §4.3 pour la définition). Puis dans [23], il démontre l'égalité entre le conducteur de Swan de la restriction à l'inertie d'une telle représentation et l'irrégularité du ∇-module correspondent. Dans la démonstration de 1.1, on se ramène, par la théorie du corps des normes, au cas d'un corps de valuation discrète complet d'égale caractéristique $p > 0$. Cependant, on ne peut pas appliquer directement le résultat de Tsuzuki, car dans notre cas, l'action de l'inertie ne se factorise pas par un quotient fini. La stratégie de la démonstration consiste à décrire la représentation de Weil-Deligne en termes de l'équation différentielle de Berger, puis de reprendre une partie de la preuve de Tsuzuki (l'induction de Brauer). Comme corollaires du théorème 1.1, on en déduit un résultat analogue pour le conducteur d'Artin (cf. 5.7) et l'égalité entre un polygone de Newton de pentes p-adiques et une limite de polygones de Newton de pentes de Swan (cf. 5.9).

Quand cet article a été déjà achevé, l'auteur a reçu une prépublication de P.Colmez [6] dont le résultat principal est une formule pour $\mathrm{sw}(V)$ en termes d'une filtration sur $\mathrm{D}_{\mathrm{dR}}(V)$. Les deux travaux sont indépendants et les méthodes utilisées semblent différentes.

Cet article est une partie de la thèse de doctorat en mathématique que je prépare à l'université de Paris 13, sous la direction d'Ahmed Abbes. Je tiens ici à le remercier pour son soutien constant tout le long de ce travail et ses lectures attentives des versions préliminaires de ce texte. Je remercie également le referee qui, par ses remarques, a amélioré ce manuscrit.

Notations

Soient k un corps parfait de caractéristique $p > 0$, $W = \mathrm{W}(k)$ (resp. $W_n = \mathrm{W}_n(k)$) l'anneau des vecteurs de Witt infinis (resp. de longueur $n \geq 1$) et $K_{\mathrm{a}} = \mathrm{Fr}\, W$ le corps des fractions de W. On note $|\cdot|$ la valeur absolue de K_{a} normalisée par $|p| = p^{-1}$ et σ l'endomorphisme de Frobenius agissant sur k,

W_n, W et K_{a}. On fixe une extension finie K/K_{a} totalement ramifiée et une clôture algébrique $\overline{K}$ de K. On note $\mathcal{O}_{\overline{K}}$ la clôture intégrale de $\mathcal{O}_K$ dans $\overline{K}$, $\overline{k}$ son corps résiduel, $\mathcal{O}_{\mathrm{C}}$ le complété p-adique de $\mathcal{O}_{\overline{K}}$ et $\mathrm{C} = \mathrm{Fr}\,\mathcal{O}_{\mathrm{C}}$. Pour toute extension finie L de K_{a}, contenue dans $\overline{K}$, on note $\mathcal{O}_L$ son anneau d'entiers, k_L son corps résiduel et $L_{\mathrm{a}} = \mathrm{Fr}\,\mathrm{W}(k_L)$. Pour tout $n \in \mathbb{N}$, soient μ_n le groupe des racines p^n-ièmes de l'unité dans $\overline{K}$ et $L_n = L(\mu_n)$. On note L_∞ la réunion des L_n, pour $n \in \mathbb{N}$, $\mathrm{H}_L = \mathrm{Gal}(\overline{K}/L_\infty)$ et $\Gamma_L = \mathrm{Gal}(L_\infty/L)$. Soit $\chi : \mathrm{G}_K \to \mathbb{Z}_p^*$ le caractère cyclotomique. Une représentation galoisienne p-adique (ou une $\mathbb{Q}_p$-représentation galoisienne) est un $\mathbb{Q}_p$-espace vectoriel de dimension finie, muni d'une action linéaire et continue de G_K. On note $\mathrm{Rep}_{\mathbb{Q}_p}(\mathrm{G}_K)$ la catégorie des $\mathbb{Q}_p$-représentations de G_K.

2 Conducteurs

On note $\mathrm{B}_{\mathrm{cris}}$ et $\mathrm{B}_{\mathrm{st}} = \mathrm{B}_{\mathrm{cris}}[X]$ les anneaux des périodes de Fontaine associés à K (cf. [12]). Soient $N : \mathrm{B}_{\mathrm{st}} \to \mathrm{B}_{\mathrm{st}}$ la $\mathrm{B}_{\mathrm{cris}}$-dérivation qui envoie X sur -1 et φ le Frobenius agissant sur $\mathrm{B}_{\mathrm{cris}}$ et B_{st}. Cettes applications vérifient $N\varphi = p\varphi N$. Ces anneaux sont munis d'une action continue de $\mathrm{G}_{K_{\mathrm{a}}}$ commutante avec φ et N. Soit L/K_{a} une extension finie contenue dans $\overline{K}$. On note $\mathrm{G}_L = \mathrm{Gal}(\overline{K}/L)$. On rappelle que $\mathrm{B}_{\mathrm{st}}^{\mathrm{G}_L} = \mathrm{B}_{\mathrm{cris}}^{\mathrm{G}_L} = L_{\mathrm{a}}$ (cf. [13, 5.1.2] et [12, 4.2.5]).
Pour tout $V \in \mathrm{Rep}_{\mathbb{Q}_p}(\mathrm{G}_K)$, Fontaine définit $\mathrm{D}_{\mathrm{cris}}(V) = (\mathrm{B}_{\mathrm{cris}} \otimes_{\mathbb{Q}_p} V)^{\mathrm{G}_K}$ et $\mathrm{D}_{\mathrm{st}}(V) = (\mathrm{B}_{\mathrm{st}} \otimes_{\mathbb{Q}_p} V)^{\mathrm{G}_K}$. Ce sont des K_{a}-espaces vectoriels de dimensions inférieures où égales à la dimension de V sur $\mathbb{Q}_p$. Le Frobenius de $\mathrm{B}_{\mathrm{cris}}$ induit un endomorphisme σ-semi-linéaire $\varphi : \mathrm{D}_{\mathrm{cris}}(V) \to \mathrm{D}_{\mathrm{cris}}(V)$, appelé Frobenius. L'espace $\mathrm{D}_{\mathrm{st}}(V)$ est muni d'un Frobenius φ et d'un endomorphisme K_{a}-linéaire N, vérifiant $N\varphi = p\varphi N$. On rappelle que V est dite cristalline (resp. semi-stable) si $\dim_{K_{\mathrm{a}}} \mathrm{D}_{\mathrm{cris}}(V) = \dim_{\mathbb{Q}_p} V$ (resp. $\dim_{K_{\mathrm{a}}} \mathrm{D}_{\mathrm{st}}(V) = \dim_{\mathbb{Q}_p} V$). Elle est dite potentiellement cristalline (resp. potentiellement semi-stable) s'il existe une extension finie K'/K telle que la restriction de V à $\mathrm{G}_{K'}$ est cristalline (resp. semi-stable). On note P le corps $K \otimes_{K_{\mathrm{a}}} \mathrm{Fr}\,\mathrm{W}(\overline{k})$. C'est le complété p-adique de l'extension maximale non-ramifiée K^{nr} de K dans $\overline{K}$. Le groupe d'inertie absolu de K est canoniquement isomorphe à G_P, le groupe de Galois absolu de P. Dans la suite, pour tout $V \in \mathrm{Rep}_{\mathbb{Q}_p}(\mathrm{G}_K)$, on considère la restriction de V à I_K comme une représentation p-adique de G_P. Par [12, 5.1.5], une représentation V est cristalline (resp. potentiellement cristalline, resp. semi-stable, resp. potentiellement semi-stable) si et seulement si sa restriction à I_K est cristalline (resp. potentiellement cristalline, resp. semi-stable, resp. potentiellement semi-stable).
Soit V une représentation p-adique potentiellement semi-stable de dimension n. Fontaine définit $\mathrm{D}_{\mathrm{pst}}(V) = \varinjlim_{G' \leq \mathrm{G}_K} (\mathrm{B}_{\mathrm{st}} \otimes_{\mathbb{Q}_p} V)^{G'}$, où la limite est prise sur les sous-groupes ouverts G' de G_K (cf. [13, 5.6.4]). C'est un $K_{\mathrm{a}}^{\mathrm{nr}}$-espace vectoriel de dimension n, muni d'une action semi-linéaire de G_K. Si L/K est une extension galoisienne finie telle que V est semi-stable comme représentation de G_L, alors $(\mathrm{B}_{\mathrm{st}} \otimes_{\mathbb{Q}_p} V)^{\mathrm{G}_L}$ est un L_{a}-espace vectoriel de dimension n et l'action

de G_K se factorise par $\mathrm{Gal}(L/K)$. On a un isomorphisme de représentations $\mathrm{D}_{\mathrm{pst}}(V) \cong K_{\mathrm{a}}^{\mathrm{nr}} \otimes_{L_{\mathrm{a}}} (\mathrm{B}_{\mathrm{st}} \otimes_{\mathbb{Q}_p} V)^{\mathrm{G}_L}$. Par conséquent la restriction $\mathrm{D}_{\mathrm{pst}}(V)_{|I_K}$ est une représentation linéaire de I_K qui se factorise à travers le sous-groupe d'inertie de $\mathrm{Gal}(L/K)$.

Définition 2.1. [10, 7.4.7] *Soit V une représentation p-adique potentiellement semi-stable. Les conducteurs de Swan et d'Artin de* $\mathrm{D}_{\mathrm{pst}}(V)_{|I_K}$ *sont aussi appelés conducteur de Swan et d'Artin de V et notés respectivement* $\mathrm{sw}(V)$ *et* $\mathrm{ar}(V)$.

Si G_K agit par un quotient fini $\mathrm{Gal}(L/K)$ sur V, alors $\mathrm{D}_{\mathrm{pst}}(V) = K_{\mathrm{a}}^{\mathrm{nr}} \otimes_{\mathbb{Q}_p} V$ et $\mathrm{sw}(V)$ et $\mathrm{ar}(V)$ coïncident avec $\mathrm{sw}(\mathrm{Gal}(L/K), V)$ et $\mathrm{ar}(\mathrm{Gal}(L/K), V)$ respectivement.
Pour une représentation V potentiellement semi-stable, on considère aussi la variante

$$\mathrm{ar}_{\mathrm{cris}}(V) = \mathrm{ar}(V) + \dim_{K_{\mathrm{a}}} \mathrm{D}_{\mathrm{st}}(V) - \dim_{K_{\mathrm{a}}} \mathrm{D}_{\mathrm{cris}}(V).$$

3 Théorie de Hodge p-adique et (φ, Γ)-modules

3.1 Les anneaux

On pose $\mathcal{O}_{\mathcal{E}} = \varprojlim_{n\in\mathbb{N}}(W_n[\![T]\!][\frac{1}{T}])$, qui est aussi la complétion p-adique de $W[\![T]\!][\frac{1}{T}]$. C'est un anneau de valuation discrète, complet, de caractéristique 0, absolument non ramifié, de corps résiduel $k(\!(T)\!)$. Son corps de fractions $\mathcal{O}_{\mathcal{E}}[1/p]$ est canoniquement isomorphe à

$$\mathcal{E} = \left\{ \sum_{n=-\infty}^{+\infty} a_n T^n \,\middle|\, a_n \in K_{\mathrm{a}}, (a_n)_{n\in\mathbb{Z}} \text{ bornée et } \lim_{n\to-\infty} |a_n| = 0 \right\}.$$

On pose

$$\mathcal{E}^{\dagger} = \left\{ \sum_{n=-\infty}^{+\infty} a_n T^n \in \mathcal{E} \,\middle|\, \exists\, 0 < \rho < 1 \text{ vérifiant } \lim_{n\to-\infty} |a_n|\rho^n = 0 \right\},$$

l'anneau des séries dans $\mathcal{E}$ qui convergent sur une couronne $\{x \in \mathrm{C} \,|\, \rho \le |x|_{\mathrm{C}} < 1\}$ pour un réel $0 < \rho < 1$. Pour tout $s \in \mathcal{E}^{\dagger}$, on note $\mathrm{v}_1(s) = \inf_{n\in\mathbb{Z}} \mathrm{v}_{K_{\mathrm{a}}}(a_n)$ la valuation de Gauss. On rappelle que cette valuation sur $\mathcal{E}^{\dagger}$ est discrète et que l'anneau de valuation $\mathcal{O}_{\mathcal{E}^{\dagger}}$ est hensélien, de corps résiduel $k(\!(T)\!)$(cf.[16, §2]).
On note aussi σ l'endomorphisme $x \mapsto x^p$ de $\mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}}$. Soit R (cf. [11, §A3.1.1]) la limite projective du système

$$\cdots \xrightarrow{\sigma} \mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}} \xrightarrow{\sigma} \mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}} \xrightarrow{\sigma} \mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}} \xrightarrow{\sigma} \mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}}.$$

C'est une $\overline{k}$-algèbre intègre, parfaite de caractéristique p. On dispose de la description suivante : $R \cong \{(x^{(n)})_{n\in\mathbb{N}} |\, x^{(n)} \in \mathcal{O}_{\mathrm{C}},\, (x^{(n+1)})^p = x^{(n)}\}$, où, à droite, la multiplication est donnée composante par composante et la somme par la formule $(x+y)^{(n)} = \lim_{m\to+\infty}\left(x^{(n+m)}+y^{(n+m)}\right)^{p^m}$. L'anneau R est complet pour la valuation (non-discète) définie, pour tout $x\in R$, par $\mathrm{v}_R(x) = \mathrm{v}_{\mathrm{C}}(x^{(0)})$, où v_{C} est la valuation de C normalisée par $\mathrm{v}_{\mathrm{C}}(p)=1$. Le corps $\operatorname{Fr} R$ est algébriquement clos (cf. [11, A3.1.6]). On rappelle que $\mathrm{W}(R) \cong \varprojlim_{n\in\mathbb{N}} \mathrm{W}_n(\mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}})$, où les applications de transition sont la composition des morphismes de troncation et du Frobenius σ de $\mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}}$. C'est une $\mathrm{W}(\overline{k})$-algèbre. Le groupe $\mathrm{G}_{K_{\mathrm{a}}}$ agit par fonctorialité sur $\mathrm{W}(R)$ et $\mathrm{W}(\operatorname{Fr} R)$. On appelle φ le Frobenius de $\mathrm{W}(R)$ (resp. $\mathrm{W}(\operatorname{Fr} R)$). On fixe une fois pour toutes un élément $\varepsilon\in R$ tel que $\varepsilon^{(0)}=1$ et $\varepsilon^{(1)}\neq 1$. Pour tout x dans R, on note $[x]$ son relèvement de Teichmüller dans $\mathrm{W}(R)$.

On vérifie facilement que $\left((\overline{\varepsilon^{(n)}},0,\ldots,0)-1\right)^{p^n} = 0$ dans $\mathrm{W}_n(\mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}})$. On en déduit, pour tout $n\in\mathbb{N}$, un morphisme continu $W[T]/T^{p^n} \to \mathrm{W}_n(\mathcal{O}_{\overline{K}}/p\mathcal{O}_{\overline{K}})$, qui envoie T sur $(\overline{\varepsilon}^{(n)},0,\ldots,0)-1$ et $w\in W$ sur $\sigma^{-n}(w)$. D'où un morphisme continu de W-algèbres $W[\![T]\!] \to \mathrm{W}(R)$, qui envoie T dans $[\varepsilon]-1$. Comme $[\varepsilon]-1$ est inversible dans $\mathrm{W}(\operatorname{Fr} R)$, on obtient un homomorphisme continu de W-algèbres $W[\![T]\!][\frac{1}{T}] \to \mathrm{W}(\operatorname{Fr} R)$, qui se factorise par complétion p-adique en

$$\begin{tikzcd}
W[\![T]\!][\frac{1}{T}] \arrow[rr] \arrow[dr, hook] & & \mathrm{W}(\operatorname{Fr} R) \\
& \mathcal{O}_{\mathcal{E}} \arrow[ur, "i"'] &
\end{tikzcd}$$

L'homomorphisme i est injectif car $i(p)\neq 0$. En inversant p, on obtient $i:\mathcal{E}\to \operatorname{Fr}\mathrm{W}(\operatorname{Fr} R)$.

Lemme 3.1. [11, A3.2.2] *Les anneaux $i(\mathcal{E})$ et $i(\mathcal{E}^\dagger)$ ne dépendent pas du choix de ε . Ils sont stables par les actions de $\mathrm{G}_{K_{\mathrm{a}}}$ et du Frobenius φ sur $\mathrm{W}(\operatorname{Fr} R)$. Les actions induites de $\mathrm{G}_{K_{\mathrm{a}}}$ et de φ sur $\mathcal{E}$ et $\mathcal{E}^\dagger$ sont données par*

$$\forall g\in \mathrm{G}_{K_{\mathrm{a}}},\qquad g(T) = (T+1)^{\chi(g)}-1,\qquad \varphi(T) = (T+1)^p-1.$$

L'action de $\mathrm{G}_{K_{\mathrm{a}}}$ se factorise par $\Gamma_{K_{\mathrm{a}}}$.

On rappelle brièvement la construction du corps des normes (cf. [25, §2.2]). L'extension maximale modérément ramifiée de K dans K_∞ est finie. Soit n_1 le plus petit entier tel que K_∞/K_{n_1} soit totalement sauvagement ramifiée. On choisit une uniformisante u de K_{n_1}. Pour tout $n\geq n_1$, le Frobenius de $\mathcal{O}_{K_{n+1}}/u\mathcal{O}_{K_{n+1}}$ se factorise à travers $\mathcal{O}_{K_n}/u\mathcal{O}_{K_n}\subset \mathcal{O}_{K_{n+1}}/u\mathcal{O}_{K_{n+1}}$. Soit $\lambda_n : \mathcal{O}_{K_{n+1}}/u\mathcal{O}_{K_{n+1}} \to \mathcal{O}_{K_n}/u\mathcal{O}_{K_n}$ le morphisme ainsi défini. On pose $\mathcal{O}_{\mathrm{E}_K} = \varprojlim_{n\in\mathbb{N}} \mathcal{O}_{K_n}/u\mathcal{O}_{K_n}$, où les applications de transitions sont les λ_n. C'est un anneau de valuation discrète, complet, de corps résiduel canoniquement isomorphe au corps résiduel de K_∞, qui est une extension finie k' de k. Il ne dépend pas du choix de u. Soit $\mathrm{E}_K = \operatorname{Fr}\mathcal{O}_{\mathrm{E}_K}$. Par fonctorialité de la construction, on associe à $\overline{K}$ une clôture séparable $\mathrm{E}_K^{\mathrm{sep}}$ de E_K et $\mathrm{Gal}(\mathrm{E}_K^{\mathrm{sep}}/\mathrm{E}_K)$ est canoniquement isomorphe à H_K. Pour tout $n \geq n_1$, on a un diagramme commutatif

$$\begin{tikzcd} \mathcal{O}_{K_{n+1}}/u\mathcal{O}_{K_{n+1}} \arrow[r, hook] \arrow[d, "\lambda_n"] & \mathcal{O}_{\overline{K}}/u\mathcal{O}_{\overline{K}} \arrow[d, "\sigma"] \\ \mathcal{O}_{K_n}/u\mathcal{O}_{K_n} \arrow[r, hook] & \mathcal{O}_{\overline{K}}/u\mathcal{O}_{\overline{K}} \end{tikzcd}$$

Comme $R \cong \varprojlim_{n\geq n_1} \mathcal{O}_{\overline{K}}/u\mathcal{O}_{\overline{K}}$, on en déduit des applications injectives $\mathcal{O}_{\mathrm{E}_K} \hookrightarrow R$ et $\mathrm{E}_K \hookrightarrow \mathrm{Fr}\, R$.

Soient $\mathcal{E}^{\mathrm{nr}}$ l'extension maximale non-ramifiée de $\mathcal{E}$ dans $\mathrm{Fr}\,\mathrm{W}(\mathrm{Fr}\, R)$ et $\mathcal{O}_{\mathcal{E}}^{\mathrm{sh}}$ son anneau d'entiers. Par fonctorialité de la hensélisation, $\mathrm{G}_{K_{\mathrm{a}}}$ et φ agissent sur $\mathcal{E}^{\mathrm{nr}}$. L'inclusion $i : \mathcal{O}_{\mathcal{E}} \hookrightarrow \mathrm{W}(\mathrm{Fr}\, R)$ induit, par réduction modulo p, un isomorphisme canonique entre les corps résiduels de $\mathcal{O}_{\mathcal{E}}$ et $\mathrm{E}_{K_{\mathrm{a}}}$. Par conséquent, $\mathrm{Gal}(\mathcal{E}^{\mathrm{nr}}/\mathcal{E})$ est canoniquement isomorphe à $\mathrm{H}_{K_{\mathrm{a}}}$. Soit L une extension finie de K_{a}. On pose $\mathcal{E}_L = (\mathcal{E}^{\mathrm{nr}})^{\mathrm{H}_L}$. C'est une extension finie non ramifiée de $\mathcal{E}$. Elle est munie d'actions naturelles de Γ_L et de φ . Pour L/K_{a} finie galoisienne, $\mathcal{E}_L$ est muni d'une action naturelle de $\mathrm{Gal}(L_\infty/K_{\mathrm{a}})$. On note $\mathcal{O}_{\mathcal{E}_L}$ l'anneau de valuation de $\mathcal{E}_L$. On note k'_L le corps résiduel de E_L et $L' = \mathrm{Fr}\,\mathrm{W}(k'_L)$. Si L est absolument non-ramifié, alors $k'_L = k_L$ et $L' = L_{\mathrm{a}} = L$ (cf. [20, Ch.IV, Prop.17]). Dans ce cas le corps $\mathcal{E}_L$ a la description simple suivante.

Lemme 3.2. *Soit L/K_{a} une extension finie non-ramifiée. Il existe un isomorphisme canonique $\mathcal{E}_L \cong \mathcal{E} \otimes_{K_{\mathrm{a}}} L$, compatible avec l'action de Γ_L et du Frobenius. Pour L/K_{a} finie galoisienne, cet isomorphisme est compatible à l'action de* $\mathrm{Gal}(L_\infty/K_{\mathrm{a}})$.

Démonstration. L'anneau $\mathcal{E}\otimes_{K_{\mathrm{a}}} L$ est un corps car K_{a} est algébriquement fermé dans $\mathcal{E}$ et p est inversible. Comme $L = L_{\mathrm{a}} \subseteq \mathrm{Fr}\,\mathrm{W}(\mathrm{Fr}\, R)$, l'inclusion i s'étend, par linéarité, en $i_L : \mathcal{E} \otimes_{K_{\mathrm{a}}} L \hookrightarrow \mathrm{Fr}\,\mathrm{W}(\mathrm{Fr}\, R)$. L'image de cette application est contenue dans $\mathcal{E}_L$. On a $|i_L(\mathcal{E} \otimes_{K_{\mathrm{a}}} L) : \mathcal{E}| = |k_L : k| = |\mathrm{E}_L : \mathrm{E}_{K_{\mathrm{a}}}| = |\mathcal{E}_L : \mathcal{E}|$, donc $i_L(\mathcal{E} \otimes_{K_{\mathrm{a}}} L) = \mathcal{E}_L$. □

Proposition 3.3. [11, A2.2.1] *Soient $\overline{\pi}$ une uniformisante de E_L et π un relèvement dans $\mathcal{O}_{\mathcal{E}_L}$. Il existe un unique isomorphisme continu de L'-algèbres, $\psi_\pi : \mathcal{E}_{L'} \to \mathcal{E}_L$ qui envoie T sur π.*

Démonstration. L'anneau $\mathcal{O}_{\mathcal{E}_L}$ est de valuation discrète, complet, absolument non-ramifié. C'est donc un anneau de Cohen (cf. [9, IV_0 19.8.5]). Par [9, IV_0 19.8.6(ii)] il existe un isomorphisme $\psi : \mathcal{O}_{\mathcal{E}_{L'}} \to \mathcal{O}_{\mathcal{E}_L}$ relevant l'isomorphisme $k'_L((T)) \to \mathrm{E}_L$ qui envoie T sur $\overline{\pi}$. On note $\omega = \pi - \psi(T) \in p\mathcal{O}_{\mathcal{E}_L}$. Soit

$s = \sum_{n\in\mathbb{Z}} a_n T^n \in \mathcal{O}_{\mathcal{E}_{L'}}$. On écrit

$$\sum_{n=0}^{m} a_n\pi^n = \sum_{n=0}^{m} a_n(\psi(T)+\omega)^n = \sum_{n=0}^{m} a_n \sum_{i=0}^{n}\binom{n}{i}\psi(T)^i\omega^{n-i} =$$
$$= \sum_{j=0}^{m}\left(\sum_{i=0}^{m-j} a_{j+i}\binom{j+i}{i}\psi(T)^i\right)\omega^j = \sum_{j=0}^{m}\psi\left(\sum_{i=0}^{m-j} a_{j+i}\binom{j+i}{i}T^i\right)\omega^j,$$

où $j = n-i$. Pour tout $j \in \mathbb{N}$, on pose $s_j = \sum_{i=0}^{+\infty} a_{j+i}\binom{j+i}{i}T^i \in \mathcal{O}_{\mathcal{E}_{L'}}$. On définit $\psi_\pi(s) = \sum_{n<0} a_n\pi^n + \sum_{j\geq 0}\psi(s_j)\omega^j$, qui converge p-adiquement car $\omega \in p\mathcal{O}_{\mathcal{E}_L}$ et $\lim_{n\to-\infty} a_n = 0$. L'application ψ_π est un isomorphisme $\mathcal{O}_{\mathcal{E}_{L'}} \to \mathcal{O}_{\mathcal{E}_L}$, qui envoie T sur π. On la prolonge en un isomorphisme $\mathcal{E}_{L'} \to \mathcal{E}_L$. L'unicité est évidente. □

Soient $\mathcal{O}^{\mathrm{sh}}_{\mathcal{E}^\dagger}$ l'hensélisé strict de $\mathcal{O}_{\mathcal{E}^\dagger}$ dans $\mathrm{W}(\mathrm{Fr}\,R)$ et $\mathcal{E}^{\dagger\,\mathrm{nr}}$ son corps des fractions. Par fonctorialité de la hensélisation, $\mathrm{G}_{K_{\mathrm{a}}}$ et φ agissent sur $\mathcal{E}^{\dagger\,\mathrm{nr}}$. Comme plus haut, l'inclusion canonique $\mathcal{O}^{\mathrm{sh}}_{\mathcal{E}^\dagger} \hookrightarrow \mathrm{W}(\mathrm{Fr}\,R)$ induit un isomorphisme $\mathrm{Gal}(\mathcal{E}^{\dagger\,\mathrm{nr}}/\mathcal{E}^\dagger) \cong \mathrm{H}_{K_{\mathrm{a}}}$. Soit L une extension finie de K_{a}. On pose $\mathcal{E}^\dagger_L = (\mathcal{E}^{\dagger\,\mathrm{nr}})^{\mathrm{H}_L}$. C'est une extension finie non ramifiée de $\mathcal{E}^\dagger$, de corps résiduel E_L. Pour L/K_{a} finie galoisienne non-ramifiée, on démontre comme dans le lemme 3.2, qu'il y a un isomorphisme canonique, $\mathcal{E}^\dagger \otimes_{K_{\mathrm{a}}} L \cong \mathcal{E}^\dagger_L$, compatible avec l'action de $\mathrm{Gal}(L_\infty/K_{\mathrm{a}})$ et du Frobenius.

Proposition 3.4. [16, Prop. 3.4] *Soit $\overline{\pi}$ une uniformisante de E_L. Il existe un relèvement π de $\overline{\pi}$ dans $\mathcal{O}_{\mathcal{E}^\dagger_L}$ tel que sous l'isomorphisme $\psi_\pi : \mathcal{E}_{L'} \to \mathcal{E}_L$, on ait $\psi_\pi(\mathcal{E}^\dagger_{L'}) = \mathcal{E}^\dagger_L$.*

On note

$$\mathcal{R} = \left\{ \sum_{n=-\infty}^{+\infty} a_n T^n \,\middle|\, a_n \in K_{\mathrm{a}},\ \exists\rho_c \in]0,1[\text{ t.q. } \forall\rho \in]\rho_c,1[\,,\ \lim_{n\to\pm\infty} |a_n|\rho^n = 0 \right\}.$$

On munit cet anneau d'une action du Frobenius et de $\Gamma_{K_{\mathrm{a}}}$ en posant :

$$\varphi(T) = (1+T)^p - 1 \text{ et } \forall\gamma \in \Gamma_{K_{\mathrm{a}}},\quad \gamma(T) = (1+T)^{\chi(\gamma)} - 1.$$

On pose $t = \log(T+1) \in \mathcal{R}$, qu'on note aussi abusivement $\log[\varepsilon]$, bien qu'il n'appartienne pas à $\mathrm{W}(\mathrm{Fr}\,R)$. On a une inclusion $\mathcal{E}^\dagger \subset \mathcal{R}$ compatible avec les actions du Frobenius et de $\Gamma_{K_{\mathrm{a}}}$.
Pour toute extension finie L/K_{a}, on pose $\mathcal{R}_L = \mathcal{R} \otimes_{\mathcal{E}^\dagger} \mathcal{E}^\dagger_L$. On vérifie que cet anneau est intègre. C'est une extension étale finie de $\mathcal{R}$. On le munit du Frobenius produit tensoriel des Frobenius sur $\mathcal{R}$ et sur $\mathcal{E}^\dagger_L$. Si L/K_{a} est galoisienne finie, alors le groupe $\mathrm{Gal}(L_\infty/K_{\mathrm{a}})$ agit sur $\mathcal{R}_L$ en agissant sur $\mathcal{R}$ à travers le quotient $\Gamma_{K_{\mathrm{a}}}$ et sur $\mathcal{E}^\dagger_L$ via son action naturelle. On a $\mathcal{R}_L^{\mathrm{Gal}(L_\infty/(K_{\mathrm{a}})_\infty)} = \mathcal{R}$. On vérifie

facilement que $\log \frac{(1+T)^p-1}{T^p} \in \mathcal{R}$ et pour tout $\gamma \in \Gamma_{K_\mathrm{a}}$, $\log \frac{(1+T)^{\chi(\gamma)}-1}{T} \in \mathcal{R}$. On prolonge les actions de φ et $\mathrm{Gal}(L_\infty/K_\mathrm{a})$ à l'anneau des polynômes $\mathcal{R}_L[X]$ par

$$\varphi_L(X) = pX + \log \frac{(1+T)^p - 1}{T^p}$$
$$\forall \gamma \in \mathrm{Gal}(L_\infty/K_\mathrm{a}), \quad \gamma(X) = X + \log \frac{(1+T)^{\chi(\gamma)} - 1}{T}.$$

On notera formellement $X = \log T$. On désigne par $\mathcal{R}[1/t]$ (resp. $\mathcal{R}[\log T][1/t]$) le localisé de $\mathcal{R}$ (resp. $\mathcal{R}[\log T]$) en t. Dans la suite, on considérera souvent sur $\mathcal{R}_L[1/t]$ et $\mathcal{R}_L[\log T][1/t]$ l'action du sous groupe Γ_L de $\mathrm{Gal}(L_\infty/K_\mathrm{a})$. On vérifie que (cf. [3, Prop. 3.3])

$$(\mathcal{R}_L[\log T][1/t])^{\Gamma_L} = L_\mathrm{a}.$$

Soit L/K_a une extension finie galoisienne, on a $\mathcal{R}_{L'} \cong \mathcal{R} \otimes_{K_\mathrm{a}} L'$. Si on prend un relevement $\pi \in \mathcal{E}_L^\dagger$ d'une uniformisante de E_L, satisfaisant la proposition 3.4, alors ψ_π se prolonge en un isomorphisme $\psi_\pi : \mathcal{R}_{L'} \to \mathcal{R}_L$.

3.2 Le théorème de comparaison de Berger

Un (φ, Γ)-module sur $\mathcal{O}_{\mathcal{E}_K}$ (resp. $\mathcal{E}_K$) est un $\mathcal{O}_{\mathcal{E}_K}$-module de type fini (resp. un $\mathcal{E}_K$-espace vectoriel de dimension finie) muni d'une action semi-linéaire et continue de Γ_K et d'un endomorphisme φ semi-linéaire par rapport au Frobenius de $\mathcal{O}_{\mathcal{E}_K}$ (resp. $\mathcal{E}_K$) commutant entre eux. On dit qu'un (φ, Γ)-module $\mathcal{M}$ sur $\mathcal{O}_{\mathcal{E}_K}$ est étale si l'image $\varphi(\mathcal{M})$ engendre $\mathcal{M}$ sur $\mathcal{O}_{\mathcal{E}_K}$.
Pour tout (φ, Γ)-module M sur $\mathcal{E}_K$ on peut choisir un réseau stable par φ et Γ_K, qui est donc un (φ, Γ)-module sur $\mathcal{O}_{\mathcal{E}_K}$. On dit qu'un (φ, Γ)-module M sur $\mathcal{E}_K$ est étale s'il existe un réseau $\mathcal{M}$ de M stable par Γ_K et φ qui est étale. On note $\Phi\Gamma^{\text{ét}}_{\mathcal{E}_K}$ la catégorie dont les objets sont les (φ, Γ)-modules étales et les morphismes sont les applications linéaires commutant avec φ et Γ_K. Dans [11], Fontaine construit une équivalence de catégories entre $\mathrm{Rep}_{\mathbb{Q}_p}(\mathrm{G}_K)$ et $\Phi\Gamma^{\text{ét}}_{\mathcal{E}_K}$. On rappelle sa construction brièvement. On note $\widehat{\mathcal{E}^{\mathrm{nr}}}$ le completé p-adique de $\mathcal{E}^{\mathrm{nr}}$. Pour toute $\mathbb{Q}_p$-représentation V de G_K, on considère le $\mathcal{E}_K$-espace vectoriel $\mathrm{D}(V) = (\widehat{\mathcal{E}^{\mathrm{nr}}} \otimes_{\mathbb{Q}_p} V)^{\mathrm{H}_K}$ muni des actions semi-linéaires de Γ_K et de φ . C'est un objet de $\Phi\Gamma^{\text{ét}}_{\mathcal{E}_K}$. Le foncteur D définit une équivalence de catégories entre $\mathrm{Rep}_{\mathbb{Q}_p}(\mathrm{G}_K)$ et $\Phi\Gamma^{\text{ét}}_{\mathcal{E}_K}$.

Théorème 3.5 (Cherbonnier-Colmez). [5, III 5.2] *Soit $V \in \mathrm{Rep}_{\mathbb{Q}_p}(\mathrm{G}_K)$. La famille des sous-$\mathcal{E}_K^\dagger$-modules de type fini de $\mathrm{D}(V)$ stables par φ et Γ_K admet un plus grand élément $\mathrm{D}^\dagger(V)$ et on a*

$$\mathrm{D}(V) = \mathcal{E}_K \otimes_{\mathcal{E}_K^\dagger} \mathrm{D}^\dagger(V).$$

Pour toute représentation p-adique V de G_K, Berger définit $\mathrm{D}^{\dagger}_{\mathrm{rig}}(V) = \mathcal{R}_K \otimes_{\mathcal{E}^{\dagger}_K} \mathrm{D}^{\dagger}(V)$ et $\mathrm{D}^{\dagger}_{\log}(V) = \mathcal{R}_K[\log T] \otimes_{\mathcal{E}^{\dagger}_K} \mathrm{D}^{\dagger}(V)$ avec les actions évidentes de φ et Γ_K (cf.[3, §3.2]). Soit L/K une extension galoisienne finie. Le module $\mathrm{D}(V_{|\mathrm{G}_L})$, est muni naturellement d'une action de $\mathrm{Gal}(L_\infty/K)$. Par le théorème 3.5, on obtient une action de $\mathrm{Gal}(L_\infty/K)$ sur $\mathrm{D}^{\dagger}(V_{|\mathrm{G}_L})$ et donc sur $\mathrm{D}^{\dagger}_{\mathrm{rig}}(V_{|\mathrm{G}_L})$ et $\mathrm{D}^{\dagger}_{\log}(V_{|\mathrm{G}_L})$.

Théorème 3.6 (Berger). [3, 3.6] *Soit V est une représentation p-adique de G_K. On a des isomorphismes canoniques :*

$$\mathrm{D}_{\mathrm{cris}}(V) \cong \left(\mathrm{D}^{\dagger}_{\mathrm{rig}}(V)[1/t]\right)^{\Gamma_K} \ \text{et}\ \mathrm{D}_{\mathrm{st}}(V) \cong \left(\mathrm{D}^{\dagger}_{\log}(V)[1/t]\right)^{\Gamma_K}.$$

Soit L/K une extension galoisienne finie. Alors les isomorphismes ci-dessus $\mathrm{D}_{\mathrm{cris}}(V_{|\mathrm{G}_L}) \cong \left(\mathrm{D}^{\dagger}_{\mathrm{rig}}(V_{|\mathrm{G}_L})[1/t]\right)^{\Gamma_L}$ et $\mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L}) \cong \left(\mathrm{D}^{\dagger}_{\log}(V_{|\mathrm{G}_L})[1/t]\right)^{\Gamma_L}$ sont équivariants pour les actions naturelles de $\mathrm{Gal}(L/K)$.

Remarque 3.7. *Dans [3, 3.6], l'assertion sur l'équivariance par rapport à* $\mathrm{Gal}(L/K)$ *n'apparait pas. C'est une conséquence immédiate de la démonstration. On l'a mise en évidence pour l'importance qu'elle jouera dans la suite.*

4 Équations différentielles p-adiques

Soit $\hat{\Omega}^1_{\mathcal{R}/K_{\mathrm{a}}}$ le module des différentielles continues de $\mathcal{R}$ sur K_{a}. C'est un $\mathcal{R}$-module libre de rang 1 de base $\frac{dT}{T+1} = \frac{d[\varepsilon]}{[\varepsilon]}$. Soient L/K_{a} une extension finie et L' l'extension finie non-ramifiée de K_{a} qui lui est associée dans la section §3.1 (au dessus de Lemme 3.2). Comme $\mathcal{R} \hookrightarrow \mathcal{R}_L$ est étale finie et L'/K_{a} est finie, on a un isomorphisme canonique $\hat{\Omega}^1_{\mathcal{R}_L/L'} \cong \mathcal{R}_L \otimes_{\mathcal{R}} \hat{\Omega}^1_{\mathcal{R}/K_{\mathrm{a}}}$. On étend la dérivation $d : \mathcal{R} \to \hat{\Omega}^1_{\mathcal{R}/K_{\mathrm{a}}}$ en $d \colon \mathcal{R}[\log T] \to \hat{\Omega}^1_{\mathcal{R}/K_{\mathrm{a}}}$, en posant $d(\log T) = \frac{1}{T}dT$. Le corps des constantes de $\mathcal{R}_L[1/t]$ et de $\mathcal{R}_L[\log T]$ est L'.
On appelle équation différentielle p-adique (ou module à connexion) sur $\mathcal{R}_K$ un $\mathcal{R}_K$-module de présentation finie muni d'une connexion. On démontre qu'un tel module est forcement libre (cf. [2, Prop. 2.3]). Une équation différentielle p-adique est dite unipotente (resp. quasi-unipotente) si elle est extension itérée d'équations différentielles triviales (resp. s'il existe une extension finie L/K telle que $M \otimes_{\mathcal{R}_K} \mathcal{R}_L$ soit unipotente). On dit qu'une équation différentielle p-adique M est munie d'une structure de Frobenius s'il existe un endomorphisme $\varphi_M : M \to M$, φ-semi-linéaire, horizontal, tel que $\varphi_M(M)$ engendre M sur $\mathcal{R}_K$.

4.1 Équation différentielle p-adique associée à une représentation de de Rham

Soit V une représentation galoisienne p-adique de G_K. On rappelle que Berger démontre que pour tout $x \in \mathrm{D}^{\dagger}_{\mathrm{rig}}(V)[1/t]$, la limite $\lim_{\gamma \to \mathrm{Id}_{\Gamma_K}} \frac{\gamma(x)-x}{\chi(\gamma)-1}$ existe

(cf. [3, §5.1]). On note cette limite $\nu(x)$. L'application $x \mapsto t^{-1}\nu(x) \otimes \frac{dT}{T+1}$ définit une connexion $\nabla_V \colon \mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t] \to \mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t] \otimes_{\mathcal{R}_K} \hat{\Omega}^1_{\mathcal{R}_K/K'}$.
Soit V une représentation de de Rham de dimension n. On rappelle que Berger montre qu'il existe un unique sous-$\mathcal{R}_K$-module libre de rang n de $\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]$ stable par $t^{-1}\nu$. On l'appelle $\mathrm{N}_{\mathrm{dR}}(V)$. Il est stable par l'action du Frobenius et de Γ_K (cf. [3, §5.4 et §5.5] ou [4, IV.4]). On vérifie qu'il est une muni d'une structure de Frobenius. Par construction, on associe à un morphisme de représentations de de Rham $V_1 \to V_2$, un morphisme d'équations différentielles p-adiques $\mathrm{N}_{\mathrm{dR}}(V_1) \to \mathrm{N}_{\mathrm{dR}}(V_2)$.

Théorème 4.1 (Berger). [3, 5.20] *On a un foncteur additif $V \mapsto \mathrm{N}_{\mathrm{dR}}(V)$, de la catégorie des représentations p-adiques de de Rham de G_K, dans la catégorie des équations différentielles p-adiques sur $\mathcal{R}_K$ munies d'une structure de Frobenius. Ce foncteur associe à une représentation de dimension n une équation différentielle de rang n. C'est un $\otimes$-foncteur exacte et fidèle. L'équation $\mathrm{N}_{\mathrm{dR}}(V)$ est quasi-unipotente si et seulement si la représentation V est potentiellement semi-stable. L'équation $\mathrm{N}_{\mathrm{dR}}(V)$ est unipotente (resp. triviale) si et seulement s'il existe n tel que la restriction de V à G_{K_n} soit semi-stable (resp. cristalline).*

On étend ∇_V en une connexion ∇_V sur $\mathrm{D}^\dagger_{\log}(V)$ et $\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_K[\log T]$, en posant $\nabla_V(\log T) = \frac{dT}{T}$.
Soit M un module à connexion sur $\mathcal{R}_K$. On vérifie aisément que la dimension sur L' des sections horizontales de $M \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T]$ est inférieure ou égale au rang de M. Si on a l'égalité on dit que M est triviale sur $\mathcal{R}_L[\log T]$. Le module M est unipotent si et seulement si M est triviale sur $\mathcal{R}_K[\log T]$.

Corollaire 4.2. *Si V est cristalline (resp. semi-stable), alors*

$$\mathrm{D}_{\mathrm{cris}}(V) \otimes_{K_{\mathrm{a}}} K' \cong \left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\nabla_V} \cong \left(\mathrm{N}_{\mathrm{dR}}(V)\right)^{\nabla_V}$$
$$\left(\textit{resp. } \mathrm{D}_{\mathrm{st}}(V) \otimes_{K_{\mathrm{a}}} K' \cong \left(\mathrm{D}^\dagger_{\log}(V)[1/t]\right)^{\nabla_V} \cong \left(\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_K[\log T]\right)^{\nabla_V}\right).$$

Démonstration. Le théorème 3.6 implique que $\mathrm{D}_{\mathrm{cris}}(V) \cong \left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\Gamma_K}$. Par définition de ∇_V, on a $\left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\Gamma_K} \subseteq \left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\nabla_V}$. Car si $x \in \left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\Gamma_K}$, alors $\nu(x) = \lim_{\gamma \to \mathrm{Id}_{\Gamma_K}} \frac{\gamma(x)-x}{\chi(\gamma)-1} = 0$. D'autre part, par définition, on a aussi, $\left(\mathrm{N}_{\mathrm{dR}}(V)\right)^{\nabla_V} \subseteq \left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\nabla_V}$. On vérifie facilement que la dimension sur K' de $\left(\mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]\right)^{\nabla_V}$ est inférieure ou égale à $\dim_{\mathbb{Q}_p} V$. Comme V est cristalline, $\dim_{\mathbb{Q}_p} V = \dim_{K_{\mathrm{a}}} \mathrm{D}_{\mathrm{cris}}(V) = \dim_{K'} \left(\mathrm{N}_{\mathrm{dR}}(V)\right)^{\nabla_V}$. D'où les premiers isomorphismes. Le cas semi-stable est analogue. $\square$

Théorème 4.3 (André, Kedlaya, Mebkhout). [1, 7.1.5] [15, 1.1] [18, 5.0-23] *Tout module à connexion sur $\mathcal{R}_K$, admettant une structure de Frobenius est quasi-unipotent.*

4.2 Rappels sur l'irrégularité d'une équation différentielle p-adique

On rappelle brièvement la définition de l'indice d'une équation différentielle p-adique introduite initialement par Robba (cf. [7, §2.3] et [8, §14]). Soient C un corps et u un endomorphisme d'un C-espace vectoriel V. On dit que u admet un indice si $\operatorname{Ker} u$ et $\operatorname{Coker} u$ sont de dimension finie et on appelle indice de u l'entier $\chi(u, V) = \dim_C \operatorname{Ker} u - \dim_C \operatorname{Coker} u$. On note $\mathcal{A} = \{\sum_{n=0}^{+\infty} a_n T^n \in \mathcal{R}\}$. C'est la sous-algèbre de $\mathcal{R}$ des séries convergentes sur le disque ouvert. On note γ_+ l'inclusion de $\mathcal{A}$ dans $\mathcal{R}$ et et γ^+ la troncation $\sum_{n\in\mathbb{Z}} a_n T^n \mapsto \sum_{n\in\mathbb{N}} a_n T^n$ de $\mathcal{R}$ sur $\mathcal{A}$. Pour tout nombre naturel n, on note abusivement γ_+ (resp. γ^+) l'inclusion de $\mathcal{A}^{\oplus n}$ dans $\mathcal{R}^{\oplus n}$ (resp. la projection de $\mathcal{R}^{\oplus n}$ sur $\mathcal{A}^{\oplus n}$). Soit M un module à connexion sur $\mathcal{R}$ de rang n. On choisit une base de M. Soient $\xi = T\frac{d}{dT}$ et G la matrice de la dérivation $\nabla_\xi : M \to M$ par rapport à cette base. Soit u le K_{a}-endomorphisme $T\frac{d}{dT} - G$ de $\mathcal{R}^{\oplus n}$. On dit que M admet un indice généralisé sur $\mathcal{A}$ si $\gamma^+ \circ u \circ \gamma_+$ admet un indice. Ceci ne dépend pas de la base choisie. On note $\widetilde{\chi}(M, \mathcal{A}) = \chi(\gamma^+ \circ u \circ \gamma_+, \mathcal{A}^{\oplus n})$. Si deux modules à connexion sur $\mathcal{R}$, M' et M'' ont un indice généralisé, alors pour toute extension M de M' par M'', $\widetilde{\chi}(M, \mathcal{A}) = \widetilde{\chi}(M', \mathcal{A}) + \widetilde{\chi}(M'', \mathcal{A})$.

Théorème 4.4 (Christol-Mebkhout). *Soit M une équation différentielle p-adique ayant une structure de Frobenius. Alors M admet un indice généralisé sur $\mathcal{A}$.*

Démonstration. L'existence d'une structure de Frobenius implique la solubilité de M et que ses exposants sont non-Liouville (cf. [7, §2.5]). Le corps de constantes K_{a} est de valuation discrète donc maximalement complet. C'est donc un cas particulier de [8, Th. 14.11]. □

Pour une équation différentielle p-adique M munie d'une structure de Frobenius, on appelle irrégularité de M et on note $\operatorname{irr}(M)$, l'indice généralisé $\widetilde{\chi}(M, \mathcal{A})$.

Remarque 4.5. *1. Soient L/K_{a} une extension finie et M un module libre de rang r sur $\mathcal{R} \otimes_{K_{\mathrm{a}}} L$ muni d'une connexion. On considère M comme une équation différentielle p-adique sur $\mathcal{R}$ de rang $r \dim_{K_{\mathrm{a}}} L$. Si M admet un indice généralisé, on pose $\operatorname{irr}(M) = (\dim_{K_{\mathrm{a}}} L)^{-1} \widetilde{\chi}(M, \mathcal{A})$.*
2. Soit M une équation différentielle p-adique sur $\mathcal{R}_K$. On choisit un élément π dans $\mathcal{O}_{\mathcal{E}_K^\dagger}$ comme dans 3.4. *Via les isomorphismes $\psi_\pi : \mathcal{R}_{K'} \to \mathcal{R}_K$ et $\mathcal{R}_{K'} \cong \mathcal{R} \otimes_{K_{\mathrm{a}}} K'$, on peut définir l'irrégularité de M. On vérifie que ceci ne dépend pas du choix de π.*
3. L'indice généralisé coïncide avec la définition d'indice sur $\mathcal{E}^\dagger$ de Tsuzuki dans [23, §1] *(cf.* [17, §8]*).*

4.3 Équations quasi-unipotentes

La catégorie des équations différentielles p-adiques quasi-unipotentes est tannakienne neutre. On rappelle ici des constructions classiques (cf. [1] et [17]).

Soit $k(\!(T)\!)^{\mathrm{sep}}$ une clôture séparable fixée de $k(\!(T)\!)$ (dans la section 5 on suppose $k(\!(T)\!)^{\mathrm{sep}} \subset \mathrm{Fr}\, R$). Soit $E/k(\!(T)\!)$ une extension séparable finie contenue dans $k(\!(T)\!)^{\mathrm{sep}}$. On note k_E son corps résiduel, $\mathrm{G}_E = \mathrm{Gal}(k(\!(T)\!)^{\mathrm{sep}}/E)$ et I_E le sous-groupe d'inertie. On pose $\mathcal{E}(E) = (\mathcal{E}^{\mathrm{nr}})^{\mathrm{G}_E}$, $\mathcal{E}^\dagger(E) = \mathcal{E}(E) \cap \mathcal{E}^{\dagger^{\mathrm{nr}}}$ et $\mathcal{R}(E) = \mathcal{R} \otimes_{\mathcal{E}^\dagger} \mathcal{E}^\dagger(E)$. Évidement, si E est égal au corps de normes E_L associé à une extension finie L/K, alors $\mathcal{E}(\mathrm{E}_L) = \mathcal{E}_L$, $\mathcal{E}^\dagger(\mathrm{E}_L) = \mathcal{E}_L^\dagger$ et $\mathcal{R}(\mathrm{E}_L) = \mathcal{R}_L$. Les propriétés qu'on a rappelées dans les sections précédentes pour $\mathcal{E}_L, \mathcal{E}_L^\dagger, \mathcal{R}_L$, en dehors de l'action de Γ_L, sont aussi valables pour $\mathcal{E}(E), \mathcal{E}^\dagger(E), \mathcal{R}(E)$. Soient F/E une extension galoisienne finie et M un $\mathcal{R}(E)$-module à connexion unipotent sur $\mathcal{R}(F)$. On considère le $\mathrm{Fr}\,\mathrm{W}(k_F)$-espace vectoriel des sections horizontales

$$S_F(M) = (M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T])^{\nabla}.$$

On le munit d'une action semi-linéaire de $\mathrm{Gal}(F/E)$ et d'un endomorphisme nilpotent de la façon suivante. Pour tout $g \in \mathrm{Gal}(F/E)$ et $x \in M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T]$, on pose $g(x) = (\mathrm{Id} \otimes g)(x)$. On considère sur $M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T]$ l'application $\mathrm{Id} \otimes N$, où N est la $\mathcal{R}(F)$-dérivation de $\mathcal{R}(F)[\log T]$ qui envoie $\log T$ sur 1. Cette application commute avec l'action de $\mathrm{Gal}(F/E)$. Les diagrammes suivantes sont commutatifs :

$$\begin{tikzcd}
M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \arrow[r, "\nabla"] \arrow[d, "\mathrm{Id}\otimes g"'] & M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \otimes_{\mathcal{R}_F} \hat{\Omega}^1_{\mathcal{R}_F/\mathrm{Fr}\,\mathrm{W}(k_F)} \arrow[d, "\mathrm{Id}\otimes g \otimes dg"] \\
M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \arrow[r, "\nabla"] & M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \otimes_{\mathcal{R}_F} \hat{\Omega}^1_{\mathcal{R}_F/\mathrm{Fr}\,\mathrm{W}(k_F)}
\end{tikzcd}$$

et

$$\begin{tikzcd}
M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \arrow[r, "\nabla"] \arrow[d, "\mathrm{Id}\otimes N"] & M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \otimes_{\mathcal{R}_F} \hat{\Omega}^1_{\mathcal{R}_F/\mathrm{Fr}\,\mathrm{W}(k_F)} \arrow[d, "\mathrm{Id}\otimes N \otimes \mathrm{Id}"] \\
M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \arrow[r, "\nabla"] & M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \otimes_{\mathcal{R}_F} \hat{\Omega}^1_{\mathcal{R}_F/\mathrm{Fr}\,\mathrm{W}(k_F)}
\end{tikzcd}$$

On en déduit sur $S_F(M)$, une action de $\mathrm{Gal}(F/E)$ et un endomorphisme, qu'on note $N_{S_F(M)}$, commutant entre eux. On vérifie facilement que $N_{S_F(M)}$ est nilpotent. On peut résumer ces données en disant que $S_F(M)$ est une $\mathrm{Fr}\,\mathrm{W}(k_F)$-représentation semi-linéaire du schéma en groupes $\mathrm{Gal}(F/E) \times \mathbb{G}_a$, où $\mathrm{Gal}(F/E)$ est considéré comme schéma en groupes constant. La dimension de $S_F(M)$ sur $\mathrm{Fr}\,\mathrm{W}(k_F)$ est par construction égale au rang de M. Inversement, soit V une $\mathrm{Fr}\,\mathrm{W}(k_F)$-représentation semi-linéaire de $\mathrm{Gal}(F/E) \times \mathbb{G}_a$. On note N_V l'endomorphisme nilpotent de V associé à l'action de $\mathbb{G}_a$. On pose

$$M_F(V) = \left\{ x \in V \otimes_{\mathrm{Fr}\,\mathrm{W}(k_F)} \mathcal{R}(F)[\log T] \;\middle|\; \begin{array}{c} \forall g \in \mathrm{Gal}(F/E),\, g(x) = x, \\ (N_V \otimes \mathrm{Id} + \mathrm{Id} \otimes N)(x) = 0 \end{array} \right\}.$$

Proposition 4.6. *Sous les hypothèses ci-dessus, $M_F(V)$ est un $\mathcal{R}(E)$-module libre de rang égal à la dimension de V.*

Démonstration. Comme $\mathcal{R}(F) = \mathcal{E}^\dagger(F) \otimes_{\mathcal{E}^\dagger(E)} \mathcal{R}(E)$ et $\mathcal{E}^\dagger(E)$ est un corps on vérifie que

$$(V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F))^{\operatorname{Gal}(F/E)} = (V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{E}^\dagger(F))^{\operatorname{Gal}(F/E)} \otimes_{\mathcal{E}^\dagger(E)} \mathcal{R}(E).$$

Comme $\operatorname{Gal}(\mathcal{E}^\dagger(F)/\mathcal{E}^\dagger(E)) = \operatorname{Gal}(F/E)$, on en déduit par un raisonnement classique (cf. [20, Ch.X Prop.3]) que $(V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F))^{\operatorname{Gal}(F/E)}$ est un $\mathcal{R}(E)$-module libre de rang égal à la dimension de V. On définit un isomorphisme

$$f : (V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F))^{\operatorname{Gal}(F/E)} \to M_F(V),$$

en posant $f(\sum_l v_l \otimes \alpha_l) = \sum_l \sum_{i=0}^{r-1} (-)^i N_V^i(v_l) \otimes \alpha_l (\log T)^i$, où r est un entier tel que $N_V^r = 0$. L'application inverse $g : M_F(V) \to (V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F))^{\operatorname{Gal}(F/E)}$, est induite par la projection $\mathcal{R}(F)[\log T] \to \mathcal{R}(F)$ qui envoie $\log T$ en 0. En fait, par construction $gf = \mathrm{Id}$ et pour conclure il suffit de montrer que g est injective. Soit $x = \sum_l v_l \otimes (\sum_{i=0}^d \alpha_{i,l} (\log T)^i)$. Supposons que $g(x) = \sum_l v_l \otimes \alpha_{0,l} = 0$. La relation $(N_V \otimes \mathrm{Id} + \mathrm{Id} \otimes N)(x) = 0$ équivaut à

$$\forall i = 0, \dots, d, \quad \sum_l N_V(v_l) \otimes \alpha_{i,l} = -\sum_l v_l \otimes (i+1)\alpha_{i+1,l}. \qquad (\star)$$

Comme $\sum_l v_l \otimes \alpha_{0,l} = 0$, en appliquant $N_V \otimes \mathrm{Id} + \mathrm{Id} \otimes N$, on obtient $\sum_l N_V(v_l) \otimes \alpha_{0,l} = 0$. D'où par $(\star)$, $\sum_l v_l \otimes \alpha_{1,l} = 0$. Ainsi, par récurrence, on montre $x = 0$. □

On munit $M_F(V)$ de la connexion induite par $d : \mathcal{R}(F)[\log T] \to \hat{\Omega}^1_{\mathcal{R}(F)/\operatorname{Fr} \mathrm{W}(k_F)}$ et la connexion triviale sur V. Par construction, il est quasi-unipotent. Les inclusions naturelles

$$S_F(M) \subseteq M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \text{ et } M_F(V) \subseteq V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F)[\log T]$$

induisent par linéarisation des isomorphismes

$$S_F(M) \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F)[\log T] \to M \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T],$$
$$M_F(V) \otimes_{\mathcal{R}(E)} \mathcal{R}(F)[\log T] \to V \otimes_{\operatorname{Fr} \mathrm{W}(k_F)} \mathcal{R}(F)[\log T]$$

compatibles aux structures supplémentaires. On en déduit que le foncteur M_F est un quasi-inverse de S_F.

On introduit une variante de ces équivalences après extension des scalaires à $\overline{K}$. On rappelle que le produit $\mathcal{R}(E)_{\overline{K}} = \mathcal{R}(E) \otimes_{\operatorname{Fr} \mathrm{W}(k_E)} \overline{K}$ est intègre car $\operatorname{Fr} \mathrm{W}(k_E)$ est algébriquement fermé dans $\mathcal{R}(E)$. Pour une extension galoisienne $E/k((T))$, on étend linéairement l'action de l'inertie $I(E/k((T)))$ à $\mathcal{R}(E)_{\overline{K}}$. Si F/E est une extension galoisienne finie, alors $\mathcal{R}(F)_{\overline{K}}^{I(F/E)} = \mathcal{R}(E)_{\overline{K}}$. Tout module à connexion M, libre de type fini sur $\mathcal{R}(E)_{\overline{K}}$, provient par extension des scalaires, d'un module M', libre de type fini sur $\mathcal{R}(E) \otimes_{\mathrm{Fr}\,\mathrm{W}(k_E)} L$, où L est une extension finie de $\mathrm{Fr}\,\mathrm{W}(k_E)$ contenue dans $\overline{K}$. On peut donc parler d'irrégularité de M (cf. Rem.4.5-1 et 4.5-2). Soit M un module à connexion sur $\mathcal{R}(E)_{\overline{K}}$, unipotent sur $\mathcal{R}(F)_{\overline{K}}$. On pose

$$S'_F(M) = (M \otimes_{\mathcal{R}(E)_{\overline{K}}} \mathcal{R}(F)_{\overline{K}}[\log T])^{\nabla}.$$

C'est une représentation linéaire de $I(F/E) \times \mathbb{G}_a$. De façon analogue à ci-dessus, le foncteur S'_F est une équivalence de catégories entre modules à connexion sur $\mathcal{R}(E)_{\overline{K}}$, unipotents sur $\mathcal{R}(F)_{\overline{K}}$ et représentations linéaires sur $\overline{K}$ du schéma en groupes $I(F/E) \times \mathbb{G}_a$. Un quasi-inverse est donné par

$$M'_F(V) = \left\{ x \in V \otimes_{\overline{K}} \mathcal{R}(F)_{\overline{K}}[\log T] \;\middle|\; \begin{array}{c} \forall g \in I(F/E),\ g(x) = x, \\ (N_V \otimes \mathrm{Id} + \mathrm{Id} \otimes N)(x) = 0 \end{array} \right\}.$$

Pour tout $\mathcal{R}(E)$-module à connexion M, on considère son extension des scalaires $M \otimes_{\mathrm{Fr}\,\mathrm{W}(k_E)} \overline{K}$ comme module à connexion sur $\mathcal{R}(E)_{\overline{K}}$. Il est évident que $S'_F(M \otimes_{\mathrm{Fr}\,\mathrm{W}(k_E)} \overline{K}) = S_F(M) \otimes_{\mathrm{Fr}\,\mathrm{W}(k_F)} \overline{K}$, où on ne considère sur $S_F(M)$ que l'action du sous-groupe $I(F/E) \times \mathbb{G}_a$ de $\mathrm{Gal}(F/E) \times \mathbb{G}_a$.
La proposition suivante est une variante d'une proposition de Tsuzuki [23].

Proposition 4.7. [1, 7.1.2] *Soient M un $\mathcal{R}(E)$-module à connexion quasi-unipotent et F/E une extension galoisienne finie telle que M soit unipotent sur $\mathcal{R}(F)$. Alors l'irrégularité de M est égale au conducteur de Swan de la représentation $S_F(M)$ du groupe d'inertie $I(F/E)$.*

Démonstration. Comme le conducteur de Swan et l'irrégularité ne varient pas par extension des scalaires on peut tensoriser avec $\overline{K}$. L'équivalence S'_F induit un isomorphisme entre le groupe de Grothendieck des $\mathcal{R}(E)_{\overline{K}}$-modules à connexion, unipotents sur $\mathcal{R}(F)_{\overline{K}}$ et le groupe de Grothendieck des représentations de $I(E/F) \times \mathbb{G}_a$ sur $\overline{K}$. Dans ce dernier la classe d'une représentation V de $I(E/F) \times \mathbb{G}_a$ est égale à la classe de sa restriction à $I(F/E)$. Comme le conducteur de Swan et l'irrégularité se factorisent par le groupe de Grothendieck, il suffit de vérifier qu'ils se correspondent par S'_F. Dans cet isomorphisme l'induction d'une représentation correspond à l'oubli de structure pour le module à connexion correspondant. Le conducteur de Swan et l'irrégularité varient de la même façon par rapport à l'induction et à l'oubli respectivement. On peut, en appliquant le théorème d'induction de Brauer (cf. [21, §10]), se réduire au cas de dimension 1. Soit M un module à connexion sur $\mathcal{R}(E)_{\overline{K}}$ de rang 1, unipotent sur $\mathcal{R}(F)_{\overline{K}}$. On considère la représentation $S = S'_F(M)$ de $I = I(F/E)$. Soit Λ l'extension finie de $\mathbb{Q}_p$, obtenue en ajoutant les racines m-ièmes de l'unité, où m est l'exposant du groupe I. Par un autre théorème de Brauer (cf.[21, §12.2, Th.24]), il existe une représentation S_0 de I sur Λ telle que $S \cong S_0 \otimes_{\Lambda} \overline{K}$. On note $\Lambda' \subseteq \overline{K}$ l'extension non-ramifié de Λ de corps résiduel k_F. On rappelle que Tsuzuki [23] associe à S_0 un module à connexion $\mathcal{D}^\dagger(S_0)$ sur $\Lambda' \otimes_{\mathrm{Fr\,W}(k_E)} \mathcal{E}^\dagger(E)$, d'irrégularité égale au conducteur de Swan de S_0 (en fait le cas de dimension 1 est dû à Matsuda [16]). On vérifie que $\mathcal{D}^\dagger(S_0) = (S_0 \otimes_\Lambda (\Lambda' \otimes_{\mathrm{Fr\,W}(k_F)} \mathcal{E}^\dagger(F)))^I$, avec la connexion induite par celle de $\mathcal{E}^\dagger(F)$. Comme S est de rang 1, on a $M'_F(S) = (S \otimes_{\overline{K}} \mathcal{R}(F)_{\overline{K}})^I$. On termine par,

$$\begin{aligned} M'_F(S) &\cong \left(S_0 \otimes_\Lambda \overline{K} \otimes_{\mathrm{Fr\,W}(k_F)} \left(\mathcal{E}^\dagger(F) \otimes_{\mathcal{E}^\dagger(E)} \mathcal{R}(E)\right)\right)^I \\ &\cong (S_0 \otimes_\Lambda \Lambda' \otimes_{\mathrm{Fr\,W}(k_F)} \mathcal{E}^\dagger(F))^I \otimes_{(\Lambda' \otimes_{\mathrm{Fr\,W}(k_E)} \mathcal{E}^\dagger(E))} (\overline{K} \otimes_{\mathrm{Fr\,W}(k_E)} \mathcal{R}(E)) \\ &\cong \mathcal{D}^\dagger(S_0) \otimes_{(\Lambda' \otimes_{\mathrm{Fr\,W}(k_E)} \mathcal{E}^\dagger(E))} \mathcal{R}(E)_{\overline{K}}. \end{aligned}$$

□

5 Preuve du théorème 1.1 et corollaires

Soient V une représentation p-adique de G_K et L/K une extension galoisienne finie. On rappelle (cf. §3.2) que les modules $\mathrm{D}(V_{|\mathrm{G}_L})$, $\mathrm{D}^\dagger(V_{|\mathrm{G}_L})$ et $\mathrm{D}^\dagger_{\mathrm{rig}}(V_{|\mathrm{G}_L})$ sont munis d'une action naturelle de $\mathrm{Gal}(L_\infty/K)$. Pour tout $x \in \mathrm{D}^\dagger_{\mathrm{rig}}(V_{|\mathrm{G}_L})[1/t]$ et $g \in \mathrm{Gal}(L_\infty/K)$, $g(t^{-1}\nu(x)) = \chi(g)^{-1} t^{-1}\nu(g(x))$. Pour V de de Rham, on en déduit une action de $\mathrm{Gal}(L_\infty/K)$ sur $\mathrm{N}_{\mathrm{dR}}(V_{|\mathrm{G}_L})$. On munit les modules $\mathrm{D}(V)$, $\mathrm{D}^\dagger(V)$, $\mathrm{D}^\dagger_{\mathrm{rig}}(V)$ et $\mathrm{N}_{\mathrm{dR}}(V)$ de l'action de $\mathrm{Gal}(L_\infty/K)$ via le quotient Γ_K.

Lemme 5.1. *Soient V une représentation p-adique de G_K et L/K une extension galoisienne finie. On a des isomorphismes canoniques :*

- *i)* $\mathrm{D}(V_{|\mathrm{G}_L}) \cong \mathcal{E}_L \otimes_{\mathcal{E}_K} \mathrm{D}(V)$;
- *ii)* $\mathrm{D}^\dagger(V_{|\mathrm{G}_L}) \cong \mathcal{E}^\dagger_L \otimes_{\mathcal{E}^\dagger_K} \mathrm{D}^\dagger(V)$;
- *iii)* $\mathrm{D}^\dagger_{\mathrm{rig}}(V_{|\mathrm{G}_L}) \cong \mathcal{R}_L \otimes_{\mathcal{R}_K} \mathrm{D}^\dagger_{\mathrm{rig}}(V)$;
- *iv) pour V de de Rham,* $\mathrm{N}_{\mathrm{dR}}(V_{|\mathrm{G}_L}) \cong \mathcal{R}_L \otimes_{\mathcal{R}_K} \mathrm{N}_{\mathrm{dR}}(V)$.

Ces isomorphismes sont compatibles avec les actions de $\mathrm{Gal}(L_\infty/K)$.

Démonstration. i) On a une application $\mathcal{E}_K$-linéaire injective $\mathrm{D}(V) \to \mathrm{D}(V_{|\mathrm{G}_L})$, compatible à l'action de $\mathrm{Gal}(L_\infty/K)$. Comme $\dim_{\mathcal{E}_K} \mathrm{D}(V) = \dim_{\mathcal{E}_L} \mathrm{D}(V_{|\mathrm{G}_L}) = \dim_{\mathbb{Q}_p} V$, elle induit un isomorphisme $\mathcal{E}_L \otimes_{\mathcal{E}_K} \mathrm{D}(V) \cong \mathrm{D}(V_{|\mathrm{G}_L})$ équivariant pour l'action de $\mathrm{Gal}(L_\infty/K)$. ii) est une conséquence du Théorème 3.5. iii) se déduit directement de ii). iv) L'injection $\mathrm{N}_{\mathrm{dR}}(V) \subseteq \mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t]$ entraîne que $\mathcal{R}_L \otimes_{\mathcal{R}_K} \mathrm{N}_{\mathrm{dR}}(V) \subseteq \mathcal{R}_L \otimes_{\mathcal{R}_K} \mathrm{D}^\dagger_{\mathrm{rig}}(V)[1/t] = \mathrm{D}^\dagger_{\mathrm{rig}}(V_{|\mathrm{G}_L})[1/t]$. C'est un sous-$\mathcal{R}_L$-module stable par $t^{-1}\nu$, par l'unicité de $\mathrm{N}_{\mathrm{dR}}(V_{|\mathrm{G}_L})$ il est égal à $\mathrm{N}_{\mathrm{dR}}(V_{|\mathrm{G}_L})$. □

Soit V une représentation potentiellement semi-stable de G_K. Choisissons une extension galoisienne finie L/K telle que la restriction de V à G_L soit semi-stable. Soit L'/K_{a} l'extension finie non-ramifiée associée à L dans la section

§3.1 (au dessus de Lemme 3.2). On considère le L'-espace vectoriel des sections horizontales

$$S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V)) = (\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T])^{\nabla_V}.$$

L'action de $\mathrm{Gal}(L_\infty/K)$ sur $\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T]$ commute avec la connexion par la définition même de cette dernière (cf. §4.1). On en déduit une action semi-linéaire de $\mathrm{Gal}(L_\infty/K)$ sur $S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V))$. La restriction de cette action au sous-groupe $\mathrm{Gal}(L_\infty/K_\infty)$ correspond via l'identification $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K) \cong \mathrm{Gal}(L_\infty/K_\infty)$ à l'action décrite au §4.3.

Proposition 5.2. *Soient V une représentation galoisienne p-adique de G_K et L/K une extension galoisienne finie telle que la restriction de V à G_L soit semi-stable. Il existe un isomorphisme canonique*

$$\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T] \cong \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L}) \otimes_{L_{\mathrm{a}}} \mathcal{R}_L[\log T]. \tag{1}$$

Le groupe $\mathrm{Gal}(L_\infty/K)$ *agit sur le deux membres de* (1) *: à gauche comme décrit plus haut et à droite diagonalement, par son quotient* $\mathrm{Gal}(L/K)$ *sur* $\mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L})$ *et par son action naturelle sur* $\mathcal{R}_L[\log T]$*. L'isomorphisme* (1) *est équivariant pour cette action. Il est aussi horizontal où on considère à droite la connexion triviale sur* $\mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L})$*. De façon équivalente, en prenant les sections horizontales, on un isomorphisme canonique* $\mathrm{Gal}(L_\infty/K)$*-équivariant*

$$S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V)) \cong \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L}) \otimes_{L_{\mathrm{a}}} L'. \tag{2}$$

Par conséquent, l'action de $\mathrm{Gal}(L_\infty/K)$ *sur* $S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V))$ *se factorise par son quotient fini* $\mathrm{Gal}(L \otimes_{L_{\mathrm{a}}} L'/K)$*. Le conducteur de Swan de V est égal à* $\mathrm{sw}(I(L \otimes_{L_{\mathrm{a}}} L'/K), S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V)))$*.*

Démonstration. Par le corollaire 4.2, on a un isomorphisme

$$\mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L}) \otimes_{L_{\mathrm{a}}} L' \cong (\mathrm{N}_{\mathrm{dR}}(V_{|\mathrm{G}_L}) \otimes_{\mathcal{R}_L} \mathcal{R}_L[\log T])^{\nabla_V}.$$

Il est équivariant par rapport à l'action de $\mathrm{Gal}(L_\infty/K)$ car il est le composé de l'isomorphisme équivariant du théorème 3.6 avec une inclusion naturelle. Cette action se factorise évidemment par $\mathrm{Gal}(L \otimes_{L_{\mathrm{a}}} L'/K)$. Le lemme 5.1-iv) donne un isomorphisme, $\mathrm{Gal}(L_\infty/K)$-équivariant,

$$(\mathrm{N}_{\mathrm{dR}}(V_{|\mathrm{G}_L}) \otimes_{\mathcal{R}_L} \mathcal{R}_L[\log T])^{\nabla_V} \cong (\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T])^{\nabla_V} = S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V)).$$

Ceci montre (2). L'isomorphisme (1) suit en tensorisant (2) par $\mathcal{R}_L[\log T]$ au dessus de L' et en composant avec l'isomorphisme canonique $S_{\mathrm{E}_L}(\mathrm{N}_{\mathrm{dR}}(V)) \otimes_{L'} \mathcal{R}_L[\log T] \cong \mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T]$. □

Remarque 5.3. *Un résultat analogue à la proposition* 5.2 *a été démontré par N. Wach pour les représentations sur un corps absolument non-ramifié, qui sont de de Rham et de hauteur finie (cf.* [24, A5 et B1.4.2] *).*

Soit L/K une extension galoisienne finie quelconque contenue dans $\overline{K}$. Pour tout $n \in \mathbb{N}$, on note $G_n = \mathrm{Gal}(L_n/K_n)$ et I_n son sous-groupe d'inertie. Pour $n \in \mathbb{N}$, la restriction induit un monomorphisme de groupes $f_n : G_{n+1} \hookrightarrow G_n$. Soit $n(L)$ le plus petit entier tel que $K_{n(L)}$ contienne l'intersection de L avec K_∞. Pour $n \geq n(L)$, f_n est un isomorphisme et le groupe G_n est canoniquement isomorphe à $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$. Par abus on note encore $f_n : \mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K) \to G_n$ cet isomorphisme. Le lemme suivant est une reformulation d'un résultat classique de Sen (cf. [19, Lemma 1, pg. 40] et [25, 3.3.2]).

Lemme 5.4. *Pour n assez grand la filtration de ramification supérieure et inférieure de G_n est stationnaire et correspond via l'isomorphisme f_n à la filtration de ramification de* $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$.

Démonstration. On utilise ici une définition différente du corps de normes mais équivalente à celle donnée dans le paragraphe 3.1. On considère la limite projective d'ensembles $\varprojlim_{n\in\mathbb{N}} \mathcal{O}_{K_n}$, où les applications de transitions sont les normes. Cet ensemble est isomorphe à $\mathcal{O}_{\mathrm{E}_K}$ muni de la multiplication composante par composante et de la somme donnée par la formule suivante. Si $x = (x_n)_{n\in\mathbb{N}}$ et $y = (y_n)_{n\in\mathbb{N}}$ appartient à $\mathcal{O}_{\mathrm{E}_K}$, alors $(x+y)_n = \lim_{m\to+\infty} N_{K_m/K_n}(x_m + y_m)$. On note $G = \mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$. Soit $\pi \in \mathrm{E}_L$ une uniformisante. Comme $(L_n/K)_{n\geq 1}$ est cofinal dans l'ensemble des sous-extension finies de L_∞/K, on peut écrire $\pi = (\pi_n)_{n\geq 1}$ avec $\pi_n \in L_n$. Soit $n' \geq n(L)$ un entier tel que $L_\infty/L_{n'}$ soit totalement ramifiée. Pour tout $n \geq n'$, π_n est une uniformisante de L_n. Pour tout $g \in G$ et $n \geq n'$, on pose $i(g) = i_G(g) = \mathrm{v}_{\mathrm{E}_L}(g(\pi) - \pi)$ et $i_n(g) = i_{G_n}(f_n(g)) = \mathrm{v}_{L_n}(f_n(g)(\pi_n) - \pi_n)$. On doit montrer que $i(g) = \lim_{n\to+\infty} i_n(g)$. En fait,

$$\begin{aligned} i(g) = \mathrm{v}_{\mathrm{E}_L}(g(\pi)-\pi) = \mathrm{v}_{L_{n'}}((g(\pi)-\pi)_{n'}) &= \\ &= \mathrm{v}_{L_{n'}}\left(\lim_{n'\leq n\to+\infty} N_{L_n/L_{n'}}(g(\pi_n)-\pi_n)\right) = \\ &= \lim_{n'\leq n\to+\infty} \mathrm{v}_{L_{n'}}(N_{L_n/L_{n'}}(g(\pi_n)-\pi_n)) = \\ &= \lim_{n'\leq n\to+\infty} \mathrm{v}_{L_n}((g(\pi_n)-\pi_n)) = \lim_{n\to+\infty} i_n(g). \end{aligned}$$

□

On rappelle qu'on a noté $V_n = V_{|\mathrm{G}_{K_n}}$.

Lemme 5.5. *Soit V une représentation potentiellement semi-stable. La suite $(\mathrm{sw}(V_n))_{n\in\mathbb{N}}$ est stationnaire.*

Démonstration. Par définition $\mathrm{sw}(V_n) = \mathrm{sw}(\mathrm{D}_{\mathrm{pst}}(V_n)_{|I_{K_n}})$. On a un monomorphisme évident $\mathrm{D}_{\mathrm{pst}}(V_n) \hookrightarrow \mathrm{D}_{\mathrm{pst}}(V)_{|\mathrm{G}_{K_n}}$, qui est un isomorphisme car $\mathrm{D}_{\mathrm{pst}}(V_n)$ et $\mathrm{D}_{\mathrm{pst}}(V)$ ont la même dimension. Donc $\mathrm{D}_{\mathrm{pst}}(V_n)_{|I_{K_n}} \cong \mathrm{D}_{\mathrm{pst}}(V)_{|I_{K_n}}$ et l'action de I_{K_n} se factorise par I_n. Par le lemme 5.4, pour n assez grand, le conducteur $\mathrm{sw}(\mathrm{D}_{\mathrm{pst}}(V)_{|I_{K_n}})$ est constant. □

Soit V une représentation de G_K qui devient semi-stable sur L. Pour un entier n assez grand, on considère $\mathrm{D}_{\mathrm{pst}}(V_n) = K_{\mathrm{a}}^{\mathrm{nr}} \otimes_{L'} \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_{L_n}})$ comme $K_{\mathrm{a}}^{\mathrm{nr}}$-représentation semi-linéaire de $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$ via l'isomorphisme f_n. Cette représentation ne dépend pas de n. Sa restriction à l'inertie est linéaire. Dans la suite on la considère comme une représentation de $\mathrm{G}_{\mathrm{E}_K}$ et on la note $\mathrm{D}_{\mathrm{pst}}^{\infty}(V)$. Par le lemme 5.5,

$$\mathrm{sw}(\mathrm{D}_{\mathrm{pst}}^{\infty}(V)) = \lim_{n\to+\infty} \mathrm{sw}(V_n). \tag{3}$$

Démonstration du théorème 1.1. Par le théorème de monodromie p-adique la représentation V est potentiellement semi-stable. Soit L/K une extension galoisienne finie telle que $V_{|\mathrm{G}_L}$ soit semi-stable. On choisit un entier n assez grand de façon que $G_n = \mathrm{Gal}(L_n/K_n)$ soit isomorphe à $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$ avec leurs filtrations de ramifications. Donc $\mathrm{D}_{\mathrm{pst}}^{\infty}(V) = \mathrm{D}_{\mathrm{pst}}(V_n)$ et $\mathrm{sw}(\mathrm{D}_{\mathrm{pst}}^{\infty}(V)) = \mathrm{sw}(I(L_n/K_n), \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_{L_n}}))$. Par la proposition 5.2, on a un isomorphisme $S_{\mathrm{E}_{L_n}}(\mathrm{N}_{\mathrm{dR}}(V)) \cong \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_{L_n}})$, équivariant par rapport à l'action de $\mathrm{Gal}(L_n/K)$. D'où, par restriction, l'égalité de $\mathrm{sw}(I(L_n/K_n), \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_{L_n}}))$ et de $\mathrm{sw}(I(\mathrm{E}_L/\mathrm{E}_K), S_{\mathrm{E}_{L_n}}(\mathrm{N}_{\mathrm{dR}}(V)))$. Comme $\mathrm{E}_L = \mathrm{E}_{L_n}$, ce dernier est égal à l'irrégularité de $\mathrm{N}_{\mathrm{dR}}(V)$, par la proposition 4.7. $\square$

Lemme 5.6. *Soient V une représentation galoisienne p-adique de G_K, L/K une extension galoisienne finie telle que $V_{|\mathrm{G}_L}$ soit semi-stable et n' le plus petit entier tel que $K_{n'} \supseteq (L \otimes_{L_{\mathrm{a}}} L') \cap K_\infty$. Alors :*

i) Pour tout $n \geq n'$, $\mathrm{D}_{\mathrm{st}}(V_n) = \mathrm{D}_{\mathrm{st}}(V_{n'}) \cong (\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_K[\log T])^{\nabla_V}$.

ii) Pour tout $n \geq n'$, $\mathrm{D}_{\mathrm{cris}}(V_n) = \mathrm{D}_{\mathrm{cris}}(V_{n'}) \cong (\mathrm{N}_{\mathrm{dR}}(V))^{\nabla_V}$.

Démonstration. On pose $D = \mathrm{D}_{\mathrm{st}}(V_{|\mathrm{G}_L})$. i) On considère l'isomorphisme (1) (prop. 5.2), $D \otimes_{L_{\mathrm{a}}} \mathcal{R}_L[\log T] \cong \mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_L[\log T]$. En prenant les sections horizontales et les points fixès par $\mathrm{Gal}(L_\infty/K_\infty)$, on obtient $\mathrm{D}_{\mathrm{st}}(V_{n'}) = (D \otimes_{L_{\mathrm{a}}} L')^{\mathrm{Gal}(L_\infty/K_\infty)} = (\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_K[\log T])^{\nabla_V}$. Pour $m \geq n'$, de façon analogue, on obtient $\mathrm{D}_{\mathrm{st}}(V_m) = (\mathrm{N}_{\mathrm{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_{K_m}[\log T])^{\nabla_V}$. Comme $\mathrm{H}_{K_m} = \mathrm{H}_K$, on a $\mathcal{R}_K = \mathcal{R}_{K_m}$ en tant qu'anneaux. D'où $\mathrm{D}_{\mathrm{st}}(V_m) = \mathrm{D}_{\mathrm{st}}(V_{n'})$.
ii) Par la proposition 5.2,

$$\begin{aligned}
(\mathrm{N}_{\mathrm{dR}}(V))^{\nabla_V} &\cong (M_{\mathrm{E}_L}(D \otimes_{L_{\mathrm{a}}} L'))^{\nabla} \\
&= \left\{ x \in D \otimes_{L_{\mathrm{a}}} \mathcal{R}_L[\log T] \;\middle|\; \begin{array}{c} \forall g \in \mathrm{Gal}(L_\infty/K_\infty),\, g(x) = x, \\ (N_D \otimes \mathrm{Id} + \mathrm{Id} \otimes N)(x) = 0, \\ (\mathrm{Id} \otimes d)(x) = 0 \end{array} \right\} \\
&= \left\{ x \in D \otimes_{L_{\mathrm{a}}} L' \;\middle|\; \begin{array}{c} \forall g \in \mathrm{Gal}(L_\infty/K_\infty),\, g(x) = x, \\ (N_D \otimes \mathrm{Id})(x) = 0 \end{array} \right\} = \mathrm{D}_{\mathrm{cris}}(V_{n'}).
\end{aligned}$$

On conclut comme dans i). $\square$

Corollaire 5.7. *Soit V une représentation de de Rham de* G_K. *Alors :*

$$\lim_{n\to+\infty} \mathrm{ar}(V_n)$$
$$= \mathrm{irr}(\mathrm{N_{dR}}(V)) + \mathrm{rg}(\mathrm{N_{dR}}(V)) - \dim_{K'} \left(\mathrm{N_{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_K[\log T]\right)^{\nabla_V}$$

et

$$\lim_{n\to+\infty} \mathrm{ar_{cris}}(V_n) = \mathrm{irr}(\mathrm{N_{dR}}(V)) + \mathrm{rg}(\mathrm{N_{dR}}(V)) - \dim_{K'} \left(\mathrm{N_{dR}}(V)\right)^{\nabla_V}.$$

Démonstration. Soit L/K une extension galoisienne finie telle que $V_{|\mathrm{G}_L}$ soit semi-stable. On pose $D = \mathrm{D_{st}}(V_{|\mathrm{G}_L})$. On rappelle que $\dim_{L_\mathrm{a}} D^{I(L/K)} = \dim_{K_\mathrm{a}} D^{\mathrm{Gal}(L/K)}$ car $H^1(\mathrm{Gal}(k_L/k), \mathrm{GL}_n(L_\mathrm{a}))$ est triviale (cf. [20, Ch.X, Prop.3]). On a $\mathrm{ar}(V) - \mathrm{sw}(V) = \dim_{L_\mathrm{a}} D - \dim_{L_\mathrm{a}} D^{I(L/K)} = \dim_{L_\mathrm{a}} D - \dim_{K_\mathrm{a}} D^{\mathrm{Gal}(L/K)} = \dim_{\mathbb{Q}_p} V - \dim_{K_\mathrm{a}} \mathrm{D_{st}}(V) = \mathrm{rg}(\mathrm{N_{dR}}(V)) - \dim_{K_\mathrm{a}} \mathrm{D_{st}}(V)$. Par le théorème 1.1 et le lemme 5.6-i) on obtient la première formule. On en déduit facilement la deuxième en utilisant la définition de $\mathrm{ar_{cris}}(V)$ et 5.6-ii). □

Remarque 5.8. *Le corollaire* 5.7 *généralise un résultat de Berger (cf. Th.* 4.1 *et* [3, Th. 5.20] *) : l'équation* $\mathrm{N_{dR}}(V)$ *est unipotente (resp. triviale) si et seulement si V est semi-stable (resp. cristalline) sur K_n, pour n assez grand. En effet V est semi-stable (resp. cristalline) sur K_n si et seulement si* $\mathrm{ar}(V_n) = 0$ *(resp.* $\mathrm{ar_{cris}}(V_n) = 0$*). L'équation* $\mathrm{N_{dR}}(V)$ *est unipotente (resp. triviale) si et seulement si* $\mathrm{rg}(\mathrm{N_{dR}}(V)) = \dim_{K'} \left(\mathrm{N_{dR}}(V) \otimes_{\mathcal{R}_K} \mathcal{R}_K[\log T]\right)^{\nabla_V}$ *(resp.* $\mathrm{rg}(\mathrm{N_{dR}}(V)) = \dim_{K'} \left(\mathrm{N_{dR}}(V)\right)^{\nabla_V}$*) et dans ce cas on a aussi* $\mathrm{irr}(\mathrm{N_{dR}}(V)) = 0$.

Soit M une équation différentielle p-adique ayant une structure de Frobenius. Dans [7], Christol et Mebkhout associent à M une décomposition en somme directe indexée par les rationnels, la décomposition par les pentes p-adiques. C'est un théorème profond de la théorie que la hauteur du polygone de Newton associé à cette décomposition est égale à l'indice $\widetilde{\chi}(M, \mathcal{A})$.

Corollaire 5.9. *Soient V une représentation galoisienne p-adique de* G_K *et L/K une extension galoisienne finie telle que la restriction de V à* G_L *soit semi-stable.*

i) On a un isomorphisme $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$*-équivariant,*

$$\mathrm{D}^\infty_{\mathrm{pst}}(V) \cong K^{\mathrm{nr}}_\mathrm{a} \otimes_{L'} S_{\mathrm{E}_L}(\mathrm{N_{dR}}(V)).$$

ii) Sous l'isomorphisme du i), la décomposition de $K^{\mathrm{nr}}_\mathrm{a} \otimes_{L'} S_{\mathrm{E}_L}(\mathrm{N_{dR}}(V))$ *induite par les pentes p-adiques de* $\mathrm{N_{dR}}(V)$*, correspond à la décomposition par les pentes de Swan de* $\mathrm{D}^\infty_{\mathrm{pst}}(V)$ *(cf.* [14, Ch. 1] *). Par conséquent, on a l'égalité des polygones de Newton associés.*

Démonstration. On déduit l'assertion i) par restriction de l'isomorphisme (2) (prop.5.2) à $\mathrm{Gal}(L_\infty/K_\infty)$. Pour ii), soit $\mathrm{N_{dR}}(V) = \oplus_{q\in\mathbb{Q}} \mathrm{N_{dR}}(V)_q$ la décomposition par les pentes p-adiques. On pose $D_q = K^{\mathrm{nr}}_\mathrm{a} \otimes_{L'} S_{\mathrm{E}_L}(\mathrm{N_{dR}}(V)_q)$.

C'est un facteur directe de $\mathrm{D}^{\infty}_{\mathrm{pst}}(V)$ de dimension égal au rang de $\mathrm{N}_{\mathrm{dR}}(V)_q$. C'est stable par φ et N. Par la proposition 4.7, on a $\mathrm{sw}(D_q) = \mathrm{irr}(\mathrm{N}_{\mathrm{dR}}(V)_q) = q \operatorname{rg} \mathrm{N}_{\mathrm{dR}}(V)_q = q \dim_{K^{\mathrm{nr}}_{\mathrm{a}}} D_q$. Pour conclure il suffit de montrer que toute pente de Swan de D_q est égale à q. Soient s une pente de Swan de D_q et $D_{q,s} \neq 0$ sa partie isopentique de pente s. Par [14, Lemma 1.8] , $D_{q,s}$ est stable par $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$. Comme φ et N commutent avec $\mathrm{Gal}(\mathrm{E}_L/\mathrm{E}_K)$, l'unique pente possible pour $\varphi(D_{q,s})$ et $N(D_{q,s})$ est s. Comme $D_{q,s}$ est maximal de pente s, on a $\varphi(D_{q,s}) \subseteq D_{q,s}$ et $N(D_{q,s}) \subseteq D_{q,s}$. On en déduit que $M_{\mathrm{E}_L}(D_{q,s})$ est un sous-module différentiel de $\mathrm{N}_{\mathrm{dR}}(V)_q$ muni d'un Frobenius. Donc $M_{\mathrm{E}_L}(D_{q,s})$ a comme seule pente q. Par 4.7,

$$s \dim_{K^{\mathrm{nr}}_{\mathrm{a}}} D_{q,s} = \mathrm{sw}(D_{q,s}) = \mathrm{irr}(M_{\mathrm{E}_L}(D_{q,s})) = q \operatorname{rg} M_{\mathrm{E}_L}(D_{q,s}) = q \dim_{K^{\mathrm{nr}}_{\mathrm{a}}} D_{q,s}.$$

D'où $s = q$. □

Références

[1] André, Y. : Filtrations de type Hasse-Arf et monodromie p-adique, *Invent. Math.* 148 (2002), 285–317.

[2] André, Y. : Représentations galoisiennes et operateurs de Bessel p-adiques, *Ann. Inst. Fourier* 52(3) (2002), 779–808.

[3] Berger, L. : Représentations p-adiques et équations différentielles, *Invent. Math.* 148 (2002), 219–284.

[4] Berger, L. : An introduction to the theory of p-adic representations, *Geometric Aspects of Dwork's Theory*, Volume I, de Gruyter Proc. in Math., Berlin (2004), 255-292.

[5] Cherbonnier, F. et Colmez, P. : Représentations p-adiques surconvergentes, *Invent. Math.* 133 (1998), 581–611.

[6] Colmez, P. : Conducteur d'Artin d'une représentation de de Rham, Prépublication (2003), 1–25.

[7] Christol, G. et Mebkhout, Z. : Sur le théorème de l'indice des équations différentielles p-adiques IV, *Invent. Math.* 143 (2001), 629–672.

[8] Christol, G. et Mebkhout, Z. : Équations différentielles p-adiqueset coefficients p-adiques sur les courbes, *Astérisque* 279 (2002), 125–183.

[9] Dieudonné, J. et Grothendieck, A. : EGA, *Publ. Math. IHES.*

[10] Fontaine, J.-M. : Modules galoisiens, modules filtré et anneaux de Barsotti-Tate, Dans : *Journées de Géométrie algébriques de Rennes (III),* Astérisque 65 (1979), 3–80.

[11] Fontaine, J.-M. : Représentations p-adiques des corps locaux, Dans : *The Grothendieck Festschrift*, Volume II, Prog. Math., Birkhäuser, Basel (1990), 249–309.

[12] Fontaine, J.-M. : Le corps de périodes p-adiques, avec une appendice par Pierre Colmez, Dans : *Périodes p-adiques,* Astérisque 223 (1994), 59–111.

[13] Fontaine, J.-M. : Représentations p-adiques semi-stables, Dans : *Périodes p-adiques,* Astérisque 223 (1994), 113–184.

[14] Katz, N. : Gauss sums, Kloosterman sums and monodromy groups, *Annals of mathematics studies* 116, Priceton Univerity press, Princeton, New Jersey (1988).

[15] Kedlaya, S.K. : A p-adic local monodromy theorem, *arXiv* : math.AG/0110124, Apparaîtra dans *Ann. of Math.*.

[16] Matsuda, S. : Local indeces of p-adic differential operators corresponding to Artin-Schreier-Witt coverings, *Duke Mathematical Journal* 77(3) (1995), 607–625.

[17] Matsuda, S. : Katz Correspondence for quasi-Unipotent Overconvergent Isocrystals, *Compositio Math.* 134 (2002), 1–34.

[18] Mebkhout, Z. : Analogue p-adique du théorème de Turrittin et le théorème de la monodromie p-adique, *Invent. Math.* 148 (2002), 319–351.

[19] Sen, S. : On automorphisms of local fields, *Ann. of Math.* 90 (1969), 33–46.

[20] Serre, J.-P. : *Corps Locaux,* deuxième édition, Hermann, Paris (1968).

[21] Serre, J.-P. : *Représentations linéaires de groupes finis,* Hermann, Paris (1967).

[22] Tsuzuki, N. : Finite local monodromy of overconvergent unit-root F-isocrystals on a curve, *Amer. J. Math.* 120 (1998), 1165–1190.

[23] Tsuzuki, N. : The local index and Swan conductor, *Compositio Math.* 111 (1998), 245–288.

[24] Wach, N. : Représentations p-adiques potentiellement cristallines, *Bull. Soc. Math. France* 124 (1996), 375–400.

[25] Wintenberger, J.-P. : Le corps des normes de certaines extensions infinies de corps locaux ; applications, *Ann. Sci. École Norm. Sup. (4)*, 16 (1983), 59–89.

Adriano Marmora
LAGA, Institut Galilée
Université Paris 13
93430 Villetaneuse (France)
marmora@math.univ-paris13.fr

On Connecting Orbits of Semilinear Parabolic Equations on S^1

Yasuhito Miyamoto

Received: August 3, 2003
Revised: June 30, 2004

Communicated by Bernold Fiedler

Abstract. It is well-known that any bounded orbit of semilinear parabolic equations of the form

$$u_t = u_{xx} + f(u, u_x), \quad x \in S^1 = \mathbb{R}/\mathbb{Z}, \quad t > 0,$$

converges to steady states or rotating waves (non-constant solutions of the form $U(x - ct)$) under suitable conditions on f. Let S be the set of steady states and rotating waves (up to shift). Introducing new concepts — the *clusters* and the *structure* of S —, we clarify, to a large extent, the heteroclinic connections within S; that is, we study which $u \in S$ and $v \in S$ are connected heteroclinically and which are not, under various conditions. We also show that $\sharp S \geq N + \sum_{j=1}^{N} [[\sqrt{(f_u(r_j, 0))_+}/(2\pi)]]$ where $\{r_j\}_{j=1}^{N}$ is the set of the roots of $f(\,\cdot\,, 0)$ and $[[y]]$ denotes the largest integer that is strictly smaller than y. In paticular, if the above equality holds or if f depends only on u, the *structure* of S completely determines the heteroclinic connections.

2000 Mathematics Subject Classification: 35B41, 34C29
Keywords and Phrases: global attractor, heteroclinic orbit, zero number, semilinear parabolic equation.

Contents

1. Introduction

We will investigate the global dynamics of semilinear parabolic partial differential equations on $S^1 = \mathbb{R}/\mathbb{Z}$ in $X = C^1(S^1)$

$$\begin{cases} u_t = u_{xx} + f(u, u_x), & x \in S^1, \\ u(x,0) = u_0(x), & x \in S^1. \end{cases} \tag{1.1}$$

The above problem is equivalent to a problem on the interval $[0,1]$ under the periodic boundary conditions $u(0,t) = u(1,t)$, $u_x(0,t) = u_x(1,t)$ for $t > 0$. Under suitable conditions on f, the solutions of (1.1) exist globally in $t > 0$. Thus (1.1) defines a global semiflow Φ_t on X. We will call each solution of (1.1) an *orbit*.

Angenent and Fiedler [AF88] and Matano [Ma88] have shown independently that any solution of (1.1) approaches as $t \to \infty$ to a solution (or a family of solutions) of the form $U(x - ct)$, where c is some real constant. Since $U(x - ct)$ is a solution to (1.1), the function $U(\zeta)$ should satisfy the following equation:

$$\frac{d^2U}{d\zeta^2} + c\frac{dU}{d\zeta} + f\left(U, \frac{dU}{d\zeta}\right) = 0, \quad \zeta \in S^1, \tag{1.2}$$

where $\zeta = x - ct$. Note that $U(\zeta + \theta)$ is a solution to (1.2) for all $\theta \in S^1$ provided that $U(\zeta)$ is a solution. If $c \neq 0$ and if $U(\zeta)$ is not a constant function, then $U(x - ct)$ is a time periodic solution called a *rotating wave* with speed c. If $c = 0$ and if $U(\zeta)$ is not a constant function, then $U(x)$ is called a *standing wave*. Thus steady states consist of both standing waves and constant steady states. By using these terms, the above assertion can be restated that any solution of (1.1) approaches either rotating waves or steady states.

Under suitable conditions on f that will be specified later, (1.1) has the set $\mathcal{A} \subset X$ called the *global attractor*. This set $\mathcal{A}$ is characterized as the maximal compact invariant set and it attracts all the orbits of (1.1).

Matano and Nakamura [MN97] have shown that the global attractor $\mathcal{A}$ of (1.1) consists of rotating waves, standing waves and *connecting orbits* that connect these waves. Therefore, in order to understand the dynamical structure of $\mathcal{A}$ it is important to know which pairs of waves are connected heteroclinically and which pairs are not. The paper [AF88] proves the existence of some connecting orbits for the problem (1.1) by using a topological method. We are interested in finding out a sharper criterion for the existence of connecting orbits.

In this paper we will give a precise lower bound for the number of mutually distinct rotating waves and steady states (Corollary B). If the Morse index of every wave is odd or zero, then certain order relations among waves defined below and the Morse index of all the waves determine which pairs of waves are connected heteroclinically and which pairs are not (Theorem A). In particular, if the actual number of the waves coincides with the lower bound given in Corollary B, then the hypothesis of Theorem A is automatically fulfilled, hence the heteroclinic connections are completely determined (Theorem C). In the special case where f depends only on u, we can completely determine which pairs of waves are connected heteroclinically and which are not (Theorems A and A'), and we will present rather simple and explicit sufficient conditions on f for the hypotheses of Theorem C to be satisfied (Proposition D).

Theorems A and C and Proposition D are proved by using the concepts of *clusters* and the *structure* which we introduce in Section 2. Let S be the set of all the waves. Roughly speaking, a *cluster* is a subset of S consisting of waves sharing certain common features, and S is expressed as a disjoint union of *clusters*. One can show that each *cluster* is a totally ordered set with respect to the following order relation

$$u \rhd v \quad \overset{def}{\Longleftrightarrow} \quad R(u) \supset R(v),$$

where $R(u)$ denotes the range of u (see Definition 2.5 and Remark 2.6). We then define the *structure* of S by associating each *cluster* with the sequence of (modified) Morse indices of its elements. Lemmas E, F and F' give fundamental properties of this sequence of modified Morse indices.

Now, many authors study the global attractor of (1.1) for the case where the boundary conditions in (1.1) is replaced by the Dirichlet or the Neumann boundary conditions. We can see [BF89] for the Dirichlet boundary conditions, [FR96] and [Wo02] for the Neumann boundary conditions and [MN97] for periodic boundary conditions. Here we recall the results of [FR96]. In the case of the Neumann boundary conditions on $[0, 1]$, the global attractor consists of the steady states and the connecting orbits between these steady states, if all the steady states are hyperbolic. Let $\{U_j(x)\}_{j=1}^n$ ($U_1(0) < U_2(0) < \cdots < U_n(0)$) be the set of all the steady states. Roughly speaking, the permutation that rearranges the sequence $(U_1(1), U_2(1), \ldots, U_n(1))$ in increasing order determines the Morse indices of all the steady states and the zero number of functions $U_j(x) - U_k(x)$ ($1 \le j < k \le n$) (In brief, the zero number of a function, which is defined in Section 2, is the number of the roots of the function). Once these Morse indices and the zero number of the difference of all the pairs among the waves are obtained, then this information tells which steady states are connected and which are not. Wolfrum [Wo02] has simplified the conditions of whether steady states are connected heteroclinically or not using the concept of *k-adjacent*. The concept of *k-adjacent* also uses the zero number of functions $U_j(x) - U_k(x)$ and the value of one of end points $U_j(0)$ (or $U_j(1)$). In the case of the periodic boundary conditions, we cannot use the method of [FR96] because the end points do not exist on S^1, therefore the Morse indices and the zero number of the difference of the pairs cannot be characterized in terms of permutation. Instead the maximum value, the minimum value and the mode of the waves play an important role in determining the Morse index of the waves and the zero number of the difference of the pairs, thereby giving the global picture of their heteroclinic connection.

This paper is organized as follows: In Section 2 we introduce some notation and definitions and state our main results (Theorems A A' and C, Corollary B, Proposition D and Lemmas E, F and F'). Roughly speaking, Corollary B gives a lower bound for the number of the waves in terms of the derivatives of f, and Lemma F is concerned with the modified Morse indices of waves and the structure of *clusters*. Theorems A, A' and C and Proposition D determine the heteroclinic connections among waves under various conditions. In Section 3 we will prove Lemma 3.1 which is the key lemma of this paper. In Section 4 we will show that each *cluster* is a totally ordered set in our order relation. We state the main results of [AF88]. We will prove Theorem C by using the results. In Section 5 we will investigate a sequence of modified Morse indices of waves in each *cluster* and prove Lemmas E and F and Corollary B. In Section 6 we will prove Theorem A, using Lemma F and main results of [AF88]. In Section 7 we consider the case where f depends only on u. We will prove Theorem A' and Lemma F'. In Section 8 we prove Proposition D, which is a special case of Theorems A and A'. We will give rather simple and explicit sufficient conditions on f under which all the *clusters* are *monotone* and *simple*, the meaning of which will be defined in Section 2. The *monotonicity* and *simplicity* of *clusters* automatically determine the Morse index of all the waves and the zero number of the difference of the pairs among the waves, hence their heteroclinic connections.

ACKNOWLEDGMENT. The author would like to thank Professor H. Matano for his valuable comments and many fruitful discussions, and thank the referee for his/her useful suggestions. He would also like to express his gratitude to Professor B. Fiedler, whose early work has given the author much inspiration.

2. NOTATION AND MAIN THEOREMS

In this paper the nonlinear term f satisfies the following assumptions:

(A1) $f: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a C^3-function.

(A2) There exists a constant $L_1 > 0$ such that $u \cdot f(u, 0) < 0$ for $|u| > L_1$, and the function $f(\,\cdot\,, 0)$ has finitely many real roots.

(A3) (i) For any solution $u(x,t)$ to (1.1), $||u(\,\cdot\,,t)||_{C^1(S^1)} := ||u(\,\cdot\,,t)||_{C^0(S^1)} + ||u_x(\,\cdot\,,t)||_{C^0(S^1)}$ remains bounded as $t \to \infty$.

(ii) There exists a constant $L_2 > 0$ such that

$$||U(\zeta)||_{C^1(\mathbb{R})} := ||U(\zeta)||_{C^0(\mathbb{R})} + ||U_\zeta(\zeta)||_{C^0(\mathbb{R})} < L_2$$

for any periodic solution or constant solution $U(\zeta)$ to the following equation:

$$\frac{d^2U}{d\zeta^2} + c\frac{dU}{d\zeta} + f\left(U, \frac{dU}{d\zeta}\right) = 0, \quad \zeta \in \mathbb{R}, \tag{2.1}$$

where c is an arbitrary real number.

The assumption (A3) (ii) will be needed in Section 3, where we study the bifucation structure of rotating waves and constant steady states. The assumption (A3) is satisfied if the following condition (A3)$'$ holds:

(A3)$'$ For any constant $M_1 > 0$, there exists a constant $L_3 > 0$ such that $f_u(u,p) \le 0$ for $|u| < M_1$ and $|p| > L_3$.

From (A1), (A2) and (A3) it follows that (1.1) defines a global semiflow Φ_t on X that is dissipative. Here a semiflow Φ_t on X is called *dissipative* if there exists a ball $B \subset X$ which satisfies the following: For any $u_0 \in X$, there exists $t_0 > 0$ such that $\Phi_t(u_0) \in B$ for all $t \ge t_0$ (see [Ma76]).

Hereafter, we assume (A1)+(A2)+(A3)$'$ throughout the present paper.

By the standard parabolic estimates, the mapping Φ_t is a compact mapping for every $t > 0$. This, together with the dissipativity of Φ_t, implies that there is the (nonempty) maximal compact invariant set $\mathcal{A} \subset X$. It is well-known from the general theory of dissipative dynamical systems that $\mathcal{A}$ is connected and attracts all the orbits of (1.1). This set $\mathcal{A}$ is called the global attractor. The Hausdorff dimension of $\mathcal{A}$ of (1.1) is $2\,[M/2]+1$ where M is the maximal generalized Morse index of the steady states or the rotating waves (see [MN97]).

Let us introduce some definitions and notation. In this paper we denote by S the set of steady states and rotating waves of (1.1). Note that if $U(x-ct)$ is a rotating wave (or a steady state in the case where $c=0$), then $U(x-ct+\theta)$ is also a rotating wave (or a steady state) for any $\theta \in S^1$. Hereafter we identify $U(\,\cdot\,)$ and $U(\,\cdot\,+\theta)$. In other words, we will understand S to be the set of equivalence classes, each of which is expressed in the form

$$\Gamma(U) := \{U(x-ct+\theta)|\ \theta \in S^1\},$$

where $U(\zeta)$ is a solution of (1.2). However in order to simplify notation, we write $U(x-ct) \in S$ to mean $[U(x-ct)] \in S$, where $[U(x-ct)]$ denotes the equivalence class to which $U(x-ct)$ belongs. Therefore $u(x,t) \in S$ shall mean that $u(x,t) = U(x-ct+\theta)$ for some $\theta \in S^1$ where $U(\zeta)$ is a solution to (1.2). Furthermore, by a heteroclinic connection from $u(x,t)(:= U(x-ct)) \in S$ to $v(x,t)(:= V(x-\tilde{c}t)) \in S$ we mean that there is an orbit $w(x,t)$ of (1.1) such that

$$\inf_{\theta_1 \in S^1} \|w(x,t) - U(x-ct+\theta_1)\|_{L^\infty(S^1)} \to 0 \ (t \to -\infty),$$
$$\inf_{\theta_2 \in S^1} \|w(x,t) - V(x-\tilde{c}t+\theta_2)\|_{L^\infty(S^1)} \to 0 \ (t \to +\infty).$$

In particular, if U and V are 'hyperbolic' (whose meaning is defined below in this section), then a heteroclinic connection from u to v automatically implies the following stronger convergence:

$$\|w(x,t) - U(x-ct+\theta_1)\|_{L^\infty(S^1)} \to 0 \ (t \to -\infty) \text{ for some } \theta_1 \in S^1,$$
$$\|w(x,t) - V(x-\tilde{c}t+\theta_2)\|_{L^\infty(S^1)} \to 0 \ (t \to +\infty) \text{ for some } \theta_2 \in S^1.$$

The number of the roots of $f(\,\cdot\,,0)$ is finite owing to (A2). Let $\{r_j\}_{j=1}^N$ ($r_1 < r_2 < \cdots < r_N$) be the roots of $f(\,\cdot\,,0)$ throughout the present paper. All the constant steady states are $u(x,t) = r_j$ ($j \in \{1,2,\ldots,N\}$).

Remark 2.1. If $f_u(r_j,0) \neq 0$ for all $j \in \{1,2,\ldots,N\}$, then N is odd because of (A2). Moreover $u(x,t) = r_j$ ($j \in \{1,3,5,\ldots,N\}$) is a stable constant steady state, while $u(x,t) = r_j$ ($j \in \{2,4,6,\ldots,N-1\}$) is an unstable constant steady state (see Remark 2.8 below).

The zero number is a powerful tool to analyze nonlinear single reaction-diffusion equations in one space dimension:

$$z(w) := \sharp \left\{x \mid w(x) = 0,\ x \in S^1\right\} \quad \text{for } w \in X,$$

where $\sharp Y$ denotes the number of elements of the set Y. It is well-known that $z(w(\,\cdot\,,t))$ is a non-increasing function of t if w is a solution of a one-dimensional linear parabolic equation (see [Ma82], [Ni62] and [St36]). Furthermore, the following proposition holds:

PROPOSITION 2.2 (Angenent and Fiedler [AF88] and Angenent [An88]). *Let $a(x,t)$ and $b(x,t)$ be C^2-functions in $(x,t) \in S^1 \times (0,\tau)$ ($\tau > 0$). Let $w(x,t) \in X$ be a solution to the following equations:*

$$w_t = w_{xx} + a(x,t)w_x + b(x,t)w, \qquad (x,t) \in S^1 \times (0,\tau).$$

Then $z(w(\,\cdot\,,t))$ is finite for every $t \in (0,\tau)$ and is non-increasing in t. Moreover $z(w(\,\cdot\,,t))$ drops at each $t = t_0$ when the function $x \longmapsto w(x,t_0)$ has a multiple zero.

Remark 2.3. Angenent and Fiedler [AF88] have proved Proposition 2.2 in the case where $a(x,t)$ and $b(x,t)$ are real analytic functions. Angenent [An88] has relaxed this analyticity assumption.

Using the moving frame with speed c, we can rewrite (1.1) as follows:

$$u_t = u_{\zeta\zeta} + cu_\zeta + f(u,u_\zeta), \tag{2.2}$$

where $\zeta = x - ct$. Let $U(x-ct) \in S$. The wave $U(\zeta)(= U(x-ct))$ is a steady state of (2.2). In order to analyze the stability of $U(\zeta)$, we define the linearized operator of (2.2) at $U(\zeta)$ by

$$L_U w = w_{\zeta\zeta} + cw_\zeta + f_u(U,U_\zeta)w + f_p(U,U_\zeta)w_\zeta, \quad \zeta \in S^1,$$

provided that U is a non-constant steady state of (2.2). Here f_p denotes the derivative of f with respect to the second variable. If U is a constant steady state of (2.2), then we define the linearized operator by

$$L_U w = w_{\zeta\zeta} + f_u(U,0)w + f_p(U,0)w_\zeta, \quad \zeta \in S^1.$$

By the standard spectral theory for ordinary differential operators of the second order, the spectrum of L_U consists of eigenvalues of finite multiplicity and has no accumulation point except ∞. Let $\{\lambda_n\}_{n=0}^{\infty}$ be the eigenvalues of L_U that are repeated according to their algebraic multiplicity. We define the Morse index of $U \in S$ by $i(U) := \sharp\{\lambda_n | \ \mathrm{Re}(\lambda_n) > 0\}$. By a Sturm-Liouville type theorem (see [AF88] and [MN97]), we have

$$\mathrm{Re}(\lambda_0) > \mathrm{Re}(\lambda_1) \geq \mathrm{Re}(\lambda_2) > \mathrm{Re}(\lambda_3) \geq \cdots \geq \mathrm{Re}(\lambda_{2j}) > \mathrm{Re}(\lambda_{2j+1}) \geq \cdots .$$

Moreover if U is a non-constant steady state, we can see

$$i(U) \in \{z(U_\zeta),\ z(U_\zeta) - 1\} \tag{2.3}$$

(see [AF88] and [MN97]). Note that $z(U_\zeta)$ is even and $z(U_\zeta)-1$ is odd since U_ζ is a periodic function of ζ. We can see the Morse index of the constant steady states by easy calculations (see Remark 2.8 below).

Next we define the hyperbolicity of $U \in S$. Because of translation equivariance of the equation (1.1), each rotating wave and each non-constant steady state form a one-dimensional manifold that is homeomorphic to S^1. This equivariance has to be taken into account when we define the hyperbolicity of those solutions.

Definition 2.4.

- (i) *Let u be a (non-constant) rotating wave ($c \neq 0$) or a non-constant steady state ($c = 0$). We say u is hyperbolic if 0 is the only eigenvalue of L_u on the imaginary axis and if 0 is a simple eigenvalue.*
- (ii) *Let u be a constant steady state (i.e. $u(x,t) = r_j$). We say u is hyperbolic if there is no eigenvalue of L_u on the imaginary axis.*

Definition 2.5. *Let $u(x,t)$ be a solution of (1.1). We define*

$$R(u(\,\cdot\,,t)) := \left\{ y \in \mathbb{R} \;\middle|\; \min_{x\in S^1} u(x,t) \leq y \leq \max_{x\in S^1} u(x,t) \right\}.$$

Remark 2.6. If $u \in S$, then $R(u(\,\cdot\,,t))$ is independent of t. Hereafter we simply write $R(u)$ if $u \in S$.

Definition 2.7. *For $u \in S$, we define its "modified Morse index" by*

$$I(u) := \begin{cases} z(u_x) & \textit{if } u \textit{ is not a constant steady state;} \\ i(u) + 1 & \textit{if } u \textit{ is an unstable constant steady state;} \\ 0 & \textit{if } u \textit{ is a stable constant steady state.} \end{cases}$$

Remark 2.8. One can calculate the Morse index of the constant steady states. Let u be a constant steady state (*i.e.* $u(x) = r_j$). Then

$$i(u) = \begin{cases} 2\left[\dfrac{\sqrt{f_u(r_j,0)}}{2\pi}\right] + 1 & \text{if } f_u(r_j,0) > 0; \\ 0 & \text{if } f_u(r_j,0) \leq 0, \end{cases}$$

where $[y]$ denotes the largest integer not exceeding y. If all the constant steady states are hyperbolic, then $i(u) = 2[\sqrt{f_u(r_j,0)}/(2\pi)]+1$ for $j \in \{2,4,6,\ldots,N-1\}$ and $i(u) = 0$ for $j \in \{1,3,5,\ldots,N\}$. Thus r_j ($j \in \{1,3,5,\ldots,N\}$) is stable and r_j ($j \in \{2,4,6,\ldots,N-1\}$) is unstable.

Note that $I(u)$ is always a non-negative even integer. From (2.3) it follows that

$$i(u) \le I(u) \le i(u) + 1.$$

Therefore $I(u)$ is a good approximation of the real Morse index $i(u)$. Clearly, $I(u) = i(u)$ if and only if $i(u)$ is even.

While the modified Morse index $I(u)$ is easily computable from Definition 2.7 and Remark 2.8, the real Morse index $i(u)$ is not always easily to determine. This is the reason why we introduce the notion modified Morse index.

Now we can define the *cluster*.

Definition 2.9. *Let $1 \le k \le l \le N$. We define the clusters by*

$$C_{kl} := \{u \in S|\ S_{kl} \subset R(u),\ (\{r_1, r_2, \ldots, r_N\} \backslash S_{kl}) \cap R(u) = \emptyset\},$$

where $S_{kl} := \{r_k, r_{k+1}, \ldots, r_l\}$.

It is not difficult to see that

$$C_{kl} \cap C_{k'l'} = \emptyset \quad \text{if } (k,l) \ne (k',l'),$$

$$S = \bigcup_{1 \le k \le l \le N} C_{kl}.$$

Furthermore one can see that, if k or l is odd, then

$$C_{kk} = \{r_k\} \text{ and } C_{kl} = \emptyset \ (k \ne l).$$

The concept of *clusters* will be useful in the phase plane analysis as we will see in Section 6.

Definition 2.10. *Let C_{kl} be a cluster. We define*

$$R(C_{kl}) := \bigcup_{u \in C_{kl}} R(u).$$

Definition 2.11. *Let $u, v \in S$. We define the order relation of S as follows:*

$$u \rhd v \quad \overset{def}{\Longleftrightarrow} \quad R(u) \supset R(v).$$

Let $u, v, w \in S$. If $u \rhd v$, then we say v is smaller than u in the order $\rhd$, and u is bigger than v in the order $\rhd$. If there is no w such that $u \rhd w \rhd v$, then we say that u is the smallest wave in the order $\rhd$ that satisfies $u \rhd v$.

We have either $R(u) \supset R(v)$ or $R(v) \supset R(u)$ provided that $R(u) \cap R(v) \ne \emptyset$. This will be shown in Corollary 4.2 in Section 4. Consequently we have either $u \rhd v$ or $v \rhd u$ if $u, v \in C_{kl}$. Thus C_{kl} is a totally ordered set. Hereafter, we number the elements of each $C_{kl} = \{u_1^{kl}, u_2^{kl}, \ldots, u_{m_{kl}}^{kl}\}$ (with $m_{kl} := \sharp C_{kl}$) in such a way that

$$u_1^{kl} \lhd u_2^{kl} \lhd \cdots \lhd u_{m_{kl}}^{kl}.$$

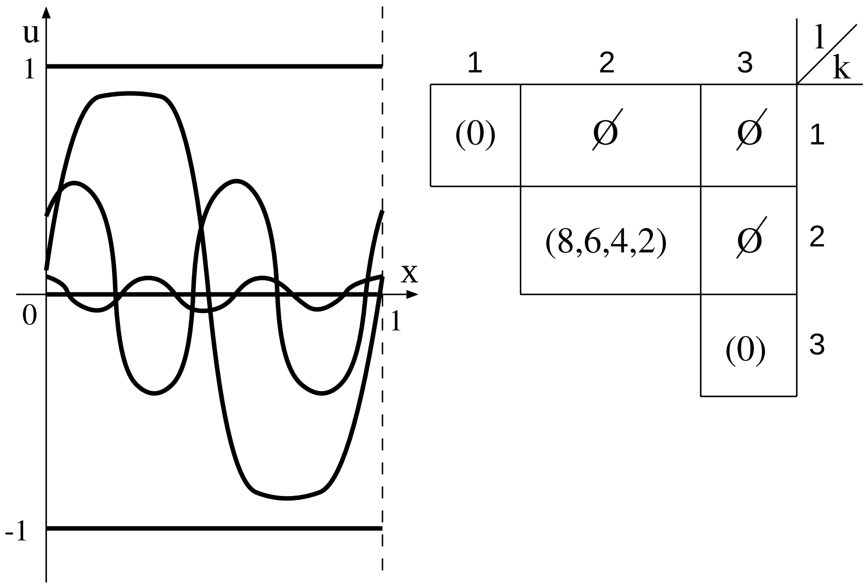


Figure 1. Wave profiles (left) and the *structure* of S (right) for equation (2.4). The three horizontal lines indicate constant steady states.

We call C_{kl} a *monotone cluster* if $I(u_1^{kl}) > I(u_2^{kl}) > \cdots > I(u_{m_{kl}}^{kl})$. The *cluster* C_{kk} is called a *simple cluster*. We call C_{kk} a *trivial cluster* provided that $\sharp C_{kk} = 1$. Note that C_{kk} always contains the constant steady state r_k, but it may contain other elements under certain circumstances.

Next we define an order relation among *clusters* in S.

Definition 2.12. *Let $C_{k_1 l_1}, C_{k_2 l_2}$ be clusters. We define the order relation $\rhd$ as follows:*

$$C_{k_1 l_1} \rhd C_{k_2 l_2} \quad \overset{def}{\Longleftrightarrow} \quad k_1 \le k_2 \text{ and } l_1 \ge l_2.$$

Let $C_{k_1 l_1}, C_{k_2 l_2}$ be *clusters*. If $C_{k_1 l_1} \rhd C_{k_2 l_2}$, then we say $C_{k_2 l_2}$ is smaller than $C_{k_1 l_1}$ in the order $\rhd$.

We define the *structure* of S.

Definition 2.13. *Let $C_{kl} := \left\{u_1^{kl}, u_2^{kl}, \ldots, u_{m_{kl}}^{kl}\right\}$ (with $m_{kl} := \sharp C_{kl}$) be a cluster. We call*

$$J_{kl} := (I(u_1^{kl}), I(u_2^{kl}), \ldots, I(u_{m_{kl}}^{kl}))$$

the sequence of modified Morse indices. We call

$$(J_{kl})_{1 \le k \le l \le N}$$

the structure of S.

Example 2.14. Let us investigate the *structure* of the waves of the following equation:

$$u_t = u_{xx} + 500(u - u^3), \quad x \in S^1. \tag{2.4}$$

Clearly, there are three constant steady states. Let $r_1 = 1$, $r_2 = 0$ and $r_3 = -1$. The nonlinear term depends only on u. Thus all the waves are standing waves (see Remark 2.18). A simple calculation reveals that the nonlinear term satisfies the hypothesis of Proposition D below. Thus we can see that all the clusters are *simple* and *monotone*, using Proposition D. Since all the *clusters* are *simple*, there are precisely three *clusters*: C_{11}, C_{22} and C_{33} (where $r_1 \in C_{11}$, $r_2 \in C_{22}$ and $r_3 \in C_{33}$). We can see that C_{11} and C_{33} are *trivial clusters*, using (i) of Lemma F. Furthermore r_1 and r_3 are stable (see Remark 2.1) and $I(r_1) = I(r_3) = 0$ (see Definition 2.7 and Remark 2.8). The *cluster* C_{22} is *monotone.* Thus Theorem C below tells us that the derivative of the nonlinear term at $u = r_2$ gives $\sharp C_{22} = 4$, because

$$3 < \frac{\sqrt{\frac{d}{du}\left\{500(u-u^3)\right\}\big|_{u=r_2}}}{2\pi} < 4.$$

Therefore C_{22} has three non-constant standing waves and one constant steady state. The profile of the waves are as shown in Figure 1. We denote by u_1^{22} the constant steady state in C_{22} and by u_2^{22}, u_3^{22} and u_4^{22} the non-constant standing waves. We can assume that $u_1^{22} \lhd u_2^{22} \lhd u_3^{22} \lhd u_4^{22}$, because all the *clusters* are totally ordered sets. Since C_{22} is *monotone*, we can see by (ii) and (v) of Lemma F that $I(u_1^{22}) = 8$, $I(u_2^{22}) = 6$, $I(u_3^{22}) = 4$ and $I(u_4^{22}) = 2$. Therefore the *structure* of S is as shown in the table in Figure 1.

We introduce some more notation to state main theorems. Let $u \in S$ and let $C(u)$ be the *cluster* containing u. Define

$$u_+ := \inf\{w \mid w > u,\ w \text{ is a constant steady state}\},$$
$$u_- := \sup\{w \mid w < u,\ w \text{ is a constant steady state}\},$$

and for each integer $n \geq 0$, define u_n to be the smallest wave in the order $\rhd$ that satisfies the following: $I(u_n) = 2n$, $u_n \rhd u$, and $u_n \in C(u)$. That is,

$$u_n = \min_{\rhd}\{v \in C(u) \mid v \rhd u,\ I(v) = 2n\}.$$

Lemma F below tells us that such u_n exists for $n \in \{1, 2, \ldots, [i(u)/2]\}$.

Roughly speaking u_+ is the constant steady state that is just above u in the usual order, and u_- is the constant steady state that is just below u in the usual order.

THEOREM A. *Suppose that all the elements of S are hyperbolic. Then*

(i) *If the wave u is not a stable constant steady state, then u connects to u_+, u_- and u_n for all $n \in \{1, 2, \ldots, I(u)/2 - 1\}$.*

(ii) *Furthermore if $i(u)$ is odd, then u does not connect to any other waves. Therefore the structure of S determines completely which $u \in S$ and $v \in S$ are connected and which are not, if the Morse index of every wave is odd or zero.*

Remark 2.15. The statement (i) of Theorem A is obtained by Angenent-Fiedler [AF88] (see Proposition 6.3 of the present paper).

Theorem A'. *Suppose that f is dependent only on u, say $f = g(u)$, and that all the waves are hyperbolic. Let u be a wave whose Morse index $i(u)$ is even. Then u connects only to u_+, u_-, u_n ($n \in \{1, 2, \cdots, I(u)/2\}$), and every $v \in S$ that satisfies the following: $v \lhd u$, $I(v) \leq I(u)$, and there is no wave w such that $u \rhd w \rhd v$, $I(u) = I(w)$, and $u \neq w \neq v$.*

Remark 2.16. The *structure* of S tells us the modified Morse index of every wave. In the case where f depends only on u, we can know the (real) Morse index of every wave by using Lemmas F and F' stated below. Thus we see by Theorems A and A' that the heteroclinic connections are determined by the *structure* of S provided that f depends only on u.

Corollary B.

$$\sharp S \geq N + \sum_{j=1}^{N} \left[\left[\frac{\sqrt{(f_u(r_j, 0))_+}}{2\pi}\right]\right],$$

where $[[y]]$ denotes the largest integer that is strictly smaller than y (i.e. $[[y]] = -[-y] - 1$) and $(y)_+ := \max\{y, 0\}$.

Remark 2.17. The hyperbolicity of the solutions is not assumed in Corollary B.

Theorem C. *Suppose that all $u \in S$ are hyperbolic. Then the following two conditions are equivalent:*

(a)

$$\sharp S = N + \sum_{j=1}^{N} \left[\left[\frac{\sqrt{(f_u(r_j, 0))_+}}{2\pi}\right]\right], \tag{2.5}$$

where $(y)_+ := \max\{y, 0\}$.

(b) *all the clusters are simple and monotone.*

Moreover, under these conditions, $i(u) = I(u) - 1 = (z(u_x) - 1)$ is odd for any non-constant $u \in S$. Thus the hypotheses of Theorem A *are satisfied. The conclusions of* Theorem A *hold. Specifically the structure of S is uniquely determined by the sequence $[[\sqrt{(f_u(r_j, 0))_+}/(2\pi)]]$ ($j = 1, 2, \ldots, N$). The global picture of heteroclinic connections in S is also uniquely determined as shown in* Figure 9.

In the case where f is dependent only on u, say $f = g(u)$, we introduce other two assumptions (A4) and $(A5)_j$ below. Let

$$G(u) = \int_0^u g(r)dr. \tag{2.6}$$

(A4) There exists an odd constant k such that $G(r_1) \le G(r_3) \le \cdots \le G(r_k) \ge G(r_{k+2}) \ge \cdots \ge G(r_N)$, $G(r_2) \le G(r_4) \le \cdots \le G(r_{k-1})$ and $G(r_{k+1}) \ge G(r_{k+3}) \ge \cdots \ge G(r_{N-1})$.

If $k = 1$ or $k = n$, then the second or the third inequalities in (A4) are not assumed respectively. We will see in Section 8 that C_{kl} ($k \ne l$) is empty provided that (A4) holds. Thus every *cluster* is *simple* (see Figures 11 and 12). We impose the other assumption: For $j \in \{2, 4, 6, \ldots, N-1\}$,

$(\mathbf{A5})_j$ $g(u)/|u|$ is decreasing for $u \in (r_{j-1}, r_j) \cup (r_j, r_{j+1})$.

The condition $(\mathrm{A5})_j$ guarantees that C_{jj} is *monotone* (Lemma 8.1). Hence we obtain the following:

Proposition D. *Suppose that f is dependent only on u, say $f = g(u)$, and that all the waves are hyperbolic. If* (A4) *holds and if* $(\mathrm{A5})_j$ *holds for all even $j \in \{2, 4, 6, \ldots, N-1\}$, then the hypotheses of* Theorem C *are satisfied. Thus the conclusions of* Theorems A, A' *and* C *hold.*

Remark 2.18. The equation (1.1) does not have rotating waves in the case where the nonlinear term f depends only on u. For the details, see the beginning of Section 7.

The next lemma is concerned with the structure of each *cluster*.

Lemma E (Cluster lemma 1). *Suppose that all $u \in S$ are hyperbolic. Let $1 \le k \le l \le N$. Let $C_{kl} = \{u^{kl}_1, u^{kl}_2, \ldots, u^{kl}_{m_{kl}}\}$ ($m_{kl} = \sharp C_{kl}$) be a cluster and let $J_{kl} = (I(u^{kl}_1), I(u^{kl}_2), \ldots, I(u^{kl}_{m_{kl}}))$ be the corresponding sequence of modified Morse indices. Then the following hold:*

- (i) *If k or l is odd and if $k \ne l$, then $C_{kl} = \emptyset$.*
- (ii) *If k is odd, then $\sharp J_{kk} = 1$. Thus C_{kk} is a trivial cluster. Moreover $I(u^{kk}_1) = 0$.*

Lemma F (Cluster lemma 2). *Under the same hypotheses of Lemma E, the following hold:*

- (i) *Every $I(u)$ is an even integer, and $I(u^{kl}_n) - I(u^{kl}_{n+1})$ is equal to -2, 0 or 2 for all $n \in \{1, 2, \ldots, m_{kl} - 1\}$.*
- (ii) *If $I(u^{kl}_{n_1-1}) < I(u^{kl}_{n_1}) = \cdots = I(u^{kl}_{n_2}) < I(u^{kl}_{n_2+1})$ $(2 \le n_1 \le n_2 \le m_{kl} - 1)$ or if $I(u^{kl}_{n_1-1}) > I(u^{kl}_{n_1}) = \cdots = I(u^{kl}_{n_2}) > I(u^{kl}_{n_2+1})$ $(2 \le n_1 \le n_2 \le m_{kl} - 1)$, then $n_2 - n_1$ is even.*
- (iii) *If $I(u^{kl}_{n_1-1}) < I(u^{kl}_{n_1}) = \cdots = I(u^{kl}_{n_2}) > I(u^{kl}_{n_2+1})$ $(2 \le n_1 \le n_2 \le m_{kl} - 1)$ or if $I(u^{kl}_{n_1-1}) > I(u^{kl}_{n_1}) = \cdots = I(u^{kl}_{n_2}) < I(u^{kl}_{n_2+1})$ $(2 \le n_1 \le n_2 \le m_{kl} - 1)$, then $n_2 - n_1$ is odd.*
- (iv) *If C_{kl} is not trivial, that is, if $\sharp J_{kl} \ge 2$, then $I(u^{kl}_{m_{kl}}) = 2$.*
- (v) *If $k \ne l$, and if $C_{kl} \ne \emptyset$, then $I(u^{kl}_1) = 2$.*

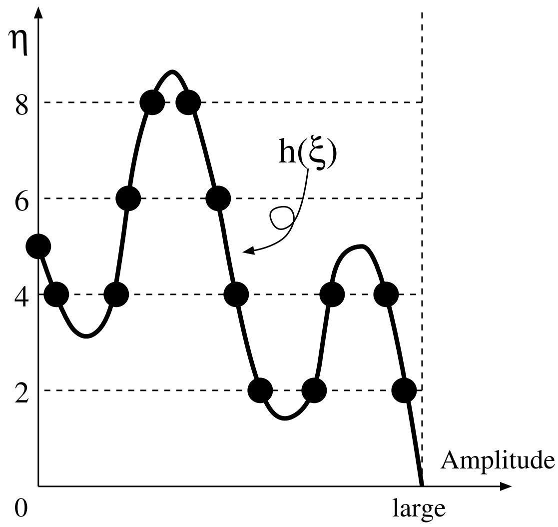


FIGURE 2. An example of $h(\xi)$ mentioned in Remark 2.19. In this case, the sequence of modified Morse indices is $(6, 4, 4, 6, 8, 8, 6, 4, 2, 2, 4, 4, 2)$.

Remark 2.19. In view of Lemma F, the sequence of modified Morse indices $J_{kk} = \big(I(u_j^{kk})\big)_{j=1}^{m_{kk}}$ may better be illustrated as the intersection points between the graph of a function $\eta = h(\xi)$ where $1/h(\xi)$ is the time-map and the horizontal lines $\eta = 2, 4, 6, 8, \cdots$ (see Figure 2). The time-map is used in Section 3 (see the definition of $T(a)$ in the statement of Lemma 3.1). This function h satisfies that $h'(\xi) \neq 0$ whenever $h(\xi)$ is an even integer, and that $h(\xi) = 0$ if ξ is large.

LEMMA F' *Suppose that f is dependent only on u, say $f = g(u)$, and that all the waves are hyperbolic. Let $\{u_{b_1}^{kl}, u_{b_2}^{kl}, \dots, u_{b_n}^{kl}\}$ $(b_1 < b_2 < \cdots < b_n)$ be the non-constant waves in a cluster C_{kl} whose modified Morse indices are the same number (i.e. $I(u_{b_1}^{kl}) = I(u_{b_2}^{kl}) = \cdots = I(u_{b_n}^{kl})$). Then $i(u_{b_{n-2j}}^{kl}) = I(u_{b_{n-2j}}^{kl}) - 1$ $(j \in \{0, 1, \dots, [(n-1)/2]\})$ and $i(u_{b_{n-2j-1}}^{kl}) = I(u_{b_{n-2j-1}}^{kl})$ $(j \in \{0, 1, \dots, [(n-2)/2]\})$.*

3. PROOF OF THE KEY LEMMA

We will also prove three lemmas which are used in the proof of main theorems. One of these lemmas (Lemma 3.1) is the key to the present paper.

In this section we assume that all the waves are hyperbolic in order to simplify notation. The number of all the constant steady states N is odd owing to the hyperbolicity and (A2) (see Remark 2.1). Using (A2), we can see that the

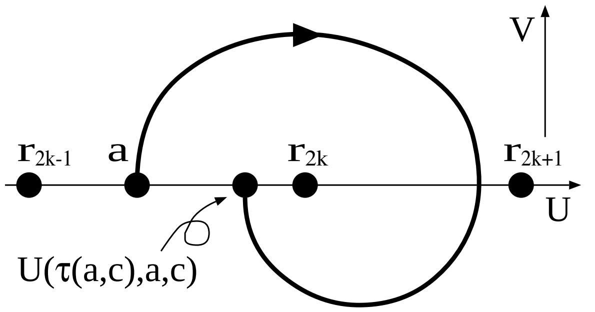


FIGURE 3. The picture denotes the arc that starts from the point $(a, 0)$ at the time 0 and arrives at a point $(b, 0)$ $(r_{2k-1} < b < r_{2k})$ at a certain positive time. The U-coordinate of the arrival point is denoted by $U(\tau(a, c), a, c)$ whose meaning is specified below. The arc deforms with respect to a and c. If we select suitable a and c, then the arrival point coincides with the starting point (*i.e.* $U(\tau(a, c), a, c) = a$), which means the arc is a closed orbit.

following hold:

$$f_u(r_{2k}, 0) < 0 \quad \text{if } k \in \{1, 2, \ldots, [N/2]\};$$
$$f_u(r_{2k-1}, 0) > 0 \quad \text{if } k \in \{1, 2, \ldots, [N/2] + 1\}.$$

Let us introduce some notation. Let $u(x, t) = U(\zeta) \in S$. The wave $U(\zeta)$ should satisfy the following equation and periodic boundary conditions:

$$\begin{cases} U_{\zeta\zeta} + cU_\zeta + f(U, U_\zeta) = 0, & \zeta \in (0, 1), \\ U(0) = U(1), \ U_\zeta(0) = U_\zeta(1). \end{cases} \tag{3.1}$$

We transform the equation of (3.1) into the normal form:

$$\begin{cases} \dfrac{dU}{d\zeta} = V \\ \dfrac{dV}{d\zeta} = -cV - f(U, V). \end{cases} \tag{3.2}$$

Let U-axis and V-axis be the horizontal and vertical axes of the phase plane respectively. First, we note that no closed orbit appears near the points $(r_{2k-1}, 0)$ $(k \in \{1, 2, \ldots, [N/2] + 1\})$, since there points are saddle points. In what follows we will construct closed orbits in a neighborhood of the points $(r_{2k}, 0)$ $(k \in \{1, 2, \ldots, [N/2]\})$ on the phase plane.

In order to explain our idea suppose that there is an orbit as shown in Figure 3. This orbit starts from the point $(a, 0)$, passes the segment $(r_{2k}, r_{2k+1}) \times \{0\}$, and arrives at a point on the segement $(r_{2k-1}, r_{2k}) \times \{0\}$.

Let $(b, 0)$ be the arrival point. As we will see in the proof of Lemma 3.1, the value of b depends continuously on a and c as far as the orbit remains within the band domain $r_{2k-1} < U < r_{2k+1}$. Hereafter by *the arc corresponding to* (a, c) we shall mean the portion of the orbit of (3.2) starting at $(a, 0)$ and ending at a point on the segment $(r_{2k-1}, r_{2k}) \times \{0\}$ as shown in Figure 3.

Let $\tau(a, c)$ be the arrival time of this arc; that is $\tau(a, c)$ is the smallest positive time τ such that $U(\tau(a, c), a, c) \in (r_{2k-1}, r_{2k})$ and $U_\zeta(\tau(a, c), a, c) = 0$ where $U(\zeta, a, c)$ denotes the solution of (3.2) with initial data $U(0) = a$, $V(0) = 0$, and U_ζ denotes the derivative of U with respect to the first variable. Clearly the arc forms a closed orbit of (3.2) if and only if

$$a = U(\tau(a, c), a, c). \tag{3.3}$$

Furthermore this closed orbit represents a solution of (3.1) if and only if

$$\tau(a, c) = \frac{1}{n}$$

for some $n \in \{1, 2, \ldots\}$.

The following lemma shows that there is a continuous family of closed orbits corresponding to varying choice of a and c.

Lemma 3.1. *For each r_{2k} $(k = 1, 2, \ldots, [N/2])$, there exists a constant $\underline{a}$ with $r_{2k-1} \leq \underline{a} < r_{2k}$ and a function $c = c(a) \in C^1((\underline{a}, r_{2k}))$ such that the following hold.*

(i) *For each $a \in (\underline{a}, r_{2k})$, the relation* (3.3) *holds if and only if $c = c(a)$.*

(ii) *Let $T(a)$ be the period of the closed orbit obtained in* (i)*, that is, $T(a) = \tau(a, c(a))$. Then*

$$\lim_{a \to \underline{a}} T(a) = \infty, \quad \lim_{a \to r_{2k}} T(a) = \frac{2\pi}{\sqrt{f_u(r_{2k}, 0)}}.$$

Proof. We begin with the outline of the proof. The proof consists of three steps. In Step 1 we will show by using the bifurcation theory that there exists a family of closed orbits of (3.2) near the point $(r_{2k}, 0)$. Thus $c(a)$ can be defined near $a = r_{2k}$. In Step 2 we will show that whenever (a_0, c_0) satisfies (3.3), a C^1-function $c(a)$ can be defined in a neighborhood of a_0 such that $c(a_0) = c_0$. We will use the implicit function theorem to show that. In Step 3 we will expand the domain of the function $c(a)$. We will define the infimum $\underline{a}$ such that $c(a)$ can be defined on the interval $(\underline{a}, r_{2k})$. We will prove $\lim_{a \to r_{2k}} T(a) = 2\pi/\sqrt{f_u(r_{2k}, 0)}$ where $T(a)$ is the period of the closed orbit corresponding to $(a, c(a))$. We will also prove $\lim_{a \to \underline{a}} T(a) = \infty$.

Step 1: We linearize (3.2) at the point $(r_{2k}, 0)$:

$$\frac{d}{d\zeta}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -f_u(r_{2k}, 0) & -c - f_p(r_{2k}, 0) \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix}, \tag{3.4}$$

where f_u and f_p indicate derivatives of f with respect to the first and the second variable respectively. Let $\nu_\pm$ be the eigenvalues of the above matrix. Then we have

$$\mathrm{Re}(\nu_\pm) = -\frac{c + f_p(r_{2k}, 0)}{2}, \ \mathrm{Im}(\nu_\pm) = \pm\sqrt{-f_u(r_{2k}, 0) + \left(\frac{c + f_p(r_{2k}, 0)}{2}\right)^2}.$$

We regard c as a parameter. If $c = -f_p(r_{2k}, 0)$, then the matrix is non-singular, has the pair of simple pure imaginary eigenvalues $\pm i\mu$ ($\mu > 0$), and has no eigenvalue of the form $\pm ik\mu$ ($k \in \mathbb{N},\ k \neq 1$). Moreover we can easily see that

$$\left.\frac{d\mathrm{Re}(\nu_\pm)}{dc}\right|_{c=-f_p(r_{2k},0)} = -\frac{1}{2} < 0.$$

Therefore, (see for example Theorem 2.6 of [AP93] (Section 7, page 144)) a Hopf bifurcation occurs at $c = -f_p(r_{2k}, 0)$. Thus there are closed orbits encircling the point $(r_{2k}, 0)$ on the phase plane that have any small amplitude.

Step 2: From Step 1, we assume that there is a closed orbit corresponding to (a_0, c_0) on the phase plane. The continuity of the arc with respect to a and c guarantees that there is a constant $\varepsilon > 0$ such that the arc corresponding to (a, c) exists as shown in Figure 3 provided that $|a - a_0| < \varepsilon$ and $|c - c_0| < \varepsilon$. Since the solution $U(\zeta)$ to (3.2) with initial data $U(0) = a$, $U_\zeta(0) = 0$ depends on a and c continuously, we write $U = U(\zeta, a, c)$. Let $F(\,\cdot\,,\,\cdot\,)$ be a function as follows:

$$F(a, c) := U(\tau(a, c), a, c) - a, \tag{3.5}$$

where $\tau(a, c)$ which is defined in the first part of Section 3 is the arrival time of *the arc corresponding to* (a, c). From (3.3), the arc corresponding to (a, c) is a closed orbit if and only if $F(a, c) = 0$. We will prove that there exists a C^1-function $c(a)$ in a neighborhood of a_0 that satisfies $F(a, c(a)) = 0$. First we see by the assumption that $F(a_0, c_0) = 0$. Second we see that $U(\zeta, a, c)$ is a C^2-function of ζ, a and c by the general theory of ordinary differential equations. Using the equation

$$U_\zeta(\tau(a + \Delta a, c), a + \Delta a, c) - U_\zeta(\tau(a, c), a, c) = 0,$$

where Δa is a small number and the definition of the derivative, we can show that $\tau(a, c)$ is of class C^1. Thus $F(a, c)$ is of class C^1. Third we will show that $F_c(a_0, c_0) \neq 0$ where

$$F_c(a, c) = U_\zeta(\tau(a, c), a, c)\tau_c(a, c) + U_c(\tau(a, c), a, c).$$

Since $U_\zeta(\tau(a_0, c_0), a_0, c_0) = 0$, we obtain

$$F_c(a_0, c_0) = U_c(\tau(a_0, c_0), a_0, c_0).$$

We will prove in Lemma 3.2 below that

$$U_c(\tau(a_0, c_0), a_0, c_0) \neq 0. \tag{3.6}$$

Now we assume that Lemma 3.2 holds. Then the implicit function theorem says that there is a C^1-function $c(a)$ that satisfies $F(a, c(a)) = 0$ for $a \in (a_0 - \tilde{\varepsilon}, a_0 + \tilde{\varepsilon})$ where $\tilde{\varepsilon}$ (> 0) is so small that $|c_0 - c(a)| < \varepsilon$ and $|a_0 - a| < \varepsilon$ for $a \in (a_0 - \tilde{\varepsilon}, a_0 + \tilde{\varepsilon})$.

We will see in Lemma 3.2 that $U(\tau(a, c), a, c)$ is non-decreasing in c and (3.6) holds. Thus $U(\tau(a, c), a, c)$ is increasing in c. For each fixed a, if there exists c satisfying (3.3), then c is uniquely determined. The function $c(a)$ is

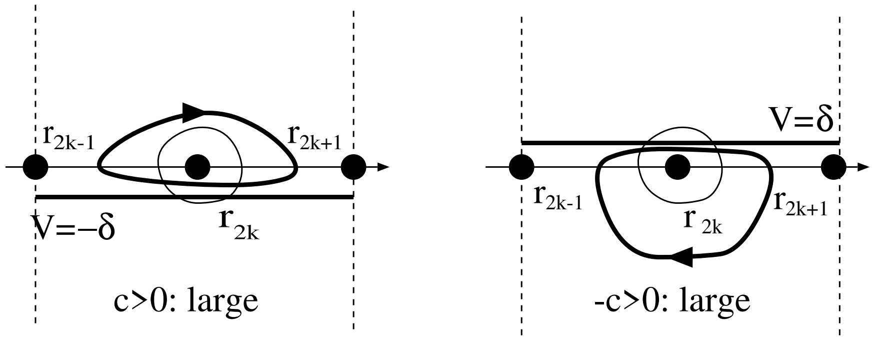


Figure 4. The phase planes of (3.2) for two extreme cases. Circles in each picture are the closed orbit corresponding to (a_0, c_0). If c is large, then the arc corresponding to (a, c) $(a < a_0)$ cannot pass the segment $(r_{2k-1}, r_{2k+1}) \times \{-\delta\}$ and the arrival point is in the inside of the closed orbit (the left picture). If $-c$ is large, then the arc corresponding to (a, c) does not pass the segment $(r_{2k}, r_{2k+1}) \times \{0\}$ (the right picture).

uniquely determined. This means that there is no closed orbit corresponding to (a_0, c_1) $(c_1 \neq c_0)$ when there is a closed orbit corresponding to (a_0, c_0).

Step 3: Hereafter we suppose that there exists the closed orbit corresponding to (a_0, c_0). We define $\underline{a}$ as follows:

$$\underline{a} := \inf\{a \in \mathbb{R} |\ c = c(\xi) \text{ can be defined for all } \xi \in (a, a_0)\}.$$

Note that there is a closed orbit corresponding to $(a, c(a))$ for all $a \in (\underline{a}, a_0)$. We will show by contradiction that the family of closed orbit corresponding to $(a, c(a))$ $(a \in (\underline{a}, a_0))$ is not uniformly away from two points $(r_{2k-1}, 0)$ and $(r_{2k+1}, 0)$. We assume that the family is uniformly away from two points.

We will show that there exists a constant $c^* > 0$ such that the following holds: if $|c| > c^*$, then a closed orbit starting from the point $(a, 0)$ $(a < a_0)$ does not exists.

For any $\delta > 0$, there is a constant c (> 0) such that $-cV - f(U, V) > 0$ on the segment $(r_{2k-1}, r_{2k+1}) \times \{-\delta\}$. The segment should intersect the closed orbit corresponding to (a_0, c_0) provided that δ is small. If there is a closed orbit corresponding to (a, c) $(a < a_0)$, then it should intersect the other closed orbit and this contradicts to Lemma 4.1. Similarly, if $-c$ (> 0) is large, then there should not exist closed orbits corresponding to (a, c) $(a < a_0)$.

If the closed orbit corresponding to (a, c) $(a < a_0)$ exists, then $c = c(a)$ is bounded.

Let $\{a_m\}_{m=1}^{\infty}$ be a sequence that satisfies the following:

$$a_m > \underline{a}, \quad a_m \to \underline{a} \ \text{ as } \ m \to \infty.$$

Since $c(a)$ is bounded, then there exists a constant c_* such that the following holds:

$$c(a_m) \to c_* \quad \text{as} \quad m \to \infty.$$

We consider the arc corresponding to $(\underline{a}, c_*)$. Let $(U(\zeta), U_\zeta(\zeta))$ be a closed orbit with period T_1. Then $U(\zeta)$ satisfies (2.1). From Lemma 3.3 below, there is a constant $M > 0$ such that $||U_\zeta(\zeta)|| \le M$. Thus any closed orbit is bounded on the phase plane.

Because of the continuity of arcs with respect to a and c, the boundedness of arcs, and the assumption that the family of closed orbit is uniformly away from the two points, the arrival point of the arc corresponding to $(\underline{a}, c_*)$ exists. Thus $U(\tau(\underline{a}, c_*), \underline{a}, c_*)$ can be defined. Using the continuity of $U(\tau(a, c), a, c)$ with respect to a and c, we can obtain a contradiction if we assume that $U(\tau(\underline{a}, c_*), \underline{a}, c_*) \ne \underline{a}$. Thus we see that

$$U(\tau(\underline{a}, c_*), \underline{a}, c_*) = \underline{a}.$$

This implies that there exists a closed orbit that contains $(\underline{a}, 0)$ on the phase plane. This is a contradiction because of the definition of $\underline{a}$ and Step 2. Thus the family is not uniformly away from the two points $(r_{2k-1}, 0)$ and $(r_{2k+1}, 0)$. This means that $\underline{a} = r_{2k-1}$ or the shortest distance of the closed orbit corresponding to $(a, c(a))$ and the point $(r_{2k+1}, 0)$ goes to zero as $a \to \underline{a}$.

We will show that $c(a)$ can be defined in $(\underline{a}, r_{2k})$. We define $\bar{a}$ as follows:

$$\bar{a} := \sup\{a \in \mathbb{R} |\ c = c(\xi) \text{ can be defined for all } \xi \in (\underline{a}, \bar{a})\}.$$

Suppose $\bar{a} < r_{2k}$. From Step 1 we can find $\tilde{a}$ with $\bar{a} < \tilde{a} < r_{2k}$ so that there is a closed orbit that contains $(\tilde{a}, 0)$ on the phase plane. Since there are closed orbits with any small amplitude encircling the point $(r_{2k}, 0)$. The function $c(a)$ can be defined at some $\tilde{a}$ for $\tilde{a} \in (\bar{a}, r_{2k})$. Using Step 2, we can expand the domain of $c(a)$ to the left. Since $c(a)$ is unique, this contradicts to the definition of $\bar{a}$. Thus $\bar{a} = r_{2k}$.

Since $c(a)$ is unique and continuous, there is precisely one closed orbit that contains the point $(a, 0)$. Thus the limit $\lim_{a \to r_{2k}} T(a)$ should coincide with the limit in the statements of Theorem 2.6 in Section 7 of [AP93]. We have

$$\lim_{a \to r_{2k}} T(a) = \frac{2\pi}{\sqrt{f_u(r_{2k}, 0)}}.$$

Hereafter we will show that $\lim_{a \to \underline{a}} T(a) = \infty$ in the case where the shortest distance of the family of periodic orbits and the point $(r_{2k+1}, 0)$ goes to zero. First, we consider the linearized eigenvalue problem of (3.2) at the point $(r_{2k+1}, 0)$. Let λ_1, λ_2 be the eigenvalues. Then we have

$$\lambda_1 = \frac{1}{2}\left\{-(c + f_p(r_{2k+1}, 0)) - \sqrt{(c + f_p(r_{2k+1}, 0))^2 - 4f_u(r_{2k+1}, 0)}\right\},$$
$$\lambda_2 = \frac{1}{2}\left\{-(c + f_p(r_{2k+1}, 0)) + \sqrt{(c + f_p(r_{2k+1}, 0))^2 - 4f_u(r_{2k+1}, 0)}\right\}.$$

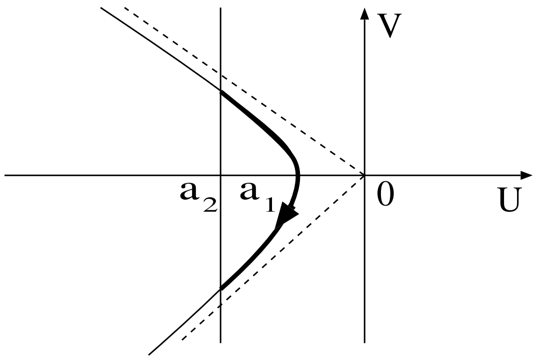


Figure 5. This picture indicates the phase plane displayed in the new coordinate. The thick curved arrow $(\tilde{U}, \tilde{V})$ is the arc that we observe. Two dashed lines are directions of the two eigenvectors of the matrix. The time required for traveling through the thick part of the arc diverges as $a_1 \to 0$.

Because $f_u(r_{2k+1}, 0) < 0$, we have $\lambda_1 < 0 < \lambda_2$. Thus the equilibrium point $(r_{2k+1}, 0)$ on the phase plane is hyperbolic. Then the Grobman-Hartman theorem says that there is a local homeomorphism Ψ such that $\phi_t \circ \Psi = \Psi \circ \tilde{\phi}_t$ and $\Psi(0,0) = (r_{2k+1}, 0)$ where $\phi_t, \tilde{\phi}_t$ are the semiflows on $\mathbb{R}^2$ formed by (3.2) and (3.4) respectively.

We can see that the time required for traveling through a neighborhood of the origin diverges as the shortest distance of the arc and the origin tends to zero. We omit the details of the proof of this fact.

We consider arcs of (3.2) in a neighborhood $(r_{2k+1}, 0)$. For each arc corresponding to (a, c), there is an orbit of (3.4) that is mapped to the arc by Ψ. Since $c(a)$ is bounded, the time which needs the orbit of (3.4) to pass a neighborhood of the origin uniformly diverges. Thus the time which needs the arc corresponding to $(a, c(a))$ to pass a neighborhood of the origin diverges as $a \to \underline{a}$. This means

$$\lim_{a \to \underline{a}} T(a) = \infty. \tag{3.7}$$

We can prove (3.7) similarly in the case where $\underline{a} = r_{2k-1}$. The proof is completed. □

Lemma 3.2. *Let $F(a, c)$ be the function defined by* (3.5)*. If there is a closed orbit corresponding to* (a_0, c_0) *on the phase plane, then* $F_c(a_0, c_0) \neq 0$.

Proof. We use the notation used in the proof of Lemma 3.1. We assume that a closed orbit corresponding to (a_0, c_0) exists. Differentiating $F(a, c) = U(\tau(a, c), a, c) - a$ with respect to c yields

$$F_c(a, c) = U_\zeta(\tau(a, c), a, c)\tau_c(a, c) + U_c(\tau(a, c), a, c).$$

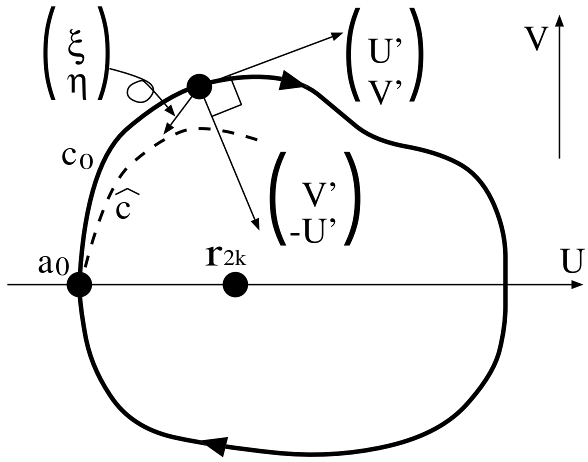


FIGURE 6. The thick closed curve represents the closed orbit corresponding to (a_0, c_0) whose starting and arrival points are $(a_0, 0)$. The dashed curve represents the arc corresponding to $(a_0, \hat{c})$ $(\hat{c} > c_0)$ whose starting point is also $(a_0, 0)$. The short arrow represents the vector (ξ, η). This picture indicates that the vector (ξ, η) points toward the interior of the closed orbit.

We have

$$F_c(a_0, c_0) = U_c(\tau(a_0, c_0), a_0, c_0),$$

because $U_\zeta(\tau(a_0, c_0), a_0, c_0) = 0$. We have to show that $U_c(\tau(a_0, c_0), a_0, c_0) \neq 0$. Let $\hat{c}$ $(> c_0)$ be a real number that is close to c_0. Using the vector $\begin{pmatrix} V \\ -c_0V - f(U, V) \end{pmatrix}$, we can see by [Du53] that the arc corresponding to (a_0, c_0) does not intersect with the arc corresponding to $(a_0, \hat{c})$ in spite that all assumptions of [Du53] are not satisfied on $\{V = 0\}$. The continuity of the arc corresponding to (a, c) with respect to c, togather with the above fact, tells us that the point $(U(\zeta, a_0, \hat{c}), V(\zeta, a_0, \hat{c}))$ $(\zeta > 0)$ is in the domain surrounded by the closed orbit corresponding to (a_0, c_0). This means that $U(\tau(a, c), a, c)$ is non-decreasing in c. We define ξ and η as follows:

$$\xi(\zeta) := U_c(\zeta, a_0, c_0), \quad \eta(\zeta) := V_c(\zeta, a_0, c_0),$$

where U_c is a derivative of U with respect to the third variable. Let $G(\zeta)$ be the inner product of $\begin{pmatrix} V_\zeta \\ -U_\zeta \end{pmatrix}$ and $\begin{pmatrix} \xi \\ \eta \end{pmatrix}$. Namely $G(\zeta) = \xi(\zeta)V_\zeta(\zeta) - \eta(\zeta)U_\zeta(\zeta)$. Then we have $G(\zeta) \geq 0$, because the vector $\begin{pmatrix} \xi \\ \eta \end{pmatrix}$ points toward the interior of the closed orbit (see Figure 6).

Differentiating (3.2) with respect to c yields

$$\begin{cases} \dfrac{d\xi}{d\zeta} = \eta \\ \dfrac{d\eta}{d\zeta} = -v - c\eta - f_u(U(\zeta), V(\zeta))\xi - f_p(U(\zeta), V(\zeta))\eta. \end{cases} \tag{3.8}$$

Using (3.2) and (3.8), we can express $G(\zeta)$, $G_\zeta(\zeta)$, $G_{\zeta\zeta}(\zeta)$ and $G_{\zeta\zeta\zeta}(\zeta)$ with ξ, η, c, V and derivatives of f as follows:

$$\begin{aligned} G(\zeta) =& -(c\xi V + \xi f + \eta V), \\ G_\zeta(\zeta) =& (c\xi V + \xi f + \eta V)(c + f_p) + V^2, \\ G_{\zeta\zeta}(\zeta) =& (c\xi V + \xi f + \eta V)\left\{f_{up}V - (cV + f)f_{pp} - (c + f_p)^2\right\} \\ & - 3cV^2 - V^2 f_p - 2Vf, \\ G_{\zeta\zeta\zeta}(\zeta) =& (c\xi V + \xi f + \eta V)\left[(cV + f)^2 f_{ppp} + \{4(c + f_p)(cV + f) - Vf_u\} f_{pp}\right. \\ & - (4cV + 3Vf_p + f_p)f_{up} - V(3cV + 1)f_{upp} \\ & \left. + V^2 f_{uup} - V^2 + (c + f_p)^3\right] \\ & - V^2(Vf_{up} - cVf_{pp}ff_{pp}) - 2V(Vf_u - cVf_p - ff_p) \\ & + 2(cV + f)(3cV + ff_pV + f). \end{aligned}$$

We suppose that $U_c(\tau(a_0, c_0), a_0, c_0) = \xi(\tau(a_0, c_0)) = 0$. Since $V(\tau(a_0, c_0), a_0, c_0) = 0$, we obtain $G(\tau) = G_\zeta(\tau) = G_{\zeta\zeta}(\tau) = 0$ and $G_{\zeta\zeta\zeta}(\tau) = 2f^2 > 0$ where $\tau = \tau(a_0, c_0)$. Therefore, there is a small constant $\delta > 0$ such that $G(P - \delta) < 0$. This is a contradiction, because $G(\zeta) \geq 0$. $\square$

Lemma 3.3. *There is a constant $M > 0$ such that* $\sup_{\zeta \in \mathbb{R}} |U_\zeta(\zeta)| \leq M$ *for any closed orbit* $(U(\zeta), U_\zeta(\zeta))$ *of* (3.2) .

Proof. Let $(U(\zeta), U_\zeta(\zeta))$ be a closed orbit of (3.2) with some c. Then $U(\zeta)$ satisfies (2.1). Thus from (A3) there is a constant $L_2 > 0$ such that $||U(\zeta)||_{C^1(S^1)} < L_2$ for any periodic solution or constant solution. The lemma is proved. $\square$

Lemma 3.2 completes the proof of Lemma 3.1.

4. Preparation for the Proof of Theorem A

In this section we will show that every *cluster* is a totally ordered set in the order $\rhd$ (Corollary 4.2). We will show that $z(u - v) = z(v_x)$ provided that $u, v \in S$ and $v \rhd u$ (Lemma 4.4). The two lemmas are used to prove Theorem A.

The following Lemma 4.1 is a generalized version of Corollary 4.2 below.

Lemma 4.1. *Let $(u(x), u_x(x))$, $(v(x), v_x(x))$ be closed orbits on the phase plane. Then the two closed orbits does not intersect.*

We can prove Lemma 4.1 by contradiction. We omit the proof.

Using a phase plane analysis and Lemma 4.1, we immediately obtain the following corollary.

COROLLARY 4.2 (Matano and Nakamura [MN97]). *Let* $u, v \in S$. *If* $R(u) \cap R(v) \neq \emptyset$, *then* $\mathrm{Int}(R(v)) \supset R(u)$ *or* $\mathrm{Int}(R(u)) \supset R(v)$ *where* $\mathrm{Int}(R(u))$ *indicates the set consists of the interior points of* $R(u)$.

Remark 4.3. Let $u, v \in S$ $(u \neq v)$. By Corollary 4.2, we can see that $u \rhd v$ means that $\mathrm{Int}(R(u)) \supset R(v)$.

Let $u, v \in S$. By using Corollary 4.2, we have either $u \rhd v$ or $v \rhd u$ provided that $R(u) \cap R(v) \neq \emptyset$.

Corollary 4.2 and the definition of the *clusters* show that every *cluster* is a totally ordered set. Thus we can number the elements of each *cluster* $\{u_1^{kl}, u_2^{kl}, \ldots, u_{m_{kl}}^{kl}\}$ in such a way that

$$u_1^{kl} \lhd u_2^{kl} \lhd \cdots \lhd u_{m_{kl}}^{kl}.$$

LEMMA 4.4 (Matano and Nakamura [MN97]). *Let* $u, v \in S$. *If* $v \rhd u$, *then* $z(u - v) = z(v_x)$.

5. PROOF OF COROLLARY B AND LEMMAS E AND F

In this section we will prove Corollary B and Lemmas E and F by using Lemma 5.1 and the results in Sections 3 and 4.

Let $c = c(a)$ be the function defined in the statement of Lemma 3.1, and let $T = T(a)$ be the period of the closed orbit corresponding to $(a, c(a))$ defined in the statement of Lemma 3.1.

LEMMA 5.1. *Let* $u \in S$ *be the closed orbit corresponding to* $(a_0, c(a_0))$ *in Section* 3. *If* u *is hyperbolic, then* $\partial_a T(a)|_{a=a_0} \neq 0$.

Proof. We will prove the lemma by contradiction. We assume that $\partial_a T(a)|_{a=a_0} = 0$. Let $u(x,t) = U(\zeta)$ $(\zeta = x - ct)$ be a rotating wave or a steady state. We can suppose that $U(0) = a$ and $U_\zeta(0) = 0$ without loss of generality. The function $U = U(\zeta, a, c(a))$ defined in Section 3 satisfies $U_{\zeta\zeta} + c(a)U_\zeta + f(U, U_\zeta) = 0$. Differentiating the equation with respect to a gives

$$\partial_{\zeta\zeta}(U_a + c_a U_c) + c\partial_\zeta(U_a + c_a U_c) + f_u \cdot (U_a + c_a U_c) + f_p \partial_\zeta(U_a + c_a U_c) = -c_a U_\zeta.$$

Let $\varphi(\zeta) = U_a(\zeta) + c_a U_c(\zeta)$. The function $\varphi(\zeta)$ satisfies the following equation:

$$\varphi_{\zeta\zeta} + c\varphi_\zeta + f_u\varphi + f_p\varphi_\zeta = -c_a U_\zeta, \quad \zeta \in S^1. \tag{5.1}$$

If $c_a(a_0) = 0$, then $\alpha \cdot U_\zeta(\zeta)$ $(\alpha \in \mathbb{R})$ are the solutions to (5.1) because of the hyperbolicity of $U(\zeta)$. If $c_a(a_0) \neq 0$, then (5.1) has no solution. Because 0 is a simple eigenvalue of the following problem:

$$\varphi_{\zeta\zeta} + c\varphi_\zeta + f_u\varphi + f_p\varphi_\zeta = \lambda\varphi, \quad \zeta \in S^1.$$

Case 1: $c_a(a_0) = 0$
Differentiating $U(0, a, c(a)) = U(T(a), a, c(a))$ with respect to a gives

$$(5.2)\quad \begin{aligned} &U_a(0, a, c(a)) + c_a(a)U_c(0, a, c(a)) \\ &= \partial_a T(a)U_\zeta(T(a), a, c(a)) + U_a(T(a), a, c(a)) + c_a(a)U_c(T(a), a, c(a)). \end{aligned}$$

Substituting $\partial_a T(a)|_{a=a_0} = 0$ and $c_a(a_0) = 0$ for (5.2) gives $U_a(0, a_0, c(a_0)) = U_a(T(a_0), a_0, c(a_0))$. Since $U_a(\,\cdot\,, a_0, c(a_0))$ is a periodic function and the period $T(a_0)$ is equal to $1/n$ for some $n \in \{1, 2, \cdots\}$, we have $U_a(0, a_0, c(a_0)) = U_a(1, a_0, c(a_0))$. Since $\varphi(\zeta) = U_a(\zeta)$, we have

$$\varphi(0) = \varphi(1). \tag{5.3}$$

We differentiate $U_\zeta(0, a, c(a)) = U_\zeta(T(a), a, c(a))$ with respect to a, and substitute a_0 for it. Then we obtain

$$U_{\zeta a}(0, a_0, c(a_0)) = U_{\zeta a}(T(a_0), a_0, c(a_0)) + \partial_a T(a)|_{a=a_0} U_{\zeta\zeta}(T(a_0), a_0, c(a_0)).$$

Since $\varphi_\zeta(\zeta) = U_{\zeta a}(\zeta, a_0, c(a_0))$, we have

$$\varphi_\zeta(0) = \varphi_\zeta(T(a_0)) + \partial_a T(a)|_{a=a_0} U_{\zeta\zeta}(T(a_0), a_0, c(a_0)).$$

Since $\partial_a T(a)|_{a=a_0} = 0$ and $\varphi_\zeta(T(a_0)) = \varphi_\zeta(1)$, we have

$$\varphi_\zeta(0) = \varphi_\zeta(1). \tag{5.4}$$

Using (5.3) and (5.4), we can see that $\varphi(\zeta)(= U_a(\zeta))$ satisfies (5.1) and periodic boundary conditions. By the hyperbolicity of $u(x, t)(= U(\zeta))$, we see that $\varphi(\zeta) = \alpha \cdot U_\zeta(\zeta)$ ($\alpha \in \mathbb{R}$) are the solutions to (5.1). On the other hand $\varphi(0) = U_a(0) = 1$. It contradicts that $U_\zeta(0) = 0$. We can see that $\partial_a T(a)|_{a=a_0} \neq 0$.
Case 2: $c_a(a_0) \neq 0$
Using the assumption of contradiction $\partial_a T(a)|_{a=a_0} = 0$, we can obtain the following two equalities in a similar way of Case 1:

$$\varphi(0) = \varphi(1), \quad \varphi_\zeta(0) = \varphi_\zeta(1). \tag{5.5}$$

Using (5.5), we can see that $\varphi(\zeta)$ satisfies (5.1) and periodic boundary conditions. The function $\varphi(\zeta)$ is a non-trivial solution to (5.1). This is a contradiction. Therefore, we obtain $\partial_a T(a)|_{a=a_0} \neq 0$. □

Hereafter, we consider the structure of each *cluster*. We divide the *clusters* in two types. One is a type of *clusters* that contain a constant steady state, and the other is a type of *clusters* that do not have a constant steady state.

First, we consider the type of *clusters* that have a constant steady state. Since the *cluster* C_{kl} has a constant steady state, we can see that $k = l$ by using a phase plane analysis. If k is odd, then $\sharp C_{kk} = 1$ and the element of C_{kk} is a stable constant steady state. If k is even, then $\sharp C_{kk} \geq 1$ and C_{kk} has precisely one unstable constant steady state.

Second, we consider the type of *clusters* C_{kl} that do not have a constant steady state. By observing the phase plane, we see that $l \geq k+2$, and k and l are even. If $u(x, t) = U(x - ct)$ is an element of C_{kl} that satisfies $(U(0), U_\zeta(0)) = (a, 0)$ and $U(\zeta) \leq a$, then we can deform the closed orbit $(U(\zeta), U_\zeta(\zeta))$ on the phase

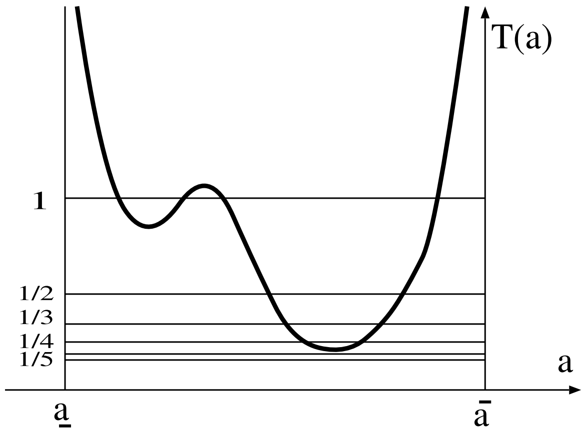


Figure 7. The picture shows the graph of $T(a)$ in the case of C_{kl} $(k \neq l)$. Each of the intersections of the curve and the lines corresponds to a rotating wave. In this case, the sequence of modified Morse indices is $(2, 2, 2, 4, 6, 8, 8, 6, 4, 2)$.

plane by using a similar way of Step 2 and Step 3 in the proof of Lemma 3.1, and enlarge the domain of $c = c(a)$. Let $(\underline{a}, \bar{a})$ be the maximal connected domain of the function $c = c(a)$. The closed orbit that corresponds to $(a, c(a))$ approaches $(r_{2k-1}, 0)$ or $(r_{2l+1}, 0)$ as $a \to \underline{a}$. Since $c(a)$ is bounded, the function $T(a)$ diverges to $+\infty$ as $a \to \underline{a}$. The function $T(a)$ diverges to $+\infty$ as $a \to \bar{a}$, because the closed orbit approaches $(r_{2k}, 0)$, $\ldots$, $(r_{2l-1}, 0)$ or $(r_{2l}, 0)$, and $c(a)$ is bounded. Hence the graph of $T(a)$ is as shown in Figure 7.

Proof of Lemma E. The statements (i) and (ii) are easily understood by observing a phase plane. $\square$

Proof of Lemma F. We can see that (i), (iv) and (v) follow from Figures 7 and 8. Lemma 3.1 implies (ii) and (iii). $\square$

Proof of Corollary B. Since $S = \bigcup_{1 \le k \le l \le N} C_{kl}$, we obtain the following:

$$\begin{aligned}\sharp S &= \sum_{1 \le k \le l \le N} \sharp C_{kl} \\ &\ge \sum_{j=1}^{N} \sharp C_{jj} \\ &\underset{\text{by Figure 8}}{\ge} N + \sum_{j=1}^{N} \left[\left[\frac{\sqrt{(f_u(r_j, 0))_+}}{2\pi} \right]\right].\end{aligned}$$

$\square$

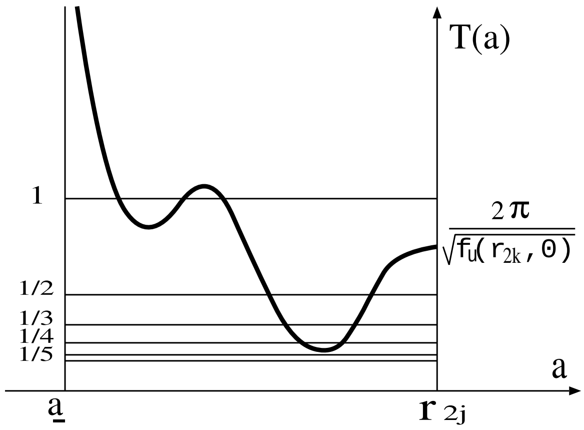


Figure 8. The picture indicates the graph of $T(a)$ in $(\underline{a}, r_{2k})$. Each of the intersections of the curve and the lines corresponds to a rotating wave. The sequence of modified Morse indices is easily computed from this picture. In this case, the sequence of modified Morse indices is $(4, 4, 6, 8, 8, 6, 4, 2, 2, 2)$.

Remark 5.2. If $\sharp S$ attains the lower bound, then every *cluster* is *simple* and *monotone*. If every *cluster* is *simple*, then the equality in the first inequality in the proof of Corollary B holds. If every *cluster* is *monotone*, then the equality in the second inequality in the proof of Corollary B holds. Therefore, $\sharp S$ attains the lower bound if and only if every *cluster* is *simple* and *monotone*.

6. Proof of Theorems A and C

In this section we will prove Theorems A and C by using Lemma 6.1, Lemma F and the main results of [AF88]. A simple example is given at the end of this section.

Lemma 6.1 (Blocking lemma). *Let $v, w \in S$ ($w \triangleright v$ and $I(w) < I(v)$). If there exists a wave $\bar{v} \in S$ such that $w \triangleright \bar{v} \triangleright v$ and $I(\bar{v}) = I(w)$, then v does not connect to w.*

The proof of Lemma 6.1 is essentially the same as the explanation after Definition 1.6 of [FR96].

Remark 6.2. Lemma 6.1 is called the *zero number blocking* (see Definition 1.6 of [FR96]).

We will use the following proposition to prove Theorem A.

Proposition 6.3 (Angenent and Fiedler [AF88]). *Let $u \in S$ with $i(u) > 0$ be hyperbolic. Then*

(i) *The wave u connects to u_+ and u_-.*

(ii) *For any* $n \in \mathbb{N}$, $0 < 2n \le i(u)$, *there exists a wave* $u^{(n)} \in S$ *such that* $u_- < u^{(n)} < u_+$, $z(u^{(n)} - u) = 2n$, *and* u *connects to* $u^{(n)}$.

We are in a position to prove Theorem A.

Proof of Theorem A. Let v be a wave in C_{kl} $(k \le l)$ and let w be a wave in C_{mn} $(m \le n)$. We prove whether v connects to w or not. When there is a connecting orbit $u = u(t)$ that connects v and w, we can suppose that $I(w) \le I(v)$, because $i(w) + 1 \le z(u - v) \le i(v)$ (see Lemma 3.7 in [AF88]). If $I(v) = 0$, then there is no connecting orbit starting from v. Thus we assume that $I(v) > 0$. We can see that k and l are odd, using a phase plane analysis.

There are two cases in general terms. In one case, w belongs to the same *cluster* as v (*i.e.* $(m, n) = (k, l)$). In the other case, w belongs to another *cluster* which does not include v (*i.e.* $(m, n) \ne (k, l)$). First, we consider the case where $w \in C_{mn}$ $((m, n) \ne (k, l))$.

Case 1: $(m, n) \ne (k, l)$
We can divide the case into four more cases.

Case 1-1: $(m, n) \in \{(k-1, k-1), (l+1, l+1)\}$
Since both $k-1$ and $l+1$ are even, the cluster C_{mn} has precisely one wave (This wave is a stable constant steady state). We can see that v connects to w by (i) of Theorem 6.3, because $w = v_+$ or $w = v_-$.

Case 1-2: $(m, n) \notin \{(k-1, k-1), (l+1, l+1)\}$ *and* $R(C_{kl}) \cap R(C_{mn}) = \emptyset$
There is a wave $\bar{w} \in S$ $(I(\bar{w}) = 0)$ between v and w in the usual order (*i.e.* $v(x) < \bar{w}(x) < w(x)$ or $w(x) < \bar{w}(x) < v(x)$). We assume that there is a connecting orbit $u(t)$ that connects v and w. The function $z(u(t) - \bar{w}(t))$ is not non-increasing in t. This is a contradiction. Therefore, the wave v does not connect to any $w \in C_{mn}$. Namely the wave v does not connect to any wave of the *above clusters* and *below clusters* in the usual order except for the two *clusters* of Case 1-1.

Case 1-3: $C_{kl} \rhd C_{mn}$
We see that $i(v) \in \{I(v), I(v) - 1\}$ generally. We have

$$i(v) = I(v) - 1,$$

in the case that $i(v)$ is odd. We suppose that there is a connecting orbit $u(t)$ that connects v and w. Then

$$z(u - v) \le i(v), \tag{6.1}$$

(see Lemma 3.7 in [AF88]). Lemma 4.4 tells us that (6.1) contradicts that $z(u(t) - v(t)) = I(v)$ for large $t > 0$. The wave v does not connect to any $w \in C_{mn}$. Namely v does not connect to any wave of the *clusters* that is smaller than C_{kl} in the order $\rhd$.

Case 1-4: $C_{mn} \rhd C_{kl}$
There is a $\bar{w} \in S$ $(I(\bar{w}) = 0)$ such that $R(v) \cap R(\bar{w}) = \emptyset$ and $w \rhd \bar{w}$. We suppose that there is a connecting orbit $u(t)$ which connects v and w. The function $z(u(t) - \bar{w}(t))$ is not non-increasing in t. This is a contradiction. Therefore, v does not connect to any $w \in C_{mn}$.

The Case 1 can be summarized as follows: If v connects to w in another *cluster*, then w should be v_+ or v_-.

Case 2: $(m, n) = (k, l)$
Let w be another wave of the same *cluster* C_{kl}. We divide this case in two more cases.

Case 2-1: $v \rhd w$
We suppose that there is a connecting orbit $u(t)$ that connects v and w. We can see that

$$I(v) = z(u(t) - v(t)) \leq i(v) \quad \text{for large } t,$$

(see Lemma 3.7 in [AF88]), because $v \rhd w$. Thus if $i(v)$ is odd (*i.e.* $i(v) = I(v) - 1$), then we obtain a contradiction. The wave u does not connect to w provided that $i(v)$ is odd.

Case 2-2: $w \rhd v$
Owing to Theorem 6.3, the wave v connects to w that attains the following minimum for each d ($d = 2, 4, 6, \ldots, I(v) - 2$):

$$\min_{I(w)=2d, w \rhd v} |R(w)|,$$

where $|R(u)| := \max_{x \in S^1} u(x,t) - \min_{x \in S^1} u(x,t)$. Suppose $i(v)$ is odd. The wave v, however, does not connect to any other w, because Lemma F tells us that there exists a wave $\bar{w}$ such that $w \rhd \bar{w} \rhd v$ and $I(w) = I(\bar{w})$. Thus we can see by Lemma 6.1 that the zero number blocking occurs.

The Case 2 can be summarized as follows. The wave v connects to $I(v)/2 - 1$ different waves that are bigger than v in the order $\rhd$ in the same *cluster*. The wave does not connect to any other wave in the same *cluster* provided that $i(v)$ is odd.

The Case 1 and the Case 2 cover all the combinations of v and w. Thus the proof is completed. □

Proof of Theorem C. We show that the hypotheses of Theorem C satisfy those of Theorem A.

Every *cluster* is *simple* and *monotone* if and only if $\sharp S$ attains the lower bound (see Remark 5.2).

We will show that the Morse index of every wave is odd or zero. Suppose that there is a wave $u \in S$ whose Morse index is even and not zero. Using Proposition 6.3, we can see that there exists a wave $v \in S$ such that $I(u) = I(v)$ and u connects to v heteroclinically. However, u and v are not in the same *cluster*, because the *cluster* is *monotone*. Thus v belongs to another *cluster*. However, there is no heteroclinic connection, because every *cluster* is *simple* and there should be a stable steady state between u and v in the usual order. This is a contradiction. Therefore all the hypotheses of Theorem A are satisfied. □

Example 6.4. Figure 9 shows the profile of every $u \in S$ and the diagram that shows which $u \in S$ and $v \in S$ are connected heteroclinically and which are not when $\{r_j\}_{j=1}^5$ are the roots of $f(\,\cdot\,, 0)$, $[[\sqrt{f_u(r_2, 0)}/(2\pi)]] = 2$, $[[\sqrt{f_u(r_4, 0)}/(2\pi)]] = 3$, $\sharp S = 10$, and all $u \in S$ are hyperbolic.

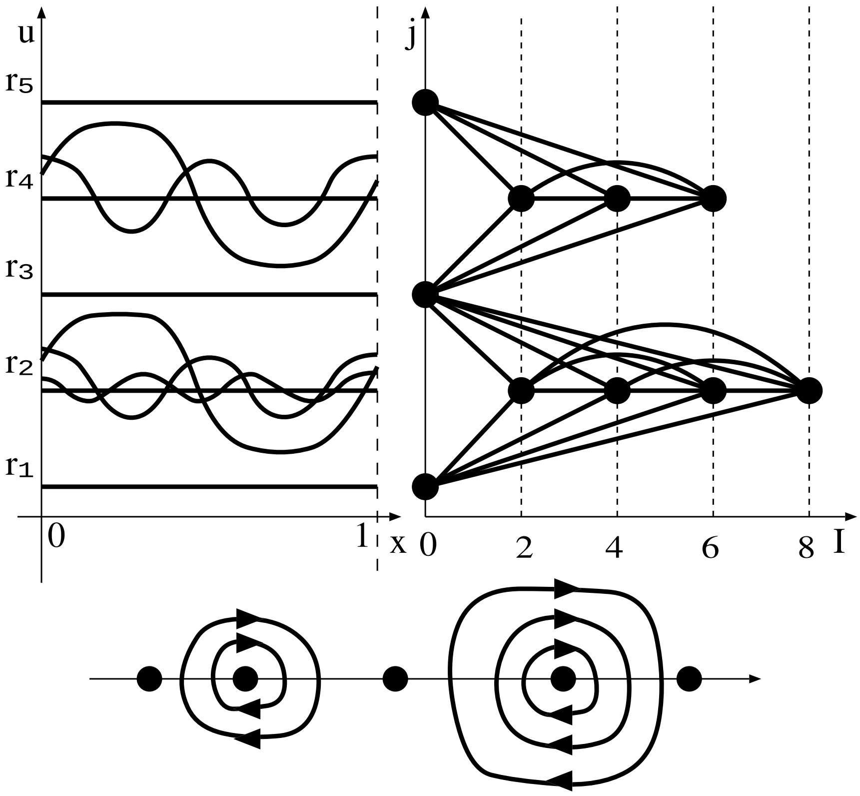


Figure 9. In the left figure, the thick curves and the lines indicate the profile of all the waves that move to the right or the left at each constant speed. In the right figure, the horizontal axis indicates the modified Morse index and the vertical axis indicates the suffix of C_{jj}. The points mean elements of S. The thick curves and the lines represent the connecting orbits. The lower figure shows closed orbits and equibrium points in the uu_x-plane. Note that they do not necessarily correspond to the same value of c.

Remark 6.5. If there is a wave $v \in S$ such that $i(v)(\neq 0)$ is even, then we cannot determine by the method used in the proof of Theorem A whether v connects to waves that are smaller than v in the order $\rhd$ or not.

Remark 6.6. We have shown Theorem A by using the *structure* and the results of [AF88]. This means that the results of [AF88] that looks a partial answer is a complete answer in some sense when the Morse index of every wave is odd or zero.

7. Proof of Theorem A' and Lemma F'

In this section we will study the case where the nonlinear term f depends only on u, and establish a sufficient condition that guarantees that all the *clusters* are *simple* and *monotone*.

We will use a character g to denote the nonlinear term (*i.e.* $f(u, u_x) = g(u)$). In this case (1.1) is written as follows:

$$\begin{cases} u_t = u_{xx} + g(u), & x \in S^1, \\ u(x,0) = u_0(x), & x \in S^1. \end{cases} \tag{7.1}$$

Matano [Ma88] showed that (1.1) does not have rotating waves provided that $f(u,p) = f(u,-p)$. Since the nonlinear term g depends only on u and satisfies this property, the equation (7.1) does not have rotating waves.

We consider the following Neumann problem:

$$\begin{cases} u_t = u_{xx} + g(u), & x \in (0, 1/2), \\ u_x(0) = 0 = u_x(1/2). \end{cases} \tag{7.2}$$

Let $u(x)$ be a wave of (7.1). Then there exists $\theta(\in S^1)$ such that $u_x(\theta) = 0$ and $u(x) \leq u(\theta)$ for all $x \in S^1$. We can see by a phase plane analysis that $u_x(\theta + 1/2) = 0$. Therefore $u(x + \theta)$ $(0 < x < 1/2)$ is a steady state of (7.2). Let $\tilde{u}(x)$ denotes $u(x + \theta)$. Thus $\tilde{u}(x)$ is a steady state of (7.2).

Next, let $v(x)$ be a non-constant steady state of (7.2) that satisfies $v(x) \leq v(0)$. Then $u(x)$ is a standing wave of (7.1) where

$$u(x) = \begin{cases} v(x), & 0 \leq x \leq 1/2; \\ v(1-x), & 1/2 \leq x \leq 1. \end{cases}$$

We can identify any wave u of (7.1) with a steady state $\tilde{u}$ of (7.2), and by *the steady state associated with u of* (7.2) we shall mean $\tilde{u}$. In short $\tilde{u} = v$.

Let v, w be steady states of (7.1) and let $\tilde{v}$, $\tilde{w}$ be steady states associated with v, w respectively. Suppose that a heteroclinic orbit $\tilde{u}(x,t)$ of (7.2) that connects $\tilde{v}$ and $\tilde{w}$ exists. Then $u(x,t)$ is a solution of (7.1) where

$$u(x,t) = \begin{cases} \tilde{u}(x,t), & 0 \leq x \leq 1/2; \\ \tilde{u}(1-x,t), & 1/2 \leq x \leq 1. \end{cases}$$

Moreover $u(\,\cdot\,,t) \to v(x)$ $(t \to -\infty)$ and $u(\,\cdot\,,t) \to w(x)$ $(t \to \infty)$. Thus $u(x,t)$ is a connecting orbit of (7.1) that connects v and w. In short, v connects to w if $\tilde{v}$ connects to $\tilde{w}$. We will use this fact to prove the existence of connecting orbits in the proof of Theorem A'.

We give two lemmas about (7.2) without proofs.

LEMMA 7.1. *Let $\{u_1^{kl}, u_2^{kl}, \ldots, u_{m_{kl}}^{kl}\}$ be a cluster and let $\{\tilde{u}_1^{kl}, \tilde{u}_2^{kl}, \ldots, \tilde{u}_{m_{kl}}^{kl}\}$ be the set of steady states of* (7.2) *associated with the waves of the cluster. Let $\{u_{b_1}^{kl}, u_{b_2}^{kl}, \ldots, u_{b_n}^{kl}\}$ $(b_1 < b_2 < \cdots < b_n)$ be the waves whose Morse indices are the same number (i.e. $I(u_{b_1}^{kl}) = I(u_{b_2}^{kl}) = \cdots = I(u_{b_n}^{kl})$). Then $i(\tilde{u}_{b_{n-2j}}^{kl}) =$*

$I(u^{kl}_{b_{n-2j}})/2$ for $j \in \{0,1,\dots,[(n-1)/2]\}$, and $i(\tilde{u}^{kl}_{b_{n-2j-1}}) = I(u^{kl}_{b_{n-2j-1}})/2+1$ for $j \in \{0,1,\dots,[(n-2)/2]\}$.

Proof. In the case of the Dirichlet problem, we can find the proof in Lemma 2.1 of [BF89]. We can prove the lemma in a similar way. □

Lemma 7.2. *Let u, v, w be waves and let $\tilde{u}$, $\tilde{v}$ be the steady states associated with u, v. If $i(u)$ is even, then the steady state $\tilde{u}$ connects to every $\tilde{v}$ that satisfies the following: $u \rhd v$, and there is no wave w such that $u \rhd w \rhd v$ and $I(u) = I(w)$.*

Proof. In the case of the Neumann problem, the problem of the heteroclinic connections are completely determined by [FR96]. We can prove the lemma by using Lemma 7.1, Definition 1.6 of [FR96] and Lemma 1.7 of [FR96]. □

Proof of Lemma F'. If $b_{n-2j} > 1$, then there exists v ($\lhd u^{kl}_{b_{n-2j}}$) such that v blocks the connections from $u^{kl}_{b_{n-2j}}$ to all the wave that are smaller than $u^{kl}_{b_{n-2j}}$ in the order $\rhd$. This means that $i(u^{kl}_{b_{n-2j}}) = I(u^{kl}_{b_{n-2j}}) - 1$. If $b_{n-2j} = 1$, then $k = l$. There also exists a wave v that satisfies the above conditions (the wave v may be a constant steady state). Thus $i(u^{kl}_{b_{n-2j}}) = I(u^{kl}_{b_{n-2j}}) - 1$. In short $i(u^{kl}_{b_{n-2j}}) = I(u^{kl}_{b_{n-2j}}) - 1$ for $j \in \{0,1,\dots,[(n-1)/2]\}$.

We consider whether $i(u^{kl}_{b_{n-2j-1}}) = I(u^{kl}_{b_{n-2j-1}}) - 1$ or $i(u^{kl}_{b_{n-2j-1}}) = I(u^{kl}_{b_{n-2j-1}})$. If $n-2j-1 > 1$, then $\tilde{u}^{kl}_{b_{n-2j-1}}$ connects to $\tilde{u}^{kl}_{b_{n-2j-2}}$. Thus $u^{kl}_{b_{n-2j-1}}$ connects to $u^{kl}_{b_{n-2j-2}}$. This means that $i(u^{kl}_{b_{n-2j-2}}) = I(u^{kl}_{b_{n-2j-1}})$. If $n-2j-1 = 1$, then there exists a wave $\tilde{v}$ such that the following hold: $v \lhd u^{kl}_{b_1}$ and $\tilde{u}^{kl}_{b_1}$ connects to $\tilde{v}$. Thus $u^{kl}_{b_1}$ connects to v. Hence $i(u^{kl}_{b_1}) = I(u^{kl}_{b_1})$. In short $i(u^{kl}_{b_{n-2j-1}}) = I(u^{kl}_{b_{n-2j-1}})$ for $j \in \{0,1,\dots,[(n-2)/2]\}$.

□

Proof of Theorem A'. Let u be a non-constant wave whose Morse index is even. In Theorem A we have identified waves that are connected by u and that satisfy $z(u-v) \le I(u) - 2$. Thus we have to check whether u connects to v or not, in the case where $z(u-v) = I(u)$.

Case 1: $v \rhd u$

Let w be a wave that satisfies the following: w is the smallest wave in the order $\rhd$ that satisfies $w \rhd u$ and $I(u) = I(w)$. Because of Lemma F', w exists in the *cluster* to which u belongs, and $i(w) = I(w) - 1$. Let $\tilde{u}$ and $\tilde{w}$ be steady states of (7.2) associated with u and w respectively. We can see that $\tilde{u}$ connects to $\tilde{w}$ (see Case 2-1 in the proof of Lemma F'). Thus u connects to w. There is no other wave that is connected by u, because w blocks other connections (see Lemma 6.1).

Case 2: $v \lhd u$

Since $v \lhd u$, it is automatically satisfied that $z(u-v) = I(u)$. If there is a wave w such that $u \rhd w \rhd v$ and $I(u) = I(w)$, then u does not connect to v because w blocks the connection (see Lemma 6.1). On the other hand, if there

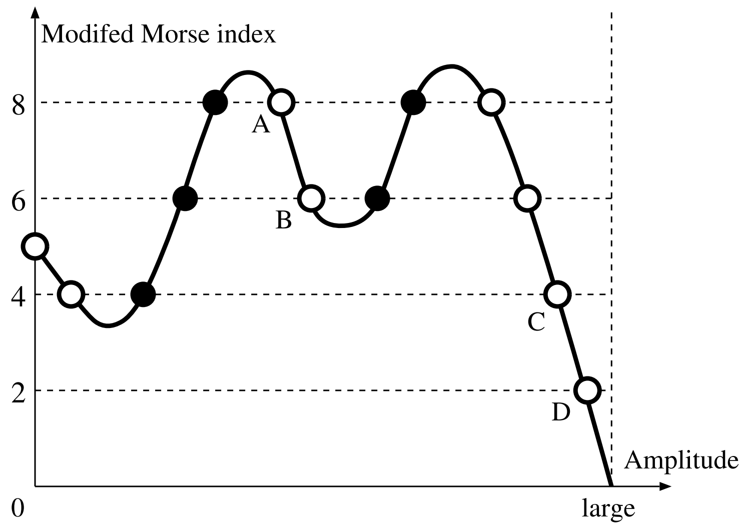


Figure 10. Each black point indicates a wave whose Morse index is even (*i.e.* $i(u_j^{kl}) = I(u_j^{kl})$) and each white point indicates a wave whose Morse index is odd (*i.e.* $i(u_j^{kl}) = I(u_j^{kl})-1$). The point A connects only to B, C, D and two constant steady states.

is no such wave, then u connects to v because $\tilde{u}$ connects to $\tilde{v}$ (see Lemma 7.2). Therefore the theorem is proved. $\square$

Example 7.3. Let $J_{kl} = \big(I(u_j^{kl})\big)_{j=1}^{m_{kl}}$ $(k \neq l)$ be a sequence of modified Morse indices. Figure 10 represents the sequence of modified Morse indices J_{kl} (see Remark 2.19). Since $k \neq l$, we see by (v) of Lemma F that $I(u_{m_{kl}}^{kl}) = 2$. If $i(u)$ is odd, all connections toward a smaller wave in the order $\rhd$ (*i.e.* toward the left in Figure 10) are blocked. If $i(u)$ is even, the connections to a smaller wave in the order $\rhd$ are not necessarily blocked.

8. Proof of Proposition D

In this section we consider the case where the nonlinear term does not depend on u_x (see (7.1)). We will use the notation used in Section 7.

We will show a sufficient condition that guarantees *clusters* to be *monotone.* The following lemma is well-known:

Lemma 8.1. *Suppose $g(\,\cdot\,)$ has exactly three roots $\{r_i\}_{i=1}^3$ and $r_1 < r_2 = 0 < r_3$. If $g(u)/|u|$ is decreasing for $u \in (r_1, 0) \cup (0, r_3)$, then there are only three monotone clusters.*

The proof of Lemma 8.1 is essentially the same as that of Theorem 5.2 of [CI74]. We omit the proof.

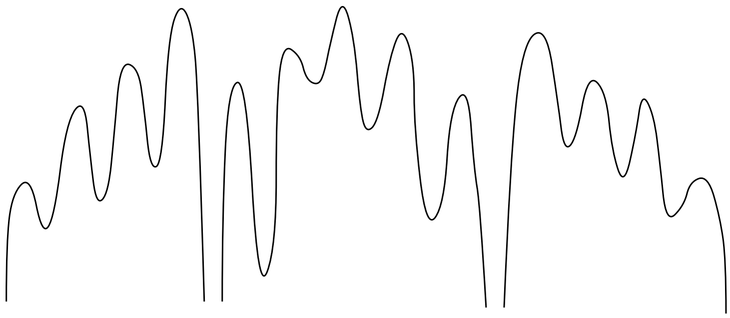


Figure 11. The graph of $G(r)$; $k = 1$ (left), $1 < k < n$ (center) and $k = n$ (right).

We will prove Proposition D after we state some definitions and notation. Hereafter, we assume that every wave of S is hyperbolic. Hence $g'(r_j) \neq 0$ for all $j \in \{1, 2, \ldots, N\}$. The point $G(r_j)$ ($j \in \{1, 3, 5, \ldots, N\}$) is a local maximum point and $G(r_j)$ ($j \in \{2, 4, 6, \ldots, N-1\}$) is a local minimum point where $G(r)$ is defined by (2.6).

First, we define a set of intervals

$$W(r) := \{\rho \mid G(\rho) < r\}.$$

We impose the following condition of the function G:

(**A6**) Let I be a bounded connected component of $W(r)$ for $r \in \mathbb{R}$. Let $J = \{r_k, r_{k+1}, \ldots, r_{l-1}, r_l\}$ $(1 \leq k \leq l \leq N)$. If $I \supset J$, then $\sharp J = 1$.

The closed curves described as $\{(u, v) \mid v^2 + 2G(u) = \text{constant}\}$ on the phase plane are candidates of steady state solutions of (7.1). If (A6) holds, then C_{kl} ($k \neq l$) is empty. Therefore, when (A6) holds, there is only one possibility which is the condition (A4) in Section 2. When (A4) is satisfied, the graph of $G(r)$ looks like one of Figure 11.

Example 8.2. If the graph of $G(r)$ is as shown in the center of Figure 11, the corresponding phase portrait is as shown in Figure 12.

If (A4) holds, then every cluster is *simple*. If $(A5)_j$ holds, then C_{jj} is *monotone*. Now we can prove Proposition D.

Proof of Proposition D. If (A4) holds, then (A6) holds. Thus every C_{kl} ($k \neq l$) is empty. Namely all the *clusters* are *simple*. After all $S = \bigcup_{j=1}^{N} C_{jj}$. Since every wave is hyperbolic, the *cluster* C_{jj} ($j \in \{1, 3, 5, \ldots, N\}$) has precisely one wave which is the stable constant steady state (see Remark 2.1). The condition $(A5)_j$ tells us that the *cluster* C_{jj} ($j \in \{2, 4, 6, \ldots, N-1\}$) is *monotone*. Thus

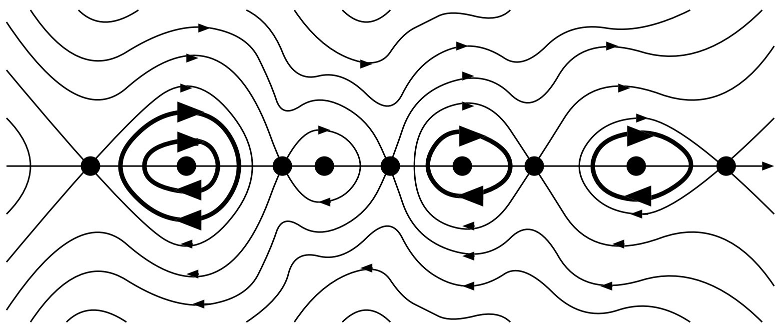


Figure 12. The picture indicates the phase portrait when $G(p)$ is as shown in the center of Figure 11. The thick closed curves indicate the closed orbits which correspond to standing waves, and the points indicate equilibrium points which correspond to constant steady states.

every *cluster* is *monotone*. Therefore, all the hypotheses of Theorem C are satisfied. The proof is completed. □

After completing this work, the author has been informed about the paper [FRW04] written by Fiedler, Rocha and Wolfrum. They have given the necessary and sufficient conditions whether any pair of waves is connected heteroclinically or not, and the method to calculate the Morse index of waves (*i.e.* the method to decide whether $i(u) = I(u)$ or $i(u) = I(u) - 1$).

References

[Am76] H. Amann, *Existence and multiplicity theorems for semi-linear elliptic boundary value problems*, Math. Z., 150 (1976), no. 3, 281-295.

[Am85] H. Amann, *Global existence for semilinear parabolic systems*, J. Reine Angew. Math., 360 (1985), 47-83.

[An86] S. Angenent, *The Morse-Smale property for a semilinear parabolic equation*, J. Diff. Eq., 62 (1986), no. 3, 427-442.

[An88] S. Angenent, *The zero set of a solution of a parabolic equation*, J. Reine Angew. Math., 390 (1988), 79-96.

[AP93] A. Ambrosetti and G. Prodi, *"A Primer of Nonlinear Analysis"*, Cambridge University Press, Cambridge, 1993.

[AF88] S. B. Angenent and B. Fiedler, *The dynamics of rotating waves in scalar reaction diffusion equations*, Trans. Amer. Math. Soc., 307 (1988), no. 2, 545-568.

[BF89] P. Brunovský and B. Fiedler, *Connecting orbits in scalar reaction diffusion equations. II. The Complete Solution*, J. Diff. Eq., 81 (1989), no. 1, 106-135.

[BF86] P. Brunovský and B. Fiedler, *Numbers of zeros on invariant manifolds in reaction-diffusion equations*, Nonlin. Anal., 10 (1986), no. 2, 179-193.

[CI74] N. Chafee and E. Infante, *A bifurcation problem for a nonlinear equation*, Applicable Anal., 4 (1974), 17-37.

[CL55] E. Coddington and N. Levinson, *"Theory of Ordinary Differential Equations"*, McGraw-Hill Book Co., Inc., New York-Toronto-London, 1955.

[Du53] G. F. Duff, *Limit-cycle and rotated vector fields*, Ann. of Math. (2) 57 (1953), 15-31.

[FM89] B. Fiedler and J. Mallet-Paret, *A Poincaré-Bendixson theorem for scalar reaction diffusion equations*, Arch. Rational Mech. Anal., 107 (1989), no. 4, 325-345.

[FR96] B. Fiedler and C. Rocha, *Heteroclinic orbits of semilinear parabolic equations*, J. Diff. Eq., 125 (1996), no. 1, 239-281.

[FRW04] B. Fiedler, C. Rocha and M. Wolfrum, *Heteroclinic orbits between rotating waves of semilinear parabolic equations on the circle*, to appear.

[Fr64] A. Friedman, *"Partial Differential Equations of Parabolic Type"*, Prentice-Hall, Englewood Cliffis, N. J., 1964.

[He84] D. Henry, *"Geometric Theory of Semilinear Parabolic Equations"*, Lect. Notes Math., Vol.840, Springer-Verlag, New York, 1984.

[He85] D. Henry, *Some infinite-dimensional Morse-Smale systems defined by parabolic partial differential equations*, J. Diff. Eq., 59 (1985), no. 2, 165-205.

[Li96] G. Lieberman, *"Second Order Parabolic Differential Equations"*, World Scientific Publishing Co., Inc., River Edge, NJ, 1996.

[Ma76] J. Mallet-Paret, *Negatively invariant sets of compact maps and an extension of a theorem of Cartwright*, J. Diff. Eq., 22 (1976), no. 2, 331-348.

[MN97] H, Matano and K. Nakamura, *The global attractor of semilinear parabolic equations on S^1*, Discrete Contin. Dynam. Systems, 3 (1997), no. 1, 1-24.

[Ma88] H. Matano, *Asymptotic behavior of solutions of semilinear heat equations on S^1*, "Nonlinear differential equations and their equilibrium states", II (W.-M. Ni, L. A. Peletier and J. Serrin, eds.), Math. Sci. Res. Inst. Publ., 13, Springer, New York 1988, pp. 139-162.

[Ma78] H. Matano, *Convergence of solutions of one dimensional parabolic equations*, J. Math. Kyoto Univ., 18 (1978), no. 2, 221-227.

[Ma82] H. Matano, *Nonincrease of the lap-number of a solution for a one-dimensional semi-linear parabolic equation*, J. Fac. Sci. Univ. Tokyo Sect. IA Math., 29 (1982), no. 2, 401-441.

[Ni62] K. Nickel, *Gestaltsaussagen über Lösungen parabolischer Differentialgleichungen*, J. Reine Angew. Math., 211 (1962), 78-94.

[Ro91] C. Rocha, *Properties of the attractor of a scalar parabolic PDE*, J. Dyn. Diff. Eq., 3 (1991), no. 4, 575-591.

[Sm83] J. SMOLLER, *"Shock Waves and Reaction-Diffusion Equations"*, Springer-Verlag, New York-Berlin, 1983.
[St36] C. STURM, *Sur une classe d'équation à différences partielles*, J. Math. Pure. Appl., 1 (1836), 373-444.
[Wo02] M. WOLFRUM, *A sequence of order relations; encoding heteroclinic connections in scalar parabolic PDE*, J. Diff. Eq., 183 (2002), 56-78.

Yasuhito Miyamoto
Reseach Institute for Electronic Science
Hokkaido University
Kita 12 Nishi 6, kita-ku,
Sapporo 060-0812, Japan
miyamoto@nsc.es.hokudai.ac.jp

Some Remarks on Morphisms between Fano Threefolds

Ekaterina Amerik

Received: June 7, 2004
Revised: October 11, 2004

Communicated by Thomas Peternell

Abstract. Let X, Y be Fano threefolds of Picard number one and such that the ample generators of Picard groups are very ample. Let X be of index one and Y be of index two. It is shown that the only morphisms from X to Y are double coverings. In fact nearly the whole paper is the analysis of the case where Y is the linear section of the Grassmannian G(1,4), since the other cases were more or less solved in another article. This remaining case is treated with the help of Debarre's connectedness theorem for inverse images of Schubert cycles.

2000 Mathematics Subject Classification: 14J45
Keywords and Phrases: Fano threefolds, connectedness

Some twenty-five years ago, Iskovskih classified the smooth complex Fano threefolds with Picard number one. Apart from $\mathbb{P}^3$ and the quadric, his list includes 5 families of Fano varieties of index two and 11 families of varieties of index one (for index one threefolds, the cube of the anticanonical divisor takes all even values from 2 to 22, except 20). Recently, the author ([A]) and C. Schuhmann ([S]) made some efforts to classify the morphisms between such Fano threefolds, the starting point being a question of Peternell: let $f : X \to Y$ be a non-trivial morphism between Fano varieties with Picard number one, is it then true that the index of X does not exceed the index of Y?
In particular, Schuhmann ([S]) proved that there are no morphisms from index-two to index-one threefolds, and that any morphism between index-two threefolds is an isomorphism (under certain additional hypotheses, some of which were handled later in [A], [IS]). As for morphisms from index-one to index-two Fano threefolds, such morphisms do exist: an index-two threefold has a double covering (branched along an anticanonical divisor) which is of index one. It is

therefore natural to ask if every morphism from index-one Fano threefold X with Picard number one to index-two Fano threefold Y with Picard number one is a double covering. In [A], I proved a theorem (Theorem 3.1) indicating that the answer should be yes, however not settling the question completely. The essential problem was that the methods of [A] would never work for $Y = V_5$, the linear section of the Grassmannian $G(1,4)$ in the Plücker embedding (all smooth three-dimensional linear sections of $G(1,4)$ are isomorphic). Though there are several ways to obtain bounds for the degree of a morphism between Fano threefolds with second Betti number one ([HM], [A]), these bounds are still too rough for our purpose.
This paper is an attempt to handle this problem. The main result is the following

Theorem *Let X be a smooth complex Fano threefold of index one and such that $Pic(X) = \mathbb{Z}$. Suppose moreover that X is anticanonically embedded. Let $f : X \to V_5$ be a non-trivial morphism. Then X is of degree 10 ("X is of type V_{10}") and f is a double covering. In other words, X is a hyperquadric section of a cone over V_5 in $\mathbb{P}^7$.*

I believe that the extra assumption made on X is purely technical and can be ruled out if one refines the arguments below. This assumption excludes two families of Fano threefolds: sextic double solids and double coverings of the quadric branched along a hyperquartic section.
Together with Theorem 3.1 of [A] and a few remarks, this theorem implies that any morphism from an index-one to an index-two threefold with cyclic Picard group is a double covering, at least under an additional assumption of the very ampleness of the generator of the Picard group of the two threefolds (see Theorem 4.1 of Section 4).
The proof, somewhat unexpectedly, relies on a connectedness theorem due to Debarre, which enables one to show that the inverse images of certain special lines on V_5 must be connected; at the same time it is well-known that, if $f : X \to V_5$ is as in our theorem, then the inverse image of a general line on V_5 is a disjoint union of conics. Starting from this, we use some Hilbert scheme manipulations to show that the connected inverse images must have very special properties, and deduce the theorem.
A smooth anticanonically embedded Fano threefold of index one and Picard number one is sometimes called *a prime Fano threefold*. We shall also call it thus throughout this paper.

Acknowledgements: It is a pleasure to thank A. Van de Ven, with whom I had some helpful discussions at the early stage of this work. I am very grateful to L. Gruson for his preliminary reading of this text and for reassuring me on certain points; and also to F. Campana for moral support.

1. Preliminaries: the geometry of V_5

Let us recall some more or less classical facts on the threefold $V_5 \subset \mathbb{P}^6$, most of which can be found in [I] or [FN]. First of all, as any Fano threefold of index two and Picard number 1, it has a two-dimensional family of lines. A general line has trivial normal bundle (call it a $(0,0)$-line), whereas there is a one-dimensional subfamily of lines with normal bundle $\mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(1)$ (call them $(-1,1)$-lines). The Hilbert scheme of lines on V_5 is isomorphic to $\mathbb{P}^2$, the curve of $(-1,1)$-lines is a conic in this $\mathbb{P}^2$, and there are 3 lines through a general point of V_5. More precisely, the $(-1,1)$-lines form the tangent surface D to a rational normal sextic B on V_5 (in particular, they never intersect), and there are three lines through any point away from D, two lines through a point on D but not on B, and one line through a point of B. The surface D is of degree 10, thus a hyperquadric section of V_5.
We shall denote by U resp. Q the restriction to V_5 of the universal bundle U_G resp. the universal quotient bundle Q_G on the Grassmannian $G(1,4)$. The cohomology groups related to those bundles are computed starting from the cohomologies of vector bundles on the Grassmannian. In particular the bundles U and Q remain stable.
We shall also use the following result from [S]: let X be a prime Fano threelold, and let $f : X \to V_5$ be a finite morphism. Let m be such that $f^*\mathcal{O}_{V_5}(1) = \mathcal{O}_X(m)$. Then the inverse image of a general line consists of $\frac{m^2 deg(X)}{10}$ disjoint conics; in general, if one replaces V_5 by another Fano threefold Y of index two with Picard number one, the inverse image of a general line shall consist of $\frac{m^2 deg(X)}{2deg(Y)}$ disjoint conics. Here by $deg(Y)$ we mean the self-intersection number of the ample generator of $Pic(Y)$.
Our starting point is the following fact, which shall be proved in the end of this section:

Proposition 1.1 *For l a $(-1,1)$-line on V_5, any irreducible projective variety X and a surjective morphism $f : X \to V_5$, $f^{-1}(l)$ is connected.*

To prove this proposition, we shall need some further details on the geometry of V_5.
Remark that the Schubert cycles of type $\sigma_{1,1}$, which are sets of points of $G(1,4)$ corresponding to lines lying in a fixed hyperplane, and are also caracterized as zero-loci of sections of the bundle dual to the universal, are 4-dimensional quadrics in the Plücker $G(1,4)$, so each of them intersects V_5 along a conic. Conversely, every smooth conic on V_5 is an intersection with such a Schubert cycle. Indeed, every conics on a Grassmannian is obviously contained in some $G(1,3)$; and if this conic is strictly contained in $G(1,3) \cap V_5$, then $G(1,3) \cap V_5$ is a surface, so the bundle U^* has a section vanishing along a surface; but this contradicts the stability of U^*.
The same is (by the same argument) true for pairs of intersecting lines on

V_5. Moreover, the correspondence between the Schubert cycles and the conics is one-to-one (it is induced by the restriction map on the global sections $H^0(G(1,4),U_G^*) \to H^0(V_5,U^*)$ which is an isomorphism).
Let us show that among these conics, there is a one-dimensional family of double lines.

Proposition 1.2 *Fix an embedding $V_5 \subset G(1,4) \subset \mathbb{P}^9$. There is a one-dimensional family of Schubert cycles Σ_t such that for each t, the intersection of V_5 and Σ_t is a (double) line. Moreover, lines on V_5 which are obtained as a set-theoretic intersection with a Schubert cycle of type $\sigma_{1,1}$, are exactly $(-1,1)$-lines.*

Proof: The three-dimensional linear sections of $G(1,4)$ in the Plücker embedding are parametrized by the Grassmann variety $G(6,9)$; let, for $P \in G(6,9)$, V_P denote the intersection of $G(1,4)$ with the corresponding linear subspace (which we will denote also by P). The Schubert cycles are parametrized by $G(3,4)=\mathbb{P}^4$; likewise, denote by Σ_t the Schubert cycle corresponding to $t \in \mathbb{P}^4$. Consider the following incidence subvariety $\mathcal{I} \subset G(6,9)\times\mathbb{P}^4$:

$$\mathcal{I} = \{(P,t) \in G(6,9)\times\mathbb{P}^4 | V_P \cap \Sigma_t \ \textit{is a line}\}.$$

The fiber $\mathcal{I}_t$ of $\mathcal{I}$ over any $t\in\mathbb{P}^4$ parametrizes the six-dimensional subspaces P of $\mathbb{P}^9$ intersecting Σ_t along a line. Σ_t is a quadric in $\mathbb{P}^5 \subset \mathbb{P}^9$, and P intersects it along a line l if and only if the plane $H = P\cap\mathbb{P}^5$ is tangent to Σ_t along l, i.e. lies in every $T_x\Sigma_t, x \in l$. The intersection of all tangent spaces to $\Sigma_t \subset \mathbb{P}^5$ along l is a three-dimensional projective space (the tangent spaces form a pencil of hyperplanes in $\mathbb{P}^5$, because Σ_t is a quadric). This means that for every l, the planes tangent to Σ_t along l form a one-dimensional family. The family of lines on a 4-dimensional quadric ($=G(1,3)$) is a 5-dimensional flag variety, so the planes in $\mathbb{P}^5$ tangent to Σ_t along a line are parametrized by a six-dimensional irreducible variety (a $\mathbb{P}^1$-bundle over a flag variety). This implies that $\mathcal{I}_t$ is irreducible of codimension 3 in $G(6,9)$, so $\mathcal{I}$ is irreducible of codimension 3 in $G(6,9)\times\mathbb{P}^4$.
We must show that the first projection $p_1 : \mathcal{I} \to G(6,9)$ is surjective and its general fiber is of dimension one. First of all, remark that there are points P in the image of p_1 such that the corresponding V_P is smooth (so, is a V_5). Indeed, fix, as above, Σ_t, $l \subset \Sigma_t$, H a plane in $\mathbb{P}^5 =< \Sigma_t >$ such that $H \cap \Sigma_t = l$; the remark will follow if we show that for a general $\mathbb{P}^6 = P \subset \mathbb{P}^9$ containing H, $G(1,4)\cap P$ is smooth. We have $H\cap G(1,4) = H \cap \Sigma_t = l$ (because $G(1,4)\cap < \Sigma_t >= \Sigma_t$), so the smoothness away from l is obvious, and one checks, again by standart dimension count, that for $x \in l$, the set $A_x = \{P|H \subset P, G(1,4)\cap P \ \textit{is singular at}\ x\}$ is of codimension two in the space of all P's containing H. Therefore for P general in the image of p_1, V_P is smooth.
It is clear that if a smooth $V_P = G(1,4)\cap P$ is such that $V_P \cap \Sigma_t = l$, then the corresponding plane H is tangent along l not only to Σ_t, but also to V_P.

Thus the normal bundle N_{l,V_P} has a subbundle $N_{l,H}$ of degree 1, and so l is of type $(-1,1)$ on V_P. Since we have only one-dimensional family of $(-1,1)$-lines on a smooth V_P, we deduce that a fiber of p_1 over a point P such that V_P is smooth, is at most one-dimensional. The irreducibility of $\mathcal{I}$ now implies that p_1 is surjective and its general fiber is of dimension one. This proves the Proposition.

Proof of Proposition 1.1:

Let us recall the following result of Debarre ([D], partial case of Théorème 8.1, Exemple 8.2 (3)):
Let X be an irreducible projective variety, and let $f : X \to G(d,n)$ be a morphism. Let Σ be a Schubert cycle of type σ_m. If in the cohomologies of $G(d,n)$, $[f(X)] \cdot \sigma_{m+1} \neq 0$, then $f^{-1}(\Sigma)$ is connected.
Let X be an irreducible projective variety and $f : X \to V_5$ be a surjective morphism. Composing with the embedding $i : V_5 \to G(1,4)$, we can view f as a morphism to $G(1,4)$. By Proposition 1.2, each $(-1,1)$-line is the intersection of $f(X)$ with a Schubert cycle of type $\sigma_{1,1}$ on our grassmannian $G(1,4) = G(1,\mathbb{P}(U))$, where U is a five-dimensional vector space. By duality, we can view this Grassmannian as $G(2,\mathbb{P}(U^*))$, and this point of view transforms the Schubert cycles of type $\sigma_{1,1}$ into Schubert cycles of type σ_2. The condition $[f(X)] \cdot \sigma_3 \neq 0$ is obviously satisfied because $f(X) = V_5$ is cut out by three hyperplanes. Thus Proposition 1.1 follows from Proposition 1.2 and Debarre's theorem.

Remark 1.3 If we knew that the inverse image of a *general* line is always connected, this would immediately solve our problem; indeed, for a Fano threefold X of index and Picard number one, the equality $\frac{m^2 deg(X)}{10} = 1$ implies that $m = 1$, $deg(X) = 10$ and f is a double covering. However, as shows an example of Peternell and Sommese, this is false in general, even if one supposes that X is a Fano threefold. In the example of [PS], X is the universal family of lines on V_5, which turns out to be a Fano threefold (of Picard number two, of course), and f is the natural triple covering. The inverse image of a general line has two connected components.

Remark 1.4 One can ask if there is a similar connectedness statement for other Fano threefolds of Picard number one and index two. Recall that these are the following: intersection of two quadrics in $\mathbb{P}^5$; cubic in $\mathbb{P}^4$; double covering of $\mathbb{P}^3$ branched in a quartic; double covering of the cone over Veronese surface branched in a hypercubic section.
Smooth quadrics in $\mathbb{P}^5$ are Grassmannians $G(1,3)$, and a smooth intersection of two quadrics in $\mathbb{P}^5$ is a quadric line complex. It is classically known (see [GH], Chapter 6) that on a quadric line complex, there is a finite (and non-zero) number of lines obtained as set-theoretic intersection with a plane in $G(1,3)$. These lines are obviously $(-1,1)$-lines, since the corresponding plane is tangent

to the quadric line complex along this line. Our intersection of two quadrics is contained in a pencil of such Grassmannians, so there is a one-dimensional family of lines on it such that each line is the intersection with a plane lying on some Grassmannian of the pencil. The curve of $(-1,1)$-lines is irreducible (it follows from the results in [GH], Chapter 6, that it is smooth and that it is an ample divisor on the Fano surface of lines, in particular, it is connected). Thus it is just the closure of that family. So that it follows again from Debarre's paper that the inverse image of a general $(-1,1)$-line is connected.
As for the cubic, even if such a connectedness statement could hold, it would not, as far as I see, follow from any well-known general result. One can, though, remark that in the examples of Peternell-Sommese type "(universal family of lines on Y)$\rightarrow$ Y", the inverse image of a $(-1,1)$-line has a tendency to be connected, whereas the inverse image of a $(0,0)$-line is certainly not connected. Indeed, it is observed in the literature that, on the threefolds as above (the cubic, the quadric line complex, V_5), a line l is in the closure of the curve C_l =\{*lines intersecting l but different from l*\} on the Hilbert scheme if and only if l is a $(-1,1)$-line.

2. A Hilbert scheme argument

The previous considerations show that on our Fano threefold X, a disjoint union of conics degenerates flatly to a connected l.c.i. scheme. Recall the following classical example: if one degenerates a disjoint union of two lines in the projective space into a pair of intersecting lines, the pair of intersecting lines shall have an embedded point at the intersection. So if one wants the limit to be a connected l.c.i., this limit must be a double line. This suggests to ask if a similar phenomenon can occur in our situation, that is: can it be true that a connected l.c.i. limit of disjoint conics is necessarily a multiple conic?
In any case it is easily checked that, say, a connected limit of pairs of disjoint conics does not have to have embedded points when the two conics become reducible and acquire a common component. So this is very probably false, and in any case there is no simple argument. In this paragraph we shall prove, though, that the inverse image of a sufficiently general $(-1,1)$-line is either a multiple conic, or supported on a union of lines, and in fact even slightly more (Proposition 2.5).
Let T be the Hilbert scheme of lines on V_5 and let $\mathcal{M} \subset T \times V_5$ be the universal family. We have the "universal family of the inverse images of lines under f"

$$\mathcal{S} = \mathcal{M} \times_{V_5} X \subset T \times X.$$

Since f is flat and $\mathcal{M}$ is flat over T, $\mathcal{S}$ is flat over T.
Let H' be the Hilbert scheme of conics on X. Consider the irreducible components of H' which are relevant for our problem, that is, the components such that their sufficiently general points correspond to conics which are in the inverse image of a sufficiently general line on V_5. Denote by H the union of all such components (each of them is, of course, two-dimensional).

For every point $x \in H$, the image of the corresponding conic C_x is a line. Indeed, "$f(C)$ is a line" is a closed condition on conics C because f is a finite morphism (for f arbitrary, "$f(C)$ is contained in a line" would be a closed condition on C).
This allows to construct a morphism $p : H \to T$ taking every conic to its image under f. Indeed,

$$\mathcal{L} = \{(C, f(x))|x \in C, C \in H\} \subset H \times V_5$$

is a family of lines over H; though apriori it is not clear that it is flat, this is a "well-defined family of algebraic cycles" in the sense of Kollar ([K], Chapter I) and so corresponds to a morphism from H to the Chow variety of lines on V_5, and this is the same as T.
We claim that p is finite. Indeed, it is clear that the only obstruction to the finiteness of p could be the existence of infinitely many double structures of arithmetic genus zero on some lines on X ("non-finiteness of the Hilbert-Chow morphism for the family of conics on X"). This obviously happens if one considers conics in $\mathbb{P}^3$ rather than conics on X. In our situation, however, this is impossible, and the Hilbert-Chow morphism is even one-to-one. Indeed, by [I], the normal bundle of a line in a prime Fano threefold is either $\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$, or $\mathcal{O}_{\mathbb{P}^1}(1) \oplus \mathcal{O}_{\mathbb{P}^1}(-2)$, and there is the following

Lemma 2.1 *Let $l \subset X$ be a line on a prime Fano threefold. If $N_{l,X} = \mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$, then there is no locally Cohen-Macaulay double structure of arithmetic genus 0 on l. If $N_{l,X} = \mathcal{O}_{\mathbb{P}^1}(1) \oplus \mathcal{O}_{\mathbb{P}^1}(-2)$, then such a structure is unique.*

Proof: All locally Cohen-Macaulay double structures on smooth curves in a threefold are obtained by a construction due to Ferrand (see for example [BF], or else [N] for details): if $Y \subset V$ is a smooth curve on a smooth threefold, and $\tilde{Y}$ is a double structure on Y, write L for $\mathcal{I}_Y/\mathcal{I}_{\tilde{Y}}$; in fact L is a locally free rang-one $\mathcal{O}_Y$-module and $\mathcal{I}_{\tilde{Y}}$ contains $\mathcal{I}_Y^2$. The double structure is thus determined by the natural surjection from the conormal bundle of Y in V to L, up to a scalar. Now take $Y = l$, $V = X$ and let L be as above; we have an exact sequence

$$0 \to L \to \mathcal{O}_{\tilde{l}} \to \mathcal{O}_l \to 0,$$

from which it is clear that $p_a(\tilde{l}) = 0$ if and only if $L = \mathcal{O}_{\mathbb{P}^1}(-1)$. Now in the first part of our assertion, there is no non-trivial surjection from $N^*_{l,X}$ to $\mathcal{O}_{\mathbb{P}^1}(-1)$, and in the second part, such a surjection is unique up to a scalar.

Note that we do not have to consider curves which are not locally Cohen-Macaulay, since, for example, the above argument shows that there are no higher genus locally Cohen-Macaulay double structures, and an embedded point decreases the genus.

Thus, for any $t \in T$, $p^{-1}(t)$ is a finite set $\{h_1, \dots, h_k\}$, and to each h_i there corresponds one conic C_i on X, mapped to l_t by f. The next step is to show that f and p "agree with each other":

Lemma 2.2 *Let $t \in T$ be any point and $l_t \in V_5$ be the corresponding line. Let $h_1, \dots, h_k$ be the points of $p^{-1}(t)$ and $C_1, \dots, C_k$ the corresponding conics on X. Then the support of $f^{-1}(l_t)$ is $\bigcup_i C_i$.*

Proof: Indeed, for a general $t \in T$, it is true: $f^{-1}(l_t) = \bigcup_i C_i$. For a special $t \in T$, choose a curve $V \subset T$ through t, such that t is the only "non-general" point of V in the above sense, and let $U = p^{-1}(V)$. Denote by $\mathcal{C}_U \subset U \times X$ the restriction to U of the universal family of conics over H. The support of the fiber over t of $(p \times id)(\mathcal{C}_U) \subset V \times X$ is equal to $\bigcup_i C_i$. But the family $\mathcal{S}|_V$ coincides with $(p \times id)(\mathcal{C}_U)$ except at t. $\mathcal{S}|_V$ being flat, it must be the scheme-theoretic closure of $(p \times id)(\mathcal{C}_U)|_{V-\{t\}}$ in $V \times X$, and thus the support of $\mathcal{S}|_V$ is $(p \times id)(\mathcal{C}_U)$, q.e.d.

Let now $t \in T$ be a point corresponding to a sufficiently general $(-1,1)$-line. We know that $f^{-1}(l_t)$ is connected. Suppose that the number k from the Lemma is > 1, so that there are several conics in the $Supp(f^{-1}(l_t))$. Decompose the set of those conics into two disjoint non-empty subsets Σ_1 and Σ_2.

Proposition 2.3 *There exists a conic in Σ_1 which has a common component with a conic in Σ_2; in other words, $(\bigcup_{C \in \Sigma_1} C) \bigcap (\bigcup_{C \in \Sigma_2} C)$ cannot be zero-dimensional.*

Proof Choose a suitable small 1-dimensional disc $(V, 0)$ centered at t. The inverse image $p^{-1}V$ is a disjoint union of two analytic sets U_1 and U_2 (U_i consists of points corresponding to conics near those of Σ_i). Repeat the procedure of the previous lemma: consider the universal families $\mathcal{C}_i$ of conics over U_i and their images $\mathcal{S}_i = (p \times id)(\mathcal{C}_i) \subset V \times X$. Let $\mathcal{S}^0$, $\mathcal{S}_i^0$ denote the restriction of our families $\mathcal{S}$, $\mathcal{S}_i$ to the punctured disc $V^0 = V - \{0\}$. The family $\mathcal{S}^0$ is just the disjoint union of $\mathcal{S}_i^0$. Now take the closure of all those (as analytic spaces) in $V \times X$: the closure of $\mathcal{S}^0$ is just $\mathcal{S}|_V$, by flatness, and the closure $\mathcal{S}_i'$ of $\mathcal{S}_i^0$ has the same support as $\mathcal{S}_i$, is contained in $\mathcal{S}|_V$ and is flat over V. The fiber of $\mathcal{S}_i'$ over 0, denoted S_i, is contained in the fiber S of $\mathcal{S}$, since the tensor multiplication preserves the surjectivity. So $f^{-1}(l_t) = S$ contains $S_1 \cup S_2$. By construction, S_i are flat limits of disjoint unions of a_i conics and S is a flat limit of disjoint unions of $a_1 + a_2$ $(= \frac{m^2 deg(X)}{10})$ conics.
If S_1 and S_2 do not have common components, then, since by flatness $deg(S) = deg(S_1) + deg(S_2)$, this implies $S = S_1 \cup S_2$, because S is purely one-dimensional (being an inverse image of a line under a finite morphism). But then we can apply the exact sequence

$$0 \to \mathcal{O}_S \to \mathcal{O}_{S_1} \oplus \mathcal{O}_{S_2} \to \mathcal{O}_{S_1 \cap S_2} \to 0$$

and get a contradiction, since by flatness $\chi(\mathcal{O}_S) = \chi(\mathcal{O}_{S_1}) + \chi(\mathcal{O}_{S_2})$, $S_1 \cap S_2$ is non-empty and it is zero-dimensional by assumption. Thus S_1 and S_2 must have common components, and, as S_i is supported on $\bigcup_{C\in\Sigma_i} C$, the Proposition is proved.

Corollary 2.4 *In the situation as above, $f^{-1}(l_t)$ is supported either on a single conic, or on a union of lines.*

Indeed, the proposition shows that if $f^{-1}(l_t)$ contains more than one conic, then any conic from $f^{-1}(l_t)$ must have a common component with the rest of these conics, that is, it must be singular.

Some results from commutative algebra allow to prove a stronger ("local") version of Proposition 2.3:

Proposition 2.5 *In the situation of Proposition 2.3, through each intersection point P of $\bigcup_{C\in\Sigma_1} C$ and $\bigcup_{C\in\Sigma_2} C$ passes some common component of $\bigcup_{C\in\Sigma_1} C$ and $\bigcup_{C\in\Sigma_2} C$.*

Proof: The family $\mathcal{S}$ is flat over T which is smooth, and the fibers are l.c.i., thus locally Cohen-Macaulay. It follows ([EGA], 6.3.1, 6.3.5) that $\mathcal{S}$ is locally Cohen-Macaulay, and that the same it true for the restriction of $\mathcal{S}$ to any smooth curve in T. Suppose that Proposition 2.5 is not true for some intersection point P. Let $x = (t, P) \in T \times X$ be the point corresponding to P in $\mathcal{S}$. Consider the restriction of $\mathcal{S}$ to a general curve through t, and an analytic neighbourhood of x in this restriction. Clearly, if one removes x, this neighbourhood becomes disconnected: there are at least two branches corresponding to $Supp\mathcal{S}_i$ as in Proposition 2.3. But this is impossible by Hartshorne's connectedness ([H]), which implies that a connected Cohen-Macaulay neighbourhood remains connected if one removes a subvariety of codimension at least two.

Remark 2.6 The argument of the Proposition is more or less the following: "if we have a disjoint union of certain smooth curves A and B, which degenerates flatly into a certain connected C in such a way that A and B do not acquire common components in the limit, then C will have embedded points at the intersection points of the limits of A and B, so this is impossible if we know that C is purely one-dimensional". Examples show that one cannot say anything reasonable if one allows A and B to acquire common components. But in fact our "C", that is, $f^{-1}(l_t)$, is more than just purely one-dimensional: it is a locally complete intersection. I do not know if its being a flat limit of disjoint unions of conics can impose stronger restrictions on its geometry.

To illustrate how we shall apply this, let us handle the case when $f^{-1}(l_t)$ is supported on a single conic.

Proposition 2.7 *In this case $X = V_{10}$ and f is a double covering.*

Proof: As the degree of the subscheme $f^{-1}(l_t)$ of X is $\frac{m^2deg(X)}{5}$, this conic is of multiplicity $\frac{m^2deg(X)}{10}$ in $f^{-1}(l_t)$. That is, the local degree of f near a general point of such a conic is also $\frac{m^2deg(X)}{10}$. Now this is the local degree of f along a certain divisor, because we have chosen the line l_t to be "sufficiently general among the $(-1,1)$ lines": it varies in a one-dimensional family. This divisor is thus a component of the ramification divisor of f, and $\frac{m^2deg(X)}{10} - 1$ is its ramification multiplicity.
Now the ramification divisor of f is an element of $|\mathcal{O}_X(2m-1)|$, and so the local degree of f at its general point is at most $2m$, and if it is $2m$, then the ramification divisor is the inverse image of the surface covered by the $(-1,1)$-lines and set-theoretically a hyperplane section of X. So we have:

$$\frac{m^2deg(X)}{10} \leq 2m,\ mdeg(X) \leq 20,$$

and if the equality holds, then f is unramified outside the inverse image of the surface of $(-1,1)$-lines. Also, $\frac{m^2deg(X)}{10}$ must be an integer. The inequality thus only holds for $deg(X) = 10$ and $m = 1$ (this is a double covering) or $m = 2$ (in this case it is an equality), and for $deg(X) = 4$ and $m = 5$ (also an equality). Let us exclude the last two cases. If f is unramified outside the inverse image of the surface of $(-1,1)$-lines, then p is $\frac{m^2deg(X)}{10}$-to-one everywhere except over the conic parametrizing the $(-1,1)$-lines on $T = \mathbb{P}^2$. It is thus a topological covering of the complement to this conic in T. But the latter is simply-connected; so that H has $\frac{m^2deg(X)}{10}$ irreducible components and each one maps one-to-one on T. Notice that the number $\frac{m^2deg(X)}{10}$ is superiour to three in both cases. But this is impossible. Indeed, on V_5 one has only 3 lines through a general point; whereas, if H has k components, each component would give at least one conic through a general point of X. Those conics are mapped to *different* lines through $f(x)$, because they intersect; thus $k \leq 3$.

3. Proof of the Theorem

We have seen that the inverse image of a general $(-1,1)$-line is supported either on one conic, or on a union of lines, and settled the first case in the end of the second section. Let us now settle the remaining case, using Proposition 2.5.
The following lemma is standart (and follows e.g. from the arguments of [M], Chapter 3):

Lemma 3.1 *Let $g : X_1 \to X_2$ be a proper morphism of complex quasiprojective varieties, which is finite of degree d. Suppose that X_2 is smooth. Then the inverse image of any point $x \in X_2$ consists of d points at most, and if there* *are exactly d points in the inverse image of all $x \in X_2$, then topologically g is a covering.*

Let H be as in the last section, and let $\mathcal{C}$ be the universal family of conics over H. Each conic of H is contained in the inverse image of some line on V_5, and set-theoretically such an inverse image is a union of conics of H. Denote by D the surface covered by $(-1,1)$-lines on V_5. Recall that through each point of the complement to D in V_5 there are three lines, that D is a tangent surface to a rational normal sextic and that there are two lines (one $(-1,1)$-line and one $(0,0)$-line) through any point of D away from this sextic and a single line through each point of the sextic. Since the inverse image of a general $(0,0)$-line is a disjoint union of conics of H, there are three conics of H through a general point of X, and at least three through any point away from $f^{-1}(D)$. The natural morphism $q : \mathcal{C} \to X$ is proper and finite of degree three. Lemma 3.1 has thus an obvious corollary:

Corollary 3.2 *There are at most three conics of H through any point of X, and exactly three conics of H through any point of X away from $f^{-1}(D)$.*

Let l be a general $(-1,1)$-line on V_5. Consider the case when $Z = f^{-1}(l)$ is a set-theoretic union of degenerate conics $C_1, ..., C_k$ of H.

Lemma 3.3 *Z contains a line which belongs to a single C_i (say C_1).*

Proof: Suppose the contrary, that is, that any component of Z is contained in at least two conics of H. Through a general point x of this component there is at least one more conic of H, coming from the inverse image of the $(0,0)$-line through $f(x)$. This implies that the morphism $q : \mathcal{C} \to X$ is three-to-one outside an algebraic subset A of codimension at least two in X. That is, $\mathcal{C} - q^{-1}(A)$ is, topologically, a covering of $X - A$. But $X - A$ is simply-connected because X is Fano and thus simply-connected. This means that $\mathcal{C}$ is reducible, consists of three components and each of them maps one-to-one to X. Since X is smooth, it must be isomorphic to each of those components (by Zariski's Main Theorem). But this is impossible because the components are fibered in conics and X has cyclic Picard group.

Before continuing our argument, let us recall some well-known facts on lines on prime Fano threefolds ([I]). Lines on our Fano threefold X are parametrized by a curve, which may of course be reducible or non-reduced. Its being reduced or not influences the geometry of the surface covered by lines on X. Namely, if a component of the Hilbert scheme of lines on X is reduced, then the natural morphism from the corresponding component of the universal family to X is an immersion along a general line; and there is a classical computation ([I], [T]) which says that if its image M is an element of $|\mathcal{O}_X(d)|$, then a general line of M intersects $d+1$ other lines of M. If a component of the Hilbert scheme

of lines is non-reduced, then the surface M covered by the corresponding lines is either a cone (but this can happen only on a quartic), or a tangent surface to a curve. One knows only one explicit example of a Fano threefold as above such that the surface covered by lines on it is a tangent surface to a curve, it is constructed by Mukai and Umemura (having been overlooked by Iskovskih) and has degree 22. The surface itself is a hyperplane section of this threefold and its lines never intersect.
The following Proposition, due to Iliev and Schuhmann, is the main result of [IS] slightly reformulated:

Proposition 3.4 *Let X be a prime Fano threefold, $\mathcal{L}$ a complete one-dimensional family of lines on X and M the surface on X covered by lines of $\mathcal{L}$. If X is different from the Mukai-Umemura threefold, then a general line of $\mathcal{L}$ intersects at least one other line of $\mathcal{L}$.*

An outline of the proof: If not, then, by what we have said above, the surface M must be a tangent surface to a curve. Studying its singularities, Iliev and Schuhmann prove that it must be a hyperplane section of X. Then they show, by case-by-case analysis (of which certain cases appear already in [A]), that the only prime Fano threefold containing a tangent surface to a curve as a hyperplane section, is the Mukai-Umemura threefold.

"Lines contained in a single C_i" cover a divisor on X as Z varies (this is the branch divisor of q). Since $(-1, 1)$-lines on V_5 never intersect, Proposition 3.4 implies that if X is not the Mukai-Umemura threefold, then in Z there are at least two lines contained in a single conic (say, $l_1 \subset C_1$ and $l_2 \subset C_2$), and that they intersect, say at the point P. Notice that C_1 is necessarily different from C_2: otherwise we get a contradiction with Proposition 2.5 by considering $\Sigma_1 = \{l_1 \cup l_2\}$, Σ_2 the set of all the other C_i and the intersection point P.

Claim 3.5 Both C_1 and C_2 are pairs of lines intersecting at the point P, and Z is supported on $C_1 \cup C_2$. Thus Z is, set-theoretically, the union of three or four lines through P.

Proof:
1) If C_1 is a double line, we get a contradiction with Proposition 2.5 by considering $\Sigma_1 = \{C_1\}$ and the point P; the same is true for C_2.
2) Let $C_1 = l_1 \cup l'_1$. If l'_1 does not pass through P, we get the contradiction in the same way, thus $P \in l'_1$. Also, $P \in l'_2$, where $C_2 = l_2 \cup l'_2$.
3)There are two possibilities:
a)If $l'_1 \neq l'_2$, then there must be another conic from Z through P, containing l'_1. Indeed, otherwise we again get a contradiction with Proposition 2.5. In the same way, there is a conic from Z through P which contains l'_2. In fact it is the same conic, because otherwise there are at least four conics through P, contradicting Corollary 3.2. Denote it by C_3. No other conic from Z passes

through P. So $C_3 = l'_1 \cup l'_2$, and l'_1, l'_2 are not contained in conics others than C_1, C_2, C_3.
b) If $l'_1 = l'_2$, then no other conic from Z contains this line (otherwise through its general point there will pass at least four conics from H, the fourth one coming from the inverse image of the correspondent $(0,0)$-line).
4) Now the union $C_1 \cup C_2 \cup C_3$ in the case a), resp. the union $C_1 \cup C_2$ in the case b), cannot have any points in common with the other components of Z; otherwise, taking $\Sigma_1 = \{C_1, C_2, C_3\}$, resp. $\Sigma_1 = \{C_1, C_2\}$, we obtain a contradiction with Proposition 2.5. But Z is connected, so Z is supported on the lines l_1, l'_1, l_2, l'_2, q.e.d.

We are now ready to finish the proof of the theorem stated in the introduction.

Proof of the theorem: If X is the Mukai-Umemura threefold, then the lines on X never intersect at all, so that $f^{-1}(l)$ must be supported on a single conic. Proposition 2.7 shows that a morphism from X to V_5 is impossible. (It should be, however, said at this point that the paper [HM] contains a better proof of the non-existence of morphisms from the Mukai-Umemura threefold onto any other smooth variety, besides $\mathbb{P}^3$!).
If X is not the Mukai-Umemura threefold and $f^{-1}(l)$ is not supported on a single conic, then we know by Claim 3.5 how $f^{-1}(l)$ looks. Remark that $f^{-1}(D)$ is a reducible divisor: it has two components, one swept out by the lines l_1 and l_2 as Z varies, another by l'_1 and l'_2 (the components are really different because, by construction, l_1 and l_2 are lines contained in a single conic of H, whereas l'_1 and l'_2 are not). Neither component is a hyperplane section: indeed, if a hyperplane section of X is covered by lines, then it is either a cone (impossible in our situation), or a general line intersects two other lines on the surface by the classical computation from [T] mentioned above, since a hyperplane section cannot be a tangent surface to a curve by [IS]. Let k be the multiplicity of the component corresponding to l_i and k' be the multiplicity of the component corresponding to l'_i. As $f^*(D)$ is a divisor from $|\mathcal{O}_X(2m)|$, $k + k' \leq m$. At the same time, Z must be of degree $\frac{m^2 deg(X)}{5}$, and thus $2k + 2k' = \frac{m^2 deg(X)}{5}$, so $m^2 deg(X) \leq 10$, leavng the only possibility $m = 1, deg(X) = 10$.

4. Concluding remarks

In this section, we shall make a further (minor) precision on Theorem 3.1 from [A].
In that theorem, it was proved that if X, Y are Fano threefolds with Picard number one and very ample generator of the Picard group, X is of index one, Y is of index two different from V_5 (that is, Y is a cubic or a quadric line complex), and $f: X \to Y$ is a surjective morphism, then f is a "projection", that is, $f^*\mathcal{O}_Y(1) = \mathcal{O}_X(1)$. The argument of the theorem also worked for Y a quartic double solid, whereas there were some problems (hopefully technical

ones) for Y a double Veronese cone and for X not anticanonically embedded. Even in the "good" cases, the theorem proves a little bit less than one would like; that is, we want f to be a double covering and we prove only that $f^*\mathcal{O}_Y(1) = \mathcal{O}_X(1)$. This still leaves the following additional possibilities:
(1) If Y is a cubic, X can be V_{12}, $deg(f) = 4$ (X cannot be V_{18} because of the Betti numbers: $b_3(V_{18}) < b_3(Y)$);
(2) If Y is an intersection of two quadrics, X can be V_{16}, $deg(f) = 4$ (here V_{12} is impossible since in this case the inverse image of a general line would consist of 3/2 conics).
The first possibility can be excluded by using an inequality of [ARV]: it says that for a finite morphism $f : X \to Y$ and a line bundle L on Y such that $\Omega_Y(L)$ is globally generated, $deg(f)c_{top}\Omega_Y(L) \leq c_{top}\Omega_X(f^*L)$, so, for X and Y of dimension three, $deg(f)(c_3(\Omega_Y) + c_2(\Omega_Y)L + c_1(\Omega_Y)L^2)$ must not exceed $c_3(\Omega_X) + c_2(\Omega_X)f^*L + c_1(\Omega_X)f^*L^2$.
Consider the situation of (1): we may take $L = \mathcal{O}_Y(2)$, and we know that $c_3(\Omega_Y) = 6$ and $c_3(\Omega_X) = 10$. Using the equalities $c_2(X)c_1(X) = c_2(Y)c_1(Y) = 24$, we arrive at $4(6 + 24 - 24) \leq 10 + 48 - 48$, which is false. So the case (1) cannot occur.
This inequality does not work in the case (2): indeed, now $c_3(\Omega_Y) = 0$, $c_3(\Omega_Y) = 2$ and the inequality reads as follows: $4(0 + 24 - 32) \leq 2 + 48 - 64$, so does not give a contradiction. However we can rule out this case by our connectedness argument. Indeed, the inverse image of a general $(-1, 1)$-line is connected (Remark 1.4) and the inverse image of a general $(0, 0)$-line consists of two disjoint conics. The results of Section 2 apply, of course, to our situation; it follows that the inverse image of a general $(-1, 1)$-line is either a double conic, or a union of two reducible conics which have a common component. In both cases, it is clear that the ramification locus of f projects *onto* the surface covered by $(-1, 1)$-lines. But the ramification divisor is a hyperplane section of V_{16}, and thus can project onto a surface from $|\mathcal{O}_Y(4)|$ at most. Whereas it is well-known (and follows for example from the results in [GH], Chapter 6) that the surface covered by $(-1, 1)$-lines on Y is an element of $|\mathcal{O}_Y(8)|$.
All this put together gives the following

Theorem 4.1 *Let X, Y be smooth complex Fano threefolds of Picard number one, X of index one, Y of index two. Assume further that the ample generators of $Pic(X)$ and $Pic(Y)$ are very ample. Then any morphism from X to Y is a double covering.*

I would like to mention that the verification of this statement without the very ampleness hypothesis amounts to a very small number of particular cases; for instance, if Y is a double Veronese cone, then already the formula of [ARV] combined with the knowledge of Betti numbers implies that for any morphism $f : X \to Y$ with X Fano of index one with cyclic Picard group, $deg(f) = 2$ and X is a sextic double solid. It seems that one could be able to work out the remaining cases without any essentially new ideas.

References

[A] E. Amerik, Maps onto certain Fano threefolds, Doc. Math. 2 (1997), 195–211.

[ARV] E. Amerik, M. Rovinsky, A. Van de Ven, A boundedness theorem for morphisms between threefolds, Ann. Inst. Fourier 49 (1999), 405-415.

[BF] C. Bănică, O. Forster, Multiplicity structures on space curves, The Lefschetz centennial conference, Part I (Mexico City, 1984), Contemp. Math., 58 (1986), 47-64.

[D] O. Debarre, Théorèmes de connexité pour les produits d'espaces projectifs et les grassmanniennes, Amer. J. Math. 118 (1996), no. 6, 1347–1367.

[EGA] A. Grothendieck, Elements de la géométrie algébrique, IV, seconde partie, Publ. Math. IHES 24 (1965).

[FN] M. Furushima, N. Nakayama: The family of lines on the Fano threefold V_5. Nagoya Math. J. 116 (1989), 111–122.

[GH] Ph. Griffiths, J. Harris, Principles of Algebraic Geometry, Wiley, 1978.

[H] R. Hartshorne, Complete intersections and connectedness, Amer. J. Math. 84 (1962), 497-508.

[HM] J.M. Hwang, N. Mok, Finite morphisms onto Fano manifolds of Picard number 1 which have rational curves with trivial normal bundles, J. Algebraic Geom. 12 (2003), no. 4, 627–651.

[I] V.A. Iskovskih, Fano threefolds I, II, Math. USSR Izv. 11 (1977), 485–527, and 12 (1978), 469-506.

[IS] A. Iliev, C. Schuhmann: Tangent scrolls in prime Fano threefolds, Kodai Math. J. 23 (2000), no.3, 411-431.

[K] J. Kollar, Rational curves on algebraic varieties, Springer-Verlag, 1996.

[M] D. Mumford, Algebraic geometry I. Complex projective varieties. Springer-Verlag, 1976.

[N] S. Nollet, The Hilbert scheme of degree three curves, Ann. Sci. ENS 30 (1997), 367-384.

[PS] Th. Peternell, A.J. Sommese, Ample vector bundles and branched coverings II, preprint math.AG/0211220.

[S] C. Schuhmann, Morphisms between Fano threefolds, J. Algebraic Geom. 8 (1999), no. 2, 221–244.

[T] A.N. Tyurin, Five lectures on three-dimensional varieties, Uspekhi Mat. Nauk 27 (1972), no. 5 (167), 3–50; translation in Russian Math. Surveys 27 (1972), no. 5, 1–53.

Ekaterina Amerik
Université Paris-Sud
Laboratoire des Mathématiques
Bâtiment 425
91405 Orsay, France
Ekaterina.Amerik@math.u-psud.fr

Projective Bundle Theorem in Homology Theories with Chern Structure

Alexander Nenashev

Received: December 17, 2003
Revised: December 5, 2004

Communicated by Ulf Rehmann

Abstract. Panin and Smirnov deduced the existence of push-forwards, along projective morphisms, in a cohomology theory with cup products, from the assumption that the theory is endowed with an extra structure called orientation. A part of their work is a proof of the Projective Bundle Theorem in cohomology based on the assumption that we have the first Chern class for line bundles. In some examples we have to consider a pair of theories, cohomology and homology, related by a cap product. It would be useful to construct transfer maps (pull-backs) along projective morphisms in homology in such a situation under similar assumptions. In this note we perform the projective bundle theorem part of this project in homology.

Keywords and Phrases: (Co)homology theory, Chern structure, projective bundle, algebraic variety

1. Introduction

Let k be a field and Sm be the category of smooth quasi-projective algebraic varieties over k. Let $\mathcal{P}$ denote the category of pairs (X, U), with $X \in Sm$ and U a Zarisky open in X, where a morphism $(X, U) \to (X', U')$ is a morphism $f : X \to X'$ in Sm such that $f(U) \subset U'$. Sm embeds into $\mathcal{P}$ by $X \mapsto (X, \emptyset)$. For any functor A defined on $\mathcal{P}$, we can compose it with this embedding and write $A(X)$ for $A(X, \emptyset)$.
For $f : (X, U) \to (X', U')$ we will denote by f_A (resp. f^A) the morphism $A(f) : A(X, U) \to A(X', U')$ (resp. $A(f) : A(X', U') \to A(X, U)$) if A is covariant (respectively, contravariant). We will call such maps push-forwards or pull-backs respectively. Note that the rule $(X, U) \mapsto (U, \emptyset)$ defines an endofunctor on $\mathcal{P}$.

Definition. *A homology theory over k with values in an abelian category $\mathcal{M}$ is a covariant functor $A. : \mathcal{P} \to \mathcal{M}$ endowed with a natural transformation $d : A.(X,U) \to A.(U)$ called the boundary homomorphism, subject to the following requirements:*
(h1) (Homotopy invariance) The arrow $p_A : A.(X \times \mathbb{A}^1) \to A.(X)$ induced by the projection $p : X \times \mathbb{A}^1 \to X$ is an isomorphism for any $X \in Sm$.
(h2) (Localization sequence) For any $(X,U) \in \mathcal{P}$, the sequence

$$\ldots \to A.(U) \to A.(X) \to A.(X,U) \xrightarrow{d} A.(U) \to A(X) \to \ldots$$

is exact.
(h3) (Nisnevich excision) Let $(X,U),(X',U') \in \mathcal{P}$, $Z = X - U$, and $Z' = X' - U'$. Then for any étale morphism $f : X' \to X$ such that $f^{-1}(Z) = Z'$ and $f : Z' \to Z$ is an isomorphism, the map $f_A : A.(X',U') \to A.(X,U)$ must be an isomorphism.

These axioms are dual to the axioms of a cohomology theory given in [PS] and [PS1].[1] The objective of [PS1] is to provide simple conditions under which one can construct transfer maps (push-forwards) along projective morphisms in a cohomology theory. This, in its turn, is a prerequisite for the proof of a very general version of the Riemann-Roch Theorem in [Pa]. All the assumptions made in [PS] and [Pa] are true for many particular cohomology theories such as, for instance, K-theory, étale cohomology, higher Chow groups, and the algebraic cobordism theory introduced by Voevodsky in [V]. We therefore get, in a very uniform way, the existence of push-forwards and the Riemann-Roch Theorem in all these theories.
In some situations we have to consider a pair of theories $(A^{\cdot}, A.)$ consisting of a cohomology and a homology theory related by a cap-product. An important example of this is given by motivic cohomology and homology introduced by Suslin and Voevodsky in [SV]. An ultimate goal in such a situation is to obtain a Poincaré duality in the sense of [PY] for the pair $(A^{\cdot}, A.)$. Among the assumptions from which the Poincaré duality is deduced in [PY], there is the assumption of existence of transfer maps in both $A^{\cdot}$ and $A.$. However, the homology part of this, i.e. the verification of existence of transfers (pull-backs) in homology is still lacking. A general objective in this context is to construct transfers along projective morphisms in a homology theory starting from simple assumptions analogous to those made in [PS] for cohomology.
The purpose of this note is to prove the Projective Bundle Theorem in homology (PBTH), which is a part of the whole program aimed towards the existence of transfer maps in homology. In Section 2 we provide definitions and state the main result (PBTH). Its proof is given in Sections 3 and 4.
A similar result was obtained independently by K. Pimenov in a slightly different framework [Pi].

[1] [PS] is a part of [PS1] which has been published already.

Acknowledgements. I wish to thank Ivan Panin for attracting my attention to the problem of existence of pull-backs in homology, for his continuous interest in my work, and for numerous useful remarks. Special thanks are due to the Alexander von Humboldt Foundation (Germany) for sponsoring my stay in Bielefeld in May 2003, where the very initial steps of the project were undertaken. I am grateful to the University of Regina (Canada) where the work was completed.

2. Definitions and the Main Result

Let $A.$ be a homology theory satisfying (h1-h3) and let $A^{\cdot}$ be a cohomology theory in the sense of [PS, Def. 2.0.1]. The latter means that $A^{\cdot}$ is a contravariant functor $\mathcal{P} \to \mathcal{M}$ equiped with a natural transformation $d : A^{\cdot}(U) \to A^{\cdot}(X, U)$ and satisfying the dual set of axioms that we will call (c1-c3). All the general properties of a cohomology theory deduced from (c1-c3) in [PS, Sect. 2.2] have their duals for a homology theory, obtained by inverting the arrows. In particular, the Mayer-Vietoris exact sequence in homology and the localization sequence for a triple can be deduced from (h1-h3).

We will use the "(co)homology with support" notation $A_Z(X) = A(X, U)$, where $Z = X - U$, for both $A.$ and $A^{\cdot}$. For simplicity, we will assume that $A.$ and $A^{\cdot}$ take their values in the category $\mathcal{A}b$ of abelian groups. From now on we will often write just A for the homology groups, while keeping the upper dot in the cohomology notation.

2.1. Product structures. We will assume that $A^{\cdot}$ is a ring cohomology theory in the sense of [PS, Sect. 2.4]. This, in particular, means that $A^{\cdot}$ is equiped with cup-products

$$\cup : A^{\cdot}_Z(X) \times A^{\cdot}_{Z'}(X) \to A^{\cdot}_{Z\cap Z'}(X)$$

that are functorial with respect to pull-backs and satisfy the following properties:

(cup1)(associativity) $(a \cup b) \cup c = a \cup (b \cup c)$ in $A^{\cdot}_{Z_1\cap Z_2\cap Z_3}(X)$ for any $a \in A^{\cdot}_{Z_1}(X)$, $b \in A^{\cdot}_{Z_2}(X), c \in A^{\cdot}_{Z_3}(X)$.

(cup2) The absolute cohomology groups $A^{\cdot}(X)$ become associative unitary rings; the pull-back maps $f^A : A^{\cdot}(X) \to A^{\cdot}(Y)$ are homomorphisms of such rings for all $f : Y \to X$.

(cup3) The groups $A^{\cdot}_Z(X)$ become two-sided unitary modules over $A^{\cdot}(X)$ for all X and closed $Z \subset X$.

We say that $a \in A^{\cdot}_Z(X)$ is a central element if $a \cup b = b \cup a$ for any $b \in A^{\cdot}(X)$. We say that a is universally central if $f^A(a) \in A^{\cdot}_{Z'}(X')$ is central for any $f : (X', X'-Z') \to (X, X-Z)$ in $\mathcal{P}$. Note that the notion of a ring cohomology theory also requires compatibility of cup-products with boundary maps, which implies compatibility of cup-products with Mayer-Vietoris arguments, etc.

We will also assume that A is a left unitary module over $A^{\cdot}$ in the sense that we have cap-products

$$\cap : A^{\cdot}_Z(X) \times A_{Z\cap Z'}(X) \to A_{Z'}(X)$$

satisfying the properties:
(cap1) $(a \cup b) \cap c = a \cap (b \cap c)$ in $A_{Z_3}(X)$ for any $a \in A^{\cdot}_{Z_1}(X)$, $b \in A^{\cdot}_{Z_2}(X)$, $c \in A_{Z_1 \cap Z_2 \cap Z_3}(X)$.
(cap2) $1 \cap a = a$ whenever defined.
(cap3) Let U and U' (resp. V and V') be Zarisky opens in X (resp. in Y). Let $Z = X - U$, $Z' = X - U'$, $T = Y - V$, $T' = Y - V'$. Then for any $f : (Y, V, V') \to (X, U, U')$ and any $a \in A^{\cdot}_Z(X)$, $b \in A_{T \cap T'}(Y)$, we have $f_A(f^A(a) \cap b) = a \cap f_A(b)$ in $A_{Z'}(X)$.
(cap4) (compatibility with boundary maps)

Chern structure. We will assume that $A^{\cdot}$ is equiped with a Chern structure in the sense of [PS, Def. 3.2.1], i.e., to any $X \in Sm$ and any line bundle L over X there is assigned a universally central element $c(L) \in A^{\cdot}(X)$ called the (first) Chern class of L, subject to the requirements:
(ch1) Functoriality with respect to pull-backs; $c(L) = c(L')$ if $L \cong L'$ over X.
(ch2) $c(\mathbf{1}_X) = 0 \in A^{\cdot}(X)$, where $\mathbf{1}_X$ denotes the trivial line bundle $X \times \mathbb{A}^1$ over X, for any X.
(ch3) For any $X \in Sm$, let $\xi = c(\mathcal{O}_{X \times \mathbb{P}^1}(-1)) \in A^{\cdot}(X \times \mathbb{P}^1)$, where $\mathcal{O}_{X \times \mathbb{P}^1}(-1) = p^*(\mathcal{O}_{\mathbb{P}^1}(-1))$, $\mathcal{O}_{\mathbb{P}^1}(-1)$ denotes the tautological line bundle over $\mathbb{P}^1$, and $p : X \times \mathbb{P}^1 \to \mathbb{P}^1$ is the projection. Define the maps $f^{(0)}, f^{(1)} : A^{\cdot}(X) \to A^{\cdot}(X \times \mathbb{P}^1)$ by $f^{(0)} = p^A$ and $f^{(1)} = (\xi \cup -) \circ p^A$. Then the map $(f^{(0)}, f^{(1)}) : A^{\cdot}(X) \oplus A^{\cdot}(X) \to A^{\cdot}(X \times \mathbb{P}^1)$ is an isomorphism.
In the homology, define the maps $f_0, f_1 : A.(X \times \mathbb{P}^1) \to A.(X)$ by $f_0 = p_A$ and $f_1 = p_A \circ (\xi \cap -)$. We will say that we have an *extended Chern structure* (extended to homology) if
(ch4) The map $(f_0, f_1) : A.(X \times \mathbb{P}^1) \to A.(X) \oplus A.(X)$ is an isomorphism for any $X \in Sm$.
The axioms (ch3) and (ch4) can be considered as a dim $= 1$ case of the PBTC and PBTH accordingly. Our goal is to show that the extended Chern structure on $(A^{\cdot}, A.)$ implies the following general version of PBTH for A.

Projective Bundle Theorem. *Let X be a smooth quasiprojective variety over k and E a vector bundle over X of rank $n+1$. Assume that the pair of theories $(A^{\cdot}, A.)$ is endowed with a product structure and an extended Chern structure. Denote $\mathbb{P}(E)$ the projectivisation of E, $\mathcal{O}(-1)$ the tautological line bundle over $\mathbb{P}(E)$, and let $\xi = c(\mathcal{O}(-1)) \in A^{\cdot}(\mathbb{P}(E))$ be its Chern class. For $0 \le i \le n$, denote $f_i = f_{n,i}$ the composite map $A.(\mathbb{P}(E)) \xrightarrow{\xi^i \cap -} A.(\mathbb{P}(E)) \xrightarrow{p_A} A.(X)$, where $p : \mathbb{P}(E) \to X$ is the natural projection. Then the map*

$$F_n := (f_0, f_1, \dots, f_n) : A.(\mathbb{P}(E)) \to A.(X) \oplus A.(X) \oplus \dots \oplus A.(X)$$

is an isomorphism.

A crucial reason for which we cannot consider the theory $A.$ separately and must rather work with the pair $(A^{\cdot}, A.)$ is that ξ lives in the cohomology. However, everything works smoothly along the same guidelines as in [PS, Sect. 3.3].

3. Proof: Part I

Localizing and applying the Mayer-Vietoris, we reduce the situation to the case of a trivial bundle $E \cong X \times \mathbb{A}^{n+1}$, $\mathbb{P}(E) \cong X \times \mathbb{P}^n$. Next we can reduce it to the case $X = pt$. We leave it to the reader to check that $X \times -$ can be inserted throughout the proof. Thus we want to prove that the map

$$(f_0, \dots, f_n) : A.(\mathbb{P}^n) \to A.(pt) \oplus \dots \oplus A.(pt)$$

is an isomorphism.

We proceed by induction on n. Choose homogeneous coordinates $[x_0 : \dots : x_n]$ in $\mathbb{P}^n$ and introduce the following notation:

(i) $0 = [1 : 0 : \dots : 0]$ the distinguished point;

(ii) for $0 \le i \le n$, $\mathbb{P}_i^{n-1}$ is the projective hyperplane $x_i = 0$;

(iii) for $1 \le i \le n$, $\mathbb{P}_i^1$ is the projective axis on which all $x_j = 0$ for $j \ne 0, i$;

(iv) $\mathbb{A}_i^n = \mathbb{P}^n - \mathbb{P}_i^{n-1}$ for $0 \le i \le n$; we will often write just $\mathbb{A}^n$ for $\mathbb{A}_0^n$;

(v) $\mathbb{A}_i^1 = \mathbb{P}_i^1 \cap \mathbb{A}^n$ and $\mathbb{A}_i^{n-1} = \mathbb{P}_i^{n-1} \cap \mathbb{A}^n$ for $1 \le i \le n$.

Consider the localization sequence of the pair $(\mathbb{P}^n, \mathbb{P}^n - 0)$:

$$\dots \to A(\mathbb{P}^n - 0) \xrightarrow{u_A} A(\mathbb{P}^n) \xrightarrow{v_A} A_0(\mathbb{P}^n) \to \dots, \tag{3.1}$$

where $u : \mathbb{P}^n - 0 \to \mathbb{P}^n$ and $v : (\mathbb{P}^n, \emptyset) \to (\mathbb{P}^n, \mathbb{P}^n - 0)$ are the natural maps. Note that $\mathbb{P}^n - 0$ can be considered as a line bundle over $\mathbb{P}_0^{n-1}$, with the projection map $t : \mathbb{P}^n - 0 \to \mathbb{P}_0^{n-1}$ given by $[x_0 : x_1 : \dots : x_n] \mapsto [0 : x_1 : \dots : x_n]$. Denote by $s : \mathbb{P}_0^{n-1} \to \mathbb{P}^n - 0$ the inclusion map, then by (h1), $s_A : A(\mathbb{P}_0^{n-1}) \to A(\mathbb{P}^n - 0)$ and $t_A : A(\mathbb{P}^n - 0) \to A(\mathbb{P}_0^{n-1})$ are inverse isomorphisms. Let $u' : \mathbb{P}_0^{n-1} \to \mathbb{P}^n$ be the inclusion map, then $u' = us$ and $u'_A = u_A s_A$. Consider the diagram

$$\begin{array}{ccc} A(\mathbb{P}_0^{n-1}) & \xrightarrow{u'_A} & A(\mathbb{P}^n) \\ {\scriptstyle F_{n-1}}\downarrow & & \downarrow{\scriptstyle F_n} \\ \bigoplus_{i=0}^{n-1} A(pt) & \xrightarrow{a_{n-1,n}} & \bigoplus_{i=0}^{n} A(pt) \end{array} \tag{3.2}$$

where $a_{n-1,n}$ maps each summand of $\bigoplus_{i=0}^{n-1} A(pt)$ to the same summand in $\bigoplus_{i=0}^{n} A(pt)$ as the identity map, the last summand in the latter group is therefore not being covered. We claim that the diagram commutes. For it suffices to prove that the diagram

$$\begin{array}{ccc} A(\mathbb{P}_0^{n-1}) & \xrightarrow{u'_A} & A(\mathbb{P}^n) \\ {\scriptstyle f_{n-1,i}}\downarrow & & \downarrow{\scriptstyle f_{n,i}} \\ A(pt) & \xrightarrow{1} & A(pt) \end{array}$$

commutes for every $0 \le i \le n-1$ and that $f_{n,n}u'_A = 0$. The first assertion follows from the commutativity of the diagram

$$\begin{tikzcd} A(\mathbb{P}^{n-1}_0) \arrow[r, "u'_A"] \arrow[d, "\xi^i_{n-1}\cap -"'] & A(\mathbb{P}^n) \arrow[d, "\xi^i_n\cap -"] \\ A(\mathbb{P}^{n-1}_0) \arrow[r, "u'_A"] & A(\mathbb{P}^n) \end{tikzcd}$$

which commutes by (cap3) since the restriction of $\mathcal{O}_{\mathbb{P}^n}(-1)$ to $\mathbb{P}^{n-1}_0$ is isomorphic to $\mathcal{O}_{\mathbb{P}^{n-1}_0}(-1)$ and $(u')^A(\xi_n) = \xi_{n-1}$. The same diagram with $i = n$ implies that the composition $f_{n,n}u'_A$ vanishes as $\xi^n_{n-1} = 0$. (See [PS, Sect. 3.3] for a standard argument that proves $\xi^n_{n-1} = 0$.)
Now consider the map $a_{n,n-1} : \bigoplus_{i=0}^n A(pt) \to \bigoplus_{i=0}^{n-1} A(pt)$ that identically maps the ith summand to the ith summand for all $0 \le i \le n-1$ and vanishes on the nth summand. As $a_{n,n-1}a_{n-1,n} = 1$, the commutativity of (3.2) implies $F_{n-1} = a_{n,n-1}a_{n-1,n}F_{n-1} = a_{n,n-1}F_n u'_A$. By the inductional hypothesis F_{n-1} is an isomorphism, whence u'_A is a split monomorphism, and so is u_A as s_A is an isomorphism. This has two important consequences:
(i) (3.1) is in fact a split short exact sequence;
(ii) the map $f_{n,n} : A(\mathbb{P}^n) \to A(pt)$ factors uniquely through v_A.
Denote by $g : A_0(\mathbb{P}^n) \to A(pt)$ the factoring map: $f_{n,n} = gv_A$. The diagram

$$\begin{tikzcd} 0 \arrow[r] & A(\mathbb{P}^{n-1}_0) \arrow[r, "u'_A"] \arrow[d, "F_{n-1}"] & A(\mathbb{P}^n) \arrow[r, "v_A"] \arrow[d, "F_n"] & A_0(\mathbb{P}^n) \arrow[r] \arrow[d, "g"] & 0 \\ 0 \arrow[r] & \bigoplus_{i=0}^{n-1} A(pt) \arrow[r, "a_{n-1,n}"] & \bigoplus_{i=0}^{n} A(pt) \arrow[r] & A(pt) \arrow[r] & 0 \end{tikzcd}$$

shows that we will be done as soon as it is proved that g is an isomorphism.
For $1 \le i \le n$, consider the cohomology localization sequence of the pair $(\mathbb{P}^n, \mathbb{A}^n_i)$:

$$A^{\cdot}_{\mathbb{P}^{n-1}_i}(\mathbb{P}^n) \xrightarrow{v_i^A} A^{\cdot}(\mathbb{P}^n) \xrightarrow{u_i^A} A^{\cdot}(\mathbb{A}^n_i)\,, \tag{3.3}$$

where $u_i : \mathbb{A}^n_i \to \mathbb{P}^n$ and $v_i : (\mathbb{P}^n, \emptyset) \to (\mathbb{P}^n, \mathbb{A}^n_i)$ are the natural maps. As $A^{\cdot}(\mathbb{A}^n_i) \cong A^{\cdot}(pt)$ by (c1), this is a split short exact sequence, the splitting for u_i^A given by $1 \mapsto 1$. The element $\xi_n \in A^{\cdot}(\mathbb{P}^n)$ maps to zero via u_i^A as the restriction of $\mathcal{O}(-1)$ to $\mathbb{A}^n_i$ is isomorphic to the trivial line bundle. Thus ξ_n comes from a uniquely determined element $\bar{\xi}_{n,i} \in A^{\cdot}_{\mathbb{P}^{n-1}_i}(\mathbb{P}^n)$. Note that $\mathbb{P}^{n-1}_1 \cap \ldots \cap \mathbb{P}^{n-1}_n = \{0\}$ and consider the diagram

$$\begin{tikzcd} A^{\cdot}_{\mathbb{P}^{n-1}_1}(\mathbb{P}^n) \oplus A^{\cdot}_{\mathbb{P}^{n-1}_2}(\mathbb{P}^n) \oplus \ldots \oplus A^{\cdot}_{\mathbb{P}^{n-1}_n}(\mathbb{P}^n) \arrow[r, "\cup"] \arrow[d, "v_1^A \oplus \ldots \oplus v_n^A"'] & A^{\cdot}_0(\mathbb{P}^n) \arrow[d, "v^A"] \\ A^{\cdot}(\mathbb{P}^n) \oplus A^{\cdot}(\mathbb{P}^n) \oplus \ldots \oplus A^{\cdot}(\mathbb{P}^n) \arrow[r, "\cup"] & A^{\cdot}(\mathbb{P}^n) \end{tikzcd}$$

which commutes since $\cup$ is compatible with pull-backs. It follows that the element

$$\bar{th}_n := \bar{\xi}_{n,1} \cup \bar{\xi}_{n,2} \cup \ldots \cup \bar{\xi}_{n,n} \in A^{\cdot}_0(\mathbb{P}^n)$$

satisfies $v^A(\bar{th}_n) = \xi_n^n$.
Now apply (cap3), with $X = Y = \mathbb{P}^n$, $U = U' = \mathbb{P}^n - 0$, $V = V' = \emptyset$ and $f = v$, to $a = \bar{th}_n$ and any $b \in A(\mathbb{P}^n)$ and get the commutativity of the diagram

$$\begin{array}{ccccc}
A(\mathbb{P}^n) & \xrightarrow{\xi_n^n \cap -} & A(\mathbb{P}^n) & \longrightarrow & A(pt) \\
{\scriptstyle v_A}\big\downarrow & & \big\downarrow{\scriptstyle 1} & & \big\downarrow{\scriptstyle 1} \\
A_0(\mathbb{P}^n) & \xrightarrow{\bar{th}_n \cap -} & A(\mathbb{P}^n) & \longrightarrow & A(pt)
\end{array}$$

The composition of the top arrows is $f_{n,n}$. As g is the unique arrow satisfying $f_{n,n} = gv_A$, we can conclude that g equals the composition of the bottom arrows.
Let $j : \mathbb{A}^n \to \mathbb{P}^n$ denote the inclusion map, and let $j_i : (\mathbb{A}^n, \mathbb{A}^n - \mathbb{A}_i^{n-1}) \to (\mathbb{P}^n, \mathbb{P}^n - \mathbb{P}_i^{n-1})$, with $1 \leq i \leq n$, and $\tilde{j} : (\mathbb{A}^n, \mathbb{A}^n - 0) \to (\mathbb{P}^n, \mathbb{P}^n - 0)$ denote the corresponding maps of pairs. Define $\xi_{n,i} \in A^{\cdot}_{\mathbb{A}_i^{n-1}}(\mathbb{A}^n)$ to be the image of $\bar{\xi}_{n,i}$ under the map $j_i^A : A^{\cdot}_{\mathbb{P}_i^{n-1}}(\mathbb{P}^n) \to A^{\cdot}_{\mathbb{A}_i^{n-1}}(\mathbb{A}^n)$. The diagram

$$\begin{array}{ccc}
A^{\cdot}_{\mathbb{P}_1^{n-1}}(\mathbb{P}^n) \oplus \ldots \oplus A^{\cdot}_{\mathbb{P}_n^{n-1}}(\mathbb{P}^n) & \xrightarrow{\cup} & A^{\cdot}_0(\mathbb{P}^n) \\
{\scriptstyle j_1^A \oplus \ldots \oplus j_n^A}\big\downarrow & & \big\downarrow{\scriptstyle \tilde{j}^A} \\
A^{\cdot}_{\mathbb{A}_1^{n-1}}(\mathbb{A}^n) \oplus \ldots \oplus A^{\cdot}_{\mathbb{A}_n^{n-1}}(\mathbb{A}^n) & \xrightarrow{\cup} & A^{\cdot}_0(\mathbb{A}^n)
\end{array}$$

shows that the element

$$th_n := \xi_{n,1} \cup \ldots \cup \xi_{n,n} \in A^{\cdot}_0(\mathbb{A}^n)$$

satisfies $\tilde{j}^A(\bar{th}_n) = th_n$.
Consider the diagram

$$\begin{array}{ccccc}
A_0(\mathbb{A}^n) & \xrightarrow{th_n \cap -} & A(\mathbb{A}^n) & \xrightarrow{\sim} & A(pt) \\
{\scriptstyle \tilde{j}_A}\big\downarrow & & \big\downarrow{\scriptstyle j_A} & & \big\downarrow{\scriptstyle 1} \\
A_0(\mathbb{P}^n) & \xrightarrow{\bar{th}_n \cap -} & A(\mathbb{P}^n) & \longrightarrow & A(pt)
\end{array}$$

that commutes by (cap3). Recall that our current goal is to prove that g, which equals the composition of the bottom arrows in the diagram, is an isomorphism. As $\tilde{j}_A$ is an isomorphism by excision and $A(\mathbb{A}^n) \to A(pt)$ is an isomorphism by homotopy invariance, it now suffices to prove that $th_n \cap - : A_0(\mathbb{A}^n) \to A(\mathbb{A}^n)$ is an isomorphism. This will be done in the next section.

4. Proof: Part II

First we will obtain another description for the elements $\xi_{n,i}$ and th_n. Consider the short exact sequnce

$$0 \to A^{\cdot}_0(\mathbb{P}^1) \to A^{\cdot}(\mathbb{P}^1) \to A^{\cdot}(\mathbb{P}^1 - 0) \to 0\,,$$

which is the one-dimensional version of (3.3). The element $\xi_1 = c(\mathcal{O}(-1)) \in A^{\cdot}(\mathbb{P}^1)$ maps to zero and comes therefore from a uniquely determined element $\bar{t} \in A^{\cdot}_0(\mathbb{P}^1)$. Let $t \in A^{\cdot}_0(\mathbb{A}^1)$ denote its image under the restriction map $A^{\cdot}_0(\mathbb{P}^1) \to A^{\cdot}_0(\mathbb{A}^1)$. (As the one-dimensional case plays a distinguished role, we change the notation and denote these elements by $\bar{t}$ and t.) If we think of $\mathbb{P}^1$ and $\mathbb{A}^1$ as coordinate axes $\mathbb{P}^1_i$ and $\mathbb{A}^1_i$ in $\mathbb{P}^n$ and $\mathbb{A}^n$ accordingly, $1 \le i \le n$, then we will denote the corresponding elements by $\bar{t}_i \in A^{\cdot}_0(\mathbb{P}^1_i)$ and $t_i \in A^{\cdot}_0(\mathbb{A}^1_i)$. Denote by $pr_i : \mathbb{A}^n \to \mathbb{A}^1_i$ the projection to the i-th coordinate and consider the map $pr_i^A : A^{\cdot}_0(\mathbb{A}^1_i) \to A^{\cdot}_{\mathbb{A}^{n-1}_i}(\mathbb{A}^n)$. It is proved in [PS] that $pr_i^A(t_i) = \xi_{n,i}$, and we can therefore rewrite th_n in the form

$$th_n = pr_1^A(t_1) \cup pr_2^A(t_2) \cup \ldots \cup pr_n^A(t_n). \tag{4.1}$$

(NB: In [PS] th_n is defined by the above formula and then it is proved that $pr_i^A(t_i)$ can be replaced by $\xi_{n,i}$, with a different notation.)
To proceed further we first need to prove a technical lemma which is the homology counterpart of [PS, Lemma 3.3.2]. Let $Y \in Sm$ and $Z \subset Y$ be a closed subset. Let $p : Y \times \mathbb{A}^1 \to Y$ and $pr : Y \times \mathbb{A}^1 \to \mathbb{A}^1$ denote the projections. Consider the map $pr^A : A^{\cdot}_0(\mathbb{A}^1) \to A^{\cdot}_{Y\times 0}(Y \times \mathbb{A}^1)$ and the image $pr^A(t)$ of $t \in A^{\cdot}_0(\mathbb{A}^1)$ under this map. The cap-product

$$\cap : A^{\cdot}_{Y\times 0}(Y \times \mathbb{A}^1) \times A_{Z\times 0}(Y \times \mathbb{A}^1) \to A_{Z\times\mathbb{A}^1}(Y \times \mathbb{A}^1)$$

induces the map $pr^A(t) \cap - : A_{Z\times 0}(Y \times \mathbb{A}^1) \to A_{Z\times\mathbb{A}^1}(Y \times \mathbb{A}^1)$.

Lemma. *The map $pr^A(t) \cap - : A_{Z\times 0}(Y \times \mathbb{A}^1) \to A_{Z\times\mathbb{A}^1}(Y \times \mathbb{A}^1)$ is an isomorphism.*

Proof. As $p_A : A_{Z\times\mathbb{A}^1}(Y \times \mathbb{A}^1) \to A_Z(Y)$ is an isomorphism by (h1), the assertion of the lemma is equivalent to the claim that the composed map

$$T := p_A \circ (pr^A(t) \cap -) : A_{Z\times 0}(Y \times \mathbb{A}^1) \to A_Z(Y)$$

is an isomorphism. It is this claim that we will actually prove.
We will make use of the localization sequence of the triple $(Y \times \mathbb{P}^1, Y \times \mathbb{P}^1 - Z \times 0, (Y - Z) \times \mathbb{P}^1)$:
(4.2)

$$\ldots \to A_{Z\times\mathbb{A}^1_\infty}(Y \times \mathbb{P}^1 - Z \times 0) \xrightarrow{\alpha_A} A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1) \xrightarrow{\beta_A} A_{Z\times 0}(Y \times \mathbb{P}^1) \to \ldots\,,$$

where $\mathbb{A}^1_\infty := \mathbb{P}^1 - 0$ and α and β are the corresponding inclusion maps of pairs.

Consider the inclusion $i : (Y \times \mathbb{A}^1_\infty, (Y - Z) \times \mathbb{A}^1_\infty) \to (Y \times \mathbb{P}^1 - Z \times 0, (Y - Z) \times \mathbb{P}^1)$. One checks that i satisfies the excision conditions (Zarisky version), whence $i_A : A_{Z\times\mathbb{A}^1_\infty}(Y \times \mathbb{A}^1_\infty) \to A_{Z\times\mathbb{A}^1_\infty}(Y \times \mathbb{P}^1 - Z \times 0)$ is an isomorphism.
Let $\tilde{p} : (Y \times \mathbb{P}^1 - Z \times 0, (Y - Z) \times \mathbb{P}^1) \to (Y, Y - Z)$ and $p' = \tilde{p}i : (Y \times \mathbb{A}^1_\infty, (Y-Z) \times \mathbb{A}^1_\infty) \to (Y, Y-Z)$ be the projections. As p'_A is an isomorphism by (h1), $p'_A = \tilde{p}_A i_A$ shows that $\tilde{p}_A : A_{Z\times\mathbb{A}^1_\infty}(Y \times \mathbb{P}^1 - Z \times 0) \to A_Z(Y)$ is an isomorphism too.
Let $\bar{p} : (Y \times \mathbb{P}^1, (Y-Z) \times \mathbb{P}^1) \to (Y, Y-Z)$ denote the projection. Then $\tilde{p} = \bar{p}\alpha$ and $\tilde{p}_A = \bar{p}_A\alpha_A$, which implies that α_A is a split monomorphism. It follows that β_A is surjective and (4.2) is a short exact sequence.
Let $\bar{pr} : Y \times \mathbb{P}^1 \to \mathbb{P}^1$ denote the projection. Consider the diagram

$$\begin{array}{ccc} A^{\cdot}_0(\mathbb{P}^1) & \longrightarrow & A^{\cdot}(\mathbb{P}^1) \\ {\scriptstyle \bar{pr}^A}\downarrow & & \downarrow{\scriptstyle \bar{pr}^A} \\ A^{\cdot}_{Y\times 0}(Y \times \mathbb{P}^1) & \longrightarrow & A^{\cdot}(Y \times \mathbb{P}^1) \end{array} \tag{4.3}$$

and the cap-products

$$\cap : A^{\cdot}_{Y\times 0}(Y \times \mathbb{P}^1) \times A_{Z\times 0}(Y \times \mathbb{P}^1) \to A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1)$$
$$\cap : A^{\cdot}(Y \times \mathbb{P}^1) \times A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1) \to A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1)\,.$$

The element $\bar{pr}^A(\bar{t})$ maps to $\bar{pr}^A(\xi_1)$ via the bottom arrow in (4.3). By (cap3), it follows that the diagram

$$\begin{array}{ccc} A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1) & \xrightarrow{\beta_A} & A_{Z\times 0}(Y \times \mathbb{P}^1) \\ {\scriptstyle \bar{pr}^A(\xi_1)\cap -}\downarrow & & \downarrow{\scriptstyle \bar{pr}^A(\bar{t})\cap -} \\ A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1) & \xrightarrow{1} & A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1) \end{array} \tag{4.4}$$

commutes. Denote $\bar{T} = \bar{p}_A \circ (\bar{pr}^A(\bar{t}) \cap -)$, $f^Z_0 = \bar{p}_A$, and $f^Z_1 = \bar{p}_A \circ (\bar{pr}^A(\xi_1) \cap -)$. We are now prepared to consider the diagram

$$\begin{array}{ccccc} A_{Z\times\mathbb{A}^1_\infty}(Y \times \mathbb{P}^1 - Z \times 0) & \xrightarrow{\alpha_A} & A_{Z\times\mathbb{P}^1}(Y \times \mathbb{P}^1) & \xrightarrow{\beta_A} & A_{Z\times 0}(Y \times \mathbb{P}^1) \\ {\scriptstyle \tilde{p}_A}\downarrow & & \downarrow{\scriptstyle \binom{f^Z_0}{f^Z_1}} & & \downarrow{\scriptstyle \bar{T}} \\ A_Z(Y) & \xrightarrow{\binom{1}{0}} & A_Z(Y) \oplus A_Z(Y) & \xrightarrow{(0,1)} & A_Z(Y) \end{array}$$

where the rows are short exact sequences, with undisplayed zeros on both sides. The right square commutes since (4.4) commutes. To prove the commutativity on the left we must check that $f_1^Z \alpha_A = 0$ (recall that $\bar{p}_A \alpha_A = \tilde{p}_A$). Consider the diagram

$$\begin{array}{ccc} A_{Z\times\mathbb{A}^1_\infty}(Y\times\mathbb{A}^1_\infty) & \xrightarrow{\alpha_A i_A} & A_{Z\times\mathbb{P}^1}(Y\times\mathbb{P}^1) \\ {\scriptstyle i^A(\bar{pr}^A(\xi_1))\cap-}\downarrow & & \downarrow{\scriptstyle \bar{pr}^A(\xi_1)\cap-} \\ A_{Z\times\mathbb{A}^1_\infty}(Y\times\mathbb{A}^1_\infty) & \xrightarrow{\alpha_A i_A} & A_{Z\times\mathbb{P}^1}(Y\times\mathbb{P}^1) \end{array}$$

which commutes by (cap3). But $i^A(\bar{pr}^A(\xi_1)) \in A^{\cdot}(Y\times\mathbb{A}^1_\infty)$ vanishes as $\mathcal{O}(-1)$ restricted to $\mathbb{A}^1$ is trivial. Thus $(\bar{pr}^A(\xi_1)\cap -)\alpha_A i_A = 0$. As i_A is an isomorphism, $(\bar{pr}^A(\xi_1)\cap -)\alpha_A = 0$, whence $f_1^Z\alpha_A = 0$ and the big diagram commutes. Now we claim that the arrow $\begin{pmatrix} f_0^Z \\ f_1^Z \end{pmatrix}$ is an isomorphism. The absolute (without supports) version of this is postulated in (ch4). The 'with supports' version can be deduced from (ch4) by applying the five-lemma to obvious localization sequences. Recall that $\tilde{p}_A$ is an isomorphism and conclude that $\bar{T}$ is an isomorphism.

To complete the proof, consider the diagram

$$\begin{array}{ccccc} A_{Z\times 0}(Y\times\mathbb{A}^1) & \xrightarrow{pr^A(t)\cap-} & A_{Z\times\mathbb{A}^1}(Y\times\mathbb{A}^1) & \longrightarrow & A_Z(Y) \\ \downarrow & & \downarrow & & \downarrow{\scriptstyle 1} \\ A_{Z\times 0}(Y\times\mathbb{P}^1) & \xrightarrow{\bar{pr}^A(\bar{t})\cap-} & A_{Z\times\mathbb{P}^1}(Y\times\mathbb{P}^1) & \longrightarrow & A_Z(Y) \end{array}$$

where $\mathbb{A}^1$ now denotes $\mathbb{P}^1 - \infty$ (as opposed to $\mathbb{A}^1_\infty$). It commutes by (cap3) as $\bar{pr}^A(\bar{t})$ maps to $pr^A(t)$ via the map $A^{\cdot}_{Y\times 0}(Y\times\mathbb{P}^1) \to A^{\cdot}_{Y\times 0}(Y\times\mathbb{A}^1)$. The top composition is T, the bottom one is $\bar{T}$, and the left vertical arrow is an isomorphism by excision. We conclude that T is an isomorphism. The lemma is proved.

Define $\mathbb{A}^{(k)}$ to be the k-dimensional affine subspace of $\mathbb{A}^n$ given by $x_1 = \ldots = x_{n-k} = 0$, for $0 \le k \le n$. Then $\mathbb{A}^{(k+1)} \cap \mathbb{A}^{n-1}_{n-k} = \mathbb{A}^{(k)}$. By (4.1) and (cap1), the map $th_n \cap - : A_0(\mathbb{A}^n) \to A(\mathbb{A}^n)$ can be decomposed as

$$A_0(\mathbb{A}^n) \xrightarrow{pr_n^A(t_n)\cap-} A_{\mathbb{A}^{(1)}}(\mathbb{A}^n) \xrightarrow{pr_{n-1}^A(t_{n-1})\cap-} A_{\mathbb{A}^{(2)}}(\mathbb{A}^n) \xrightarrow{pr_{n-2}^A(t_{n-2})\cap-} \ldots \xrightarrow{pr_1^A(t_1)\cap-} A(\mathbb{A}^n)\,.$$

A generic step of this decomposition is a map

$$pr_{n-k}^A(t_{n-k}) : A_{\mathbb{A}^{(k)}}(\mathbb{A}^n) \to A_{\mathbb{A}^{(k+1)}}(\mathbb{A}^n)\,. \tag{4.5}$$

In the notation of the lemma, put $Y = \mathbb{A}^{n-1}_{n-k}$, $Z = \mathbb{A}^{(k)}$, and think of $\mathbb{A}^1$ as $\mathbb{A}^1_{n-k}$. Then $Y\times\mathbb{A}^1$ can be identified with $\mathbb{A}^n$, $Z\times\mathbb{A}^1$ with $\mathbb{A}^{(k+1)}$, $pr^A(t)$ becomes $pr_{n-k}^A(t_{n-k})$, and we get that (4.5) is an isomorphism. The theorem is proved.

Applying (h2) we obtain

Corollary (PBTH with supports). *If Z is a closed subvariety in a smooth X, E is a vector bundle over X of rank $n+1$, and E_Z is its restriction to Z, then the map*

$$F_n = (f_0, \dots , f_n) : A_{\mathbb{P}(E_Z)}(\mathbb{P}(E)) \to A_Z(X) \oplus \dots \oplus A_Z(X)$$

defined the same way as F_n in PBTH is an isomorphism.

References

[PS] I. Panin (after I. Panin and A. Smirnov), *Oriented cohomology theories of algebraic varieties*, K-Theory **30** (2003), 263-312.

[PS1] I. Panin (after I. Panin and A. Smirnov), *Push-forwards in oriented cohomology theories of algebraic varieties: II*, preprint at http://www.math.uiuc.edu/K-theory/0619/ (2003).

[Pa] I. Panin (after I. Panin and A. Smirnov), *Riemann-Roch theorems for oriented cohomology*, in Proceedings of the NATO ASI "Axiomatic, enriched and motivic homotopy theory" (Ed. J. P. C. Greenlees), NATO Sci. Ser. II: Math. Phys. and Chemistry, Vol. 131 (2004), 261-333.

[Pi] K. Pimenov, *Traces in oriented homology theories of algebraic varieties*, preprint at http://www.math.uiuc.edu/K-theory/0667/ (2003).

[PY] I. Panin and S. Yagunov, *Poincaré duality for projective varieties*, preprint at http://www.math.uiuc.edu/K-theory/0576/ (2002).

[SV] A. Suslin and V. Voevodsky, *Bloch-Kato conjecture and motivic cohomology with finite coefficients*, The Arithmetic and Geometry of Algebraic Cycles, 117-189; NATO Sci. Ser. C: Math. Phys. Sci., Vol. 548, Kluwer Acad. Publ., Dortrecht (2002).

[V] V. Voevodsky, $\mathbb{A}^1$*-homotopy theory*, Doc. Math., Extra Vol. ICM 1998 (I), 417-442.

Alexander Nenashev
York University - Glendon College
Department of Mathematics
2275 Bayview Av., Toronto ON
Canada M4N 3M6
nenashev@glendon.yorku.ca

Theoremes d'Annulation et Lieux de Degenerescence en Petit Corang

Pierre-Emmanuel Chaput

Received: April 19, 2004

Communicated by Thomas Peternell

Abstract. After giving an explicit description of all the non vanishing Dolbeault cohomology groups of ample line bundles on grassmannians, I give two series of vanishing theorems for ample vector bundles on a smooth projective variety. They imply a part of a conjecture by Fulton and Lazarsfeld about the connectivity of some degeneracy loci.

2000 Mathematics Subject Classification: 14F17
Keywords and Phrases: Théorèmes d'annulation, grassmannienne.

Introduction

Soit X une variété projective complexe, E un fibré vectoriel sur X et L un fibré en droites. Supposons que l'on ait une forme quadratique sur E à valeurs dans L, soit une section de $S^2E^* \otimes L$. Si k est un entier, on note $D_k(E)$ le sous-schéma de X où cette section est au plus de rang k. Dans [FL 81, Remark 2,p.50], on peut lire la conjecture suivante ($t(x) := \frac{x(x+1)}{2}$) :

Conjecture 1 *Soit E un fibré vectoriel de rang e, sur une variété X lisse, projective, connexe et de dimension n. Supposons que E est muni d'une forme quadratique à valeurs dans un fibré en droites L. Soit k un entier et supposons que*
– $\dim D_k(E) = \rho := n - t(e-k) \geq 1$.
– $S^2E^* \otimes L$ *est ample.*
Alors, $D_k(E)$ est connexe.

Dans cet article, je montre cette conjecture sous l'hypothèse supplémentaire que $e - k \leq 4$ et $\rho \geq 2$. J'obtiens en fait les résultats plus précis suivants :

Théorème 3 : *Sous les hypothèses de la conjecture précédante, à part que X n'est plus supposée connexe, et si de plus $\rho \geq 2$ et $e-k \leq 4$, alors l'application de restriction $H^q(X, \mathcal{O}_X) \to H^q(D_k(E), \mathcal{O}_{D_k(E)})$ est un isomorphisme pour $0 \leq q < \rho - 1$, et est injective pour $q = \rho - 1$.*

La conjecture de Fulton et Larzarsfeld a été résolue en utilisant une technique différente par Tu [Tu 89], dans le cas où k est pair. Par ailleurs, [Tu 88] montre aussi la connexité, mais à condition que $\rho \geq e-k$. Si $e-k=3$ ou 4, ma borne, 2, est donc meilleure. J'espère aussi que ce travail est un pas de plus vers une parfaite compréhension des phénomènes combinatoires qui permettent d'établir les théorèmes d'annulation. Mentionnons enfin que [Lay 96, prop 2.3] donne en utilisant la même technique que moi la connexité d'un lieu de dimension strictement positive, mais pour $e-k \leq 2$, ce qui, comme le lecteur pourra le mesurer, simplifie considérablement le problème.

Il est bien connu que la conjecture découle de théorèmes d'annulation adéquats [Man 97] : il suffit en effet d'utiliser une résolution du faisceau structural de $D_k(E)$ par des fibrés vectoriels qui sont des puissances de Schur de E, et d'appliquer les théorèmes d'annulation aux termes de cette résolution. Dans le cas où $e-k \leq 2$, seuls des crochets, à savoir des partitions dont seule la première part est éventuellement supérieure à 2, interviennent, et le théorème 2.1 dans [LN 02] convient. Dans cet article, je propose une généralisation de ce résultat pour des produits tensoriels de crochets. Cette généralisation donne bien entendu un théorème d'annulation pour toutes les partitions ; malheureusement, la borne obtenue est insuffisante pour établir la conjecture. Je propose donc dans cet article une méthode un peu nouvelle, de "comparaison de suites spectrales", pour établir des théorèmes d'annulation plus puissants. Le théorème 3 est alors conséquence des théorèmes d'annulation obtenus par cette méthode.

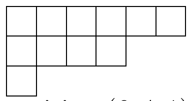
L'efficacité des théorèmes d'annulation obtenus dépend directement de notre compréhension de la cohomologie des fibrés en droites homogènes amples sur une grassmannienne ; Snow [Sno 86] a donné pour la calculer une méthode diagrammatique commode. L'inconvénient de cette méthode est qu'il est difficile d'en déduire, pour $r < e$ et l des entiers fixés, l'entier maximal p tel qu'il existe q avec $H^{p,q}[G(r,e), \mathcal{O}(l)] \neq 0$ ($G(r,n)$ désigne la grassmannienne des r-plans vectoriels dans un espace vectoriel de dimension n et $\mathcal{O}(l)$ est la l-ième puissance du déterminant du fibré quotient). Pourtant, comme nous allons le voir, la détermination de cet entier est cruciale pour obtenir des théorèmes d'annulation. Je propose donc une description nouvelle de la cohomologie de $\mathcal{O}(l)$ sur une grassmannienne, basée sur celle de Snow, et j'en déduis le théorème 1 qui permet de déterminer cet entier.

Je remercie mon directeur de thèse Laurent Manivel pour son aide tout au long de l'élaboration de cet article.

1 Cohomologie de Dolbeault des fibres $\mathcal{O}(l)$ sur une grassmannienne

1.1 Description explicite de toutes les partitions admissibles

Rappelons tout d'abord la description de Snow [Sno 86] des groupes de cohomologie non nuls sur une grassmannienne. Une partition est suite décroissante finie d'entiers. Je note λ_i le i-ième entier de λ ; c'est par convention la i-ième part de λ. Le poids de λ, noté $|\lambda|$, est la somme de ses parts. Sa longueur, $l(\lambda)$, est le nombre de parts non nulles. Dans de nombreuses circonstances, il est pratique de représenter les partitions par un diagramme comme le suggère l'exemple qui suit :



partition $(6, 4, 1)$.

A chaque partition λ correspond un foncteur de la catégorie des espaces vectoriels dans elle-même que j'appelle "de Schur" et que je note S_λ [FH 91].
Lorsque λ est une partition, on peut attribuer à chaque case de cette partition son nombre de crochet qui est le nombre de cases de λ situées en-dessous ou à droite de cette case, cette case comprise. Par exemple, les cases de la partition suivante ont été numérotées par leurs nombres de crochets :

6	4	2	1
3	1		
1			

.

Convenons alors, si l est un entier, que λ sera appelée l-admissible si toutes les cases reçoivent un numéro différent de l.
Snow [Sno 86] a montré qu'il existe une bijection, à $r < e$ et l fixés, entre les composantes des groupes de cohomologie $H^{p,q}[G(r,e), \mathcal{O}(l)]$ et les partitions de longueur inférieure ou égale à r dont toutes les parts sont inférieures ou égales à $e - r$ (par la suite je les appellerai simplement partitions de taille $(e-r, r)$) qui sont l-admissibles. Si λ est une partition l-admissible, soit ${}^h\lambda^-$ (respectivement ${}^v\lambda^-$) la suite d'entiers telle que ${}^h\lambda_i^-$ (respectivement ${}^v\lambda_j^-$) est le nombre de cases situées sur la i-ième ligne (respectivement sur la j-ème colonne) dont le nombre de crochet est strictement inférieur à l. Nous allons voir que ${}^h\lambda^-$ et ${}^v\lambda^-$ sont des partitions. Si $p = |\lambda|$ et q est le nombre de cases qui ont un nombre de crochets strictement supérieur à l, alors $H^{p,q}[G(r,e), \mathcal{O}(l)]$ contient $S_\mu \mathbb{C}^e$, si μ est la partition obtenue en réordonnant toutes les parts $(l - {}^h\lambda_i^-)$ et ${}^v\lambda_j^-$.
Le seul inconvénient de cette méthode est que l'ensemble des partitions l-admissibles n'est pas aisé à décrire. Je propose dans ce paragraphe de montrer qu'il est en bijection avec l'ensemble de toutes les partitions de taille $(l-1, r)$, et d'étudier des propriétés de cette bijection.

Proposition 1 *L'application $\lambda \mapsto {}^h\lambda^-$ est une bijection de l'ensemble des partitions l-admissibles dans l'ensemble des partitions de taille $(l-1, r)$.*

Demonstration : Tout d'abord si λ est l-admissible, alors $\forall i, {}^h\lambda_i^- < l$ puisque les nombres de crochet sont strictement décroissants sur une ligne. Introduisons quelques notations.

Definition 1

- *Soit $h_{i,j}$ le nombre de crochet de la case située sur la i-ième ligne et la j-ème colonne. Convenons que $h_{i,0} = +\infty$.*
- *Pout tout i, soit $(g_{i,j})_{j\in\mathbb{N}}$ la suite strictement croissante d'entiers dont l'image est le complémentaire dans $\mathbb{N}$ de l'ensemble des $h_{i,j}$.*
- *Notons δ_i l'entier $l - {}^h\lambda_i^-$, tel que $g_{i,\delta_i} = l$. Notons enfin $x_{i,j}$ le numéro de colonne de la dernière case sur la i-ième ligne de nombre de crochet supérieur à $g_{i,j}$.*

Ainsi, on a $g_{i,0} = 0$, $x_{i,0} = \lambda_i$ et pour $j >> 1$, $x_{i,j} = 0$, par nos conventions. Alors, par la définition des nombres de crochet, on a $h_{i-1,j} = h_{i,j} + \lambda_{i-1} - \lambda_i + 1$. Les nombres de crochet qui apparaissent sur la $(i-1)$-ième ligne sont donc $1, \cdots, \lambda_{i-1} - \lambda_i$ et les $h_{i,j} + \lambda_{i-1} - \lambda_i + 1$. Ainsi, $g_{i-1,j+1} = g_{i,j} + \lambda_{i-1} - \lambda_i + 1$. En particulier, $g_{i-1,\delta_i+1} > l$, donc $\delta_{i-1} \leq \delta_i$ et donc ${}^h\lambda^-$ est une partition. On démontre maintenant un résultat plus précis que la proposition 1 :

Lemme 1 *Pout toute partition ν de taille $(l-1, r)$, il existe une unique partition l-admissible λ telle que ${}^h\lambda^- = \nu$; pour celle-ci, on a :*
- *$\lambda_i = \lambda_{i+l-\nu_i} + \nu_i$.*
- *$x_{i,j} = \lambda_{i+j}$.*

Demonstration : On démontre par récurrence descendante sur i_0 qu'une partition a ses nombres de crochet différents de l à partir de la ligne i_0 et vérifie ${}^h\lambda_i^- = \nu_i$ pour $i \geq i_0$ si et seulement si elle vérifie les deux points du lemme pour $i \geq i_0$.
On a $\lambda_r = \nu_r, x_{r,0} = \lambda_r$ et $x_{r,j} = 0$ pour $j > 0$, de sorte que la propriété de récurrence est vraie pour $i_0 = r$. Soit maintenant i_0 fixé tel que cette hypothèse est vérifiée pour $i_0 + 1$. Alors posons $\delta = \delta_{i_0} (= l - \nu_{i_0})$ et $d = \lambda_{i_0} - \lambda_{i_0+1}$ (d dépend de λ_{i_0} qui est pour l'instant inconnu). On a vu que $g_{i_0,0} = 0$ et $g_{i_0,j+1} = g_{i_0,j} + d + 1$ pour $j \geq 0$. Les nombres de crochet de λ en-dessous et sur la i_0-ième ligne sont différents de l si et seulement si $g_{i_0,\delta} = l$, ce qui équivaut donc à $d = l - 1 - g_{i_0+1,\delta_{i_0}-1}$. Il existe donc exactement une possibilité pour λ_{i_0}. De plus, on a toujours $x_{i_0,j+1} = x_{i_0+1,j} = \lambda_{i_0+j+1}$ et le fait qu'effectivement ${}^h\lambda_{i_0}^- = \nu_{i_0}$ implique que $\lambda_{i_0} = x_{i_0,\delta} + \nu_{i_0} = \lambda_{i_0+\delta} + \nu_{i_0}$. •

Notation 2 *Si ν est une partition de taille $(l-1, r)$, je noterai $\widehat{\nu}$ la partition λ telle que ${}^h\lambda^- = \nu$. Par aileurs je noterai $p(\nu)$ et $q(\nu)$ le poids et le nombre de cases de nombre de crochet supérieur à l de λ.*

J'ai affirmé qu'aussi ${}^v\lambda^-$ est une partition, cela découle du fait que λ^* est l-admissible et ${}^v\lambda^- = {}^h(\lambda^*)^-$, où λ^* désigne la partition dont la i-ième part est la longueur de la i-ième colonne du diagramme représentant λ.

Si λ_1, λ_2 sont des partitions, on note $\lambda_1 \subset \lambda_2$ le fait que $\forall i, \lambda_{1,i} \leq \lambda_{2,i}$. Par une récurrence immédiate, le lemme 1 implique :

Proposition 2 *Si $\nu_1 \subset \nu_2$, alors $\widehat{\nu}_1 \subset \widehat{\nu}_2$.*

Proposition 3 *Soient k et l deux entiers et λ une partition. Si λ est l-admissible, alors elle est (kl)-admissible.*

Demonstration : En effet, si G_i désigne l'ensemble des entiers naturels qui n'apparaissent pas parmi les nombres de crochet sur la i-ième ligne, on sait que pour tout i, $l \in G_i$ et il existe d_i tel que $G_{i-1} = (d_i + 1 + G_i) \cup \{0\}$. Une récurrence immédiate prouve alors que pour tout i, G_i vérifie la propriété $x \in G_i \Rightarrow x + l \in G_i$. La proposition en découle.
•

La contraposée de cette proposition implique que pour toute partition λ exactement inscrite dans un rectangle $r \times (e - r)$ (c'est-à-dire que $\lambda_1 = e - r$ et $l(\lambda) = r$), il existe une case dont le nombre de crochet est d pour tout d diviseur de $e - 1$. Une telle partition n'est donc pas d-admissible. Si λ correspond à une composante non nulle de $H^{*,*}[G(r, \mathbb{C}^e), \mathcal{O}(d)]$, on a donc soit $\lambda_1 < e - r$ soit $l(\lambda) < r$. Ainsi, si $S_\mu \mathbb{C}^e \subset H^{*,*}[G(r, \mathbb{C}^e), \mathcal{O}(d)]$, on a soit $\mu_e = 0$, soit $\mu_1 = d$.

1.2 Partitions admissibles de poids maximal

Comme nous le verrons au paragraphe suivant, pour démontrer des théorèmes d'annulation pour la cohomologie de fibrés vectoriels amples, il est utile de bien comprendre la cohomologie des fibrés homogènes sur une grassmannienne, et plus précisément quelle est, pour l, n et r des entiers quelconques, la valeur maximale de p telle qu'il existe un entier q avec $H^{p,q}(G(r, n), \mathcal{O}(l))$ non nul. Si $n > rl$, cet entier a été déterminé par Laurent Manivel [Man 92, proposition 1.2.1, p.111] et cela a permis de démontrer de nombreux théorèmes d'annulation, dont celui de F. Laytimi et W. Nahm concernant les partitions en forme de crochets [LN 02]. Lorsque n, r et l ne satisfont plus cette inégalité, la combinatoire des partitions l-admissibles devient nettement plus compliquée et aucun résultat ne donne cet entier p.
Soit n, r, l des entiers fixés. Si $a, \alpha, \beta, c, \gamma$ sont des entiers, notons $\mu(a, \alpha, \beta, c, \gamma)$ la partition μ définie par :

$$\begin{cases} \forall\ i \leq \alpha(l - a), & \mu_i = a \\ \forall\ \alpha(l - a) < i \leq \alpha(l - a) + \beta(l - a + 1), & \mu_i = a - 1 \\ \forall\ \alpha(l - a) + \beta(l - a + 1) < i \leq \alpha(l - a) + \beta(l - a + 1) + \gamma, & \mu_i = c \end{cases}$$

Cette partition est de longueur $(\alpha + \beta)(l - a) + \beta + \gamma$ et un calcul direct utilisant le lemme 1 montre que $\widehat{\mu}_1(a, \alpha, \beta, c, \gamma) = (\alpha + \beta)a - \beta + c$.
Notons $\delta(\alpha, \beta) = \min\{\alpha, \beta\}$ si $\beta \neq 0$, et $\delta(\alpha, 0) = \alpha$. Rappelons que si μ est une partition, $\widehat{\mu}$ a été définie comme l'unique partition telle que ${}^h\widehat{\mu}^- = \mu$ (cf lemme 1).

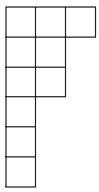
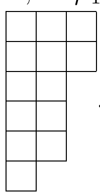
THEOREME 1 *Soient n, r et l des entiers, et λ une partition l-admissible de longueur r et telle que $\lambda_1 \leq n-r$.*
Alors, soit λ est de la forme $\widehat{\mu}(a,\alpha,\beta,c,\gamma)$ avec $\gamma \leq l-a$, soit il existe des entiers a,α,β,c,γ avec $\gamma \leq l-a$ et tels que $\widehat{\mu}(a,\alpha,\beta,c,\gamma)$ soit aussi l-admissible de longueur r et de première part inférieure ou égale à $n-r$, et que de plus $|\widehat{\mu}(a,\alpha,\beta,c,\gamma)| \geq |\lambda| + \delta(\alpha,\beta)$. De plus, ces entiers sont tels que $\alpha+\beta$ et $\gamma+c$ sont le quotient et le reste de la division euclidienne de n par l.

DEMONSTRATION : Considérons $n > r$ et l trois entiers, et soit λ une partition l-admissible de longueur au plus r et telle que $\lambda_1 \leq n-r$. On va raisonner sur la partition ${}^h\lambda^-$: nous allons voir que dans trois cas, on peut modifier cette partition pour en obtenir une de poids supérieur et encore de longueur au plus r et de même première part. Définissons une suite a_i par

$$\begin{cases} a_1 = 1 \\ a_{i+1} = \min\{a_i + l - {}^h\lambda^-_{a_i}, r+1\}. \end{cases}$$

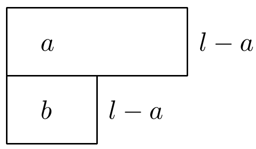
Soit A le plus grand entier i tel que $a_i \leq r$. Le lemme 1 montre que pour $i \leq A$,

$$\lambda_{a_i} = \sum_{i \leq j \leq A} {}^h\lambda^-_{a_j}.$$

En particulier, si μ est une partition telle que $\forall i \leq A, \mu_{a_i} = {}^h\lambda^-_{a_i}$, alors on a $\widehat{\mu}_1 = \lambda_1$. Par exemple, si μ la partition définie par $\forall i \leq A, \forall j \in [a_i, a_{i+1}-1]$, $\mu_j = {}^h\lambda^-_{a_i}$, on a, par la proposition 2, $\widehat{\mu} \supset \lambda$, et $\widehat{\mu}_1 = \lambda_1$. Par exemple, pour $l = 5$, on remplace la partition  par  .

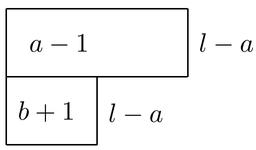
Supposons qu'il existe des entiers i, a et b avec $b \leq a-2$ et tels que

$$\begin{cases} \forall j \in [i+1, i+l-a], & {}^h\lambda^-_j = a \\ \forall j \in [i+l-a+1, i+2(l-a)+1], & {}^h\lambda^-_j = b. \end{cases}$$



Alors, considérons la partition μ définie par

$$\begin{cases} \forall j \in [i+1, i+l-a], & \mu_j = a-1 \\ \forall j \in [i+l-a+1, i+2(l-a)+1], & \mu_j = b+1 \\ \forall j \notin [i+1, i+2(l-a)+1], & \mu_j = {}^h\lambda^-_j. \end{cases}$$



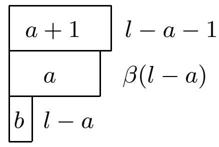
Le lemme 1 implique que

$$\begin{cases} \forall j \in [i+1, i+l-a], & \widehat{\mu}_j = \lambda_j \\ \forall j \in [i+l-a+1, i+2(l-a)+1], & \widehat{\mu}_j \geq \lambda_j + 1 \\ \forall j \notin [i+1, i+2(l-a)+1], & \widehat{\mu}_j = \lambda_j. \end{cases}$$

Par conséquent, on a $|\widehat{\mu}| \geq |\lambda| + l - a + 1$. Appelons "transformation A" le fait de remplacer ${}^h\lambda^-$ par μ.

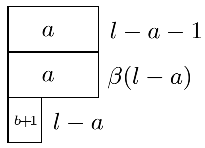
Supposons enfin l'existence d'entiers $i, a, b, \beta \geq 1$ tels que

$$\begin{cases} \forall j \in [i+1, i+l-a-1], & {}^h\lambda_j^- = a+1 \\ \forall j \in [i+l-a, i+(1+\beta)(l-a)-1], & {}^h\lambda_j^- = a \\ \forall j \in [i+(1+\beta)(l-a), i+(2+\beta)(l-a)-1], & {}^h\lambda_j^- = b. \end{cases}$$



Alors, considérons la partition μ définie par

$$\begin{cases} \forall j \in [i+1, i+(1+\beta)(l-a)-1], & \mu_j = a \\ \forall j \in [i+(1+\beta)(l-a), i+(2+\beta)(l-a)-1], & \mu_j = b+1 \\ \forall j \notin [i+1, i+(2+\beta)(l-a)-1], & \mu_j = {}^h\lambda_j^-. \end{cases}$$



Le lemme 1 implique que

$$\begin{cases} \forall j \in [i, i+l-a-1], & \widehat{\mu}_j = \lambda_j \\ \forall j \in [i+l-a, i+(2+\beta)(l-a)-2], & \widehat{\mu}_j = \lambda_j + 1 \\ \forall j \notin [i+1, i+(2+\beta)(l-a)-1], & \widehat{\mu}_j = \lambda_j. \end{cases}$$

En particulier, on a donc $|\widehat{\mu}| \geq |\lambda| + 2(l-a)$. Appelons "transformation B" le fait de remplacer ${}^h\lambda^-$ par μ.

On peut donc construire une suite finie de partitions μ^i par récurrence : tout d'abord, on pose $\mu^0 = {}^h\lambda^-$. Ensuite, on impose que pour passer de μ^i à μ^{i+1}, soit on ajoute une case, soit on fait l'une des deux transformations A ou B, lorsque cela est possible de façon à ce que $\widehat{\mu}^{i+1}$ reste dans un rectangle de taille $r \times (n-r)$. A un certain indice N, on ne peut plus faire aucune de ces trois opérations. Soit (b_i) les nombres définis comme a_i pour la partition μ^N :

$$\begin{cases} b_1 = 1 \\ b_{i+1} = \inf\{b_i + l - \mu_{b_i}^N, r+1\}. \end{cases}$$

Comme on ne peut plus ajouter de cases, on a $\mu_j^N = \mu_{b_i}^N$ pour $b_i \leq j \leq b_{i+1} - 1$, et comme aucune des deux configurations étudiées ne peut avoir lieu, μ^N est de type $\mu(a, \alpha, \beta, c, \gamma)$ avec $\gamma \leq l - a$. De plus, puisque l'on ne peut pas ajouter une case, on a soit $l[\mu(a, \alpha, \beta, c, \gamma)] = r$, soit $c = 0$. Si $c = 0$, $\mu(a, \alpha, \beta, c, \gamma)$ ne dépend pas de γ ; on peut donc dans tous les cas supposer que

$$\alpha(l-a) + \beta(l-a+1) + \gamma = r.$$

Pour montrer que $|\mu^N| \geq |\lambda| + \delta(\alpha, \beta)$, il suffit de montrer que $|\mu^N| \geq |\mu^{N-1}| + \delta(\alpha, \beta)$. Or, quand on applique l'opération A, on obtient une partition qui a exactement $l - a$ parts égales à $a - 1$, et quand on applique l'opération B, on obtient une partition qui a exactement $2(l-a) - 1$ parts égales à a : ce ne sont

donc pas des partitions de type $\mu(a,\alpha,\beta,c,\gamma)$. On en déduit donc que pour passer de μ^{N-1} à $\mu^N = \mu(a,\alpha,\beta,c,\gamma)$, on a ajouté une case. Soit cette case se trouvait sur la r-ième ligne, et on a $|\mu^N|-|\mu^{N-1}| = \alpha+\beta+1$; soit, si $\beta > 0$, elle se trouvait sur la $\alpha(l-a)+\beta(l-a+1)$-ième ligne, et on a $|\mu^N|-|\mu^{N-1}| = \beta$; soit, enfin, elle se trouvait sur la $\alpha(l-a)$-ième ligne, et dans ce cas $|\mu^N|-|\mu^{N-1}| = \alpha$; on voit que dans tous les cas, $|\mu^N| - |\mu^{N-1}| \geq \delta(\alpha,\beta)$.
Pour achever la preuve du théorème, il ne reste plus qu'à calculer $\alpha+\beta$ et $\gamma+c$. Or, on a vu que $\widehat{\mu}_1(a,\alpha,\beta,c,\gamma) = (\alpha+\beta)a - \beta + c$; comme les transformations effectuées ne changent pas la première part, on en déduit $\lambda_1 = (\alpha+\beta)a-\beta+c$. Par ailleurs, on a $r = (\alpha+\beta)(l-a)+\beta+\gamma$. On en déduit donc que

$$n = r + \lambda_1 = (\alpha+\beta)\, l + (\gamma + c). \tag{1}$$

Or, on a vu que $\gamma \leq l-a$ et $c \leq a$. Par convention, on exclut les égalités $\gamma = l-a$ et $c = a$, car $\mu(a,\alpha,0,a,l-a) = \mu(a,\alpha+1,0,0,0)$; on a donc $\gamma + c < l$. La formule (1) exprime donc la division euclidienne de n par l, concluant la preuve du théorème.

•

2 Suite spectrale de Borel Le-Potier

Dans ce paragraphe, je rappelle la définition de la suite spectrale de Borel Le-Potier, qui est à la base de mes théorèmes d'annulation.
Soit $\pi : Y \to X$ une fibration localement triviale propre et E un fibré vectoriel sur Y. Soit $\Omega^i_{Y/X}$ le fibré des i-formes sur Y relatives à π, défini par $\Omega^i_{Y/X} = \Lambda^i \Omega_{Y/X}$ et la suite exacte de fibrés sur Y suivante :

$$0 \to \pi^*\Omega X \xrightarrow{\pi^*} \Omega Y \to \Omega_{Y/X} \to 0.$$

Soient aussi les fibrés $G^{t,p} := \Omega^{p-t}_{Y/X} \otimes \pi^*\Omega^t_X$. Pour chaque p, on a [LP 77] une suite spectrale aboutissant sur $H^{p,q}(Y,E)$ et dont les termes d'ordre 1 sont :

$${}^pE_1^{t,q-t} = H^q(Y, G^{t,p} \otimes E).$$

Pour calculer les groupes de cohomologie $H^q(Y, G^{t,p} \otimes E)$, on utilise une suite spectrale de Leray. La suite spectrale de termes d'ordre 2

$${}^{p,t}E_2^{k,j-k} = H^k(X, R^{j-k}\pi_* G^{t,p}) = H^{t,k}[X, R^{j-k}\pi_*(E \otimes \Omega^{p-t}_{Y/X})] \tag{2}$$

aboutit sur ${}^pE_1^{t,j}$. Introduisons enfin une notation :

Notation 3 *Si $\pi : Y \to X$ est une fibration et E un fibré vectoriel sur Y, notons $R^{p,q}\pi_* E$ le faisceau $R^q\pi_*(E \otimes \Omega^p_{Y/X})$.*

3 Theoreme d'annulation pour un produit tensoriel de crochets

W. Nahm et F. Laytimi ont montré un théorème d'annulation pour la cohomologie d'une puissance de Schur correspondant à un crochet d'un fibré ample [LN 02, th 2.1]. Dans ce paragraphe, je généralise leur résultat à un produit de crochets. Si α et β sont des entiers, notons $Z^{\alpha,\beta}$ le foncteur S_λ pour la partition λ de longueur $\alpha+1$ et de poids β telle que $\lambda_1 = \beta-\alpha$ et $\forall 1 < i \leq \alpha+1, \lambda_i = 1$.

Proposition 4 *Soit E un fibré ample de rang e sur une variété X projective et lisse de dimension n. Soit $a \in \mathbb{N}$, $(k_i)_{1\leq i\leq a}$ et $(\alpha_i)_{1\leq i\leq a}$ des entiers tels que $\alpha_i < k_i$; soit $\sigma = \sum \alpha_i$ et $k = \sum k_i$; alors,*

$H^{p,q}(X, \otimes_i Z^{k_i-\alpha_i-1,k_i}E) = 0$ *si* $q > n-p+[\delta(n-p)+\sigma][ae-k+2\sigma]-\sigma(\sigma+1)$.

Posons $Q(p,\sigma) = n-p+[\delta(n-p)+\sigma][ae-k+2\sigma]-\sigma(\sigma+1)$.

Demonstration : Reprenant les idées de [LN 02], soit $C_2^k = \frac{k(k-1)}{2}$ et δ la fonction définie par $\forall n \in \mathbb{N}, C_2^{\delta(n)} \leq n < C_2^{\delta(n)+1}$. On peut alors définir un ordre sur $\mathbb{N}^2$. Celui-ci est donné par l'ordre lexicographique sur $\mathbb{N}^3$ et l'injection $(x,\sigma) \mapsto (\delta(x)+\sigma, x - C_2^{\delta(x)}, \sigma)$. Notons $\mathcal{N}$ l'ensemble $\mathbb{N}^2$ muni de cet ordre. L'intérêt principal de $\mathcal{N}$ est qu'il vérifie :

Lemme 2 *[LN 02] : Pour tous $(x,\sigma) \in \mathcal{N}, \mu \in \mathbb{Z}-\{0\}$, si $r = \delta(x)$ et si $x+\mu r \in \mathbb{N}$, alors $(x+\mu r, \sigma-\mu) < (x,\sigma)$.*

On démontre la proposition par récurrence sur les couples $(n-p,\sigma) \in \mathcal{N}$. Cette récurrence peut sembler peu naturelle et troubler le lecteur ; aussi, je donne les 10 premiers éléments de $\mathcal{N}$ pour qu'il puisse suivre plus aisément les premiers pas de la récurrence.

	$n-p$	σ	$\delta(n-p)+\sigma$	$n-p-C_2^{\delta(n-p)}$	σ
1	0	0	1	0	0
2	1	0	2	0	0
3	0	1	2	0	1
4	2	0	2	1	0
5	3	0	3	0	0
6	1	1	3	0	1
7	0	2	3	0	2
8	4	0	3	1	0
9	2	1	3	1	1
10	5	0	3	2	0

Si on place ces numéros d'apparition sur un plan repéré par $n-p$ selon l'axe des abscisses et σ selon l'axe des ordonnées, on obtient le schéma suivant :

$$\begin{array}{cccccccc}
 & & & & & & & \sigma \\
 & & & & & & & \uparrow \\
 & & & & & & & 7 \\
 & & & & & 9 & 6 & 3 \\
n-p & \leftarrow & 10 & 8 & 5 & 4 & 2 & 1
\end{array}$$

La méthode employée est celle maintenant classique qui consiste à constater que les groupes de cohomologie que l'on cherche à annuler apparaissent dans une suite spectrale qui calcule la cohomologie d'un fibré en droites ample sur une variété adéquate $\mathcal{Y}$ elle-même fibrée au-dessus de X. Ce dernier groupe est nul par le théorème de Kodaira si q est suffisant ; il suffit donc de s'assurer que ces groupes "passent à travers" la suite spectrale, ce qui fait intervenir une récurrence.
Soit donc $X^n, E^e, p, q, (k_i)$ et σ_0 vérifiant les hypothèses de la proposition fixés. On va montrer l'annulation de $H^{p,q}(X, \oplus_{\alpha_i:\sum \alpha_i=\sigma_0} \otimes_i Z^{k_i-\alpha_i-1,k_i}E)$ en supposant la proposition vraie pour tous les couples (p',σ) tels que $(n-p',\sigma) < (n-p,\sigma_0)$. Supposons la suite (k_i) croissante. Soit $r=\delta(n-p)$; soit aussi l_i et s_i les quotients et restes de la division de k_i par r : $k_i = rl_i + s_i$. Un premier lemme assure que l'on peut supposer que $e > r$ tous les l_i sont supérieurs ou égaux à

$$l_0 = \begin{cases} \frac{r\sigma_0+(n-p)}{r-1} & si \quad r > 2 \\ \sigma+1 & si \quad r=1 \end{cases}$$

Lemme 3 *Soit l un entier tel que la proposition 4 soit vraie pour une certaine valeur de $n-p$ et de σ, et pour toutes les variétés et tous les fibrés vectoriels amples, à la condition que tous les l_i soient supérieurs ou égaux à l et que $e > \delta(n-p)$. Alors cette proposition est vraie sans restrictions pour ces valeurs de e et de σ.*

Demonstration : Soit E un fibré sur une variété X et L un fibré en droites ample sur X. On peut supposer $l > \delta(n-p)$. La proposition est vraie pour $l'_i = l_i + l$ (c'est-à-dire, puisque $k'_i = rl'_i + s_i$, pour $k'_i = k_i + rl$) et les fibrés amples E' de rang supérieur à $\delta(n-p)$. Pour le fibré ample $E' = E \oplus L^{\oplus rl}$ de rang $e+rl$, et pour $k'_i = k_i + rl$, on a donc l'annulation de $H^{p,q}(X, \otimes_i Z^{k'_i-\alpha_i-1,k'_i}E')$ si

$$\begin{array}{rcl} q & > & [\delta(n-p)+\sigma][a(e+rl)-k'+2\sigma]-\sigma(\sigma+1) \\ & = & [\delta(n-p)+\sigma][ae-k+2\sigma]-\sigma(\sigma+1). \end{array}$$

Mais comme $Z^{k'_i-\alpha_i-1,k'_i}E' \supset Z^{k_i-\alpha_i-1,k_i}E \otimes L^{\otimes rl}$, on en déduit que pour toute variété X, tous fibrés E et L, $H^{p,q}(X, \otimes_i Z^{k_i-\alpha_i-1,k_i}E \otimes L^{\otimes arl}) = 0$ si $q > [\delta(n-p)+\sigma][ae-k+2\sigma]-\sigma(\sigma+1)$. En vertu du lemme suivant, ce résultat reste vrai si L est trivial, et le lemme 3 est donc prouvé. •

Lemme 4 *Soient n, p, q_0, e des entiers et λ une partition tels que $H^{p,q}(X, S_\lambda E \otimes L) = 0$ pour toute variété projective lisse X de dimension n, tout fibré vectoriel ample E de rang e et tout fibré en droites ample L sur X, si $q > q_0$.*
Alors $H^{p,q}(X, S_\lambda E \otimes L) = 0$ sous les mêmes conditions, sauf que L est seulement nef.

Demonstration : La démonstration du lemme 1.5.1 p.128 dans [Man 97] est valable sous ces hypothèses. •

On suppose donc dorénavant que $l_1 \geq l_0$ et $e > r$. Si $0 < s < r < e$ sont des entiers et V un espace vectoriel de dimension e, on notera $M_{s,r}(V)$ la variété de drapeaux absolue constituée de l'ensemble des couples (V_r, V_s) de sous-espaces vectoriels de V avec

$$0 \subset V_r \subset V_s \subset V, \quad codim(V_r) = r, \quad codim(V_s) = s.$$

Soit alors $\mathcal{M}_{s,r}(E)$ la variété de drapeaux relative à E, c'est-à-dire la variété fibrée au-dessus de X dont la fibre au-dessus du point x s'identifie à la variété $M_{s,r}(E_x)$ des drapeaux de la forme

$$(0 \subset E_r \subset E_s \subset E_x), \quad codim(V_r) = r, \quad codim(V_s) = s.$$

Soit aussi $\mathcal{Q}^{l+1,l}$ le fibré en droites sur $\mathcal{M}_{s,r}(V)$ dont la fibre au-dessus du drapeau précédent s'identifie à $(\det(E_x/E_s))^{l+1} \otimes (\det(E_s/E_r))^l = \det(E/E_s) \otimes (\det(E/E_r))^l$. Considérons alors le produit au-dessus de X défini par $\mathcal{Y} := \times M_{s_i,r}(E)$ et le fibré en droites $\mathcal{L}$ au-dessus de $\mathcal{Y}$ égal au produit $\otimes \pi_i^* \mathcal{Q}_i$, si π_i désigne la projection de $\mathcal{Y}$ sur $\mathcal{M}_{s_i,r_i}(E)$ et $\mathcal{Q}_i$ le fibré $\mathcal{Q}^{l_i+1,l_i}$ sur cette variété de drapeaux relative. Comme $e > r > s_i$ pour tout i, cette variété a bien un sens.

Laurent Manivel a étudié une partie de la cohomologie des fibrés en droites $Q^{l+1,l}$ sur une variété de drapeaux absolue, partie suffisante pour établir notre proposition. Néanmoins, j'ai besoin de généraliser ses résultats à un produit de variétés de drapeaux. Cette généralisation sera une conséquence facile de la formule de Kunneth.

Soit $M_{s,r}(V)$ une variété de drapeaux absolue et $Q^{l+1,l}$ comme précédemment. Notons $\lambda = l - 1, t(r) = \frac{r(r+1)}{2}$ et $k = rl + s$. Notons $\pi_{r,s} = \frac{s(2r-s+1)}{2}$ et pour π un entier,

$$n_s(\pi) = \#\{c \in \{0,1\}^r, |c| = s \text{ et } \sum ic_i = \pi\}$$

Ainsi $n_s(\pi) = 0$ si $\pi < 0$ ou $\pi > \pi_{r,s}$.
Alors, on a [Man 92, proposition 1.2.1, p.111 et lemme 1.3.1, p.114] :

Proposition 5 *Supposons $p = \lambda.t(r) + \pi$ avec $\pi \geq \pi_{r,s} - k + l$. Alors*

$$H^{p,q}(M_{s,r}(V), Q^{l+1,l}) = \bigoplus_{\alpha=0}^{l} n_s(\pi + r\alpha)\delta_{q,p-r\lambda-s+\alpha} Z^{k-\alpha-1,k} V.$$

Considérons maintenant un produit $Y = \times M_{s_i,r}(V)$ de variétés de drapeaux et le fibré en droites $L = \otimes \pi_i^* Q_i$, avec Q_i le fibré Q^{l_i+1,l_i} sur $M_{s_i,r}(V)$ et π_i la projection naturelle de Y sur $M_{s_i,r}$. Je propose alors la généralisation suivante de la proposition 5, si λ_i désigne $l_i - 1$, $\lambda = \sum \lambda_i$, $l = \sum l_i$, $s = \sum s_i$, et $k = \sum k_i$:

Proposition 6 *Supposons que* $p = \lambda.t(r) + \pi$ *avec* $\pi \geq \sum \pi_{r,s_i} - k + l$. *Soit* $\sigma = q + r\lambda + s - p$. *Alors*

$$H^{p,q}(Y,L) = n_s(\pi + r\sigma) \bigoplus_{\alpha_i : \sum \alpha_i = \sigma} \otimes_i Z^{k_i - \alpha_i - 1, k_i} V.$$

Dans cette proposition, $n_s(\pi)$ désigne le nombre

$$n_s(\pi) := \sum_{\sum \pi_i = \pi} n_{s_i}(\pi_i) = \#\{c_{i,j} \in \{0,1\}^{r^2} : \forall i \ |c_i| = s_i \text{ et } \sum_{i,j} j c_{i,j} = \pi\}.$$

Cette proposition montre que le fait qu'un groupe $\otimes_i Z^{k_i - \alpha_i - 1, k_i} V$ soit ou non dans $H^{p,q}(Y,L)$ ne dépend que de la somme σ des α_i, et non des valeurs de tous les α_i ; c'est ce qui fait que l'on obtient un théorème d'annulation où la borne ne dépend aussi que de σ.

Demonstration : On applique tout d'abord la formule de Kunneth :

$$H^{p,q}(Y,L) = \bigoplus_{\sum p_i = p, \sum q_i = q} \otimes H^{p_i,q_i}(M_{s_i,r}, Q^{l_i,l_i+1}).$$

Ecrivons chaque p_i comme $\lambda_i t(r) + \pi_i$. Si une telle suite d'entiers p_i fournit un groupe non nul, alors par la proposition 5, on a pour tout i la relation $\pi_i \leq \pi_{r,s_i}$. Si donc il existe i tel que $\pi_i < \pi_{r,i} - k_i + l_i$, comme $\pi = \sum \pi_i$, on en déduit que $\pi < \sum \pi_{r,i} - k + l$, ce qui contredit l'hypothèse de la proposition. Ainsi, pour tout i on a $\pi_{r,s_i} - k_i + l_i \leq \pi_i \leq \pi_{r,s_i}$, et donc on peut en déduire par la proposition 5 que

$$H^{p_i,q_i}(M_{s_i,r}, Q^{l_i,l_i+1}) = \oplus_{\alpha_i} n_s(\pi_i + r\alpha_i) \delta_{q_i, p_i - r\lambda_i - s_i + \alpha_i} Z^{k_i - \alpha_i - 1, k_i} V.$$

Si donc $\otimes_i Z^{k_i - \alpha_i - 1, k_i} V$ apparaît dans notre groupe de cohomologie, cela implique que $q_i = p_i - r\lambda_i - s_i + \alpha_i$, et par sommation que $q = p - r\lambda - s + \sigma$. Enfin, la multiplicité de ce groupe est bien la somme sur les π_i des produits des multiplicités de $Z^{k_i - \alpha_i - 1, k_i} V$ dans $H^{p_i,q_i}(M_{s_i,r}, Q^{l_i,l_i+1})$, soit $n_s(\pi)$. •

Notons $P_{\max} = \lambda.t(r) + \sum \pi_{r,s_i}$ et $Q_{\max} = \lambda.t(r-1) + \sum \pi_{r,s_i} - S$. Une conséquence de la proposition 6 est :

Proposition 7 *Soit* p *et* q *deux entiers. Alors :*
– *Si* $p > P_{max}$ *ou* $q > Q_{max}$, *alors* $H^{p,q}(Y,L) = 0$.
– *Si* $p \geq P_{max} - r\sigma$ *ou* $q \geq Q_{max} - (r-1)\sigma$, *alors*

$$H^{p,q}(Y,L) \subset \bigoplus_{\sum \alpha_i \leq \sigma} \otimes_i Z^{k_i - \alpha_i - 1, k_i} V.$$

On considère maintenant la fibration en variétés de drapeaux $\pi : \mathcal{Y} \to X$ introduite précédemment. Soit ${}^{P}E_m^{i,j}$ la suite spectrale de Borel-Le Potier, avec $P = p + P_{\max} - r\sigma_0$. Le résultat souhaité va être conséquence de propriétés de cette suite spectrale. La première exhibe un terme de la suite spectrale de Borel-Le Potier qui contient le groupe que l'on veut annuler :

Lemme 5 *Soit* $q_0 = Q_{max} - (r-1)\sigma_0$. *Alors*

$$ {}^{P}E_1^{p,q+q_0-p} \supset H^{p,q}(X, \oplus_{\alpha_i : \sum \alpha_i = \sigma_0} \otimes_i Z^{k_i - \alpha_i - 1, k_i} E). $$

Demonstration : Comme on a supposé que $l_1(r-1) \geq r\sigma_0 + (n-p)$, on a

$$ \begin{aligned} n + P_{\max} - k_1 + l_1 &= n + P_{\max} - l_1(r-1) - s_1 \\ &\leq n + P_{\max} - r\sigma_0 - (n-p) \\ &= P_{\max} + p - r\sigma_0 = P. \end{aligned} $$

Ainsi, si $p \leq n$, alors $P - p \geq P - n \geq P_{\max} - k_1 + l_1$. Si $p > n$, il est clair que ${}^{P}E_1^{p,q+q_0-p} = 0$. On peut donc utiliser les propositions 6 et 7 pour obtenir des renseignements sur les $H^{P-p,q'}(Y, L)$ pour tous q'. Pour tous les entiers q_1 et q_2, on a (cf notation 3)

$$ {}^{P,p}E_2^{q_1,q_2} = H^{p,q_1}[X, R^{P-p,q_2}\pi_*\mathcal{L}]. $$

Si $f_{i,j}$ sont les applications de changement de cartes de E, on a vu que $R^{P-p,q_2}\pi_*\mathcal{L}$ est un fibré vectoriel de fibre type $H^{P-p,q_2}(Y,L)$ (soit $x \in X$; $Y = \pi^{-1}(x)$ et L est la restriction du fibré en droites $\mathcal{L}$ à Y), et dont les changements de carte sont induits par $f_{i,j}$. Supposons que $H^{P-p,q_2}(Y,L) = S_\lambda E_x$. Nous savons que $|\lambda| = \sum rl_i + s_i$. On en déduit donc que pour $g \in GL(E_x)$, l'application induite par g sur $H^{P-p,q_2}(Y,L)$ est $S_\lambda g$. En effet, c'est vrai $g \in SL(E_x)$ par le théorème de Bott et pour g une homothétie, puisque l'application induite par $\lambda.Id$ est $\lambda^{\sum rl_i + s_i}$ (l'action de $\lambda.Id$ sur le fibré tangent est triviale).
Les applications de changement de cartes de $R^{P-p,q_2}\pi_*\mathcal{L}$ sont donc les applications $S_\lambda f_{i,j}$, si $f_{i,j}$ est une application de changement de cartes de E. On a donc $R^{P-p,q_2}\pi_*\mathcal{L} = S_\lambda E$.

Or $H^{P-p,q_0}(Y,L) = \oplus_{\alpha_i : \sum \alpha_i = \sigma_0} \otimes_i Z^{k_i-\alpha_i-1,k_i} V$. En effet, si un terme de la forme $\oplus_{\alpha_i : \sum \alpha_i = \sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i} V$ est une composante de ce groupe, alors on doit avoir $\sigma = q_0 + rL + S - (P-p) = \sigma_0$. Par ailleurs, il est clair que $n_s(\sum \pi_{r,s_i}) = 1$. On en déduit donc que ${}^{P,p}E_2^{q,q_0} = H^{p,q}(X, \oplus_{\alpha_i : \sum \alpha_i = \sigma_0} \otimes_i Z^{k_i-\alpha_i-1,k_i} E)$.
Par ailleurs, si $q_2 > q_0$ et q_1 est quelconque, alors la proposition 7 montre que ${}^{P,p_1}E_2^{q_1,q_2} = 0$. Enfin, si $q_1 > q$ et $q_2 < q_0$, et si $H^{p,q_1}(X, \oplus_{\alpha_i : \sum \alpha_i = \sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i} E) \subset {}^{P,p_1}E_2^{q_1,q_2}$, alors on a $\sigma < \sigma_0$ et $H^{p,q_1}(X, \oplus_{\alpha_i : \sum \alpha_i = \sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i} E) = 0$ par l'hypothèse de récurrence.

Toutes les différentielles d_m issus de ou aboutissant sur ${}^{P,p}E_m^{q_0,q}$ sont donc nulles, et ${}^{P,p}E_\infty^{q_0,q} = H^{p_1,q_1}(X, \oplus_{\alpha_i:\sum \alpha_i=\sigma_0} \otimes_i Z^{k_i-\alpha_i-1,k_i}E)$, ce qui implique notre lemme. •

Pour vérifier que ce groupe passe à travers la suite spectrale de Borel-Le Potier, il suffit de montrer l'annulation des groupes ${}^{P}E_1^{p+m,q-p-m+\mathrm{sgn}(m)+q_0}$, pour tous les entiers m non nuls. Soit donc m un tel entier et p_1, q_1 et q_2 des entiers tels que ${}^{P,p_1}E_\infty^{q_1,q_2}$ soit un élément d'une filtration de ${}^{P}E_1^{p+m,q-p-m+\mathrm{sgn}(m)+q_0}$, c'est-à-dire des entiers tels que :
$p_1 = p+m$ et $q_1+q_2-p_1 = q-p-m+\mathrm{sgn}(m)+q_0$. Remarquons que de manière équivalente, on a $p_1 = p+m$ et $q_1+q_2 = q+q_0+\mathrm{sgn}(m)$.
De nouveau, il suffit de montrer que tous les groupes ${}^{P,p_1}E_2^{q_1,q_2}$ sont nuls. Or ceux-ci valent $H^{p_1,q_1}[X, H^{P-p_1,q_2}(\mathcal{Y},\mathcal{L})]$. Supposons que

$$\oplus_{\alpha_i:\sum \alpha_i=\sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i}V \subset H^{P-p_1,q_2}(Y,L).$$

Alors par la proposition 7, $P-p_1 \leq P_{\max} - r\sigma = (P-p) - r(\sigma-\sigma_0)$, donc $p_1 \geq p + r(\sigma-\sigma_0)$, soit $m \geq r(\sigma-\sigma_0)$. Par ailleurs cette même proposition assure que $q_2 \leq Q_{\max} - (r-1)\sigma = q_0 - (r-1)(\sigma-\sigma_0)$.
L'égalité $q_1+q_2 = q+q_0+\mathrm{sgn}(m)$ donne alors $q_1 \geq q+\mathrm{sgn}(m)+(r-1)(\sigma-\sigma_0)$. Notons $\mu = \sigma-\sigma_0$, comme nous avons vu que $m \geq r\mu$, $\mathrm{sgn}(m) \geq \mathrm{sgn}(\mu)$ et nous avons donc établi :

$$\begin{gathered}\sigma = \sigma_0 + \mu \\ p_1 \geq p + \mu r \\ q_1 \geq q + \mathrm{sgn}(\mu) + (r-1)\mu.\end{gathered}$$

Sous ces hypothèses, il ne nous reste plus, en utilisant l'hypothèse de récurrence, qu'à prouver l'annulation de $H^{p_1,q_1}(X, \oplus_{\alpha_i:\sum \alpha_i=\sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i}E)$. Remarquons tout d'abord que l'on peut supposer que $ae-k+\sigma \leq 0$, car sinon $\oplus_{\alpha_i:\sum \alpha_i=\sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i}E = 0$.

Lemme 6 *Dans l'ensemble ordonné $\mathcal{N}$, on a :* $(n-p_1,\sigma) < (n-p,\sigma_0)$

Demonstration : Cet ordre étant croissant selon les deux coordonnées (c'est-à-dire que si $x \leq x', y \in \mathbb{N}$, alors $(x,y) \leq (x',y)$ et $(y,x) \leq (y,x')$), cela découle du lemme 2. •

Lemme 7 *Si $q > Q(p,\sigma_0)$ alors $q_1 > Q(p_1,\sigma)$.*

Demonstration : On peut supposer que $p_1 = p+\mu r$ et $q_1 = q+\mu(r-1)+\mathrm{sgn}(\mu)$. Calculons alors $Q(p,\sigma_0) - Q(p+\mu r, \sigma_0+\mu) + \mathrm{sgn}(\mu) + (r-1)\mu$. Cette quantité vaut :
$2\mu r + 2\mu\sigma_0 + \mu^2 + \mathrm{sgn}(\mu) + (r+\sigma_0)(ae-k+2\sigma_0) - [\delta(n-p-\mu r)+\sigma_0+\mu][ae-k+2\sigma_0+2\mu]$. En remarquant que $2\mu r + 2\mu\sigma_0 = 2\mu(r+\sigma_0)$ et en factorisant par $r+\sigma_0$, on en déduit que cette quantité égale

$(ae-k+2\sigma_0+2\mu)(r+\sigma_0)-(ae-k+2\sigma_0-2\mu)[\delta(n-p+\mu r)+\sigma_0-\mu]+\mu^2+\mathrm{sgn}(\mu)$, soit $(ae-k+2\sigma)[r+\mu-\delta(r+\mu r)]+\mu^2+\mathrm{sgn}(\mu)$.
Comme $ae-k+\sigma \geq 0$, il découle de la définition de δ que ce nombre est positif. $\bullet$

Enfin, on a l'inégalité suffisante pour assurer que ${}^{P}E_{\infty}^{p,q-p+q_0} = 0$:

Lemme 8 *On a* $P+q+q_0-\dim\mathcal{Y} > 0$.

Demonstration : Tout d'abord,

$$2\sum \pi_{r,s_i} - S = 2rS + S - \sum s_i^2 - S \geq rS \text{ et } k = (L+a)r + S.$$

Ainsi :

$$\begin{array}{rl} & P+q+q_0-\dim\mathcal{Y} \\ = & p+q+Lt(r)+Lt(r-1)+2\sum \pi_{r,s_i} - S - (2r-1)\sigma_0 - n - ar(e-r) \\ \geq & \sigma_0(ae-k+\sigma_0) \end{array}$$

$\bullet$

Nous avons ainsi (péniblement) achevé la preuve de la proposition 4. Nous allons maintenant montrer une autre proposition qui rétablit la symétrie entre p et q :

Proposition 8 *Soit E un fibré ample de rang e sur une variété X projective lisse et de dimension n. Soit $a \in \mathbb{N}$, $(k_i)_{1\leq i\leq a}$ et $(\alpha_i)_{1\leq i\leq a}$ des entiers tels que $\alpha_i < k_i$; soit $\sigma = \sum \alpha_i$ et $k = \sum k_i$, alors*

$H^{p,q}(X, \otimes_i Z^{k_i-\alpha_i-1,k_i}E) = 0$ *si* $p > n-q+[\delta(n-q)+\sigma][ae-k+2\sigma]-\sigma(\sigma+1)$.

Demonstration : Posons $P(q,\sigma) = n-q+[\delta(n-q)+\sigma][ae-k+2\sigma]-\sigma(\sigma+1)$. La démonstration est très similaire à celle de la proposition 4. En effet, soit $r = \delta(n-q)$ et comme précédemment $\mathcal{Y} := \times M_{s_i,r}(E)$ et le fibré en droites $\mathcal{L}$ au-dessus de $\mathcal{Y}$ égal au produit $\otimes \pi_i^* \mathcal{Q}_i$, où l_i et s_i sont le quotient et le reste de la division euclidienne de k_i par r et $\mathcal{Q}_i = \mathcal{Q}^{l_i,l_i+1}$. Soit aussi $p+P_{max}-r\sigma_0$, le lemme 5 reste vrai et pour voir que ce groupe passe à travers la suite spectrale, il suffit comme précédemment de montrer en utilisant l'hypothèse de récurrence que si

$$\begin{gathered} \sigma = \sigma_0 + \mu \\ p_1 \geq p + \mu r \\ q_1 \geq q + \mathrm{sgn}(\mu) + (r-1)\mu, \end{gathered}$$

alors $H^{p_1,q_1}(X, \oplus_{\alpha_i:\sum\alpha_i=\sigma} \otimes_i Z^{k_i-\alpha_i-1,k_i}E) = 0$.
Ceci découle du fait que le lemme 8 reste vrai et que l'on a l'analogue des lemmes 6 et 7 :

Lemme 9 *Dans l'ensemble ordonné $\mathcal{N}$, on a :* $(n-q_1,\sigma) < (n-q,\sigma_0)$.

Demonstration : On peut aussi supposer que $q_1 = q + \operatorname{sgn}(\mu) + (r-1)\mu$ et $\sigma = \sigma_0 + \mu$. Alors si $\mu < 0$, $q_1 \geq q + r\mu$ et le lemme 2 s'applique. Si $\mu = 1$, ce même lemme fonctionne car $q_1 = q + r$. Pour $\mu > 1$, en posant $x = n - (q+r)$, ce lemme donne $(n-(q+r)-(\mu-1)\delta(n-(q+r)), \sigma+\mu) < (n-(q+r), \sigma+1)$. Comme $\delta(n-(q+r)) \leq \delta(n-q)-1$, on a $(n-(q+r)-(\mu-1)(r-1), \sigma+\mu) < (n-(q+r), \sigma+1)$.L'égalité $n-q_1 = n-q-1-\mu(r-1) = n-(q+r)-(\mu-1)(r-1)$ donne alors $(n-q_1, \sigma) < (n-q, \sigma_0)$, et le lemme est démontré. •

Lemme 10 *Si $p > P(q, \sigma_0)$ alors $p_1 > P(q_1, \sigma)$.*

Demonstration : On peut supposer que $p_1 = p + \mu r$ et $q_1 = q + \mu(r-1) + \operatorname{sgn}(\mu)$. Calculons alors $P(q, \sigma_0) - P(q + \mu(r-1) + \operatorname{sgn}(\mu), \sigma_0 + \mu) + \mu r$. Cette quantité vaut :

$$\begin{gathered} 2\mu r + 2\mu\sigma_0 + \mu^2 + \operatorname{sgn}(\mu) + (r+\sigma_0)(ae - k + 2\sigma_0) \\ -[\delta(n - q - \operatorname{sgn}(\mu) - \mu(r-1)) + \sigma_0 + \mu][ae - k + 2\sigma_0 + 2\mu]. \end{gathered}$$

En remarquant que $2\mu r + 2\mu\sigma_0 = 2\mu(r+\sigma_0)$ et en factorisant par $r + \sigma_0$, on voit que cette quantité égale

$$\begin{gathered} \mu^2 + \operatorname{sgn}(\mu) + (ae - k + 2\sigma_0 + 2\mu)(r + \sigma_0) \\ -(ae - k + 2\sigma_0 + 2\mu)[\delta(n - q - \operatorname{sgn}(\mu) - \mu(r-1)) + \sigma_0 + \mu], \end{gathered}$$

soit $(ae - k + 2\sigma)[r - \mu - \delta(n - q - \operatorname{sgn}(\mu) - \mu(r-1))] + \mu^2 - \operatorname{sgn}(\mu)$.
Comme $ae - k + \sigma \geq 0$ et que le lemme 9 implique
$\delta(n - q - \operatorname{sgn}(\mu) - \mu(r-1)) + \mu \leq \delta(n-q) = r$, ce dernier nombre est donc positif. •

On peut maintenant regrouper les propositions 4 et 8 sous la forme du

Theoreme 2 *Soit E un fibré ample de rang e sur une variété X de dimension n, $a \in \mathbb{N}$, $(k_i)_{1 \leq i \leq a}$ et $(\alpha_i)_{1 \leq i \leq a}$ des entiers tels que $\alpha_i < k_i$; soit $\sigma = \sum \alpha_i$ et $k = \sum k_i$, alors*

$$\begin{gathered} H^{p,q}(X, \otimes_i Z^{k_i - \alpha_i - 1, k_i} E) = 0 \\ si \\ p + q > n + [\min[\delta(n-p), \delta(n-q)] + \sigma][ae - k + 2\sigma] - \sigma(\sigma+1). \end{gathered}$$

4 Resultats topologiques en petit co-rang

4.1 Resultats topologiques

L'objet de ce paragraphe est la démonstration du

Theoreme 3 *Soit E un fibré vectoriel de rang e, sur une variété X lisse, projective et de dimension n. Supposons que E est muni d'une forme quadratique à valeurs dans un fibré en droites L. Soit k un entier et supposons que*
– $\dim D_k(E) = \rho := n - t(e-k)$.
– $S^2E^* \otimes L$ *est ample.*
– $e - k \leq 4$.
Alors, l'application de restriction $H^q(X, \mathcal{O}_X) \to H^q(D_k(E), \mathcal{O}_{D_k(E)})$ est un isomorphisme pour $0 \leq q < \rho - 1$, et est injective pour $q = \rho - 1$.

Demonstration : Notons $D := D_k(E)$. Tout d'abord, il existe une résolution du faisceau $\mathcal{O}_D$ sur $\mathcal{O}_X$ par des fibrés vectoriels : le i-ième terme R^i d'une telle résolution est donné par

$$R^i := \bigoplus_{\substack{\lambda = (2l, \mu, \mu^*) \\ |\mu| + l(2l-1) = i}} S_{\lambda(k-1)} E \otimes L^{-l(2l+k-1)}.$$

Dans cette formule, l'expression $\lambda = (2l, \mu, \mu^*)$ signifie que la partition λ est de rang $2l$, que $\lambda_i = 2l + \mu_i$ si $1 \leq i \leq 2l$ et $\lambda_{i+2l} = \mu_i^*$. Si λ est une partition, $\lambda(k-1)$ est obtenue en intercalant $k-1$ parts égales au rang de λ à λ : si par exemple $\lambda = \ydiagram{3,2,1}$, alors $\lambda(2) = \ydiagram{3,2,2,2,1}$. Si $k = 0$, alors $(2l, \mu, \mu^*)(-1)$ est la partition ν telle que $\nu_i = 2l + \mu_i$ pour $1 \leq i \leq 2l-1$, $\nu_{2l} = \mu_1^* + \mu_{2l}$, et $\nu_{2l-1+i} = \mu_i^*$ pour $i \geq 2$.

L'existence de cette résolution, bien connue des spécialistes, peut se justifier de la façon suivante : soit Y l'espace total du fibré $S^2E^* \otimes L$ et $D \subset Y$ le schéma des formes symétriques de rang au plus k. Par [JPW 81, théorème 3.19, p.139] et [Nie 81], on a une résolution de $\mathcal{O}_D$ par des $\mathcal{O}_Y$-modules localement libres analogue à celle que j'ai décrite. La section s de $S^2E^* \otimes L$ induit un morphisme $X \to Y$ et, comme l'explique Nielsen [Nie 81], on peut tirer en arrière cette résolution pour obtenir la résolution de $\mathcal{O}_{D_k(E)}$ souhaitée. Pour cela, il suffit en effet d'annuler les faisceaux $\underline{Tor}_i^Y(\mathcal{O}_D, \mathcal{O}_X)$ pour $i > 0$, ce qui résulte d'une version du corollaire 1.10 de [PS 73], où l'on peut remplacer, avec leurs notations, l'hypothèse $prof(B_\mathfrak{p}) \geq prof(A_{\varphi^{-1}(\mathfrak{p})})$ par l'hypothèse $prof(B_\mathfrak{p}) \geq dp_A(M)$. Pour annuler $\underline{Tor}_i^Y(\mathcal{O}_D, \mathcal{O}_X)$, on applique ce corollaire à $Spec(B)$ un ouvert affine de X trivialisant $S^2E^* \otimes L$, $Spec(A) \subset Y$ l'image réciproque de $Spec(B)$ par la projection $Y \to X$ et M le A-module définissant le faisceau $\mathcal{O}_D$ sur $Spec(A)$. Pour tout idéal $\mathfrak{p} \subset B$ correspondant à un point du support de $M \otimes_A B$, on a $prof(B_\mathfrak{p}) \geq t(e-k)$, puisque X est supposée lisse et $\dim D_k(E) = n - t(e-k)$. L'hypothèse $prof(B_\mathfrak{p}) \geq dp_A(M)$ est donc bien vérifiée.

Convenons qu'une partition λ de la forme $(2l, \mu, \mu^*)(k-1)$ sera dite $(k-1)$-*symétrique* (si $k-1=0$, ceci signifie que $\lambda^* = \lambda$). Si λ est $(k-1)$-symétrique, notons $i(\lambda, k)$ l'entier tel que $S_\lambda E^* \otimes L^{-l(2l+k-1)} \subset R^{i(\lambda,k)}$.

En utilisant cette résolution, il est facile de montrer que le théorème 3 est conséquence de la proposition suivante (cet argument est par exemple expliqué dans [Man 97]) :

Proposition 9 *Soit F un fibré vectoriel ample de rang e sur une variété X projective lisse de dimension n. Soit k un entier tel que $0 < e-k \leq 4$ et λ une partition $(k-1)$-symétrique. Alors, $H^{n,q}(X, S_\lambda F) = 0$ si $q > t(e-k)+1-i(\lambda, k)$.*

Demonstration : Puisque $e-k \leq 4$, le rang $2l$ d'une partition $(k-1)$-symétrique et de longueur inférieure ou égale à e est nécessairement inférieur ou égal à 4 ; on a donc $l = 1$ ou $l = 2$. Les partitions obtenues avec $l = 2$ ne nous posent pas de problème, car le théorème A' de [Man 97] donne la borne $10 - i$. Il suffit donc de traiter le cas $l = 1$.
Pour simplifier les notations, j'ai supposé, dans le tableau suivant, que de plus $k = 1$, et l'ai listé toutes les partitions λ à considérer, et indiqué, en face, l'entier q_0 tel que si $q > q_0$, alors $H^{n,q}(X, S_\lambda E) = 0$ pour tout fibré E ample de rangs 4 et 5. La dernière colonne indique le numéro du lemme qui montre l'annulation du groupe de cohomologie pour la borne indiquée. L'indication A' renvoie au théorème A' de [Man 97]. La colonne intermédiaire indique la valeur $t(e-k)+1-i$ de q_0 maximale pour montrer le théorème 3. Lorsque k prend d'autres valeurs que 1 mais que $e-k$ reste égal à 3 ou 4, puisque les lemmes indiqués donnent une borne qui ne dépend que de $e-k$, on obtient encore la proposition 9.

$rang(E) = 4$

partition	q_0	7-i	lemme
$\ydiagram{2,2}$	4	6	A'
$\ydiagram{3,2,1}$	5, 4	5	11, 12
$\ydiagram{3,3,2}$	4	4	15
$\ydiagram{4,2,1,1}$	2	4	A'
$\ydiagram{4,4,2,2}$	0	2	A'
$\ydiagram{4,3,2,1}$	2	3	17

$rang(E) = 5$

partition	q_0	11-i	lemme
$\ydiagram{5,5,2,2,2}$	0	4	A'
$\ydiagram{5,2,1,1,1}$	3	7	A'
$\ydiagram{2,2}$	6	10	A'
$\ydiagram{3,2,1}$	8	9	11
$\ydiagram{5,4,2,2,1}$	4, 3	5	13, 14
$\ydiagram{3,3,2}$	8	8	15
$\ydiagram{4,2,1,1}$	8	8	15
$\ydiagram{4,4,2,2}$	6	6	16
$\ydiagram{5,3,3,1,1}$	4	6	17
$\ydiagram{4,3,2,1}$	7, 6	7	19, 20

4.2 Theoremes d'annulation

Dans ce paragraphe, je montre les lemmes annoncés dans le précédent. Soit X une variété projective lisse de dimension n, E un fibré ample de rang e sur X, et L un fibré en droites nef.
Pour des raisons qui sont expliquées après le lemme 13, je note $\lambda = [a,b,c,d]$ la partition de rang 2 telle que $\lambda_1 = a+2, \lambda_2 = b+2, \lambda_1^* = e-c, \lambda_2^* = e-d$. Par exemple, $S_{[0,0,0,0]}E = (\det E)^2$. De même, je note $\lambda = [a,b]$ la partition de rang 1 telle que $\lambda_1 = a+1$ et $\lambda_1^* = e-b$. Par la suite, $a \leq b, c \leq d, t, q$ sont des entiers positifs ou nuls.

Lemme 11 $$\begin{array}{lll} H^{n,q}(X, S_{[1,c+1]}E \otimes S_{[0,d]}E \otimes L^t) & = 0 & \textit{si } q > 2c+d+2, \textit{ et} \\ H^{n-1,q}(X, S_{[0,c]}E \otimes S_{[0,d]}E \otimes L^t) & = 0 & \textit{si } q > 2c+d+1. \end{array}$$

Remarque : Le théorème 2 donne la borne $2c+2d+2$ au lieu de $2c+d+2$ et les autres résultats que je connais ont une borne qui dépend de k (par exemple, [Man 97, théorème A] donne la borne $k+c+d$) ; il est clair que pour démontrer le théorème 3 par cette méthode, il faut une borne indépendante de k.
Demonstration : Pour montrer ce lemme, je vais appliquer la technique, nouvelle à ma connaissance, de "comparaison de suites spectrales".
Tout d'abord, on peut supposer que $e-c$ est pair. En effet, supposons que ce lemme est vrai pour $e-c$ pair. Si $e-c$ est impair et M est un fibré en droites ample sur X, alors on peut considérer le fibré vectoriel $E \oplus M$ de rang $e+1$; comme $e+1-c$ est alors pair, on peut appliquer le lemme, et comme pour le lemme 3, on en déduit que $H^{n,q}(X, S_{[1,c+1]}E \otimes S_{[0,d]}E \otimes L^t) = 0$ si $q > 2c+d+2$ et $H^{n-1,q}(X, S_{[0,c]}E \otimes S_{[0,d]}E \otimes L^t) = 0$ si $q > 2c+d+1$.
Soit alors un entier l tel que $e-c = 2l$. Considérons d'abord la variété $Y = \mathbb{P}E^* \times_X \mathbb{P}E^*$, notons $\pi : Y \to X$ la projection naturelle, et considérons le fibré en droites $\mathcal{L} = \mathcal{O}(2l, e-d) \otimes \pi^* L^t$ sur Y. Comme $\mathcal{O}(2l, e-d)$ est ample (en effet, E est ample) et $\pi^* L$ nef, $\mathcal{L}$ est ample. Soit aussi $P = n+(2l-1)+(e-d-1)-1$. Déterminons les groupes ${}^{P,t}E_2^{i,j}$ de la suite spectrale de Leray introduite au paragraphe 2. Ceux-ci valent, par la formule (2), $H^{t,i}[X, R^{P-t,j}\pi_* \mathcal{O}(2l, e-d) \otimes L^t]$. Ces groupes sont donc nuls si $t > n$ ou $P-t > (2l-1)+(e-d-1)$. En effet, si V est un espace vectoriel fixé et l et m sont des entiers positifs, alors $H^{i,j}[\mathbb{P}V, \mathcal{O}(l)] = 0$ si $i \geq l$ (conséquence directe de la proposition 1), et donc $H^{i,j}[\mathbb{P}V \times \mathbb{P}V, \mathcal{O}(l,m)] = 0$ si $i \geq l+m$ (formule de Künneth). Les seules valeurs possibles sont donc $t = n$ ou $t = n-1$, et on a

$$\begin{array}{rcl} {}^{P,n-1}E_2^{i,0} & = & H^{n-1,i}(X, S_{[0,c]}E \otimes S_{[0,d]}E \otimes L^t), \quad \text{et} \\ {}^{P,n}E_2^{i,0} & = & H^{n,i}(X, L^t \otimes R^{(2l-1)+(e-d-1)-1,0}\pi_* \mathcal{O}(2l, e-d)) \\ & = & H^{n,i}(X, S_{[1,c+1]}E \otimes S_{[0,d]}E \otimes L^t) \\ & \oplus & H^{n,i}(X, S_{[0,c]}E \otimes S_{[1,d+1]}E \otimes L^t). \end{array}$$

Ces suites spectrales sont donc dégénérées, et on en déduit que

$$\begin{array}{rcl} {}^{P}E_1^{n-1,q-n+1} & = & H^{n-1,q}(X, S_{[0,c]}E \otimes S_{[0,d]}E \otimes L^t), \\ {}^{P}E_1^{n,q-n} & = & H^{n,q}(X, S_{[1,c+1]}E \otimes S_{[0,d]}E \otimes L^t) \\ & \oplus & H^{n,q}(X, S_{[0,c]}E \otimes S_{[1,d+1]}E \otimes L^t), \end{array}$$

et les autres termes sont nuls. Par ailleurs, ${}^{P}E_{\infty}^{p,q-p}$ est une composante de $H^{P,q}(X,\mathcal{L})$; il est donc nul par le théorème de Kodaira si

$$P+q>n+2(e-1),\ \text{soit}\ q>n+2(e-1)-P=c+d+1.$$

On en déduit donc que dans ce cas, la flêche connectant ${}^{P}E_1^{n-1,q-n+1}$ et ${}^{P}E_1^{n,q-n+1}$ est un isomorphisme, ce qui conduit à :

$$\begin{array}{ll} & H^{n-1,q}(X,S_{[0,c]}E\otimes S_{[0,d]}E\otimes L^t)\\ = & H^{n,q+1}(X,S_{[1,c+1]}E\otimes S_{[0,d]}E\otimes L^t)\\ \oplus & H^{n,q+1}(X,S_{[0,c]}E\otimes S_{[1,d+1]}E\otimes L^t)\ \text{si}\ q>c+d+1.\end{array}$$

En considérant une autre suite spectrale, on va obtenir une autre égalité faisant intervenir ces termes ; en comparant ces égalités, on prouvera qu'ils sont nuls. Soit donc maintenant $Y=G_2(E^*)\times_X \mathbb{P}E^*$ et $\mathcal{L}=\mathcal{O}(l,e-d)$ sur Y, et posons $P=n+3(l-1)+(e-d-1)-1$. On a alors de façon similaire ${}^{P}E_1^{n-1,q-n+l}=H^{n-1,q}(X,S_{[0,c]}E\otimes S_{[0,d]}E\otimes L^t)$, mais par contre, ${}^{P}E_1^{n,q-n+l-1}=H^{n,q}(X,S_{[0,c]}E\otimes S_{[1,d+1]}E\otimes L^t)$. Ceci conduit à l'égalité

$$\begin{array}{rcl} H^{n-1,q}(X,S_{[0,c]}E\otimes S_{[0,d]}E\otimes L^t) & = & H^{n,q+1}(X,S_{[0,c]}E\otimes S_{[1,d+1]}E\otimes L^t)\\ & \text{si} & q>2c+d+1.\end{array}$$

En comparant ces égalités, on obtient

$$H^{n,q}(X,S_{[1,c+1]}E\otimes S_{[0,d]}E\otimes L^t)=0,$$

soit la première affirmation du lemme. En utilisant la règle de Littlewood-Richardson, on va montrer l'autre résultat du lemme. En effet, rappelons que $c\leq d$; si $c=d$, on peut conclure ; supposons donc $c<d$. Cette règle implique que

$$\begin{array}{rcl} S_{[1,c+1]}E\otimes S_{[0,d]}E & = & \bigoplus_{0\leq x\leq c+1} S_{[1,0,x,d+c+1-x]}E\bigoplus R_1\\ S_{[0,c]}E\otimes S_{[1,d+1]}E & = & \bigoplus_{0\leq x\leq c} S_{[1,0,x,c+d+1-x]}E\bigoplus R_2,\end{array}$$

où R_1 et R_2 sont des sommes de composantes de type $S_{[0,0,x,y]}$, que l'on sait déjà annuler. L'annulation de la cohomologie de $S_{[1,c+1]}E\otimes S_{[0,d]}E\otimes L^t$ entraine donc celle de $S_{[0,c]}E\otimes S_{[1,d+1]}E\otimes L^t$, et le lemme est alors conséquence d'une des égalités démontrées.

•

Notons que ce lemme montre par exemple $H^{n,q}(X,S_{[1,0,c,c+1]}E)=0$ si $q>3c+2$, car $S_{[1,0,c,c+1]}E\subset S_{[1,c+1]}E\otimes S_{[0,c]}E$. Ce n'est pas la borne indiquée dans le tableau p. 517. En effet, pour cette partition particulière, on peut un peu raffiner le raisonnement précédent. Auparavant, je souhaite faire une remarque d'ordre général qui allègera les calculs.

En généralisant le raisonnement utilisé pour le lemme précédent, on voit que l'on obtient simultanément l'annulation de $H^{p_1,q}(X,S_{\lambda_1}E)$ et celle de

$H^{p_2,q}(X, S_{\lambda_2}E)$; supposons que $p_1 < p_2$: la borne q_1 obtenue pour le groupe $H^{p_1,q}(X, S_{\lambda_1}E)$ est la différence entre la dimension de Y et la somme $i+j+p_1$, où i et j sont les entiers tels que si V est un espace vectoriel fixé, Y_V la fibre de la projection $Y \to X$, et $\mathcal{L}_V$ la restriction du fibré en droites $\mathcal{L}$ sur une telle fibre, alors $H^{i,j}(Y_V, \mathcal{L}_V) = S_\lambda V$. La borne q_2 pour $H^{p_2,q}(X, S_{\lambda_2}E)$ vaut $q_1 + t$, où t est l'entier tel que les flêches de la suite spectrale connectent $H^{p_1,q}(X, S_{\lambda_1}E)$ et $H^{p_2,q+t}(X, S_{\lambda_2}E)$. Attention, cette recette n'est bien sûr valable que si aucun autre terme de la suite spectrale ne vient compenser les termes que l'on veut annuler. Par ailleurs, on peut facilement exprimer cette différence :

Remarque 4 *Une suite spectrale donnée par un fibré en droites sur $G_r(E^*) \times_X G_s(E^*)$ peut montrer l'annulation de $H^{p,q}(X, S_{[a,b]}E \otimes S_{[c,d]}E \otimes L^t)$ pour $q > (r-1)a + rb + (s-1)c + sd$.*

Demonstration : En effet, comme je l'ai expliqué, il s'agit d'évaluer la différence entre $\dim G_r(V)$ et $i + j$, où i et j sont les entiers tels que $H^{i,j}(G_r(V), \mathcal{O}(l)) = S_{[a,b]}V$. Cette différence vaut $ra + (r-1)b$, car $i = (l-1)t(r) - ar, j = (l-1)t(r-1) - a(r-1)$, et $\dim V = rl - a + b$. •

Dans le lemme qui suit, je traite des partitions qui sont un cas particulier de celles traitées par le lemme 11 ; pour ces partitions particulières, je peux réduire de 1 la borne donnée par le lemme 11.

Lemme 12 *Supposons $c > 0$.*

$$\begin{array}{lll} H^{n,q}(X, S_{[1,0,c+1,c]}E \otimes L^t) & = 0 & \textit{si } q > 3c+1, \textit{ et} \\ H^{n-1,q}(X, S_{[0,0,c,c]}E \otimes L^t) & = 0 & \textit{si } q > 3c. \end{array}$$

Demonstration : Tout d'abord, on peut supposer que $e - c$ est impair et $e - c >> 0$. Dans la démonstration du lemme 11, j'ai montré comment tenir compte de L^t ; pour alléger les notations, je suppose dorénavant que $t = 0$.
Soit alors l un entier tel que $e - c = 2l + 1$. Considérons d'abord la variété $Y = \mathbb{P}E^* \times_X \mathbb{P}E^*$, et le fibré $\mathcal{O}(2l+1, 2l+1)$ sur Y. Comme dans la démonstration du lemme précédent, on montre que

$$2H^{n,q+1}(X, S_{[1,c+1]}E \otimes S_{[0,c]}) = H^{n-1,q}(X, S_{[0,c]}E \otimes S_{[0,c]}E) \text{ si } q > 2c+1.$$

Soit maintenant $\mathcal{O}(2l, l+1) \to \mathbb{P}E^* \times_X G_2(E^*)$. On a alors de façon similaire $H^{n,q+1}(X, S_{[1,c+2]}E \otimes S_{[0,c-1]}E) = H^{n-1,q}(X, S_{[0,c+1]}E \otimes S_{[0,c-1]}E)$ si $q > 3c$.
Par la règle de Littlewood-Richardson, $S_{[1,c+1]}E \otimes S_{[0,c]}E = S_{[1,c+2]}E \otimes S_{[0,c-1]}E \oplus S_{[1,0,c,c+1]}E$ (à des termes de type $S_{[0,0,x,y]}E$ près), et $S_{[0,c]}E \otimes S_{[0,c]}E = S_{[0,c-1]}E \otimes S_{[0,c+1]}E \oplus S_{[0,0,c,c]}E$.
Je donne maintenant une dernière égalité faisant intervenir les groupes $H^{n,q+1}(X, S_{[1,0,c,c+1]}E)$ et $H^{n-1,q}(S_{[0,0,c,c]}E)$, qui permettra de conclure. Considérons $\mathcal{O}(e-c) \to G_2(E^*)$. La proposition 1 montre que toute partition λ, de taille au plus $(e-2) \times 2$ et $(e-c)$-admissible, vérifie soit

$\lambda_1, \lambda_2 \leq e-c-2$, soit $\lambda_1 = \lambda_2 + e - c - 1$. Dans le deuxième cas, on a $|\lambda| = \lambda_1 + \lambda_2 \leq [e-2-(e-c-1)] + [e-2] = e+c-3$. Par contre, la partition $(e-c-2, e-c-2)$ est de poids $2(e-c-2)$. Il est clair qu'à c fixé, si e est grand, le poids de cette partition est strictement supérieur à toute constante plus le poids d'une partition du deuxième type. Quitte à augmenter e, on peut donc faire comme si seules les partitions du premier type intervenaient. En posant $P = n + 2(e-c-2) - 1$, on montre alors que

$$H^{n,q+1}(X, S_{[1,0,c,c+1]}E) = H^{n-1,q}(X, S_{[0,0,c,c]}E) \text{ si } q > 2c+1.$$

Les trois égalités montrent que si $q > 3c$, alors $H^{n,q+1}(X, S_{[1,c+1]}E \otimes S_{[0,c]}E) = 0$. La première implique alors le reste du lemme. •

Lemme 13 $\quad$ $H^{n,q}(X, S_{[a,0]}E \otimes S_{[b+1,1]}E \otimes L^t) = 0$ *si* $q > a+2$, *et* $H^{n-1,q}(X, S_{[a,0]} \otimes S_{[b,0]}E \otimes L^t) = 0$ *si* $q > a+1$.

Remarque : Ce lemme me donne l'occasion de justifier mes notations. Les partitions les plus faciles à traiter sont celles qui correspondent à une puissance du déterminant : en effet, puisque $(\det E)^l$ est ample pour tout l, le théorème de Kodaira donne directement que $H^{p,q}(X, (\det E)^l) = 0$ si $p+q > n$. Lorsqu'on s'écarte de cette partition, je vais expliquer qu'il existe une symétrie entre le fait d'allonger les premières parts et le fait de creuser le bas de la partition. Autrement dit, dans la notation $\lambda = [a,b,c,d]$, a et c, et b et d, jouent des rôles symétriques. On peut d'ailleurs remarquer que le théorème 2 n'échappe pas à ce principe : si l'on note $r = \min\{\delta(n-p), \delta(n-q)\}$, alors la borne vaut $q_0 = n - p + (r+\sigma)(ae - k + 2\sigma) - \sigma(\sigma+1)$. Or, avec mes notations, on a $Z^{k_i - \alpha_i - 1, k_i} = [\alpha_i, e - k_i + \alpha_i]$. Notons $\beta_i = e - k_i + \alpha_i$; on a $ae - k + \sigma = \sum_i (e - k_i + \alpha_i) = \sum_i \beta_i =: \tau$. Ainsi, on a donc

$$q_0 = n - p + (r+\sigma)(\tau+\sigma) - \sigma(\sigma+1) = n - p + r(\sigma+\tau) + \sigma(\tau-1). \quad (3)$$

On voit bien que cette formule est symétrique en σ et τ (-1). Le lecteur intrigué pourra s'amuser de constater que la démonstration du théorème 2 fonctionne de façon tout à fait similaire si l'on fait jouer par τ le rôle joué par σ, et donne le même résultat. Il est vraisemblable que le démonstration de théorèmes d'annulation efficaces pour des partitions de rang strictement supérieur à 1 utilise une récurrence qui fasse intervenir de façon combinée σ et τ.
Le lemme 13, pour cette symétrie, est le symétrique du lemme 11 ; il se démontre de façon tout à fait similaire :
Demonstration : On se ramène comme précédemment au cas $e+a$ pair, puis on considère le fibré en droites $\mathcal{O}(e+a, e+b)$ sur $\mathbb{P}E^* \times_X \mathbb{P}E^*$, et le fibré en droites $\mathcal{O}(e+b, \frac{e+a}{2})$ sur $\mathbb{P}E^* \times_X G_2(E^*)$. On utilise aussi, comme pour démontrer le lemme 11, la règle de Littlewood-Richardson. En tenant compte de la remarque 4, on trouve ainsi $H^{n-1,q}(X, S_{[a,0]} \otimes S_{[a+1,1]}E) = 0$ si $q > a+1$. Dans la suite spectrale, il y a une flèche entre $H^{n,q+1}(X, S_{[a,0]}E \otimes S_{[a+1,1]}E)$ et $H^{n-1,q}(X, S_{[a,0]} \otimes S_{[a+1,1]}E)$; $H^{n,q}(X, S_{[a,0]}E \otimes S_{[a+1,1]}E)$ s'annule donc si $q > a+2$.

•

Par ailleurs, on peut, comme le fait le lemme 12, raffiner un peu le lemme précédent pour traiter certains cas particuliers :

Lemme 14 *Supposons* $a > 0$.

$$\begin{array}{lll} H^{n,q}(X, S_{[a+1,a,0,1]}E \otimes L^t) & = 0 & \textit{si } q > a+1, \textit{ et} \\ H^{n-1,q}(X, S_{[a,a,0,0]}E \otimes L^t) & = 0 & \textit{si } q > a. \end{array}$$

Demonstration : Similaire à celle du lemme 12. •

Lemme 15 $H^{n,q}(X, S_{[1,c+1]}E \otimes S_{[1,c+1]}E \otimes L^t) = H^{n,q}(X, S_{[0,c+2]}E \otimes S_{[0,c]}E \otimes L^t) = 0$ *si* $q > 4c+4$*, et* $H^{n-1,q}(X, S_{[1,c+1]} \otimes S_{[0,c]}E \otimes L^t) = 0$ *si* $q > 4c+3$.

Demonstration : Comme précédemment, on peut supposer que $e - c$ est pair et $t = 0$; soit donc l tel que $e - c = 2l$. Considérons le fibré en droites $\mathcal{O}(2l, 2l)$ sur $\mathbb{P}E^* \times_X \mathbb{P}E^*$, et la suite spectrale correspondante pour $P = n + 2(2l-1) - 2$. Les termes ${}^P E_1^{i,j}$ ne peuvent être non nuls que si i vaut $n, n-1$, ou $n-2$. Or ${}^P E_1^{n-2,q-n+2} = H^{n-2,q}(X, S_{[0,c]}E \otimes S_{[0,c]}E)$, qui est nul si $q > 4c+2$ par le théorème 2. On en déduit donc que si $q > 4c+2$, alors

$$\begin{array}{ll} & H^{n,q+2}(X, S_{[1,c+1]}E \otimes S_{[1,c+1]}E) \oplus 2H^{n,q+2}(X, S_{[2,c+2]}E \otimes S_{[0,c]}E) \\ = & 2H^{n-1,q+1}(X, S_{[1,c+1]} \otimes S_{[0,c]}E). \end{array}$$

De même, en considérant $\mathcal{O}(2l, l) \to \mathbb{P}E^* \times_X G_2(E^*)$, on montre que $H^{n,q+2}(X, S_{[2,c+2]}E \otimes S_{[0,c]}E) = H^{n-1,q+1}(X, S_{[1,c+1]} \otimes S_{[0,c]}E)$ si $q > 4c+2$. Ce système donne $H^{n,q+2}(X, S_{[1,c+1]}E \otimes S_{[1,c+1]}E) = 0$. Mais comme

$$S_{[1,c+1]}E \otimes S_{[1,c+1]}E \supset S_{[2,c+2]}E \otimes S_{[0,c]}E,$$

on en déduit qu'aussi $H^{n,q+2}(X, S_{[2,c+2]}E \otimes S_{[0,c]}E) = 0$, et donc $H^{n-1,q+1}(X, S_{[1,c+1]} \otimes S_{[0,c]}E) = 0$. •

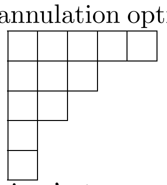
Le symétrique du lemme précédent est :

Lemme 16 $H^{n,q}(X, S_{[a+1,1]}E \otimes S_{[a+1,1]}E \otimes L^t) = H^{n,q}(X, S_{[a+2,2]}E \otimes S_{[a,0]}E \otimes L^t) = 0$ *si* $q > 2a+4$ *et* $H^{n-1,q}(X, S_{[a+1,1]} \otimes S_{[a,0]}E \otimes L^t) = 0$ *si* $q > 2a+3$.

Demonstration : Comme pour le lemme précédent, on pose $2l = a+e$ et $t = 0$, et on compare les suites spectrales correspondant à $\mathcal{O}(2l, 2l) \to \mathbb{P}E^* \times_X \mathbb{P}E^*$ et $\mathcal{O}(2l, l) \to \mathbb{P}E^* \times_X G_2(E^*)$. •

Jusqu'à maintenant, nous avons considéré des partitions de type $[a, \epsilon] \otimes [b, \epsilon']$ ($\epsilon, \epsilon' = 0$ ou 1), ou de type $[\epsilon, a] \otimes [\epsilon', b]$. Etudions finalement des partitions "panachées", c'est-à-dire de type $[a, \epsilon, b, \epsilon']$.

Lemme 17 *Si a et c ont même parité, alors* $H^{n,q}(X, S_{[a,0]}E \otimes S_{[1,c+1]}E \otimes L^t) = 0$ *si* $q > 2c+2$, *et* $H^{n,q}(X, S_{[a+1,1]}E \otimes S_{[0,c]}E \otimes L^t) = 0$ *si* $q > a+2$. *Si* $q > \max(a, 2c)+1$, *alors* $H^{n-1,q}(X, S_{[a,0]}E \otimes S_{[0,c]}E \otimes L^t) = 0$.

Remarque : Ce lemme illustre la difficulté du problème qui consiste à obtenir des théorèmes d'annulation optimaux : le lemme 16 avec $a = 1$ donne la borne 6 pour la partition  (si $e = 5$) ; cette borne convient pour démontrer le théorème 3 mais n'est pas optimale, puisque le lemme 17 donne la borne 4 ($a = 3, c = 1$). Par contre, j'aurai besoin de cette bonne borne pour montrer le lemme 19. Pour démontrer un analogue du théorème 3 en tous rangs, il faut probablement prendre en compte ce genre de subtilités.

Demonstration : On suppose que $a + e = 2l, e - c = 2m$ et $t = 0$. Si l'on considère $\mathcal{O}(2l, 2m) \to \mathbb{P}E^* \times_X \mathbb{P}E^*$, alors on obtient

$$H^{n-1,q}(X, S_{[a,0]}E \otimes S_{[0,c]}E) =$$
$$H^{n,q+1}(X, S_{[a,0]}E \otimes S_{[1,c+1]}E) \oplus H^{n,q+1}(X, S_{[a+1,1]}E \otimes S_{[0,c]}E) \text{ si } q > c+1.$$

Avec $\mathcal{O}(2l, m) \to \mathbb{P}E^* \times_X G_2(E^*)$, on a

$$H^{n-1,q}(X, S_{[a,0]}E \otimes S_{[0,c]}E) = H^{n,q+1}(X, S_{[a+1,1]}E \otimes S_{[0,c]}E) \text{ si } q > 2c+2.$$

Enfin, $\mathcal{O}(l, 2m) \to G_2(E^*) \times_X \mathbb{P}E^*$, donne

$$H^{n-1,q}(X, S_{[a,0]}E \otimes S_{[0,c]}E) = H^{n,q+1}(X, S_{[a,0]}E \otimes S_{[1,c+1]}E) \text{ si } q > a+2.$$

Le lemme découle de ces trois remarques.

•

Lemme 18 *Si a et c ont même parité, alors*

$$H^{n-2,q}(X, S_{[a,0]}E \otimes S_{[0,c]}E \otimes L^t) = 0 \text{ si } q > a+2c+2.$$

Demonstration : Si $t = 0, 2l = e + a$ et $2m = e - c$, on regarde la suite spectrale associée à $\mathcal{O}(l, m) \to G_2(E^*) \times_X G_2(E^*)$, pour $P = n + (2e - l - 3) + (3(m-1)) - 2$. Par la remarque 4, pour $q > a + 2c + 2$, celle-ci compare $H^{n-2,q}(X, S_{[a,0]}E \otimes S_{[0,c]}E)$ et $H^{n,q+2}(X, S_{[a+1,1]}E \otimes S_{[0,c]}E) \oplus H^{n,q+2}(X, S_{[a,0]}E \otimes S_{[1,c+1]}E)$. Or, ces deux derniers groupes sont nuls, grâce au lemme 17, si $q > \max\{a, 2c\}$.

•

Lemme 19 *Si a et c ont même parité, alors*

$$H^{n,q}(X, S_{[a+1,1]}E \otimes S_{[1,c+1]}E \otimes L^t) = 0 \text{ si } q > a+2c+4.$$

Demonstration : Si $t = 0$, $2l = e + a$ et $2m = e - c$, alors $\mathcal{O}(2l, 2m) \to \mathbb{P}E^* \times_X \mathbb{P}E^*$ donne, pour $q > a + 2c + 2$,

$$\begin{array}{ll} & H^{n,q+2}(X, S_{[a+1,1]}E \otimes S_{[1,c+1]}E) \oplus H^{n,q+2}(X, S_{[a+2,2]}E \otimes S_{[0,c]}E) \\ \oplus & H^{n,q+2}(X, S_{[0,a]}E \otimes S_{[2,c+2]}E) \\ = & H^{n-1,q+1}(X, S_{[1,a+1]}E \otimes S_{[0,c]}E) \oplus H^{n-1,q}(X, S_{[0,a]}E \otimes S_{[1,c+1]}E). \end{array}$$

En effet, $H^{n-2,q}(X, S_{[a,0]} \otimes S_{[0,c]}E) = 0$ si $q > a + 2c + 2$ par le lemme 18. De même, $\mathcal{O}(l, 2m) \to G_2(E^*) \times_X \mathbb{P}E^*$ et $\mathcal{O}(2l, m) \to \mathbb{P}E^* \times_X G_2(E^*)$ donnent respectivement, pour $q > 2a + c + 2$,

$$\begin{array}{lcl} H^{n,q+2}(X, S_{[a+2,2]}E \otimes S_{[0,c]}E) & = & H^{n-1,q+1}(X, S_{[1,a+1]}E \otimes S_{[0,c]}E) \text{ et} \\ H^{n,q+2}(X, S_{[0,a]}E \otimes S_{[2,c+2]}E) & = & H^{n-1,q}(X, S_{[0,a]}E \otimes S_{[1,c+1]}E). \end{array}$$

•

Remarquons enfin que l'on peut améliorer la borne pour $S_{[2,1]}E \otimes S_{[1,2]}E$. En effet, en généralisant la démonstration du lemme précédent, on voit que l'annulation $H^{n-2,q}(X, S_{[1,0]}E \otimes S_{[0,1]}E) = 0$ si $q > q_0$ implique $H^{n,q}(X, S_{[2,1]}E \otimes S_{[1,2]}E) = 0$ si $q > q_0 + 2$. Le lemme précédent utilisait la borne $q_0 = 5$ montrée dans le lemme 18 ; je vais montrer que $H^{n-2,q}(X, S_{[1,0]}E \otimes S_{[0,1]}E) = 0$ si $q > 4$ par un argument qui ne fonctionne que pour $S_{[1,0]}E \otimes S_{[0,1]}E$. La remarque spécifique est que, par la règle de Littlewood-Richardson,

$$S_{[1,0]}E \otimes S_{[0,1]}E = (\det E)^2 \oplus \det E \otimes S_{[1,1]}E.$$

Puisqu'on a $H^{n-2,q}(X, (\det E)^2) = 0$ si $q > 2$, il suffit de montrer que $H^{n-2,q}(X, \det E \otimes S_{[1,1]}E) = 0$ si $q > 4$. Pour cela, on peut supposer que le rang de E est multiple de 4, $e = 4f$, et on regarde la suite spectrale correspondant à $\mathcal{O}(f, 4f) \to G_4(E) \times_X \mathbb{P}E$. Dans cette suite spectrale, notre groupe, $H^{n-2,q}(X, \det E \otimes S_{[1,1]}E) = 0$, est connecté à $H^{n-3,q-1}(X, (\det E)^2)$ (qui s'annule si $q - 1 > 3$), $H^{n-1,q+1}(X, \det E \otimes S_{[2,2]}E)$ (nul pour $q + 1 > 5$: lemme 17), et enfin $H^{n,q+1}(X, \det E \otimes S_{[3,3]}E) = 0$ (nul pour $q + 1 > 3$ par le théorème A' de [Man 97]). On a donc démontré

Lemme 20 *$H^{n,q}(X, S_{[2,1]}E \otimes S_{[1,2]}E) = 0$ si $q > 6$.*

Ceci illustre à nouveau la complexité du sujet : on obtient de meilleures bornes pour des partitions spécifiques que celles obtenues par la méthode générale. Par ailleurs, comme nous utilisons une récurrence, le fait de ne pas obtenir le théorème d'annulation optimal pour une partition précise se répercute sur de nombreuses autres partitions, entrainant progressivement la "catastrophe".

Références

[FH 91] W. Fulton, J. Harris, Representation theory. A first course. Graduate Texts in Mathematics, 129. Readings in Mathematics, Springer-Verlag, New York, 1991.

[FL 81] W. Fulton et R. Lazarsfeld, Connectivity and its applications in algebraic geometry. Algebraic geometry (Chicago, Ill., 1980), pp. 26–92, Lecture Notes in Math., 862, Springer, Berlin-New York, 1981.

[JPW 81] T. Józefiak, P. Pragacz, J. Weyman, *Resolutions of determinantal varieties and tensor complexes associated with symmetric and antisymmetric matrices.* Young tableaux and Schur functors in algebra and geometry (Toruń, 1980), Astérisque, pp. 109–189, 87-88, Soc. Math. France, Paris, 1981.

[Lay 96] F. Laytimi, *On degeneracy loci.* Internat. J. Math. 7 (1996), no. 6, 745–754

[LN 02] F. Laytimi, W. Nahm, *A generalization of Le Potier's vanishing theorem.* math.AG/0210010.

[LP 77] J. Le Potier, *Cohomologie de la grassmannienne à valeurs dans les puissances extérieures et symétriques du fibré universel.* Math. Ann. 226 (1977), no. 3, 257–270.

[Man 97] L. Manivel, *Vanishing theorems for ample vector bundles.* Invent. Math. 127 (1997), no. 2, 401–416.

[Man 92] L. Manivel, *Théorèmes d'annulation pour la cohomologie des fibrés vectoriels amples*, thèse.

[Nie 81] H.A. Nielsen, *Free resolutions of tensor forms.* Young tableaux and Schur functors in algebra and geometry 289–302, Astérisque, 87-88.

[PS 73] C. Peskine ; L. Szpiro, *Dimension projective finie et cohomologie locale. Applications à la démonstration de conjectures de M. Auslander, H. Bass et A. Grothendieck.* Inst. Hautes Études Sci. Publ. Math. 42, (1973), 47–119.

[Sno 86] D.M. Snow, *Cohomology of twisted holomorphic forms on Grassmann manifolds and quadric hypersurfaces.* Math. Ann. 276 (1986), no. 1, 159–176.

[Tu 89] L.W. Tu, *The connectedness of symmetric and skew-symmetric degeneracy loci : even ranks.* Trans. Amer. Math. Soc. 313 (1989), 381–392.

[Tu 88] L.W. Tu, J. Harris, *The connectedness of symmetric degeneracy loci : odd ranks.* Banach Center of publications, 26 : Topics in algebra, Part 2 (1988), 249–256.

Pierre-Emmanuel Chaput
Université Grenoble 1
Institut Fourier
UFR de Mathématiques
UMR 5582
BP 74
38402 Saint-Martin-d'Heres CEDEX
France

Stabilité en Niveau 0, pour les Groupes Orthogonaux Impairs p-Adiques

Colette Moeglin

Received: December 16, 2003
Revised: December 3, 2004

Communicated by Peter Schneider

Abstract. The general problem we discuss in this paper is how to prove stability properties for a linear combination of characters of irreductible discrete series of p-adic groups. Here we give ideas on how to reduce the case where the Langlands parameter is trivial on the wild ramification group to the case where this Langlands parameter factorizes through the Frobenius; we handle only the case of an odd orthogonal group. The principal result is that the localization commutes with the Lusztig's induction.

2000 Mathematics Subject Classification: 22E50
Keywords and Phrases: representations of p-adic groups, stability, discrete series, Langands parameter

Précisons tout de suite que dans ce qui suit, F est un corps extension finie de $\mathbb{Q}_p$ avec $p \neq 2$ et même pour le théorème principal p grand. Le but de ce travail est de produire des fonctions sur les groupes p-adiques orthogonaux impairs dont les intégrales orbitales sur les éléments elliptiques réguliers ne dépendent que des classes de conjugaison stable. Au passage, on produit aussi des fonctions dont la somme des intégrales orbitales à l'intérieur d'une classe stable fixée est nulle. A la fin du papier, on interprète ce résultat en terme de stabilité des représentations elliptiques de niveau 0 pour ces groupes orthogonaux. Si l'on a bien prédit les signes qui dépendent encore de la traduction en terme d'algèbre de Hecke de l'induction de Lusztig (ceci devrait être l'objet de [2]), c'est la somme des représentations dans un paquet qui est stable et ces combinaisons linéaires engendrent l'espace distributions stables combinaison linéaires de représentations elliptiques.

Dans le détail, on commence par rappeler ce qu'est une représentation de niveau zéro et comment on peut lui associer un pseudo-coefficient; ce n'est pas nouveau et n'est pas au cœur du papier. C'est quand on passe à la description des paramètres de ces représentations que l'on entre dans le vif du sujet. On décrit ces paramètres en terme d'ensemble d'orbites unipotentes de groupes complexes convenables et de systèmes locaux sur ces orbites; la façon classique de faire cela est de considérer le morphisme de Langlands-Lusztig, ψ, de $W_F \times SL(2,\mathbb{C})$ dans $Sp(2n,\mathbb{C})$ (le groupe dual) et de décomposer d'abord la restriction de ψ à W_F. Ici on décompose la restriction de ψ au sous-groupe de W_F noyau de l'application $W_F \to < Fr >$ (où Fr est un Frobénius) et c'est dans le commutant de l'image par ψ de ce sous-groupe que vivent les orbites unipotentes et systèmes locaux ci-dessus. Pour faire cette classification, on utilise le fait que ψ est de niveau 0 (cf la définition donnée dans le texte) mais ceci n'est pas une hypothèse importante à cet endroit. On peut alors utiliser les méthodes de Lusztig et la représentation de Springer généralisée pour associer aux systèmes locaux trouvés, des fonctions sur des groupes finis, qui sont en fait des parahoriques en réduction du groupe orthogonal de départ; c'est là que l'hypothèse de niveau 0 intervient. Ce sont les faisceaux caractères de Lusztig. On remonte ces fonctions sur le groupe parahorique en les rendant invariantes par le radical pro-p-unipotent et on les prolonge par zéro pour en faire des fonctions sur le groupe orthogonal. Ce sont ces fonctions pour lesquelles on peut calculer le comportement des intégrales sur les classes de conjugaison d'éléments elliptiques à l'intérieur d'une classe de conjugaison stable. Avec cette méthode, ce ne sont pas des combinaisons linéraires de paramètres de Langlands qui donnent des objets stables mais directement certains paramètres; c'était déjà le cas en [8] et ici c'est expliqué en 5.1. On revient à des combinaisons linéraires plus habituelles en faisant une opération style transformation de Fourier, cf. 6.1.

Pour finir, on veut interpréter ces fonctions comme un ensemble de pseudo coefficients pour les représentations elliptiques (elliptiques au sens d'Arthur) de niveau zéro (cf. 6.2); vu ce qui est rappelé au début de ce papier, pour le faire on doit calculer ce que l'on appelle la restriction aux parahoriques des représentations, c'est-à-dire calculer l'action de chaque parahorique dans la sous-représentation formée par les vecteurs invariants sous l'action du radical pro-p-unipotent du dit parahorique. Pour cela, on a besoin de 2 résultats. Le premier est un résultat non disponible (cf. [2]) qui ramène l'étude de l'algèbre de Hecke des représentations induites de cuspidales de niveau 0 pour un groupe réductif à des algèbres de Hecke de représentations induites à partir de cuspidales de réduction unipotente pour des groupes réductifs convenables; ici ce résultat fera intervenir d'autres groupes orthogonaux et des groupes unitaires et il y aura sans doute un signe et même si on voit les idées le calcul est loin d'être acquis.

Il faut ensuite savoir calculer la restriction aux parahoriques des représentations de réduction unipotente pour ces groupes orthogonaux et unitaires. Pour les groupes orthogonaux, c'est essentiellement fait en [12] (nous l'avons déjà utilisé dans [8]) et pour les groupes unitaires c'est fait dans [7]. Moyennant ces restrictions, on a donc une bonne description de pseudo coefficient pour les représentations elliptiques de niveau 0 des groupes orthogonaux considérés ici. Et en utilisant les résultats d'Arthur ([1]) qui ramènent les problèmes de stabilité pour des représentations elliptiques à la stabilité des intégrales orbitales en les éléments elliptiques de leurs pseudo-coefficients, on en déduit une description des paquets stables de représentations elliptiques de niveau zéro. Tout ceci est réminiscent de [8].

Pour finir cette introduction, je remercie Anne-Marie Aubert pour les conversations que nous avons eues et le texte qu'elle a écrit pour moi, ainsi que Jean-Loup Waldspurger qui m'a fait un nombre certain de calculs.

1 Représentations de niveau zéro.

En suivant les définitions usuelles et en particulier celles de [8], on appelle réseau presque autodual un réseau L de V tel que

$$\omega_F \tilde{L} \subset L \subset \tilde{L}.$$

Pour L un réseau presque autodual comme ci-dessus, on note $K(L)$ le stabilisateur du réseau et $U(L)$ le radical pro-p-unipotent. Plus généralement, on appelle chaîne de réseaux presque autoduaux une famille $L_. = (L_0, L_1, \cdots, L_r)$ de réseaux de V telle que:

$$\omega_F \tilde{L}_r \subset L_r \subset \cdots \subset L_1 \subset L_0 \subset \tilde{L}_0.$$

Et on généralise de façon évidente la définition de $K(L_.)$ et $U(L_.)$. Une description totalement explicite de ces objets a été donnée en [8] 1.2. On y a en particulier défini la notion d'association.

Soit π une représentation irréductible de $G(F)$; on dit que π est de niveau 0 s'il existe un réseau presque autodual L tel que l'espace des invariants sous $U(L)$, $\pi^{U(L)}$ est non nul; c'est une définition standard reprise en particulier de [3]. On note alors π_L la représentation de $K(L)/U(L)$ dans ces invariants; c'est une représentation (non irréductible en général) d'un groupe réductif sur le corps $\mathbb{F}_q$.

La notion de représentation de niveau zéro pour les groupes GL est de même ordre et nous l'utiliserons.

La construction de pseudo-coefficients pour les séries discrètes (ou plus généralement les représentations elliptiques au sens d'Arthur) faite en [8] 1.9 (iv) s'étend au cadre des représentations elliptiques de niveau 0. C'est ce que nous allons expliquer dans cette partie.

1.1 Support cuspidal

Soit π une représentation irréductible de $G(F)$; on appelle support cuspidal de π la donnée d'un sous-groupe de Levi M de G, qui est le Levi d'un parabolique défini sur F et une représentation cuspidale π_{cusp} de $M(F)$ tel que π soit

quotient de l'induite de π_{cusp} (grâce à un parabolique de Levi M). La donnée de (M, π_{cusp}) est alors définie à conjugaison près par le groupe de Weyl de G. A partir de maintenant, on suppose que π est de niveau zéro. Il résulte de [9] 6.11 que π_{cusp} l'est aussi (le groupe G étant remplacé par M). On vérifie alors que l'on peut construire une chaîne de réseaux presque autoduaux, $L_.$ de V, en bonne position par rapport à M, telle que π_{cusp} ait des invariants (non nuls) sous $U(L_.) \cap M(F) =: U_M(L_.)$; on note $\pi_{cusp}^{U_M(L_.)}$ cet espace d'invariants. On note $K_M(L_.) := K(L_.) \cap M(F)$ et $\pi_{cusp}^{U_M(L_.)}$ est naturellement une représentation du groupe fini $K_M(L_.)/U_M(L_.)$; elle est cuspidale. On note χ_{cusp} la donnée de cette représentation et du groupe fini qui opère; cela sous-entend que la donnée d'une chaîne de réseaux presque autoduaux, $L_{.,cusp}$ ait été faite. Un tel choix n'est pas unique mais il l'est à association près.
Il résulte de [9] que pour L un réseau presque autodual de V, la représentation π_L n'est pas nulle si et seulement si $K(L)$ est associé à un sous-groupe parahorique contenant $K(L_{.,cusp})$ et le support cuspidal de π_L comme représentation du groupe fini $K(L)/U(L)$ est conjuguée de χ_{cusp}. On note $C(K(L))_{\chi_{cusp}}$ l'ensemble des fonctions sur le groupe fini $K(L)/U(L)$ engendré par les caractères des représentations irréductible ayant un conjugué de χ_{cusp} comme support cuspidal. Et on note $C(K(L))_{cusp,\chi_{cusp}}$ la projection de l'espace $C(K(L))_{\chi_{cusp}}$ sur l'ensemble des fonctions cuspidales.
Lusztig associe aux représentations des groupes finis et donc à χ_{cusp} un élément semi-simple s_χ dans un certain groupe sur $\overline{\mathbb{F}}_q$, le groupe $Sp(2n', \overline{\mathbb{F}}_q) \times O(2n'', \overline{\mathbb{F}}_q)$. La classe de conjugaison de s_χ dans $GL(2n, \overline{\mathbb{F}}_q)$ est elle bien définie. On dira que s_χ, ou plutôt sa classe de conjugaison, est la classe de conjugaison semi-simple associée à χ_{cusp}.

1.2 Représentation elliptique

Comme en [8] 1.7, on utilise la notion de représentation elliptique telle que précisée par Arthur; c'est, une représentation elliptique est une combinaison linéaire de représentations tempérées. Si l'on fixe un Levi M de G, Levi d'un sous-groupe parabolique de G défini sur F et une représentation cuspidale π_{cusp} de $M(F)$, on dit que la représentation elliptique, π, est de support cuspidal (M, π_{cusp}) si toutes les représentations irréductibles qui interviennent dans sa combinaison linéaire ont cette propriété; et on dit qu'elle est de niveau 0 si le support cuspidal est de niveau 0.

1.3 Pseudo-coefficients

On reprend les notations de 1.1, en particulier χ_{cusp} et $C(K(L))_{cusp,\chi_{cusp}}$. En suivant [8], on remonte tout élément f de $C(K(L))_{cusp,\chi_{cusp}}$, en une fonction sur $K(L)$ invariante par $U(L)$ et on la prolonge en une fonction notée f^G sur $G(F)$ en l'étendant par 0 hors de $K(L)$. Par cette procédure, on obtient une fonction cuspidale sur $G(F)$, c'est-à-dire une fonction dont les intégrales orbitales sur les éléments semi-simples non elliptiques sont nulles. Quand on somme cette

construction sur tous les supports cuspidaux de niveau 0, on construit ainsi un morphisme de $\oplus_{L/\sim} C(K(L))_{cusp}$ dans l'ensemble des fonctions cuspidale sur $G(F)$.
On note $Ps^G_{ell,\chi_{cusp}} := \oplus_{L/\sim} C(K(L))_{cusp,\chi_{cusp}}$ que l'on munit d'un produit scalaire. Pour définir ce produit scalaire, il faut fixer un ensemble de représentants des groupes $K(L)$ modulo association (les sommes sur K ci-dessous signifient la somme sur un ensemble de représentants)

$$(\sum_K \phi_K, \sum_K \phi'_K) = \sum_K w(K)^{-1}(\phi_K, \phi'_K)_K,$$

où le dernier produit scalaire est le produit scalaire usuel sur un groupe fini et où $w(K)$ est un volume décrit en [8] 1.6. Quand on somme cette construction sur l'ensemble des supports cuspidaux de niveau 0, on définit $Ps^G_{ell,0}$ muni d'un produit scalaire. La décomposition suivant les supports cuspidaux (modulo conjugaison) est une somme directe.
On rappelle la construction des pseudo coefficients pour les représentations elliptiques de niveau 0 faite (essentiellement) en [8] 1.9. et qui repose sur les travaux de [10] et de [4]
Pour tout K comme ci-dessus, définissons $B(K)$ (resp. $B(K)_{\chi_{cusp}}$) une base de $C(K)_{cusp}$ (resp. $C(K)_{cusp,\chi_{cusp}}$) et pour toute représentation virtuelle tempérée, D, posons

$$\phi_D := \oplus_K \sum_{f\in B(K)} w(K)\overline{D(f^G)}f;$$

$$\phi_{D,\chi_{cusp}} := \oplus_K \sum_{f\in B(K)_{\chi_{cusp}}} w(K)\overline{D(f^G)}f.$$

Théorème. *L'application $D \mapsto \phi_D$ induit un isomorphisme de l'espace engendré par les caractères des représentations elliptiques de niveau 0 sur $Ps^G_{ell,0}$. Cet isomorphisme est compatible à la décomposition suivant le support cuspidal.*
On peut récrire ϕ_D sous la forme la plus utilisable. On pose:

$$\varphi_D := \sum_K w(K)tr_K(D),$$

où $tr_K(D)$ est la trace pour la représentation de K/U_K dans l'espace des invariants de la représentation D sous U_K (le radical pro p-unipotent de K). On note $proj_{ell}\varphi_D$ la projection de φ_D sur les fonctions cuspidales (cette projection se fait pour chaque parahorique K individuellement). On sait alors définir $(proj_{ell}\varphi_D)^G$ qui est un pseudo coefficient de D si D est elliptique (cf [8] 1.9 (iv)). Si D a χ_{cusp} (cf. ci-dessus) pour support cuspidal, alors $proj_{ell}\varphi_D \in \oplus_K C[K]_{cusp}$
Remarque. *Avec les notations ci-dessus, $(proj_{ell}\varphi_D)^G$ est un pseudo-coefficient de la représentation elliptique D.*

On rappelle aussi que d'après Arthur [1] 6.1, 6.2 (on a enlevé l'hypothèse relative au lemme fondamental en [8] 4.6) une combinaison linéaire D de représentations elliptiques est stable si et seulement si les intégrales orbitales de φ_D^G sont constantes sur les classes de conjugaison stable d'éléments elliptiques réguliers.

2 Classification des paramètres discrets de niveau 0

On considère les couples (ψ, ϵ) de morphismes continus suivants:

$$\psi : W_F \times SL(2,\mathbb{C}) \to Sp(2n,\mathbb{C})$$
$$\epsilon : Cent_{Sp(2n,\mathbb{C})}(\psi) \to \{\pm 1\},$$

où, en notant I_F le sous-groupe de ramification de W_F, la restriction de ψ à I_F est triviale sur le groupe de ramification sauvage et où le centralisateur de ψ dans $Sp(2n,\mathbb{C})$ n'est pas inclus dans un sous-groupe de Levi de $Sp(2n,\mathbb{C})$. De tels couples sont appelés des paramètres discrets de niveau 0; comme la condition ne porte que sur ψ et non sur ϵ, on peut dire aussi que ψ est discret de niveau 0 sans référence à ϵ.

2.1 Morphismes de paramétrisation

Pour donner la classification des morphismes comme ci-dessus, il est plus simple d'avoir fixé un générateur du groupe abélien I_F/P_F, où P_F est le groupe de ramification sauvage. Et pour cela, il est plus simple de fixer une extension galoisienne modérément ramifiée E/F et de ne considérer que les morphismes ψ qui se factorisent par le groupe de Weil relatif $W_{E/F}$. On fixe Fr une image réciproque d'un Frobénius de l'extension non ramifiée dans $W_{E/F}$ et s_E un générateur du groupe multiplicatif du corps résiduel de E. Dans ce cas la restriction de ψ à I_F est déterminée par l'image de s_E. A conjugaison près c'est donc la donnée des valeurs propres de la matrice image de s_E par ψ qui détermine cette restriction. On va donc fixer cette restriction en la notant χ, c'est à dire fixer une matrice de $Sp(2n,\mathbb{C})$ dont les valeurs propres sont des racines de l'unité d'ordre premier à p. On peut donc oublier E et garder χ et considérer que χ est déterminé par une collection de racines de l'unité, l'ensemble des valeurs propres ensemble que l'on note $VP(\chi)$. Pour $u \in VP(\chi)$ on note $mult(u)$ la multiplicité de u en tant que valeur propre. L'action du Frobenius transforme χ en χ^q, ainsi si $u \in VP(\chi)$ alors $u^q \in VP(\chi)$ et $mult(u) = mult(u^q)$. Comme χ est à valeurs dans $Sp(2n,\mathbb{C})$, l'espace propre pour la valeur propre u est en dualité avec l'espace propre pour la valeur propre u^{-1}, d'où aussi $mult(u) = mult(u^{-1})$. A l'intérieur de $VP(\chi)$ on définit l'équivalence engendrée par la relation élémentaire $u \sim u^q$. On note $[VP(\chi)]$ les classes d'équivalence et si $u \in VP(\chi)$, on note $[u]$ sa classe d'équivalence. On vérifie que s'il existe $u \in VP(\chi)$ tel que $u^{-1} \notin [u]$ alors le centralisateur de ψ est inclus dans un sous-groupe de Levi de $Sp(2n,\mathbb{C})$; on suppose donc que $u^{-1} \in [u]$ pour tout u. Pour tout $u \in VP(\chi)$, $u \notin \{\pm 1\}$, on définit $\ell_{[u]}$ comme le plus petit entier tel que $u^{-1} = u^{q^{\ell_{[u]}}}$; le cardinal de la classe $[u]$ est alors $2\ell_{[u]}$. On pose: $m([u]) := mult(u)$ où $u \in VP(\chi)$ dans la classe de $[u]$ comme la notation le suggère. On remarque que

$$2n = m(1) + m(-1) + \sum_{[u]\in[VP(\chi)], u\notin\{\pm 1\}} m([u])2\ell_{[u]}.$$

Soit M un entier et U une orbite unipotente de $GL(M,\mathbb{C})$; on dit que U est symplectique (resp. orthogonale) si tous les blocs de Jordan sont pairs (resp. impairs) et on dit qu'elle est discrète si son nombre de blocs de Jordan d'une taille donnée est au plus 1. D'où la notation symplectique discrète et orthogonale discrète qui allie les 2 définitions.

Proposition. *L'ensemble des homomorphismes ψ ci-dessus (c'est-à-dire discrets et de niveau 0), pris à conjugaison près, dont la restriction à I_F est conjuguée de χ est en bijection avec l'ensemble des collections d'orbites unipotentes $\{U_{[u],\zeta}, [u] \in [VP(\chi)], \zeta \in \{\pm 1\}\}$, de groupe $GL(m([u],\zeta),\mathbb{C})$, ce qui définit l'entier $m([u],\zeta)$ (éventuellement 0) avec les propriétés suivantes:*

$$\forall [u] \in [VP(\chi)],\ m([u],+) + m([u],-) = m([u]);$$

pour tout $[u] \in [VP(\chi)], u \neq \pm 1$, l'orbite $U_{[u],+}$ est une orbite symplectique discrète, l'orbite $U_{[u],-}$ est une orbite orthogonale discrète et les orbites $U_{[\pm 1],\pm}$ sont des orbites symplectiques discrètes.

L'intérêt de ramener la classification à une collection d'orbites unipotentes est de pouvoir ensuite utiliser la représentation de Springer généralisée pour construire des représentations de groupes de Weyl, puisque l'on aura aussi des systèmes locaux sur ces orbites.

On a décrit avant l'énoncé comment on comprenait la restriction de ψ à I_F; pour avoir la restriction de ψ à W_F, il faut encore décrire l'image du relèvement du Frobénius, Fr, à conjugaison près. Par commodité et uniquement dans cette démonstration, on note V l'espace vectoriel $\mathbb{C}^{2n}$ et pour $u \in VP(\chi)$, on note $V[u]$ l'espace propre correspondant à cette valeur propre. Les conditions que doivent vérifier $\psi(Fr)$ sont: être une matrice symplectique et induire un isomorphise entre $V[u]$ et $V[u^q]$ pour tout $u \in VP(\chi)$.

Pour traduire ces conditions, fixons $u \in VP(\chi)$. Il faut distinguer les 2 cas:

premier cas: $u \neq \pm 1$. On remarque que $\psi(Fr^{q^{2\ell_u}})$ induit un isomorphisme de $V[u]$ dans lui-même. On note F_u cet homomorphisme. Le groupe $GL(V[u])$ s'identifie naturellement à un sous-groupe de $Sp(2n,\mathbb{C})$. Comme on ne cherche à classifier les morphismes ψ qu'à conjugaison près, on peut encore conjuguer sous l'action de $GL(V[u])$; cela se traduit sur F_u par la conjugaison habituelle. A conjugaison près F_u est donc déterminé par ses valeurs propres dont on note $VP(F_u)$ l'ensemble. On vérifie encore que si $VP(F_u)$ contient un élément autre que ± 1, alors l'image de ψ est incluse dans un sous-groupe de Levi propre de $Sp(2n,\mathbb{C})$. Pour $\zeta \in \{\pm 1\}$, on note $V[u,\zeta]$ l'espace propre pour la valeur propre ζ de F_u. On remarque pour la suite que $V[u,\zeta]$ est muni du produit scalaire:

$$\forall v, v' \in V[u,\zeta], < v, v' >_u := < v, \psi(Fr)^{q^{\ell_u}} v' > .$$

Et, pour v et v' comme ci-dessus:

$$\begin{aligned} < v, \psi(Fr)^{q^{\ell_u}} v' > &= < \psi(Fr)^{-q^{\ell_u}} v, v' > \\ &= \zeta < \psi(Fr)^{-q^{\ell_u}} F_u v, v' > \\ &= \zeta < \psi(Fr)^{q^{\ell_u}} v, v' > \qquad (1) \\ &= -\zeta < v', \psi(Fr)^{q^{\ell_u}} v > \\ &= -\zeta < v', v >_u . \end{aligned} \qquad (2)$$

Ainsi, la forme $<,>_u$ est symplectique pour $\zeta = 1$ et orthogonale pour $\zeta = -1$. Il est facile de vérifier que cette forme est non dégénérée. Ces constructions se font donc pour tout $u \in VP(\chi)$ différent de ± 1. De plus $\psi(Fr)$ induit une isométrie de $V[u, \zeta]$ sur $V[u^q, \zeta]$; ceci permet de définir intrinsèquement l'espace orthogonal ou symplectique $V([u], \zeta)$ pour tout $[u] \in [VP(\chi)]$ muni du produit scalaire $<,>_{[u]}$.
deuxième cas: $u \in \{\pm 1\}$. On définit ici F_u comme l'action de $\psi(Fr)$ comme automorphisme de $V[u]$. On vérifie comme ci-dessus que si F_u a des valeurs propres autres que ± 1, l'image de ψ se trouve dans un Levi de $Sp(2n, \mathbb{C})$; on définit donc encore $V[u, \zeta]$ pour $\zeta = \pm 1$ les valeurs propres de F_u. Mais ces espaces sont ici des espaces symplectiques par restriction de la forme symplectique.
Comme $\psi(SL(2, \mathbb{C}))$ commute à $\psi(W_F)$ les images des éléments unipotents de $SL(2, \mathbb{C})$ s'identifient à des éléments unipotents des automorphismes des espaces $V([u], \zeta)$ pour tout $[u] \in [VP(\chi)]$ et tout $\zeta \in \{\pm 1\}$. A conjugaison près, le morphisme ψ restreint à $SL(2, \mathbb{C})$ est même uniquement déterminé par l'orbite de ces éléments. Ce sont ces orbites qui sont notées $U_{[u],\zeta}$ dans l'énoncé. Comme les éléments de $\psi(SL(2, \mathbb{C}))$ commutent à $\psi(Fr)$ et respectent la forme symplectique, ils respectent chaque forme $<,>_u$. Ce sont donc des orbites unipotentes du groupe d'automorphismes de la forme. Il reste à remarquer que si l'une de ces orbites a 2 blocs de Jordan de même taille, alors l'image de ψ est incluse dans un Levi. Réciproquement la donnée des orbites permet de reconstruire (à conjugaison près) l'homomorphisme ψ.
REMARQUE. *Soit χ comme ci-dessus et identifions les racines de l'unité d'ordre premier à p de $\mathbb{C}$ avec leurs analogues dans $\overline{\mathbb{F}}_q$. Les éléments de $VP(\chi)$ avec leur multiplicité définissent donc un élément de $GL(2n, \overline{\mathbb{F}}_q)$ dont la classe de conjugaison est bien définie.*
Avec cette remarque, on peut associer à χ un élément semi-simple s_χ bien défini à conjugaison près dans $GL(2n, \overline{\mathbb{F}}_q)$. C'est l'analogue du s_χ de 1.1.

2.2 SYSTÈME LOCAL

On fixe ψ, ϵ commme dans l'introduction de cette section et on reprend les notations de la preuve précédente en notant $Jord(U_{[u],\zeta})$, où $[u] \in [VP(\chi)]$ et

$\zeta \in \{\pm 1\}$, l'ensemble des blocs de Jordan des orbites unipotentes associées à ψ.

Remarque. *Le centralisateur de ψ est isomorphe à*

$$\prod_{[u]\in VP(\chi)} \prod_{\alpha\in Jord(U_{[u],\zeta})} \{\pm 1\}.$$

L'image du centre de $Sp(2n,\mathbb{C})$ dans ce commutant est l'élément -1 diagonal. Ainsi ϵ s'identifie à une application de $\cup_{[u]\in[VP(\chi)];\zeta\in\{\pm 1\}} Jord(U_{[u],\zeta})$ dans $\{\pm 1\}$.

En reprenant la preuve précédente, on voit que le commutant de ψ s'identifie au commutant de $\psi(SL(2,\mathbb{C}))$ vu comme sous-ensemble de $\times_{[u]\in[VP(\chi)]}\times_{\zeta=\pm} Aut(V([u],\zeta),<\,,\,>_u)$. On sait calculer ce commutant. C'est alors un produit de groupes orthogonaux $\times_{[u],\zeta}\times_{\alpha\in Jord(U_{[u],\zeta})} O(mult_\alpha,\mathbb{C})$, où $mult_\alpha$ est la multiplicité de α comme bloc de Jordan de l'orbite en question; pour ψ discret cette multiplicité est 1. Pour avoir ce résultat la seule hypothèse utilisée est que $U_{[u],\zeta}$ est symplectique si $<\,,\,>_u$ est symplectique et orthogonale sinon. On aura aussi à regarder le cas elliptique où cette hypothèse sur le type de $U_{[u],\zeta}$ est satisfaite mais pas la multiplicité 1; on utilisera alors cette description. Dans le cas de la multiplicité 1, le groupe orthogonal se réduit à $\{\pm 1\}$; d'où l'énoncé, l'identification du centre étant immédiate.

Remarquons encore que quelle que soit la multiplicité, on peut voir le ϵ comme une application de $\times_{[u],\zeta} Jord(U_{[u],\zeta})$ dans $\{\pm 1\}$.

3 Faisceaux caractères.

3.1 Construction de fonctions.

Soit $m \in \mathbb{N}$; on utilisera fréquemment la notation $D(m)$ pour l'ensemble des couples d'entiers (m',m'') tels que $m = m' + m''$. On fixe χ un morphisme comme en 2.1. On reprend les notations $[VP(\chi)]$ de 2.1. Pour tout $[u] \in [VP(\chi)]$ avec $[u] \neq \pm 1$, on a défini les entiers $m([u])$ (qui sont les multiplicités des valeurs propres). On pose $n([u]) = m([u])$ si $[u] \neq \pm 1$ et $n(1) = m(1)/2$, $n(-1) = m(-1)/2$. Pour $(n'_{[u]}, n''_{[u]}) \in D(n([u]))$, on note $\mathbb{C}[\hat{W}_{n'_{[u]},n''_{[u]}}] := \mathbb{C}[\hat{\mathfrak{S}}_{n'[u]}] \otimes \mathbb{C}[\hat{\mathfrak{S}}_{n''[u]}]$ et $\mathbb{C}[\hat{W}_{D([u])}]$ l'espace vectoriel $\oplus_{(n'_{[u]},n''_{[u]})\in D(n([u])}\mathbb{C}[\hat{W}_{n'_{[u]},n''_{[u]}}]$, où les chapeaux représentent les classes d'isomorphie de représentations du groupe chapeauté. Pour $u = \pm 1$, la situation est plus compliquée à cause de l'existence de faisceaux caractères cuspidaux. On garde la même notation (pour unifier) mais on remplace $\hat{\mathfrak{S}}_{n'[u]}$ et $\hat{\mathfrak{S}}_{n''[u]}$ par l'ensemble des symboles de rang $n'[u]$ respectivement $n''[u]$ de défaut impair respectivement pair; il est rappelé en [13] 2.2, 2.3 comment ces symboles paramétrisent aussi des représentations irréductibles de groupes; un symbole de défaut impair, $I =: 2h+1$, et de rang $n'([u])$ paramétrise une représentation du groupe de Weyl de type C et de rang $n'([u]) - h^2 - h$. Dans le cas du défaut pair, il faut admettre les défauts négatifs; dans la référence donnée tout est

expliqué avec précision, les difficultés venant de la non connexité des groupes orthogonaux pairs et du fait que pour un tel groupe il faut regarder simultanément la forme déployée et celle qui ne l'est pas. Grosso modo, un symbole de défaut pair, $2h''$, et de rang $n''([u])$ paramétrise une représentation d'un groupe de Weyl de type C de rang $n''([u]) - (h'')^2$.
Fixons maintenant un ensemble de paires $(n'[u], n''[u]) \in D(n([u])$. On pose:

$$n' := \sum_{[u]\in[VP(\chi)];[u]\neq[\pm 1]} n'[u]\ell_{[u]} \; + (n'[1] + n'[-1]),$$

$$n'' := \sum_{[u]\in[VP(\chi)];[u]\neq[\pm 1]} n''[u]\ell_{[u]} \; + (n''[1] + n''[-1]).$$

On pose $\sharp = iso$ si G est déployé et $\sharp = an$ sinon. On note alors $K_{n',n''}$ un sous-groupe parahorique (non connexe) de G dont le groupe en réduction, $\overline{K}_{n',n''}$ est isomorphe à $SO(2n'+1, \mathbb{F}_q) \times O(2n'', \mathbb{F}_q)_\sharp$ (cf. 1.1). Il est bien défini à association près. On note M un sous-groupe de $\overline{K}_{n',n''}$ isomorphe à

$$\times_{[u]\neq[\pm 1]} U(n'[u], \mathbb{F}_{q^{2\ell_{[u]}}}/\mathbb{F}_{q^{\ell_u}}) \times SO(2(n'[1] + n'[-1]) + 1, \mathbb{F}_q)$$
$$\times_{[u]\neq[\pm 1]} U(n''[u], \mathbb{F}_{q^{2\ell_{[u]}}}/\mathbb{F}_{q^{\ell_u}}) \times O(2(n''[1] + n''[-1]), \mathbb{F}_q)_\sharp;$$

ci-dessus, on n'a pas précisé le plongement car cela n'a pas d'importance, sur les corps finis il n'y a qu'une classe de formes unitaires. On note $\underline{n'}$ et $\underline{n''}$ les collections $(n'[u])$ et $(n''[u])$ comme ci-dessus. Grâce à Lusztig (étendu au cas non connexe cf. [13] 3.1 et 3.2), on sait associer à un élément de $\mathbb{C}[\hat{W}_{\underline{n'},\underline{n''}}]$ et à χ une fonction sur M, la trace du faisceau caractère associé. Puis on définit cette fonction sur $\overline{K}_{n',n''}$ (par induction); c'est une fonction invariante par conjugaison.
En sommant sur toutes les décompositions $D(\chi)$, on construit ainsi une application de $\mathbb{C}[\hat{W}_{D(\chi)}]$ dans l'ensemble des fonctions $\oplus_{n',n''\in D(n)}\mathbb{C}[\overline{K}_{n',n''}]$. On remonte ensuite de telles fonctions en des fonctions sur $K_{n',n''}$ par invariance et on les prolonge à $SO(2n+1, F)_\sharp$ par 0. On note $k_{\sharp,\chi}$ cette application. Quand on fait une somme directe de $\sharp = iso$ avec $\sharp = an$, on la note k_χ.

3.2 Support cuspidal des faisceaux caractères

Dans cette section, on fixe quelques notations relatives aux faisceaux caractères quadratiques unipotents; elles viennent essentiellement (à des modifications formelles près) de [13] 3.1 et 3.7 lui-même fortement inspiré de Lusztig. Les difficultés viennent de la présence de faisceaux caractères cuspidaux; c'est le cas des groupes orthogonaux impairs et pairs qui a été sommairement expédié ci-dessus qu'il faut préciser.
Pour les groupes orthogonaux impairs, $SO(2m'+1, \mathbb{F}_q)$ on forme les faisceaux caractères quadratiques unipotents avec la donnée d'un couple ordonné de 2 symboles de défaut impair dont la somme des rangs est m'. Notons Λ'_+, Λ'_- ces 2 symboles et I_+, I_- leurs défauts. On écrit encore $I_\pm =: 2h'_\pm + 1$ en utilisant

le fait que les défauts sont impairs. On retrouve alors à peu près les notations de [13] 3.7. On considère le couple d'entier $(h'_+ + h'_- + 1, |h'_+ - h'_-|)$ et le signe $\sigma' := +$ si $h'_+ \geq h'_-$ et $-$ sinon. Dans ce couple d'entiers, l'un des nombres est pair et l'autre est impair; on note r'_p celui qui est pair et r'_{im} celui qui est impair. On note $n'_\pm$ le rang de $\Lambda'_\pm$ et on pose $N'_\pm := n'_\pm - h'_\pm(h'_\pm + 1)$. Ainsi $\Lambda'_\pm$ paramétrise une représentation du groupe de Weyl de type C de rang $N'_\pm$. Tandis que le couple $r'_{im}, \sigma' r'_p$ détermine un faisceau cuspidal pour le groupe $SO(r'^2_{im} + r'^2_p, \mathbb{F}_q)$ et l'on a: $2N'_+ + 2N'_- + r'^2_{im} + r'^2_p = 2m' + 1$.

Pour les groupes orthogonaux pairs, $O(2m'', \mathbb{F}_q)_\sharp$, on forme un faisceau caractère quadratique unipotent à l'aide d'un couple ordonné de 2 symboles eux-mêmes ordonnés au sens qu'un symbole est formé de 2 ensembles de nombres (avec des propriétés). Au sens habituel, l'ordre des ensembles n'a pas d'importance et le défaut est la différence entre le cardinal de l'ensemble ayant le plus d'élément (au sens large) et celui de l'ensemble ayant le moins d'éléments (au sens large). Ici, les 2 ensembles sont ordonnés et le défaut est la différence entre le cardinal du premier ensemble et celui du deuxième, ainsi le défaut peut-être négatif. On demande uniquement que les défauts soient pairs (0 est un nombre pair). On note Λ''_+, Λ''_- le couple des 2 symboles et P_+, P_- la valeur absolue de leur défaut et ζ_+, ζ_- les signes des défauts; on fera une convention sur le signe quand le défaut est 0 ci-dessous, pour le moment on n'en a pas besoin. Ainsi $P_\pm$ sont des nombres positifs ou nuls pairs. On pose encore $r''_\pm := (\zeta_+ P_+ \pm \zeta_- P_-)/2$; on a ainsi 2 éléments de $\mathbb{Z}$ de même parité. On note $n''_\pm$ le rang de $\Lambda''_\pm$ et $N''_\pm := n''_\pm - (h''_\pm)^2$ (où $h''_\pm = 1/2 P_\pm$). Ainsi $\Lambda''_\pm$ paramétrise une représentation du groupe de Weyl de type C de rang $N''_\pm$. Tandis que le couple r''_+, r''_- détermine un faisceau cuspidal pour le groupe $O(r''^2_+ + r''^2_-, \mathbb{F}_q)$ (cf. [13] 3.1) et l'on a: $2N'_+ + 2N'_- + r''^2_+ + r''^2_- = 2m''$.

On aura à considérer simultanément 2 couples ordonnés formé chacun de 2 symboles $(\Lambda'_\epsilon, \Lambda''_\epsilon); \epsilon \in \{\pm\}$ où, pour $\epsilon = +$ ou $-$, Λ'_ϵ est de défaut impair, I_ϵ et Λ''_ϵ est de défaut pair $\zeta_\epsilon P_\epsilon$ avec $P_\epsilon \in \mathbb{N}$ et $\zeta_\epsilon \in \{\pm\}$ avec ici la convention que si $P_\epsilon = 0$ alors $\zeta_\epsilon = (-1)^{(I_\epsilon - 1)/2}$.

Pour $\epsilon = +1$ ou -1, on pose précisément $\hat{W}_{D(n[\epsilon])}$ l'ensemble des couples de symboles $\Lambda'_\epsilon, \Lambda''_\epsilon$ comme ci-dessus dont la somme des rangs vaut $n[\epsilon]$. Ainsi $\hat{W}_{D(n[+1])} \times \hat{W}_{D(n[-1])}$ est un ensemble en bijection avec l'ensemble des quadruplets de symboles ordonnés dont le premier et le troisième sont de défaut impair et les 2 autres de défaut pair avec des conditions sur la somme des rangs. On pourra donc interpréter cet ensemble en utilisant ce qui est ci-dessus comme un ensemble des couples de représentations quadratiques unipotentes des groupes $SO(2m' + 1, \mathbb{F}_q) \times O(2m'', \mathbb{F}_q)$ où $m' + m'' = n[+1] + n[-1]$. Avec cette interprétation et ce que l'on a vu ci-dessus, les défauts des symboles déterminent des faisceaux cuspidaux, c'est-à-dire, combinatoirement, des nombres entiers $\underline{r} := (r'_+, r''_+, r'_-, r''_-)$ et un signe σ' avec r'_+ positif et impair, r'_- positif ou nul et pair et r''_+, r''_- des entiers relatifs de même parité. On pose alors $|\underline{r}|$ le quadruplet $(r'_+, |r''_+|, r'_-, |r''_-|)$. C'est lui qui permet de construire des fonctions de Green utiles pour la localisation (cf. 4.1).

On note $\mathbb{C}[\hat{W}_{D(n[+1])}] \otimes \mathbb{C}[\hat{W}_{D(n[-1])}]$ l'espace vectoriel de base $\hat{W}_{D(n[+1])} \times \hat{W}_{D(n[-1])}$.

3.3 Représentation de Springer-Lusztig

On fixe (ψ, ϵ) un paramètre discret de niveau 0 et on note encore χ la restriction de ψ au groupe de ramification de W_F. A un tel paramètre, on a associé une collection d'orbites $U_{[u],\zeta}$ où $[u] \in [VP(\chi)]$ et $\zeta \in \{\pm 1\}$ et ϵ s'identifie à un caractère du groupe des composantes du centralisateur d'un élément de $U_{[u],\zeta}$; on voit donc ϵ comme un morphisme de $\cup_{[u],\zeta} Jord(U_{[u],\zeta})$ dans $\{\pm 1\}$ (cf. 2.2). Pour $[u] \in [VP(\chi)], [u] \neq \pm 1$, on pose $U_{[u]} := U_{[u],+} \cup U_{[u],-}$ ou plutôt l'orbite unipotente de $GL(m([u]), \mathbb{C})$ engendrée et on pose:

$$n'[u]_{\psi,\epsilon} := \sum_{\alpha \in Jord(U_{[u]}); \epsilon(\alpha)=+1} \alpha, \qquad n''[u]_{\psi,\epsilon} := \sum_{\alpha \in Jord(U_{[u]}); \epsilon(\alpha)=-1} \alpha.$$

On définit alors $U'_{[u]}$ comme l'orbite unipotente de $GL(n'([u])_{\psi,\epsilon}, \mathbb{C})$ ayant comme bloc de Jordan l'ensemble des α blocs de Jordan de $U_{[u]}$ pour lesquels $\epsilon(\alpha) = +$. On définit de même $U''_{[u]}$.
Pour $u = \pm 1$, on pose:

$$n'[u]_{\psi,\epsilon} = \sum_{\alpha \in Jord(U_{[u],+})} \alpha, \qquad n''[u]_{\psi,\epsilon} = \sum_{\alpha \in Jord(U_{[u],-})} \alpha.$$

Pour unifier les notations, on pose ici aussi $U'_{[u]} := U_{[u],+}$ et $U''_{[u]} := U_{[u],-}$. Cette collection de paires $(n'[u]_{\psi,\epsilon}, n''[u]_{\psi,\epsilon})$ est naturellement notée $\underline{n'}_{\psi,\epsilon}, \underline{n''}_{\psi,\epsilon}$ et on voit la représentation de Springer-Lusztig comme l'élément de $\mathbb{C}[\hat{W}_{\underline{n'}_{\psi,\epsilon}, \underline{n''}_{\psi,\epsilon}}]$ défini ainsi:
soit $[u] \in [VP(\chi)], [u] \neq \pm 1$; Springer a associé à l'orbite $U'_{[u]}$ une représentation de $\mathfrak{S}_{n'[u]_{\psi,\epsilon}}$, non irréductible en général, dans la cohomologie de la variété des Borel (on regarde toute la représentation pas seulement celle en degré maximal). Cela définit donc un élément de $\mathbb{C}[\hat{\mathfrak{S}}_{n'[u]_{\psi,\epsilon}}]$. On fait la même construction en remplaçant $U'_{[u]}$ par $U''_{[u]}$ et on obtient un élément de $\mathbb{C}[\hat{\mathfrak{S}}_{n''[u]_{\psi,\epsilon}}]$.
Soit maintenant $u = \pm 1$. Ce sont les constructions de Lusztig qui sont rappelées en [8] 5.5 (et [12] 5.1). Ici la situation est un peu plus compliquée puisque l'on a 4 orbites les $U_{u,\epsilon'}$, pour $u, \epsilon' \in \{\pm 1\}$ avec des systèmes locaux et non pas 2 comme dans [8]. A chacune de ces orbites, $U_{u,\epsilon'}$ avec son système local est associé par la correspondance de Springer généralisée, un entier noté $k_{u,\epsilon'}$ et une représentation non irréductible en général du groupe de Weyl de type C, $W_{N_{u,\epsilon'}}$, où l'on a posé $N_{u,\epsilon'} := 1/2(\sum_{\alpha \in Jord(U_{u,\epsilon'})} \alpha - k_{u,\epsilon'}(k_{u,\epsilon'} + 1))$.
On considère les 2 couples indexés par le choix d'un élément u dans $\{\pm 1\}$ $(k_{u,+} + k_{u,-} + 1, |k_{u,+} - k_{u,-}|)$ et les 2 signes ζ_u qui sont le signe de $k_{u,+} - k_{u,-}$ quand ce nombre est non nul; s'il est nul le signe est $(-1)^{k_{u,+}}$ par convention. Dans les couples l'un des nombres est impair et on le note I_u et l'autre est

pair et est noté P_u. En regardant le produit tensoriel de la représentation de $W_{N_{u,+}}$ avec celle de $W_{N_{u,-}}$, on obtient une représentation du produit qui se traduit en terme de couples de symboles dont le premier est de défaut I_u et le deuxième de défaut $\zeta_u P_u$. C'est donc ainsi que l'on construit un élément de $\mathbb{C}[\hat{\mathcal{W}}_{n'(u)_{\psi,\epsilon},n''(u)_{\psi,\epsilon}}]$.

3.4 Induction, restriction

Je ne connais pas d'autre justification aux constructions faites ci-dessous que le fait que le résultat énoncé en [8] 5.5 et démontré en [12] suggère la conjecture de 6.2.

Il s'agit de construire une application $\rho \circ \iota$ de $\mathbb{C}[\hat{W}_{D(\chi)}]$ dans lui-même. Cela provient d'un produit tensoriel d'applications $\rho_{[u]} \circ \iota_{[u]}$ de même nature pour tout $[u] \in [VP(\chi)]$. Ces applications sont définies en [8] 3.18 pour $[u] = [\pm 1]$ et [8] 3.1 et 3.2 dans le cas de $[u] \neq [\pm 1]$; on en rappelle la définition d'autant que l'on en donne une présentation un peu différente.

Considérons le cas où $[u] \neq \pm 1$; on note $W_{m[u]}$ le groupe de Weyl de type C et de rang $m([u])$. Pour $(n'[u], n''[u]) \in D(m[u])$, on définit de même $W_{n'[u]}, W_{n''[u]}$; il existe une application naturelle de $W_{n'[u]} \times W_{n''[u]}$ sur $\mathfrak{S}_{n'[u]} \times \mathfrak{S}_{n''[u]}$. On peut ainsi remonter des représentations de $\mathfrak{S}_{n'[u]} \times \mathfrak{S}_{n''[u]}$ en des représentations de $W_{n'[u]} \times W_{n''[u]}$; ensuite on tensorise la représentation obtenue par le caractère sgn_{CD} de $W_{n''[u]}$. Puis on induit pour trouver un élément de $\mathbb{C}[\hat{W}_{m[u]}]$. L'application $\iota_{[u]}$ est la somme sur toutes les paires dans $D(m[u])$ de toutes ces opérations; $\iota_{[u]}$ définit alors un isomorphisme de

$$\oplus_{(n'[u],n''[u]) \in D(m[u])} \mathbb{C}[\hat{\mathfrak{S}}_{n'[u]} \times \hat{\mathfrak{S}}_{n''[u]}] \rightarrow \mathbb{C}[\hat{W}_{m[u]}].$$

Fixons encore $(n'[u], n''[u]) \in D(m[u])$. On voit maintenant $\mathfrak{S}_{n'[u]} \times \mathfrak{S}_{n''[u]}$ comme un sous-ensemble de $\mathcal{W}_{n'[u]} \times \mathcal{W}_{n''[u]}$. Il y a en fait 2 façons presque naturelles d'envoyer le groupe $\mathfrak{S}_m$ dans le groupe $\mathcal{W}_m$ ($m \in \mathbb{N}$); la première est l'homomorphisme évident $\sigma \mapsto w$ avec $w(\pm i) = \pm\sigma(i)$ pour tout $i \in [1, m]$. La deuxième façon n'est pas un homomorphisme de groupe car elle est définie par $\sigma \mapsto w$ avec $w(\pm i) = \mp\sigma(i)$; bien que cette application n'est pas un morphisme de groupe, elle est équivariante pour l'action adjointe. En revenant à notre inclusion cherchée c'est le produit de la première façon appliquée à $\mathfrak{S}_{n'[u]}$ avec la deuxième appliquée à $\mathfrak{S}_{n''[u]}$. Cela permet alors de restreindre des éléments de $\mathbb{C}[\hat{W}_{m[u]}]$ en des éléments de $\mathbb{C}[\hat{\mathfrak{S}}_{n'[u]} \times \hat{\mathfrak{S}}_{n''[u]}]$. En sommant ces constructions sur toutes les paires dans $D(m[u])$, on obtient $\rho_{[u]}$. Contrairement à $\iota_{[u]}$, $\rho_{[u]}$ n'est pas un isomorphisme mais ce qui est important mais qui n'intervient que de façon cachée dans 6.2 est que le composé $\rho_{[u]} \circ \iota_{[u]}$ est un isomorphisme si on se limite aux fonctions à support dans les éléments U-elliptiques, c'est-à-dire aux permutations qui se décomposent en produit de cycles de longueur impaire (cf. loc. cit.).

On va décrire d'une autre façon cette application $\rho \circ \iota$ précisément quand on se limite aux permutations qui se décomposent en produit de cycles de longueur impaire, en utilisant le fait qu'induire puis restreindre peut aussi se faire en sens inverse, d'abord restreindre puis induire. Pour cela soit $m \in \mathbb{N}$; on note $D(m)$ l'ensemble des couples (m', m'') tels que $m = m'+m''$ et $DD(m)$ l'ensembles des quadruplets $(m^{i,j}; i,j \in \{',''\})$ tels que $\sum_{i,j} m^{i,j} = m$. Soit $(m',m'') \in D(m)$ et $(m^{i,j}) \in DD(m)$; on dit que $(m^{i,j}) <_d (m',m'')$ si $m' = m^{','} + m^{','''}$ et $(m^{i,j}) <_e (m',m'')$ si $m' = m^{','} + m^{'','}$ (d est pour direct et e pour entrelacé). On notera plus génériquement $\underline{m}$ et $\underline{\underline{m}}$ les éléments de $D(m)$ et de $DD(m)$. Pour $\underline{m} \in D(m)$, on considère de façon évidente le groupe $\mathfrak{S}_{\underline{m}} = \mathfrak{S}_{m'} \times \mathfrak{S}_{m''}$; on définit de même les groupes $\mathfrak{S}_{\underline{\underline{m}}}$ pour $\underline{\underline{m}} \in DD(m)$. On note $\chi_{\underline{\underline{m}}}$ le signe $(-1)^{m^{'','''}}$.

Fixons $\underline{m} \in D(m)$; pour $\underline{\underline{m}} \in D(m)$, on définit l'application $res_{d,\underline{m},\underline{\underline{m}}}$ de $\mathbb{C}[\hat{\mathfrak{S}}_{\underline{m}}]$ dans $\mathbb{C}[\hat{\mathfrak{S}}_{\underline{\underline{m}}}]$ comme l'application de restriction évidente si $\underline{\underline{m}} <_d \underline{m}$ et 0 sinon. Et on note $ind_{e,\underline{m},\underline{\underline{m}}}$ l'application de $\mathbb{C}[\hat{\mathfrak{S}}_{\underline{\underline{m}}}]$ dans $\mathbb{C}[\hat{\mathfrak{S}}_{\underline{m}}]$ qui est l'induction si $\underline{\underline{m}} <_e \underline{m}$ et 0 sinon; ici il faut, pour l'induction, considérer l'inclusion naturelle de $\mathfrak{S}_{\underline{\underline{m}}}$ dans $\mathfrak{S}_{\underline{m}}$ où l'on échange d'abord 2e et 3e facteur.

Remarque. *Fixons $\underline{m}_0 \in D(m)$ et considérons $\rho \circ \iota$ comme une application de $\mathbb{C}[\hat{\mathfrak{S}}_{\underline{m}_0}]$ dans $\oplus_{\underline{m} \in D(m)} \mathbb{C}[\hat{\mathfrak{S}}_{\underline{m}}]$. On a alors, en se limitant aux fonctions invariantes de support l'ensemble des permutations ayant des cycles de longueur impaire:*

$$\rho \circ \iota = \oplus_{\underline{m} \in D(m)} \sum_{\underline{\underline{m}} \in DD(m)} ind_{e,\underline{m},\underline{\underline{m}}} \circ \left(\chi_{\underline{\underline{m}}} res_{d,\underline{m}_0,\underline{\underline{m}}} \right)$$

Le seul point est de remarquer que sur les permutations n'ayant que des cycles de longueur impaire le signe $\chi_{\underline{\underline{m}}}$ coïncide avec sgn_{CD} tel que définit ci-dessus.

Définition. *Dans la suite, on définit $\rho \circ \iota$ comme dans la remarque ci-dessus.*

Les définitions du cas $[u] = [\pm 1]$ sont plus compliquées (cf. [8] en particulier 3.18 et 3.19) à cause de la partie cuspidale. On les présente ainsi. On fixe $\epsilon \in \{\pm 1\}$; on doit définir une application de $\hat{W}_{D(n[\epsilon])}$ dans $\mathbb{C}[\hat{W}_{D(n[\epsilon])}]$ que l'on prolongera linéairement en un endomorphisme de $\mathbb{C}[\hat{W}_{D(n[\epsilon])}]$. Et on veut l'interpréter comme une restriction suivie d'une induction tordue par un caractère. Un élément de $\hat{W}_{D(n[\epsilon])}$ est la donnée de deux symboles l'un de défaut impair et l'autre de défaut pair dont la somme des rangs est $n'[\epsilon]$. De façon beaucoup plus compliquée mais équivalente (et qui permet de parler de représentations) c'est la donnée:

d'un entier impair I_ϵ, d'un entier pair P_ϵ, d'un signe ζ_ϵ, de 2 entiers $N'_\epsilon, N''_\epsilon$ tous ces nombres vérifiant l'égalité $N'_\epsilon + N''_\epsilon + (I_\epsilon^2 + P_\epsilon^2 - 1)/4 = n[\epsilon]$ et de 2 représentations irréductibles l'une de $W_{N'_\epsilon}$ et l'autre de $W_{N''_\epsilon}$, les groupes de Weyl de type C et de rang écrit en indice. On suppose que $\zeta_\epsilon = (-1)^{(I_\epsilon - 1)/2}$ si $P_\epsilon = 0$.

On pose $\tilde{\zeta}_\epsilon := (-1)^{(I_\epsilon - 1)/2} \zeta_\epsilon$; en particulier, on a $\tilde{\zeta} = 1$ si $P_\epsilon = 0$.

On note $\tilde{\chi}$ le caractère trivial si $I_\epsilon > P_\epsilon$ et le caractère sgn_{CD} sinon.

On définit une application de $\mathbb{C}[\hat{W}_{N'_\epsilon} \times \hat{W}_{N''_\epsilon}]$ dans

$$\oplus_{(M'_\epsilon, M''_\epsilon) | M'_\epsilon + M''_\epsilon = N'_\epsilon + N''_\epsilon} \mathbb{C}[\hat{W}_{M'_\epsilon} \times \hat{W}_{M''_\epsilon}]$$

par restriction puis induction tordue de façon similaire le cas des groupes symétriques; c'est-à-dire que l'on fixe $M'_\epsilon, M''_\epsilon$ avec $M'_\epsilon + M''_\epsilon = N'_\epsilon + N''_\epsilon$ et considère les quadruplets $N_\epsilon^{i,j}$ où $i, j \in \{', ''\}$ vérifiant $N_\epsilon^{i,'} + N_\epsilon^{i,''} = N_\epsilon^i$ pour $i =' $ et $''$ et $N_\epsilon^{',j} + N_\epsilon^{'',j} = M_\epsilon^j$ pour $j =' $ et $''$, s'il en existe (sinon on ne fait rien pour ce choix de $M'_\epsilon, M''_\epsilon$). On restreint la représentation de $W_{\tilde{N'}_\epsilon} \times W_{\tilde{N}''_\epsilon}$ au groupe $\times_{i,j} W_{N_\epsilon^{i,j}}$, on tensorise la restriction par le caractère de ce groupe qui vaut: $sgn_{CD}^{(1-\zeta_\epsilon)/2}$ sur $W_{N'_\epsilon,''}$, $\tilde{\chi}$ sur $W_{N''_\epsilon,'}$, 1 sur $W_{N'_\epsilon,'}$ et $\tilde{\chi} sgn_{CD}^{(1+\zeta_\epsilon)/2}$ sur $W_{N''_\epsilon,''}$.
On induit au groupe $W_{M'_\epsilon} \times W_{M''_\epsilon}$ après avoir échangé les 2e et 3e facteurs, c'est-à-dire $W_{N'_\epsilon,''}$ et $W_{N''_\epsilon,'}$. Puis on somme sur tous les quadruplets. Ensuite on identifie $\oplus_{(M'_\epsilon, M''_\epsilon) | M'_\epsilon + M''_\epsilon = N'_\epsilon + N''_\epsilon} \mathbb{C}[\hat{W}_{M'_\epsilon} \times \hat{W}_{M''_\epsilon}]$ à l'ensemble des combinaisons linéaires de base les couples de symboles ordonnés le premier de défaut impair égal à I_ϵ et le deuxième de défaut pair égal à $\tilde{\zeta}_\epsilon P_\epsilon$.
C'est la construction de [8] 3.18 que l'on a complètement explicitée. C'est assez compliqué; remarquons que la présence du caractère $\tilde{\chi}$ n'a joué de rôle dans [8] qu'en 5.5. Il en est de même ici, ce caractère ne joue aucun rôle sauf dans l'énoncé de la conjecture 6.2. La prise en compte de ζ_ϵ dans la définition, elle joue un rôle mais dans la définition de k_χ, ζ_ϵ aussi joue un rôle et en fait ces 2 prises en compte se compensent en grande partie (cf. la preuve de 4.2).

4 Localisation

4.1 Localisation des faisceaux caractères

On va avoir besoin d'une formule due à Lusztig qui calcule les faisceaux caractères au voisinage des points semi-simples en terme de fonctions de Green. Elle est écrite en toute généralité dans [5] et explicitée dans certains cas dans [13] et [8]; c'est la présentation de [13] par.7 que l'on reprend. On fixe un élément semi-simple g_s de $SO(2n+1, F)_\sharp$ et on suppose que toutes les valeurs propres de g_s sont des racines de l'unité d'ordre premier à p. On fixe aussi g_u un élément topologiquement unipotent de $SO(2n+1, F)_\sharp$ commutant à g_s. On suppose qu'il existe un parahorique $K_{n',n''}$ (pour n', n'' convenables) contenant $g_s g_u$ et on note s, u les réductions de g_s et g_u modulo le radical pro-p-unipotent. On se donne aussi $\underline{n} \in D(\chi)$ que l'on suppose relatif à (n', n'') au sens $\sum_{[u] \in [VP(\chi)]} n'([u]) \ell_{[u]} = n'$ et $\sum_{[u] \in [VP(\chi)]} n''([u]) \ell_{[u]} = n''$.
Dans la suite, on fixe, pour $\epsilon \in \{\pm 1\}$, des données $I_\epsilon, P_\epsilon, \tilde{\zeta}_\epsilon$ comme dans les paragraphes précédents, c'est-à-dire une donnée cuspidale $\tilde{cusp}$; on a choisi cette notation pour qu'elle soit analogue à celle de la fin de 3.4. Donc en particulier, la propriété de $\tilde{\zeta}_\epsilon$ est de vérifier $\tilde{\zeta}_\epsilon = +$ si $P_\epsilon = 0$. Et dans l'espace vectoriel

$$\oplus_{\epsilon \in \{\pm 1\}} \oplus_{m'(\epsilon), m''(\epsilon); m'(\epsilon) + m''(\epsilon) = m(\epsilon)} \mathbb{C}[\hat{\mathcal{W}}_{m'(\epsilon)}] \otimes \mathbb{C}[\hat{\mathcal{W}}_{m'(\epsilon)}]$$

on ne regarde que les sous-espaces vectoriels correspondant aux symboles relatifs à ces donnés cuspidales. Pour manifester cette restriction, on ajoute $c\tilde{usp}$ en indice. En fonction de ce que l'on a rappelé en 3.2 cela revient au même que de regarder les représentations des différents groupes $\times_{\epsilon\in\{\pm 1\}} W_{N'(\epsilon)}\times W_{N''(\epsilon)}$, où W est un groupe de Weyl de type C et où les nombres $N'(\epsilon), N''(\epsilon)$ vérifient les conditions:

$$2N'(\epsilon)+2N''(\epsilon)+(I_\epsilon^2+P_\epsilon^2)/2=m(\epsilon).$$

On utilise la convention que si $P_\epsilon=0$ alors $\tilde{\zeta}=+1$. On pose $\zeta_\epsilon=(-1)^{(I_\epsilon-1)/2}\tilde{\zeta}_\epsilon$ et on retrouve la convention de 3.4 que si $P_\epsilon=0$, alors $\zeta_\epsilon=(-1)^{(I_\epsilon-1)/2}$.
Soit alors ϕ un élément de $\mathbb{C}[\hat{\mathcal{W}}_{\underline{n},c\tilde{usp}}]$. Le point est de calculer $k_\chi(\phi)(g_sg_u)$ à l'aide des fonctions de Green du commutant de s dans le groupe $K_{n',n''}$ en réduction. Pour le faire, on suppose s elliptique.
On décrit d'abord le commutant de g_s dans $SO(2n+1,F)$; on note $[VP(g_s)]$ l'ensemble des valeurs propres de g_s regroupées en paquets λ et λ' sont dans le même paquet s'il existe $a\in\mathbb{N}$ tel que $\lambda'=\lambda^{q^a}$. On note $m([\lambda])$ la multiplicité de λ. Si $\lambda\neq\pm 1$, on note $\ell_{[\lambda]}:=1/2\,|[\lambda]|$. Pour unifier pour $\epsilon=\pm$, on pose $\ell_\epsilon=1$. Pour $\lambda\neq\pm 1$, on note $F_{2\ell_{[\lambda]}}$ l'extension non ramifiée de F de degré $2\ell_{[\lambda]}$ et il existe une forme hermitienne (pour l'extension $F_{2\ell_{[\lambda]}}/F_{\ell_{[\lambda]}}$) $<,>_{[\lambda]}$ sur l'espace vectoriel sur $F_{2\ell_{[\lambda]}}$ de dimension $m([\lambda])$ tel que la partie du commutant de g_s relative à la valeur propre λ soit précisément le groupe unitaire de cette forme. Des formes hermitiennes, comme ci-dessus, il y en a exactement 2 qui se distinguent par la parité de la valuation du déterminant, donc par un signe que nous noterons $\epsilon_{[\lambda]}$. Si l'on note $U_{\epsilon_{[\lambda]}}$ le groupe de la forme correspondant à $\epsilon_{[\lambda]}$, on rappelle que $U_{\epsilon_{[\lambda]}}\simeq U_{-\epsilon_{[\lambda]}}$ si $m([\lambda])$ est impair; mais nos constructions dépendront de $\epsilon_{[\lambda]}$ comme en [8] 3.3 et suivants, et il faut donc garder la distinction.
Une valeur propre dans $\{\pm 1\}$ introduit une forme orthogonale $<,>_\pm$ ($\pm$ est ici le signe de la valeur propre considérée) sur un F-espace vectoriel de dimension $m(\pm 1)$ et la partie du commutant qui lui correspond est le groupe orthogonal de la forme; on utilise la notation $\eta_\pm$ et $\epsilon_\pm$ pour le signe du discriminant et l'invariant de Hasse; il y a toujours un problème sur la normalisation du discriminant et ici on suit les conventions de [8], c'est-à-dire que le discriminant est invariant par ajout de plans hyperboliques mais il n'est donc pas additif. Le signe du discriminant est l'image du discriminant par le caractère quadratique non ramifié de F^* et la non additivité n'est un problème pour le signe du discriminant que si -1 n'est pas un carré.
Remarquons tout de suite, puisque l'on en aura besoin, que toutes les quantités qui viennent d'être introduites sont constantes sur la classe de conjugaison stable de g_s sauf la famille des $\epsilon_{[\lambda]}$ pour $[\lambda]\in[VP(g_s)]$. Cette famille est soumise à la condition $\times_{[\lambda]}\epsilon_{[\lambda]}=\sharp$ et l'ensemble de ces familles soumises à cette condition paramétrise l'ensemble des classes de conjugaison dans la classe de conjugaison stable de g_s dans $SO(2n+1,F)_\sharp$. Si l'on enlève la condition de produit, l'ensemble plus grand paramétrise la classe de conjugaison stable de g_s dans $SO(2n+1,F)_{iso}\cup SO(2n+1,F)_{an}$.

Revenons au parahorique $K_{n',n''}$ que l'on a fixé au début et qui contient g_s et g_u. On définit s, u comme ci-dessus. Décrivons le commutant de s dans la partie réductive du parahorique; on écrit $s = (s', s'')$ avec $s' \in SO(2n'+1, \mathbb{F}_q)$ et $s'' \in O(2n'', \mathbb{F}_q)$. Les valeurs propres de s sont les "mêmes" que celles de g_s ; pour $[\lambda] \in [VP(g_s)]$, on note $m'([\lambda])$ la multiplicité d'un élément $\lambda \in [\lambda]$ comme valeur propre de s' et $m''([\lambda])$ l'analogue pour s''. On a $m'([\lambda]) + m''([\lambda]) = m([\lambda])$. De même pour $[\lambda] \in [VP(g_s)]$, $\lambda \neq \pm 1$, il existe une forme hermitienne $<,>'_{[\lambda]}$ (resp. une forme hermitienne $<,>''_{[\lambda]}$) sur un $\mathbb{F}_{2\ell_{[\lambda]}}$ espace vectoriel de dimension $m'([\lambda])$ (resp. $m''([\lambda])$) et des formes orthogonales pour $\lambda = \pm 1$, telles que le commutant de s' (resp. s'') dans $SO(2n'+1, \mathbb{F}_q)$ (resp. $O(2n'', \mathbb{F}_q)$) soit les éléments de déterminant 1 dans le produit (resp. le produit) des groupes de ces formes. Il y a évidemment des rapports entre $<,>'_{[\lambda]}, <,>''_{[\lambda]}$ et $<,>_{[\lambda]}$.
Supposons d'abord que $\lambda \notin \{\pm 1\}$. Si $\epsilon_{[\lambda]} = 1$ (défini ci-dessus), alors $m''([\lambda])$ est nécessairement pair alors que ce nombre est impair si $\epsilon_{[\lambda]} = -1$, ensuite $<,>'_{[\lambda]} \otimes <,>''_{[\lambda]}$ est obtenu par réduction (à l'aide d'un réseau convenable) de $<,>_{[\lambda]}$.
Supposons maintenant que $\lambda \in \{\pm 1\}$. Alors $m''([\lambda])$ a la même parité que $v_F(\eta_{[\lambda]})$. On note $\eta'_{[\lambda]}$ et $\eta''_{[\lambda]}$ les discriminants des formes $<,>'_{[\lambda]}$ et $<,>''_{[\lambda]}$; on les voit comme des signes, c'est-à-dire qu'au lieu de regarder le discriminant comme un élément de $\mathbb{F}_q$ modulo les carrés, on regarde sont image dans $\{\pm 1\}$. Si $v_F(\eta_{[\lambda]})$ est pair, l'image de $\eta''_{[\lambda]}$ dans $\mathbb{F}_q^*/\mathbb{F}_q^{*2}$ est $\epsilon_{[\lambda]}$ tandis que si $v_F(\eta_{[\lambda]})$ est impair c'est l'image de $\eta'_{[\lambda]}$ dans le même groupe qui est $\epsilon_{[\lambda]}$. De plus le signe du discriminant de $<,>_{[\lambda]}$ vérifie:
$sgn(\eta_{[\lambda]}) = sgn(-1)^{m'([\lambda])m''([\lambda])}(-1)^{v_F(\eta_{[\lambda]})}\epsilon_{[\lambda]}\eta^i_{[\lambda]}$, où $i([\lambda]) ='$ si $v_F(\eta_{[\lambda]})$ est pair et $''$ sinon.
On pose ici pour $[\lambda] \neq [\pm 1]$, $\mathcal{W}_{\underline{m}_{g_s}(\lambda)} := \mathfrak{S}_{m'_{g_s}(\lambda)} \times \mathfrak{S}_{m''_{g_s}(\lambda)}$. Pour $\lambda = \pm 1$, il y a des difficultés liées à l'existence de faisceaux cuspidaux; ici on ne s'intéresse qu'aux fonctions de Green et les paramètres pour la partie cuspidales sont alors des quadruplets d'entiers positifs ou nuls. On écrit les choses comme on en aura besoin; on avait fixé ci-dessus une donnée cuspidale, $\tilde{cusp}$; on note $|cusp|$ un quadruplet d'entiers positifs ou nuls, $|r|^i_{\epsilon'}$ pour $i \in \{', ''\}$ et $\epsilon' \in \{\pm 1\}$, qui est la partie cuspidale pour les fonctions de Green. Ce quadruplet dépend de $\tilde{cusp}$ par les formules:
$|r|'_{+1}, |r|'_{-1}$ est à l'ordre près le couple $I_+ + I_-, |I_+ - I_-|$ avec $|r|'_{+1}$ impair par hypothèse et $|r|''_{+1}, |r|''_{-1}$ est à l'ordre près le couple $P_+ + P_-, |P_+ - P_-|$ avec $|r|''_{+1} \geq |r|''_{-1}$ si et seulement si $\zeta_+\zeta_- = (-1)^{1+(I_+ + I_-)/2}$.
On pose alors pour $\lambda \in \{\pm 1\}$ que l'on note plutôt ϵ', et pour un couple d'entier $m'(\epsilon'), m''(\epsilon')$ vérifiant $m'(\epsilon') + m''(\epsilon') = m(\epsilon')$

$$\mathcal{W}_{m'(\epsilon'), m''(\epsilon'), |cusp|} := W_{1/2(m'(\epsilon') - |r'_{\epsilon'}|^2)} \times W_{1/2(m''(\epsilon') - |r''_{\epsilon'}|^2)}$$

étant entendu que ce groupe est nul si l'un des indices n'est pas un entier positif ou nul.
Pour la donnée d'un ensemble de couple $\underline{m}_{g_s} := \{m'(\lambda), m''(\lambda) \in D(m_{g_s}(\lambda))\}$, on pose $\mathcal{W}_{\underline{m}_{g_s}, |cusp|}$ le produit des groupes définis ci-dessus. Et les fonctions

de Green donnent une application de $\mathbb{C}[\hat{\mathcal{W}}_{\underline{m}_{g_s}}]$ dans l'ensemble des fonctions à support unipotent du commutant de g_s dans $SO(2n'+1,\mathbb{F}_q)\times O(2n'',\mathbb{F}_q)$; cela s'interprète encore comme des fonctions à support dans l'ensemble des éléments topologiquement unipotents de $K_{n',n''}$ commutant à g_s. Quand on somme sur tous les $\underline{m}_{g_s}$, on définit $\mathbb{C}[\hat{\mathcal{W}}_{D(g_s),|cusp|}]$ et on s'autorisera la suppression du *cusp* quand il n'y a pas d'ambiguité.

On revient maintenant à ϕ et donc pour tout $[u]\in[VP(\chi)]$ on a un élément $(n'[u],n''[u])\in D(m([u]))$ de telle sorte que $\sum_{[u]\in[VP(\chi)]} n'[u]\ell_{[u]} = n'$ et on rappelle que l'on a aussi fixé une donnée cuspidale, $\tilde{cusp}$. On note $\underline{n}$ l'ensemble de ces paires et on a défini $\mathcal{W}_{\underline{n},\tilde{cusp}}$. On va définir une application de $\mathbb{C}[\hat{\mathcal{W}}_{\underline{n}}]$ dans $\mathbb{C}[\hat{\mathcal{W}}_{\underline{m}_{g_s}}]$. Mais on a besoin d'un certain nombre d'objets intermédiaires. Soit une collection de paires $\underline{\nu}([u],[\lambda]) = (\nu'([u],[\lambda]),\nu''([u],[\lambda]))$ soumises aux conditions, où $\ell_{[u]}$ et $\ell_{[\lambda]}$ sont comme ci-dessus (pour a,b des entiers, on note (a,b) le pgcd de ces nombres):

$$\sum_{[u]} \tilde{\nu}([u],[\lambda])\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]}) = \underline{m}_{g_s}([\lambda]);$$

la somme de couples, se fait terme à terme. Et on a aussi:

$$\sum_{[\lambda]} \nu([u],[\lambda])/(\ell_{[u]},\ell_{[\lambda]}) = \underline{n}([u]).$$

On pose $\mathcal{W}_{\nu([u],[\lambda])} := \mathfrak{S}_{\nu'([u],[\lambda])}\times\mathfrak{S}_{\nu''([u],[\lambda])}$ si soit $[u]$ soit $[\lambda]$ n'est un élément de $\{\pm1\}$. Pour traiter le cas de ±1, on réutilise la donnée de la partie cuspidale $|cusp|$ et on pose, pour $\epsilon\in\{\pm1\}$ et $\epsilon'\in\{\pm1\}$ $\mathcal{W}_{\underline{\nu}([\epsilon],[\epsilon'])} =$ $W_{1/2(\nu'([\epsilon],[\epsilon'])-(|r|'_{\epsilon'})^2)}\times W_{1/2(\nu''([\epsilon],[\epsilon'])-(|r|''_{\epsilon'})^2)}$; en particulier cela sous-entend que ces nombres écrits en indice sont des entiers positifs ou nuls (sinon le groupe défini est 0). Et on note $\mathcal{W}_{\underline{\nu}(\chi,g_s),cusp}$ l'union de tous ces groupes.

Pour $\delta ='$ ou $''$ et pour $[u]\in[VP(\chi)]$, $[\lambda]\in[VP(g_s)]$, on a besoin de définir une application de $\mathcal{W}_{\nu^\delta([u],[\lambda])}$ dans $\mathcal{W}_{n^\delta([u])_{[\lambda]}}$ et dans $\mathcal{W}_{n^\delta([\lambda])_{[u]}}$ on précisera dans chaque cas les rapports entre les entiers $\nu^\delta([u],[\lambda]), n^\delta([u])_{[\lambda]}$ et $\nu^\delta([\lambda])_{[u]}$. Cette application n'est pas un morphisme de groupes mais simplement compatible à l'action adjointe; au passage, on définit aussi une fonction invariante par conjugaison sur $\mathcal{W}_{\nu^\delta([u],[\lambda])}$ que l'on note $\chi^\delta_{[u],[\lambda]}$. Le point qui n'est pas nouveau est que pour les groupes unitaires qui interviennent soit pour $[u]$ quand $[u]\neq\pm1$ soit pour $[\lambda]$ quand $[\lambda]\neq\pm1$, un tore n'est pas vraiment associé à un élément du groupe symétrique convenable mais au produit d'un tel élément par un Frobénius.

On décrit ces constructions au cas par cas: premier cas: $[u],[\lambda]\in\pm1$; ici, on veut $\nu^\delta([u],[\lambda]) = n^\delta([u])_{[\lambda]}$ et l'application de $\mathcal{W}_{\nu^\delta([u],[\lambda])} = W_{\nu^\delta([u],[\lambda])}$ dans $\mathcal{W}_{n^\delta([u])_{[\lambda]}} = W_{n^\delta([u])_{[\lambda]}}$ est l'identité. On fait une construction analogue en remplaçant $[u]$ par $[\lambda]$. Quand à la fonction $\chi^\delta_{[u],[\lambda]}$ c'est l'identité sauf si $u=\lambda=-1$ où c'est le $sgn_{CD}(-1)^{\nu^\delta([u],[\lambda])(q-1)/2}$.

deuxième cas: $[u] \neq [\pm 1]$, $[\lambda] \neq [\pm 1]$ et $\ell_{[u]}$ et $\ell_{[\lambda]}$ sont divisibles par la même puissance de 2 (cette dernière condition donne des renseignements sur les tores des groupes unitaires intervenant). L'application de $\mathcal{W}_{\nu^\delta([u],[\lambda])} = \mathfrak{S}_{\nu^\delta([u],[\lambda])}$ dans $\mathfrak{S}_{n^\delta([u])_{[\lambda]}}$ se décrit quand $n^\delta([u])_{[\lambda]} = \nu^\delta([u],[\lambda])\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]})$. A une permutation dont les cycles sont $(\alpha_1,\cdots,\alpha_r)$ on associe la permutation de cycles $(\alpha_1\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]}),\cdots,\alpha_r\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]}))$. Quand on considère $[\lambda]$ au lieu de $[u]$, on échange simplement les rôles de $\ell_{[u]}$ et $\ell_{[\lambda]}$. Pour décrire la fonction $\chi^\delta_{[u],[\lambda]}$, on a besoin d'une notation auxiliaire. On pose, pour $\alpha \in \mathbb{N}$, $y \in \mathbb{Q}$:

$$\phi_{\alpha,[\lambda],y} := y^{-1}(\ell_{[\lambda]})^{-1}(\lambda^\alpha + \lambda^{\alpha q} + \cdots + \lambda^{\alpha q^{\ell_{[\lambda]}-1}} + \lambda^{-\alpha} + \lambda^{-\alpha q} + \cdots + \lambda^{-\alpha q^{\ell_{[\lambda]}-1}}),$$

où λ est n'importe quel élément dans $[\lambda]$. Alors sur l'élément w associé aux cycles $\alpha_1,\cdots,\alpha_r$, $\chi^\delta_{[u],[\lambda]}(w) = \prod_{s\in[1,r]} \phi_{\alpha_s,[\lambda],(\ell_{[u]},\ell_{[\lambda]})}$.
troisième cas: $[u] \neq [\pm 1]$, $[\lambda] \neq [\pm 1]$ et $\ell_{[u]}$ et $\ell_{[\lambda]}$ ne sont pas divisibles par la même puissance de 2. Ici on veut alors, $n^\delta([u])_{[\lambda]} = 2\nu^\delta([u],[\lambda])\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]})$. A une permutation dont les cycles sont $(\alpha_1,\cdots,\alpha_r)$ on associe, ici, la permutation de cycles $(2\alpha_1\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]}),\cdots,2\alpha_r\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]}))$. Quand on considère $[\lambda]$ au lieu de $[u]$, on échange simplement les rôles de $\ell_{[u]}$ et $\ell_{[\lambda]}$. Et la fonction $\chi^\delta([u],[\lambda])$ vaut sur ce w, $\prod_{s\in[1,r]} \phi_{\alpha_s,[\lambda],(\ell_{[u]},\ell_{[\lambda]})/2}$.
quatrième cas: $u = \pm 1$ et $[\lambda] \neq [\pm 1]$. Ici on veut $n^\delta([u])_{[\lambda]} = \nu^\delta([u],[\lambda])\ell_{[\lambda]}$ et $n^\delta([\lambda])_{[u]} = \nu^\delta([u],[\lambda])$. L'application de $\mathfrak{S}_{\nu^\delta([u],[\lambda])}$ dans $\mathfrak{S}_{n^\delta([\lambda])_{[u]}}$ est l'identité. L'application de $\mathfrak{S}_{\nu^\delta([u],[\lambda])}$ dans $W_{n^\delta([u])_{[\lambda]}}$ envoie l'élément w de cycle $(\alpha_1,\cdots,\alpha_r,\alpha'_1,\cdots,\alpha'_{r'})$, où les α sont pairs et les α' impairs sur l'élément de $W_{n^\delta([u])_{[\lambda]}}$ qui correspond à 2 partitions $(\alpha_1\ell_{[\lambda]},\cdots,\alpha_r\ell_{[\lambda]};\alpha'_1\ell_{[\lambda]},\cdots,\alpha'_{r'}\ell_{[\lambda]})$. Quand à la fonction $\chi^\delta([u],[\lambda])$, c'est l'identité si $[u] = [+1]$ et est constante de valeur $(-1)^{\nu^\delta([u],[\lambda])(1+q^{\ell_{[\lambda]}})/2}$ pour $[u] = [-1]$.
cinquième cas: $[u] \neq [\pm 1]$ et $\lambda = \pm 1$. On échange les rôles de $[u]$ et $[\lambda]$ dans le cas ci-dessus.
Il faut aussi prendre en compte une contribution de la partie cuspidale; là il n'y a pas de groupes mais simplement une fonction à définir, $\chi^\delta_{cusp,\underline{\nu}}$ (où $\delta \in \{','' \}$ comme ci-dessus) qui dépend de $\underline{\nu} \in D(\chi, g_s)$ et de $w \in \mathcal{W}_{\underline{\nu}}$. Pour cela, on pose $\chi^\delta_{+,\underline{\nu}}(w) := (-1)^{\sum_{[u]\in[VP(\chi)]-\{\pm 1\}} \nu^\delta([u],+)} sgn_{CD} w_{\nu^\delta(+,+)} w_{\nu^\delta(-,+)}$ et une définition analogue, $\chi^\delta_{-,\nu}$ en remplaçant $+$ par $-$. On note aussi ϵ_I (resp. ϵ_P) le signe tel que $I_{\epsilon_I} - I_{-\epsilon_I} > 0$ (resp. $P_{\epsilon_P} - P_{-\epsilon_P} > 0$); si les 2 nombres sont égaux à 0, on prend le signe $\epsilon_I = \epsilon_P = +$ par convention et si l'un seulement des nombres vaut 0, alors on prend par convention $\epsilon_I = \epsilon_P$. On pose r'_{im} l'élément impair du couple $(I_+ + I_-)/2, |I_+ - I_-|/2$ et l'on a ($x' \in \mathbb{N}$ ne nous sert à rien mais est le $\delta(r', r'')$ de [13] et il en est de même pour x''):

$$\chi'_{cusp,\underline{\nu}}(w) = q^{x'}(-1)^{(q-1)(r'-1)/4} \times$$
$$(\chi'_{+,\underline{\nu}}(w)\chi'_{-,\underline{\nu}}(w)\eta'_+(g_s)\eta'_-(g_s))^{(I_{\epsilon_I}-1)/2} \begin{cases} 1 \text{ si } \epsilon_I = + \\ \eta'_-(g_s)\chi'_{-,\underline{\nu}}(w) \text{ si } \epsilon_I = -. \end{cases}$$

ci-dessous y'' est r''_+ si ce nombre est pair, et $r''_- - 1$ sinon,

$$\chi''_{cusp,\underline{\nu}}(w) = q^{x''}(-1)^(y''(q-1)/4)(\eta''_+(g_s)\eta''_-(g_s)\chi''_{+,\underline{\nu}}\chi''_{-,\underline{\nu}})^{(I_{\epsilon_P}-1)/2}$$

$$\begin{cases} 1 \text{ si } \zeta_{\epsilon_P} > 0 \text{ et } \epsilon_P = + \\ \eta''_-(g_s)\chi''_{-,\underline{\nu}}(w) \text{ si } \zeta_{\epsilon_P} = + \text{ et } \epsilon_P = - \\ \eta''_+(g_s)\chi''_{+,\underline{\nu}}(w) \text{ si } \zeta_{\epsilon_P} = - \text{ et } \epsilon_P = - \\ \eta''_+(g_s)\eta''_-(g_s)\chi''_{+,\underline{\nu}}(w)\chi''_{-,\underline{\nu}}(w) \text{ si } \zeta_{\epsilon_P} = - \text{ et } \epsilon_P = +. \end{cases}$$

Le produit de ces 2 fonctions se simplifie un peu. On écrit ce produit comme le produit des 3 termes
C_{cusp} qui est une constante,
$c_{cusp}(g_s) := (\eta'_+(g_s)\eta'_-(g_s))^{(I_{\epsilon_I}-1)/2}(\eta''_+(g_s)\eta''_-(g_s))^{(I_{\epsilon_P}-1)/2}\times$

$$\begin{cases} 1 \text{ si } \zeta_{\epsilon_P} = + \\ \eta'_+(g_s)\eta'_-(g_s) \text{ si } \zeta_{\epsilon_P} = - \end{cases} \times \begin{cases} 1 \text{ si } \epsilon_I = + \\ \eta'_-(g_s) \text{ si } \epsilon_I = - \end{cases} \times \begin{cases} 1 \text{ si } \epsilon_P = + \\ \eta''_-(g_s) \text{ si } \epsilon_P = - \end{cases}$$

$\chi_{cusp,\underline{\nu}}(w) := (\chi'_{+,\underline{\nu}}(w)\chi'_{-,\underline{\nu}}(w))^{(I_{\epsilon_I}-1)/2}(\chi''_{+,\underline{\nu}}(w)\chi''_{-,\underline{\nu}}(w))^{(I_{\epsilon_P}-1)/2}\times$

$$\begin{cases} 1 \text{ si } \zeta_{\epsilon_P} = + \\ \chi'_{+\underline{\nu}}(w)\chi'_{-,\underline{\nu}}(w) \text{ si } \zeta_{\epsilon_P} = - \end{cases} \times \begin{cases} 1 \text{ si } \epsilon_I = + \\ \chi'_{-,\underline{\nu}}(w) \text{ si } \epsilon_I = - \end{cases} \times \begin{cases} 1 \text{ si } \epsilon_P = + \\ \chi''_{-,\underline{\nu}}(w) \text{ si } \epsilon_P = - \end{cases}$$

Supposons maintenant donné $\underline{n} \in D(\chi)$ et $\underline{m} \in D(g_s)$. Soit $\underline{\nu} \in D(\chi, g_s)$ et on suppose que les égalités suivantes sont vérifiées:
pour $\delta \in \{',''\}$, pour tout $[u] \in [VP(\chi)]$,

$$\sum_{[\lambda]\in[VP(g_s)]} \nu^\delta([u],[\lambda])2^{x_{[u],[\lambda]}}\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]}) = n^\delta([u]),$$

où $x_{[u],[\lambda]} = 0$ sauf si $\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]})$ ou $\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]})$ est pair, où il vaut 1.
Et pour $\delta \in \{',''\}$, pour tout $[\lambda] \in [VP(g_s)]$,

$$\sum_{[u]\in[VP(\chi)]} \nu^\delta([u],[\lambda])2^{x_{[u],[\lambda]}}\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]}) = m^\delta([\lambda]),$$

où $x_{[u],[\lambda]}$ est comme ci-dessus.
En faisant des produits convenables des constructions ci-dessus, on a une application de $\mathcal{W}_{\underline{\nu}}$ d'une part dans $\mathcal{W}_{\underline{n}}$ et d'autre part dans $\mathcal{W}_{\underline{m}}$. Partant donc d'un élément de $\mathbb{C}[\hat{\mathcal{W}}_{\underline{n}}]$ on peut le restreindre en un élément de $\mathbb{C}[\hat{\mathcal{W}}_{\underline{\nu}}]$, le multiplier par le $\chi'_{cusp,\underline{\nu}}\chi''_{cusp,\underline{\nu}}\prod_{[u]\in[VP(\chi)],[\lambda]\in[VP(g_s)]}\chi'([u],[\lambda])\chi''([u],[\lambda])$ puis l'induire en un élément de $\mathbb{C}[\hat{\mathcal{W}}_{\underline{m}}]$.
L'application cherchée est la somme sur toutes les collections $\underline{\nu}(\chi, g_s)$ comme ci-dessus. Elle est notée, $loc_{g_s,\underline{m}}$.
Définition. *Pour g_s comme ci-dessus, on note $loc_{g_s} := \sum_{\underline{m}\in D(g_s)} loc_{g_s;\underline{m}}$.*

On rappelle que les fonctions de Green définissent, pour $\underline{m} \in D(g_s)$, une application de $\hat{\mathcal{W}}_{\underline{m}}$ dans les fonctions sur le centralisateur de s (la réduction de g_s) dans la réduction de $K_{n',n''}$; on remarque que n' et n'' sont déterminés par $\underline{m}$. On remonte ensuite en des fonctions sur un sous-groupe compact ouvert convenable du centralisateur de g_s dans $SO(2n+1, F)_\sharp$ par invariance sous le radical pro-p-unipotent puis on les prolonge par zéro en des fonctions sur ce centralisateur. On note Q cette application. On peut donc définir $Q(loc_{g_s}(\rho))$ pour $\rho \in \hat{\mathcal{W}}_{\underline{n}}$ où $\underline{n} \in D(\chi)$ (avec une donnée cuspidale fixée)

Remarque. *La définition précédente ne dépend de g_s dans sa classe de conjugaison stable que par le facteur $c_{cusp}(g_s)$.*

Pour cette section, plutôt que de travailler avec k_χ qui a été normalisé pour avoir de bonnes propriétés (cf. [13]), on énonce un résultat pour un analogue de k_χ non normalisé, c'est-à-dire que les fonctions de Green généralisées associent, pour $\underline{n}$ fixé dans $D(\chi)$ et une donnée cuspidale fixée $c\tilde{u}sp$, à un élément du groupe $w \in \mathcal{W}_{\underline{n},c\tilde{u}sp}$ une fonction sur un groupe fini convenable. Et, pour ϕ une représentation de $\mathcal{W}_{\underline{n},c\tilde{u}sp}$, on a défini $k_\chi(\phi) := \sum_{w \in \mathcal{W}_{\underline{n},c\tilde{u}sp}} \lambda(w) tr(\phi)(w) k_\chi(w)$, où $\lambda(w)$ est un caractère qui dépend de la donnée cuspidale. On utilise ici simplement $k_\chi^{nn}(\phi)$ l'analogue en supprimant λ. La raison est que $\rho \circ \iota$ dépend aussi du support cuspidal et que l'on ne s'intéresse qu'au composé $k_\chi \circ \rho \circ \iota$; les normalisations s'annulent partiellement et il vaut donc mieux ne pas se fatiguer à les faire. Et on a, pour g_u un élément topologiquement unipotent qui commute à g_s:

Lemme: *Pour $c\tilde{u}sp$, $\underline{n}$ et ϕ comme ci-dessus, (i)$k_\chi^{nn}(\phi)(g_s g_u) = Q(loc_{g_s;\underline{m}}(\phi))(g_u)$.*

(ii)On a les égalités d'intégrales orbitales (où le groupe est mis en exposant)

$$I^{SO(2n+1,F)_\sharp}(g_s, k_\chi^{nn}(\phi)) = I^{Cent^0_{SO(2n+1,F)}g_s}(g_u, Q(loc_{g_s}(\phi)).$$

(i) C'est un problème sur le groupe fini $SO(2n'+1, \mathbb{F}_q) \times O(2n'', \mathbb{F}_q)$ où il s'agit de localiser au voisinage de la réduction de g_s le faisceau caractère associé à $\underline{\rho}$. Cela a été fait en toute généralité par Lusztig et il faut expliciter ses formules. En [8] 2.16, on a traité le cas où $[VP(\chi)]$ est réduit à $+1$ mais il n'y a pas de restriction sur g_s. En [13] est traité le cas où $[VP(\chi)]$ et $[VP(g_s)]$ contiennent $+1$ et -1. La formule de Lusztig s'applique aux faisceaux caractères associés à des éléments de $\mathcal{W}_{D(\chi)}$ et non pas aux représentations de ce groupe. La démonstration de [13] 7.1 qui se place dans ce cadre est très générale et elle montre que le seul point est le calcul de la constante notée z_2 en loc.cit. Cette constante est la somme des valeurs du caractère déterminé par χ sur les conjugués de s. Le calcul est plus compliqué qu'en loc.cite mais c'est le calcul des $\chi'([u],[\lambda])$ et $\chi''([u],[\lambda])$. Ensuite [13] 7.2 déduit le résultat cherché.

(ii) pour pouvoir utiliser (i), on décompose l'orbite de g_s sous $SO(2n+1, F)_\sharp$ en orbites sous $K_{n',n''}$ (le parahorique qui sert à la définition de $k_\chi(\underline{\rho})$). Ces orbites sont paramétrées par les éléments de $D(g_s)$; seuls comptent les orbites qui coupent $K_{n',n''}$ et pour cela il faut la relation:

$$\sum_{[\lambda]\in[VP(\chi)]} m'([\lambda])\ell_{[\lambda]} = n'.$$

Si cette relation n'est pas satisfaite, on remarque que les sommes intervenant sont vides et $loc_{g_s;\underline{m}}(\rho) = 0$. On n'a donc pas à se préoccuper de cette condition. Ensuite on calcule l'intégrale sous chaque $K_{n',n''}$-orbite en utilisant (i). Il y a clairement des mesures à prendre en compte; c'est fait en [8] 3.17, le $|W(d)|$ n'intervient pas pour nous car il a été pris en compte quand on travaille avec des représentations des groupes $\mathcal{W}$ et non les éléments de ces groupes et les constantes $(c(\gamma), c(\gamma)_\sharp$ de loc.cit) ont été mises dans la définition de loc_{g_s}.

4.2 Restriction des représentations et localisation des faisceaux caractères

Dans cette section, il s'agit de montrer que l'opération de restriction aux parahoriques des représentations commute à l'action de restriction des caractères auprès des éléments semi-simples elliptiques, même si l'on ne peut l'exprimer en ces termes tant que la conjecture 6.2 n'est pas démontrée. En plus comme on peut s'y attendre, vu la complexité des formules, il n'y a vraiment commutation que dans les cas favorables.

On fixe une donnée cuspidale, *cusp*, pour les faisceaux caractères, c'est-à-dire pour nous, 2 entiers impairs, I_+ et I_- et 2 entiers pairs P_+, P_- ainsi que 2 signes ζ_+, ζ_- avec la convention que si pour $\epsilon = \pm$, $P_\epsilon = 0$ alors $\zeta_\epsilon = (-1)^{(I_\epsilon-1)/2}$. Dans $\mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}]$, on ne considère que la partie relative à cette donnée cuspidale; on définit comme en 3.4 une autre donnée cuspidale $\tilde{cusp}$ simplement en changeant les signes, $\tilde{\zeta}_\pm := (-1)^{(I_\pm-1)/2}\zeta_\pm$ et on reprend la notation $\tilde{\chi}$ de loc.cite; on note $\tilde{}$ l'application évidente de la partie de $\mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}]$ relative à *cusp* dans son homologue relative à $\tilde{cusp}$ qui en terme de représentation de groupes de Weyl est tout simplement l'identité (mais on a changé le support cuspidal) composé avec la multiplication par le caractère $\tilde{\chi}$.

Soit g_s un élément semi-simple elliptique dont les valeurs propres sont des racines de l'unité d'ordre premier à p. On reprend la notation $[VP(g_s)]$ pour signifier l'ensemble des valeurs propres de g_s regroupées en paquets λ, λ' sont dans le même paquet s'il existe $a \in \mathbb{N}$ tel que $\lambda' = \lambda^{q^a}$; la multiplicité d'une valeur propre λ est notée $m([\lambda])$ car elle ne dépend que du paquet auquel λ appartient.

On note $D(g_s)$ l'ensemble des décompositions $\{(m'([\lambda]), m''([\lambda])) \in D(m([\lambda]))\}_{[\lambda]\in[VP(g_s)]}$; pour chaque élément $\underline{m} \in D(g_s)$, on a une classe d'association de parahorique $K_{n',n''}$ où $n' = \sum_{[\lambda]} m'([\lambda])$ et à l'intérieur de ce parahorique une classe de conjugaison d'éléments semi-simples de réduction semi-simple elliptique incluse dans la classe de conjugaison de g_s; on connait les valeurs propres des éléments dans cette classe de conjugaison. Pour $\underline{m} \in D(g_s)$, on reprend la notation $\mathcal{W}_{\underline{m}}$ de 4.1. On a défini ci-dessus des opérations de localisation de $\mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}]$ dans $\mathbb{C}[\hat{\mathcal{W}}_{\underline{m}}]$; en sommant sur tous les

éléments de $D(g_s)$, on définit donc une application de localisation de $\mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}]$ dans $\mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}]$ que l'on note loc_{g_s}.
On remarque que $\rho \circ \iota$ se définit aussi de $\mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}]$ dans lui-même; ce sont exactement les définitions de [8] 3.1, 3.2 et 3.9, 3.10. On les présente un peu différemment de façon similaire aux formules de 3.4. Précisément, on note $DD(g_s)$ l'ensemble des u-plets $\{m^{i,j}([\lambda]); i,j \in \{','' \}, [\lambda] \in [VP(g_s)]\}$. Pour $\underline{\underline{m}} \in DD(g_s)$, on définit $\hat{\mathcal{W}}_{\underline{\underline{m}}}$ de façon similaire à 3.4, la partie cuspidale possible étant simplement donnée comme en 4.1 (notée $|cusp|$). Et ici, $\rho \circ \iota$ est la somme sur tous les $\underline{\underline{m}} \in DD(g_s)$ de la restriction à $\mathcal{W}_{\underline{\underline{m}}}$ suivie de l'induction après avoir tordu par $(-1)^{\sum_{[\lambda] \notin \{\pm 1\}} m^{'','' }([\lambda])}(sgn_{CD})_{|W_{N_+^{'',''}} \times W_{N_-^{'',''}}}$, les notations et les inclusions entre les groupes étant celles décrites en 3.4.
Pour traiter tous les cas, on pose encore quelques définitions. Pour $\epsilon' \in \{\pm 1\}$, on note $X_{\epsilon'}$ l'endomorphisme de $\mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}]$ qui est la tensorisation par le caractère trivial sur tous les facteurs sauf $\hat{\mathcal{W}}_{m''(\epsilon')}$ où il vaut sgn_{CD}, suivie par l'inversion des facteurs relatifs $m'(\epsilon')$ et $m''(\epsilon')$ (à ce stade cette inversion est assez formelle mais elle a de l'importance quand ensuite on applique $\rho \circ \iota$). On rappelle la donnée cuspidale fixée et on reprend les notations ϵ_I et ϵ_P de 4.1. On note X_{cusp} l'endormophisme de $\mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}]$ défini par:

$$X_{cusp} := \begin{cases} 1 \text{ si } \epsilon_I = \epsilon_P \text{ et } \zeta_{\epsilon_P} = + \\ X_+ X_- \text{ si } \epsilon_I = \epsilon_P \text{ et } \zeta_{\epsilon_P} = - \\ X_{(-1)^{1+(I_+ + I_-)/2}} \text{ si } \epsilon_I \neq \epsilon_P \text{ et } \zeta_{\epsilon_P} = - \\ X_{(-1)^{(I_+ + I_-)/2}} \text{ si } \epsilon_I \neq \epsilon_P \text{ et } \zeta_{\epsilon_P} = +. \end{cases}$$

Lemme: *Fixons la donnée cuspidale comme ci-dessus. Alors le diagramme ci-dessous est commutatif pour tout élément g_s comme en 4.1:*

$$\begin{array}{ccc} \mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}] & \stackrel{\rho\circ\iota}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}] \\ \downarrow X_{cusp} \circ loc_{g_s} \circ \tilde{} & & \downarrow loc_{g_s} \\ \mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}] & \stackrel{\rho\circ\iota}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}] \end{array}$$

Pour démontrer ce lemme, on réintroduit le groupe auxiliaire $\mathcal{W}_{DD(\chi)}$ et son avatar $\mathcal{W}_{DD(g_s)}$ qui permettent de calculer $\rho \circ \iota$. De même, on réintroduit $\mathcal{W}_{D(\chi,g_s)}$. On espère que le lecteur voit une localisation de $\mathcal{W}_{DD(\chi)}$ vers $\mathcal{W}_{DD(g_s)}$ qui utilise le groupe $\mathcal{W}_{DD(\chi,g_s))}$ suggérée par les notations; ce groupe est construit comme tous les groupes de même type mais en utilisant des collections d'entiers $\nu^{i,j}([u],[\lambda])$ pour $i,j \in \{','' \}$, $[u] \in [VP(\chi)]$ et $[\lambda] \in [VP(g_s)]$ qui vérifient:

$$\sum_{i,j,[u]} \nu^{i,j}([u],[\lambda]) \ell_{[u]} (\ell_{[u]}, \ell_{[\lambda]})^{-1} = m_{g_s}([\lambda]) \qquad (*)$$

et une égalité de même type en échangeant les rôles de $[u]$ et de $[\lambda]$.

On écrit le diagramme:

$$\begin{array}{ccccc}
\mathbb{C}[\hat{\mathcal{W}}_{\underline{n}}] & \stackrel{res}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{DD_{\underline{n}}(\chi)}] & \stackrel{ind}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{D(\chi)}] \\
\downarrow res & & \downarrow res & & \downarrow res \\
\mathbb{C}[\hat{\mathcal{W}}_{D(\chi,g_s)}] & \stackrel{res}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{D_{\underline{n}}(D(\chi,g_s))}] & \stackrel{ind}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{D(\chi,g_s)}] \\
\downarrow ind & & \downarrow ind & & \downarrow ind \\
\mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}] & \stackrel{res}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{DD(g_s)}] & \stackrel{ind}{\to} & \mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}]
\end{array}$$

Expliquons ce qu'est l'objet central, le $\mathcal{W}_{D_{\underline{n}}(D(\chi,g_s))}$ est un sous-groupe de $\mathcal{W}_{D(D(\chi,g_s))}$; les collections $\nu^{i,j}([u],[\lambda])$ qui servent à le construire sont astreintes aux relations de (*) mais aussi à, pour tout $i \in \{','' \}$ et pour tout $[u] \in [VP(\chi)]$:

$$\sum_{j,[\lambda]} \nu^{ij}([u],[\lambda])\ell_{[\lambda]}(\ell_{[u]},\ell_{[\lambda]})^{-1} = n^i[u].$$

Un diagramme comme celui-ci est commutatif, le seul point est une formule à la Mackey du genre $res \circ ind = ind \circ res$; une telle formule nécessite des sommes: précisément considérons un groupe fini H avec des sous-groupes H', H''. Soit aussi une représentation de dimension finie, ρ' de H' et on calcule la restriction à H'' de l'induite de ρ' à H. Cette restriction est isomorphe à la somme sur γ dans un ensemble de représentants des doubles classes $H'\backslash H/H''$ des induites à H'' de la représentation ρ' transportée par γ et restreinte au groupe $\gamma^{-1}H'\gamma \cap H''$ (dans [8], ce raisonnement est utilisé en 3.19 ce qui suit (4)). Comme en loc.cit il y a la difficulté que les inclusions ne sont pas complètement évidentes. On applique cette formule 2 fois, pour le carré en bas à gauche, et pour le carré en haut à droite. Pour le carré en bas à gauche, on l'applique avec $H' = \mathcal{W}_{D(\chi,g_s)}$, $H = \mathcal{W}_{D(g_s)}$ et $H'' = \mathcal{W}_{DD(g_s)}$. Les doubles classes sont précisément indexées par $DD(\chi, g_s)$; en effet, pour $[\lambda] \neq [\pm 1], [\lambda] \in [VP(g_s)]$, on a à considérer les doubles classes:

$$\times_{[u]\in[VP(\chi)]}\mathfrak{S}_{\nu'([u],[\lambda]}\backslash\mathfrak{S}_{m'([\lambda])}/\mathfrak{S}_{m',' ([\lambda])} \times \mathfrak{S}_{m','' (\lambda)}$$

et un objet analogue où $'$ est remplacé par $''$, en tenant compte du fait que l'inclusion de $\mathfrak{S}_{\nu'([u],[\lambda])}$ dans $\mathfrak{S}_{m'([\lambda])}$ est décrite dans ce qui précède l'énoncé (il faut multiplier les cycles des permutations par $\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]})$). L'ensemble de ces doubles classes est bien indexé par les collections $(\nu^{i,j}([u],[\lambda]); i,j \in {}','', [u] \in [VP(\chi)]$ soumises aux conditions:

$$\forall [u] \in [VP(\chi)], \forall i \in \{','' \}, \nu^{i,'}([u],[\lambda]) + \nu^{i,''}([u],[\lambda]) = \nu^i([u],[\lambda]),$$

$$\forall i,j \in \{','' \}, \sum_{[u]\in[VP(\chi)]} \nu^{i,j}([u],[\lambda])\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]}) = \nu^{i,j}([\lambda]).$$

Et ensuite il reste à identifier $\times_{(\nu^{i,j}([u],[\lambda]);i,j\in ','',[u]\in[VP(\chi)]} \times_{i,j\in ','',[u]\in[VP(\chi)]} \mathfrak{S}_{\nu^{i,j}([u],[\lambda])}$ avec $\times_\gamma \gamma^{-1}H'\gamma \cap H''$ (avec les notations précédentes). Si $\lambda = \pm 1$, dans les objets ci-dessus, il faut remplacer certains groupes symétriques par des groupes de Weyl de type C; cela ne change rien.

On a un raisonnement du même type à faire pour le carré en haut à droite du diagramme.

Le point maintenant à considérer est que $\rho \circ \iota$ n'est pas exactement $ind \circ res$ écrit sur les lignes; il faut tordre les éléments de la colonne du milieu. Le même phénomène se produit pour loc_{g_s} et c'est ce qui motive l'introduction de l'endomorphisme X_{cusp}: pour $[u] \in [VP(\chi)]$ et $[\lambda] \in [VP(g_s)]$ regardons par quelle fonction il faut multiplier le facteur $\mathbb{C}[\otimes_{i,j\in\{','',\}}\hat{\mathcal{W}}_{\nu^{i,j}([u],[\lambda])}]$ avant d'induire pour arriver dans $\mathbb{C}[\hat{\mathcal{W}}_{D(g_s)}]$ pour que l'on obtienne le même résultat qu'en faisant le chemin, première ligne horizontale et dernière ligne verticale. Dans toute la discussion ci-dessous, on néglige, dans les flèches verticales tous les termes dépendant symétriquement de $\nu^{i,j}(.,.)$, symétriquement en i,j; c'est ce que l'on peut appeler de la torsion symétrique car elle ne gêne pas la commutation du diagramme.

Supposons d'abord que $\lambda \neq \pm 1$. Si $[u] \neq \pm 1$, la ligne horizontale multiplie par le signe $(-1)^{\nu^{'','',}([u],[\lambda])}$ et la ligne verticale n'introduit que de la torsion symétrique; si on fait le chemin de gauche, i.e. première ligne verticale et dernière ligne horizontale, c'est pareil et l'on n'a pas de problème de commutation.

Si $[u] = \pm 1$, la ligne horizontale du haut tensorise par $sgn_{CD} w_{\nu^{'','',}([u],[\lambda])}\tilde{\chi}$ si $\zeta_u = (-1)^{(I_u-1)/2}$; si cette égalité n'est pas vérifiée c'est une autre torsion mais il faut alors aussi tenir compte de la torsion dans la définition de k_χ et la combinaison des 2 ramènent à la formule $sgn_{CD} w_{\nu^{'','',}([u],[\lambda])}\tilde{\chi}$. Quand on fait l'autre chemin, on trouve la multiplication par $\tilde{\chi}$ qui est introduite par l'application $\tilde{}$ puir le signe $(-1)^{\nu^{'','',}([u],[\lambda])}$; ces 2 signes coïncident grâce à la définition de l'inclusion donnée en 4.1.

Reste le cas où $\lambda = \pm 1$; on note alors ϵ' au lieu de λ; on rappelle que les inclusions des groupes $\mathfrak{S}_m$ dans W_m (pour m un entier) considérées sont telles que $(-1)^m$ est aussi la valeur du sgn_{CD} de l'image par l'inclusion. On ne parlera donc que de sgn_{CD}. A priori il y a une différence quand $[u] \neq \pm 1$ et son contraire mais comme ci-dessus, cette différence s'efface quand on tient compte de la défintion de k_χ; on oublie aussi le signe $\tilde{\chi}$ qui est pris en compte par l'application $\tilde{}$. Ainsi la première ligne horizontale et la définition de k_χ introduisent la multiplication par $sgn_{CD}(w_{\nu^{'','',}([u],\epsilon')})$; la ligne verticale multiplie par le signe de la forme

$$\prod_{i=','',} (sgn_{CD} w_{\nu^{i,'}([u],\epsilon')})^{(I_{\epsilon_I}-1)/2} (sgn_{CD} w_{\nu^{i,''}([u],\epsilon')})^{(I_{\epsilon_P}-1)/2} \times \begin{cases} 1 \text{ si } \zeta_{\epsilon_P} = + \\ w_{\nu^{i,''}([u],\epsilon')} \text{ si } \zeta_{\epsilon_P} = - \end{cases}$$

et si $\epsilon' = -$ il faut encore multiplier par le caractère

$$\begin{cases} 1 \text{ si } \epsilon_I = + \\ \prod_{i=',''} sgn_{CD} w_{\nu^{i,'}([u],-)} \text{ si } \epsilon_I = - \end{cases} \times \begin{cases} 1 \text{ si } \epsilon_P = + \\ \\ \prod_{i=',''} sgn_{CD} w_{\nu^{i,''}([u],-)} \text{ si } \epsilon_P = - \end{cases} \times$$

$$\begin{cases} 1 \text{ si } (I_{\epsilon_I} + I_{\epsilon_P})/2 \text{ est impair} \\ \prod_{i=',''} sgn_{CD} w_{\nu^{i,\delta}([u],\epsilon'} \end{cases}$$

où $\delta =\,'$ si $(I_{\epsilon_I} - 1)/2$ est impair et $(I_{\epsilon_P} - 1)/2$ est pair et $\delta =\,''$ sinon.
Par l'autre chemin, l'application verticale introduit un caractère similaire à celui qui vient d'être écrit sauf que ce qui était $\nu^{i,'}$ devient $n^{',i}$ et ce qui était $\nu^{i,''}$ devient $n^{'',i}$. Il n'y a donc pas de difficulté quand ce qui intervient vraiment est un produit sur $(i,j) \in \{','' \}$. C'est le cas quand $\epsilon_I = \epsilon_P$ et $\zeta_{\epsilon_P} = +$. L'introduction du X_{cusp} est exactement fait pour résoudre les autres cas. Vérifions la commutativité du diagramme; on pose $\zeta = 0$ si $\zeta_{\epsilon_P} = +$ et 1 sinon et on pose aussi $\epsilon = 0$ si $\epsilon_I = \epsilon_P$ et 1 sinon et finalement, on pose $\Sigma := (I_{\epsilon_I} + I_{\epsilon_P})/2)$. On vérfie que X_{cusp} n'est autre que le produit $X_+^{1+\zeta+\Sigma} X_-^{1+\epsilon+\zeta+\Sigma}$. On étudie le chemin horizontal puis vertical; il s'introduit donc, d'abord le signe $sgn_{CD} w_{\nu'','' ([u],\epsilon')}$ puis par la dernière flèche verticale, un signe χ_{cusp}. Mais pour les problèmes de commutation, on peut multiplier ce signe par n'importe quel signe de la forme $\prod_{i,j \in \{','' \}} sgn_{CD} w_{\nu^{i,j}([u],\epsilon'_0)}$, où $\epsilon'_0 \in \{\pm 1\}$ comme expliqué ci-dessus. Ce qui veut dire qu'au lieu d'utiliser χ_{cusp} tel qu'il a été écrit, on peut utiliser

$$(\chi''_+ \chi''_-)^{1+\Sigma+\delta} (\chi''_-)^{\epsilon} = (\chi''_+)^{1+\delta+\Sigma} (\chi''_-)^{1+\delta+\epsilon+\Sigma}. \qquad (*)$$

Quand on fait la dernière flèche verticale, en terme de $w_{\nu^{i,j}}$ cela devient un produit sur tout $[u] \in [VP(\chi)]$ de

$$\prod_{i \in \{','' \}} sgn_{CD} w^{1+\delta+\Sigma}_{\nu^{i,''}([u],+)} \prod_{i \in \{','' \}} sgn_{CD} w^{1+\delta+\epsilon+\Sigma}_{\nu^{i,''}([u],-)}.$$

En incorporant le signe de la ligne horizontale, on trouve, un produit sur tout $[u]$

$$sgn_{CD} w^{1+\delta+\Sigma}_{\nu','' ([u],+)} sgn_{CD} w^{\delta+\Sigma}_{\nu'','' ([u],+)} sgn_{CD} w^{1+\delta+\epsilon+\Sigma}_{\nu','' ([u],-)} sgn_{CD} w^{\delta+\epsilon+\Sigma}_{\nu'','' ([u],-)}. \qquad (**)$$

On examine maitenant le chemin utilisant d'abord la première flèche verticale puis l'action de X_{cusp} et la dernière ligne horizontale. Dans X_{cusp} on commence par multiplier par un caractère qui est exactement le caractère (*) qui s'introduit par la flèche verticale après la simplification effectuée ci-dessus. Finalement, pour ce chemin, il suffit de regarder le caractère de la dernière ligne horizontale en tenant compte de l'inversion éventuelle. Or on a inversion entre $m'(+)$ et $m''(+)$ par hypothèse si $\delta + \Sigma$ est impaire et inversion entre $m'(-)$ et $m''(-)$ si $\delta + \epsilon + \Sigma$ est impaire. L'inversion entre $'$ et $''$ a pour effet que $\rho \circ \iota$ introduit le signe $sgn_{CD} w_{\nu','' ([u],\epsilon')}$ au lieu de $sgn_{CD} w_{\nu'','' ([u],\epsilon')}$. On trouve donc exactement le caractère (**). Cela termine la preuve.

5 Stabilité

5.1 Stabilité, définition

On reprend encore les définitions de [8]; soit G un groupe classique qui est donc le groupe des automorphismes d'une forme (ici orthogonale ou unitaire); on doit considérer simultanément 2 formes de ce groupe correspondant à 2 formes orthogonale ou unitaire, séparée dans le cas orthogonal par l'invariant de Hasse et dans le cas unitaire par la parité de la valuation du déterminant. On note ces 2 formes G_{iso} et G_{an} en imposant que G_{iso} est la forme quasidéployée et que G_{an} est l'autre groupe; dans tous les cas, G_{an} est une forme intérieure de G_{iso} mais éventuellement, on a même un isomorphisme $G_{an} \simeq G_{iso}$; ces 2 formes interviennent dans le calcul du centralisateur d'un élément semi-simple tel que fait dans 4.1 et les constructions dépendent de la forme orthogonale ou unitaire qui intervient et pas seulement de son groupe d'automorphismes, d'où la nécessité de garder la différence dans les notations. On sait définir la classe de conjugaison stable de tout élément fortement régulier de $G_\sharp$ pour $\sharp = iso$ ou an et on sait aussi définir une inclusion de l'ensemble des classes de conjugaison stable de G_{an} dans l'ensemble des classes de conjugaison stable dans G_{iso}. Soit $\phi = (\phi_{iso}, \phi_{an})$ une fonction dans $C_c^\infty(G_{iso}) \oplus C_c^\infty(G_{an})$; on dit qu'elle est stable si les intégrales orbitales de ϕ_{iso} et de ϕ_{an} sont constantes sur les classes de conjugaison stable et si les intégrales orbitales de ϕ_{iso} et de ϕ_{an} se correspondent pour l'inclusion des classes stables pour G_{an} dans les classes stables de G_{iso} et ϕ_{iso} a une intégrale nulle sur les classes stables de G_{iso} ne provenant pas de G_{an}); il a évidemment fallu fixer des mesures cohérentes. On dit que ϕ est semi-stable si ϕ_{iso} et ϕ_{an} sont stables mais si pour tout γ fortement régulier dans G_{an}, l'intégrale orbitale de ϕ_{an} sur la classe de conjugaison stable de γ est l'opposée de l'intégrale de ϕ_{iso} sur la classe de conjugaison stable dans G_{iso} correspondant à celle de γ. On dit que ϕ_{iso} est instable si pour tout γ fortement régulier l'intégrale sur la classe de conjugaison stable de γ est nulle; on définit de même ϕ_{an} instable et on dit que ϕ est instable si ϕ_{iso} et ϕ_{an} sont instables.

Soit $\underline{n} \in D(\chi)$ avec une donnée cuspidale *cusp*; on dit que

1. $\underline{n}, cusp$ est stable si $\epsilon_I = \epsilon_P$, $\zeta_+ = \zeta_- = +$, $|I_\epsilon - P_\epsilon| = 1$ pour $\epsilon = \pm$ et $n''[u] = 0$ pour tout $[u] \in [VP(\chi)]$.
2. On dit que $\underline{n}, cusp$ est semi-stable si $\epsilon_I = \epsilon_P$, $\zeta_+ = \zeta_- = -$, $|I_\epsilon - P_\epsilon| = 1$ pour $\epsilon = \pm$ et $n'([u]) = 0$ pour tout $[u] \in [VP(\chi)]$.
3. On dit que $\underline{n}, cusp$ est instable dans tous les autres cas.

Remarque. *Soit (ψ, ϵ) un paramètre discret de niveau zéro. Le couple $\underline{n}_{\psi,\epsilon}, cusp$ qui lui est associé avec la représentation de Springer-Lusztig est stable si et seulement si pour tout $[u] \in [VP(\chi)] \neq [\pm 1]$, $\epsilon_{[u]}$ est le caractère trivial et si $U_{[\pm 1],-} = \emptyset$ (avec les notations de 2.1); $\underline{n}_{\psi,\epsilon}, cusp$ est semi-stable si*

$\epsilon_{[u]} \equiv -1$ *pour tout* $[u] \in [VP(\chi)] \neq [\pm 1]$ *et si* $U_{[\pm 1],+} = \emptyset$. *Et* $\underline{n}_{\psi,\epsilon}, cusp$ *est instable dans tous les autres cas.*

On rappelle les formules données dans 3.3. Pour $[u] \neq [\pm 1]$, la traduction de $n'([u]) = 0$ ou $n''([u]) = 0$ en terme du caractère du groupe des composantes est claire .
Le cas de $u = \pm 1$ est plus compliqué. On regarde d'abord la partie cuspidale; à chaque orbite $U_{u,\epsilon'}$ munie de son caractère du groupe des composantes est associé un entier $k_{u,\epsilon'}$ par la représentation de Springer généralisée. On fixe $u, \epsilon' \in \{\pm 1\}$ et on montre d'abord l'équivalence:

$$k_{u,\epsilon'} = 0 \Leftrightarrow |I_u - P_u| = 1 \text{ et } \zeta_u = \epsilon'.$$

En effet, on vérifie d'après les formules données que $k_{u,\zeta_u} = (I_u + P_u - 1)/2$ et $k_{u,-\zeta_u} = (|I_u - P_u| - 1)/2$. Et l'équivalence est alors claire, en tenant compte du fait que I_u est impair alors que P_u est pair par hypothèse. Ensuite, c'est presque les définitions que $U_{u,\epsilon'} = 0$ est équivalente à $k_{u,\epsilon'} = 0$ et $n^\delta(u) = 0$ où $\delta ='$ si $\epsilon' = +$ et $''$ si $\epsilon' = -$.

5.2 Stabilité, théorème

On fixe une donnée cuspidale $cusp$ et $\underline{n} \in D(\chi)$.
Théorème. *soit* $\phi \in \mathbb{C}[\hat{\mathcal{W}}_{\underline{n},cusp}]$ *et soit* $\Phi := k_\chi\, \rho \circ \iota(\phi)$. *Alors* Φ *est stable si et seulement si* $\underline{n}, cusp$ *est stable; de même* Φ *est semi-stable si et seulement si* $\underline{n}, cusp$ *est semi-stable et* Φ *est instable si et seulement si* $\underline{n}, cusp$ *est instable.*
On suit la méthode de [8] 3.20 (qui démontre le même théorème dans le cas où $[VP(\chi)] = [1]$. On écrit $\Phi := (\Phi_{iso}, \Phi_{an})$. On fixe un élément semi-simple fortement régulier $g \in SO(2n+1,F)_{iso}$ et on étudie les intégrales orbitales de Φ_{iso} pour les éléments de la classe de conjugaison stable de g ainsi que celles de Φ_{an} pour la classe de conjugaison stable dans $SO(2n+1,F)_{an}$ quand elle existe. Il est clair que ces intégrales orbitales sont nulles si g n'est pas elliptique et compact. On écrit $g = g_s g_u$ comme en 4.1. L'ensemble $[VP(g_s)]$ est indépendant de g dans sa classe de conjugaison stable et quand g varie dans sa classe de conjugaison stable vue dans $SO(2n+1,F)_{iso} \cup SO(2n+1,F)_{an}$ g_s varie exactement dans sa classe de conjugaison stable dans $SO(2n+1,F)_{iso} \cup SO(2n+1,F)_{an}$. Les classes de conjugaison dans la classe de conjugaison stable de g_s sont paramétrées ([11] 1.7) par les collections $\{\sharp_{[\lambda]} \in \{+1,-1\} \simeq \{iso, an\}\}_{[\lambda]\in[VP(g_s)]}$ de telle sorte que si g_s correspond à la collection $\{\sharp_{[\lambda]}(g_s); \lambda \in [VP(g_s)]\}$ le commutant de g_s est isomorphe au produit $Aut((F'_{[\lambda]})^{m_{[\lambda]}}, <\,,\,>_{\sharp_{[\lambda]}(g_s)})$ où $F'_{[\lambda]}$ est une extension non ramifiée de degré 2 de $F_{[\lambda]}$ l'extension non ramifiée de F de degré $\ell_{[\lambda]}$ ($2\ell_{[\lambda]}$ est le cardinal de l'ensemble $[\lambda]$) et où $<\,,\,>_{\sharp_{[\lambda]}(g_s)}$ est une forme unitaire (pour l'extension $F'_{[\lambda]}$ de $F_{[\lambda]}$) dont le déterminant est de valuation paire ou impaire suivant que $\sharp_{[\lambda]}(g_s) = 1$ ou -1, si $[\lambda] \neq \pm 1$ et est une forme orthogonale si $[\lambda] = [\pm 1]$ (il n'y a alors pas d'extension de degré 2 à considérer); dans ce dernier cas, on a la même propriété que précédemment mais "parité de la valuation du

déterminant" étant remplacé par invariant de Hasse. Pour décrire les classes de conjugaison dans la classe de conjugaison stable de g, il faut encore décrire où varie g_u quand g_s est fixé. Comme on appliquera le début de la section 3 de [8] tel quel, nous n'avons pas besoin de faire cette description et on renvoie le lecteur à loc. cite.

Fixons maintenant $g = g_s g_u \in SO(2n+1,F)_{iso} \cup SO(2n+1,F)_{an}$ comme ci-dessus et calculons $I_g(\Phi)$. On note $\sharp(g)$ l'élément *iso* ou *an* tel que $g \in SO(2n+1,F)_{\sharp(g)}$. On a défini la fonction de Green $Q(loc_{g_s}(\Phi))$ en 4.1; c'est une fonction à support les éléments topologiquement unipotents sur le groupe

$$\times_{[\lambda]\in[VP(g_s)],[\lambda]\neq\pm 1} U(m(\lambda), F'_{[\lambda]}/F_{[\lambda]})_{\sharp_{[\lambda]}(g_s)} \times \\ SO(m([1]),F)_{\sharp_{[1]}(g_s)} \times O(m([-1]),F)_{\sharp_{[-1]}(g_s)},$$

où les notations sont celles de 4.1. L'élément g_u définit une classe de conjugaison d'éléments topologiquement unipotents dans ce groupe. Pour la suite on notera $(loc_{g_s}(\Phi))_{[\lambda]}$ la fonction sur le groupe indexé par $[\lambda]$ définie par $(loc_{g_s}(\Phi))$ quand les points dans les groupes indexés par $[\lambda'] \neq [\lambda]$ sont fixés.

Fixons une classe de conjugaison stable dans $SO(2n+1,F)$ d'éléments semi-simples réguliers $\mathcal{C}_{st}$; sans restreindre la généralité, on les suppose compacts (sinon les intégrales orbitales sont nulles). On note génériquement $\mathcal{C}$ les classes de conjugaison incluses dans $\mathcal{C}_{st}$; les classes de conjugaison le sont pour un groupe, c'est-à-dire qu'une telle classe $\mathcal{C}$ correspond à une valeur de $\sharp$ qui est notée $\sharp(\mathcal{C})$. Pour chaque classe $\mathcal{C}$, on fixe un élément $g(\mathcal{C}) \in \mathcal{C}$. On a à calculer pour Φ comme ci-dessus et pour $\sharp$ fixé $I^{st,\sharp}(\mathcal{C}_{st},\Phi) := \sum_{\mathcal{C}\in\mathcal{C}_{st},\sharp(\mathcal{C})=\sharp} I^{SO(2n+1,F)_\sharp}(g(\mathcal{C}),\Phi)$. On écrit chaque $g(\mathcal{C}) = g_s(\mathcal{C})g_u(\mathcal{C})$. On établit une relation d'équivalence entre les $\mathcal{C} \in \mathcal{C}_{st}$ par $\mathcal{C} \sim \mathcal{C}'$ si $g_s(\mathcal{C})$ est conjugué de $g_s(\mathcal{C}')$; on les supposera alors égaux. On écrira donc $g_s([\mathcal{C}])$ plutôt que $g_s(\mathcal{C})$.

Il existe une classe stable d'éléments semi-simples dont les valeurs propres sont des racines de l'unité d'ordre premier à p, $\mathcal{C}_{s,st}$ tel que $g_s([\mathcal{C}])$ soit un représentant des classes de conjugaison dans $\mathcal{C}_{s,st}$. On l'utilisera plus bas mais tout d'abord, on écrit:

$$I^{st,\sharp}(\mathcal{C}_{st},\Phi) := \sum_{[\mathcal{C}]\in\mathcal{C}_{st}/\sim,\sharp(\mathcal{C})=\sharp} \sum_{\mathcal{C}\in[\mathcal{C}]} I^{Cent_{SO(2n+1,F)_\sharp}(g_s([\mathcal{C}]))}(g_u(\mathcal{C}), loc_{g_s([\mathcal{C}])}\rho\circ\iota\,\phi).$$

On utilise 4.2 pour récrire, pour $[\mathcal{C}]$ fixée:

$$\sum_{\mathcal{C}\in[\mathcal{C}]} I^{Cent_{SO(2n+1,F)_\sharp}(g_s([\mathcal{C}]))}(g_u(\mathcal{C}), loc_{g_s([\mathcal{C}])}\rho\circ\iota\,\phi) \\ = \sum_{\mathcal{C}\in[\mathcal{C}]} I^{Cent_{SO(2n+1,F)_\sharp}(g_s([\mathcal{C}]))}(g_u(\mathcal{C}), Q\,\rho\circ\iota\, X_{cusp} \widetilde{loc_{g_s([\mathcal{C}])}}(\phi)).$$

On utilise tout de suite le fait que $[VP(g_s([\mathcal{C}]))]$ est indépendant de $[\mathcal{C}]$ dans $\mathcal{C}_{st}$; ce qui varie sont les invariants des formes $<\,,\,>_{[\lambda]}$, cf. 4.1 et ci-dessus. On remplace donc la notation $[VP(g_s)]$ par $[VP(\mathcal{C}_{st})]$. On décompose la somme ci-dessus en produit sur les $[\lambda] \in [VP(\mathcal{C}_{st})]$ et on constate qu'elle est nulle si l'une des composantes $Q\,\rho \circ \iota X_{cusp}(loc_{g_s([\mathcal{C}]}\tilde{(}\phi))_{[\lambda]}$ est instable. La condition d'instabilité pour ce genre de fonction (c'est à dire pour des fonctions dans l'image de $Q\,\rho\circ\iota$) est décrite dans [8] 3.4 pour les groupes unitaires (c'est-à-dire ici pour $[\lambda] \neq \pm 1$) et en [8] 3.12 pour les groupes orthogonaux (c'est-à-dire ici pour $[\lambda] = [\pm 1]$). Pour $[\lambda] \neq [\pm 1]$, on a instabilité si $m'([\lambda])m''([\lambda]) \neq 0$. Pour $\lambda \in \{\pm 1\}$, on a les données pour le support cuspidal des fonctions de Green qui sont $|c\tilde{u}sp|$ c'est-à-dire: $|r'_\epsilon| := (I_+ + \epsilon\delta I_-)/2$, où $\delta = +$ ou $-$ de façon à ce que $|r'_+|$ soit impair et $|r''_\epsilon| := |\big((-1)^{(I_+-1)/2}\zeta_+ P_+ + \epsilon(-1)^{(I_--1)/2}\zeta_- P_-\big)/2|$. Ainsi $\delta = (-1)^{1+(I_++I_-)/2}$ et $|r''_\epsilon| = |(P_+ + \epsilon\delta\zeta_+\zeta_- P_-)|$.

Pour que les intégrales ne soient pas nulles, il faut que $M'(\lambda) := m'(\lambda) - (r'_\lambda)^2$ soit un entier pair (≥ 0) et la même propriété pour $M''(\lambda) := m''(\lambda) - (r''_\lambda)^2$. Pour $\epsilon = \pm$, le terme correspondant à $\lambda = \epsilon$ est instable si et seulement si soit $||r'_\epsilon| - |r''_\epsilon|| > 1$ soit $r'_\epsilon r''_\epsilon M''(\epsilon)$ ou soit $M'(\epsilon)M''(\epsilon) = 0$. On remarque que

$$|r'_\epsilon| - |r''_\epsilon| = |I_+ + \epsilon\delta I_-| - |P_+ + \epsilon\delta\zeta_+\zeta_- P_-|. \qquad (*)$$

On a donc instabilité si l'une des conditions $I_+ + I_- - |P_+ + \zeta_+\zeta_- P_-| \in \{-2,0,2\}$, $|I_- - I_-| - |P_+ - \zeta_+\zeta_- P_-| \in \{-2,0,2\}$ n'est pas satisfaite. On remarque que $I_+ + I_- - |I_+ - I_-| \geq 2$ et que si $P_+P_- \neq 0$, $P_+ + P_- - |P_+ - P_-| \geq 4$ pour des questions de parité. Ainsi si $\zeta_+\zeta_- \neq +$, c'est-à-dire vaut - et si $P_+P_- \neq 0$ la différence entre $I_+ + I_- - |P_+ - P_-|$ et $|I_+ - I_-| - (P_+ + P_-)$ est au moins 6. On ne peut donc avoir les deux conditions satisfaites en même temps. Ainsi, on a instabilité si $P_+P_- \neq 0$ mais $\zeta_+\zeta_- = -$. Si $P_+P_- = 0$, en reprenant les notations, ϵ_I et ϵ_P de 4.1, on a donc $P_{-\epsilon_P} = 0$ et on vérifie que nécessairement $I_{-\epsilon_I} = 1$.

On a donc déjà démontré que l'on a instabilité sauf si soit $\zeta_+\zeta_- = +$ soit $P_{-\epsilon_P} = 0$ et $I_{-\epsilon_I} = 1$.

Récrivons les conditions (*) ci-dessus sous la forme plus simple $(I_{\epsilon_I} - P_{\epsilon_P}) \pm (I_{-\epsilon_I} - P_{-\epsilon_P}) \in \{-2,0,2\}$ et encore

$$(I_{\epsilon_I} - P_{\epsilon_P}) \in \{-1,1\}; (I_{-\epsilon_I} - P_{-\epsilon_P}) \in \{-1,1\}. \qquad (*)_{cusp}$$

On rappelle la convention que $\epsilon_I = \epsilon_P$ si $(I_+ - I_-)(P_+ - P_-) = 0$ et que $\epsilon_I = \epsilon_P = +$ si $I_+ = I_-$ et $P_+ = P_-$.

On a donc instabilité au moins s'il existe $[\lambda] \in [VP(\mathcal{C}_{st})]$ tel que $m'([\lambda])m''([\lambda]) \neq 0$ (en remplaçant m' par M' et m'' par M'' si $\lambda = \pm 1$ ou si $(*)_{cusp}$ n'est pas satisfaite ou encore s'il existe $\epsilon = \pm$ tel que $M'([\epsilon]) = 0$, $M''(\epsilon) \neq 0$ et $r'_\epsilon r''_\epsilon \neq 0$.

Par les références déjà données, on sait aussi quand ces sommes partielles ne dépendent que de l'invariant $\sharp_{[\lambda]}$ de $<\,,\,>_{[\lambda]}$ ou sont indépendantes du choix de $[\mathcal{C}]$ dans $\mathcal{C}_{g_s}$. Il faut simplement faire attention que $loc_{g_s[\mathcal{C}]}$ dépend de $[\mathcal{C}]$ dans $\mathcal{C}_{st}$ par ce qui est noté $c_{cusp}(g_s)$ en 4.1 et donc (en supprimant ce qui est encore indépendant) par le signe:

$$(\eta'_+\eta'_-)^{(I_{\epsilon_I}-1)/2}(\eta''_+\eta''_-)^{(I_{\epsilon_P}-1)/2}\times\begin{cases}1 \text{ si } \zeta_{\epsilon_P}=+\\ \eta''_+\eta''_- \text{ si } \zeta_{\epsilon_P}=-\end{cases}\times\begin{cases}1 \text{ si } \epsilon_I=\epsilon_P=+\\ \eta'_-\eta''_- \text{ si } \epsilon_I=\epsilon_P=-\\ \eta'_- \text{ si } \epsilon_I=-,\epsilon_P=+\\ \eta''_- \text{ si } \epsilon_I=+,\epsilon_P=-.\end{cases}$$

Il est temps d'utiliser les propriétés des formes $<,>_\epsilon$ pour $\epsilon=\pm$ et de leur "réduction" (rappelées en [8] 3.11): $\eta'_\epsilon\eta''_\epsilon$ est calculé par le discriminant de la forme $<,>_\epsilon$ et η'_ϵ comme η''_ϵ sont soit l'invariant de Hasse soit son opposé (le choix dépend du discriminant). Ainsi, dans la formule ci-dessus, si $\epsilon_I\neq\epsilon_P$, $c_{cusp}(g_s([\mathcal{C}]))$ dépend de l'invariant de Hasse de l'une des formes $<,>_\epsilon$ et pas de l'autre: en effet le premier terme dépend du produit des 2 invariants de Hasse, le deuxième terme dépend soit du produit soit vaut 1 et le troisième terme dépend de l'un des invariants de Hasse exactement. Par contre si $\epsilon_I=\epsilon_P$ alors $c_{cusp}(\mathcal{C})$ dépend du produit des invariants de Hasse quand $\zeta_{\epsilon_P}=-$ et est constant si $\zeta_{\epsilon_P}=+$.

Comme la seule chose qui est fixée pour les classes de conjugaison dans la classe de conjugaison stable de g_s à l'intérieur d'un groupe $SO(2n+1,F)_\sharp$ où $\sharp$ est fixé est le produit sur tous les $[\lambda]$ des invariants $\sharp_{[\lambda]}$ (qui sont les invariants de Hasse pour $\lambda=\pm1$), on voit aisément que l'on a encore instabilité s'il existe $[\lambda]\in[VP(\mathcal{C}_{st})]$ tel que le terme correspondant dépend de l'invariant $\sharp_{[\lambda]}$ de la forme $<,>_{[\lambda]}$ et qu'il existe $\lambda'\in[VP(\mathcal{C}_{st}]$ tel que le terme correspondant ne dépend pas de l'invariant de la forme $<,>_{[\lambda']}$. Et on aura stabilité si aucun des termes n'en dépend et semi-stabilité si tous les termes en dépendent. Ainsi la stabilité se produit quand $m''([\lambda])=0$ pour tout $[\lambda]\in[VP(\mathcal{C}_{st})]-\{-1,+1\}$ et si le produit des termes correspondant à $+1$ et -1 est aussi indépendant des invariants de Hasse. Comme on l'a vu ci-dessus, il faut distinguer le cas $\epsilon_I\neq\epsilon_P$ du cas où l'on a égalité. Supposons d'abord que $\epsilon_I=\epsilon_P$; dans ce cas, si $\zeta_{\epsilon_P}=+$, $c_{cusp}(\mathcal{C})$ est indépendant des invariants de Hasse, il faut donc aussi que les intégrales en soient indépendantes et donc que $M''(+1)=M''(-1)=0$.

Par contre si $\zeta_{\epsilon_P}=-$, toujours sous l'hypothèse $\epsilon_I=\epsilon_P$, il faut $|r'_\epsilon||r''_\epsilon|=0$ et $M'(\epsilon)=0$ pour $\epsilon=\pm$. Pour avoir stabilité, on a déjà vu qu'il faut $P_{-\epsilon_P}=0$ et $I_{-\epsilon_I}=1$. La condition $|r'_\epsilon||r''_\epsilon|=0$ écrite pour $\epsilon=\delta$, donne $(I_++I_-)|(P_+--P_-)|=0$. D'où $P_+=P_-=0$ et on retrouve aussi $I_{\epsilon_I}=1$ en utilisant $(*)_{cusp}$. D'où par convention $\zeta_+=\zeta_-=+$ ce qui contredit $\zeta_{\epsilon_P}=-$.

Terminons le cas de la stabilité quand $\epsilon_I=\epsilon_P$; la stabilité est alors équivalente à $m''([\lambda])=0$ pour tout $[\lambda]\neq\pm1$ et pour $\lambda=\pm1$, il faut $M''(\lambda)=0$, $\zeta_+=\zeta_-=+$ et $(*)_{cusp}$. On remarque que la condition $\epsilon_I=\epsilon_P$ couplée avec $(*)_{cusp}$ est équivalente à $|I_+-P_+|=1$ et $|I_--P_-|=1$. Il faut encore utiliser le fait que la localisation est non nulle si l'ensemble $D(\chi,g_s)$ n'est pas vide;

c'est-à-dire qu'il existe donc une collection $\nu'([u],[\lambda]), \nu''([u],[\lambda])$ satisfaisant à:

$$m'([\lambda]) = \sum_{[u]\in[VP(\chi)]} \nu'([u],[\lambda])\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]}),$$

$$m''([\lambda]) = \sum_{[u]\in[VP(\chi)]} \nu''([u],[\lambda])\ell_{[u]}/(\ell_{[u]},\ell_{[\lambda]}),$$

pour tout $[\lambda] \neq [\pm 1]$ et une formule analogue quand $\lambda \in \{\pm 1\}$. On a aussi, avec les mêmes notations, pour tout $[u] \in [VP(\chi)] \neq \pm 1$:

$$n'([u]) = \sum_{[\lambda]\in[VP([\mathcal{C}])]} \nu'([u],[\lambda])\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]}),$$

$$\nu''([u]) = \sum_{[\lambda]\in[VP([\mathcal{C}])]} \nu''([u],[\lambda])\ell_{[\lambda]}/(\ell_{[u]},\ell_{[\lambda]});$$

et une formule analogue pour $u \in \{\pm 1\}$. On en déduit que les conditions $n''([u]) = 0$ (si $u \neq \pm 1$) et $N''(\pm 1) = 0$ sont équivalentes à leurs analogues pour m', m'' et $[\lambda]$.
Pour la stabilité, reste à voir le cas où $\epsilon_I \neq \epsilon_P$. On a vu que $c_{cusp}(g_s(\mathcal{C}))$ dépend de l'un des invariants de Hasse; il faut donc que les intégrales dépendent elles aussi de l'un des invariants de Hasse exactement. Mais on doit donc avoir l'une des conditions $r'_\epsilon r''_\epsilon = 0$ qui nécessairement entraîne soit $I_+ = I_-$ soit $P_+ = P_-$. Et on a donc immédiatement une impossibilité avec les conventions sur ϵ_I et ϵ_P.
Cela termine la preuve en ce qui concerne la stabilité.
Pour la semi-stabilité: le raisonnement est du même type, il faut pour tout $\lambda \neq \pm 1$, $m'(\lambda) = 0$. Pour $\lambda = \pm 1$, on vérifie qu'il faut $\epsilon_I = \epsilon_P$ (c'est comme ci-dessus), puis $M''(\lambda) = 0$ et $\zeta_+ = \zeta_- = -$. Ensuite, on se rappelle des échanges induits par X_{cusp}; on doit échanger $\nu'([u],\epsilon')$ et $\nu''([u],\epsilon')$ pour tout $[u] \in [VP(\chi)]$ et tout $\epsilon' = \pm 1$ (cf. 4.1). On en déduit que la semi-stabilité est équivalente à ce que $n'([u]) = 0$ pour tout $[u] \neq \pm 1$ dans $[VP(\chi)]$, $n'(\pm 1) = 0$ ainsi que les conditions déjà écrites sur la partie cuspidale. Cela termine la preuve.

5.3 Traduction en termes de paramètres

La remarque ci-dessous vient d'idées de Lusztig avec des compléments pour les groupes non connexes de [13]. On considère l'ensemble des quadruplets d'entiers positifs ou nuls introduit dans 3.3, pour $u = \pm$ et $\epsilon' = \pm$, $k_{u,\epsilon'}$ et on pose I_ϵ l'entier impair du couple $(k_{u,+} + k_{u,-} + 1, |k_{u,+} - k_{u,-}|)$ et P_ϵ l'entier pair. On pose aussi ζ_u le signe de $k_{u,+} - k_{u,-}$ avec la convention que si ce nombre est nul $\zeta_u = (-1)^{k_{u,+}}$ ce qui est compatible avec la convention de 4.1 car $(I_u - 1)/2 = k_{u,+}$ dans ce cas.
Remarque. *Avec les notations précédentes, on a l'équivalence des conditions:*

$$\forall u \in \{\pm\}; |I_u - P_u| = 1 \text{ et } \zeta_u = + \Leftrightarrow \forall u \in \{\pm\}\, k_{u,-} = 0;$$

$$\forall u \in \{\pm\}; |I_u - P_u| = 1 \textit{ et } \zeta_u = - \Leftrightarrow \forall u \in \{\pm\}\, k_{u,+} = 0.$$

On a pour $u = \pm$, $|I_u - P_u| = 1 + k_{u,-\zeta_\epsilon}$ et la remarque en découle.
La remarque précédente motive la définition:
Définition. *Soit ψ, ϵ un paramètre discret de niveau 0. On reprend les notations de 2.1 et 3.3. On dit qu'il est stable si pour tout $[u] \in [VP(\chi)]$ l'orbite $U'_{[u]} = 0$; il est semi-stable si pour tout $[u] \in [VP(\chi)$ l'orbite $U''_{[u]} = 0$ et il est instable sinon.*
Corollaire. *L'espace des fonctions associées via la représentation de Springer-Lusztig et les faisceaux caractères à un paramètre discret de niveau zéro est formé de fonctions stables, semi-stables ou instables si et seulement si le paramètre est stable, semi-stable ou instable.*
Avec la remarque, cela résulte de 5.2 et de la définition ci-dessus.

6 Interprétation

6.1 Transformation de Fourier

Dans les conjectures de Langlands, les propriétés de stabilité ne s'expriment pas comme dans le corollaire ci-dessus. Ce ne sont pas certains paramètres qui sont stables mais au contraire ce sont des combinaisons linéaires. Les 2 façons d'exprimer le résultat se déduisent l'une de l'autre par une transformation style transformation de Fourier. Plus précisément, on fixe χ et on fixe des orbites $U_{[u]}$ pour tout $[u] \in [VP(\chi)]$ vérifiant les conditions de 2.1 et on note $\underline{U} := \{U_{[u]}; [u] \in [VP(\chi)]\}$. On note $\mathfrak{P}_{\chi,\underline{U}}$ l'espace vectoriel complexe de base l'ensemble des paramètres discrets de niveau 0 tels que la restriction de ψ à $I_F \times SL(2,\mathbb{C})$ (cf. 2.1) soit déterminée par χ et $\underline{U}$. Le principe est de définir un produit scalaire sur cet espace, $<,>$ et de définir la transformation $\mathcal{F}$, en posant:

$$\forall p \in \mathfrak{P}_{\chi,\underline{U}}; \mathcal{F}(p) := \sum_{p' \in \mathfrak{P}_{\chi,\underline{U}}} < p, p' > p'$$

Et si on a donné les bonnes définitions, on doit obtenir que l'application $\mathcal{F}$ transforme les paramètres stables au sens de 5.3 en des combinaisons linéaires stables à la Langlands; on renvoie aux paragraphes suivants pour expliquer cette dernière notion. Cela a déjà été fait dans [8] par. 6 dans ce qui est, en fait, le cas le plus difficile; en effet la difficulté vient de ce qu'il faut travailler avec des paramètres elliptiques et non pas des paramètres discrets et cette difficulté n'apparaît vraiment que quand $[VP(\chi)]$ contient $+1$ et/ou -1.
On dit qu'une orbite unipotente d'un groupe linéaire complexe est elliptique symplectique (resp. orthogonale) si ses blocs de Jordan sont tous pairs (resp. impairs) intervenant avec multiplicité au plus 2; pour le calcul du commutant, il n'y a pas de changement majeur au lieu d'avoir des groupes $O(1)$, on a un groupe $O(2)$ chaque fois qu'il y a multiplicité 2 (cf. 2.1). Disons qu'un paramètre ψ, ϵ est elliptique de niveau 0 si $\psi_{|W_F}$ est modérément ramifié comme en 2.1 et avec les notations de loc.cite, pour tout $[u] \in [VP(\chi)]$ et pour tout

$\zeta=\pm 1$, l'orbite $U_{[u],\zeta}$ est elliptique symplectique ou orthogonale (la différence entre symplectique et orthogonale étant comme en 2.1). On remarque qu'il y a une différence pour $[u]\neq[\pm 1]$ et pour $[u]=[\pm 1]$. En effet dans le premier cas, il revient au même de dire que $U_{[u]}$ est discrète (resp. elliptique) que de dire que chaque $U_{[u],\zeta}$, pour $\zeta=\pm 1$ est discrète ou elliptique et de plus, ce qui est le plus intéressant est que $U_{[u]}$ détermine chaque $U_{[u],\zeta}$. Ce n'est plus le cas si $[u]=[\pm 1]$; dans ce deuxième cas la multiplicité d'un bloc de Jordan dans $U_{[u]}$ a comme seule obligation d'être inférieure ou égale à 4 et le point le plus grave est que $U_{[u]}$ ne détermine pas chaque $U_{[u],\pm}$.
Ici $u=\pm 1$. On considère l'espace vectoriel complexe de base les quadruplets $(U_{[u],\pm},\epsilon_{[u],\pm})$ et on note $\mathbb{C}[Ell_{[u]}]$ son sous-espace vectoriel engendré par les éléments

$$\sum_{\epsilon_{[u],+},\epsilon_{[u],-}}\Big(\prod_{\substack{\alpha_+\in Jord(U_{[u],+});mult_+(\alpha_+)=2,\\ \alpha_-\in Jord(U_{[u],-});mult_-(\alpha_-)=2}}\epsilon_{[u],+}(\alpha_+)\epsilon_{[u],-}(\alpha_-)\Big)(U_{[u],\pm},\epsilon_{[u],\pm})$$

où $U_{[u],\pm}$ est fixé et la somme ne porte que sur les $\epsilon_{[u],\pm}$ fixés sur l'ensemble des $\alpha_\pm$ dans $Jord(U_{[u],\pm})$ dont la multiplicité $mult_{[u],\pm}(\alpha)$ est 1.
On a défini en [8] 6.11 une involution de $\mathbb{C}[Ell_{[u]}]$; il faut transporter le $\mathcal{F}$ du (i) de loc.cit par la bijection rea du (ii) de loc. cit.. C'était même une isométrie, mais on n'insiste pas la-dessus ici. C'est trop technique pour qu'on redonne la définition. On note $\mathcal{F}_{[u]}$ cette involution.
Considérons maintenant le cas de $[u]\neq\pm 1$; on pose ici $\mathbb{C}[Disc_{[u]}]$, l'espace vectoriel complexe de base les éléments $U_{[u]},\epsilon_{[u]}$ où est $U_{[u]}$ est discrète (c'est-à-dire que tous ses blocs de Jordan ont multiplicité 1). Pour définir l'application $\mathcal{F}_{[u]}$, on définit le produit scalaire:

$$<(U_{[u]},\epsilon_{[u]}),(U'_{[u]},\epsilon'_{[u]})>_{[u]}:=\begin{array}{l}0,\text{ si } U_{[u]}\neq U'_{[u]},\\ \sigma(U_{[u]})\sigma(\epsilon_{[u]})\sigma_{[u]}(\epsilon'_{[u]})\displaystyle\prod_{\substack{\alpha\in Jord(U_{[u]})\\ \epsilon_{[u]}(\alpha)=-1}}\epsilon'_{[u]}(\alpha)\text{ sinon,}\end{array}$$

où tous les σ sont des signes dépendant de l'objet dans la parenthèse; ici on n'a besoin que de $\sigma_{[u]}(\epsilon'_{[u]})$. On le prend égal à $\times_{\alpha\in Jord(U_{[u]}),\alpha\equiv 1[2]}\epsilon'_{[u]}(\alpha)$.
On pose alors $\mathcal{F}_{[u]}(U_{[u]},\epsilon_{[u]}):=\sum_{\epsilon'_{[u]}}<(U_{[u]},\epsilon_{[u]}),(U_{[u]},\epsilon'_{[u]})>_{[u]}(U_{[u]},\epsilon'_{[u]})$.
On remarque aisément que $\mathcal{F}^2_{[u]}=2^{|Jord(U_{[u]})|}\mathcal{F}_{[u]}$.
Pour homogénéiser, on définit aussi $\mathbb{C}[Ell_{[u]}]$; c'est l'espace vectoriel engendré par les éléments:

$$\sum_{\epsilon_{[u]}}\Big(\prod_{\alpha\in Jord(U_{[u]});mult(\alpha)=2}\epsilon_{[u]}(\alpha)\Big)(U_{[u]},\epsilon_{[u]})$$

où $U_{[u],\pm}$ est fixé et la somme ne porte que sur les $\epsilon_{[u],\pm}$ fixés sur l'ensemble des $\alpha_\pm$ dans $Jord(U_{[u],\pm})$ dont la multiplicité $mult_{[u],\pm}(\alpha)$ est 1. On étend $\mathcal{F}_{[u]}$

à $\mathbb{C}[Ell_{[u]}]$ en étendant la formule déjà donnée en précisant simplement que le produit ne porte que sur les α dont la multiplicité comme bloc de Jordan est 1.
On pose $\mathbb{C}[Ell_\chi] := \otimes_{[u]\in[VP(\chi)];[u]}\mathbb{C}[Ell_{[u]}]$. Et on définit $\mathcal{F} := \otimes_{[u]\in[VP(\chi)]}\mathcal{F}_{[u]}$

6.2 Restriction aux parahoriques des représentations

On a défini la représentation de Springer-Lusztig en 3.3; suivant [7] 6.5, on modifie légèrement cette définition dans le cas des groupes unitaires, on la note alors SpL_{ell}; cela induit alors un changement dans 3.3 que l'on marque par le changement de notation de SpL_{ell}. Soit (ψ,ϵ) un paramètre discret (ou elliptique) de niveau 0; on note ϵ_Z la restriction de ϵ à l'élément non trivial du centre de $Sp(2n,\mathbb{C})$. Pour $\sharp = iso$ ou an, on dit que $\epsilon_Z = \sharp$ si $\epsilon_Z = 1$ quand $\sharp = iso$ et -1 sinon. On note $|D|$ l'involution de [2] et [10] qui envoie une représentation irréductible sur une représentation irréductible.
Conjecture: *Il existe une bijection entre les paramètres discrets de niveau 0 ayant χ comme restriction à I_F et vérifiant $\epsilon_Z = \sharp$ et les séries discrètes de niveau zéro du groupe $SO(2n+1,F)_\sharp$ ayant χ comme élément semi-simple de leur support cuspidal: $(\psi,\epsilon)\mapsto \pi_{\psi,\epsilon}$ qui s'étend en une bijection, notée rea, entre $\mathbb{C}[Ell_\chi]$ et l'espace vectoriel complexe engendré par les représentations elliptiques au sens d'Arthur ayant χ comme élément semi-simple de leur support cuspidal avec la propriété: pour tout paramètre discret de niveau zéro, (ψ,ϵ), $k_\chi(\rho\circ\iota)\, SpL_{ell}(\psi,\epsilon)$ est un pseudo-coefficient de $|D| rea\,\mathcal{F}(\psi,\epsilon)$ (ou plus exactement a les mêmes intégrales orbitales qu'un pseudo-coefficient en les points semi-simples réguliers elliptiques de $SO(2n+1,F)_\sharp$).*
Cette conjecture est démontrée dans [12] pour les représentations de réduction unipotente.

Cette conjecture est motivée par ([7] 7.) bien qu'il y ait une différence entre les signes; cette différence doit traduire un signe provenant de la traduction en terme d'algèbre de Hecke de l'induction de Lusztig (ce qui devrait être l'objet de [2]). En [7], le signe qui s'introduit est $\prod_{u\neq\pm1}\prod_{\alpha\in Jord(U_{[u]});\alpha\equiv m([u])+1[2]}\epsilon'(\alpha)$ alors qu'ici on a fait le produit sur les blocs de Jordan impair.

6.3 Interprétation des résultats de stabilité

On a vu en [8] 4.6 (à la suite d'Arthur) que la stabilité des représentations elliptiques se lit sur les intégrales orbitales des pseudo-coefficients, en admettant la conjecture de 6.2 on peut décrire les combinaisons linéaires de représentations discrètes qui sont stables. Soit ψ,ϵ un paramètre discret de niveau 0 de restriction le caractère χ à I_F. On note encore ϵ_Z la restriction de ϵ au centre de $SO(2n+1,F)_\sharp$ (où $\sharp = iso$ ou an) et on dit que $\epsilon_Z = \sharp$ si $\epsilon_Z = +$ quand $\sharp = iso$ et $-$ quand $\sharp = an$.

Théorème. *Ici on admet la conjecture de 6.2. Soient* $\sharp = iso$ *ou* an *et* ψ *un paramètre discret de niveau 0.*
(i) La combinaison linéaire:

$$\sum_{\epsilon;\epsilon_Z=\sharp} \pi_{\psi,\epsilon}$$

est stable pour le groupe $SO(2n+1,F)_\sharp$. *De plus dans le transfert entre* $SO(2n+1,F)_{an}$ *et* $SO(2n+1,F)_{iso}$*, les combinaisons linéaires* $\epsilon_Z \sum_{\epsilon;\epsilon_Z=\sharp} \pi_{\psi,\epsilon}$ *se correspondent (ici* $\sharp$ *est vu comme un élément de* ± 1*).*
(ii) toute combinaison linéaire des représentations $\pi_{\psi,\epsilon}$ *pour* ϵ *variant avec* $\epsilon_Z = \sharp$ *est instable si elle n'est pas proportionnelle à la combinaison écrite en (i).*
On fixe $\sharp = iso$ ou an et on note $\mathbb{C}[Ell_\chi]_{stable,\sharp}$ le sous-espace de $\mathbb{C}[Ell_\chi]$ formé de l'image par rea des combinaisons linéaires stables de représentations elliptiques pour $SO(2n+1,F)_\sharp$. Et on note $\mathbb{C}[Ell_\chi]_{st,sst}$ le sous-espace de $\mathbb{C}[Ell_\chi]$ engendré par les paramètres elliptiques de niveau 0, stables ou semi-stables. On reprend les notations $U'([u])$ et $U''([u])$ de 3.3, l'espace ci-dessus est donc naturellement la somme directe des 2 sous-espaces, l'un (resp. l'autre) engendré par les paramètres (ψ,ϵ) tels que $U''_{[u]} = 0$ (resp. $U'_{[u]} = 0$) pour tout $[u] \in [VP(\chi)]$. Avec la conjecture et le théorème de 5.2 (complété par la remarque de 5.1), on sait que $rea \circ \mathcal{F}$ induit entre $\mathbb{C}[Ell_\chi]_{st,sst}$ et $\mathbb{C}[Ell_\chi]_{stable,iso} \oplus \mathbb{C}[Ell_\chi]_{stable,an}$. Il suffit donc de calculer l'image par $\mathcal{F}$ d'un paramètre (ψ,ϵ) elliptique de niveau 0 qui soit stable ou semi-stable et de reprojeter sur l'espace vectoriel engendré par les paramètres (ψ,ϵ) vérifiant $\epsilon_Z = \sharp$ quand $\sharp$ est fixé. Fixons donc $\zeta = \pm$ et calculons l'image $\mathcal{F}(\psi,\epsilon)$ en supposant que pour tout $[u] \in [VP(\chi)]$ l'orbite $U^\delta_{[u]} = 0$, où $\delta ='$ si $\zeta = +$ et $\delta ='' $ si $\zeta = -$. On espère que le lecteur comprendra une décomposition $\mathcal{F}(\psi,\epsilon) = \times_{[u]\in[VP(\chi)]}\mathcal{F}_{[u]}(\psi_{[u]},\epsilon_{[u]})$. Et on calcule $\mathcal{F}_{[u]}(\psi_{[u]},\epsilon_{[u]})$ en supposant d'abord que $[u] \neq [\pm 1]$; d'après la définition, on a

$$\mathcal{F}_{[u]}(\psi_{[u]},\epsilon_{[u]}) = \sigma(U_{[u]})\sigma_{[u]}(\epsilon_{[u]}) \sum_{\epsilon'_{[u]}} \sigma_{[u]}(\epsilon'_{[u]}) \Big(\prod_{\substack{\alpha \in Jord(U_{[u]}) \\ \epsilon_{[u]}(\alpha)=-1}} \epsilon'_{[u]}(\alpha) \Big) (\psi_{[u]},\epsilon'_{[u]}).$$

Or $\epsilon_{[u]}(\alpha) = \zeta$ par hypothèse pour tout α; la formule ci-dessus se simplifie donc si $\zeta = +$ en

$$\mathcal{F}_{[u]}(\psi_{[u]},\epsilon_{[u]}) = \sigma_{[u]}(\epsilon_{[u]}) \sum_{\epsilon'_{[u]}} \sigma_{[u]}(\epsilon'_{[u]})(\psi_{[u]},\epsilon'_{[u]}).$$

Par contre si $\zeta = -$, elle se simplifie en:

$$\mathcal{F}_{[u]}(\psi_{[u]},\epsilon_{[u]}) = \sigma_{[u]}(\epsilon_{[u]}) \sum_{\epsilon'_{[u]}} \sigma_{[u]}(\epsilon'_{[u]}) \Big(\prod_{\alpha \in Jord(U_{[u]})} \epsilon'_{[u]}(\alpha) \Big) (\psi_{[u]},\epsilon'_{[u]}).$$

Le cas de $[u] = [\pm 1]$ est exactement celui traité en [8] 6.12, et le résultat est analogue à ci-dessus.

En revenant au produit, on obtient dans le cas $\zeta = +$, avec σ un signe qui dépend de ψ, ϵ:

$$\mathcal{F}(\psi, \epsilon) = \sigma \sum_{\epsilon'} (\psi, \epsilon').$$

Dans le cas $\zeta = -$, dans la formule s'ajoute $\prod_{\alpha \in \cup_{[u]} Jord(U_{[u]})} \epsilon'(\alpha)$ qui n'est autre que ϵ'_Z. Quand on sépare les 2 morceaux, celui correspondant à $\sharp = iso$ et $\sharp = an$, ϵ'_Z est constant dans chaque morceau.
Pour pouvoir en déduire le résultat de stabilité cherchée, il faut utiliser la conjecture 6.2 qui permet de calculer les intégrales orbitales des caractères des représentations pour les éléments elliptiques. Mais il faut d'abord enlever $|D|$. Or pour toute représentation π irréductible et pour tout élément elliptique γ de G le caractère de π et de $|D|\pi$ coïncident en γ au signe $(-1)^{rg_G - rg_{P_{cusp,\pi}}}$, où $P_{cusp,\pi}$ est le sous-groupe parabolique de G minimal pour la propriété que $res_P(\pi)$ est non nulle. Pour π de la forme $\pi(\psi, \epsilon)$, $(-1)^{rg_G - rg_{P_{cusp,\pi}}} = \prod_{\alpha \in \cup_{[u]} Jord(U_{[u]}); \alpha \equiv 0[2]} \epsilon(\alpha)$. On a ainsi démontré que la distribution

$$\sum_{\epsilon'} \Big(\prod_{\alpha \in \cup_{[u]} Jord(U_{[u]})} \epsilon'(\alpha) \Big) \pi(\psi, \epsilon')$$

est stable et que pour $\sharp = iso$ ou an les seules distributions stables sont les sous-sommes de la somme ci-dessus où l'on ne somme que sur les ϵ' tels que $\epsilon'_Z = \sharp$. Cela donne le résultat annoncé.

References

[1] Arthur J. : *On local character relations,* Selecta Math. 2, 1996, pp. 501-579

[2] Aubert A.-M. : *Dualité dans le groupe de Grothendieck de la catégorie des représentations lisses de longueur finie d'un groupe réductif p-adique,*TAMS, 347, 1995, pp. 2179-2189 avec l'erratum publié dans TAMS, 348, 1996, pp. 4687-4690

[3] Aubert A.-M., Kutzko P., Morris L. : *Algèbres de Hecke des représentations de niveau zéro des groupes réductifs p-adiques. Applications,* version très préliminaire communiquée à l'auteur

[4] Courtès F. : *Distributions invariantes,* prépublication

[5] Lusztig G. : *Intersection cohomology complexes on a reductive group,* Inventiones 75, 1984, pp. 205-272

[6] Lusztig G. : *Green functions and character sheaves,* Ann of Math 131, 1990, pp. 355-408

[7] Mœglin C.: *Stabilité pour les représentations elliptiques de réduction unipotente: le cas des groupes unitaires,* prépublication Février 2003

[8] Mœglin C., Waldspurger J.-L.: *Paquets stables de représentations tempérées et de réduction unipotente pour $SO(2n+1)$* prépublication, Inst. Math. Jussieu 294, 2001.

[9] Moy A., Prasad G. : *Jacquet functors and unrefined minimal K-types,* Comment. Math. Helvetici 71, 1996, pp. 98-121

[10] Schneider M., Stuhler U. : *Representation theory and sheaves on the Bruhat-Tits building* Publ. Math. IHES 85, 1997, pp. 97-191

[11] Waldspurger J.-L. : *Intégrales orbitales unipotentes et endoscopie pour les groupes classiques non ramifiés* Astérisque 269, 2001

[12] Waldspurger J.-L. : *Représentations de réduction unipotente pour $SO(2n+1)$: quelques conséquences d'un article de Lusztig* prépublication Inst. Math. Jussieu 298, 2001, à paraître dans recueil en l'honneur de Shalika.

[13] Waldspurger J.-L. : *Une conjecture de Lusztig pour les groupes classiques* prépublication, 2002

Colette Moeglin
CNRS, Institut de mathématiques
de JUSSIEU
Paris
moeglin@math.jussieu.fr

Non-Hausdorff Groupoids, Proper Actions and K-Theory

Jean-Louis Tu

Received: March 29, 2004
Revised: June 21, 2004

Communicated by Joachim Cuntz

Abstract. Let G be a (not necessarily Hausdorff) locally compact groupoid. We introduce a notion of properness for G, which is invariant under Morita-equivalence. We show that any generalized morphism between two locally compact groupoids which satisfies some properness conditions induces a C^*-correspondence from $C_r^*(G_2)$ to $C_r^*(G_1)$, and thus two Morita equivalent groupoids have Morita-equivalent C^*-algebras.

2000 Mathematics Subject Classification: 22A22 (Primary); 46L05, 46L80, 54D35 (Secondary).
Keywords and Phrases: groupoid, C^*-algebra, K-theory.

Introduction

Very often, groupoids that appear in geometry, such as holonomy groupoids of foliations, groupoids of inverse semigroups [15, 6] and the indicial algebra of a manifold with corners [10] are not Hausdorff. It is thus necessary to extend various basic notions to this broader setting, such as proper action and Morita equivalence. We also show that a generalized morphism from G_2 to G_1 satisfying certain properness conditions induces an element of $KK(C_r^*(G_2), C_r^*(G_1))$.

In Section 2, we introduce the notion of proper groupoids and show that it is invariant under Morita-equivalence.
Section 3 is a technical part of the paper in which from every locally compact topological space X is canonically constructed a locally compact Hausdorff space $\mathcal{H}X$ in which X is (not continuously) embedded. When G is a groupoid (locally compact, with Haar system, such that $G^{(0)}$ is Hausdorff), the closure X' of $G^{(0)}$ in $\mathcal{H}G$ is endowed with a continuous action of G and plays an important technical rôle.
In Section 4 we review basic properties of locally compact groupoids with Haar system and technical tools that are used later.

In Section 5 we construct, using tools of Section 3, a canonical $C^*_r(G)$-Hilbert module $\mathcal{E}(G)$ for every (locally compact...) proper groupoid G. If $G^{(0)}/G$ is compact, then there exists a projection $p \in C^*_r(G)$ such that $\mathcal{E}(G)$ is isomorphic to $pC^*_r(G)$. The projection p is given by $p(g) = (c(s(g))c(r(g)))^{1/2}$, where $c\colon G^{(0)} \to \mathbb{R}_+$ is a "cutoff" function (Section 6). Contrary to the Hausdorff case, the function c is not continuous, but it is the restriction to $G^{(0)}$ of a continuous map $X' \to \mathbb{R}_+$ (see above for the definition of X').
In Section 7, we examine the question of naturality $G \mapsto C^*_r(G)$. Recall that if $f\colon X \to Y$ is a continuous map between two locally compact spaces, then f induces a map from $C_0(Y)$ to $C_0(X)$ if and only if f is proper. When G_1 and G_2 are groups, a morphism $f\colon G_1 \to G_2$ does not induce a map $C^*_r(G_2) \to C^*_r(G_1)$ (when $G_1 \subset G_2$ is an inclusion of discrete groups there is a map in the other direction). When $f\colon G_1 \to G_2$ is a groupoid morphism, we cannot expect to get more than a C^*-correspondence from $C^*_r(G_2)$ to $C^*_r(G_1)$ when f satisfies certain properness assumptions: this was done in the Hausdorff situation by Macho-Stadler and O'Uchi ([11, Theorem 2.1], see also [7, 13, 17]), but the formulation of their theorem is somewhat complicated. In this paper, as a corollary of Theorem 7.8, we get that (in the Hausdorff situation), if the restriction of f to $(G_1)^K_K$ is proper for each compact set $K \subset (G_1)^{(0)}$ then f induces a correspondence $\mathcal{E}_f$ from $C^*_r(G_2)$ to $C^*_r(G_1)$. In fact we construct a C^*-correspondence out of any groupoid generalized morphism ([5, 9]) which satisfies some properness conditions. As a corollary, if G_1 and G_2 are Morita equivalent then $C^*_r(G_1)$ and $C^*_r(G_2)$ are Morita-equivalent C^*-algebras.

Finally, let us add that our original motivation was to extend Baum, Connes and Higson's construction of the assembly map μ to non-Hausdorff groupoids; however, we couldn't prove μ to be an isomorphism in any non-trivial case.

1. Preliminaries

1.1. Groupoids. Throughout, we will assume that the reader is familiar with basic definitions about groupoids (see [16, 15]). If G is a groupoid, we denote by $G^{(0)}$ its set of units and by $r\colon G \to G^{(0)}$ and $s\colon G \to G^{(0)}$ its range and source maps respectively. We will use notations such as $G_x = s^{-1}(x)$, $G^y = r^{-1}(y)$, $G^y_x = G_x \cap G^y$. Recall that a topological groupoid is said to be *étale* if r (and s) are local homeomorphisms.

For all sets X, Y, T and all maps $f\colon X \to T$ and $g\colon Y \to T$, we denote by $X \times_{f,g} Y$, or by $X \times_T Y$ if there is no ambiguity, the set $\{(x,y) \in X \times Y \mid f(x) = g(y)\}$.

Recall that a (right) action of G on a set Z is given by

(a) a ("momentum") map $p\colon Z \to G^{(0)}$;
(b) a map $Z \times_{p,r} G \to Z$, denoted by $(z,g) \mapsto zg$

with the following properties:

(i) $p(zg) = s(g)$ for all $(z,g) \in Z \times_{p,r} G$;
(ii) $z(gh) = (zg)h$ whenever $p(z) = r(g)$ and $s(g) = r(h)$;

(iii) $zp(z) = z$ for all $z \in Z$.

Then the crossed-product $Z \rtimes G$ is the subgroupoid of $(Z \times Z) \times G$ consisting of elements (z, z', g) such that $z' = zg$. Since the map $Z \rtimes G \to Z \times G$ given by $(z, z', g) \mapsto (z, g)$ is injective, the groupoid $Z \rtimes G$ can also be considered as a subspace of $Z \times G$, and this is what we will do most of the time.

1.2. Locally compact spaces. A topological space X is said to be quasi-compact if every open cover of X admits a finite sub-cover. A space is compact if it is quasi-compact and Hausdorff. Let us recall a few basic facts about locally compact spaces.

Definition 1.1. *A topological space X is said to be locally compact if every point $x \in X$ has a compact neighborhood.*

In particular, X is locally Hausdorff, thus every singleton subset of X is closed. Moreover, the diagonal in $X \times X$ is locally closed.

Proposition 1.2. *Let X be a locally compact space. Then every locally closed subspace of X is locally compact.*

Recall that $A \subset X$ is locally closed if for every $a \in A$, there exists a neighborhood V of a in X such that $V \cap A$ is closed in V. Then A is locally closed if and only if it is of the form $U \cap F$, with U open and F closed.

Proposition 1.3. *Let X be a locally compact space. The following are equivalent:*

(i) *there exists a sequence (K_n) of compact subspaces such that $X = \cup_{n \in \mathbb{N}} K_n$;*
(ii) *there exists a sequence (K_n) of quasi-compact subspaces such that $X = \cup_{n \in \mathbb{N}} K_n$;*
(iii) *there exists a sequence (K_n) of quasi-compact subspaces such that $X = \cup_{n \in \mathbb{N}} K_n$ and $K_n \subset \mathring{K}_{n+1}$ for all $n \in \mathbb{N}$.*

Such a space will be called σ-compact.

Proof. (i) $\Longrightarrow$ (ii) is obvious. The implications (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i) follow easily from the fact that for every quasi-compact subspace K, there exists a finite family $(K_i)_{i \in I}$ of compact sets such that $K \subset \cup_{i \in I} \mathring{K}_i$. □

1.3. Proper maps.

Proposition 1.4. [2, Théorème I.10.2.1] *Let X and Y be two topological spaces, and $f\colon X \to Y$ a continuous map. The following are equivalent:*

(i) *For every topological space Z, $f \times \mathrm{Id}_Z\colon X \times Z \to Y \times Z$ is closed;*
(ii) *f is closed and for every $y \in Y$, $f^{-1}(y)$ is quasi-compact.*

A map which satisfies the equivalent properties of Proposition 1.4 is said to be *proper*.

Proposition 1.5. [2, Proposition I.10.2.6] *Let X and Y be two topological spaces and let $f\colon X \to Y$ be a proper map. Then for every quasi-compact subspace K of Y, $f^{-1}(K)$ is quasi-compact.*

Proposition 1.6. *Let X and Y be two topological spaces and let $f\colon X \to Y$ be a continuous map. Suppose Y is locally compact, then the following are equivalent:*

(i) *f is proper;*
(ii) *for every quasi-compact subspace K of Y, $f^{-1}(K)$ is quasi-compact;*
(iii) *for every compact subspace K of Y, $f^{-1}(K)$ is quasi-compact;*
(iv) *for every $y \in Y$, there exists a compact neighborhood K_y of y such that $f^{-1}(K_y)$ is quasi-compact.*

Proof. (i) $\Longrightarrow$ (ii) follows from Proposition 1.5. (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (iv) are obvious. Let us show (iv) $\Longrightarrow$ (i).
Since $f^{-1}(y)$ is closed, it is clear that $f^{-1}(y)$ is quasi-compact for all $y \in Y$. It remains to prove that for every closed subspace $F \subset X$, $f(F)$ is closed. Let $y \in \overline{f(F)}$. Let $A = f^{-1}(K_y)$. Then $A \cap F$ is quasi-compact, so $f(A \cap F)$ is quasi-compact. As $f(A \cap F) \subset K_y$, it is closed in K_y, i.e. $K_y \cap \overline{f(A \cap F)} = K_y \cap f(A \cap F)$. We thus have $y \in K_y \cap \overline{f(A \cap F)} = K_y \cap f(A \cap F) \subset f(F)$. It follows that $f(F)$ is closed. □

2. Proper groupoids and proper actions

2.1. Locally compact groupoids.

Definition 2.1. *A topological groupoid G is said to be locally compact (resp. σ-compact) if it is locally compact (resp. σ-compact) as a topological space.*

Remark 2.2. *The definition of a locally compact groupoid in* [15] *corresponds to our definition of a locally compact, σ-compact groupoid with Haar system whose unit space is Hausdorff, thanks to Propositions 2.5 and 2.8.*

Example 2.3. *Let Γ be a discrete group, H a closed normal subgroup and let G be the bundle of groups over $[0, 1]$ such that $G_0 = \Gamma$ and $G_t = \Gamma/H$ for all $t > 0$. We endow G with the quotient topology of $([0, 1] \times \Gamma) / ((0, 1] \times H)$. Then G is a non-Hausdorff locally compact groupoid such that $(t, \bar{\gamma})$ converges to $(0, \gamma h)$ as $t \to 0$, for all $\gamma \in \Gamma$ and $h \in H$.*

Example 2.4. *Let Γ be a discrete group acting on a locally compact Hausdorff space X, and let $G = (X \times \Gamma)/\sim$, where (x, γ) and (x, γ') are identified if their germs are equal, i.e. there exists a neighborhood V of x such that $y\gamma = y\gamma'$ for all $y \in V$. Then G is locally compact, since the open sets $V_\gamma = \{[(x, \gamma)]|\ x \in X\}$ are homeomorphic to X and cover G.*
Suppose that X is a manifold, M is a manifold such that $\pi_1(M) = \Gamma$, $\tilde{M}$ is the universal cover of M and $V = (X \times \tilde{M})/\Gamma$, then V is foliated by $\{[x, \tilde{m}]|\ \tilde{m} \in \tilde{M}\}$ and G is the restriction to a transversal of the holonomy groupoid of the above foliation.

Proposition 2.5. *If G is a locally compact groupoid, then $G^{(0)}$ is locally closed in G, hence locally compact. If furthermore G is σ-compact, then $G^{(0)}$ is σ-compact.*

Proof. Let Δ be the diagonal in $G \times G$. Since G is locally Hausdorff, Δ is locally closed. Then $G^{(0)} = (\mathrm{Id}, r)^{-1}(\Delta)$ is locally closed in G.
Suppose that $G = \cup_{n\in\mathbb{N}} K_n$ with K_n quasi-compact, then $s(K_n)$ is quasi-compact and $G^{(0)} = \cup_{n\in\mathbb{N}} s(K_n)$. □

Proposition 2.6. *Let Z a locally compact space and G be a locally compact groupoid acting on Z. Then the crossed-product $Z \rtimes G$ is locally compact.*

Proof. Let $p\colon Z \to G^{(0)}$ be the momentum map of the action of G. From Proposition 2.5, the diagonal $\Delta \subset G^{(0)} \times G^{(0)}$ is locally closed in $G^{(0)} \times G^{(0)}$, hence $Z \rtimes G = (p, r)^{-1}(\Delta)$ is locally closed in $Z \times G$. □

Let T be a space. Recall that there is a groupoid $T \times T$ with unit space T, and product $(x, y)(y, z) = (x, z)$.
Let G be a groupoid and T be a space. Let $f\colon T \to G^{(0)}$, and let $G[T] = \{(t', t, g) \in (T\times T)\times G|\ g \in G_{f(t)}^{f(t')}\}$. Then $G[T]$ is a subgroupoid of $(T\times T)\times G$.

Proposition 2.7. *Let G be a topological groupoid with $G^{(0)}$ locally Hausdorff, T a topological space and $f\colon T \to G^{(0)}$ a continuous map. Then $G[T]$ is a locally closed subgroupoid of $(T \times T) \times G$. In particular, if T and G are locally compact, then $G[T]$ is locally compact.*

Proof. Let $F \subset T\times G^{(0)}$ be the graph of f. Then $F = (f\times\mathrm{Id})^{-1}(\Delta)$, where Δ is the diagonal in $G^{(0)} \times G^{(0)}$, thus it is locally closed. Let $\rho\colon (t', t, g) \mapsto (t', r(g))$ and $\sigma\colon (t', t, g) \mapsto (t, s(g))$ be the range and source maps of $(T \times T) \times G$, then $G[T] = (\rho, \sigma)^{-1}(F \times F)$ is locally closed. □

Proposition 2.8. *Let G be a locally compact groupoid such that $G^{(0)}$ is Hausdorff. Then for every $x \in G^{(0)}$, G_x is Hausdorff.*

Proof. Let $Z = \{(g, h) \in G_x \times G_x|\ r(g) = r(h)\}$. Let $\varphi\colon Z \to G$ defined by $\varphi(g, h) = g^{-1}h$. Since $\{x\}$ is closed in G, $\varphi^{-1}(x)$ is closed in Z, and since $G^{(0)}$ is Hausdorff, Z is closed in $G_x \times G_x$. It follows that $\varphi^{-1}(x)$, which is the diagonal of $G_x \times G_x$, is closed in $G_x \times G_x$. □

2.2. Proper groupoids.

Definition 2.9. *A topological groupoid G is said to be proper if $(r, s)\colon G \to G^{(0)} \times G^{(0)}$ is proper.*

Proposition 2.10. *Let G be a topological groupoid such that $G^{(0)}$ is locally compact. Consider the following assertions:*

(i) *G is proper;*
(ii) *(r, s) is closed and for every $x \in G^{(0)}$, G_x^x is quasi-compact;*
(iii) *for all quasi-compact subspaces K and L of $G^{(0)}$, G_K^L is quasi-compact;*
(iii)' *for all compact subspaces K and L of $G^{(0)}$, G_K^L is quasi-compact;*

(iv) *for every quasi-compact subspace K of $G^{(0)}$, G_K^K is quasi-compact;*

(v) *$\forall x,\ y \in G^{(0)}$, $\exists K_x$, L_y compact neighborhoods of x and y such that $G_{K_x}^{L_y}$ is quasi-compact.*

Then (i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iii) $\Longleftrightarrow$ (iii)' $\Longleftrightarrow$ (v) $\Longrightarrow$ (iv). If $G^{(0)}$ is Hausdorff, then (i)–(v) are equivalent.

Proof. (i) $\Longleftrightarrow$ (ii) follows from Proposition 1.4, and from the fact that G_x^x is homeomorphic to G_x^y if $G_x^y \neq \emptyset$. (i) $\Longrightarrow$ (iii) and (v) $\Longrightarrow$ (i) follow Proposition 1.6 and the formula $G_K^L = (r,s)^{-1}(L \times K)$. (iii) $\Longrightarrow$ (iii)' $\Longrightarrow$ (v) and (iii) $\Longrightarrow$ (iv) are obvious. If $G^{(0)}$ is Hausdorff, then (iv) $\Longrightarrow$ (v) is obvious. $\square$

Note that if $G = G^{(0)}$ is a non-Hausdorff topological space, then G is not proper (since (r,s) is not closed), but satisfies property (iv).

PROPOSITION 2.11. *Let G be a topological groupoid. If $r\colon G \to G^{(0)}$ is open then the canonical mapping $\pi\colon G^{(0)} \to G^{(0)}/G$ is open.*

Proof. Let $V \subset G^{(0)}$ be an open subspace. If r is open, then $r(s^{-1}(V)) = \pi^{-1}(\pi(V))$ is open. Therefore, $\pi(V)$ is open. $\square$

PROPOSITION 2.12. *Let G be a topological groupoid such that $G^{(0)}$ is locally compact and $r\colon G \to G^{(0)}$ is open. Suppose that $(r,s)(G)$ is locally closed in $G^{(0)} \times G^{(0)}$, then $G^{(0)}/G$ is locally compact. Furthermore,*

(a) *if $G^{(0)}$ is σ-compact, then $G^{(0)}/G$ is σ-compact;*

(b) *if $(r,s)(G)$ is closed (for instance if G is proper), then $G^{(0)}/G$ is Hausdorff.*

Proof. Let $R = (r,s)(G)$. Let $\pi\colon G^{(0)} \to G^{(0)}/G$ be the canonical mapping. By Proposition 2.11, π is open, therefore $G^{(0)}/G$ is locally quasi-compact. Let us show that it is locally Hausdorff. Let V be an open subspace of $G^{(0)}$ such that $(V \times V) \cap R$ is closed in $V \times V$. Let Δ be the diagonal in $\pi(V) \times \pi(V)$. Then $(\pi\times\pi)^{-1}(\Delta) = (V\times V)\cap R$ is closed in $V\times V$. Since $\pi\times\pi\colon V\times V \to \pi(V)\times\pi(V)$ is continuous open surjective, it follows that Δ is closed in $\pi(V) \times \pi(V)$, hence $\pi(V)$ is Hausdorff. This completes the proof that $G^{(0)}/G$ is locally compact and of assertion (b).
Assertion (a) follows from the fact that for every $x \in G^{(0)}$ and every compact neighborhood K of x, $\pi(K)$ is a quasi-compact neighborhood of $\pi(x)$. $\square$

2.3. PROPER ACTIONS.

DEFINITION 2.13. *Let G be a topological groupoid. Let Z be a topological space endowed with an action of G. Then the action is said to be proper if $Z \rtimes G$ is a proper groupoid. (We will also say that Z is a proper G-space.)*

A subspace A of a topological space X is said to be relatively compact (resp. relatively quasi-compact) if it is included in a compact (resp. quasi-compact) subspace of X. This does not imply that $\overline{A}$ is compact (resp. quasi-compact).

Proposition 2.14. *Let G be a topological groupoid. Let Z be a topological space endowed with an action of G. Consider the following assertions:*

(i) *G acts properly on Z;*

(ii) *$(r,s)\colon Z \rtimes G \to Z \times Z$ is closed and $\forall z \in Z$, the stabilizer of z is quasi-compact;*

(iii) *for all quasi-compact subspaces K and L of Z, $\{g \in G|\ Lg \cap K \neq \emptyset\}$ is quasi-compact;*

(iii)' *for all compact subspaces K and L of Z, $\{g \in G|\ Lg \cap K \neq \emptyset\}$ is quasi-compact;*

(iv) *for every quasi-compact subspace K of Z, $\{g \in G|\ Kg \cap K \neq \emptyset\}$ is quasi-compact;*

(v) *there exists a family $(A_i)_{i\in I}$ of subspaces of Z such that $Z = \cup_{i\in I} \mathring{A}_i$ and $\{g \in G|\ A_i g \cap A_j \neq \emptyset\}$ is relatively quasi-compact for all $i, j \in I$.*

Then (i) $\iff$ (ii) $\implies$ (iii) $\implies$ (iii)' and (iii) $\implies$ (iv). If Z is locally compact, then (iii)' $\implies$ (v) and (iv) $\implies$ (v). If $G^{(0)}$ is Hausdorff and Z is locally compact Hausdorff, then (i)–(v) are equivalent.

Proof. (i) $\iff$ (ii) follows from Proposition 2.10[(i) $\iff$ (ii)]. Implication (i) $\implies$ (iii) follows from the fact that if $(Z \rtimes G)_K^L$ is quasi-compact, then its image by the second projection $Z \rtimes G \to G$ is quasi-compact. (iii) $\implies$ (iii)' and (iii) $\implies$ (iv) are obvious.

Suppose that Z is locally compact. Take $A_i \subset Z$ compact such that $Z = \cup_{i\in I} \mathring{A}_i$. If (iii)' is true, then $\{g \in G|\ A_i g \cap A_j \neq \emptyset\}$ is quasi-compact, hence (v). If (iv) is true, then $\{g \in G|\ A_i g \cap A_j \neq \emptyset\}$ is a subset of the quasi-compact set $\{g \in G|\ Kg \cap K \neq \emptyset\}$, where $K = A_i \cup A_j$, hence (v).

Suppose that Z is locally compact Hausdorff and that $G^{(0)}$ is Hausdorff. Let us show (v) $\implies$ (ii). Let C_{ij} be a quasi-compact set such that $\{g \in G|\ A_i g \cap A_j \neq \emptyset\} \subset C_{ij}$.

Let $z \in Z$. Choose $i \in I$ such that $z \in A_i$. Since Z and $G^{(0)}$ are Hausdorff, $\operatorname{stab}(z)$ is a closed subspace of C_{ii}, therefore it is quasi-compact.

It remains to prove that the map $\Phi\colon Z \times_{G^{(0)}} G \to Z \times Z$ given by $\Phi(z,g) = (z, zg)$ is closed. Let $F \subset Z \times_{G^{(0)}} G$ be a closed subspace, and $(z,z') \in \overline{\Phi(F)}$. Choose i and j such that $z \in \mathring{A}_i$ and $z' \in \mathring{A}_j$. Then $(z,z') \in \overline{\Phi(F) \cap (A_i \times A_j)} \subset \overline{\Phi(F \cap (A_i \times_{G^{(0)}} C_{ij}))} \subset \overline{\Phi(F \cap (Z \times_{G^{(0)}} C_{ij}))}$. There exists a net $(z_\lambda, g_\lambda) \in F \cap (Z \times_{G^{(0)}} C_{ij})$ such that (z,z') is a limit point of $(z_\lambda, z_\lambda g_\lambda)$. Since C_{ij} is quasi-compact, after passing to a universal subnet we may assume that g_λ converges to an element $g \in C_{ij}$. Since $G^{(0)}$ is Hausdorff, $F \cap (Z \times_{G^{(0)}} C_{ij})$ is closed in $Z \times C_{ij}$, so (z,g) is an element of $F \cap (Z \times_{G^{(0)}} C_{ij})$. Using the fact that Z is Hausdorff and Φ is continuous, we obtain $(z,z') = \Phi(z,g) \in \Phi(F)$. $\square$

Remark 2.15. *It is possible to define a notion of slice-proper action which implies properness in the above sense. The two notions are equivalent in many cases* [1, 3].

Proposition 2.16. *Let G be a locally compact groupoid. Then G acts properly on itself if and only if $G^{(0)}$ is Hausdorff. In particular, a locally compact space is proper if and only if it is Hausdorff.*

Proof. It is clear from Proposition 2.10(ii) that G acts properly on itself if and only if the product $\varphi\colon G^{(2)} \to G \times G$ is closed. Since φ factors through the homeomorphism $G^{(2)} \to G \times_{r,r} G$, $(g,h) \mapsto (g, gh)$, G acts properly on itself if and only if $G \times_{r,r} G$ is a closed subset of $G \times G$.
If $G^{(0)}$ is Hausdorff, then clearly $G \times_{r,r} G$ is closed in $G \times G$. Conversely, if $G^{(0)}$ is not Hausdorff, then there exists $(x,y) \in G^{(0)} \times G^{(0)}$ such that $x \neq y$ and (x,y) is in the closure of the diagonal of $G^{(0)} \times G^{(0)}$. It follows that (x,y) is in the closure of $G \times_{r,r} G$, but $(x,y) \notin G \times_{r,r} G$, therefore $G \times_{r,r} G$ is not closed. □

2.4. Permanence properties.

Proposition 2.17. *If G_1 and G_2 are proper topological groupoids, then $G_1 \times G_2$ is proper.*

Proof. Follows from the fact that the product of two proper maps is proper [2, Corollaire I.10.2.3]. □

Proposition 2.18. *Let G_1 and G_2 be two topological groupoids such that $G_1^{(0)}$ is Hausdorff and G_2 is proper. Suppose that $f\colon G_1 \to G_2$ is a proper morphism. Then G_1 is proper.*

Proof. Denote by r_i and s_i the range and source maps of G_i $(i = 1,2)$. Let $\bar{f}$ be the map $G_1^{(0)} \times G_1^{(0)} \to G_2^{(0)} \times G_2^{(0)}$ induced from f. Since $\bar{f} \circ (r_1, s_1) = (r_2, s_2) \circ f$ is proper and $G_1^{(0)}$ is Hausdorff, it follows from [2, Proposition I.10.1.5] that (r_1, s_1) is proper. □

Proposition 2.19. *Let G_1 and G_2 be two topological groupoids such that G_1 is proper. Suppose that $f\colon G_1 \to G_2$ is a surjective morphism such that the induced map $f'\colon G_1^{(0)} \to G_2^{(0)}$ is proper. Then G_2 is proper.*

Proof. Denote by r_i and s_i the range and source maps of G_i $(i = 1,2)$. Let $F_2 \subset G_2$ be a closed subspace, and $F_1 = f^{-1}(F_2)$. Since G_1 is proper, $(r_1,s_1)(F_1)$ is closed, and since $f' \times f'$ is proper, $(f' \times f') \circ (r_1, s_1)(F_1)$ is closed. By surjectivity of f, we have $(r_2, s_2)(F_2) = (f' \times f') \circ (r_1, s_1)(F_1)$. This proves that (r_2, s_2) is closed. Since for every topological space T, the assumptions of the proposition are also true for the morphism $f \times 1\colon G_1 \times T \to G_2 \times T$, the above shows that $(r_2, s_2) \times 1_T$ is closed. Therefore, (r_2, s_2) is proper. □

Proposition 2.20. *Let G be a topological groupoid with $G^{(0)}$ Hausdorff, acting on two spaces Y and Z. Suppose that the action of G on Z is proper, and that Y is Hausdorff. Then G acts properly on $Y \times_{G^{(0)}} Z$.*

Proof. The groupoid $(Y \times_{G^{(0)}} Z) \rtimes G$ is isomorphic to the subgroupoid $\Gamma = \{(y, y', z, g) \in (Y \times Y) \times (Z \rtimes G) |\ p(y) = r(g),\ y' = yg\}$ of the proper groupoid

$(Y \times Y) \times (Z \rtimes G)$. Since Y and $G^{(0)}$ are Hausdorff, Γ is closed in $(Y \times Y) \times (Z \rtimes G)$, hence by Proposition 2.10(ii), $(Y \times_{G^{(0)}} Z) \rtimes G$ is proper. □

Corollary 2.21. *Let G be a proper topological groupoid with $G^{(0)}$ Hausdorff. Then any action of G on a Hausdorff space is proper.*

Proof. Follows from Proposition 2.20 with $Z = G^{(0)}$. □

Proposition 2.22. *Let G be a topological groupoid and $f\colon T \to G^{(0)}$ be a continuous map.*

(a) *If G is proper, then $G[T]$ is proper.*
(ii) *If $G[T]$ is proper and f is open surjective, then G is proper.*

Proof. Let us prove (a). Suppose first that T is a subspace of $G^{(0)}$ and that f is the inclusion. Then $G[T] = G_T^T$. Since (r_T, s_T) is the restriction to $(r,s)^{-1}(T \times T)$ of (r,s), and (r,s) is proper, it follows that (r_T, s_T) is proper. In the general case, let $\Gamma = (T \times T) \times G$ and let $T' \subset T \times G^{(0)}$ be the graph of f. Then Γ is a proper groupoid (since it is the product of two proper groupoids), and $G[T] = \Gamma[T']$.

Let us prove (b). The only difficulty is to show that (r,s) is closed. Let $F \subset G$ be a closed subspace and $(y,x) \in \overline{(r,s)(F)}$. Let $\tilde{F} = G[T] \cap (T \times T) \times F$. Choose $(t',t) \in T \times T$ such that $f(t') = y$ and $f(t) = x$. Denote by $\tilde{r}$ and $\tilde{s}$ the range and source maps of $G[T]$. Then $(t',t) \in \overline{(\tilde{r},\tilde{s})(\tilde{F})}$. Indeed, let $\Omega \ni (t',t)$ be an open set, and $\Omega' = (f \times f)(\Omega)$. Then Ω' is an open neighborhood of (y,x), so $\Omega' \cap (r,s)(F) \neq \emptyset$. It follows that $\Omega \cap (\tilde{r},\tilde{s})(\tilde{F}) \neq \emptyset$.
We have proved that $(t',t) \in \overline{(\tilde{r},\tilde{s})(\tilde{F})} = (\tilde{r},\tilde{s})(\tilde{F})$, so $(y,x) \in (r,s)(F)$. □

Corollary 2.23. *Let G be a groupoid acting properly on a topological space Z, and let Z_1 be a saturated subspace. Then G acts properly on Z_1.*

Proof. Use the fact that $Z_1 \rtimes G = (Z \rtimes G)[Z_1]$. □

2.5. Invariance by Morita-equivalence. In this section, we will only consider groupoids whose range maps are open. We thus need a stability lemma:

Lemma 2.24. *Let G be a topological groupoid whose range map is open. Let Z be a G space and $f\colon T \to G^{(0)}$ be a continuous open map. Then the range maps for $Z \rtimes G$ and $G[T]$ are open.*

To prove Lemma 2.24 we need a preliminary result:

Lemma 2.25. *Let X, Y, T be topological spaces, $g\colon Y \to T$ an open map and $f\colon X \to T$ continuous. Let $Z = X \times_T Y$. Then the first projection $pr_1\colon X \times_T Y \to X$ is open.*

Proof. Let $\Omega \subset Z$ open. There exists an open subspace Ω' of $X \times Y$ such that $\Omega = \Omega' \cap Z$. Let Δ be the diagonal in $X \times X$. One easily checks that $(\mathrm{pr}_1, \mathrm{pr}_1)(\Omega) = (1 \times f)^{-1}(1 \times g)(\Omega') \cap \Delta$, therefore $(\mathrm{pr}_1, \mathrm{pr}_1)(\Omega)$ is open in Δ. This implies that $\mathrm{pr}_1(\Omega)$ is open in X. □

Proof of Lemma 2.24. This is clear for $Z \rtimes G = Z \times_{G^{(0)}} G$ using Lemma 2.25. For $G[T]$, first use Lemma 2.25 to prove that $T \times_{f,s} G \xrightarrow{pr_2} G$ is open. Since the range map is open by assumption, the composition $T \times_{f,s} G \xrightarrow{pr_2} G \xrightarrow{r} G^{(0)}$ is open. Using again Lemma 2.25, $G[T] \simeq T \times_{f,r\circ pr_2} (T \times_{f,s} G) \xrightarrow{pr_1} T$ is open. ☐

In order to define the notion of Morita-equivalence for topological groupoids, we introduce some terminology:

Definition 2.26. *Let G be a topological groupoid. Let T be a topological space and $\rho\colon G^{(0)} \to T$ be a G-invariant map. Then G is said to be ρ-proper if the map $(r,s)\colon G \to G^{(0)} \times_T G^{(0)}$ is proper. If G acts on a space Z and $\rho\colon Z \to T$ is G-invariant, then the action is said to be ρ-proper if $Z \rtimes G$ is ρ-proper.*

It is clear that properness implies ρ-properness. There is a partial converse:

Proposition 2.27. *Let G be a topological groupoid, T a topological space, $\rho\colon G^{(0)} \to T$ a G-invariant map. If G is ρ-proper and T is Hausdorff, then G is proper.*

Proof. Since T is Hausdorff, $G^{(0)} \times_T G^{(0)}$ is a closed subspace of $G^{(0)} \times G^{(0)}$, therefore (r,s), being the composition of the two proper maps $G \to G^{(0)} \times_T G^{(0)} \to G^{(0)} \times G^{(0)}$, is proper. ☐

Remark 2.28. *When T is locally Hausdorff, one easily shows that G is ρ-proper iff for every Hausdorff open subspace V of T, $G_{\rho^{-1}(V)}^{\rho^{-1}(V)}$ is proper.*

Proposition 2.29. [14] *Let G_1 and G_2 be two topological (resp. locally compact) groupoids. Let r_i, s_i $(i = 1,2)$ be the range and source maps of G_i, and suppose that r_i are open. The following are equivalent:*

(i) *there exist a topological (resp. locally compact) space T and $f_i\colon T \to G_i^{(0)}$ open surjective such that $G_1[T]$ and $G_2[T]$ are isomorphic;*
(ii) *there exists a topological (resp. locally compact) space Z, two continuous maps $\rho\colon Z \to G_1^{(0)}$ and $\sigma\colon Z \to G_2^{(0)}$, a left action of G_1 on Z with momentum map ρ and a right action of G_2 on Z with momentum map σ such that*
   (a) *the actions commute and are free, the action of G_2 is ρ-proper and the action of G_1 is σ-proper;*
   (b) *the natural maps $Z/G_2 \to G_1^{(0)}$ and $G_1\backslash Z \to G_2^{(0)}$ induced from ρ and σ are homeomorphisms.*

Moreover, one may replace (b) by

(b)' *ρ and σ are open and induce bijections $Z/G_2 \to G_1^{(0)}$ and $G_1\backslash Z \to G_2^{(0)}$.*
In (i), *if T is locally compact then it may be assumed Hausdorff.*

If G_1 and G_2 satisfy the equivalent conditions in Proposition 2.29, then they are said to be Morita-equivalent. Note that if $G_i^{(0)}$ are Hausdorff, then by Proposition 2.27, one may replace "ρ-proper" and "σ-proper" by "proper".
To prove Proposition 2.29, we need preliminary lemmas:

LEMMA 2.30. *Let G be a topological groupoid. The following are equivalent:*

(i) *$r\colon G \to G^{(0)}$ is open;*
(ii) *for every G-space Z, the canonical mapping $\pi\colon Z \to Z/G$ is open.*

Proof. To show (ii) $\Longrightarrow$ (i), take $Z = G$: the canonical mapping $\pi\colon G \to G/G$ is open. Therefore, for every open subspace U of G, $r(U) = G^{(0)} \cap \pi^{-1}(\pi(U))$ is open.
Let us show (i) $\Longrightarrow$ (ii). By Lemma 2.24, the range map $r\colon Z \rtimes G \to Z$ is open. The conclusion follows from Proposition 2.11. $\square$

LEMMA 2.31. *Let G be a topological groupoid such that the range map $r\colon G \to G^{(0)}$ is open. Let X be a topological space endowed with an action of G and T a topological space. Then the canonical map*

$$f\colon (X \times T)/G \to (X/G) \times T$$

is an isomorphism.

Proof. Let $\pi\colon X \to X/G$ and $\pi'\colon X \times T \to (X \times T)/G$ be the canonical mappings. Since π is open (Lemma 2.30), $f \circ \pi' = \pi \times 1$ is open. Since π' is continuous surjective, it follows that f is open. $\square$

LEMMA 2.32. *Let G be a topological groupoid whose range map is open and $f\colon Y \to Z$ a proper, G-equivariant map between two G-spaces. Then the induced map $\bar f\colon Y/G \to Z/G$ is proper.*

Proof. We first show that $\bar f$ is closed. Let $\pi\colon Y \to Y/G$ and $\pi'\colon Z \to Z/G$ be the canonical mappings. Let $A \subset Y/G$ be a closed subspace. Since f is closed and π is continuous, $(\pi')^{-1}(\bar f(A)) = f(\pi^{-1}(A))$ is closed. Therefore, $\bar f(A)$ is closed.
Applying this to $f \times 1$, we see that for every topological space T, $(Y \times T)/G \to (Z \times T)/G$ is closed. By Lemma 2.31, $\bar f \times 1_T$ is closed. $\square$

LEMMA 2.33. *Let G_2 and G_3 be topological groupoids whose range maps are open. Let Z_1, Z_2 and X be topological spaces. Suppose there are maps*

$$X \xleftarrow{\rho_1} Z_1 \xrightarrow{\sigma_1} G_2^{(0)} \xleftarrow{\rho_2} Z_2 \xrightarrow{\sigma_2} G_3^{(0)},$$

a right action of G_2 on Z_1 with momentum map σ_1, such that ρ_1 is G_2-invariant and the action of G_2 is ρ_1-proper, a left action of G_2 on Z_2 with momentum map ρ_2 and a right ρ_2-proper action of G_3 on Z_2 with momentum map σ_2 which commutes with the G_2-action.
Then the action of G_3 on $Z = Z_1 \times_{G_2} Z_2$ is ρ_1-proper.

Proof. Let $\varphi\colon Z_2 \rtimes G_3 \to Z_2 \times_{G_2^{(0)}} Z_2$ be the map $(z_2,\gamma)\mapsto(z_2,z_2\gamma)$. By assumption, φ is proper, therefore $1_{Z_1}\times\varphi$ is proper. Let $F=\{(z_1,z_2,z_2')\in Z_1\times Z_2\times Z_2 |\ \sigma_1(z_1)=\rho_2(z_2)=\rho_2(z_2')\}$. Then $1_{Z_1}\times\varphi\colon(1\times\varphi)^{-1}(F)\to F$ is proper, i.e. $Z_1\times_{G_2^{(0)}}(Z_2\rtimes G_3)\to Z_1\times_{G_2^{(0)}}(Z_2\times_{G_2^{(0)}}Z_2)$ is proper. By Lemma 2.32, taking the quotient by G_2, we get that the map

$$\alpha\colon Z\rtimes G_3\to Z_1\times_{G_2}(Z_2\times_{G_2^{(0)}}Z_2)$$

defined by $(z_1,z_2,\gamma)\mapsto(z_1,z_2,z_2\gamma)$ is proper.
By assumption, the map $Z_1\rtimes G_2\to Z_1\times_X Z_1$ given by $(z_1,g)\mapsto(z_1,z_1g)$ is proper. Endow $Z_1\rtimes G_2$ with the following right action of $G_2\times G_2$: $(z_1,g)\cdot(g',g'')=(z_1g',(g')^{-1}gg'')$. Using again Lemma 2.32, the map

$$\begin{aligned}\beta\colon Z_1\times_{G_2}(Z_2\times_{G_2^{(0)}}Z_2)&=(Z_1\rtimes G_2)\times_{G_2\times G_2}(Z_2\times Z_2)\\ &\to(Z_1\times_X Z_1)\times_{G_2\times G_2}(Z_2\times Z_2)\simeq Z\times_X Z\end{aligned}$$

is proper. By composition, $\beta\circ\alpha\colon Z\rtimes G_3\to Z\times_X Z$ is proper. $\square$

Proof of Proposition 2.29. Let us treat the case of topological groupoids. Assertion (b') follows from the fact that the canonical mappings $Z\to Z/G_2$ and $Z\to G_1\backslash Z$ are open (Lemma 2.30).
Let us first show that (ii) is an equivalence relation. Reflexivity is clear (taking $Z=G$, $\rho=r$, $\sigma=s$), and symmetry is obvious. Suppose that (Z_1,ρ_1,σ_2) and (Z_2,ρ_2,σ_2) are equivalences between G_1 and G_2, and G_2 and G_3 respectively. Let $Z=Z_1\times_{G_2}Z_2$ be the quotient of $Z_1\times_{G_2^{(0)}}Z_2$ by the action $(z_1,z_2)\cdot\gamma=(z_1\gamma,\gamma^{-1}z_2)$ of G_2. Denote by $\rho\colon Z\to G_1^{(0)}$ and $\sigma\colon Z\to G_3^{(0)}$ the maps induced from $\rho_1\times 1$ and $1\times\sigma_2$. By Lemma 2.25, the first projection $pr_1\colon Z_1\times_{G_2^{(0)}}Z_2\to Z_1$ is open, therefore $\rho=\rho_1\circ pr_1$ is open. Similarly, σ is open. It remains to show that the actions of G_3 and G_1 are ρ-proper and σ-proper respectively. For G_3, this follows from Lemma 2.33 and the proof for G_1 is similar.
This proves that (ii) is an equivalence relation. Now, let us prove that (i) and (ii) are equivalent.
Suppose (ii). Let $\Gamma=G_1\ltimes Z\rtimes G_2$ and $T=Z$. The maps $\rho\colon T\to G_1^{(0)}$ and $\sigma\colon T\to G_2^{(0)}$ are open surjective by assumption. Since $G_1\ltimes Z\simeq Z\times_{G_2^{(0)}}Z$ and $Z\rtimes G_2\simeq Z\times_{G_1^{(0)}}Z$, we have $G_2[T]=(T\times T)\times_{G_2^{(0)}\times G_2^{(0)}}G_2\simeq(Z\rtimes G_2)\times_{s\circ pr_2,\sigma}Z\simeq(Z\times_{G_1^{(0)}}Z)\times_{\sigma\circ pr_2,\sigma}Z=Z\times_{G_1^{(0)}}(Z\times_{G_2^{(0)}}Z)\simeq Z\times_{G_1^{(0)}}(G_1\ltimes Z)\simeq G_1\ltimes(Z\times_{G_1^{(0)}}Z)\simeq G_1\ltimes(Z\rtimes G_2)=\Gamma$. Similarly, $\Gamma\simeq G_1[T]$, hence (i).
Conversely, to prove $(i)\implies(ii)$ it suffices to show that if $f\colon T\to G^{(0)}$ is open surjective, then G and $G[T]$ are equivalent in the sense (ii), since we know that (ii) is an equivalence relation. Let $Z=T\times_{r,f}G$.
Let us check that the action of G is pr_1-proper. Write $Z\rtimes G=\{(t,g,h)\in T\times G\times G |\ f(t)=r(g)\text{ and } s(g)=r(h)\}$. One needs to check that the map $Z\rtimes G\to(T\times_{f,r}G)^2$ defined by $(t,g,h)\mapsto(t,g,t,h)$ is a homeomorphism onto its image. This follows easily from the facts that the diagonal map $T\to T\times T$

and the map $G^{(2)} \to G \times G$, $(g,h) \mapsto (g,gh)$ are homeomorphisms onto their images.
Let us check that the action of $G[T]$ is $s \circ pr_2$-proper. One easily checks that the groupoid $G' = G[T] \ltimes (T \times_{f,r} G)$ is isomorphic to a subgroupoid of the trivial groupoid $(T \times T) \times (G \times G)$. It follows that if r' and s' denote the range and source maps of G', the map (r', s') is a homeomorphism of G' onto its image.

Let us now treat the case of locally compact groupoids. In the proof that (ii) is a transitive relation, it just remains to show that Z is locally compact.
Let U_3 be a Hausdorff open subspace of $G_3^{(0)}$. We show that $\sigma^{-1}(U_3)$ is locally compact. Replacing G_3 by $(G_3)_{U_3}^{U_3}$, we may assume that G_2 acts freely and properly on Z_2. Let Γ be the groupoid $(Z_1 \times_{G_2^{(0)}} Z_2) \rtimes G_2$, and $R = (r,s)(\Gamma) \subset (Z_1 \times_{G_2^{(0)}} Z_2)^2$. Since the action of G_2 on Z_2 is free and proper, there exists a continuous map $\varphi\colon Z_2 \times_{G_3^{(0)}} Z_2 \to G_2$ such that $z_2 = \varphi(z_2, z_2')z_2'$. Then $R = \{(z_1, z_2, z_1', z_2') \in (Z_1 \times_{G_2^{(0)}} Z_2)^2;\ z_1' = z_1\varphi(z_2, z_2')\}$ is locally closed. By Proposition 2.12, $Z = (Z_1 \times_{G_2^{(0)}} Z_2)/G$ is locally compact.
Finally, if (i) holds with $T = \cup_i V_i$ with V_i open Hausdorff, let $T' = \amalg V_i$. It is clear that $G_1[T'] \simeq G_2[T']$. $\square$

Let us examine standard examples of Morita-equivalences:

Example 2.34. *Let G be a topological groupoid whose range map is open. Let $(U_i)_{i\in I}$ be an open cover of $G^{(0)}$ and $\mathcal{U} = \amalg_{i\in I} U_i$. Then $G[\mathcal{U}]$ is Morita-equivalent to G.*

Example 2.35. *Let G be a topological groupoid, and let H_1, H_2 be subgroupoids such that the range maps $r_i\colon H_i \to H_i^{(0)}$ are open. Then $(H_1\backslash G_{s(H_2)}^{s(H_1)}) \rtimes H_2$ and $H_1 \ltimes (G_{s(H_2)}^{s(H_1)}/H_2)$ are Morita-equivalent.*

Proof. Take $Z = G_{s(H_2)}^{s(H_1)}$ and let $\rho\colon Z \to Z/H_2$ and $\sigma\colon H_1\backslash Z$ be the canonical mappings. The fact that these maps are open follows from Lemma 2.30. $\square$

The following proposition is an immediate consequence of Proposition 2.22.

Proposition 2.36. *Let G and G' be two topological groupoids such that the range maps of G and G' are open. Suppose that G and G' are Morita-equivalent. Then G is proper if and only if G' is proper.*

Corollary 2.37. *With the notations of Example 2.34, G is proper if and only if $G[\mathcal{U}]$ is proper.*

3. A topological construction

Let X be a locally compact space. Since X is not necessarily Hausdorff, a filter[1] $\mathcal{F}$ on X may have more than one limit. Let S be the set of limits of a convergent filter $\mathcal{F}$. The goal of this section is to construct a Hausdorff space

[1] or a net; we will use indifferently the two equivalent approaches

$\mathcal{H}X$ in which X is (not continuously) embedded, and such that $\mathcal{F}$ converges to S in $\mathcal{H}X$.

3.1. The space $\mathcal{H}X$.

Lemma 3.1. *Let X be a topological space, and $S \subset X$. The following are equivalent:*

(i) *for every family $(V_s)_{s\in S}$ of open sets such that $s \in V_s$, and $V_s = X$ except perhaps for finitely many s's, one has $\cap_{s\in S} V_s \neq \emptyset$;*
(ii) *for every finite family $(V_i)_{i\in I}$ of open sets such that $S \cap V_i \neq \emptyset$ for all i, one has $\cap_{i\in I} V_i \neq \emptyset$.*

Proof. (i) $\Longrightarrow$ (ii): let $(V_i)_{i\in I}$ as in (ii). For all i, choose $s(i) \in S \cap V_i$. Put $W_s = \cap_{s=s(i)} V_i$, with the convention that an empty intersection is X. Then by (i), $\emptyset \neq \cap_{s\in S} W_s = \cap_{i\in I} V_i$.
(ii) $\Longrightarrow$ (i): let $(V_s)_{s\in S}$ as in (i), and let $I = \{s \in S |\ V_s \neq X\}$. Then $\cap_{s\in S} V_s = \cap_{i\in I} V_i \neq \emptyset$. □

We shall denote by $\mathcal{H}X$ the set of non-empty subspaces S of X which satisfy the equivalent conditions of Lemma 3.1, and $\hat{\mathcal{H}}X = \mathcal{H}X \cup \{\emptyset\}$.

Lemma 3.2. *Let X be a locally Hausdorff space. Then every $S \in \mathcal{H}X$ is locally finite. More precisely, if V is a Hausdorff open subspace of X, then $V \cap S$ has at most one element.*

Proof. Suppose $a \neq b$ and $\{a, b\} \subset V \cap S$. Then there exist V_a, V_b open disjoint neighborhoods of a and b respectively; this contradicts Lemma 3.1(ii). □

Suppose that X is locally compact. We endow $\hat{\mathcal{H}}X$ with a topology. Let us introduce the notations $\Omega_V = \{S \in \mathcal{H}X |\ V \cap S \neq \emptyset\}$ and $\Omega^Q = \{S \in \mathcal{H}X |\ Q \cap S = \emptyset\}$. The topology on $\hat{\mathcal{H}}X$ is generated by the Ω_V's and Ω^Q's (V open and Q quasi-compact). More explicitly, a set is open if and only if it is a union of sets of the form $\Omega^Q_{(V_i)_{i\in I}} = \Omega^Q \cap (\cap_{i\in I} \Omega_{V_i})$ where $(V_i)_{i\in I}$ is a finite family of open Hausdorff sets and Q is quasi-compact.

Proposition 3.3. *For every locally compact space X, the space $\hat{\mathcal{H}}X$ is Hausdorff.*

Proof. Suppose $S \not\subset S'$ and $S, S' \in \hat{\mathcal{H}}X$. Let $s \in S - S'$. Since S' is locally finite and since every singleton subspace of X is closed, there exist V open and K compact such that $s \in V \subset K$ and $K \cap S' = \emptyset$. Then Ω_V and Ω^K are disjoint neighborhoods of S and S' respectively. □

For every filter $\mathcal{F}$ on $\hat{\mathcal{H}}X$, let

$$L(\mathcal{F}) = \{a \in X |\ \forall V \ni a \text{ open}, \Omega_V \in \mathcal{F}\}. \quad (1)$$

Lemma 3.4. *Let X be a locally compact space. Let $\mathcal{F}$ be a filter on $\hat{\mathcal{H}}X$. Then $\mathcal{F}$ converges to $S \in \hat{\mathcal{H}}X$ if and only if properties (a) and (b) below hold:*

(a) *$\forall V$ open, $V \cap S \neq \emptyset \implies \Omega_V \in \mathcal{F}$;*

(b) $\forall Q$ *quasi-compact,* $Q \cap S = \emptyset \implies \Omega^Q \in \mathcal{F}$.

If $\mathcal{F}$ *is convergent, then* $L(\mathcal{F})$ *is its limit.*

Proof. The first statement is obvious, since every open set in $\hat{\mathcal{H}}X$ is a union of finite intersections of Ω_V's and Ω^Q's.

Let us prove the second statement. It is clear from (a) that $S \subset L(\mathcal{F})$. Conversely, suppose there exists $a \in L(\mathcal{F}) - S$. Since S is locally finite and every singleton subspace of X is closed, there exists a compact neighborhood K of a such that $K \cap S = \emptyset$. Then $a \in L(\mathcal{F})$ implies $\Omega_K \in \mathcal{F}$, and condition (b) implies $\Omega^K \in \mathcal{F}$, thus $\emptyset = \Omega^K \cap \Omega_K \in \mathcal{F}$, which is impossible: we have proved the reverse inclusion $L(\mathcal{F}) \subset S$. □

Remark 3.5. *This means that if* $S_\lambda \to S$, *then* $a \in S$ *if and only if* $\forall\lambda$ *there exists* $s_\lambda \in S_\lambda$ *such that* $s_\lambda \to a$.

Example 3.6. *Consider Example 2.3 with* $\Gamma = \mathbb{Z}_2$ *and* $H = \{0\}$. *Then* $\mathcal{H}G = G \cup \{S\}$ *where* $S = \{(0,0),(0,1)\}$. *The sequence* $(1/n, 0) \in G$ *converges to* S *in* $\mathcal{H}G$, *and* $(0,0)$ *and* $(0,1)$ *are two isolated points in* $\mathcal{H}G$.

Proposition 3.7. *Let* X *be a locally compact space and* $K \subset X$ *quasi-compact. Then* $L = \{S \in \mathcal{H}X |\ S \cap K \neq \emptyset\}$ *is compact. The space* $\mathcal{H}X$ *is locally compact, and it is* σ*-compact if* X *is* σ*-compact.*

Proof. We show that L is compact, and the two remaining assertions follow easily. Let $\mathcal{F}$ be a ultrafilter on L. Let $S_0 = L(\mathcal{F})$. Let us show that $S_0 \cap K \neq \emptyset$: for every $S \in L$, choose a point $\varphi(S) \in K \cap S$. By quasi-compactness, $\varphi(\mathcal{F})$ converges to a point $a \in K$, and it is not hard to see that $a \in S_0$.

Let us show $S_0 \in \mathcal{H}X$: let (V_s) $(s \in S_0)$ be a family of open subspaces of X such that $s \in V_s$ for all $s \in S_0$, and $V_s = X$ for every $s \notin S_1$ ($S_1 \subset S_0$ finite). By definition of S_0, $\Omega_{(V_s)_{s \in S_1}} = \cap_{s \in S_1} \Omega_{V_s}$ belongs to $\mathcal{F}$, hence it is non-empty. Choose $S \in \Omega_{(V_s)_{s \in S_1}}$, then $S \cap V_s \neq \emptyset$ for all $s \in S_1$. By Lemma 3.1(ii), $\cap_{s \in S_1} V_s \neq \emptyset$. This shows that $S_0 \in \mathcal{H}X$.

Now, let us show that $\mathcal{F}$ converges to S_0.

- If V is open Hausdorff such that $S_0 \in \Omega_V$, then by definition $\Omega_V \in \mathcal{F}$.
- If Q is quasi-compact and $S_0 \in \Omega^Q$, then $\Omega^Q \in \mathcal{F}$, otherwise one would have $\{S \in \mathcal{H}X |\ S \cap Q \neq \emptyset\} \in \mathcal{F}$, which would imply as above that $S_0 \cap Q \neq \emptyset$, a contradiction.

From Lemma 3.4, $\mathcal{F}$ converges to S_0. □

Proposition 3.8. *Let* X *be a locally compact space. Then* $\hat{\mathcal{H}}X$ *is the one-point compactification of* $\mathcal{H}X$.

Proof. It suffices to prove that $\hat{\mathcal{H}}X$ is compact. The proof is almost the same as in Proposition 3.7. □

Remark 3.9. *If* $f\colon X \to Y$ *is a continuous map from a locally compact space* X *to any* Hausdorff *space* Y, *then* f *induces a continuous map* $\mathcal{H}f\colon \mathcal{H}X \to Y$. *Indeed, for every open subspace* V *of* Y, $(\mathcal{H}f)^{-1}(V) = \Omega_{f^{-1}(V)}$ *is open.*

Proposition 3.10. *Let G be a topological groupoid such that $G^{(0)}$ is Hausdorff, and $r\colon G \to G^{(0)}$ is open. Let Z be a locally compact space endowed with a continuous action of G. Then $\mathcal{H}Z$ is endowed with a continuous action of G which extends the one on Z.*

Proof. Let $p\colon Z \to G^{(0)}$ such that G acts on Z with momentum map p. Since p has a continuous extension $\mathcal{H}p\colon \mathcal{H}Z \to G^{(0)}$, for all $S \in \mathcal{H}Z$, there exists $x \in G^{(0)}$ such that $S \subset p^{-1}(x)$. For all $g \in G^x$, write $Sg = \{sg|\ s \in S\}$.
Let us show that $Sg \in \mathcal{H}Z$. Let V_s ($s \in S$) be open sets such that $sg \in V_s$. By continuity, there exist open sets $W_s \ni s$ and $W_g \ni g$ such that for all $(z,h) \in W_s \times_{G^{(0)}} W_g$, $zh \in V_s$. Let $V'_s = W_s \cap p^{-1}(r(W_g))$. Then V'_s is an open neighborhood of s, so there exists $z \in \cap_{s\in S} V'_s$. Since $p(z) \in r(W_g)$, there exists $h \in W_g$ such that $p(z) = r(h)$. It follows that $zh \in \cap_{s\in S} V_s$. This shows that $Sg \in \mathcal{H}Z$.
Let us show that the action defined above is continuous. Let $\Phi\colon \mathcal{H}Z \times_{G^{(0)}} G \to \mathcal{H}Z$ be the action of G on $\mathcal{H}Z$. Suppose that $(S_\lambda, g_\lambda) \to (S,g)$ and let $S' = L((S_\lambda, g_\lambda))$. Then for all $a \in S$ there exists $s_\lambda \in S_\lambda$ such that $s_\lambda \to a$. This implies $s_\lambda g_\lambda \to ag$, thus $ag \in S'$. The converse may be proved in a similar fashion, hence $Sg = S'$.
Applying this to any universal net (S_λ, g_λ) converging to (S,g) and knowing from Proposition 3.8 that $\Phi(S_\lambda, g_\lambda)$ is convergent in $\hat{\mathcal{H}}Z$, we find that $\Phi(S_\lambda, g_\lambda)$ converges to $\Phi(S,g)$. This shows that Φ is continuous in (S,g). □

3.2. The space $\mathcal{H}'X$. Let X be a locally compact space. Let $\Omega'_V = \{S \in \mathcal{H}X|\ S \subset V\}$. Let $\mathcal{H}'X$ be $\mathcal{H}X$ as a set, with the coarsest topology such that the identity map $\mathcal{H}'X \to \mathcal{H}X$ is continuous, and Ω'_V is open for every relatively quasi-compact open set V. The space $\mathcal{H}'X$ is Hausdorff since $\mathcal{H}X$ is Hausdorff, but it is usually not locally compact.

Lemma 3.11. *Let X be a locally compact space. Then the map*

$$\mathcal{H}'X \to \mathbb{N}^* \cup \{\infty\}, \quad S \mapsto \#S$$

is upper semi-continuous.

Proof. Let $S \in \mathcal{H}'X$ such that $\#S < \infty$. Let V_s ($s \in S$) be open relatively compact Hausdorff sets such that $s \in V_s$, and let $W = \cup_{s\in S} V_s$. Then $S' \in \mathcal{H}'X$ implies $\#(S' \cap V_s) \le 1$, therefore $S' \in \Omega'_W$ implies $\#S' \le \#S$. □

Proposition 3.12. *Let X be a locally compact space such that the closure of every quasi-compact subspace is quasi-compact. Then*

(a) *the natural map $\mathcal{H}'X \to \mathcal{H}X$ is a homeomorphism,*
(b) *for every compact subspace $K \subset X$, there exists $C_K > 0$ such that*

$$\forall S \in \mathcal{H}X,\ S \cap K \neq \emptyset \implies \#S \le C_K,$$

(c) *If G is a locally compact proper groupoid with $G^{(0)}$ Hausdorff then G satisfies the above properties.*

Proof. To prove (b), let K_1 be a quasi-compact neighborhood of K and let $K' = \overline{K}_1$. Let $a \in K \cap S$ and suppose there exists $b \in S - K'$. Then $\mathring{K}_1$ and $X - K'$ are disjoint neighborhoods of a and b respectively, which is impossible. We deduce that $S \subset K'$.
Now, let $(V_i)_{i\in I}$ be a finite cover of K' by open Hausdorff sets. For all $b \in S$, let $I_b = \{i \in I |\ b \in V_i\}$. By Lemma 3.2, the I_b's ($b \in S$) are disjoint, whence one may take $C_K = \#I$.

To prove (a), denote by $\Delta \subset X \times X$ the diagonal. Let us first show that $pr_1 \colon \overline{\Delta} \to X \times X$ is proper.
Let $K \subset X$ compact. Let $L \subset X$ quasi-compact such that $K \subset \mathring{L}$. If $(a,b) \in \overline{\Delta} \cap (K \times X)$, then $b \in \overline{L}$: otherwise, $L \times L^c$ would be a neighborhood of (a,b) whose intersection with Δ is empty. Therefore, $pr_1^{-1}(K) = \overline{\Delta} \cap (K \times \overline{L})$ is quasi-compact, which shows that pr_1 is proper.
It remains to prove that Ω'_V is open in $\mathcal{H}X$ for every relatively quasi-compact open set $V \subset X$. Let $S \in \Omega'_V$, $a \in S$ and K a compact neighborhood of a. Let $L = pr_2(\overline{\Delta} \cap (K \times X))$. Then $Q = L - V$ is quasi-compact, and $S \in \Omega^Q_{\mathring{K}} \subset \Omega'_V$, therefore Ω'_V is a neighborhood of each of its points.

To prove (c), let $K \subset G$ be a quasi-compact subspace. Then $L = r(K) \cup s(K)$ is quasi-compact, thus G^L_L is also quasi-compact. But $\overline{K}$ is closed and $\overline{K} \subset G^L_L$, therefore $\overline{K}$ is quasi-compact. □

4. HAAR SYSTEMS

4.1. THE SPACE $C_c(X)$. For every locally compact space X, $C_c(X)_0$ will denote the set of functions $f \in C_c(V)$ (V open Hausdorff), extended by 0 outside V. Let $C_c(X)$ be the linear span of $C_c(X)_0$. Note that functions in $C_c(X)$ are not necessarily continuous.

PROPOSITION 4.1. *Let X be a locally compact space, and let $f \colon X \to \mathbb{C}$. The following are equivalent:*

(i) $f \in C_c(X)$;
(ii) $f^{-1}(\mathbb{C}^*)$ *is relatively quasi-compact, and for every filter $\mathcal{F}$ on X, let $\tilde{\mathcal{F}} = i(\mathcal{F})$, where $i \colon X \to \mathcal{H}X$ is the canonical inclusion; if $\tilde{\mathcal{F}}$ converges to $S \in \mathcal{H}X$, then $\lim_{\mathcal{F}} f = \sum_{s\in S} f(s)$.*

Proof. Let us show (i) $\Longrightarrow$ (ii). By linearity, it is enough to consider the case $f \in C_c(V)$, where $V \subset X$ is open Hausdorff. Let K be the compact set $\overline{f^{-1}(\mathbb{C}^*)} \cap V$. Then $f^{-1}(\mathbb{C}^*) \subset K$. Let $\mathcal{F}$ and S as in (ii). If $S \cap V = \emptyset$, then $S \in \Omega^K$, hence $\Omega^K \in \tilde{\mathcal{F}}$, i.e. $X - K \in \mathcal{F}$. Therefore, $\lim_{\mathcal{F}} f = 0 = \sum_{s\in S} f(s)$. If $S \cap V = \{a\}$, then a is a limit point of $\mathcal{F}$, therefore $\lim_{\mathcal{F}} f = f(a) = \sum_{s\in S} f(s)$.
Let us show (ii) $\Longrightarrow$ (i) by induction on $n \in \mathbb{N}^*$ such that there exist $V_1, \dots V_n$ open Hausdorff and K quasi-compact satisfying $f^{-1}(\mathbb{C}^*) \subset K \subset V_1 \cup \dots \cup V_n$.

For $n = 1$, for every $x \in V_1$, let $\mathcal{F}$ be a ultrafilter convergent to x. By Proposition 3.8, $\tilde{\mathcal{F}}$ is convergent; let S be its limit, then $\lim_{\mathcal{F}} f = \sum_{s\in S} f(s) = f(x)$, thus $f_{|V_1}$ is continuous.
Now assume the implication is true for $n-1$ ($n \geq 2$) and let us prove it for n. Since K is quasi-compact, there exist $V_1', \ldots, V_n'$ open sets, $K_1 \ldots, K_n$ compact such that $K \subset V_1' \cup \cdots \cup V_n'$ and $V_i' \subset K_i \subset V_i$. Let $F = (V_1' \cup \cdots \cup V_n') - (V_1' \cup \cdots \cup V_{n-1}')$. Then F is closed in V_n' and $f_{|F}$ is continuous. Moreover, $f_{|F} = 0$ outside $K' = K - (V_1' \cup \cdots \cup V_{n-1}')$ which is closed in K, hence quasi-compact, and Hausdorff, since $K' \subset V_n'$. Therefore, $f_{|F} \in C_c(F)$. It follows that there exists an extension $h \in C_c(V_n')$ of $f_{|F}$. By considering $f - h$, we may assume that $f = 0$ on F, so $f = 0$ outside $K' = K_1 \cup \cdots \cup K_{n-1}$. But $K' \subset V_1 \cup \cdots \cup V_{n-1}$, hence by induction hypothesis, $f \in C_c(X)$. □

COROLLARY 4.2. *Let X be a locally compact space, $f\colon X \to \mathbb{C}$, $f_n \in C_c(X)$. Suppose that there exists fixed quasi-compact set $Q \subset X$ such that $f_n^{-1}(\mathbb{C}^*) \subset Q$ for all n, and f_n converges uniformly to f. Then $f \in C_c(X)$.*

LEMMA 4.3. *Let X be a locally compact space. Let $(U_i)_{i\in I}$ be an open cover of X by Hausdorff subspaces. Then every $f \in C_c(X)$ is a finite sum $f = \sum f_i$, where $f_i \in C_c(U_i)$.*

Proof. See [6, Lemma 1.3]. □

LEMMA 4.4. *Let X and Y be locally compact spaces. Let $f \in C_c(X \times Y)$. Let V and W be open subspaces of X and Y such that $f^{-1}(\mathbb{C}^*) \subset Q \subset V \times W$ for some quasi-compact set Q. Then there exists a sequence $f_n \in C_c(V) \otimes C_c(W)$ such that $\lim_{n\to\infty} \|f - f_n\|_\infty = 0$.*

Proof. We may assume that $X = V$ and $Y = W$. Let (U_i) (resp. (V_j)) be an open cover of X (resp. Y) by Hausdorff subspaces. Then every element of $C_c(X\times Y)$ is a linear combination of elements of $C_c(U_i \times V_j)$ (Lemma 4.3). The conclusion follows from the fact that the image of $C_c(U_i)\otimes C_c(V_j) \to C_c(U_i\times V_j)$ is dense. □

LEMMA 4.5. *Let X be a locally compact space and $Y \subset X$ a closed subspace. Then the restriction map $C_c(X) \to C_c(Y)$ is well-defined and surjective.*

Proof. Let $(U_i)_{i\in I}$ be a cover of X by Hausdorff open subspaces. The map $C_c(U_i) \to C_c(U_i \cap Y)$ is surjective (since Y is closed), and $\oplus_{i\in I} C_c(U_i \cap Y) \to C_c(Y)$ is surjective (Lemma 4.3). Therefore, the map $\oplus_{i\in I} C_c(U_i) \to C_c(Y)$ is surjective. Since it is also the composition of the surjective map $\oplus_{i\in I} C_c(U_i) \to C_c(X)$ and of the restriction map $C_c(X) \to C_c(Y)$, the conclusion follows. □

4.2. HAAR SYSTEMS. Let G be a locally compact proper groupoid with Haar system (see definition below) such that $G^{(0)}$ is Hausdorff. If G is Hausdorff, then $C_c(G^{(0)})$ is endowed with the $C_r^*(G)$-valued scalar product $\langle \xi, \eta\rangle(g) = \overline{\xi(r(g))}\eta(s(g))$. Its completion is a $C_r^*(G)$-Hilbert module. However, if G is not Hausdorff, the function $g \mapsto \overline{\xi(r(g))}\eta(s(g))$ does not necessarily belong to

$C_c(G)$, therefore we need a different construction in order to obtain a $C^*_r(G)$-module.

Definition 4.6. [16, pp. 16-17] *Let G be a locally compact groupoid such that G^x is Hausdorff for every $x \in G^{(0)}$. A Haar system is a family of positive measures $\lambda = \{\lambda^x |\ x \in G^{(0)}\}$ such that $\forall x, y \in G^{(0)}$, $\forall \varphi \in C_c(G)$,*

(i) $\operatorname{supp}(\lambda^x) = G^x$;
(ii) $\lambda(\varphi)\colon x \mapsto \int_{g\in G^x} \varphi(g)\,\lambda^x(\mathrm{d}g) \quad \in C_c(G^{(0)})$;
(iii) $\int_{h\in G^x} \varphi(gh)\,\lambda^x(\mathrm{d}h) = \int_{h\in G^y} \varphi(h)\,\lambda^y(\mathrm{d}h)$.

Note that G^x is automatically Hausdorff if $G^{(0)}$ is Hausdorff (Prop. 2.8). Recall also [15, p. 36] that the range map for G is open.

Lemma 4.7. *Let G be a locally compact groupoid with Haar system. Then for every quasi-compact subspace K of G, $\sup_{x\in G^{(0)}} \lambda^x(K \cap G^x) < \infty$.*

Proof. It is easy to show that there exists $f \in C_c(G)$ such that $1_K \le f$. Since $\sup_{x\in G^{(0)}} \lambda(f)(x) < \infty$, the conclusion follows. □

Lemma 4.8. *Let G be a locally compact groupoid with Haar system such that $G^{(0)}$ is Hausdorff. Suppose that Z is a locally compact space and that $p\colon Z \to G^{(0)}$ is continuous. Then for every $f \in C_c(Z \times_{p,r} G)$, $\lambda(f)\colon z \mapsto \int_{g\in G^{p(z)}} f(z, g)\,\lambda^{p(z)}(\mathrm{d}g)$ belongs to $C_c(Z)$.*

Proof. By Lemma 4.5, f is the restriction of an element of $C_c(Z \times G)$.
If $f(z, g) = f_1(z)f_2(g)$, then $\psi(x) = \int_{g\in G^x} f_2(g)\,\lambda^x(dg)$ belongs to $C_c(G^{(0)})$, therefore $\psi \circ p \in C_b(Z)$. It follows that $\lambda(f) = f_1(\psi \circ p)$ belongs to $C_c(Z)$.
By linearity, if $f \in C_c(Z) \otimes C_c(G)$, then $\lambda(f) \in C_c(Z)$.
Now, for every $f \in C_c(Z \times G)$, there exist relatively quasi-compact open subspaces V and W of Z and G and a sequence $f_n \in C_c(V) \otimes C_c(W)$ such that f_n converges uniformly to f. From Lemma 4.7, $\lambda(f_n)$ converges uniformly to $\lambda(f)$, and $\lambda(f_n) \in C_c(Z)$. From Corollary 4.2, $\lambda(f) \in C_c(Z)$. □

Proposition 4.9. *Let G be a locally compact groupoid with Haar system such that $G^{(0)}$ is Hausdorff. If G acts on a locally compact space Z with momentum map $p\colon Z \to G^{(0)}$, then $(\lambda^{p(z)})_{z\in Z}$ is a Haar system on $Z \rtimes G$.*

Proof. Results immediately from Lemma 4.8. □

5. The Hilbert module of a proper groupoid

5.1. The space X'. Before we construct a Hilbert module associated to a proper groupoid, we need some preliminaries. Let G be a locally compact groupoid such that $G^{(0)}$ is Hausdorff. Denote by X' the closure of $G^{(0)}$ in $\mathcal{H}G$.

Lemma 5.1. *Let G be a locally compact groupoid such that $G^{(0)}$ is Hausdorff. Then for all $S \in X'$, S is a subgroup of G.*

Proof. Since r and $s\colon G \to G^{(0)}$ extend continuously to maps $\mathcal{H}G \to G^{(0)}$, and since $r = s$ on $G^{(0)}$, one has $\mathcal{H}r = \mathcal{H}s$ on X', i.e. $\exists x_0 \in G^{(0)}$, $S \subset G_{x_0}^{x_0}$.
Let $\mathcal{F}$ be a filter on $G^{(0)}$ whose limit is S. Then $a \in S$ if and only if a is a limit point of $\mathcal{F}$. Since for every $x \in G^{(0)}$ we have $x^{-1}x = x$, it follows that for every a, $b \in S$ one has $a^{-1}b \in S$, whence S is a subgroup of $G_{x_0}^{x_0}$. □

Denote by $q\colon X' \to G^{(0)}$ the map such that $S \subset G_{q(S)}^{q(S)}$. The map q is continuous since it is the restriction to X' of $\mathcal{H}r$.

LEMMA 5.2. *Let G be a locally compact proper groupoid such that $G^{(0)}$ is Hausdorff. Let $\mathcal{F}$ be a filter on X', convergent to S. Suppose that $q(\mathcal{F})$ converges to $S_0 \in X'$. Then S_0 is a normal subgroup of S, and there exists $\Omega \in \mathcal{F}$ such that $\forall S' \in \Omega$, S' is group-isomorphic to S/S_0. In particular, $\{S' \in X' |\ \#S = \#S_0 \#S'\} \in \mathcal{F}$.*

Proof. Using Proposition 3.12, we see that S is finite.
We shall use the notation $\tilde{\Omega}_{(V_i)_{i\in I}} = \Omega_{(V_i)_{i\in I}} \cap \Omega'_{\cup_{i\in I} V_i}$. Let $V'_s \subset V_s$ $(s \in S)$ be Hausdorff, open neighborhoods of s, chosen small enough so that for some $\Omega \in \mathcal{F}$,

(a) $\Omega \subset \tilde{\Omega}_{(V'_s)_{s\in S}}$;
(b) $V'_{s_1} V'_{s_2} \subset V_{s_1 s_2}$, $\forall s_1$, $s_2 \in S$.
(c) $\forall s \in S - S_0$, $\forall S' \in \Omega$, $q(S') \notin V_s$;
(d) $q(\Omega) \subset \tilde{\Omega}_{(V_s)_{s\in S_0}}$;

Let $S' \in \Omega$. Let $\varphi\colon S \to S'$ such that $\{\varphi(s)\} = S' \cap V'_s$. Then φ is well-defined since $S' \cap V'_s \neq \emptyset$ (see (a)) and V'_s is Hausdorff.
If s_1, $s_2 \in S$ then $\varphi(s_i) \in S' \cap V'_{s_i}$. By (b), $\varphi(s_1)\varphi(s_2) \in S' \cap V_{s_1 s_2}$. Since $V_{s_1 s_2}$ is Hausdorff and also contains $\varphi(s_1 s_2) \in S'$, we have $\varphi(s_1 s_2) = \varphi(s_1)\varphi(s_2)$. This shows that φ is a group morphism.
The map φ is surjective, since $S' \subset \cup_{s\in S} V'_s$ (see (a)).
By (c), $\ker(\varphi) \subset S_0$ and by (d), $S_0 \subset \ker(\varphi)$. □

Suppose now that the range map $r\colon G \to G^{(0)}$ is open. Then X' is endowed with an action of G (Prop. 3.10) defined by $S \cdot g = g^{-1}Sg = \{g^{-1}sg |\ s \in S\}$.

5.2. CONSTRUCTION OF THE HILBERT MODULE. Now, let G be a locally compact, proper groupoid. Assume that G is endowed with a Haar system, and that $G^{(0)}$ is Hausdorff. Let

$$\mathcal{E}^0 = \{f \in C_c(X') |\ f(S) = \sqrt{\#S} f(q(S))\ \forall S \in X'\}.$$

($q(S) \in G^{(0)}$ is identified to $\{q(S)\} \in X'$.)
Define, for all ξ, $\eta \in \mathcal{E}^0$ and $f \in C_c(G)$: $\langle \xi, \eta\rangle(g) = \overline{\xi(r(g))}\eta(s(g))$ and

$$(\xi f)(S) = \int_{g\in G^{q(S)}} \xi(g^{-1}Sg) f(g^{-1})\, \lambda^x(dg).$$

PROPOSITION 5.3. *With the above assumptions, the completion $\mathcal{E}(G)$ of $\mathcal{E}^0$ with respect to the norm $\|\xi\| = \|\langle \xi, \xi\rangle\|^{1/2}$ is a $C_r^*(G)$-Hilbert module.*

We won't give the direct proof here since this is a particular case of Theorem 7.8 (see Example 7.7(c)).

6. Cutoff functions

If G is a locally compact Hausdorff proper groupoid with Haar system. Assume for simplicity that $G^{(0)}/G$ is compact. Then there exists a so-called "cutoff" function $c \in C_c(G^{(0)})_+$ such that for every $x \in G^{(0)}$, $\int_{g\in G^x} c(s(g))\,\lambda^x(\mathrm{d}g) = 1$, and the function $g \mapsto \sqrt{c(r(g))c(s(g))}$ defines projection in $C_r^*(G)$. However, if G is not Hausdorff, then the above function does not belong to $C_c(G)$ is general, thus we need another definition of a cutoff function.

Let $X'_{\geq k} = \{S \in X' |\ \#S \geq k\}$. By Lemma 3.11, $X'_{\geq k}$ is closed.

Lemma 6.1. *Let G be a locally compact, proper groupoid with $G^{(0)}$ Hausdorff. Let $X_{\geq k} = q(X'_{\geq k})$. Then $X_{\geq k}$ is closed in $G^{(0)}$.*

Proof. It suffices to show that for every compact subspace K of $G^{(0)}$, $X_{\geq k} \cap K$ is closed. Let $K' = G_K^K$. Then K' is quasi-compact, and from Proposition 3.7, $K'' = \{S \in \mathcal{H}G |\ S \cap K' \neq \emptyset\}$ is compact. The set $q^{-1}(K) \cap X'_{\geq k} = K'' \cap X'_{\geq k}$ is closed in K'', hence compact; its image by q is $X_{\geq k} \cap K$. ☐

Lemma 6.2. *Let G be a locally compact, proper groupoid, with $G^{(0)}$ Hausdorff. Let $\alpha \in \mathbb{R}$. For every compact set $K \subset G^{(0)}$, there exists $f \colon X'_K \to \mathbb{R}_+^*$ continuous, where $X'_K = q^{-1}(K) \subset X'$, such that*

$$\forall S \in X'_K, \quad f(S) = f(q(S))(\#S)^\alpha.$$

Proof. Let $K' = G_K^K$. It is closed and quasi-compact. From Proposition 3.7, X'_K is quasi-compact. For every $S \in X'_K$, we have $S \subset K'$. By Proposition 3.12, there exists $n \in \mathbb{N}^*$ such that $X'_{\geq n+1} \cap X'_K = \emptyset$. We can thus proceed by reverse induction: suppose constructed $f_{k+1} \colon X'_K \cap q^{-1}(X_{\geq k+1}) \to \mathbb{R}_+^*$ continuous such that $f_{k+1}(S) = f_{k+1}(q(S))(\#S)^\alpha$ for all $S \in X'_K \cap q^{-1}(X_{\geq k+1})$.
Since $X'_K \cap q^{-1}(X_{\geq k+1})$ is closed in the compact set $X'_K \cap q^{-1}(X_{\geq k})$, there exists a continuous extension $h \colon X'_K \cap q^{-1}(X_{\geq k}) \to \mathbb{R}$ of f_{k+1}. Replacing $h(x)$ by $\sup(h(x), \inf f_{k+1})$, we may assume that $h(X'_K \cap q^{-1}(X_{\geq k})) \subset \mathbb{R}_+^*$. Put $f_k(S) = h(q(S))(\#S)^\alpha$. Let us show that f_k is continuous.
Let $\mathcal{F}$ be a ultrafilter on $X'_K \cap q^{-1}(X_{\geq k})$, and let S be its limit. Since $q(\mathcal{F})$ is a ultrafilter on K, it has a limit $S_0 \in X'_K$.
For every $S_1 \in q^{-1}(X_{\geq k})$, choose $\psi(S_1) \in X'_{\geq k}$ such that $q(S_1) = q(\psi(S_1))$. Let $S' \in X'_K \cap X'_{\geq k}$ be the limit of $\psi(\mathcal{F})$.
From Lemma 5.2, $\Omega_1 = \{S_1 \in X'_K \cap q^{-1}(X_{\geq k}) |\ \#S = \#S_0 \#S_1\}$ is an element of $\mathcal{F}$, and $\Omega_2 = \{S_2 \in X'_{\geq k} |\ \#S' = \#S_0 \#S_2\}$ is an element of $\psi(\mathcal{F})$.

- If $\#S_0 > 1$, then $S' \in X_{\geq k+1}$, so S and S_0 belong to $q^{-1}(X_{\geq k+1})$. Therefore, $f_k(S_1) = (\#S_1)^\alpha h(q(S_1))$ converges with respect to $\mathcal{F}$ to

$$\begin{aligned}\frac{(\#S)^\alpha}{(\#S_0)^\alpha}h(S_0) = \frac{(\#S)^\alpha}{(\#S_0)^\alpha}f_{k+1}(S_0) = f_{k+1}(S)\\ = f_{k+1}(q(S))(\#S)^\alpha = h(q(S))(\#S)^\alpha = f_k(S).\end{aligned}$$

- If $S_0 = \{q(S)\}$, then $f_k(S_1) = (\#S_1)^\alpha h(q(S_1))$ converges with respect to $\mathcal{F}$ to $(\#S)^\alpha h(q(S)) = f_k(S)$.

Therefore, f_k is a continuous extension of f_{k+1}. $\square$

Theorem 6.3. *Let G be a locally compact, proper groupoid such that $G^{(0)}$ is Hausdorff and $G^{(0)}/G$ is σ-compact. Let $\pi\colon G^{(0)} \to G^{(0)}/G$ be the canonical mapping. Then there exists $c\colon X' \to \mathbb{R}_+$ continuous such that*

(a) *$c(S) = c(q(S))\#S$ for all $S \in X'$;*
(b) *$\forall \alpha \in G^{(0)}/G$, $\exists x \in \pi^{-1}(\alpha)$, $c(x) \neq 0$;*
(c) *$\forall K \subset G^{(0)}$ compact, $\operatorname{supp}(c) \cap q^{-1}(F)$ is compact, where $F = s(G^K)$.*

If moreover G admits a Haar system, then there exists $c\colon X' \to \mathbb{R}_+$ continuous satisfying (a), (b), (c) and

(d) $\forall x \in G^{(0)}$, $\displaystyle\int_{g\in G^x} c(s(g))\,\lambda^x(dg) = 1.$

Proof. There exists a locally finite cover (V_i) of $G^{(0)}/G$ by relatively compact open subspaces. Since π is open and $G^{(0)}$ is locally compact, there exists $K_i \subset G^{(0)}$ compact such that $\pi(K_i) \supset V_i$. Let (φ_i) be a partition of unity associated to the cover (V_i). For every i, from Lemma 6.2, there exists $c_i\colon X'_{K_i} \to \mathbb{R}_+^*$ continuous such that $c_i(S) = c_i(q(S))\#S$ for all $S \in X'_{K_i}$. Let

$$c(S) = \sum_i c_i(S)\varphi_i(\pi(q(S))).$$

It is clear that c is continuous from X' to $\mathbb{R}_+$, and that $c(S) = c(q(S))\#S$.
Let us prove (b): let $x_0 \in G^{(0)}$. There exists i such that $\varphi_i(\pi(x_0)) \neq 0$. Choose $x \in K_i$ such that $\pi(x) = \pi(x_0)$, then $c(x) \geq c_i(x)\varphi_i(\pi(x_0)) > 0$.
Let us show (c). Note that $F = \pi^{-1}(\pi(K))$ is closed, so $q^{-1}(F)$ is closed. Let K_1 be a compact neighborhood of K and $F_1 = \pi^{-1}(\pi(K_1))$. Let $J = \{i|\ V_i \cap \pi(K_1) \neq \emptyset\}$. Then for all $i \notin J$, $c_i(\varphi_i \circ \pi \circ q) = 0$ on $q^{-1}(F_1)$, therefore $c = \sum_{j\in J} c_j(\varphi_j \circ \pi \circ q)$ in a neighborhood of $q^{-1}(F)$. Since for all i, $\operatorname{supp}(c_i(\varphi_i \circ \pi \circ q))$ is compact and since J is finite, $\operatorname{supp}(c) \cap q^{-1}(F) \subset \cup_{i\in J}\operatorname{supp}(c_i(\varphi_i \circ \pi \circ q))$ is compact.
Let us show the last assertion. Let $\varphi(g) = c(s(g))$. Let $\mathcal{F}$ be a filter on G convergent in $\mathcal{H}G$ to $A \subset G$. Choose $a \in A$ and let $S = a^{-1}A$. Then $s(\mathcal{F})$ converges to S in $\mathcal{H}G$, hence

$$\lim_{\mathcal{F}} \varphi = \#Sc(s(a)) = \sum_{g\in S} c(s(g)) = \sum_{g\in S}\varphi(g).$$

For every compact set $K \subset G^{(0)}$,

$$\begin{aligned}&\{g \in G |\ r(g) \in K \text{ and } \varphi(g) \neq 0\}\\ &\quad \subset \quad \{g \in G |\ r(g) \in K \text{ and } s(g) \in \mathrm{supp}(c)\}\\ &\quad \subset \quad G^K_{q(\mathrm{supp}(c)\cap q^{-1}(F))},\end{aligned}$$

so $G^K \cap \{g \in G |\ \varphi(g) \neq 0\}$ is included in a quasi-compact set. Therefore, for every $l \in C_c(G^{(0)})$, $g \mapsto l(r(g))\varphi(g)$ belongs to $C_c(G)$. It follows that $h(x) = \int_{g\in G^x} \varphi(g)\,\lambda^x(dg)$ is a continuous function. Moreover, for every $x \in G^{(0)}$ there exists $g \in G^x$ such that $\varphi(g) \neq 0$, so $h(x) > 0\ \forall x \in G^{(0)}$. It thus suffices to replace $c(x)$ by $c(x)/h(x)$. □

Example 6.4. *In Example 2.3 with $\Gamma = \mathbb{Z}_n$ and $H = \{0\}$, the cutoff function is the unique continuous extension to X' of the function $c(x) = 1$ for $x \in (0,1]$, and $c(0) = 1/n$.*

Proposition 6.5. *Let G be a locally compact, proper groupoid with Haar system such that $G^{(0)}$ is Hausdorff and $G^{(0)}/G$ is compact. Let c be a cutoff function. Then the function $p(g) = \sqrt{c(r(g))c(s(g))}$ defines a selfadjoint projection $p \in C^*_r(G)$, and $\mathcal{E}(G)$ is isomorphic to $pC^*_r(G)$.*

Proof. Let $\xi_0(x) = \sqrt{c(x)}$. Then one easily checks that $\xi_0 \in \mathcal{E}^0$, $\langle \xi_0, \xi_0\rangle = p$ and $\xi_0\langle\xi_0,\xi_0\rangle = \xi_0$, therefore p is a selfadjoint projection in $C^*_r(G)$. The maps

$$\begin{aligned}\mathcal{E}(G) \to pC^*_r(G), &\qquad \xi \mapsto \langle \xi_0, \xi\rangle = p\langle\xi_0,\xi\rangle\\ pC^*_r(G) \to \mathcal{E}(G), &\qquad a \mapsto \xi_0 a = \xi_0 p a\end{aligned}$$

are inverses from each other. □

7. Generalized morphisms and C^*-algebra correspondences

Until the end of the paper, all groupoids are assumed locally compact, with open range map. In this section, we introduce a notion of generalized morphism for locally compact groupoids which are not necessarily Hausdorff, and a notion of locally proper generalized morphism.
Then, we show that a locally proper generalized morphism from G_1 to G_2 which satisfies an additional condition induces a $C^*_r(G_1)$-module $\mathcal{E}$ and a $*$-morphism $C^*_r(G_2) \to \mathcal{K}(\mathcal{E})$, hence an element of $KK(C^*_r(G_2), C^*_r(G_1))$.

7.1. Generalized morphisms.

Definition 7.1. [4, 5, 8, 9, 12, 14] *Let G_1 and G_2 be two groupoids. A generalized morphism from G_1 to G_2 is a triple (Z,ρ,σ) where*

$$G_1^{(0)} \xleftarrow{\rho} Z \xrightarrow{\sigma} G_2^{(0)},$$

Z is endowed with a left action of G_1 with momentum map ρ and a right action of G_2 with momentum map σ which commute, such that

(a) *the action of G_2 is free and ρ-proper,*
(b) *ρ induces a homeomorphism $Z/G_2 \simeq G_1^{(0)}$.*

In Definition 7.1, one may replace (b) by (b)' or (b)" below:

(b)' ρ is open and induces a bijection $Z/G_2 \to G_1^{(0)}$.
(b)" the map $Z \rtimes G_2 \to Z \times_{G_1^{(0)}} Z$ defined by $(z,\gamma) \mapsto (z, z\gamma)$ is a homeomorphism.

Example 7.2. *Let G_1 and G_2 be two groupoids.If $f\colon G_1 \to G_2$ is a groupoid morphism, let $Z = G_1^{(0)} \times_{f,r} G_2$, $\rho(x,\gamma) = x$ and $\sigma(x,\gamma) = s(\gamma)$. Define the actions of G_1 and G_2 by $g \cdot (x,\gamma) \cdot \gamma' = (r(g), f(g)\gamma\gamma')$. Then (Z,ρ,σ) is a generalized morphism from G_1 to G_2.*

That ρ is open follows from the fact that the range map $G_2 \to G_2^{(0)}$ is open and from Lemma 2.25. The other properties in Definition 7.1 are easy to check.

7.2. Locally proper generalized morphisms.

Definition 7.3. *Let G_1 and G_2 be two groupoidsA generalized morphism from G_1 to G_2 is said to be* locally proper *if the action of G_1 on Z is σ-proper.*

Our terminology is justified by the following proposition:

Proposition 7.4. *Let G_1 and G_2 be two groupoids such that $G_2^{(0)}$ is Hausdorff. Let $f\colon G_1 \to G_2$ be a groupoid morphism. Then the associated generalized groupoid morphism is locally proper if and only if the map $(f,r,s)\colon G_1 \to G_2 \times G_1^{(0)} \times G_1^{(0)}$ is proper.*

Proof. Let $\varphi\colon G_1 \times_{f\circ s,r} G_2 \to (G_2 \times_{s,s} G_2) \times_{r\times r, f\times f} (G_1^{(0)} \times G_1^{(0)})$ defined by $\varphi(g_1,g_2) = (f(g_1)g_2, g_2, r(g_1), s(g_1))$. By definition, the action of G_1 on Z is proper if and only if φ is a proper map. Consider $\theta\colon G_2 \times_{s,s} G_2 \to G_2^{(2)}$ given by $(\gamma,\gamma') = (\gamma(\gamma')^{-1}, \gamma')$. Let $\psi = (\theta \times 1) \circ \varphi$. Since θ is a homeomorphism, the action of G_1 on Z is proper if and only if ψ is proper.
Suppose that (f,r,s) is proper. Let $f' = (f,r,s) \times 1\colon G_1 \times G_2 \to G_2 \times G_1^{(0)} \times G_1^{(0)} \times G_2$. Then f' is proper. Let $F = \{(\gamma, x, x', \gamma') \in G_2 \times G_1^{(0)} \times G_1^{(0)} \times G_2 |\ s(\gamma) = r(\gamma') = f(x'),\ r(\gamma) = f(x)\}$. Then $f'\colon (f')^{-1}(F) \to F$ is proper, i.e. ψ is proper.
Conversely, suppose that ψ is proper. Let $F' = \{(\gamma, y, x, x') \in G_2 \times G_2^{(0)} \times G_1^{(0)} \times G_1^{(0)} |\ s(\gamma) = y\}$. Then $\psi\colon \psi^{-1}(F') \to F'$ is proper, therefore (f,r,s) is proper. $\square$

Our objective is now to show the

Proposition 7.5. *Let G_1, G_2, G_3 be groupoidsLet (Z_1,ρ_1,σ_1) and (Z_2,ρ_2,σ_2) be two generalized groupoid morphisms from G_1 to G_2 and from G_2 to G_3 respectively. Then $(Z,\rho,\sigma) = (Z_1 \times_{G_2} Z_2, \rho_1 \times 1, 1 \times \sigma_2)$ is a generalized groupoid morphism. If (Z_1,ρ_1,σ_1) and (Z_2,ρ_2,σ_2) are locally proper, then (Z,ρ,σ) is locally proper.*

Proposition 7.5 shows that groupoids form a category whose arrows are generalized morphisms, and that two groupoids are isomorphic in that category if and only if they are Morita-equivalent. Moreover, the same conclusions hold for the category whose arrows are locally proper generalized morphisms. In particular, local properness of generalized morphisms is invariant under Morita-equivalence.

All the assertions of Proposition 7.5 follow from Lemma 2.33.

7.3. Proper generalized morphisms.

Definition 7.6. *Let G_1 and G_2 be groupoids. A generalized morphism (Z,ρ,σ) from G_1 to G_2 is said to be* proper *if it is locally proper, and if for every quasi-compact subspace K of $G_2^{(0)}$, $\sigma^{-1}(K)$ is G_1-compact.*

Examples 7.7. (a) *Let X and Y be locally compact spaces and $f\colon X\to Y$ a continuous map. Then the generalized morphism (X,Id,f) is proper if and only if f is proper.*
(b) *Let $f\colon G_1\to G_2$ be a continuous morphism between two locally compact groups. Let $p\colon G_2\to\{*\}$. Then (G_2,p,p) is proper if and only if f is proper and $f(G_1)$ is co-compact in G_2.*
(c) *Let G be a locally compact proper groupoid with Haar system such that $G^{(0)}$ is Hausdorff, and let $\pi\colon G^{(0)}\to G^{(0)}/G$ be the canonical mapping. Then $(G^{(0)},\mathrm{Id},\pi)$ is a proper generalized morphism from G to $G^{(0)}/G$.*

7.4. Construction of a C^*-correspondence. Until the end of the section, our goal is to prove:

Theorem 7.8. *Let G_1 and G_2 be locally compact groupoids with Haar system such that $G_1^{(0)}$ and $G_2^{(0)}$ are Hausdorff, and (Z,ρ,σ) a locally proper generalized morphism from G_1 to G_2. Then one can construct a $C_r^*(G_1)$-Hilbert module $\mathcal{E}_Z$ and a map $\pi\colon C_r^*(G_2)\to\mathcal{L}(\mathcal{E}_Z)$. Moreover, if (Z,ρ,σ) is proper, then π maps to $\mathcal{K}(\mathcal{E}_Z)$. Therefore, it gives an element of $KK(C_r^*(G_2),C_r^*(G_1))$.*

Corollary 7.9. *(see* [14]*) Let G_1 and G_2 be locally compact groupoids with Haar system such that $G_1^{(0)}$ and $G_2^{(0)}$ are Hausdorff. If G_1 and G_2 are Morita-equivalent, then $C_r^*(G_1)$ and $C_r^*(G_2)$ are Morita-equivalent.*

Corollary 7.10. *Let $f\colon G_1\to G_2$ be morphism between two locally compact groupoids with Haar system such that $G_1^{(0)}$ and $G_2^{(0)}$ are Hausdorff. If the restriction of f to $(G_1)_K^K$ is proper for each compact set $K\subset(G_1)^{(0)}$ then f induces a correspondence $\mathcal{E}_f$ from $C_r^*(G_2)$ to $C_r^*(G_1)$. If in addition for every compact set $K\subset G_2^{(0)}$ the quotient of $G_1^{(0)}\times_{f,r}(G_2)_K$ by the diagonal action of G_1 is compact, then $C_r^*(G_2)$ maps to $\mathcal{K}(\mathcal{E}_f)$ and thus f defines a KK-element $[f]\in KK(C_r^*(G_2),C_r^*(G_1))$.*

Proof. See Proposition 7.4 and Definition 7.6 applied to the generalized morphism $Z_f=G_1^{(0)}\times_{f,r}G_2$ as in Example 7.2 $\square$

The rest of the section is devoted to proving Theorem 7.8.

Let us first recall the construction of the correspondence when the groupoids are Hausdorff [11]. It is the closure of $C_c(Z)$ with the $C^*_r(G_1)$-valued scalar product

$$\langle \xi, \eta\rangle(g) = \int_{\gamma\in(G_2)^{\sigma(z)}} \overline{\xi(z\gamma)}\eta(g^{-1}z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma), \tag{2}$$

where z is an arbitrary element of Z such that $\rho(z) = r(g)$. The right $C^*_r(G_1)$-module structure is defined $\forall \xi \in C_c(Z)$, $\forall a \in C_c(G_1)$ by

$$(\xi a)(z) = \int_{g\in(G_1)^{\rho(z)}} \xi(g^{-1}z)a(g^{-1})\,\lambda^{\rho(z)}(\mathrm{d}g), \tag{3}$$

and the left action of $C^*_r(G_2)$ is

$$(b\xi)(z) = \int_{\gamma\in(G_2)^{\sigma(z)}} b(\gamma)\xi(z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma) \tag{4}$$

for all $b \in C_c(G_2)$.

We now come back to non-Hausdorff groupoids. For every open Hausdorff set $V \subset Z$, denote by V' its closure in $\mathcal{H}((G_1 \ltimes Z)^V_V)$, where $z \in V$ is identified to $(\rho(z), z) \in \mathcal{H}((G_1 \ltimes Z)^V_V)$. Let $\mathcal{E}^0_V$ be the set of $\xi \in C_c(V')$ such that $\xi(z) = \dfrac{\xi(S\times\{z\})}{\sqrt{\#S}}$ for all $S \times \{z\} \in V'$.

Lemma 7.11. *The space $\mathcal{E}^0_Z = \sum_{i\in I}\mathcal{E}^0_{V_i}$ is independent of the choice of the cover (V_i) of Z by Hausdorff open subspaces.*

Proof. It suffices to show that for every open Hausdorff subspace V of Z, one has $\mathcal{E}^0_V \subset \sum_{i\in I}\mathcal{E}^0_{V_i}$. Let $\xi \in \mathcal{E}^0_V$. Denote by $q_V\colon V' \to V$ the canonical map defined by $q_V(S\times\{z\}) = z$. Let $K \subset V$ compact such that $\mathrm{supp}(\xi) \subset q_V^{-1}(K)$. There exists $J \subset I$ finite such that $K \subset \cup_{j\in J}V_j$. Let $(\varphi_j)_{j\in J}$ be a partition of unity associated to that cover, and $\xi_j = \xi.(\varphi_j \circ q_V)$. One easily checks that $\xi_j \in \mathcal{E}^0_{V_j}$ and that $\xi = \sum_{j\in J}\xi_j$. □

We now define a $C^*_r(G_1)$-valued scalar product on $\mathcal{E}^0_Z$ by Eqn. (2) where z is an arbitrary element of Z such that $\rho(z) = r(g)$. Our definition is independent of the choice of z, since if z' is another element, there exists $\gamma' \in G_2$ such that $z' = z\gamma'$, and the Haar system on G_2 is left-invariant.
Moreover, the integral is convergent for all $g \in G_1$ because the action of G_2 on Z is proper.
Let us show that $\langle\xi, \eta\rangle \in C_c(G_1)$ for all ξ, $\eta \in \mathcal{E}^0_Z$. We need a preliminary lemma:

Lemma 7.12. *Let X and Y be two topological spaces such that X is locally compact and $f\colon X \to Y$ proper. Let $\mathcal{F}$ be a ultrafilter such that f converges to $y \in Y$ with respect to $\mathcal{F}$. Then there exists $x \in X$ such that $f(x) = y$ and $\mathcal{F}$ converges to x.*

Proof. Let $Q = f^{-1}(y)$. Since f is proper, Q is quasi-compact. Suppose that for all $x \in Q$, $\mathcal{F}$ does not converge to x. Then there exists an open neighborhood V_x of x such that $V_x^c \in \mathcal{F}$. Extracting a finite cover $(V_1, \ldots, V_n)$ of Q, there exists an open neighborhood V of Q such that $V^c \in \mathcal{F}$. Since f is closed, $f(V^c)^c$ is a neighborhood of y. By assumption, $f(V^c)^c \in f(\mathcal{F})$, i.e. $\exists A \in \mathcal{F}$, $f(A) \subset f(V^c)^c$. This implies that $A \subset V$, therefore $V \in \mathcal{F}$: this contradicts $V^c \in \mathcal{F}$.
Consequently, there exists $x \in Q$ such that $\mathcal{F}$ converges to x. ☐

To show that $\langle \xi, \eta \rangle \in C_c(G_1)$, we can suppose that $\xi \in \mathcal{E}_U^0$ and $\eta \in \mathcal{E}_V^0$, where U and V are open Hausdorff. Let $F(g,z) = \overline{\xi(z)}\eta(g^{-1}z)$, defined on $\Gamma = G_1 \times_{r,\rho} Z$. Since the action of G_1 on Z is proper, F is quasi-compactly supported. Let us show that $F \in C_c(\Gamma)$.
Let $\mathcal{F}$ be a ultrafilter on Γ, convergent in $\mathcal{H}\Gamma$. Since $G_1^{(0)}$ is Hausdorff, its limit has the form $S = S'g_0 \times S''$ where $S' \subset (G_1)_{r(g_0)}^{r(g_0)}$, $S'' \subset \rho^{-1}(r(g_0))$. Moreover, S' is a subgroup of $(G_1)_{r(g)}^{r(g)}$ by the proof of Lemma 5.1.
Suppose that there exist $z_0, z_1 \in S''$ and $g_1 \in S'g_0$ such that $z_0 \in U$ and $g_1^{-1}z_1 \in V$. By Lemma 7.12 applied to the proper map $G_1 \rtimes Z \to Z \times Z$, there exists $s_0 \in S'$ such that $z_0 = s_0 z_1$. We may assume that $g_0 = s_0 g_1$. Then $\sum_{s\in S} F(s) = \sum_{s'\in S'} \overline{\xi(z_0)}\eta(g_0^{-1}(s')^{-1}z_0)$. If $s' \notin \mathrm{stab}(z_0)$, then $g_0^{-1}(s')^{-1}z_0 \notin V$ since $g_0^{-1}z_0$ and $g_0^{-1}(s')^{-1}z_0$ are distinct limits of $(g,z) \mapsto g^{-1}z$ with respect to $\mathcal{F}$ and V is Hausdorff. Therefore,

$$\begin{aligned}
\sum_{s\in S} F(s) &= \#(\mathrm{stab}(z_0) \cap S')\overline{\xi(z_0)}\eta(g_0^{-1}z_0) \\
&= \overline{\sqrt{\#(\mathrm{stab}(z_0) \cap S')}\xi(z_0)}\sqrt{\#(\mathrm{stab}(g_0^{-1}z_0) \cap (g_0^{-1}S'g_0))}\eta(z_0) \\
&= \lim_{\mathcal{F}} \overline{\xi(z)}\eta(g^{-1}z) = \lim_{\mathcal{F}} F(g,z).
\end{aligned}$$

If for all z_0, $z_1 \in S''$ and all $g_1 \in S'g_0$, $(z_0, g_1^{-1}z_1) \notin U \times V$, then $\sum_{s\in S} F(g,z) = 0 = \lim_{\mathcal{F}} F(g,z)$.
By Proposition 4.1, $F \in C_c(\Gamma)$.
Since $\langle \xi, \eta \rangle(g) = \int_{\gamma \in (G_2)^{\sigma(z)}} F(g, z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma)$, to prove that $\langle \xi, \eta \rangle \in C_c(G_1)$ it suffices to show:

Lemma 7.13. *Let G_1 and G_2 be two locally compact groupoids with Haar system such that $G_i^{(0)}$ are Hausdorff. Let (Z, ρ, σ) be a generalized morphism from G_1 to G_2. Let $\Gamma = G_1 \times_{r,\rho} Z$. Then for every $F \in C_c(\Gamma)$, the function*

$$g \mapsto \int_{\gamma \in (G_2)^{\sigma(z)}} F(g, z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma),$$

where $z \in Z$ is an arbitrary element such that $\rho(z) = r(g)$, belongs to $C_c(G_1)$.

Proof. Suppose first that $F(g,z) = f(g)h(z)$, where $f \in C_c(G_1)$ and $h \in C_c(Z)$. Let $H(z) = \int_{\gamma \in (G_2)^{\sigma(z)}} h(z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma)$. By Lemma 7.14 below (applied to the groupoid $Z \rtimes G_2$), H is continuous. It is obviously G_2-invariant, therefore $H \in C_c(Z/G_2)$. Let $\tilde{H} \in C_c(G_1^{(0)}) \simeq C_c(Z/G_2)$ correspond to H. The map

$$g \mapsto \int_{\gamma \in (G_2)^{\sigma(z)}} F(g, z\gamma)\, \lambda^{\sigma(z)}(\mathrm{d}\gamma) = f(g)\tilde{H}(s(g))$$

thus belongs to $C_c(G_1)$.

By linearity, the lemma is true for $F \in C_c(G_1) \otimes C_c(Z)$. By Lemma 4.4 and Lemma 4.5, F is the uniform limit of functions $F_n \in C_c(G_1) \otimes C_c(Z)$ which are supported in a fixed quasi-compact set $Q = Q_1 \times Q_2 \subset G_1 \times Z$. Let $Q' \subset Z$ quasi-compact such that $\rho(Q') \supset r(Q_1)$. Since the action of G_2 on Z is proper, $K = \{\gamma \in G_2 |\ Q'\gamma \cap Q_2 \neq \emptyset\}$ is quasi-compact. Using the fact that $G_1^{(0)} \simeq Z/G_2$, it is easy to see that

$$\sup_{(g,z) \in \Gamma} \int_{\gamma \in (G_2)^{\sigma(z)}} 1_Q(g, z\gamma)\, \lambda^{\sigma(z)}(\mathrm{d}\gamma) \le \sup_{z \in Q'} \int_{\gamma \in G_2^{\sigma(z)}} 1_{Q_2}(z\gamma) \lambda^{\sigma(z)}(\mathrm{d}\gamma)$$

$$\le \sup_{x \in G_2^{(0)}} \int_{\gamma \in G_2^x} 1_K(\gamma) \lambda^x(\mathrm{d}\gamma) < \infty$$

by Lemma 4.7. Therefore,

$$\lim_{n \to \infty} \sup_{g \in G_1} \left| \int_{\gamma \in G_2^{\sigma(z)}} F(g, z\gamma) - F_n(g, z\gamma)\, \lambda^{\sigma(z)}(\mathrm{d}\gamma) \right| = 0.$$

The conclusion follows from Corollary 4.2. □

In the proof of Lemma 7.13 we used the

LEMMA 7.14. *Let G be a locally compact, proper groupoid with Haar system, such that G^x is Hausdorff for all $x \in G^{(0)}$, and $G_x^x = \{x\}$ for all $x \in G^{(0)}$. We do not assume $G^{(0)}$ to be Hausdorff. Then $\forall f \in C_c(G^{(0)})$,*

$$\varphi \colon G^{(0)} \to \mathbb{C}, \quad x \mapsto \int_{g \in G^x} f(s(g))\, \lambda^x(dg)$$

is continuous.

Proof. Let V be an open, Hausdorff subspace of $G^{(0)}$. Let $h \in C_c(V)$. Since $(r, s) \colon G \to G^{(0)} \times G^{(0)}$ is a homeomorphism from G onto a closed subspace of $G^{(0)} \times G^{(0)}$, and $(x, y) \mapsto h(x)f(y)$ belongs to $C_c(G^{(0)} \times G^{(0)})$, the map $g \mapsto h(r(g))f(s(g))$ belongs to $C_c(G)$, therefore by definition of a Haar system, $x \mapsto \int_{g \in G^x} h(r(g))f(s(g))\, \lambda^x(dg) = h(x)\varphi(x)$ belongs to $C_c(G^{(0)})$.

Since $h \in C_c(V)$ is arbitrary, this shows that $\varphi_{|V}$ is continuous, hence φ is continuous on $G^{(0)}$. □

Now, let us show the positivity of the scalar product. Recall that for all $x \in G_1^{(0)}$ there is a representation $\pi_{G_1,x} \colon C^*(G_1) \to \mathcal{L}(L^2(G_1^x))$ such that for all $a \in C_c(G_1)$ and all $\eta \in C_c(G_1^x)$,

$$(\pi_{G_1,x}(a)\eta)(g) = \int_{h \in G_1^{s(g)}} a(h)\eta(gh)\, \lambda^{s(g)}(\mathrm{d}h).$$

By definition, $\|a\|_{C_r^*(G_1)} = \sup_{x\in G_1^{(0)}} \|\pi_{G_1,x}(a)\|$.

$$\begin{aligned}\langle \eta, \pi_{G_1,x}(a)\eta\rangle &= \int_{g\in G_1^x,\ h\in G_1^{s(g)}} \overline{\eta(g)}a(h)\eta(gh)\,\lambda^{s(g)}(\mathrm{d}h)\lambda^x(dg)\\ &= \int_{g\in G_1^x,\ h\in G^{s(g)}} \overline{\eta(g)}a(g^{-1}h)\eta(h)\,\lambda^x(dg)\lambda^x(dh).\end{aligned}$$

Fix $z\in Z$ such that $\rho(z)=x$. Replacing $a(g^{-1}h)$ by

$$\langle \xi,\xi\rangle(g^{-1}h) = \int_{\gamma\in G_2^{\sigma(z)}} \overline{\xi(g^{-1}z\gamma)}\xi(h^{-1}z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma),$$

we get

$$\langle \eta, \pi_{G_1,x}(\langle\xi,\xi\rangle)\eta\rangle = \int_{\gamma\in G_2^{\sigma(z)}} \lambda^{\sigma(z)}(\mathrm{d}\gamma)\left|\int_{g\in G^x} \eta(g)\xi(g^{-1}z\gamma)\,\lambda^x(dg)\right|^2. \tag{5}$$

It follows that $\pi_{G_1,x}(\langle\xi,\xi\rangle)\geq 0$ for all $x\in G_1^{(0)}$, so $\langle\xi,\xi\rangle\geq 0$ in $C_r^*(G_1)$.

Now, let us define a $C_r^*(G_1)$-module structure on $\mathcal{E}_Z^0$ by Eqn.(3) for all $\xi\in\mathcal{E}_Z^0$ and $a\in C_c(G_1)$.
Let us show that $\xi a\in\mathcal{E}_Z^0$. We need a preliminary lemma:

Lemma 7.15. *Let X and Y be quasi-compact spaces, (Ω_k) an open cover of $X\times Y$. Then there exist finite open covers (X_i) and (Y_j) of X and Y such that $\forall i,j\ \exists k$, $X_i\times Y_j\subset\Omega_k$.*

Proof. For all $(x,y)\in X\times Y$ choose open neighborhoods $U_{x,y}$ and $V_{x,y}$ of x and y such that $U_{x,y}\times V_{x,y}\subset\Omega_k$ for some k. For y fixed, there exist $x_1,\ldots,x_n$ such that $(U_{x_i,y})_{1\leq i\leq n}$ covers X. Let $V_y=\cap_{i=1}^n U_{x_i,y}$. Then for all $(x,y)\in X\times Y$, there exists an open neighborhood $U'_{x,y}$ of x and k such that $U'_{x,y}\times V_y\subset\Omega_k$. Let $(V_1,\ldots,V_m)=(V_{y_1},\ldots,V_{y_m})$ such that $\cup_{1\leq j\leq m}V_j=Y$. For all $x\in X$, let $U'_x=\cap_{j=1}^m U'_{x,y_j}$. Let $(U_1,\ldots,U_p)$ be a finite sub-cover of $(U'_x)_{x\in X}$. Then for all i and for all j, there exists k such that $U_i\times V_j\subset\Omega_k$. □

Let Q_1 and Q_2 be quasi-compact subspaces of G_1 of Z respectively such that $a^{-1}(\mathbb{C}^*)\subset Q_1$ and $\xi^{-1}(\mathbb{C}^*)\subset Q_2$. Let Q be a quasi-compact subspace of Z such that $\forall g\in Q_1$, $\forall z\in Q_2$, $g^{-1}z\in Q$. Let (U_k) be a finite cover of Q by Hausdorff open subspaces of Z. Let $Q'=Q_1\times_{r,\rho}Q_2$. Then Q' is a closed subspace of $Q_1\times Q_2$. Let $\Omega'_k=\{(g,z)\in Q' |\ g^{-1}z\in U_k\}$. Then (Ω'_k) is a finite open cover of Q'. Let Ω_k be an open subspace of $Q_1\times Q_2$ such that $\Omega'_k=\Omega_k\cap Q'$. Then $\{Q_1\times Q_2-Q'\}\cup\{\Omega_k\}$ is an open cover of $Q_1\times Q_2$. Using Lemma 7.15, there exist finite families of Hausdorff open sets (W_i) and (V_j) which cover Q_1 and Q_2, such that for all i, j and for all $(g,z)\in W_i\times_{G_1^{(0)}}V_j$, there exists k such that $g^{-1}z\in U_k$.
Thus, we can assume by linearity and by Lemmas 4.3 and 7.11 that $\xi\in\mathcal{E}_V^0$, $a\in C_c(W)$, $U=W^{-1}V$, and U, V and W are open and Hausdorff.

Let $\Omega = \{(g,S) \in W^{-1} \times U' |\ g^{-1}q_U(S) \in V\}$. Then the map $(g,S) \mapsto (g^{-1}, g^{-1}S)$ is a homeomorphism from Ω onto $W \times_{r,\rho\circ q_V} V'$. Therefore, the map $(g,z) \mapsto \xi(g^{-1}z)a(g^{-1})$ belongs to $C_c(\Omega) \subset C_c(G_1 \times_{r,\rho\circ q_V} U')$. By Lemma 4.8,

$$S \mapsto (\xi a)(S) = \int_{g \in G_1^{\rho\circ q_V(S)}} \xi(g^{-1}S)a(g^{-1})\,\lambda^{\rho\circ q_V(S)}(\mathrm{d}g)$$

belongs to $C_c(U')$. It is immediate that $(\xi a)(S) = \sqrt{\#S}(\xi a)(q(S))$ for all $S \in U'$, therefore $\xi a \in \mathcal{E}_U^0$. This completes the proof that $\xi a \in \mathcal{E}_Z^0$.

Finally, it is not hard to check that $\langle \xi, \eta a\rangle = \langle \xi, \eta\rangle * a$. Therefore, the completion $\mathcal{E}_Z$ of $\mathcal{E}_Z^0$ with respect to the norm $\|\xi\| = \|\langle \xi,\xi\rangle\|^{1/2}$ is a $C_r^*(G_1)$-Hilbert module.

Let us now construct a morphism $\pi\colon C_r^*(G_2) \to \mathcal{L}(\mathcal{E}_Z)$. For every $\xi \in \mathcal{E}_Z^0$ and every $b \in C_c(G_2)$, define $b\xi$ by Eqn.(4). Let us check that $b\xi \in \mathcal{E}_Z^0$. As above, by linearity we may assume that $\xi \in \mathcal{E}_V^0$, $b \in C_c(W)$ and $VW^{-1} \subset U$, where $V \subset Z$, $U \subset Z$ and $W \subset G_2$ are open and Hausdorff.
Let $\Phi(S,\gamma) = (S\gamma, \gamma)$. Then Φ is a homeomorphism from $\Omega = \{(S,\gamma) \in U' \times_{\sigma\circ q_U, r} W |\ q_U(S)\gamma \in V\}$ onto $V' \times_{\sigma\circ q_V, s} W$. Let $F(z,\gamma) = b(\gamma)\xi(z\gamma)$. Since $F = (\xi \otimes b)\circ\Phi$, F is an element of $C_c(\Omega) \subset C_c(U' \times_{\sigma\circ q_U, r} W)$. By Lemma 4.8, $b\xi \in C_c(U')$.
It is immediate that $(b\xi)(S) = \sqrt{\#S}(b\xi)(q(S))$. Therefore, $b\xi \in \mathcal{E}_U^0 \subset \mathcal{E}_Z^0$.
Let us prove that $\|b\xi\| \le \|b\|\,\|\xi\|$. Let

$$\zeta(\gamma) = \int_{g\in G_1^x} \eta(g)\xi(g^{-1}z\gamma)\,\lambda^x(dg),$$

where $z \in Z$ such that $\rho(z) = r(g)$ is arbitrary. From (5),

$$\langle \eta, \pi_{G_1,x}(\langle \xi,\xi\rangle)\eta\rangle = \|\zeta\|^2_{L^2(G_2^{\sigma(z)})}.$$

A similar calculation shows that

$$\begin{aligned}\langle \eta, \pi_{G_1,x}(\langle b\xi,b\xi\rangle)\eta\rangle &= \int_{\gamma\in G_2^{\sigma(z)}} \lambda^{\sigma(z)}(\mathrm{d}\gamma)\left|\int_{g\in G_1^x}\eta(g)\xi(g^{-1}z\gamma\gamma')b(\gamma')\,\lambda^{s(\gamma)}(\mathrm{d}\gamma')\right|^2\\ &= \langle b\zeta, b\zeta\rangle \le \|b\|^2\|\zeta\|^2.\end{aligned}$$

By density of $C_c(G_2^x)$ in $L^2(G_2^x)$, $\|\pi_{G_1,x}(\langle b\xi,b\xi\rangle)\| \le \|b\|^2\|\pi_{G_1,x}(\langle\xi,\xi\rangle)\|$. Taking the supremum over $x \in G_1^{(0)}$, we get $\|b\xi\| \le \|b\|\,\|\xi\|$. It follows that $b \mapsto (\xi \mapsto b\xi)$ extends to a $*$-morphism $\pi\colon C_r^*(G_2) \to \mathcal{L}(\mathcal{E}_Z)$.

Finally, suppose now that (Z,ρ,σ) is proper, and let us show that $C_r^*(G_2)$ maps to $\mathcal{K}(\mathcal{E}_Z)$.
For every η, $\zeta \in \mathcal{E}_Z^0$, denote by $T_{\eta,\zeta}$ the operator $T_{\eta,\zeta}(\xi) = \eta\langle\zeta,\xi\rangle$. Compact operators are elements of the closed linear span of $T_{\eta,\zeta}$'s. Let us write an explicit formula for $T_{\eta,\zeta}$:

$$\begin{aligned}T_{\eta,\zeta}(\xi)(z) &= \int_{g\in G_1^{\rho(z)}} \eta(g^{-1}z)\langle\zeta,\xi\rangle(g^{-1})\,\lambda^{\rho(z)}(\mathrm{d}g)\\ &= \int_{g\in G_1^{\rho(z)}} \eta(g^{-1}z)\int_{\gamma\in G_2^{\sigma(z)}} \overline{\zeta(g^{-1}z\gamma)}\xi(z\gamma)\,\lambda^{\sigma(z)}(\mathrm{d}\gamma)\lambda^{\rho(z)}(\mathrm{d}g).\end{aligned}$$

Let $b \in C_c(G_2)$, let us show that $\pi(b) \in \mathcal{K}(\mathcal{E}_Z)$. Let K be a quasi-compact subspace of G_2 such that $b^{-1}(\mathbb{C}^*) \subset K$. Since (Z, ρ, σ) is a proper generalized morphism, there exists a quasi-compact subspace Q of Z such that $\sigma^{-1}(r(K)) \subset G_1\mathring{Q}$. Before we proceed, we need a lemma:

Lemma 7.16. *Let G_2 be a locally compact groupoid acting freely and properly on a locally compact space Z with momentum map $\sigma\colon Z \to G_2^{(0)}$. Then for every $(z_0, \gamma_0) \in Z \rtimes G_2$, there exists a Hausdorff open neighborhood Ω_{z_0,γ_0} of (z_0, γ_0) such that*

- *$U = \{z_1\gamma_1 | (z_1, \gamma_1) \in \Omega_{z_0,\gamma_0}\}$ is Hausdorff;*
- *there exists a Hausdorff open neighborhood W of γ_0 such that $\forall \gamma \in G_2$, $\forall z \in pr_1(\Omega_{z_0,\gamma_0})$, $\forall z' \in U$, $z' = z\gamma \implies \gamma \in W$.*

Proof. Let $R = \{(z, z') \in Z \times Z | \exists \gamma \in G_2,\ z' = z\gamma\}$. Since the G_2-action is free and proper, there exists a continuous function $\phi\colon R \to G_2$ such that $\phi(z, z\gamma) = \gamma$. Let W be an open Hausdorff neighborhood of γ_0. By continuity of ϕ, there exist open Hausdorff neighborhoods V and U_0 of z_0 and $z_0\gamma_0$ such that for all $(z, z') \in R \cap (V \times U_0)$, $\phi(z, z') \in W$. By continuity of the action, there exists an open neighborhood Ω_{z_0,γ_0} of (z_0, γ_0) such that $\forall (z_1, \gamma_1) \in \Omega_{z_0,\gamma_0}$, $z_1\gamma_1 \in U_0$ and $z_1 \in V$. ∎

By Lemma 7.15, there exist finite covers (V_i) of Q and (W_j) of K such that for every i, j, $(Z \times_{G_2^{(0)}} G_2) \cap (V_i \times W_j) \subset \Omega_{z_0,\gamma_0}$ for some (z_0, γ_0).

By Lemma 6.2 applied to the groupoid $(G_1 \ltimes Z)_{V_i}^{V_i}$, for all i there exists $c_i' \in C_c(V_i')_+$ such that $c_i'(S) = (\#S)c_i'(q_{V_i}(S))$ for all $S \in V_i'$, and such that $\sum_i c_i' \geq 1$ on Q. Let

$$f_i(z) = \int_{g \in G_1^{\rho(z)}} c_i'(g^{-1}z)\, \lambda^{\rho(z)}(\mathrm{d}g)$$

and let $f = \sum_i f_i$. As in the proof of Theorem 6.3, one can show that for every Hausdorff open subspace V of Z and every $h \in C_c(V)$, $(g, z) \mapsto h(z)c_i'(g^{-1}z)$ belongs to $C_c(G \ltimes Z)$, therefore hf_i is continuous on V. Since h is arbitrary, it follows that f_i is continuous, thus f is continuous. Moreover, f is G_1-equivariant, nonnegative, and $\inf_Q f > 0$. Therefore, there exists $f_1 \in C_c(G_1 \backslash Z)$ such that $f_1(z) = 1/f(z)$ for all $z \in Q$. Let $c_i(z) = f_1(z)c_i'(z)$. Let

$$T_i(\xi)(z) = \int_{g \in G_1^{\rho(z)}} \int_{\gamma \in G_2^{\sigma(z)}} c_i(g^{-1}z) b(\gamma) \xi(z\gamma)\, \lambda^{\rho(z)}(\mathrm{d}g) \lambda^{\sigma(z)}(\mathrm{d}\gamma).$$

Then $\pi(b) = \sum_i T_i$, therefore it suffices to show that T_i is a compact operator for all i.

By linearity and by Lemma 4.3, one may assume that $b \in C_c(W_j)$ for some j. Then, by construction of V_i (see Lemma 7.16), there exist open Hausdorff sets $U \subset Z$ and $W \subset G_2$ such that $\{\gamma \in G_2 | \exists (z, z') \in V_i \times U,\ z' = z\gamma\} \subset W$, and $\{z\gamma | (z, \gamma) \in V_i \times_{\sigma,r} W\} \subset U$.

The map $(z, z\gamma) \mapsto c(z)b(\gamma)$ defines an element of $C_c(V_i' \times U)$. Let $L_1 \times L_2 \subset V_i \times U$ compact such that $(z, z\gamma) \mapsto c(z)b(\gamma)$ is supported on $q_{V_i}^{-1}(L_1) \times L_2$.

By Lemma 6.2 applied to the groupoids $(G_1 \ltimes Z)_{V_i}^{V_i}$ and $(G_1 \ltimes Z)_U^U$, there exist $d_1 \in C_c(V_i')_+$ and $d_2 \in C_c(U')_+$ such that $d_1 > 0$ on L_1 and $d_2 > 0$ on L_2, $d_1(S) = \sqrt{\#S} d_1(q_{V_i}(S))$ for all $S \in V_i'$, and $d_2(S) = \sqrt{\#S} d_2(q_U(S))$ for all $S \in U'$. Let

$$f(z, z\gamma) = \frac{c(z)b(\gamma)}{d_1(z)d_2(z\gamma)}.$$

Then $f \in C_c(V_i \times_{G_1^{(0)}} U)$. Therefore, f is the uniform limit of a sequence $f_n = \sum \alpha_{n,k} \otimes \overline{\beta_{n,k}}$ in $C_c(V_i) \otimes C_c(U)$ such that all the f_n are supported in a fixed compact set. Then T_i is the norm-limit of $\sum_k T_{d_1\alpha_{n,k}, d_2\beta_{n,k}}$, therefore it is compact.

Remark 7.17. *The construction in Theorem 7.8 is functorial with respect to the composition of generalized morphisms and of correspondences. We don't include a proof of this fact, as it is tedious but elementary. It is an easy exercise when G_1 and G_2 are Hausdorff.*

References

[1] Abels, H. Parallelizability of proper actions, global K-slices and maximal compact subgroups. *Math. Ann.* 212 (1974/75), 1–19.

[2] Bourbaki, N.: *Topologie Générale.* Masson, Paris (1990).

[3] Chabert, J.; Echterhoff, S.; Meyer, R. Deux remarques sur l'application de Baum-Connes. *C. R. Acad. Sci. Paris Sér. I Math.* 332 (2001), no. 7, 607–610.

[4] Haefliger, A. *Groupoïdes d'holonomie et classifiants*, Astérisque 116 (1984), 70–97.

[5] Hilsum, M. and Skandalis, G.: Morphismes K-orientés d'espaces de feuilles et fonctorialité en théorie de Kasparov (d'après une conjecture d'A. Connes). *Ann. Sci. École Norm. Sup.* (4) 20(1987), no. 3, 325–390.

[6] Khoshkam, M.; Skandalis, G. Regular representation of groupoid C^*-algebras and applications to inverse semigroups. *J. Reine Angew. Math.* 546 (2002), 47–72.

[7] Landsman, N. P. Operator algebras and Poisson manifolds associated to groupoids. *Comm. Math. Phys.* 222 (2001), no. 1, 97–116.

[8] Landsman, N. P. Quantized reduction as a tensor product. Quantization of singular symplectic quotients, 137–180, *Progr. Math.*, 198, Birkhäuser, Basel, 2001.

[9] Le Gall, P.Y.: Théorie de Kasparov équivariante et groupoïdes. I. *K-Theory* 16(1999), no. 4, 361–390.

[10] Le Gall, P.Y.; Monthubert, B. K-theory of the indicial algebra of a manifold with corners. *K-Theory* 23 (2001), no. 2, 105–113.

[11] Macho-Stadler, M. and O'Uchi, M.: Correspondence of groupoid C^*-algebras, *J. Operator Theory* 42(1999), 103-119.

[12] Moerdijk, I. *Toposes and groupoids*, in *Categorical algebra and its applications*, 280–298, LNM 1348, Springer, Berlin, 1988.

[13] Mrčun, J., Functoriality of the bimodule associated to a Hilsum-Skandalis map, *K-Theory* 18 (1999), 235–253.
[14] Muhly, P., Renault, J. and Williams, D.: Equivalence and isomorphism for groupoid C^*-algebras, *J. Operator Theory* 17(1987), 3–22.
[15] Paterson, A: *Groupoids, inverse semigroups, and their operator algebras.* Progress in Mathematics, 170. Birkhäuser Boston, Inc., Boston, MA, 1999.
[16] Renault, J: *A groupoid approach to C^*-algebras*, Springer, 1980, LNM no. 793.
[17] Stachura, P. C^*-algebra of a differential groupoid. With an appendix by S. Zakrzewski. Banach Center Publ., 51, Poisson geometry (Warsaw, 1998), 263–281, Polish Acad. Sci., Warsaw, 2000.

Jean-Louis Tu
University Paris VI
Institut de Mathématiques
175, rue du Chevaleret
75013 Paris, France.
tu@math.jussieu.fr

The Centre of Completed Group Algebras of Pro-p Groups

Konstantin Ardakov

Received: December 13, 2004

Communicated by Peter Schneider

Abstract. We compute the centre of the completed group algebra of an arbitrary countably based pro-p group with coefficients in $\mathbb{F}_p$ or $\mathbb{Z}_p$. Some other results are obtained.

2000 Mathematics Subject Classification: 16S34, 20E18

1 Introduction

Let G be a pro-p group. In this paper we investigate some rings related to the completed group algebra of G over $\mathbb{F}_p$, which we denote by Ω_G:

$$\Omega_G = \mathbb{F}_p[[G]] := \varprojlim_{N \lhd_o G} \mathbb{F}_p[G/N].$$

When G is analytic in the sense of [3], Ω_G and its p-adic analogue Λ_G defined by

$$\Lambda_G = \mathbb{Z}_p[[G]] := \varprojlim_{N \lhd_o G} \mathbb{Z}_p[G/N]$$

are right and left Noetherian rings, which are in general noncommutative. If in addition G is torsion free, the results of Brumer, Neumann and others show that Ω_G and Λ_G have finite global dimension and have no zero-divisors; for an overview, see [1]. Moreover, under the name of Iwasawa algebras, these rings are frequently of interest to number theorists (see [2] for more details).
Our main result is

Theorem A. *Let G be a countably based pro-p group. Then the centre of Ω_G is equal to the closure of the centre of $\mathbb{F}_p[G]$:*

$$Z(\Omega_G) = \overline{Z(\mathbb{F}_p[G])}.$$

Similarly, the centre of Λ_G is equal to the closure of the centre of $\mathbb{Z}_p[G]$:

$$Z(\Lambda_G) = \overline{Z(\mathbb{Z}_p[G])}.$$

When G is p-valued in the sense of Lazard [5, III.2.1.2], we obtain a cleaner result.

Corollary A. *Let G be a countably based p-valued pro-p group with centre Z. Then*

$$Z(\Omega_G) = \Omega_Z \quad \textit{and} \quad Z(\Lambda_G) = \Lambda_Z.$$

The class of p-valued pro-p groups is rather large; for example, every closed subgroup of a uniform pro-p group [3, 4,1] is p-valued. Also, any pro-p subgroup of $GL_n(\mathbb{Z}_p)$ is p-valued when $p > n+1$ [5, p. 101].

We remark that when G is an open pro-p subgroup of $\mathrm{GL}_2(\mathbb{Z}_p)$, a version of the above result was proved by Howson [4, 4.2] using similar techniques.

We also use the method used in the proof of Theorem A to compute endomorphism rings of certain induced modules for Ω_G, when G is an analytic pro-p group.

Theorem B. *Let H be a closed subgroup of an analytic pro-p group G. Let $M = \mathbb{F}_p \otimes_{\Omega_H} \Omega_G$ and write $R = \mathrm{End}_{\Omega_G}(M)$. Then R is finite-dimensional over $\mathbb{F}_p$ if and only if $N_{\mathfrak{g}}(\mathfrak{h}) = \mathfrak{h}$, where $\mathfrak{h}$ and $\mathfrak{g}$ denote the $\mathbb{Q}_p$-Lie algebras of H and G, respectively.*

The author would like to thank Chris Brookes and Simon Wadsley for many valuable discussions. This research was financially supported by EPSRC grant number 00802002.

2 Fixed Points

Let X be a group. For any (right) X-space S, let $S^X = \{s \in S : s.X = s\}$ denote the set of fixed points of X in S. Also, let $\mathcal{O}(S)$ denote the collection of all *finite* X-orbits on S, and for any orbit $\mathcal{C} \in \mathcal{O}(S)$, let $\hat{\mathcal{C}}$ denote the orbit sum

$$\hat{\mathcal{C}} = \sum_{s \in \mathcal{C}} s,$$

viewed as an element of the permutation module $\mathbb{F}_p[S]$. Thus, $\mathbb{F}_p[S]^X$ is spanned by all the $\hat{\mathcal{C}}$ as $\mathcal{C}$ ranges over $\mathcal{O}(S)$: $\mathbb{F}_p[S]^X = \mathbb{F}_p[\hat{\mathcal{O}(S)}]$.
Now let X be a pro-p group. Assume we are given an inverse system

$$\ldots \overset{\pi_{n+1}}{\twoheadrightarrow} A_n \overset{\pi_n}{\twoheadrightarrow} A_{n-1} \overset{\pi_{n-1}}{\twoheadrightarrow} \ldots \overset{\pi_2}{\twoheadrightarrow} A_1$$

of finite X-spaces. We can consider the natural inverse system of permutation modules associated with the A_i:

$$\ldots \overset{\pi_{n+1}}{\twoheadrightarrow} \mathbb{F}_p[A_n] \overset{\pi_n}{\twoheadrightarrow} \mathbb{F}_p[A_{n-1}] \overset{\pi_{n-1}}{\twoheadrightarrow} \ldots \overset{\pi_2}{\twoheadrightarrow} \mathbb{F}_p[A_1]$$

where we keep the same notation for the connecting maps π_n. Now, form the inverse limit

$$Y = \varprojlim A_n,$$

this is clearly an X-space. We can also form the inverse limit

$$\Omega_Y = \varprojlim \mathbb{F}_p[A_n],$$

which is easily seen to be an Ω_X-module.
Note that Ω_Y is a compact metric space, with metric given by $d(\alpha, \beta) = \|\alpha - \beta\|$, where $\|.\|$ is a norm on Ω_Y given by

$$\|(\alpha_n)_n\| = \sup\{p^{-n} : \alpha_n \neq 0\} \quad \text{and} \quad \|0\| = 0.$$

We are interested in the fixed points of Ω_Y, viewed as an X-space. It is straightforward to see that there is a natural embedding of $\mathbb{F}_p[Y]$ into Ω_Y and that $\mathbb{F}_p[Y]^X \subseteq \Omega_Y^X$.

Proposition 2.1. *With the notations above,* $\Omega_Y^X = \overline{\mathbb{F}_p[Y]}^X = \overline{\mathbb{F}_p[Y]^X}$.

Proof. Because the action of X on Ω_Y is continuous, it is clear that Ω_Y^X is a closed subset of Ω_Y, so by the above remarks $\overline{\mathbb{F}_p[Y]^X} \subseteq \Omega_Y^X$. Let $\alpha = (\alpha_n)_n \in \Omega_Y^X$. Since the natural maps $\Omega_Y \twoheadrightarrow \mathbb{F}_p[A_n]$ are maps of X-spaces, we see that each α_n lies in $\mathbb{F}_p[A_n]^X$.

Let the integer r be least such that $\alpha_r \neq 0$. Consider $\alpha_r \in \mathbb{F}_p[A_r]^X$; thus $\alpha_r = \sum_{\mathcal{C}\in\mathcal{O}(A_r)} \lambda_{\mathcal{C}}\hat{\mathcal{C}}$ and not all the $\lambda_{\mathcal{C}}$ are zero.

Pick a $\mathcal{C} \in \mathcal{O}(A_r)$ with $\lambda_{\mathcal{C}} \neq 0$. Since π_{r+1} is a map of X-spaces, $\pi_{r+1}^{-1}(\mathcal{C}) = \mathcal{D}_1 \cup \mathcal{D}_2 \cup \ldots \cup \mathcal{D}_k$ is a union of X-orbits, with $\pi_{r+1}(\mathcal{D}_j) = \mathcal{C}$ for $j = 1, \ldots, k$ and $\pi_{r+1}(\mathcal{D}_j) \cap \mathcal{C} = \emptyset$ for $j > k$, if we let $\mathcal{D}_{k+1}, \ldots, \mathcal{D}_m$ denote the remaining elements of $\mathcal{O}(A_{r+1})$.

We claim we can find a $\mathcal{D}_j$ with $1 \leq j \leq k$ such that $|\mathcal{D}_j| = |\mathcal{C}|$.

For, suppose not. Then $|\mathcal{D}_j| > |\mathcal{C}|$ for each $j = 1, \ldots, k$. As $\pi_{r+1} : \mathcal{D}_j \twoheadrightarrow \mathcal{C}$ is a surjective map of finite transitive X-spaces, and because X is a pro-p group, we deduce that each fibre $(\pi_{r+1}|\mathcal{D}_j)^{-1}(s)$ for $s \in \mathcal{C}$ has size a power of p greater than 1. But then, because we are working over $\mathbb{F}_p$, we must have $\pi_{r+1}(\hat{\mathcal{D}}_j) = 0$, for each $1 \leq j \leq k$.

Now, since $\alpha_{r+1} \in \mathbb{F}_p[A_{r+1}]^X$, we can write $\alpha_{r+1} = \sum_{j=1}^m \mu_j \hat{\mathcal{D}}_j$ for some $\mu_j \in \mathbb{F}_p$. So, $\alpha_r = \pi_{r+1}(\alpha_{r+1}) = \sum_{j=k+1}^m \mu_j \pi_{r+1}(\hat{\mathcal{D}}_j)$. But $\pi_{r+1}(\mathcal{D}_j) \cap \mathcal{C} = \emptyset$ for all $j > k$, contradicting the fact that $\mathcal{C} \subseteq \mathrm{supp}(\alpha_\mathrm{r})$.

Hence, we can find $\mathcal{C}_{r+1} \in \mathcal{O}(A_{r+1})$ with $|\mathcal{C}_{r+1}| = |\mathcal{C}_r|$ and $\pi_{r+1}(\mathcal{C}_{r+1}) = \mathcal{C}_r$, where we set $\mathcal{C}_r$ to be $\mathcal{C}$. It is clear that we can continue this process of "lifting" the X-orbits, without ever increasing the sizes. Thus, we get a sequence

$$\cdots \overset{\pi_{n+2}}{\twoheadrightarrow} \mathcal{C}_{n+1} \overset{\pi_{n+1}}{\twoheadrightarrow} \mathcal{C}_n \overset{\pi_n}{\twoheadrightarrow} \ldots \overset{\pi_{r+1}}{\twoheadrightarrow} \mathcal{C}_r$$

of X-orbits, each having the same size as $\mathcal{C}_r$.

Now, pick some $s_r \in \mathcal{C}_r$ and inductively choose lifts $s_n \in \mathcal{C}_n$ for each $n \geq r$. Let s be the element of Y determined by these lifts. It is then straightforward to see that the X-orbit of s in Y is finite and that the image of this orbit in A_r equals $\mathcal{C}$. Let $\mathcal{F}_{\mathcal{C}}$ denote this element of $\mathcal{O}(Y)$.

Finally, we can consider the element $\beta = \sum_{\mathcal{C}\in\mathcal{O}(A_r)} \lambda_{\mathcal{C}}\hat{\mathcal{F}}_{\mathcal{C}}$. Obviously β lies in $\mathbb{F}_p[Y]^X$, and the image of β in $\mathbb{F}_p[A_r]$ coincides with

α_r. Hence, $\alpha - \beta$ has norm strictly smaller than that of α and also lies in Ω_Y^X. Applying the argument above to $\alpha - \beta$ instead of α and iterating, we see that α can be approximated arbitrarily closely by elements of $\mathbb{F}_p[Y]^X$. $\square$

Next we turn to the analogous proposition over the p-adics. Let

$$\Lambda_Y = \varprojlim \mathbb{Z}_p[A_n].$$

This is naturally a Λ_X-module and there is a natural isomorphism $\Lambda_Y/p\Lambda_Y \cong \Omega_Y$ of Λ_X-modules. Moreover, since $\Lambda_Y \cong \varprojlim(\mathbb{Z}/p^n\mathbb{Z})[A_n]$ is a countably based pro-p group, it is a compact metric space.

Proposition 2.2. *With the notations above, $\Lambda_Y^X = \overline{\mathbb{Z}_p[Y]^X}$.*

Proof. As in the proof of Proposition 2.1, the inclusion $\overline{\mathbb{Z}_p[Y]^X} \subseteq \Lambda_Y^X$ is clear. Let $^-\colon \Lambda_Y \to \Omega_Y$ denote reduction mod p.
Let $\alpha \in \Lambda_Y^X$ so that $\overline{\alpha} \in \Omega_Y^X$. By Proposition 2.1, $\overline{\alpha} = \lim_{n\to\infty} u_n$ for some $u_n \in \mathbb{F}_p[Y]^X$. Since $\mathbb{F}_p[Y]^X = \mathbb{F}_p[\hat{\mathcal{O}(Y)}]$ and $\mathbb{Z}_p[Y]^X = \mathbb{Z}_p[\hat{\mathcal{O}(Y)}]$, we can choose $v_n \in \mathbb{Z}_p[Y]^X$ such that $\overline{v_n} = u_n$ for all n. Since Λ_Y is compact, by passing to a convergent subsequence we may assume that v_n converges to $\beta_0 \in \overline{\mathbb{Z}_p[Y]^X}$. Now $\overline{\alpha} = \lim_{n\to\infty} \overline{v_n} = \overline{\beta_0}$, so $\alpha - \beta_0 \in p\Lambda_Y \cap \Lambda_Y^X = p\Lambda_Y^X$, since Λ_Y is p-torsion free.
Hence we can write $\alpha = \beta_0 + p\alpha_1$ where $\alpha_1 \in \Lambda_Y^X$. Iterating the above argument, we obtain elements $\beta_1, \beta_2, \ldots \in \overline{\mathbb{Z}_p[Y]^X}$ and $\alpha_1, \alpha_2, \ldots \in \Lambda_Y^X$ such that $\alpha_n = \beta_n + p\alpha_{n+1}$ for all $n \geq 1$. So $\alpha = \sum_{n=0}^{\infty} \beta_n p^n \in \overline{\mathbb{Z}_p[Y]^X}$. $\square$

3 Main Results

We immediately make use of the above Propositions.

Proof of Theorem A. Since G is countably based, we can write G as an inverse limit of the countable system $A_n = G/G_n$, for some suitable open normal subgroups $G_1 \supset G_2 \supset \ldots \supset G_n \supset \ldots$ of G. Each A_n is a finite G-space, where G acts by conjugation. Now apply Propositions 2.1 and 2.2. $\square$

Proof of Corollary A. $Z(\mathbb{F}_p[G])$ is spanned over $\mathbb{F}_p$ by all conjugacy class sums $\hat{\mathcal{C}}$, where $\mathcal{C}$ is a *finite* conjugacy class of G. Let $\mathcal{C}$ be such a conjugacy class and let $x \in \mathcal{C}$. Then the centralizer $C_G(x)$ of x in G is a closed subgroup of G of finite index.
Let $y \in G$; then $y^{p^n} \in C_G(x)$ for some n, so $(x^{-1}yx)^{p^n} = y^{p^n}$. Since G is p-valued, the map $g \mapsto g^p$ is injective on G by [5, Chapter III, Proposition 2.1.4], so $x^{-1}yx = y$ and $\mathcal{C} \subseteq Z$.
Hence $Z(\mathbb{F}_p[G]) = \mathbb{F}_p[Z]$, and similarly $Z(\mathbb{Z}_p[G]) = \mathbb{Z}_p[Z]$. The result follows from Theorem A. $\square$

We remark that Corollary A does not extend to arbitrary torsion free analytic pro-p groups. This can be easily checked for the group given in [5, Chapter III, Example 3.2.5].
We now turn to the proof of Theorem B. Let $G = \varprojlim G/G_n$ be a pro-p group, H a closed subgroup. Let $M = \mathbb{F}_p \otimes_{\Omega_H} \Omega_G$ be the induced module from the trivial module for Ω_H. G acts on the coset space $Y = H\backslash G$ by right translation and we can write $Y = \varprojlim HG_n\backslash G$ as an inverse limit of finite G-spaces. It is easy to see that Ω_Y is then naturally isomorphic to M.
Let R denote the endomorphism ring $\operatorname{End}_{\Omega_G} M$ of M. Each element $f \in R$ gives rise to a trivial Ω_H-submodule of M generated by $f(1 \otimes 1)$, when we view M as an Ω_H-module by restriction. This gives an isomorphism of $\mathbb{F}_p-$vector spaces

$$R = \operatorname{Hom}_{\Omega_G}(\mathbb{F}_p \otimes_{\Omega_H} \Omega_G, M) \cong \operatorname{Hom}_{\Omega_H}(\mathbb{F}_p, M),$$

expressing the fact "induction is left adjoint to restriction".
Now $\operatorname{Hom}_{\Omega_H}(\mathbb{F}_p, M)$ can be thought of as the sum of all trivial Ω_H-submodules of M, which is precisely the set $M^H = \Omega_Y^H$, where H acts on Y by right translation. In view of Proposition 2.1, we are interested in the finite $H-$orbits on Y; these are given by those double cosets of H in G which are finite unions of left cosets of H. Suppose HxH is such a double coset; then $\operatorname{Stab}_H(Hx) = \{h \in H : Hxh = Hx\} = H \cap H^x$ has finite index in H, so the set $\mathcal{N}_G(H) = \{x \in G : H \cap H^x \leq_o H\}$ is of interest; we observe that it contains the usual normalizer $N_G(H)$ of H in G. This set is sometimes called the *commensurator* of H in G.

Proof of Theorem B. Recall [3, 9.5] that G contains an open normal uniform subgroup K and that the $\mathbb{Q}_p$-Lie algebra of G can be defined by

$$\mathcal{L}(G) = K \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$$

where K is viewed as a $\mathbb{Z}_p$-module of finite rank [3, 4.17]. The conjugation action of G on K is $\mathbb{Z}_p$-linear and therefore extends to an action of G on $\mathcal{L}(G)$, which is easily checked to be independent of the choice of K. This is just the adjoint action of G on $\mathcal{L}(G)$.
Next, we observe that when $x \in G$,

$$\begin{aligned} H \cap H^x \leq_o H \quad &\Leftrightarrow \quad \mathcal{L}(H \cap H^x) = \mathcal{L}(H) \cap \mathcal{L}(H)^x = \mathcal{L}(H) \\ &\Leftrightarrow \quad \mathcal{L}(H)^x = \mathcal{L}(H), \end{aligned}$$

so $N := \mathcal{N}_G(H) = \mathrm{Stab}_G \mathfrak{h}$ *is* a (closed) subgroup of G.
By [3, Exercise 9.10], we see that the Lie algebra of N is equal to the normalizer $N_{\mathfrak{g}}(\mathfrak{h})$ of $\mathfrak{h}$ in $\mathfrak{g}$. We remark in passing that this implies that $N_G(H)$ has finite index in N when dealing with analytic pro-p groups; this is not true in general.
Now, by Proposition 2.1 and the above remarks, R is finite dimensional over $\mathbb{F}_p$ if and only if the number of finite H-orbits on $Y = H\backslash G$ is finite.
Clearly $\{Hx : HxH \text{ is a finite } H\text{-orbit}\} = H\backslash N$, so the number of finite H-orbits on Y is finite if and only if H has finite index in N. This happens if and only if $\mathfrak{h} = \mathcal{L}(H) = \mathcal{L}(N) = N_{\mathfrak{g}}(\mathfrak{h})$, as required. $\square$

References

[1] K. Ardakov, K.A. Brown, *Primeness, semiprimeness and localisation in Iwasawa algebras*, preprint.

[2] J. Coates, P. Schneider, R. Sujatha, *Modules over Iwasawa algebras*, Journal of the Inst. of Math. Jussieu (2003), 2(1), 73-108.

[3] J. D.Dixon, M.P.F. Du Sautoy, A.Mann, D.Segal, *Analytic pro-p groups*, 2nd edition, CUP (1999).

[4] S. Howson, *Structure of central torsion Iwasawa modules*, Bull. Soc. Math. France 130 (2002), no. 4, 507-535.

[5] M. Lazard, *Groupes analytiques p-adiques*, Publ. Math. IHES 26 (1965), 389-603.

Konstantin Ardakov
Christ's College
Cambridge CB2 3BU
k.ardakov@dpmms.cam.ac.uk

Some Sharp Weighted Estimates for Multilinear Operators[1]

Liu Lanzhe

Received: August 8, 2004
Revised: November 25, 2004

Communicated by Heinz Siedentop

Abstract. In this paper, we establish a sharp inequality for some multilinear operators related to certain integral operators. The operators include Calderón-Zygmund singular integral operator, Littlewood-Paley operator, Marcinkiewicz operator and Bochner-Riesz operator. As application, we obtain the weighted norm inequalities and $L \log L$ type estimate for the multilinear operators.

2000 Mathematics Subject Classification: 42B20, 42B25.
Keywords and Phrases: Multilinear Operator; Calderón-Zygmund operator; Littlewood-Paley operator; Marcinkiewicz operator; Bochner-Riesz operator; Sharp estimate; BMO.

1. Introduction

Let T be a singular integral operator. In[1][2][3], Cohen and Gosselin studied the $L^p(p > 1)$ boundedness of the multilinear singular integral operator T^A defined by

$$T^A(f)(x) = \int_{R^n} \frac{R_{m+1}(A; x, y)}{|x - y|^m} K(x, y) f(y) dy.$$

In[6], Hu and Yang obtain a variant sharp estimate for the multilinear singular integral operators. The main purpose of this paper is to prove a sharp inequality for some multilinear operators related to certain non-convolution type integral operators. In fact, we shall establish the sharp inequality for the multilinear operators only under certain conditions on the size of the integral operators. The integral operators include Calderón-Zygmund singular integral operator,

[1]Supported by the NNSF (Grant: 10271071)

Littlewood-Paley operator, Marcinkiewicz operator and Bochner-Riesz operator. As applications, we obtain weighted norm inequalities and $L\log L$ type estimates for these multilinear operators.

2. Notations and Results

First, let us introduce some notations(see[6][12-14]). Throughout this paper, Q will denote a cube of R^n with side parallel to the axes. For any locally integrable function f, the sharp function of f is defined by

$$f^{\#}(x) = \sup_{x\in Q} \frac{1}{|Q|}\int_Q |f(y) - f_Q|dy,$$

where, and in what follows, $f_Q = |Q|^{-1}\int_Q f(x)dx$. It is well-known that(see[6])

$$f^{\#}(x) = \sup_{x\in Q} \inf_{c\in C} \frac{1}{|Q|}\int_Q |f(y) - c|dy.$$

We say that f belongs to $BMO(R^n)$ if $f^{\#}$ belongs to $L^\infty(R^n)$ and $||f||_{BMO} = ||f^{\#}||_{L^\infty}$. For $0 < r < \infty$, we denote $f_r^{\#}$ by

$$f_r^{\#}(x) = [(|f|^r)^{\#}(x)]^{1/r}.$$

Let M be the Hardy-Littlewood maximal operator defined by $M(f)(x) = \sup_{x\in Q}|Q|^{-1}\int_Q |f(y)|dy$, we write $M_p(f) = (M(f^p))^{1/p}$ for $0 < p < \infty$; For $k \in N$, we denote by M^k the operator M iterated k times, i.e., $M^1(f)(x) = M(f)(x)$ and $M^k(f)(x) = M(M^{k-1}(f))(x)$ for $k \geq 2$. Let B be a Young function and $\tilde{B}$ be the complementary associated to B, we denote that, for a function f

$$||f||_{B,\ Q} = \inf\left\{\lambda > 0 : \frac{1}{|Q|}\int_Q B\left(\frac{|f(y)|}{\lambda}\right)dy \leq 1\right\}$$

and the maximal function by

$$M_B(f)(x) = \sup_{x\in Q} ||f||_{B,\ Q};$$

The main Young function to be using in this paper is $B(t) = t(1 + log^+ t)$ and its complementary $\tilde{B}(t) = expt$, the corresponding maximal denoted by M_{LlogL} and M_{expL}. We have the generalized Hölder's inequality(see[12])

$$\frac{1}{|Q|}\int_Q |f(y)g(y)|dy \leq ||f||_{B,\ Q}||g||_{B,\ Q}$$

and the following inequality (in fact they are equivalent), for any $x \in R^n$,

$$M_{LlogL}(f)(x) \leq CM^2(f)(x)$$

and the following inequalities, for all cubes Q any $b \in BMO(R^n)$,

$$||b - b_Q||_{\exp L,\, Q} \leq C||b||_{BMO}, \ |b_{2^{k+1}Q} - b_{2Q}| \leq 2k||b||_{BMO}.$$

We denote the Muckenhoupt weights by A_p for $1 \leq p < \infty$(see[6]).
We are going to consider some integral operators as following.
Let m be a positive integer and A be a function on R^n. We denote that

$$R_{m+1}(A; x, y) = A(x) - \sum_{|\alpha| \leq m} \frac{1}{\alpha!} D^\alpha A(y)(x-y)^\alpha.$$

Definition 1. Let S and S' be Schwartz space and its dual and $T : S \to S'$ be a linear operator. Suppose there exists a locally integrable function $K(x, y)$ on $R^n \times R^n$ such that

$$T(f)(x) = \int_{R^n} K(x, y) f(y) dy$$

for every bounded and compactly supported function f. The multilinear operator related to the integral operator T is defined by

$$T^A(f)(x) = \int_{R^n} \frac{R_{m+1}(A; x, y)}{|x-y|^m} K(x, y) f(y) dy.$$

Definition 2. Let $F(x, y, t)$ defined on $R^n \times R^n \times [0, +\infty)$. Set

$$F_t(f)(x) = \int_{R^n} F(x, y, t) f(y) dy$$

for every bounded and compactly supported function f and

$$F_t^A(f)(x) = \int_{R^n} \frac{R_{m+1}(A; x, y)}{|x-y|^m} F(x, y, t) f(y) dy.$$

Let H be a Banach space of functions $h : [0, +\infty) \to R$. For each fixed $x \in R^n$, we view $F_t(f)(x)$ and $F_t^A(f)(x)$ as a mapping from $[0, +\infty)$ to H. Then, the multilinear operators related to F_t is defined by

$$S^A(f)(x) = ||F_t^A(f)(x)||;$$

We also define that $S(f)(x) = ||F_t(f)(x)||$.
Note that when $m = 0$, T^A and S^A are just the commutators of T, S and A. While when $m > 0$, it is non-trivial generalizations of the commutators. It is well known that multilinear operators are of great interest in harmonic analysis and have been widely studied by many authors (see [1-5][7]). The main purpose of this paper is to prove a sharp inequality for the multilinear operators T^A and S^A. We shall prove the following theorems in Section 3.
Theorem 1. Let $D^\alpha A \in BMO(R^n)$ for all α with $|\alpha| = m$. Suppose that T is the same as in Definition 1 such that T is bounded on $L^p(w)$ for all $w \in A_p$

with $1 < p < \infty$ and weak bounded of $(L^1(w), L^1(w))$ for all $w \in A_1$. If T^A satisfies the following size condition:

$$|T^A(f)(x) - T^A(f)(x_0)| \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} M^2(f)(x)$$

for any cube $Q = Q(x_0, d)$ with $supp f \subset (2Q)^c$, $x \in Q = Q(x_0, d)$. Then for any $0 < r < 1$, there exists a constant $C > 0$ such that for any $f \in C_0^\infty(R^n)$ and any $x \in R^n$,

$$(T^A(f))_r^\#(x) \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} M^2(f)(x).$$

THEOREM 2. Let $D^\alpha A \in BMO(R^n)$ for all α with $|\alpha| = m$. Suppose that S is the same as in Definition 2 such that S is bounded on $L^p(w)$ for all $w \in A_p$, $1 < p < \infty$ and weak bounded of $(L^1(w), L^1(w))$ for all $w \in A_1$. If S^A satisfies the following size condition:

$$||F_t^A(f)(x) - F_t^A(f)(x_0)|| \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} M^2(f)(x)$$

for any cube $Q = Q(x_0, d)$ with $supp f \subset (2Q)^c$, $x \in Q = Q(x_0, d)$. Then for any $0 < r < 1$, there exists a constant $C > 0$ such that for any $f \in C_0^\infty(R^n)$ and any $x \in R^n$,

$$(S^A(f))_r^\#(x) \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} M^2(f)(x).$$

From the theorems, we get the following

COROLLARY. Let $D^\alpha A \in BMO(R^n)$ for all α with $|\alpha| = m$. Suppose that T^A, T and S^A, S satisfy the conditions of Theorem 1 and Theorem 2.

(A). If $w \in A_p$ for $1 < p < \infty$. Then T^A and S^A are all bounded on $L^p(w)$, that is

$$||T^A(f)||_{L^p(w)} \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} ||f||_{L^p(w)}$$

and

$$||S^A(f)||_{L^p(w)} \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} ||f||_{L^p(w)}.$$

(B). If $w \in A_1$. Then there exists a constant $C > 0$ such that for each $\lambda > 0$,

$$w(\{x \in R^n : |T^A(f)(x)| > \lambda\}) \\ \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} \int_{R^n} \frac{|f(x)|}{\lambda} \left(1 + \log^+\left(\frac{|f(x)|}{\lambda}\right)\right) w(x)dx$$

and

$$w(\{x\in R^n:|S^A(f)(x)|>\lambda\})$$
$$\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\int_{R^n}\frac{|f(x)|}{\lambda}\left(1+\log^+\left(\frac{|f(x)|}{\lambda}\right)\right)w(x)dx.$$

3. Proof of Theorem

To prove the theorems, we need the following lemmas.

Lemma 1 (Kolmogorov, [6, p.485]). Let $0<p<q<\infty$ and for any function $f\geq 0$. We define that, for $1/r=1/p-1/q$

$$||f||_{WL^q}=\sup_{\lambda>0}\lambda|\{x\in R^n:f(x)>\lambda\}|^{1/q},N_{p,q}(f)=\sup_E||f\chi_E||_{L^p}/||\chi_E||_{L^r},$$

where the sup is taken for all measurable sets E with $0<|E|<\infty$. Then

$$||f||_{WL^q}\leq N_{p,q}(f)\leq(q/(q-p))^{1/p}||f||_{WL^q}.$$

Lemma 2([12, p.165]) Let $w\in A_1$. Then there exists a constant $C>0$ such that for any function f and for all $\lambda>0$,

$$w(\{y\in R^n:M^2f(y)>\lambda\})\leq C\lambda^{-1}\int_{R^n}|f(y)|(1+\log^+(\lambda^{-1}|f(y)|))w(y)dy.$$

Lemma 3.([3, p.448]) Let A be a function on R^n and $D^\alpha A\in L^q(R^n)$ for all α with $|\alpha|=m$ and some $q>n$. Then

$$|R_m(A;x,y)|\leq C|x-y|^m\sum_{|\alpha|=m}\left(\frac{1}{|\tilde{Q}(x,y)|}\int_{\tilde{Q}(x,y)}|D^\alpha A(z)|^q dz\right)^{1/q},$$

where $\tilde{Q}$ is the cube centered at x and having side length $5\sqrt{n}|x-y|$.

Proof of Theorem 1. It suffices to prove for $f\in C_0^\infty(R^n)$ and some constant C_0, the following inequality holds:

$$\left(\frac{1}{|Q|}\int_Q|T^A(f)(x)-C_0|^r dx\right)^{1/r}\leq CM^2(f).$$

Fix a cube $Q=Q(x_0,d)$ and $\tilde{x}\in Q$. Let $\tilde{Q}=5\sqrt{n}Q$ and $\tilde{A}(x)=A(x)-\sum_{|\alpha|=m}\frac{1}{\alpha!}(D^\alpha A)_{\tilde{Q}}x^\alpha$, then $R_m(A;x,y)=R_m(\tilde{A};x,y)$ and $D^\alpha\tilde{A}=D^\alpha A-(D^\alpha A)_{\tilde{Q}}$ for $|\alpha|=m$. We write, for $f_1=f\chi_{\tilde{Q}}$ and $f_2=f\chi_{R^n\setminus\tilde{Q}}$,

$$\begin{aligned}T^A(f)(x)&=\int_{R^n}\frac{R_{m+1}(A;x,y)}{|x-y|^m}K(x,y)f(y)dy\\&=\int_{R^n}\frac{R_{m+1}(A;x,y)}{|x-y|^m}K(x,y)f_2(y)dy\end{aligned}$$

$$+\int_{R^n}\frac{R_m(\tilde{A};x,y)}{|x-y|^m}K(x,y)f_1(y)dy$$
$$-\sum_{|\alpha|=m}\frac{1}{\alpha!}\int_{R^n}\frac{K(x,y)(x-y)^\alpha}{|x-y|^m}D^\alpha\tilde{A}(y)f_1(y)dy$$

then

$$\begin{aligned}
&\left|T^A(f)(x)-T^A(f_2)(x_0)\right|\\
\le\ &\left|T\left(\frac{R_m(\tilde{A};x,\cdot)}{|x-\cdot|^m}f_1\right)(x)\right|+\sum_{|\alpha|=m}\frac{1}{\alpha!}\left|T\left(\frac{(x-\cdot)^\alpha}{|x-\cdot|^m}D^\alpha\tilde{A}f_1\right)(x)\right|\\
&\quad+\left|T^A(f_2)(x)-T^A(f_2)(x_0)\right|\\
:=\ &I(x)+II(x)+III(x),
\end{aligned}$$

thus,

$$\begin{aligned}
&\left(\frac{1}{|Q|}\int_Q\left|T^A(f)(x)-T^A(f_2)(x_0)\right|^r dx\right)^{1/r}\\
\le\ &\left(\frac{C}{|Q|}\int_Q I(x)^r dx\right)^{1/r}+\left(\frac{C}{|Q|}\int_Q II(x)^r dx\right)^{1/r}+\left(\frac{C}{|Q|}\int_Q III(x)^r dx\right)^{1/r}\\
:=\ &I+II+III.
\end{aligned}$$

Now, let us estimate I, II and III, respectively. First, for $x\in Q$ and $y\in\tilde{Q}$, using Lemma 3, we get

$$R_m(\tilde{A};x,y)\le C|x-y|^m\sum_{|\alpha|=m}||D^\alpha A||_{BMO},$$

thus, by Lemma 1 and the weak type (1,1) of T, we get

$$\begin{aligned}
I\ &\le\ C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}|Q|^{-1}\frac{||T(f_1)\chi_Q||_{L^r}}{||\chi_Q||_{L^{r/(1-r)}}}\\
&\le\ C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}|Q|^{-1}||T(f_1)||_{WL^1}\\
&\le\ C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}|\tilde{Q}|^{-1}\int_{\tilde{Q}}|f(y)|dy\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(\tilde{x});
\end{aligned}$$

For II, similar to the proof of I, we get

$$\begin{aligned}
II\ &\le\ C\sum_{|\alpha|=m}|Q|^{-1}\frac{||T(D^\alpha\tilde{A}f_1)\chi_Q||_{L^r}}{||\chi_Q||_{L^{r/(1-r)}}}\le C\sum_{|\alpha|=m}|Q|^{-1}||T(D^\alpha\tilde{A}f_1)||_{WL^1}\\
&\le\ C\sum_{|\alpha|=m}|\tilde{Q}|^{-1}\int_{\tilde{Q}}|D^\alpha\tilde{A}(y)||f(y)|dy\le C\sum_{|\alpha|=m}||D^\alpha A||_{\exp L,\tilde{Q}}||f||_{LlogL,\tilde{Q}}\\
&\le\ C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M_{L\log L}(f)(\tilde{x})\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(\tilde{x});
\end{aligned}$$

For III, using Hölder' inequality and the size condition of T, we have

$$III \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} M^2(f)(\tilde{x}).$$

This completes the proof of Theorem 1.

Proof of Theorem 2. It is only to prove for $f \in C_0^\infty(R^n)$ and some constant C_0, the following inequality holds:

$$\left(\frac{1}{|Q|}\int_Q |S^A(f)(x) - C_0|^r dx\right)^{1/r} \le CM^2(f).$$

Fix a cube $Q = Q(x_0, d)$ and $\tilde{x} \in Q$. Let $\tilde{Q}$ and $\tilde{A}(x)$ be the same as the proof of Theorem 1. We write, for $f_1 = f\chi_{\tilde{Q}}$ and $f_2 = f\chi_{R^n\setminus\tilde{Q}}$,

$$\begin{aligned} F_t^A(f)(x) &= \int_{R^n} \frac{R_m(\tilde{A}; x, y)}{|x-y|^m} F(x,y,t) f_1(y) dy \\ &\quad - \sum_{|\alpha|=m} \frac{1}{\alpha!} \int_{R^n} \frac{F(x,y,t)(x-y)^\alpha}{|x-y|^m} D^\alpha \tilde{A}(y) f_1(y) dy \\ &\quad + \int_{R^n} \frac{R_{m+1}(A; x, y)}{|x-y|^m} F(x,y,t) f_2(y) dy, \end{aligned}$$

then

$$\begin{aligned} & |S^A(f)(x) - S^A(f_2)(x_0)| = \Big| ||F_t^A(f)(x)|| - ||F_t^A(f_2)(x_0)|| \Big| \\ \le\ & ||F_t^A(f)(x) - F_t^A(f_2)(x_0)|| \\ \le\ & \left\| F_t\left(\frac{R_m(\tilde{A}; x, \cdot)}{|x-\cdot|^m} f_1\right)(x)\right\| + \sum_{|\alpha|=m} \frac{1}{\alpha!} \left\| F_t\left(\frac{(x-\cdot)^\alpha}{|x-\cdot|^m} D^\alpha \tilde{A} f_1\right)(x)\right\| \\ & + ||F_t^A(f_2)(x) - F_t^A(f_2)(x_0)|| \\ :=\ & J(x) + JJ(x) + JJJ(x), \end{aligned}$$

thus,

$$\begin{aligned} & \left(\frac{1}{|Q|}\int_Q |S^A(f)(x) - S^A(f_2)(x_0)|^r dx\right)^{1/r} \\ \le\ & \left(\frac{C}{|Q|}\int_Q J(x)^r dx\right)^{1/r} + \left(\frac{C}{|Q|}\int_Q JJ(x)^r dx\right)^{1/r} + \left(\frac{C}{|Q|}\int_Q JJJ(x)^r dx\right)^{1/r} \\ :=\ & J + JJ + JJJ. \end{aligned}$$

Now, similar to the proof of Theorem 1, we have

$$J \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} \frac{1}{|\tilde{Q}|}\int_{\tilde{Q}} |f(x)| dx \le C \sum_{|\alpha|=m} ||D^\alpha A||_{BMO} M(f)(\tilde{x})$$

and

$$\begin{aligned} JJ &\le C\sum_{|\alpha|=m}|Q|^{-1}\frac{||S(D^\alpha\tilde{A}f_1)\chi_Q||_{L^r}}{||\chi_Q||_{L^{r/(1-r)}}}\le C\sum_{|\alpha|=m}|Q|^{-1}||S(D^\alpha\tilde{A}f_1)||_{WL^1}\\ &\le C\sum_{|\alpha|=m}|\tilde{Q}|^{-1}\int_{\tilde{Q}}|D^\alpha\tilde{A}(y)||f(y)|dy\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(\tilde{x}); \end{aligned}$$

For JJJ, using the size condition of S, we have

$$JJJ\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(\tilde{x}).$$

This completes the proof of Theorem 2.
From Theorem 1, 2 and the weighted boundedness of T and S, we may obtain the conclusion of Corollary(a).
From Theorem 1, 2 and Lemma 2, we may obtain the conclusion of Corollary(b).

4. Applications

In this section we shall apply the Theorem 1, 2 and Corollary of the paper to some particular operators such as the Calderón-Zygmund singular integral operator, Littlewood-Paley operator, Marcinkiewicz operator and Bochner-Riesz operator.

Application 1. Calderón-Zygmund singular integral operator.
Let T be the Calderón-Zygmund operator(see[6][14][15]), the multilinear operator related to T is defined by

$$T^A(f)(x)=\int\frac{R_{m+1}(A;x,y)}{|x-y|^m}K(x,y)f(y)dy.$$

Then it is easily to see that T satisfies the conditions in Theorem 1 and Corollary. In fact, it is only to verify that T^A satisfies the size condition in Theorem 1, which has done in [6](see also [12][13]). Thus the conclusions of Theorem 1 and Corollary hold for T^A.

Application 2. Littlewood-Paley operator.
Let $\varepsilon>0$ and ψ be a fixed function which satisfies the following properties:
(1) $\int_{R^n}\psi(x)dx=0$,
(2) $|\psi(x)|\le C(1+|x|)^{-(n+1)}$,
(3) $|\psi(x+y)-\psi(x)|\le C|y|^\varepsilon(1+|x|)^{-(n+1+\varepsilon)}$ when $2|y|<|x|$;
The multilinear Littlewood-Paley operator is defined by(see[8])

$$g_\psi^A(f)(x)=\left(\int_0^\infty|F_t^A(f)(x)|^2\frac{dt}{t}\right)^{1/2},$$

where

$$F_t^A(f)(x)=\int_{R^n}\frac{R_{m+1}(A;x,y)}{|x-y|^m}\psi_t(x-y)f(y)dy$$

and $\psi_t(x) = t^{-n}\psi(x/t)$ for $t > 0$. We write $F_t(f) = \psi_t * f$. We also define that

$$g_\psi(f)(x) = \left(\int_0^\infty |F_t(f)(x)|^2 \frac{dt}{t}\right)^{1/2},$$

which is the Littlewood-Paley operator(see [15]);
Let H be a space of functions $h : [0,+\infty) \to R$, normed by $||h|| = \left(\int_0^\infty |h(t)|^2 dt/t\right)^{1/2} < \infty$. Then, for each fixed $x \in R^n$, $F_t^A(f)(x)$ may be viewed as a mapping from $[0,+\infty)$ to H, and it is clear that

$$g_\psi(f)(x) = ||F_t(f)(x)|| \text{ and } g_\psi^A(f)(x) = ||F_t^A(f)(x)||.$$

It is known that g_ψ is bounded on $L^p(w)$ for all $w \in A_p$, $1 < p < \infty$ and weak $(L^1(w), L^1(w))$ bounded for all $w \in A_1$. Thus it is only to verify that g_ψ^A satisfies the size condition in Theorem 2. In fact, we write, for a cube $Q = Q(x_0, d)$ with $suppf \subset (\tilde{Q})^c$, $x \in Q = Q(x_0, d)$,

$$\begin{aligned}
&F_t^A(f)(x) - F_t^A(f)(x_0) \\
= &\int_{R^n} \left(\frac{\psi_t(x-y)}{|x-y|^m} - \frac{\psi_t(x_0-y)}{|x_0-y|^m}\right) R_m(\tilde{A}; x, y) f(y) dy \\
&+ \int_{R^n} \frac{\psi_t(x_0-y)}{|x_0-y|^m} (R_m(\tilde{A}; x, y) - R_m(\tilde{A}; x_0, y)) f(y) dy \\
&- \sum_{|\alpha|=m} \frac{1}{\alpha!} \int_{R^n} \left(\frac{(x-y)^\alpha \psi_t(x-y)}{|x-y|^m} - \frac{(x_0-y)^\alpha \psi_t(x_0-y)}{|x_0-y|^m}\right) D^\alpha \tilde{A}(y) f(y) dy \\
:= &I_1 + I_2 + I_3.
\end{aligned}$$

By Lemma 3 and the following inequality(see[14])

$$|b_{Q_1} - b_{Q_2}| \le C \log(|Q_2|/|Q_1|)||b||_{BMO}, for Q_1 \subset Q_2,$$

we know that, for $x \in Q$ and $y \in 2^{k+1}Q \setminus 2^k Q$ with $k \ge 1$,

$$\begin{aligned}
|R_m(\tilde{A}; x, y)| &\le C|x-y|^m \sum_{|\alpha|=m} (||D^\alpha A||_{BMO} + |(D^\alpha A)_{\tilde{Q}(x,y)} - (D^\alpha A)_{\tilde{Q}}|) \\
&\le Ck|x-y|^m \sum_{|\alpha|=m} ||D^\alpha A||_{BMO}.
\end{aligned}$$

Note that $|x-y| \sim |x_0-y|$ for $x \in Q$ and $y \in R^n \setminus Q$. By the condition on ψ and Minkowski' inequality , we obtain

$$\begin{aligned}
||I_1|| \le\ & C \int_{R^n} \frac{|R_m(\tilde{A}; x, y)||f(y)|}{|x_0-y|^m} \\
& \left[\int_0^\infty \left(\frac{t|x-x_0|}{|x_0-y|(t+|x_0-y|)^{n+1}} + \frac{t|x-x_0|^\varepsilon}{(t+|x_0-y|)^{n+1+\varepsilon}}\right)^2 \frac{dt}{t}\right]^{1/2} dy
\end{aligned}$$

$$\begin{aligned}
&\leq C\int_{(2Q)^c}\left(\frac{|x-x_0|}{|x_0-y|^{m+n+1}}+\frac{|x-x_0|^{\varepsilon}}{|x_0-y|^{m+n+\varepsilon}}\right)|R_m(\tilde{A};x,y)||f(y)|dy\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\\
&\qquad\sum_{k=1}^{\infty}\int_{2^{k+1}Q\setminus 2^kQ}k\left(\frac{|x-x_0|}{|x_0-y|^{n+1}}+\frac{|x-x_0|^{\varepsilon}}{|x_0-y|^{n+\varepsilon}}\right)|f(y)|dy\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\sum_{k=1}^{\infty}k(2^{-k}+2^{-\varepsilon k})|2^{k+1}Q|^{-1}\int_{2^{k+1}Q}|f(y)|dy\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x);
\end{aligned}$$

For I_2, by the formula (see [3]):

$$R_m(\tilde{A};x,y)-R_m(\tilde{A};x_0,y)=\sum_{|\beta|<m}\frac{1}{\beta!}R_{m-|\beta|}(D^\beta\tilde{A};x,x_0)(x-y)^\beta$$

and Lemma 3, we have

$$|R_m(\tilde{A};x,y)-R_m(\tilde{A};x_0,y)|\leq C\sum_{|\beta|<m}\sum_{|\alpha|=m}|x-x_0|^{m-|\beta|}|x-y|^{|\beta|}||D^\alpha A||_{BMO},$$

similar to the estimates of I_1, we get

$$\begin{aligned}
||I_2|| &\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\sum_{k=1}^{\infty}\int_{2^{k+1}\setminus 2^kQ}\frac{k|x-x_0|}{|x_0-y|^{n+1}}|f(y)|dy\\
&\leq C||D^\alpha A||_{BMO}\sum_{k=1}^{\infty}k2^{-k}|2^{k+1}Q|^{-1}\int_{2^{k+1}Q}|f(y)|dy\\
&\leq C||D^\alpha A||_{BMO}M(f)(x);
\end{aligned}$$

For I_3, similar to the proof of I_1, we obtain

$$\begin{aligned}
||I_3|| &\leq C\sum_{|\alpha|=m}\sum_{k=1}^{\infty}\int_{2^{k+1}\setminus 2^kQ}\left(\frac{|x-x_0|}{|x_0-y|^{n+1}}+\frac{|x-x_0|^{\varepsilon}}{|x_0-y|^{n+\varepsilon}}\right)|D^\alpha\tilde{A}(y)||f(y)|dy\\
&\leq C\sum_{|\alpha|=m}\sum_{k=1}^{\infty}k(2^{-k}+2^{-\varepsilon k})\frac{1}{|2^{k+1}Q|}\int_{2^{k+1}Q}|D^\alpha\tilde{A}(y)||f(y)|dy\\
&\leq C\sum_{|\alpha|=m}\sum_{k=1}^{\infty}k(2^{-k}+2^{-\varepsilon k})\\
&\qquad\left(||D^\alpha A||_{\exp L,2^{k+1}Q}||f||_{LlogL,2^{k+1}Q}+||D^\alpha A||_{BMO}M(f)(x)\right)\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}(M_{L\log L}(f)(x)+M(f)(x))\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(x).
\end{aligned}$$

From the above estimates, we know that Theorem 2 and Corollary hold for g_ψ^A.
Application 3. Marcinkiewicz operator.
Let Ω be homogeneous of degree zero on R^n and $\int_{S^{n-1}} \Omega(x')d\sigma(x') = 0$. Assume that $\Omega \in Lip_\gamma(S^{n-1})$ for $0 < \gamma \le 1$, that is there exists a constant $M > 0$ such that for any $x, y \in S^{n-1}$, $|\Omega(x) - \Omega(y)| \le M|x-y|^\gamma$. The multilinear Marcinkiewicz operator is defined by(see[9])

$$\mu_\Omega^A(f)(x) = \left(\int_0^\infty |F_t^A(f)(x)|^2 \frac{dt}{t^3}\right)^{1/2},$$

where

$$F_t^A(f)(x) = \int_{|x-y|\le t} \frac{\Omega(x-y)}{|x-y|^{n-1}} \frac{R_{m+1}(A;x,y)}{|x-y|^m} f(y)dy,$$

we write that

$$F_t(f)(x) = \int_{|x-y|\le t} \frac{\Omega(x-y)}{|x-y|^{n-1}} f(y)dy.$$

We also define that

$$\mu_\Omega(f)(x) = \left(\int_0^\infty |F_t(f)(x)|^2 \frac{dt}{t^3}\right)^{1/2},$$

which is the Marcinkiewicz operator(see [16]);
Let H be a space of functions $h : [0, +\infty) \to R$, normed by $||h|| = \left(\int_0^\infty |h(t)|^2 dt/t^3\right)^{1/2} < \infty$. Then, it is clear that

$$\mu_\Omega(f)(x) = ||F_t(f)(x)|| \text{ and } \mu_\Omega^A(f)(x) = ||F_t^A(f)(x)||.$$

Now, we will verify that μ_Ω^A satisfies the size condition in Theorem 2. In fact, for a cube $Q = Q(x_0, d)$ with $suppf \subset (2Q)^c$, $x \in Q = Q(x_0, d)$, we have

$$\begin{aligned}
&||F_t^A(f)(x) - F_t^A(f)(x_0)|| \\
\le\ & \left(\int_0^\infty \left| \int_{|x-y|\le t} \frac{\Omega(x-y)R_m(\tilde{A};x,y)}{|x-y|^{m+n-1}} f(y)dy \right.\right. \\
&\qquad \left.\left. - \int_{|x_0-y|\le t} \frac{\Omega(x_0-y)R_m(\tilde{A};x_0,y)}{|x_0-y|^{m+n-1}} f(y)dy \right|^2 \frac{dt}{t^3}\right)^{1/2} \\
&+ \sum_{|\alpha|=m} \left(\int_0^\infty \left| \int_{|x-y|\le t} \left(\frac{\Omega(x-y)(x-y)^\alpha}{|x-y|^{m+n-1}} \right.\right.\right. \\
&\qquad \left.\left.\left. - \int_{|x_0-y|\le t} \frac{\Omega(x_0-y)(x_0-y)^\alpha}{|x_0-y|^{m+n-1}}\right) D^\alpha \tilde{A}(y) f(y)dy \right|^2 \frac{dt}{t^3}\right)^{1/2}
\end{aligned}$$

$$\begin{aligned}
&\leq \left(\int_0^\infty\left[\int_{|x-y|\le t,\ |x_0-y|>t}\frac{|\Omega(x-y)||R_m(\tilde{A};x,y)|}{|x-y|^{m+n-1}}|f(y)|dy\right]^2\frac{dt}{t^3}\right)^{1/2}\\
&+\left(\int_0^\infty\left[\int_{|x-y|>t,\ |x_0-y|\le t}\frac{|\Omega(x_0-y)||R_m(\tilde{A};x_0,y)|}{|x_0-y|^{m+n-1}}|f(y)|dy\right]^2\frac{dt}{t^3}\right)^{1/2}\\
&+\left(\int_0^\infty\left[\int_{|x-y|\le t,|x_0-y|\le t}\left|\frac{\Omega(x-y)R_m(\tilde{A};x,y)}{|x-y|^{m+n-1}}\right.\right.\right.\\
&\qquad\left.\left.\left.-\frac{\Omega(x_0-y)R_m(\tilde{A};x_0,y)}{|x_0-y|^{m+n-1}}\right||f(y)|dy\right]^2\frac{dt}{t^3}\right)^{1/2}\\
&+\sum_{|\alpha|=m}\left(\int_0^\infty\left|\int_{|x-y|\le t}\left(\frac{\Omega(x-y)(x-y)^\alpha}{|x-y|^{m+n-1}}\right.\right.\right.\\
&\qquad\left.\left.\left.-\int_{|x_0-y|\le t}\frac{\Omega(x_0-y)(x_0-y)^\alpha}{|x_0-y|^{m+n-1}}\right)D^\alpha\tilde{A}(y)f(y)dy\right|^2\frac{dt}{t^3}\right)^{1/2}\\
&:= J_1+J_2+J_3+J_4
\end{aligned}$$

and

$$\begin{aligned}
J_1 &\leq C\int_{R^n}\frac{|f(y)||R_m(\tilde{A};x,y)|}{|x-y|^{m+n-1}}\left(\int_{|x-y|\le t<|x_0-y|}\frac{dt}{t^3}\right)^{1/2}dy\\
&\leq C\int_{R^n}\frac{|f(y)||R_m(\tilde{A};x,y)|}{|x-y|^{m+n-1}}\left(\frac{1}{|x-y|^2}-\frac{1}{|x_0-y|^2}\right)^{1/2}dy\\
&\leq C\int_{(2Q)^c}\frac{|f(y)||R_m(\tilde{A};x,y)|}{|x-y|^{m+n-1}}\frac{|x_0-x|^{1/2}}{|x-y|^{3/2}}dy\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\sum_{k=1}^\infty k2^{-k/2}|2^{k+1}Q|^{-1}\int_{2^{k+1}Q}|f(y)|dy\\
&\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x),
\end{aligned}$$

similarly, we have $J_2\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x)$;
For J_3, by the following inequality (see [16]):

$$\left|\frac{\Omega(x-y)}{|x-y|^{n-1}}-\frac{\Omega(x_0-y)}{|x_0-y|^{n-1}}\right|\le C\left(\frac{|x-x_0|}{|x_0-y|^n}+\frac{|x-x_0|^\gamma}{|x_0-y|^{n-1+\gamma}}\right),$$

we gain

$$\begin{aligned} J_3 &\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\int_{(2Q)^c}\left(\frac{|x-x_0|}{|x_0-y|^n}+\frac{|x-x_0|^\gamma}{|x_0-y|^{n-1+\gamma}}\right) \\ &\qquad\qquad\left(\int_{|x_0-y|\leq t,|x-y|\leq t}\frac{dt}{t^3}\right)^{1/2}|f(y)|dy \\ &\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}\sum_{k=1}^\infty k(2^{-k}+2^{-\gamma k})M(f)(x) \\ &\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x); \end{aligned}$$

For J_4, similar to the proof of J_1, J_2 and J_3, we obtain

$$\begin{aligned} ||J_4|| &\leq C\sum_{|\alpha|=m}\sum_{k=1}^\infty\int_{2^{k+1}Q\setminus 2^kQ}\left(\frac{|x-x_0|}{|x_0-y|^{n+1}}\right. \\ &\qquad\left.+\frac{|x-x_0|^{1/2}}{|x_0-y|^{n+1/2}}+\frac{|x-x_0|^\gamma}{|x_0-y|^{n+\gamma}}\right)|D^\alpha\tilde{A}(y)||f(y)|dy \\ &\leq C\sum_{|\alpha|=m}\sum_{k=1}^\infty k(2^{-k}+2^{-k/2}+2^{-\gamma k})\frac{1}{|2^{k+1}Q|}\int_{2^{k+1}Q}|D^\alpha\tilde{A}(y)||f(y)|dy \\ &\leq C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(x). \end{aligned}$$

Thus, Theorem 2 and Corollary hold for μ_Ω^A.
Application 4. Bochner-Riesz operator.
Let $B_t^\delta(f)\hat{}(\xi)=(1-t^2|\xi|^2)_+^\delta\hat{f}(\xi)$. Denote

$$B_{\delta,\ t}^A(f)(x)=\int_{R^n}\frac{R_{m+1}(A;x,y)}{|x-y|^m}B_t^\delta(x-y)f(y)dy,$$

where $B_t^\delta(z)=t^{-n}B^\delta(z/t)$ for $t>0$. The maximal multilinear Bochner-Riesz operator is defined by(see[9])

$$B_{\delta,*}^A(f)(x)=\sup_{t>0}|B_{\delta,\ t}^A(f)(x)|.$$

We also define

$$B_*^\delta(f)(x)=\sup_{t>0}|B_t^\delta(f)(x)|,$$

which is the maximal Bochner-Riesz operator (see [10][11]).
Let H be the space of functions $h(t)$ such that $||h||=\sup_{t>0}|h(t)|<\infty$, where $h(t)$ maps $[0,+\infty)$ to H. Then it is clear that

$$B_*^\delta(f)(x)=||B_t^\delta(f)(x)|| \text{ and } B_{\delta,*}^A(f)(x)=||B_{\delta,\ t}^A(f)(x)||.$$

Now, we will verify that $B^A_{\delta,*}$ satisfies the size condition in Theorem 2. In fact, for a cube $Q = Q(x_0, d)$ with $suppf \subset (2Q)^c$, $x \in Q = Q(x_0, d)$, we have

$$B^{\tilde{A}}_{t,\delta}(f)(x) - B^{\tilde{A}}_{t,\delta}(f)(x_0) = \int_{R^n}\left[\frac{B^\delta_t(x-y)}{|x-y|^m} - \frac{B^\delta_t(x_0-y)}{|x_0-y|^m}\right] R_m(\tilde{A}; x, y) f(y)dy$$
$$+ \int_{R^n}\frac{B^\delta_t(x_0-y)}{|x_0-y|^m}[R_m(\tilde{A}; x, y) - R_m(\tilde{A}; x_0, y)]f(y)dy$$
$$- \sum_{|\alpha|=m}\frac{1}{\alpha!}\int_{R^n}\left(\frac{B^\delta_t(x-y)(x-y)^\alpha}{|x-y|^m} - \frac{B^\delta_t(x_0-y)(x_0-y)^\alpha}{|x_0-y|^m}\right) D^\alpha\tilde{A}(y)f(y)dy$$
$$= L_1 + L_2 + L_3.$$

Consider the following two cases:
Case 1. $0 < t \le d$. In this case, notice that (see [11])

$$|B^\delta(z)| \le c(1+|z|)^{-(\delta+(n+1)/2)},$$

we obtain

$$\begin{aligned}
|L_1| &\le Ct^{-n}\int_{R^n\setminus\tilde{Q}}\frac{|f(y)||R_m(\tilde{A}; x, y)|}{|x_0-y|^m}(1+|x-y|/t)^{-(\delta+(n+1)/2)}dy\\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}t^{-n}\sum_{k=0}^\infty k\int_{2^{k+1}\tilde{Q}\setminus 2^k\tilde{Q}}|f(y)||(1+|x-y|/t)^{-(\delta+(n+1)/2)}dy\\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}(t/d)^{\delta-(n-1)/2}\sum_{k=1}^\infty k2^{k((n-1)/2-\delta)}M(f)(x)\\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x),\\
|L_2| &\le Ct^{-n}\int_{R^n\setminus\tilde{Q}}\frac{|f(y)||R_m(\tilde{A}; x, y) - R_m(\tilde{A}; x_0, y)|}{|x_0-y|^m}(1+|x-y|/t)^{-(\delta+(n+1)/2)}dy\\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}t^{-n}\\
&\qquad\sum_{k=0}^\infty\int_{2^{k+1}\tilde{Q}\setminus 2^k\tilde{Q}}\frac{|x-x_0||f(y)|}{|x_0-y|}(1+|x-y|/t)^{-(\delta+(n+1)/2)}dy\\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x),\\
|L_3| &\le C\sum_{|\alpha|=m}t^{-n}\sum_{k=0}^\infty\int_{2^{k+1}\tilde{Q}\setminus 2^k\tilde{Q}}|f(y)||D^\alpha\tilde{A}(y)|(1+|x_0-y|/t)^{-(\delta+(n+1)/2)}dy\\
&\le C\sum_{|\alpha|=m}(t/d)^{\delta-\frac{n-1}{2}}\sum_{k=0}^\infty 2^{k(\frac{n-1}{2}-\delta)}\frac{1}{|2^{k+1}\tilde{Q}|}\int_{2^{k+1}\tilde{Q}}|f(y)||D^\alpha A(y) - (D^\alpha A)_{\tilde{Q}}|dy\\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(x).
\end{aligned}$$

Case 2. $t > d$. In this case, we choose δ_0 such that $(n-1)/2 < \delta_0 < \min(\delta, (n+1)/2)$, notice that (see [11])

$$|B^\delta(x-y) - B^\delta(x_0-y)| \le C|x-x_0|(1+|x-y|)^{-(\delta+(n+1)/2)},$$

similar to the proof of Case 1, we obtain

$$\begin{aligned}
|L_1| &\le Ct^{-n}\int_{R^n\setminus\tilde{Q}} \frac{|f(y)||R_m(\tilde{A};x,y)|}{|x_0-y|^{m+1}}|x_0-x|(1+|x_0-y|/t)^{-(\delta_0+(n+1)/2)}dy \\
&\quad +Ct^{-n-1}\int_{R^n\setminus\tilde{Q}} \frac{|f(y)||R_m(\tilde{A};x,y)|}{|x_0-y|^{m}}|x_0-x|(1+|x_0-y|/t)^{-(\delta_0+(n+1)/2)}dy \\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}(d/t)^{(n+1)/2-\delta_0}\sum_{k=1}^\infty k2^{k((n-1)/2-\delta_0)}M(f)(x) \\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x), \\
|L_2| &\le Ct^{-n}\int_{R^n\setminus\tilde{Q}} \frac{|f(y)||R_m(\tilde{A};x,y)-R_m(\tilde{A};x_0,y)|}{|x_0-y|^{m}}(1+|x_0-y|/t)^{-(\delta_0+(n+1)/2)}dy \\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}(d/t)^{(n+1)/2-\delta_0}\sum_{k=1}^\infty 2^{k((n-1)/2-\delta_0)}M(f)(x) \\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M(f)(x), \\
|L_3| &\le C\sum_{|\alpha|=m}(d/t)^{(n+1)/2-\delta_0}\sum_{k=0}^\infty 2^{k((n-1)/2-\delta_0)}\frac{1}{|2^{k+1}\tilde{Q}|}\int_{2^{k+1}\tilde{Q}}|f(y)||D^\alpha\tilde{A}(y)|dy \\
&\le C\sum_{|\alpha|=m}\sum_{k=1}^\infty k2^{k((n-1)/2-\delta_0)}\frac{1}{|2^{k+1}\tilde{Q}|}\int_{2^{k+1}\tilde{Q}}|f(y)||D^\alpha A(y)-(D^\alpha A)_{\tilde{Q}}|dy \\
&\le C\sum_{|\alpha|=m}||D^\alpha A||_{BMO}M^2(f)(x).
\end{aligned}$$

Thus, Theorem 2 and Corollary hold for $B^A_{\delta,*}$.

Acknowledgement. The author would like to express his deep gratitude to the referee for his valuable comments and suggestions.

References

[1] J.Cohen, A sharp estimate for a multilinear singular integral on R^n, Indiana Univ. Math. J., 30(1981), 693-702.

[2] J.Cohen and J.Gosselin, On multilinear singular integral operators on R^n, Studia Math., 72(1982), 199-223.

[3] J.Cohen and J.Gosselin, A BMO estimate for multilinear singular integral operators, Illinois J. Math., 30(1986), 445-465.

[4] R.Coifman and Y.Meyer, Wavelets, Calderón-Zygmund and multilinear operators, Cambridge Studies in Advanced Math.48, Cambridge University Press, Cambridge , 1997.

[5] Y.Ding and S.Z.Lu, Weighted boundedness for a class rough multilinear operators, Acta Math. Sinica, 17(2001), 517-526.

[6] J.Garcia-Cuerva and J.L.Rubio de Francia, Weighted norm inequalities and related topics, North-Holland Math.16, Amsterdam, 1985.

[7] G.Hu and D.C.Yang, A variant sharp estimate for multilinear singular integral operators, Studia Math., 141(2000), 25-42.

[8] L.Z. Liu, Weighted weak type estimates for commutators of Littlewood-Paley operator, Japanese J. of Math., 29(1)(2003), 1-13.

[9] L.Z. Liu, The continuity of commutators on Triebel-Lizorkin spaces, Integral Equations and Operator Theory, 49(2004), 65-75.

[10] L.Z. Liu and S.Z. Lu, Weighted weak type inequalities for maximal commutators of Bochner-Riesz operator, Hokkaido Math. J., 32(1)(2003), 85-99.

[11] S.Z.Lu, Four lectures on real H^p spaces, World Scientific, River Edge, NI, 1995.

[12] C.Perez, Endpoint estimate for commutators of singular integral operators, J. Func. Anal., 128(1995), 163-185.

[13] C.Perez and G. Pradolini, Sharp weighted endpoint estimates for commutators of singular integral operators, Michigan Math. J., 49(2001), 23-37.

[14] E.M.Stein, Harmonic Analysis: real variable methods, orthogonality and oscillatory integrals, Princeton Univ. Press, Princeton NJ, 1993.

[15] A.Torchinsky, The real variable methods in harmonic analysis, Pure and Applied Math., 123, Academic Press, New York, 1986.

[16] A.Torchinsky and S.Wang, A note on the Marcinkiewicz integral, Colloq. Math., 60/61(1990), 235-24.

Liu Lanzhe
College of Mathematics
and Computer
Changsha University
of Science and Technology
Changsha 410077
P.R. of China
lanzheliu@263.net

Rational Curves on Homogeneous Cones

Nicolas Perrin

Received: October 18, 2004
Revised: December 20, 2004

Communicated by Thomas Peternell

Abstract. A homogeneous cone X is the cone over a homogeneous variety G/P embedded thanks to an ample line bundle L. In this article, we describe the irreducible components of the scheme of morphisms of class $\alpha \in A_1(X)$ from a rational curve to X.

The situation depends on the line bundle L : if the projectivised tangent space to the vertex contains lines then the irreducible components are described by the difference between Cartier and Weil divisors. On the contrary if there is no line in the projectivised tangent space to the vertex then there are new irreducible components corresponding to the multiplicity of the curve through the vertex.

2000 Mathematics Subject Classification: 14C05, 14M17.
Keywords and Phrases: homogeneous cone, scheme of morphisms, rational curves.

In this text we study the scheme of morphisms from $\mathbb{P}^1$ to any homogeneous cone that is to say a cone X over a homogeneous variety G/P.

This study is motivated by the more general problem of describing the irreducible components of the scheme of morphisms from $\mathbb{P}^1$ to a variety X endowed with the action of a solvable group with finitely many orbits (for example homogeneous varieties, Schubert varieties or spherical varieties). For homogeneous varieties, this is already known (see [Th], [KP] or [P1]). For Schubert varieties it is not known. We describe in [P2] the irreducible components of this scheme of morphisms when X is a minuscule Schubert variety. The singularities of minuscule Schubert varities are locally isomorphic to cones over homogeneous varieties (see [BP]) so it is natural to address this problem on a general cone over a homogeneous variety. Moreover some non minuscule Schubert varieties are cones over homogeneous varieties. We prove (theorem 0.1) that the situation is more complicated than in the minuscule case but still has a nice description in terms of the combinatorial data of G.

More precisely, let G/P be a homogeneous variety and let L be a very ample divisor on G/P. We may embed G/P in $\mathbb{P}(H^0L)$. If V is a n-dimensional vector space, it defines a linear subspace $\mathbb{P}(V)$ in $\mathbb{P}(H^0L \oplus V)$. Let us denote by $X = \mathfrak{C}_{L,n}(G/P)$ the cone above G/P whose vertex is $\mathbb{P}(V)$. Now let U be the open subset of X complementary to $Y = \mathbb{P}(V)$. We have a surjective morphism (see paragraph 1):

$$s : \mathrm{Pic}(U)^\vee \to A_1(X).$$

For any class $\alpha \in A_1(X)$, we can consider the following morphism:

$$i : \coprod_{s(\beta)=\alpha} \mathbf{Hom}_\beta(\mathbb{P}^1, U) \to \mathbf{Hom}_\alpha(\mathbb{P}^1, X)$$

where $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ is the scheme of morhisms $f : \mathbb{P}^1 \to X$ with $f_*[\mathbb{P}^1] = \alpha$ and $\mathbf{Hom}_\beta(\mathbb{P}^1, U)$ is the scheme of morhisms $g : \mathbb{P}^1 \to U$ such that $[g] = \beta$ where $[g]$ is the linear function $L \mapsto \deg(g^*L)$ on $\mathrm{Pic}(U)$. As $Y = X \setminus U$ lies in codimension 2, we expect the image of this morphism to be dense. For example we prove in [P2] that it is true for X a minuscule Schubert variety and U the smooth locus.

In our case the situation will be more complicated. Let us first describe the "expected" components in the case where i is dominant. In this case we may apply the results of [P1] to prove that $\mathbf{Hom}_\beta(\mathbb{P}^1, U)$ is irreducible as soon as it is non empty and the images of the irreducible varieties $\mathbf{Hom}_\beta(\mathbb{P}^1, U)$ will give the irreducible components of $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$. The expected components are thus indexed by the subset $\mathfrak{ne}(\alpha)$ of $\mathrm{Pic}(U)^\vee$ given by elements β such that $s(\beta) = \alpha$ and $\mathbf{Hom}_\beta(\mathbb{P}^1, U)$ is non empty.

This set can be discribed in terms of roots: the ample divisor L is a dominant weight in the facet of the parabolic P. An element $\alpha \in A_1(X)$ is completely determined by $\alpha \cdot L = d \in \mathbb{Z}$. Denote by $\mathfrak{ne}_B(\alpha)$ the set of all elements β in the cone generated by the positive roots such that $\langle \beta^\vee, L \rangle = d$. This is a subset of $A_1(G/B)$. Then $\mathfrak{ne}(\alpha)$ is its image in $A_1(G/P)$ (see paragraph 1 for a more details). We prove the

THEOREM 0.1. — *Let R be the root lattice.*

(i) *If $L(R) = \mathbb{Z}$, then the irreducible components of the scheme of morphisms $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ are indexed by $\mathfrak{ne}(\alpha)$.*

(ii) *If $L(R) \neq \mathbb{Z}$ (i.e. if we have $L > c_1(T_{G/P})$), then the irreducible components of the scheme $\mathbf{Hom}_\alpha(\mathbb{P}^1, \mathfrak{C}(G/P))$ are indexed by $\coprod_{\alpha' \leq \alpha} \mathfrak{ne}(\alpha')$.*

We will see (paragraph 1) that $A_1(X) \simeq \mathbb{Z}$ so that $\alpha' \leq \alpha$ in $A_1(X)$ means that the same inequality holds in $\mathbb{Z}$. In the second case, a general curve can meet the vertex of the cone. The integer $\alpha - \alpha'$ is then the multiplicity of the curve at the vertex.

Remark 0.2. — (i) *The condition $L(R) = \mathbb{Z}$ is exactly equivalent to the fact that there exists lines on G/P embedded with L. In other words there exists lines in the projectivized tangent cone to the singularity.*

We studied in [P2] the same problem for minuscule Schubert varieties where the multiplicity in the singularity did not appear. If one consider more generaly quasi-minuscule Schubert varieties of non minuscule type (see [LMS] for a definition, the case of quasi-minuscule Schubert varieties of minuscule type should be very similar to the case of minuscule Scubert varieties) we recover this condition of the existence of lines in the projectivised tangent cone to the singularity.

(ii) *If $P = B$ is a Borel subgroup and if we choose for L the Plücker embedding (or equivalently $L = \rho$ as a weight where ρ is half the sum of the positive roots) then $L(R) = \mathbb{Z}$ and the set $\mathfrak{ne}(\alpha)$ is in bijection with the set of irreducible integrable representations of level exactly $\alpha \cdot L$ of the affine Lie algebra $\hat{\mathfrak{g}}$ (see paragraph 1 for the general case).*

Here is an outline of the paper. In the first paragraph we define the surjective map s of the introduction and the set $\mathfrak{ne}(\alpha)$ for a homogeneous cone X. In the second paragraph we study the scheme of morphisms from $\mathbb{P}^1$ to the blowing-up $\widetilde{X}$ of the cone X and prove a smoothing result. In the last paragraph we prove our main result.

The key point as indicated above is to study the surjectivity of the map i that is to say to study the following problem: can any morphism $f : \mathbb{P}^1 \to X$ be factorised in U (modulo deformation). We do this by lifting f in $\widetilde{f}$ on $\widetilde{X}$ and the problem becomes: does the lifted curve $\widetilde{f}$ of a general curve f meet the exceptional divisor E. If it is the case then we add a "line" $\Gamma \subset E$ (this is possible only when $L(R) = \mathbb{Z}$) with $\Gamma \cdot E = -1$ and smooth the union $\widetilde{f}(\mathbb{P}^1) \cup \Gamma$. The intersection with E is lowered by one in the operation. We conclude by induction on the number of intersection of $\widetilde{f}$ with E.

We end with a discussion on the dimensions of the components, in particular the variety $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ is equidimensional if and only if $L = \frac{1}{2}c_1(G/P)$ or $L = c_1(G/P)$.

1 Preliminary

In this paragraph we explain the results on cycles used in the introduction. We describe the surjective morphism $s : \mathrm{Pic}(U)^\vee \to A_1(X)$ and define the set of classes $\mathfrak{ne}(\alpha)$ for $\alpha \in A_1(X)$.

If X is a scheme of dimension n, we denote by $Z_*(X)$ the group of 1-cycles on X and by $Z_*^{\equiv}(X)$ and $Z_*^r(X)$ the subgroups of cycles trivial for the numerical and rational equivalence. Let us denote by $N_*(X)$ and $A_*(X)$ the corresponding quotients. The Picard group is the image in $A_{n-1}(X)$ of the subgroup of Cartier divisors in $Z_{n-1}(X)$.

Lemma 1.1. — *Let $X = \mathfrak{C}_{L,n}(G/P)$ be a cone over a homogeneous variety G/P then*

(i) $\mathrm{Pic}(X) \simeq N^1(X)$,

(ii) $A_1(X) \simeq N_1(X)$.

In particular we have $A_1(X) \simeq \mathrm{Pic}(X)^\vee$ and they are isomorphic to $\mathbb{Z}$.

Proof. Consider the decomposition $V \oplus H^0L$. The following group

$$G' = \begin{pmatrix} GL(V) & \mathrm{Hom}(H^0L, V) \\ 0 & G \end{pmatrix}$$

acts on X and the unipotent part $U(G')$ acts on X with finitely many orbits. Remark that it is not the case if we only take the unipotent part U of $GL(V)\times G$. Indeed, take for example $V = \mathbb{C}$ and $G = GL(V) = GL_1(\mathbb{C})$. Then G/P is a point and X is a projective line but U is the trivial group and has infinitely many orbits in X. The same problem appears in the general case.

(i) Thanks to the results of [FMcPSS] the groups $A_*(X)$ are free generated by invariant subvarieties. The Picard group is contained in $A_{n-1}(X)$ and is in particular free. Thanks to [Fu] Example 19.3.3. this implies that $\mathrm{Pic}(X) \simeq N^1(X)$.

(ii) The results of [FMcPSS] also imply that $A_1(X)$ is generated by the one-dimensional invariant subvarieties. The only such subvariety is

- the fibre of the cone over the 0-dimensional orbit in G/P if $\dim V = 1$;
- the 1 dimensional orbit in $\mathbb{P}(V)$ if $\dim V \geq 2$.

We get the isomorphism $A_1(X) \simeq \mathbb{Z}$ with this variety as generator. This generator is clearly numerically free (for example its degree is 1) so we get the result.

The duality comes from general duality between $N_1(X)$ and $N^1(X)$. □

Let U be the smooth locus of X, it is also the dense orbit under G' in X. Let Y be the complementary of U in X, it is of codimension at least 2 (at least when $\dim(G/P) > 0$). This in particular implies that $\mathrm{Pic}(U) = A_{n-1}(U) \simeq A_{n-1}(X)$. We now have the following inclusion:

$$\mathrm{Pic}(X) \subset A_{n-1}(X) \simeq \mathrm{Pic}(U)$$

giving the surjection

$$s : \mathrm{Pic}(U)^\vee \to A_1(X).$$

With these notations we make the following:

Definition 1.2. — *Let $\alpha \in A_1(X)$. We define the set $\mathfrak{ne}(\alpha) \subset A_{n-1}(X)^\vee$. Let us make the identification $A_{n-1}(X) \simeq \mathrm{Pic}(U)$. The elements of $\mathfrak{ne}(\alpha)$ are the elements $\beta \in \mathrm{Pic}(U)^\vee$ such that $s(\beta) = \alpha$ and there exists a complete curve $C \subset U$ with $[C] = \beta$ as a linear form on $\mathrm{Pic}(U)$ (β is effective).*

Let us describe $\mathfrak{ne}(\alpha)$ explicitly: the smooth part U is a vector bundle over G/P. In particular we have $\mathrm{Pic}(U) \simeq \mathrm{Pic}(G/P)$.
Let us fix T a maximal Torus in G, fix B a Borel subgroup containing T and suppose that $B \subset P$. Let us denote by Δ the set of all roots, by Δ^+ (resp. Δ^-) the set of positive (resp. negative) roots and by S the set of simple roots associated to the data (G, T, B).
Denote by $\mathfrak{g}$, $\mathfrak{t}$ and $\mathfrak{p}$ the Lie algebras of G, T and P and define

$$\alpha(\mathfrak{p}) = \Big\{ \alpha \in S \ / \ \mathfrak{g}_\alpha \subset \mathfrak{p} \text{ and } \mathfrak{g}_{-\alpha} \not\subset \mathfrak{p} \Big\}.$$

Now set $\mathfrak{t}(\mathfrak{p})^*$ as the subvector space of $\mathfrak{t}^*$ generated by the roots in $\alpha(\mathfrak{p})$, we have

$$\mathrm{Pic}(G/P) \simeq \mathfrak{t}(\mathfrak{p}) \cap Q$$

where $\mathfrak{t}(\mathfrak{p})$ is the dual of $\mathfrak{t}(\mathfrak{p})^*$ in $\mathfrak{t}$ and Q is the weight lattice. The Picard group of X in $\mathrm{Pic}(U) \simeq \mathrm{Pic}(G/P) \simeq \mathfrak{t}(\mathfrak{p}) \cap Q$ is given by the the line generated by λ (the weight associated to L). We have

$$\mathrm{Pic}(U)^\vee \simeq \mathfrak{t}^*/\mathfrak{t}(\mathfrak{p})^* \cap R$$

where R is the root lattice. Furthermore, an element $\beta \in \mathrm{Pic}(U)^\vee$ gives an effective element if and only if it is in the image of the cone generated by positive roots i.e. in the cone $\mathfrak{t}^*/\mathfrak{t}(\mathfrak{p})^* \cap R^+$ (see [P1]). Then we have

$$\mathfrak{ne}(\alpha) = \big\{ \beta \in \mathfrak{t}^*/\mathfrak{t}(\mathfrak{p})^* \cap R^+ \ / \ \langle \beta^\vee, \lambda \rangle = \alpha \cdot L \big\}$$

where the integer $\langle \beta^\vee, \lambda \rangle$ is well defined because $\lambda \in \mathfrak{t}(\mathfrak{p}) \cap Q$.

Example 1.3. — *Choose for L (or for λ) the smallest ample sheave on X. This is possible: the picard group $\mathrm{Pic}(U) = \mathfrak{t}(\mathfrak{p}) \cap Q$ is a direct sum of weight lattices of semi-simple Lie algebras $(\mathfrak{g}_i)_{i \in [1,r]}$. We just have to take*

$$\lambda = \sum_{i \in [1,r]} \rho_i$$

where ρ_i is half the sum of positive roots in $\mathfrak{g}_i$.
Let us denote by $\mathfrak{ir}_{\hat{\mathfrak{g}}_i}(\ell)$ the set of isomorphism classes of irreductible integrable representations of level exactely ℓ of the affine Lie algebra $\hat{\mathfrak{g}}_i$. Then we have

$$\mathfrak{ne}(\alpha) = \prod_{\ell_1 + \cdots + \ell_r = \alpha \cdot L} \mathfrak{ir}_{\hat{\mathfrak{g}}_i}(\ell_i).$$

In particular if $P = B$ is a Borel subgroup of G then $r = 1$ and $G_1 = G$ and we recover the example of the introduction:

$$\mathfrak{ne}(\alpha) = \mathfrak{ir}_{\hat{\mathfrak{g}}}(\alpha \cdot L).$$

Remark 1.4. — (i) *The scheme $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ is the scheme of morphisms from $\mathbb{P}^1$ to X of class α (for more details see [Gr] and [Mo]).*
In general, this will just mean that $\alpha \in A_1(X)$ and that $f_*[\mathbb{P}^1] = \alpha$ but sometimes (in particular in the introduction for the open part U) we consider $\alpha \in \mathrm{Pic}(X)^\vee$ and the class of a morphism $f : \mathbb{P}^1 \to X$ will be the linear form $\mathrm{Pic}(X) \to \mathbb{Z}$ given by $L \mapsto \deg(f^*L)$. In the case of a homogeneous cone X the two notion coincide because of the previous lemma.
(ii) *If X is a variety, $\alpha \in A_1(X)$ and F a vector bundle on X we will denote $\alpha \cdot F = \int_\alpha c_1(F)$ by abuse of notation.*

2 Resolution

Recall that we denote by X the cone $\mathfrak{C}_{L,n}(G/P)$. Let $\widetilde{X}$ be the blowing-up of X in $\mathbb{P}(V)$. It is smooth and isomorphic to

$$\mathbb{P}_{G/P}((V \otimes \mathcal{O}_{G/P}) \oplus L).$$

Let us denote by p the projection from $\widetilde{X}$ to G/P and by $\pi : \widetilde{X} \to X$ the blowing-up. The morphism p has natural sections given by points of $\mathbb{P}(V)$ or equivalently by surjective morphisms $L \oplus (V \otimes \mathcal{O}_{G/P}) \to V \otimes \mathcal{O}_{G/P} \to \mathcal{O}_{G/P}$.

2.1 Cycles on $\widetilde{X}$

Lemma 2.1. — (i) *Rational and numerical equivalences coincide on $\widetilde{X}$. In particular we have $A_1(\widetilde{X}) \simeq \mathrm{Pic}(\widetilde{X})^\vee \simeq A_{n-1}(\widetilde{X})^\vee$.*
(ii) *We have $\mathrm{Pic}(\widetilde{X}) \simeq \mathrm{Pic}(G/P) \oplus \mathbb{Z}$ with the factor $\mathbb{Z}$ generated by a p-relative ample class.*

Proof. (i) Rational and numerical equivalence coincide on G/P. Moreover the fibration in projetive spaces $\widetilde{X} \to G/P$ has sections so that rational and numerical equivalences coincide on $\widetilde{X}$. This in particular implies that $\mathrm{Pic}(\widetilde{X}) = A_{n-1}(\widetilde{X}) = N^1(\widetilde{X})$ and $A_1(\widetilde{X}) = N_1(\widetilde{X})$ and the duality follows.
(ii) The variety $\widetilde{X}$ is a $\mathbb{P}^n$-bundle over G/P with sections so we get that

$$\mathrm{Pic}(\widetilde{X}) \simeq \mathrm{Pic}(G/P) \oplus \mathbb{Z}$$

with the factor $\mathbb{Z}$ generated by a p-relative ample class, the relative tangent sheaf T_p of p is $n+1$ times this class on the factor $\mathbb{Z}$. $\square$

Any element $\widetilde{\alpha} \in A_1(\widetilde{X}) \simeq \mathrm{Pic}(\widetilde{X})^\vee$ is given by the class $\beta = p_*\widetilde{\alpha} \in A_1(G/P)$ and the relative degree $d = \widetilde{\alpha} \cdot T_p$. We will use the notation $\ell = \beta \cdot L = \widetilde{\alpha} \cdot p^*L$. Let us denote by E the exceptional divisor on $\widetilde{X}$, it is a trivial $\mathbb{P}^{n-1}$ bundle over G/P given by the surjection $L \oplus (V \otimes \mathcal{O}_{G/P}) \to V \otimes \mathcal{O}_{G/P}$. Then we have:

$$\widetilde{\alpha} \cdot E = \frac{d - n\ell}{n+1},$$

it has to be an integer so that $d \equiv n\ell \mod n+1$.

Let us consider the following morphism still denoted by p:

$$p : \mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X}) \to \mathbf{Hom}_\beta(\mathbb{P}^1, G/P).$$

Proposition 2.2. — *Thanks to the morphism p, the scheme $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})$ is an open subset of a projective bundle over $\mathbf{Hom}_\beta(\mathbb{P}^1, G/P)$.*

Proof. This generalises proposition 4 of [P1] in the case where the relative degree $\widetilde{\alpha} \cdot T_p$ is negative. This is possible because the vector bundle associated to the $\mathbb{P}^n$ fibration has a decomposition $L \oplus (V \otimes \mathcal{O}_{G/P})$. We only describe the fibers, for the structure of projective bundle see [P1] proposition 4.
Let $f : \mathbb{P}^1 \to G/P$, we have to calculate the fiber of p above f. The fiber is given by sections of the $\mathbb{P}^n$-bundle $f^*(p) : \mathbb{P}_{\mathbb{P}^1}((V \otimes \mathcal{O}_{\mathbb{P}^1}) \oplus \mathcal{O}_{\mathbb{P}^1}(\ell)) \to \mathbb{P}^1$ whose relative degree is d. In other words the fiber is given by surjections $(V \otimes \mathcal{O}_{\mathbb{P}^1}) \oplus \mathcal{O}_{\mathbb{P}^1}(\ell) \to \mathcal{O}_{\mathbb{P}^1}(x)$ modulo scalar multiplication where $d = (n+1)x - \ell$. The fiber is therefore isomorphic to an open subset of $\mathbb{P}(\mathrm{Hom}((V \otimes \mathcal{O}_{\mathbb{P}^1}) \oplus \mathcal{O}_{\mathbb{P}^1}(\ell), \mathcal{O}_{\mathbb{P}^1}(x))$.
Let us remark that if $\mathbf{Hom}_\beta(\mathbb{P}^1, G/P)$ is not empty then we have $\ell \geq 0$ and in this case $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})$ is not empty if and only if $x \geq 0$ when $n \geq 2$ and if and only if $x = 0$ or $x \geq \ell$ when $n = 1$. In terms of d this means that $d = -\ell$ or $d \geq n\ell$ if $n = 1$ and $d \geq -\ell$ if $n \geq 2$. In any cases, if $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})$ is not empty then $x \geq 0$.
There are two cases:

- If $x < \ell$ then any section is included in the exceptional divisor and the dimension of the fiber is:

$$\frac{n}{n+1}(\ell + d) + n - 1.$$

- If $x \geq \ell$ then the fiber is of dimension $d + n$. $\square$

Let $\widetilde{\alpha} \in A_1(\widetilde{X})$ such that $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})$ is not empty. This is equivalent to the fact that $\beta \in A_1(G/P)$ is positive (see [P1], it is equivalent to the fact that $\mathbf{Hom}_\beta(\mathbb{P}^1, G/P)$ is non empty) and such that $d = -\ell$ or $d \geq n\ell$ if $n = 1$, $d \geq -\ell$ if $n \geq 2$ (recall that $\ell = \beta \cdot L$).

Corollary 2.3. — *The scheme* $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ *is irreducible of dimension*

- $\displaystyle\int_{\widetilde{\alpha}} c_1(T_{\widetilde{X}})+\dim(\widetilde{X})$ *if* $d\geq n\ell$
- $\displaystyle\int_{\widetilde{\alpha}} c_1(T_{\widetilde{X}})+\dim(\widetilde{X})-\widetilde{\alpha}\cdot E-1$ *if* $d<n\ell$.

Proof. We just use the preceding proposition and the fact proved in [P1] that the scheme $\mathbf{Hom}_{\beta}(\mathbb{P}^1,G/P)$ is irreductible of dimension $\int_\beta c_1(T_{G/P})+\dim(G/P)$. Remark that in the last case we have $n\ell>d$ so that the dimension of $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ is still greater than the expected dimension $\displaystyle\int_{\widetilde{\alpha}} c_1(T_{\widetilde{X}})+\dim(\widetilde{X})$. □

2.2 Smoothing curves on $\widetilde{X}$

In this paragraph we will prove some results on curves on $\widetilde{X}$.

Proposition 2.4. — *Assume that* $L(R)=\mathbb{Z}$.
Let $\widetilde{\alpha}\in A_1(\widetilde{X})$, $\widetilde{f}\in\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ *such that* $\widetilde{f}(\mathbb{P}^1)\not\subset E$ *and* $\widetilde{\alpha}\cdot E>0$. *Assume that the image of* $p\circ\widetilde{f}:\mathbb{P}^1\to G/P$ *is not a line in the embedding given by* L. *Then there exists a deformation* $\widetilde{f}'$ *of* $\widetilde{f}$ *and a curve* $\Gamma\subset\widetilde{X}$ *contracted by* π *with* $\Gamma\cdot E=-1$ *such that the curve* $\widetilde{f}'(\mathbb{P}^1)\cup\Gamma$ *can be smoothed. The smoothed curve is the image of a morphism* $\widehat{f}:\mathbb{P}^1\to\widetilde{X}$.

Proof. Let $(x,v)\in E\simeq G/P\times\mathbb{P}(V)$ be a point in the intersection $\widetilde{f}(\mathbb{P}^1)\cap E$. Let us first remark that we may assume (after deformation) that $\widetilde{f}(\mathbb{P}^1)$ is a nodal curve. Indeed, because $p\circ\widetilde{f}:\mathbb{P}^1\to G/P$ is not a line, this implies in particular that G/P is not $\mathbb{P}^1$ so it is of dimension at least 2. The results of [P1] prove that a general curve in G/P is nodal and so is $\widetilde{f}(\mathbb{P}^1)$ if $\widetilde{f}$ is general.

Lemma 2.5. — *There exists a deformation* $\widetilde{f}'$ *of* $\widetilde{f}$ *and a rational curve* Γ *in* $\widetilde{X}$ *such that* $[\Gamma]\cdot E=-1$, $[\Gamma]\cdot L=1$ *and meeting* $\widetilde{f}'(\mathbb{P}^1)$ *in exactly one point.*

Proof. Let us consider the lines in G/P that is to say the rational curves Γ' in G/P such that $[\Gamma']\cdot L=1$. Such curves exist because we have $L(R)=\mathbb{Z}$. Let Γ' be such a line passing through $p(x,v)=x\in G/P$ and let Γ be the section of Γ' in E given by the point $v\in\mathbb{P}(V)$. This curve is contracted by π to the point $v\in\mathbb{P}(V)$, its intersection with E is given by $-[\Gamma']\cdot L=-1$.
As we assumed that $p\circ\widetilde{f}(\mathbb{P}^1)$ is not a line then Γ' meets $p\circ\widetilde{f}(\mathbb{P}^1)$ in a finite number of points: x and other points (x_i). The morphism $\widetilde{f}$ is given by a section of the projective bundle over $p\circ\widetilde{f}$ that is to say by a surjection

$$s:(V\otimes\mathcal{O}_{\mathbb{P}^1})\oplus\mathcal{O}_{\mathbb{P}^1}(\ell)\to\mathcal{O}_{\mathbb{P}^1}\left(\frac{d+\ell}{n+1}\right).$$

To deform $\widetilde{f}$ we can deform this surjection s in s' such that at x, we have $s'_x = s_x$ and at x_i we have $s'_{x_i} \neq s_{x_i}$ for all i. This gives the required deformation. $\square$

Lemma 2.6. — *The curve $\widetilde{f}'(\mathbb{P}^1) \cup \Gamma$ can be smoothed. The smoothed curve is the image of a morphism $\widehat{f} : \mathbb{P}^1 \to \widetilde{X}$ of class $\widehat{\alpha}$ with*

$$\widehat{\alpha} \cdot (p^*L + E) = \widetilde{\alpha} \cdot (p^*L + E) \quad \text{and} \quad \widehat{\alpha} \cdot E < \widetilde{\alpha} \cdot E.$$

Proof. If the smoothing exists then we have $\widehat{\alpha} = \widetilde{\alpha} + [\Gamma]$ so $\widehat{\alpha} \cdot E = \widetilde{\alpha} \cdot E - 1$. Furthermore we have $p^*L + E = \pi^*L$ so that

$$\widehat{\alpha} \cdot (p^*L + E) = \widehat{\alpha} \cdot \pi^*L = \pi_*\widehat{\alpha} \cdot L = \pi_*\widetilde{\alpha} \cdot L = \widetilde{\alpha} \cdot \pi^*L = \widetilde{\alpha} \cdot (p^*L + E).$$

This simply comes from the fact that $\pi_*[\Gamma] = 0$. Let us note that the curves $f' = \pi \circ \widetilde{f}$ and $f'' = \pi \circ \widehat{f}$ have the same degrees but the curve f'' meets the vertex in one point less than f'.

To smooth $\widetilde{f}'(\mathbb{P}^1) \cup \Gamma$ we use the following result proved in [HH] for $\mathbb{P}^3$ but valid for any smooth projective variety:

Theorem 2.7. — *Let Z be a smooth projective variety and let C be a nodal curve in Z. Assume that the cohomology group $H^1T_Z|_C$ is trivial then C can be smoothed.*

As $\widetilde{f}(\mathbb{P}^1)$ is nodal and thanks to the previous lemma we know that $\widetilde{f}'(\mathbb{P}^1) \cup \Gamma$ is nodal. We just have to prove that the cohomology group $H^1(T_{\widetilde{X}}|_{\widetilde{f}'(\mathbb{P}^1)\cup\Gamma})$ is trivial. We have the exact sequences

$$0 \to T_p \to T_{\widetilde{X}} \to p^*T_{G/P} \to 0 \quad \text{and} \quad 0 \to \mathcal{O}_{\widetilde{f}'(\mathbb{P}^1)}(-Q) \to \mathcal{O}_{\widetilde{f}'(\mathbb{P}^1)\cup\Gamma} \to \mathcal{O}_\Gamma \to 0$$

where Q is the intersection point of $\widetilde{f}'(\mathbb{P}^1)$ and Γ. We just have to prove the vanishing of the following cohomology groups:

$$H^1(p^*T_{G/P}|_\Gamma) \ ; \ H^1(p^*T_{G/P}|_{\widetilde{f}'(\mathbb{P}^1)}(-Q)) \ ; \ H^1(T_p|_\Gamma) \text{ and } H^1(T_p|_{\widetilde{f}'(\mathbb{P}^1)}(-Q)).$$

The first two groups are respectively equal to $H^1(T_{G/P}|_{\Gamma'})$ and $H^1(T_{G/P}|_{p(\widetilde{f}'(\mathbb{P}^1))}(-Q))$ where we denoted $\Gamma' = p(\Gamma)$. They are trivial because $T_{G/P}$ is globally generated and Γ' and $p(\widetilde{f}'(\mathbb{P}^1))$ are rational curves. Let us denote by $\mathcal{O}_p(1)$ the tautological quotient of the projective bundle associated to $(V \otimes \mathcal{O}_X) \oplus L$, the relative tangent sheaf is given by $T_p = \mathrm{Coker}(\mathcal{O}_{\widetilde{X}} \to ((V^\vee \otimes \mathcal{O}_{\widetilde{X}}) \oplus L^\vee) \otimes \mathcal{O}_p(1))$. In particular we have:

$$T_p|_\Gamma = \mathrm{Coker}(\mathcal{O}_{\mathbb{P}^1} \to (V^\vee \otimes \mathcal{O}_{\mathbb{P}^1}) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)) \qquad \text{and}$$

$$T_p|_{\widetilde{f}'(\mathbb{P}^1)} = \mathrm{Coker}\left(\mathcal{O}_{\mathbb{P}^1} \to \left(V^\vee \otimes \mathcal{O}_{\mathbb{P}^1}\left(\frac{d+\ell}{n+1}\right)\right) \oplus \mathcal{O}_{\mathbb{P}^1}\left(\frac{d-n\ell}{n+1}\right)\right).$$

This proves that the group $H^1(T_\pi|_\Gamma)$ vanishes. Furthermore, since $\widetilde{f}$ exists we must have $\frac{d+\ell}{n+1} \geq 0$ (see proposition 2.2) and $\frac{d-n\ell}{n+1} = \widetilde{\alpha} \cdot E > 0$ so that $H^1(T_p|_{\widetilde{f}'(\mathbb{P}^1)}(-Q))$ also vanishes. $\square$

3 Homogeneous cones

Recall that we denote by X the cone $\mathfrak{C}_{L,n}(G/P)$. In this paragraph we study the irreducible components of the scheme $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ where $\alpha \in A_1(X)$. Recall that $A_1(X) \simeq \mathbb{Z}$ and under this identification α is just the degree of the corresponding curve.

3.1 The case $L(R) = \mathbb{Z}$

Theorem 3.1. — *Assume that $L(R) = \mathbb{Z}$, let $\alpha \in A_1(X)$ and $f \in \mathbf{Hom}_\beta(\mathbb{P}^1, X)$. Then there exists a deformation f' of f such that f' does not meet the vertex $\mathbb{P}(V)$ of the cone X.*

Proof. Let us begin with the following:

Lemma 3.2. — *Let $f \in \mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ such that f factors through the vertex $\mathbb{P}(V)$ of the cone. Then there exists a deformation f' of f in $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ such that $f'(\mathbb{P}^1)$ does not factor through the vertex.*

Proof. Let $x \in G/P$ and consider the linear subspace generated by x and $\mathbb{P}(V)$. It is a projective space contained in X containing $\mathbb{P}(V)$ as a hyperplane and containing $f(\mathbb{P}^1)$. In this projective space we can deform the morphism f so that it does not factor through $\mathbb{P}(V)$ any more. $\square$

A general morphism $f \in \mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ does not factor through the vertex $\mathbb{P}(V)$ of the cone so it can be lifted in a morphism $\widetilde{f} : \mathbb{P}^1 \to \widetilde{X}$. Let $\widetilde{\alpha} \in A_1(\widetilde{X})$ the class of $\widetilde{f}$, we have $\pi_*\widetilde{\alpha} = \alpha$. Because f does not factor through the vertex, the morphism $\widetilde{f}$ does not factor through the exceptional divisor E so we have: $\widetilde{\alpha} \cdot E \geq 0$. If $\widetilde{\alpha} \cdot E = 0$, then $\widetilde{f}(\mathbb{P}^1)$ does not meet E thus f does not meet the vertex and we are done. Let us assume that $\widetilde{\alpha} \cdot E > 0$. We proceed by induction on $\widetilde{\alpha} \cdot E$. Consider the morphism $p \circ \widetilde{f} : \mathbb{P}^1 \to G/P$.

Lemma 3.3. — *If the image of $p \circ \widetilde{f}$ is a line in the projective embedding given by L then there exists a deformation $f' \in \mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ of f not meeting the vertex.*

Proof. Indeed, if the image of $p \circ \widetilde{f}$ is a line then f factors through the linear subspace generated by the vertex and this line. It is a $\mathbb{P}^{n+1}$ and the vertex is a linear subspace of codimension 2. There exists a deformation f' of f in this projective space not meeting the vertex. $\square$

Let us now assume that the image of $p \circ \widetilde{f}$ is not a line, we may apply proposition 2.4 so that there exists a deformation $\widetilde{f}'$ of $\widetilde{f}$ and a curve $\Gamma \subset \widetilde{X}$ contracted by π with $\Gamma \cdot E = -1$ such that the curve $\widetilde{f}'(\mathbb{P}^1) \cup \Gamma$ can be smoothed. The

smoothed curve is the image of a morphism $\widehat{f}:\mathbb{P}^1\to\widetilde{X}$ of class $\widehat{\alpha}$. Let us consider $f'=\pi\circ\widetilde{f}'$ and $f''=\pi\circ\widehat{f}$. Then f' is a deformation of f and because Γ is contracted by π the map f'' is a deformation of f' and a fortiori of f.
We have to prove the result on f'' whose lifting is $\widehat{f}$ of class $\widehat{\alpha}$. But we have $\widehat{\alpha}\cdot E=\widetilde{\alpha}\cdot E-1$ so the result is true by induction. □

Theorem 3.4. — *Assume $L(R)=\mathbb{Z}$ and let $\alpha\in A_1(X)$ then the irreducible components of the scheme $\mathbf{Hom}_\alpha(\mathbb{P}^1,X)$ are indexed by $\mathfrak{ne}(\alpha)$. For $\widetilde{\alpha}\in\mathfrak{ne}(\alpha)$ the dimension of the corresponding component is*

$$\int_{\widetilde{\alpha}}c_1(T_{\widetilde{X}})+\dim(X).$$

Proof. Theorem 3.1 proves that the set of morphisms $f:\mathbb{P}^1\to X$ whose image does not meet the vertex $\mathbb{P}(V)$ is a dense open subset of $\mathbf{Hom}_\alpha(\mathbb{P}^1,X)$. It is enough to study this open set. Any curve is this open set comes from a unique lifting $\widetilde{f}:\mathbb{P}^1\to\widetilde{X}$ whose image does not meet E. Let $\widetilde{\alpha}\in A_1(\widetilde{X})$ the class of $\widetilde{f}$, since $\widetilde{\alpha}\cdot E=0$ we have $\widetilde{\alpha}\in\mathrm{Pic}(U)^\vee$ and in fact $\widetilde{\alpha}\in\mathfrak{ne}(\alpha)$. The morphism

$$\pi_*:\coprod_{\widetilde{\alpha}\in\mathfrak{ne}(\alpha)}\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})\to\mathbf{Hom}_\alpha(\mathbb{P}^1,X)$$

is thus dominant and birational (the inverse is given by lifting morphisms). What is left to prove is that for each $\widetilde{\alpha}\in\mathfrak{ne}(\alpha)$ the image of $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ (which is an irreducible scheme) forms an irreducible component of $\mathbf{Hom}_\alpha(\mathbb{P}^1,X)$. To prove this it is enough to prove that for any $\widetilde{\alpha}$ and $\widetilde{\alpha}'$ in $\mathfrak{ne}(\alpha)$ the image of $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ is not contained in the closure of $\mathbf{Hom}_{\widetilde{\alpha}'}(\mathbb{P}^1,\widetilde{X})$ in $\mathbf{Hom}_\alpha(\mathbb{P}^1,X)$. This would be trivial if the scheme $\coprod_{\widetilde{\alpha}\in\mathfrak{ne}(\alpha)}\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ was equidimensional (it is the case if $L=\frac{1}{2}c_1(G/P)$). In general, suppose there exist $f\in\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ such that f does not meet the vertex and such that f is the limit of a familly f'_t of morphisms in $\mathbf{Hom}_{\widetilde{\alpha}'}(\mathbb{P}^1,\widetilde{X})$. Because the condition of meeting the vertex is closed me may assume that the elements f'_t do not meet the vertex. In particular projecting on G/P gives a deformation from $p(f'_t)$ to $p(f)$. This implies that $p_*\widetilde{\alpha}=p_*\widetilde{\alpha}'$ but as $\widetilde{\alpha}\cdot E=0=\widetilde{\alpha}'\cdot E$ we have $\widetilde{\alpha}=\widetilde{\alpha}'$. The dimension comes from corollary 2.3. □

3.2 The case $L(R)\neq\mathbb{Z}$

We begin with the following lemma on root systems:

Lemma 3.5. — *Let G be a semi-simple Lie group, $P\subset G$ a parabolic subgroup, L a dominant weight in the facet defined by P and R the root lattice, then we have the equivalence*

$$L(R)\neq\mathbb{Z}\Longleftrightarrow L\geq c_1(G/P)$$

where $c_1(G/P) \in \mathrm{Pic}(G/P)$ is considered as a weight and the order is given by the positivity on simple roots.

PROOF. Let us first describe $c_1(G/P)$ as a weight. Consider the set $\alpha(\mathfrak{p})$ of simple root and the lattice $\mathfrak{t}(\mathfrak{p})\cap Q$ (which is isomorphic to $\mathrm{Pic}(G/P)$) defined in paragraph 1. The lattice $\mathfrak{t}(\mathfrak{p})\cap Q$ decomposes into a direct sum of root lattices R_i. Let ρ_i be half the sum of positive roots of the root system corresponding to R_i. Then we have

$$c_1(G/P) = 2\sum_i \rho_i.$$

If $L \geq c_1(G/P)$ then for any simple root α we have

$$\langle \alpha^\vee, L\rangle \geq \langle \alpha^\vee, c_1(G/P)\rangle = \sum_i \langle \alpha^\vee, \rho_i\rangle = \left\{ \begin{array}{ll} 0 & \text{if } \alpha \notin \alpha(\mathfrak{p}) \\ 2 & \text{if } \alpha \in \alpha(\mathfrak{p}) \end{array}\right.$$

and in particular $1 \notin L(R)$.
Conversely, suppose that $L(R) \neq \mathbb{Z}$. Because L is in the facet of P we have $\langle \alpha^\vee, L\rangle = 0$ for any simple root $\alpha \notin \alpha(\mathfrak{p})$. If α is a simple root in $\alpha(\mathfrak{p})$ then $\langle \alpha^\vee, L\rangle \geq 2$ (otherwise $L(R) = \mathbb{Z}$). We see that for any simple root $\langle \alpha^\vee, L\rangle \geq \langle \alpha^\vee, c_1(G/P)\rangle$ thus $L \geq c_1(G/P)$. $\square$

REMARK 3.6. — *Let $\widetilde{\alpha} \in A_1(\widetilde{X})$ such that $\widetilde{\alpha}\cdot E \geq 0$. Recall the notations $\beta = p_*\widetilde{\alpha}$, $d = \widetilde{\alpha}\cdot T_p$ is the relative degree and $\ell = \widetilde{\alpha}\cdot p^*L = \beta\cdot L$. Let $\alpha = \pi_*\widetilde{\alpha}$ considered as an integer. Then the dimension of $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})$ is given by*

$$\begin{aligned}
\int_{\widetilde{\alpha}} c_1(T_{\widetilde{X}}) + \dim(\widetilde{X}) &= \int_\beta c_1(T_{G/P}) + d + \dim(\widetilde{X}) \\
&= \int_\beta c_1(T_{G/P}) + (n+1)\widetilde{\alpha}\cdot E + n\ell + \dim(\widetilde{X}) \\
&= \beta\cdot(c_1(T_{G/P}) - L) + (n+1)\widetilde{\alpha}\cdot(E + p^*L) + \dim(\widetilde{X}) \\
&= \beta\cdot(c_1(T_{G/P}) - L) + (n+1)\alpha + \dim(\widetilde{X}).
\end{aligned}$$

So we have the formula

$$\dim(\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})) = \int_{\widetilde{\alpha}} p^*(c_1(T_{G/P}) - L) + (n+1)\alpha + \dim(\widetilde{X}).$$

THEOREM 3.7. — *Assume $L(R) \neq \mathbb{Z}$ and let $\alpha \in A_1(X)$. Then the irreducible components of $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ are indexed by $\coprod_{\alpha' \leq \alpha} \mathfrak{ne}(\alpha')$.*

PROOF. Thanks to lemma 3.2 (this lemma works without the hypothesis $L(R) = \mathbb{Z}$) there exists a dense open subset of $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ given by morphisms f that do not factor through the vertex of the cone. It is enough to study this open set. In particular we know that the morphism

$$\pi_* : \coprod_{\widetilde{\alpha}\in A_1(\widetilde{X}),\ \pi_*\widetilde{\alpha}=\alpha} \mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X}) \to \mathbf{Hom}_{\alpha}(\mathbb{P}^1,X)$$

is dominant. The classes $\widetilde{\alpha}$ can even be choosen such that $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ is not empty. However the intersection $\widetilde{\alpha}\cdot E$ need not to be 0. In particular the classes $\widetilde{\alpha}$ can be choosen in

$$A(\alpha) = \coprod_{\alpha'\leq\alpha} \mathfrak{ne}(\alpha')$$

where $\alpha' = p_*\widetilde{\alpha}\cdot L$ and $\alpha-\alpha' = \widetilde{\alpha}\cdot E$. Indeed let $\widetilde{\alpha}\in A_1(\widetilde{X})$ and set as usual $\beta = p_*\widetilde{\alpha}$. Then there exists a unique element $\widetilde{\alpha}'\in A_1(\widetilde{X})$ such that $p_*\widetilde{\alpha}' = \beta$ and $\widetilde{\alpha}'\cdot E = 0$ (take $n\beta\cdot L$ for the relative degree). If $\widetilde{\alpha}$ is such that $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ is not empty then β is effective and because of the value of the relative degree we have that $\mathbf{Hom}_{\widetilde{\alpha}'}(\mathbb{P}^1,\widetilde{X})$ is not empty. In particular $\widetilde{\alpha}'\in\mathfrak{ne}(\alpha')$ for $\alpha' = \pi_*\widetilde{\alpha}' = p_*\widetilde{\alpha}\cdot L$ and we have $\widetilde{\alpha}\cdot E = \pi_*\widetilde{\alpha} - \pi_*\widetilde{\alpha}'$. The element $\widetilde{\alpha}$ is uniquely determined by $\widetilde{\alpha}'$ and $\widetilde{\alpha}\cdot E$.

It is enough to prove that the images by π_* of the irreducible schemes $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ for $\widetilde{\alpha}\in A(\alpha)$ are the irreducible components. In other words we have to prove that for any $\widetilde{\alpha}$ and $\widehat{\alpha}$ in $A(\alpha)$ the image of $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ is not contained in the closure of the image of $\mathbf{Hom}_{\widehat{\alpha}}(\mathbb{P}^1,\widetilde{X})$ in $\mathbf{Hom}_{\alpha}(\mathbb{P}^1,X)$.

Let $\widetilde{f}\in\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ a generic point and $\widehat{f}_t$ a familly of morphisms in $\mathbf{Hom}_{\widehat{\alpha}}(\mathbb{P}^1,\widetilde{X})$ such that $\pi\circ\widehat{f}_t$ converges to $\pi\circ\widetilde{f}$. In the compactification of $\mathbf{Hom}_{\widehat{\alpha}}(\mathbb{P}^1,\widetilde{X})$ by stable maps (see for example [FP]), the familly $\widehat{f}_t$ has a limit say $\widehat{f}$ which is a morphism from a tree $\cup_i D_i$ of rational curves to $\widetilde{X}$. Then we must have $\pi\circ\widehat{f} = \pi\circ\widetilde{f}$ as stable maps. In particular all but one of the images by $\widehat{f}$ of the irreducible components of the tree are contracted by π. To fix notation say that D_i is contracted by π for $i\geq 2$ and $\pi\circ\widehat{f}|_{D_1} = \pi\circ\widetilde{f}$. Because $\widetilde{f}$ is generic, it is not contained in the exceptional divisor so that the equality $\pi\circ\widehat{f}|_{D_1} = \pi\circ\widetilde{f}$ implies that $\widehat{f}|_{D_1} = \widetilde{f}$. We see that $\widehat{f}_*[D_1] = \widetilde{\alpha}$ so that

$$\widehat{\alpha} = \widehat{f}_*[D_1] + \sum_{i\geq 2}\widehat{f}_*[D_i] = \widetilde{\alpha} + \sum_{i\geq 2}\widehat{f}_*[D_i].$$

In particular we have $\widehat{\beta} = p_*\widehat{\alpha} \geq p_*\widetilde{\alpha} = \widetilde{\beta}$ and because $L(R)\neq\mathbb{Z}$ we know thanks to lemma 3.5 that $L\geq c_1(G/P)$ and we get

$$\widehat{\beta}\cdot(c_1(T_{G/P})-L) \leq \widetilde{\beta}\cdot(c_1(T_{G/P})-L).$$

As $\alpha = \pi_*\widehat{\alpha} = \pi_*\widetilde{\alpha}$ we see that

$$\dim(\mathbf{Hom}_{\widehat{\alpha}}(\mathbb{P}^1,\widetilde{X})) \leq \dim(\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})).$$

But the morphism π_* is generically injective on $\mathbf{Hom}_{\widehat{\alpha}}(\mathbb{P}^1,\widetilde{X})$ and $\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X})$ so that the scheme $\pi_*(\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1,\widetilde{X}))$ cannot be in the closure of $\pi_*(\mathbf{Hom}_{\widehat{\alpha}}(\mathbb{P}^1,\widetilde{X}))$. □

REMARK 3.8. — *Let us end with a discussion on the dimensions of the irreducible components of* $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ *for* $\alpha \in A_1(X)$.

(I) *In the first case:* $L(R) = \mathbb{Z}$, *these irreducible components are indexed by elements* $\widetilde{\alpha}$ *in* $\mathfrak{ne}(\alpha)$. *For such an element we have* $\widetilde{\alpha} \cdot E = 0$ *and the dimension of the component is given by*

$$\dim(\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})) = \int_{\widetilde{\alpha}} p^*(c_1(T_{G/P}) - L) + (n+1)\alpha + \dim(\widetilde{X}).$$

The "variable" part in this dimension is the first one and it is given by

$$\beta \cdot (c_1(T_{G/P}) - L)$$

with $\beta = p_*\widetilde{\alpha}$ *and we have* $\alpha = \beta \cdot L$ *so that the "variable" part is* $\beta \cdot c_1(T_{G/P})$. *The element* β *ranges in the subset of the positive cone in the root lattice* R *(in the projection of* R *in* $\mathrm{Pic}(G/P)$*) given by the condition* $\beta \cdot L = \alpha$. *In particular if* L *is not collinear to* $c_1(G/P)$ *the dimensions of the irreducible components are not equal. In this case the variety* $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ *is equidimensional if and only if* $L = \frac{1}{2}c_1(G/P)$.

(II) *In the second case:* $L(R) \neq \mathbb{Z}$, *these irreducible components are indexed by elements* $\widetilde{\alpha} \in \coprod_{\alpha' \leq \alpha} \mathfrak{ne}(\alpha')$. *For such an element we have* $\widetilde{\alpha} \cdot E \geq 0$ *and the dimension of the component is given by*

$$\dim(\mathbf{Hom}_{\widetilde{\alpha}}(\mathbb{P}^1, \widetilde{X})) = \int_{\widetilde{\alpha}} p^*(c_1(T_{G/P}) - L) + (n+1)\alpha + \dim(\widetilde{X}).$$

The "variable" part in this dimension is the first one and it is given by

$$\beta \cdot (c_1(T_{G/P}) - L)$$

with $\beta = p_*\widetilde{\alpha}$. *In this case we have* $\beta \cdot L = \alpha' \leq \alpha$. *The element* β *ranges in the subset of the positive cone in the root lattice* R *(in the projection of* R *in* $\mathrm{Pic}(G/P)$*) given by the condition* $\beta \cdot L \leq \alpha$. *In particular if* L *is not collinear to* $c_1(G/P)$ *the dimensions of the irreducible components are not equal (look at the* β *such that* $\beta \cdot L = \alpha$*). Furthermore even if* L *is collinear to* $c_1(G/P)$ *the dimensions of the irreducible components are not equal unless* $L = c_1(G/P)$. *In this case the variety* $\mathbf{Hom}_\alpha(\mathbb{P}^1, X)$ *is equidimensional if and only if* $L = c_1(G/P)$.

REFERENCES

[BP] *Michel Brion and Patrick Polo*: Generic singularities of certain Schubert varieties. Math. Z. 231 (1999), no. 2, 301–324.

[Fu] *William Fulton*: Intersection theory. Second edition. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 2. Springer-Verlag, Berlin (1998).

[FMcPSS] *William Fulton, Robert MacPherson, Frank Sottile and Bernd Sturmfels*: Intersection theory on spherical varieties. J. Algebraic Geom. 4 (1995), no. 1.

[FP] *William Fulton, Rahul Pandharipande*: Notes on stable maps and quantum cohomology. Algebraic geometry—Santa Cruz 1995, 45–96, Proc. Sympos. Pure Math., 62, Part 2, Amer. Math. Soc., Providence, RI, 1997.

[Gr] *Alexander Grothendieck*: Techniques de construction et théorèmes d'existence en géométrie algébrique. IV. Les schémas de Hilbert. (French). Séminaire Bourbaki, Vol. 6, Exp. No. 221, 249–276, Soc. Math. France, Paris, (1995).

[HH] *Robin Hartshorne and André Hirschowitz*: Smoothing algebraic space curves. Algebraic geometry, Sitges (Barcelona), 98–131, L.N.M., 1124, Springer, Berlin, (1985).

[KP] *Bumsiq Kim and Rahul Pandharipande*: The connectedness of the moduli space of maps to homogeneous spaces. Symplectic geometry and mirror symmetry (Seoul, 2000), 187–201, World Sci. Publishing, River Edge, NJ, 2001.

[LMS] *Venkatramani Lakshmibai, Chitikila Musili and Conjeerveram S. Seshadri*: Geometry of G/P. III. Standard monomial theory for a quasi-minuscule P. Proc. Indian Acad. Sci. Sect. A Math. Sci. 88 (1979), no. 3, 93–177.

[Mo] *Shigefumi Mori*: Projective manifolds with ample tangent bundles. Ann. of Math. (2) 110 (1979), no. 3, 593–606.

[P1] *Nicolas Perrin*: Courbes rationnelles sur les variétés homogènes. Annales de l'Institut Fourier, 52, no.1 (2002), pp 105-132.

[P2] *Nicolas Perrin*: Rational curves on minuscule Schubert varieties, preprint math.AG/0407123.

[Th] *Jesper Funch Thomsen*: Irreducibility of $\overline{M}_{0,n}(G/P,\beta)$, Internat. J. Math. 9 (1998) No. 3.

Nicolas Perrin
Institut de Mathématiques de Jussieu
175 rue du Chevaleret
75013 Paris, France
nperrin@math.jussieu.fr

Matsumoto K-Groups Associated to Certain Shift Spaces

Toke Meier Carlsen and Søren Eilers

Received: May 7, 2004
Revised: December 30, 2004

Communicated by Joachim Cuntz

Abstract. In [24] Matsumoto associated to each shift space (also called a subshift) an Abelian group which is now known as Matsumoto's K_0-group. It is defined as the cokernel of a certain map and resembles the first cohomology group of the dynamical system which has been studied in for example [2], [28], [13], [16] and [11] (where it is called the dimension group).

In this paper, we will for shift spaces having a certain property $(*)$, show that the first cohomology group is a factor group of Matsumoto's K_0-group. We will also for shift spaces having an additional property $(**)$, describe Matsumoto's K_0-group in terms of the first cohomology group and some extra information determined by the left special elements of the shift space.

We determine for a broad range of different classes of shift spaces if they have property $(*)$ and property $(**)$ and use this to show that Matsumoto's K_0-group and the first cohomology group are isomorphic for example for finite shift spaces and for Sturmian shift spaces.

Furthermore, the ground is laid for a description of the Matsumoto K_0-group as an *ordered* group in a forthcoming paper.

2000 Mathematics Subject Classification: Primary 37B10, Secondary 54H20, 19K99.
Keywords and Phrases: Shift spaces, subshifts, symbolic dynamics, Matsumoto's K-groups, dimension groups, cohomology, special elements.

1 Introduction

Invariants for symbolic dynamical systems in the form of Abelian groups have a fruitful history. Important examples are the dimension group defined by Krieger in [19] and [20], and the Bowen-Franks group defined in [1] by Bowen and Franks.

In [24] Matsumoto generalized the definition of dimension groups and Bowen-Franks groups to the whole class of shift spaces and introduced what is now known as Matsumoto's K-groups.

In another direction, Putnam [29], Herman, Putnam and Skau [16], Giordano, Putnam and Skau [15], Durand, Host and Skau [11] and Forrest [13] studied what they called the dimension group (it is not the same as Krieger's or Matsumoto's dimension group) for Cantor minimal systems. The same group has for a broader class of topological dynamical systems been studied in [2], [28] and [27] where it is shown that it is the first cohomology group of the standard suspension of the dynamical system in question.

It turns out that Matsumoto's K_0-group and the first cohomology group are closely related. We will for shift spaces having a certain property $(*)$, show that the first cohomology group is a factor group of Matsumoto's K_0-group, and we will also for shift spaces having an additional property $(**)$, describe Matsumoto's K_0-group in terms of the first cohomology group and some extra information determined by the left special elements of the shift space.

We will for a broad range of different classes of shift spaces, which includes shift of finite types, finite shift spaces, Sturmian shift spaces, substitution shift spaces and Toeplitz shift spaces, determine if they have property $(*)$ and property $(**)$. This will allow us to show that Matsumoto's K_0-group and the first cohomology group are isomorphic for example for finite shift spaces and for Sturmian shift spaces and to describe Matsumoto's K_0-group for substitution shift spaces in such a way that we in [8] can for every shift space associated with a aperiodic and primitive substitution present Matsumoto's K_0-group as a stationary inductive limit of a system associated to an integer matrix defined from combinatorial data which can be computed in an algorithmic way (cf. [6], [7]).

Since both Matsumoto's K_0-group and the first cohomology group are K_0-groups of certain C^*-algebras they come with a natural (pre)order structure. All the results presented in this paper hold not just in the category of Abelian groups, but also in the category of preordered groups. Since we do not know how to prove this without involving C^*-algebras we have decided to defer this to [9], where we also show that Matsumoto's K_0-group with order is a finer invariant than Matsumoto's K_0-group without order.

We wish to thank Yves Lacroix for helping us understand Toeplitz sequences and the referee for constructive criticism.

2 Preliminaries and notation

Throughout this paper $\mathbb{Z}$ will denote the set of integers, $\mathbb{N}_0$ will denote the set of non-negative integers and $-\mathbb{N}$ will denote the negative integers.
The symbol Id will always denote the identity map. For a map ϕ between two sets X and Y, we will by $\phi^\star$ denote the map which maps a function f on Y to the function $f \circ \phi$ on X.
Let $\mathfrak{a}$ be a finite set of symbols, and let $\mathfrak{a}^\sharp$ denote the set of finite, nonempty words with letters from $\mathfrak{a}$. Thus with ϵ denoting the *empty word*, $\epsilon \notin \mathfrak{a}^\sharp$. By $|\mu|$ we denote the *length* of a finite word μ (i.e. the number of letters in μ). The length of ϵ is 0.

2.1 Shift spaces

We equip

$$\mathfrak{a}^{\mathbb{Z}}, \mathfrak{a}^{\mathbb{N}_0}, \mathfrak{a}^{-\mathbb{N}}$$

with the product topology from the discrete topology on $\mathfrak{a}$. We will strive to denote elements of $\mathfrak{a}^{\mathbb{Z}}$ by z, elements of $\mathfrak{a}^{\mathbb{N}_0}$ by x and elements of $\mathfrak{a}^{-\mathbb{N}}$ by y. If $x \in \mathfrak{a}^{\mathbb{N}_0}$ and $y \in \mathfrak{a}^{-\mathbb{N}}$, then we will by $y.x$ denote the element z of $\mathfrak{a}^{\mathbb{Z}}$ where

$$z_n = \begin{cases} y_n & \text{if } n < 0, \\ x_n & \text{if } n \geq 0. \end{cases}$$

We define $\sigma : \mathfrak{a}^{\mathbb{Z}} \to \mathfrak{a}^{\mathbb{Z}}$, $\sigma_+ : \mathfrak{a}^{\mathbb{N}_0} \to \mathfrak{a}^{\mathbb{N}_0}$, and $\sigma_- : \mathfrak{a}^{-\mathbb{N}} \to \mathfrak{a}^{-\mathbb{N}}$ by

$$(\sigma(z))_n = z_{n+1} \qquad (\sigma_+(x))_n = x_{n+1} \qquad (\sigma_-(y))_n = y_{n-1}.$$

Such maps we will refer to as *shift maps.*
A *shift space* is a closed subset of $\mathfrak{a}^{\mathbb{Z}}$ which is mapped into itself by σ. We shall refer to such spaces by "$\underline{\mathsf{X}}$".
With the obvious restriction maps

$$\pi_+ : \underline{\mathsf{X}} \to \mathfrak{a}^{\mathbb{N}_0} \qquad \pi_- : \underline{\mathsf{X}} \to \mathfrak{a}^{-\mathbb{N}}$$

we get

$$\sigma_+ \circ \pi_+ = \pi_+ \circ \sigma \qquad \sigma_- \circ \pi_- = \pi_- \circ \sigma^{-1}.$$

We denote $\pi_+(\underline{\mathsf{X}})$, respectively $\pi_-(\underline{\mathsf{X}})$, by $\underline{\mathsf{X}}^+$, respectively $\underline{\mathsf{X}}^-$, and notice that $\sigma_+(\underline{\mathsf{X}}^+) = \underline{\mathsf{X}}^+$ and $\sigma_-(\underline{\mathsf{X}}^-) = \underline{\mathsf{X}}^-$. For $z \in \mathfrak{a}^{\mathbb{Z}}$ and $n \in \mathbb{Z}$, we write

$$z_{[n,\infty[} = \pi_+(\sigma^n(z)) \text{ and } z_{]-\infty,n[} = \pi_-(\sigma^n(z)).$$

The *language* of a shift space is the subset of $\mathfrak{a}^\sharp \cup \{\epsilon\}$ given by

$$\mathcal{L}(\underline{\mathsf{X}}) = \{z_{[n,m]} \mid z \in \underline{\mathsf{X}}, n \leq m \in \mathbb{Z}\}$$

where the interval subscript notation should be self-explanatory. A compactness argument shows that an element $z \in \mathfrak{a}^{\mathbb{Z}}$ (respectively $z \in \mathfrak{a}^{\mathbb{N}_0}$, $z \in \mathfrak{a}^{-\mathbb{N}}$)

is in $\underline{\mathsf{X}}$ (respectively $\underline{\mathsf{X}}^+$, $\underline{\mathsf{X}}^-$) if and only if $z_{[n,m]} \in \mathcal{L}(\underline{\mathsf{X}})$ for all $n < m \in \mathbb{Z}$ (respectively $n < m \in \mathbb{N}_0$, $n < m \in -\mathbb{N}$) (cf. [21, Corollary 1.3.5 and Theorem 6.1.21]).

We say that shift spaces are *conjugate*, denoted by "$\simeq$", when they are homeomorphic via a map which intertwines the relevant shift maps. The concept of conjugacy also makes sense for the "one-sided" shift spaces $\underline{\mathsf{X}}^+$. If $\underline{\mathsf{X}}^+ \simeq \underline{\mathsf{Y}}^+$, then we say that $\underline{\mathsf{X}}$ and $\underline{\mathsf{Y}}$ are *one-sided conjugate*. It is not difficult to see that $\underline{\mathsf{X}}^+ \simeq \underline{\mathsf{Y}}^+ \Rightarrow \underline{\mathsf{X}} \simeq \underline{\mathsf{Y}}$ (cf. [21, §13.8]).

Finally we want to draw attention to a third kind of equivalence between shift spaces, called *flow equivalence*, which we denote by $\cong_f$. We will not define it here (see [26], [14], [2] or [21, §13.6] for the definition), but just notice that $\underline{\mathsf{X}} \simeq \underline{\mathsf{Y}} \Rightarrow \underline{\mathsf{X}} \cong_f \underline{\mathsf{Y}}$.

A *flow invariant* of a shift space $\underline{\mathsf{X}}$ is a mapping associating to each shift space another mathematical object, called the *invariant*, in such a way that flow equivalent shift spaces give isomorphic invariants. In the same way, a *conjugacy invariant* of $\underline{\mathsf{X}}$, respectively $\underline{\mathsf{X}}^+$, is a mapping associating to each shift space an invariant in such a way that conjugate, respectively one-sided conjugate, shift spaces give isomorphic invariants.

Since $\underline{\mathsf{X}} \simeq \underline{\mathsf{Y}} \Rightarrow \underline{\mathsf{X}} \cong_f \underline{\mathsf{Y}}$, a flow invariant of $\underline{\mathsf{X}}$ is also a conjugacy invariant of $\underline{\mathsf{X}}$, and since $\underline{\mathsf{X}}^+ \simeq \underline{\mathsf{Y}}^+ \Rightarrow \underline{\mathsf{X}} \simeq \underline{\mathsf{Y}}$, a conjugacy invariant of $\underline{\mathsf{X}}$ is also a conjugacy invariant of $\underline{\mathsf{X}}^+$.

2.2 Special elements

We say (cf. [17]) that $z \in \underline{\mathsf{X}}$ is *left special* if there exists $z' \in \underline{\mathsf{X}}$ such that

$$z_{-1} \neq z'_{-1} \qquad \pi_+(z) = \pi_+(z').$$

It follows from [4, Proposition 2.4.1] (cf. [3, Theorem 3.9]) that a sufficient condition for a shift space $\underline{\mathsf{X}}$ to have a left special element is that $\underline{\mathsf{X}}$ is infinite. Conversely, the following proposition shows that this condition is necessary.

Proposition 2.1. *Let $\underline{\mathsf{X}}$ be a finite shift space. Then $\underline{\mathsf{X}}$ contains no left special element.*

Proof: Since $\underline{\mathsf{X}}$ is finite, every $z \in \underline{\mathsf{X}}$ is periodic. Hence if $\pi_+(z) = \pi_+(z')$, then $z = z'$. □

We say that the left special word z is *adjusted* if $\sigma^{-n}(z)$ is not left special for any $n \in \mathbb{N}$, and that z is *cofinal* if $\sigma^n(z)$ is not left special for any $n \in \mathbb{N}$. Thinking of left special words as those which are not deterministic from the right at index -1, the adjusted and cofinal left special words are those where this is the *leftmost* and *rightmost* occurrence of nondeterminacy, respectively.

Let $z, z' \in \underline{\mathsf{X}}$. If there exist an n and an M such that $z_m = z'_{n+m}$ for all $m > M$ then we say that z and z' are *right shift tail equivalent* and write $z \sim_r z'$. We will denote the right shift tail equivalence class of z by $\mathbf{z}$.

2.3 The first cohomology group

The first cohomology group (cf. [2]) of a shift space $\underline{\mathsf{X}}$ is the group

$$C(\underline{\mathsf{X}},\mathbb{Z})/(\operatorname{Id}-(\sigma^{-1})^{\star})(C(\underline{\mathsf{X}},\mathbb{Z})).$$

Notice that usually σ is used instead of σ^{-1}, but for our purpose it is more natural to use σ^{-1}, and we of course get the same group. The group $C(\underline{\mathsf{X}},\mathbb{Z})/(\operatorname{Id}-(\sigma^{-1})^{\star})(C(\underline{\mathsf{X}},\mathbb{Z}))$ is the first Čech cohomology group of the standard suspension of $(\underline{\mathsf{X}},\sigma)$ (cf. [27, IV.15. Theorem]). It is also isomorphic to the homotopy classes of continuous maps from the standard suspension of $(\underline{\mathsf{X}},\sigma)$ into the circle (cf. [27, page 60]).
It is proved in [2, Theorem 1.5] that $C(\underline{\mathsf{X}},\mathbb{Z})/(\operatorname{Id}-(\sigma^{-1})^{\star})(C(\underline{\mathsf{X}},\mathbb{Z}))$ is a flow invariant of $\underline{\mathsf{X}}$ and thus also a conjugacy invariant of $\underline{\mathsf{X}}$ and $\underline{\mathsf{X}}^+$.

2.4 Past equivalence and Matsumoto's K_0-group

Let $\underline{\mathsf{X}}$ be a shift space. For every $x\in\underline{\mathsf{X}}^+$ and every $k\in\mathbb{N}$ we set

$$\mathcal{P}_k(x)=\{\mu\in\mathcal{L}(\underline{\mathsf{X}})\mid \mu x\in\underline{\mathsf{X}}^+,\ |\mu|=k\},$$

and define for every $l\in\mathbb{N}$ an equivalence relation $\sim_l$ on $\underline{\mathsf{X}}^+$ by

$$x\sim_l x' \iff \mathcal{P}_l(x)=\mathcal{P}_l(x').$$

Likewise we let for every $x\in\underline{\mathsf{X}}^+$

$$\mathcal{P}_\infty(x)=\{y\in\underline{\mathsf{X}}^-\mid y.x\in\underline{\mathsf{X}}\},$$

and define an equivalence relation $\sim_\infty$ on $\underline{\mathsf{X}}^+$ by

$$x\sim_\infty x' \iff \mathcal{P}_\infty(x)=\mathcal{P}_\infty(x').$$

The set

$$\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)=\{x\in\underline{\mathsf{X}}^+\mid \exists k\in\mathbb{N}:\#\mathcal{P}_k(x)>1\}$$

then consists exactly of all words on the form $z_{[n,\infty[}$ where z is left special and $n\in\mathbb{N}_0$.
Following Matsumoto ([23]), we denote by $[x]_l$ the equivalence class of x and refer to the relation as *l-past equivalence*.
Obviously the set of equivalence classes of the l-past equivalence relation $\sim_l$ is finite. We will denote the number of such classes $m(l)$ and enumerate them $\mathcal{E}^l_s$ with $s\in\{1,\dots,m(l)\}$. For each $l\in\mathbb{N}$, we define an $m(l+1)\times m(l)$-matrix I^l by

$$(\mathsf{I}^l)_{rs}=\begin{cases}1 & \text{if } \mathcal{E}^{l+1}_r\subseteq\mathcal{E}^l_s\\ 0 & \text{otherwise,}\end{cases}$$

and note that I^l induces a group homomorphism from $\mathbb{Z}^{m(l)}$ to $\mathbb{Z}^{m(l+1)}$. We denote by $\mathbb{Z}_{\underline{\mathsf{X}}}$ the group given by the inductive limit

$$\varinjlim(\mathbb{Z}^{m(l)}, \mathsf{I}^l).$$

For a subset $\mathcal{E}$ of $\underline{\mathsf{X}}^+$ and a finite word μ we let $\mu\mathcal{E} = \{\mu x \in \underline{\mathsf{X}}^+ \mid x \in \mathcal{E}\}$. For each $l \in \mathbb{N}$ and $a \in \mathfrak{a}$ we define an $m(l+1) \times m(l)$-matrix

$$(\mathsf{L}_a^l)_{rs} = \begin{cases} 1 & \text{if } \emptyset \neq a\mathcal{E}_r^{l+1} \subseteq \mathcal{E}_s^l \\ 0 & \text{otherwise,} \end{cases}$$

and letting $\mathsf{L}^l = \sum_{a\in\mathfrak{a}} \mathsf{L}_a^l$ we get a matrix inducing a group homeomorphism from $\mathbb{Z}^{m(l)}$ to $\mathbb{Z}^{m(l+1)}$. Since one can prove that $\mathsf{L}^{l+1} \circ \mathsf{I}^l = \mathsf{I}^{l+1} \circ \mathsf{L}^l$, a group endomorphism λ on $\mathbb{Z}_{\underline{\mathsf{X}}}$ is induced.

Theorem 2.2 (Cf. [24], [25, Theorem]). *Let $\underline{\mathsf{X}}$ be a shift space. The group*

$$K_0(\underline{\mathsf{X}}) = \mathbb{Z}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \lambda)\mathbb{Z}_{\underline{\mathsf{X}}},$$

called Matsumoto's K_0-group*, is a conjugacy invariant of $\underline{\mathsf{X}}$ and $\underline{\mathsf{X}}^+$, and a flow invariant of $\underline{\mathsf{X}}$.*

2.5 The space $\Omega_{\underline{\mathsf{X}}}$

We will now give an alternative description of $K_0(\underline{\mathsf{X}})$. The group $K_0(\underline{\mathsf{X}})$ is defined by taking a inductive limits of $\mathbb{Z}^{m(l)}$, where $\mathbb{Z}^{m(l)}$ could be thought of as $C(\underline{\mathsf{X}}^+/\sim_l, \mathbb{Z})$.

We will now do things in different order. First we will take the projective limit of $\underline{\mathsf{X}}^+/\sim_l$ and then look at the continuous functions from the projective limit to $\mathbb{Z}$.

Since $\sim_l$ is coarser than $\sim_{l+1}$, there is a projection π_l of $\underline{\mathsf{X}}^+/\sim_{l+1}$ onto $\underline{\mathsf{X}}^+/\sim_l$.

Definition 2.3 (Cf. [23, page 682]). *Let $\underline{\mathsf{X}}$ be a shift space. We then define $\Omega_{\underline{\mathsf{X}}}$ to be the compact topological space given by the projective limit*

$$\varprojlim(\underline{\mathsf{X}}^+/\sim_l, \pi_l).$$

We will identify $\Omega_{\underline{\mathsf{X}}}$ with the closed subspace

$$\{([x_n]_n)_{n\in\mathbb{N}_0} \mid \forall n \in \mathbb{N}_0 : x_{n+1} \sim_n x_n\}$$

of $\prod_{l=0}^{\infty} \underline{\mathsf{X}}^+/\sim_l$, where $\prod_{l=0}^{\infty} \underline{\mathsf{X}}^+/\sim_l$ is endowed with the product of the discrete topologies.

Notice that if we identify $C(\underline{\mathsf{X}}^+/\sim_l, \mathbb{Z})$ with $\mathbb{Z}^{m(l)}$, then I^l is the map induced by π_l, so $C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})$ can be identified with $\mathbb{Z}_{\underline{\mathsf{X}}}$.

If $([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$, then

$$\{([x'_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}} \mid x'_1 \sim_1 x_1\}$$

is a clopen subset of $\Omega_{\underline{\mathsf{X}}}$, and if $a \in \mathcal{P}_1(x_1)$, then $([ax'_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$ for every $([x'_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$ with $x'_1 \sim_1 x_1$, and the map

$$([x'_n]_n)_{n\in\mathbb{N}_0} \mapsto ([ax'_n]_n)_{n\in\mathbb{N}_0}$$

is a continuous map on $\{([x'_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}} \mid x'_1 \sim_1 x_1\}$. This allows us to define a map $\lambda_{\underline{\mathsf{X}}} : C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}) \to C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})$ in the following way:

Definition 2.4. *Let $\underline{\mathsf{X}}$ be a shift space, $h \in C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})$ and $([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$. Then we let*

$$\lambda_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}) = \sum_{a\in\mathcal{P}_1(x_1)} h([ax_n]_{n\in\mathbb{N}_0}).$$

Under the identification of $C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})$ and $\mathbb{Z}_{\underline{\mathsf{X}}}$, $\lambda_{\underline{\mathsf{X}}}$ is equal to λ, thus we have the following proposition:

Proposition 2.5. *Let $\underline{\mathsf{X}}$ be a shift space. Then $K_0(\underline{\mathsf{X}})$ and*

$$C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})/(\mathrm{Id} - \lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}))$$

are isomorphic as groups.

3 Property (*) and (**)

We will introduce the properties $(*)$ and $(**)$ and show that they are invariant under flow equivalence and thus under conjugacy. At the end of the section, we will for various examples of shift spaces determine if they have property $(*)$ and $(**)$.

Definition 3.1. *We say that a shift space $\underline{\mathsf{X}}$ has* property $(*)$ *if for every $\mu \in \mathcal{L}(\underline{\mathsf{X}})$ there exists an $x \in \underline{\mathsf{X}}^+$ such that $\mathcal{P}_{|\mu|}(x) = \{\mu\}$.*

Definition 3.2. *We say that a shift space $\underline{\mathsf{X}}$ has* property $(**)$ *if it has property $(*)$ and if the number of left special words of $\underline{\mathsf{X}}$ is finite, and no such left special word is periodic.*

Since flow equivalence is generated by conjugacy and symbolic expansion (cf. [25, Lemma 2.1] and [26]), it is, in order to prove the following proposition, enough to check that $(*)$ and $(**)$ are invariant under symbolic expansion and conjugacy.

Proposition 3.3. *The properties $(*)$ and $(**)$ are invariant under flow equivalence.*

Example 3.4. It follows from Proposition 2.1 that if a shift space $\underline{\mathsf{X}}$ is finite, then it contains no left special element, and thus has property $(**)$.

Example 3.5. An infinite shift of finite type does not have property $(*)$.

Proof: Let $\underline{\mathsf{X}}$ be a shift of finite type. This means (cf. [21, Chapter 2]) that there is a $k \in \mathbb{N}_0$ such that

$$\underline{\mathsf{X}} = \{z \in \mathfrak{a}^{\mathbb{Z}} \mid \forall n \in \mathbb{Z} : z_{[n,n+k]} \in \mathcal{L}(\underline{\mathsf{X}})\}.$$

Suppose that $\underline{\mathsf{X}}$ has property $(*)$. Let $\mathcal{L}(\underline{\mathsf{X}})_k = \{\mu \in \mathcal{L}(\underline{\mathsf{X}}) \mid |\mu| = k\}$, and notice that if $\mu, \nu, \omega \in \mathcal{L}(\underline{\mathsf{X}})_k$ and $\mu\nu, \nu\omega \in \mathcal{L}(\underline{\mathsf{X}})$, then $\mu\nu\omega \in \mathcal{L}(\underline{\mathsf{X}})$.
Let $\mu \in \mathcal{L}(\underline{\mathsf{X}})_k$. Then there is a $x \in \underline{\mathsf{X}}^+$ such that $\mathcal{P}_{|\mu|}(x) = \{\mu\}$. Let $\mu' = x_{[0,k[}$, and suppose that $\nu \in \mathcal{L}(\underline{\mathsf{X}})_k$ and $\nu\mu' \in \mathcal{L}(\underline{\mathsf{X}})$. Then $\nu x \in \underline{\mathsf{X}}^+$, so ν must be equal to μ. Thus there is for every $\mu \in \mathcal{L}(\underline{\mathsf{X}})_k$ a $\mu' \in \mathcal{L}(\underline{\mathsf{X}})_k$ such that

$$\nu \in \mathcal{L}(\underline{\mathsf{X}})_k \wedge \nu\mu' \in \mathcal{L}(\underline{\mathsf{X}}) \iff \nu = \mu.$$

Since $\mathcal{L}(\underline{\mathsf{X}})_k$ is finite and the map $\mu \mapsto \mu'$ is injective, there is for every $\nu \in \mathcal{L}(\underline{\mathsf{X}})_k$ a $\mu \in \mathcal{L}(\underline{\mathsf{X}})_k$ such that $\nu = \mu'$. Hence there is for every $\mu \in \mathcal{L}(\underline{\mathsf{X}})_k$ a unique $\mu' \in \mathcal{L}(\underline{\mathsf{X}})_k$ such that $\mu\mu' \in \mathcal{L}(\underline{\mathsf{X}})$ and a unique $\mu'' \in \mathcal{L}(\underline{\mathsf{X}})_k$ such that $\mu''\mu \in \mathcal{L}(\underline{\mathsf{X}})$. Thus every $z \in \underline{\mathsf{X}}$ is determined by $z_{[0,k[}$, but since $\mathcal{L}(\underline{\mathsf{X}})_k$ is finite, this implies that $\underline{\mathsf{X}}$ is finite. □

Example 3.6. An infinite minimal shift space (cf. [21, §13.7]) $\underline{\mathsf{X}}$ has property $(**)$ precisely when the number of left special words of $\underline{\mathsf{X}}$ is finite.

Proof: Since no elements in such a shift space is periodic, we only need to prove that property $(*)$ follows from finiteness of the number of left special elements. Let $\mu \in \mathcal{L}(\underline{\mathsf{X}})$ and pick any $x \in \underline{\mathsf{X}}^+$. Since $\underline{\mathsf{X}}^+$ is infinite and minimal, x is not periodic, and since the set of left special words is finite there exists $N \in \mathbb{N}$ such that $\sigma^n(x)$ is not left special for any $n \geq N$. Since $\underline{\mathsf{X}}^+$ is minimal there exists a $k \geq N$ such that $x_{[k+1,k+|\mu|]} = \mu$. Hence $\mathcal{P}_{|\mu|}(\sigma^{k+|\mu|+1}(x)) = \{\mu\}$. □

Example 3.7. If z is a non-periodic, non-regular Toeplitz sequence (cf. [32, pp. 97 and 99]), then the shift space

$$\overline{\mathcal{O}(z)} = \overline{\{\sigma^n(z) \mid n \in \mathbb{Z}\}},$$

where $\overline{X}$ denotes the closure of X, has property $(*)$.

Proof: Let $\mu \in \mathcal{L}(\overline{\mathcal{O}(z)})$. Since $\overline{\mathcal{O}(z)}$ is minimal (cf. [32, page 97]), there is an $m \in \mathbb{N}$ such that $z_{[-m-|\mu|,-m[} = \mu$. We claim that $\mathcal{P}_{|\mu|}(z_{[-m,\infty[}) = \{\mu\}$.
Assume that $z' \in \overline{\mathcal{O}(z)}$ and $z'_{[-m,\infty[} = z_{[-m,\infty[}$. Then $\pi(z') = \pi(z)$, where π is the factor map of $\overline{\mathcal{O}(z)}$ onto its maximal equicontinuous factor $(G, \hat{1})$ (cf. [32, Theorem 2.2]), because since $z'_{[-m,\infty[} = z_{[-m,\infty[}$, the distance between $\sigma^n(z')$ and $\sigma^n(z)$, and thus the distance between $\hat{1}^n(\pi(z'))$ and $\hat{1}^n(\pi(z))$, goes to 0 as n goes to infinity, but since $\hat{1}$ is equicontinuous, this implies that $\pi(z') = \pi(z)$. Since z is a Toeplitz sequence, it follows from [32, Corollary 2.4]) that $z' = z$. Thus $\mathcal{P}_{|\mu|}(z_{[-m,\infty[}) = \{\mu\}$. □

The following example shows that property $(**)$ does not follow from property $(*)$.

Example 3.8. We will construct a non-regular Toeplitz sequence $z \in \{0,1\}^{\mathbb{Z}}$ such that the shift space

$$\overline{\mathcal{O}(z)} = \overline{\{\sigma^n(z) \mid n \in \mathbb{Z}\}}$$

has infinitely many left special elements and thus does not have property ($**$). We will construct z by using the technique introduced by Susan Williams in [32, Section 4]. We will use the same notation as in [32, Section 4]. We let Y be the full 2-shift $\{0,1\}^{\mathbb{Z}}$ and defined $(p_i)_{i\in\mathbb{N}}$ recursively by setting $p_1 = 3$ and $p_{i+1} = 3^{r_i+i}p_i$ for $i \in \mathbb{N}$, where r_i is as defined in [32, Section 4]. We then have that

$$\frac{p_i\beta_{r_i}}{p_{i+1}} = \frac{2^{r_i}}{3^{r_i+i}} < 3^{-i},$$

so

$$\sum_{i=1}^{\infty} \frac{p_i\beta_{r_i}}{p_{i+1}}$$

converges, and z is non-regular by [32, Proposition 4.1].

Claim. *The shift space $\overline{\mathcal{O}(z)}$ has infinitely many left special elements.*

Proof: Let D be as defined on [32, page 103]. If

$$g \in \pi(\{z' \in D \mid -1 \in \mathrm{Aper}(z')\}),$$

$y, y' \in Y$, $y_{[0,\infty[} = y'_{[0,\infty[}$ and $y_{-1} \neq y'_{-1}$, then $\phi(g,y)_{[0,\infty[} = \phi(g,y')_{[0,\infty[}$ and $\phi(g,y)_{-1} \neq \phi(g,y')_{-1}$, where ϕ is the map define on [32, page 103]. Thus $\phi(g,y)$ and $\phi(g,y')$ are left special elements, and since

$$\pi(\{z' \in D \mid -1 \in \mathrm{Aper}(z')\}) \times \{y \in Y \mid y \text{ is left special}\}$$

is infinite and contained in $\pi(D) \times Y$, on which ϕ is $1-1$, $\overline{\mathcal{O}(z)}$ has infinitely many left special elements. $\square$

4 The first cohomology group is a factor of $K_0(\underline{\mathsf{X}})$

We will now show that if a shift space $\underline{\mathsf{X}}$ has property ($*$), then the first cohomology group is a factor group of $K_0(\underline{\mathsf{X}})$.

Suppose that a shift space $\underline{\mathsf{X}}$ has property ($*$). We can then define a map $\iota_{\underline{\mathsf{X}}}$ from $\underline{\mathsf{X}}^-$ into $\Omega_{\underline{\mathsf{X}}}$ in the following way: For each $y \in \underline{\mathsf{X}}^-$ and each $n \in \mathbb{N}_0$ we choose an $x_n \in \underline{\mathsf{X}}^+$ such that $\mathcal{P}_n(x_n) = \{y_{[-n,-1]}\}$. Then $([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$, and we denote this element by $\iota_{\underline{\mathsf{X}}}(y)$. The map $\iota_{\underline{\mathsf{X}}}$ is obviously injective and continuous.

We denote the map

$$(\iota_{\underline{\mathsf{X}}} \circ \pi_-)^\star : C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}) \to C(\underline{\mathsf{X}}, \mathbb{Z})$$

by κ.

Proposition 4.1. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(*)$. Then there is a surjective group homomorphism $\bar{\kappa}$ from $C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})/(\mathrm{Id}-\lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z}))$ to $C(\underline{\mathsf{X}},\mathbb{Z})/(\mathrm{Id}-(\sigma^{-1})^\star)(C(\underline{\mathsf{X}},\mathbb{Z}))$ which makes the following diagram commute:*

$$\begin{tikzcd}
C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z}) \arrow[r, "\kappa"] \arrow[d] & C(\underline{\mathsf{X}},\mathbb{Z}) \arrow[d] \\
C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})/(\mathrm{Id}-\lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})) \arrow[r, "\bar{\kappa}"] & C(\underline{\mathsf{X}},\mathbb{Z})/(\mathrm{Id}-(\sigma^{-1})^\star)(C(\underline{\mathsf{X}},\mathbb{Z}))
\end{tikzcd}$$

Proof: Let q be the quotient map from $C(\underline{\mathsf{X}},\mathbb{Z})$ to

$$C(\underline{\mathsf{X}},\mathbb{Z})/(\mathrm{Id}-(\sigma^{-1})^\star)(C(\underline{\mathsf{X}},\mathbb{Z})).$$

We will show that 1) $q\circ\kappa$ is surjective and 2) $(\mathrm{Id}-\lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z}))\subseteq\ker(q\circ\kappa)$. This will prove the existence and surjectivity of $\bar{\kappa}$.

1) *$q\circ\kappa$ is surjective:* Given $f\in C(\underline{\mathsf{X}},\mathbb{Z})$. Our goal is to find a function $g\in C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})$ which is mapped to $q(f)$ by $q\circ\kappa$.
Since f is continuous, there are $k,m\in\mathbb{N}$ such that

$$z_{[-k,m]}=z'_{[-k,m]}\Rightarrow f(z)=f(z').$$

Thus

$$z_{[-k-m-1,-1]}=z'_{[-k-m-1,-1]}\Rightarrow f\circ\sigma^{-(m+1)}(z)=f\circ\sigma^{-(m+1)}(z').$$

Define a function g from $\Omega_{\underline{\mathsf{X}}}$ to $\mathbb{Z}$ by

$$g(([x_n]_n)_{n\in\mathbb{N}_0})=\begin{cases} f\circ\sigma^{-(m+1)}(z) & \text{if } \mathcal{P}_{k+m+1}(x_{k+m+1})=\{z_{[-k-m-1,-1]}\},\\ 0 & \text{if } \#\mathcal{P}_{k+m+1}(x_{k+m+1})>1.\end{cases}$$

Then $g\in C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})$, and $g\circ\iota_{\underline{\mathsf{X}}}\circ\pi_-=f\circ\sigma^{-(m+1)}$, so $q\circ\kappa(g)=q(f)$.

2) *$(\mathrm{Id}-\lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z}))\subseteq\ker(q\circ\kappa)$:* Let $g\in C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})$ and $y\in\underline{\mathsf{X}}^-$. Then $\lambda_{\underline{\mathsf{X}}}(g)(\iota_{\underline{\mathsf{X}}}(y))=g(\iota_{\underline{\mathsf{X}}}(\sigma_-(y)))$, so

$$\kappa(\lambda_{\underline{\mathsf{X}}}(g))=g\circ\iota_{\underline{\mathsf{X}}}\circ\pi_-\circ\sigma^{-1},$$

which shows that $(\mathrm{Id}-\lambda_{\underline{\mathsf{X}}})(g)\in\ker(q\circ\kappa)$. □

The following corollary now follows from Proposition 2.5:

Corollary 4.2. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(*)$. Then $C(\underline{\mathsf{X}},\mathbb{Z})/(\mathrm{Id}-(\sigma^{-1})^\star)(C(\underline{\mathsf{X}},\mathbb{Z}))$ is a factor group of $K_0(\underline{\mathsf{X}})$.*

5 K_0 of shift spaces having property (∗∗)

We saw in the last section that if a shift space $\underline{\mathsf{X}}$ has property (∗), then the first cohomology group is a factor group of $K_0(\underline{\mathsf{X}})$. This stems from the fact that property (∗) causes an inclusion of $\underline{\mathsf{X}}^-$ into $\Omega_{\underline{\mathsf{X}}}$, and thus a surjection of $C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})$ onto $C(\underline{\mathsf{X}}^-,\mathbb{Z})$. We will now for shift spaces having property (∗∗) describe K_0 in terms of the first cohomology group and some extra information determined by the left special elements of the shift space.
We will first define the group $\mathcal{G}_{\underline{\mathsf{X}}}$ which is a subgroup of the external direct product of $C(\underline{\mathsf{X}}^-,\mathbb{Z})$ and an infinite product of copies of $\mathbb{Z}$, and isomorphic to $C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})$. Next, we will define the group $G_{\underline{\mathsf{X}}}$ which is the external direct product of $C(\underline{\mathsf{X}},\mathbb{Z})$ and an infinite sum of copies of $\mathbb{Z}$, and has a factor group which is isomorphic to $K_0(\underline{\mathsf{X}})$. We will round off by relating this with the fact that the first cohomology group is a factor group of $K_0(\underline{\mathsf{X}})$ and look at some examples.

Lemma 5.1. *Let $\underline{\mathsf{X}}$ be a shift space which has property (∗). Then*

$$\iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-) = \{([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}} \mid \forall n \in \mathbb{N}_0 : \#\mathcal{P}_n(x_n) = 1\}.$$

Proof: Clearly

$$\iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-) \subseteq \{([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}} \mid \forall n \in \mathbb{N}_0 : \#\mathcal{P}_n(x_n) = 1\}.$$

Suppose $([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$ and $\mathcal{P}_n(x_n) = \{\mu_n\}$ for every $n \in \mathbb{N}_0$. Let for every $n \in \mathbb{N}$, y_{-n} be the first letter of μ_n. Since $y_{[-n,-1]} = \mu_n$ for every $n \in \mathbb{N}$, $y \in \underline{\mathsf{X}}^-$, and clearly $\iota_{\underline{\mathsf{X}}}(y) = ([x_n]_n)_{n\in\mathbb{N}_0}$. □

Denote by $\mathcal{I}_{\underline{\mathsf{X}}}$ the set $\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)/\sim_\infty$ (cf. Section 2.4). We will now define a map $\phi_{\underline{\mathsf{X}}}$ from $\mathcal{I}_{\underline{\mathsf{X}}}$ to $\Omega_{\underline{\mathsf{X}}}$. We see that for $x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$, $([x]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$, and we notice that $x \sim_\infty \tilde{x}$, if and only if $([x]_n)_{n\in\mathbb{N}_0} = ([\tilde{x}]_n)_{n\in\mathbb{N}_0}$. So if we let

$$\phi_{\underline{\mathsf{X}}}([x]_\infty) = ([x]_n)_{n\in\mathbb{N}_0},$$

then $\phi_{\underline{\mathsf{X}}}$ is a well-defined and injective map from $\mathcal{I}_{\underline{\mathsf{X}}}$ to $\Omega_{\underline{\mathsf{X}}}$.

Lemma 5.2. *Let $\underline{\mathsf{X}}$ be a shift space which has property (∗). Then $\iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-) \cap \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}}) = \emptyset$, and if $\underline{\mathsf{X}}$ has property (∗∗), then $\iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-) \cup \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}}) = \Omega_{\underline{\mathsf{X}}}$.*

Proof: If $([x_n]_n)_{n\in\mathbb{N}_0} \in \iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-)$, then according to Lemma 5.1, $\#\mathcal{P}_n(x_n) = 1$ for every $n \in \mathbb{N}_0$, and if $([x_n]_n)_{n\in\mathbb{N}_0} \in \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}})$, then $\#\mathcal{P}_n(x_n) > 1$ for some $n \in \mathbb{N}_0$. Hence $\iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-) \cap \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}}) = \emptyset$.
Suppose that $\underline{\mathsf{X}}$ has property (∗∗). If $([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}} \setminus \iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-)$, then according to Lemma 5.1, there is an $n \in \mathbb{N}_0$ such that $\#\mathcal{P}_n(x_n) > 1$, and since there only are finitely many left special words, $[x_n]_n$ must be finite. Since $[x_k]_k \neq \emptyset$ and $[x_{k+1}]_{k+1} \subseteq [x_k]_k$ for every $k \in \mathbb{N}_0$, this implies that $\bigcap_{k\in\mathbb{N}_0}[x_k]_k$ is not empty. Let $x \in \bigcap_{k\in\mathbb{N}_0}[x_k]_k$. Since $\#\mathcal{P}_n(x) = \#\mathcal{P}_n(x_n) > 1$, $x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$, and since $([x_n]_n)_{n\in\mathbb{N}_0} = \phi_{\underline{\mathsf{X}}}([x]_\infty)$, we have that $([x_n]_n)_{n\in\mathbb{N}_0} \in \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}})$. □

5.1 The group $\mathcal{G}_{\underline{\mathsf{X}}}$

We will from now on assume that $\underline{\mathsf{X}}$ has property $(**)$. Let for every function $h : \Omega_{\underline{\mathsf{X}}} \to \mathbb{Z}$,

$$\gamma_{\underline{\mathsf{X}}}(h) = (h \circ \iota_{\underline{\mathsf{X}}}, (h(\phi_{\underline{\mathsf{X}}}(\mathsf{i})))_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}).$$

It follows from Lemma 5.2 that $\gamma_{\underline{\mathsf{X}}}$ is a bijective correspondence between functions from $\Omega_{\underline{\mathsf{X}}}$ to $\mathbb{Z}$ and pairs $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})$, where g is a function from $\underline{\mathsf{X}}^-$ to $\mathbb{Z}$ and each α_{i} is an integer.

Lemma 5.3. *Let g be a function from $\underline{\mathsf{X}}^-$ to $\mathbb{Z}$ and let for every $\mathsf{i} \in \mathcal{I}_{\underline{\mathsf{X}}}$, α_{i} be an integer. Then $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \gamma_{\underline{\mathsf{X}}}(C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}))$ if and only if there is an $N \in \mathbb{N}_0$ such that*

1. $\forall y, y' \in \underline{\mathsf{X}}^- : y_{[-N,-1]} = y'_{[-N,-1]} \Rightarrow g(y) = g(y')$,
2. $\forall x, x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) : [x]_N = [x']_N \Rightarrow \alpha_{[x]_\infty} = \alpha_{[x']_\infty}$,
3. $\forall x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+),\ y \in \underline{\mathsf{X}}^- : \mathcal{P}_N(x) = \{y_{[-N,-1]}\} \Rightarrow \alpha_{[x]_\infty} = g(y)$.

Proof: A function from $\Omega_{\underline{\mathsf{X}}}$ to $\mathbb{Z}$ is continuous if and only if there is an $N \in \mathbb{N}_0$ such that

$$[x_N]_N = [x'_N]_N \Rightarrow h(([x_n]_n)_{n\in\mathbb{N}_0}) = h(([x'_n]_n)_{n\in\mathbb{N}_0}),$$

for $([x_n]_n)_{n\in\mathbb{N}_0}, ([x'_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$, and since we have that if $y, y' \in \underline{\mathsf{X}}^-$, and $([x_n]_n)_{n\in\mathbb{N}_0} = \iota_{\underline{\mathsf{X}}}(y)$ and $([x'_n]_n)_{n\in\mathbb{N}_0} = \iota_{\underline{\mathsf{X}}}(y')$, then

$$[x_N]_N = [x'_N]_N \iff y_{[-N,-1]} = y'_{[-N,-1]},$$

and if $x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$, $y \in \underline{\mathsf{X}}^-$ and $([x'_n]_n)_{n\in\mathbb{N}_0} = \iota_{\underline{\mathsf{X}}}(y)$, then

$$[x]_N = [x'_N]_N \iff \mathcal{P}_N(x) = \{y_{[-N,-1]}\},$$

the conclusion follows. □

Definition 5.4. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(**)$. We denote $\gamma_{\underline{\mathsf{X}}}(C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}))$ by $\mathcal{G}_{\underline{\mathsf{X}}}$, and we let for every function $g : \underline{\mathsf{X}}^- \to \mathbb{Z}$ and $(\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}} \in \mathbb{Z}^{\mathcal{I}_{\underline{\mathsf{X}}}}$,*

$$\mathcal{A}_{\underline{\mathsf{X}}}(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) = (g \circ \sigma_-, (\tilde{\alpha}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}),$$

where

$$\tilde{\alpha}_{[x]_\infty} = \sum_{\substack{x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) \\ \sigma_+(x') = x}} \alpha_{[x']_\infty} + \sum_{\substack{z \in \underline{\mathsf{X}} \\ z_{[0,\infty[} \notin \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) \\ z_{[1,\infty[} = x}} g(\pi_-(z)).$$

Lemma 5.5. *The map $\mathcal{A}_{\underline{\mathsf{X}}}$ maps $\mathcal{G}_{\underline{\mathsf{X}}}$ into $\mathcal{G}_{\underline{\mathsf{X}}}$, and the following diagram commutes:*

$$\begin{tikzcd}
C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z}) \arrow[r, "\gamma_{\underline{\mathsf{X}}}"] \arrow[d, "\lambda_{\underline{\mathsf{X}}}"'] & \mathcal{G}_{\underline{\mathsf{X}}} \arrow[d, "\mathcal{A}_{\underline{\mathsf{X}}}"] \\
C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z}) \arrow[r, "\gamma_{\underline{\mathsf{X}}}"] & \mathcal{G}_{\underline{\mathsf{X}}}
\end{tikzcd}$$

Proof: Let $h \in C(\Omega_{\underline{\mathsf{X}}},\mathbb{Z})$ and $([x_n]_n)_{n\in\mathbb{N}_0} \in \Omega_{\underline{\mathsf{X}}}$. Then

$$\lambda_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}) = \sum_{a\in\mathcal{P}_1(x_1)} h([ax_n]_{n\in\mathbb{N}_0}).$$

We will show that $\lambda_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}) = \gamma_{\underline{\mathsf{X}}}^{-1} \circ \mathcal{A}_{\underline{\mathsf{X}}} \circ \gamma_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0})$. It will then follow that $\mathcal{A}_{\underline{\mathsf{X}}} = \gamma_{\underline{\mathsf{X}}} \circ \lambda_{\underline{\mathsf{X}}} \circ \gamma_{\underline{\mathsf{X}}}^{-1}$, and thus that $\mathcal{A}_{\underline{\mathsf{X}}}$ maps $\mathcal{G}_{\underline{\mathsf{X}}}$ into $\mathcal{G}_{\underline{\mathsf{X}}}$, and the diagram commutes.
Assume first that $([x_n]_n)_{n\in\mathbb{N}_0} \in \iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-)$. Then $\#\mathcal{P}_1(x_1) = 1$ and

$$\iota_{\underline{\mathsf{X}}}(\sigma_-(\iota_{\underline{\mathsf{X}}}^{-1}(([x_n]_n)_{n\in\mathbb{N}_0}))) = [ax_n]_{n\in\mathbb{N}_0},$$

where $a \in \mathcal{P}_1(x_1)$. Thus

$$\lambda_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}) = h(([ax_n]_n)_{n\in\mathbb{N}_0}) = \gamma_{\underline{\mathsf{X}}}^{-1} \circ \mathcal{A}_{\underline{\mathsf{X}}} \circ \gamma_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}).$$

Now assume that $([x_n]_n)_{n\in\mathbb{N}_0} \in \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}})$ and choose $x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$ such that $\phi_{\underline{\mathsf{X}}}([x]_\infty) = ([x_n]_n)_{n\in\mathbb{N}_0}$. We claim that

$$\sum_{a\in\mathcal{P}_1(x_1)} h([ax_n]_{n\in\mathbb{N}_0}) = \sum_{\substack{x'\in\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)\\ \sigma_+(x')=x}} h(\phi_{\underline{\mathsf{X}}}([x']_\infty)) + \sum_{\substack{z\in\underline{\mathsf{X}}\\ z_{[0,\infty[}\notin\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)\\ z_{[1,\infty[}=x}} h(\iota_{\underline{\mathsf{X}}}(z_{[-\infty,-1]})). \quad (1)$$

To see this let $a \in \mathcal{P}_1(x_1)$. Assume first that $([ax_n]_n)_{n\in\mathbb{N}_0} \in \iota_{\underline{\mathsf{X}}}(\underline{\mathsf{X}}^-)$, and let z be the element of $\mathfrak{a}^{\mathbb{Z}}$ satisfying $z_{]-\infty,0[} = \iota_{\underline{\mathsf{X}}}^{-1}(([ax_n]_n)_{n\in\mathbb{N}_0})$, $z_0 = a$, and $z_{[1,\infty[} = x$. Then $z \in \underline{\mathsf{X}}$, $z_{[0,\infty[} \notin \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$, $z_{[1,\infty[} = x$, and $\iota_{\underline{\mathsf{X}}}(z_{]-\infty,-1]}) = [ax_n]_{n\in\mathbb{N}_0}$. Let us then assume that $([ax_n]_n)_{n\in\mathbb{N}_0} \in \phi_{\underline{\mathsf{X}}}(\mathcal{I}_{\underline{\mathsf{X}}})$. Then $ax \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$, $\sigma_+(ax) = x$, and $\phi_{\underline{\mathsf{X}}}([ax]_\infty) = [ax_n]_{n\in\mathbb{N}_0}$.
If on the other hand z is an element of $\underline{\mathsf{X}}$ which satisfies $z_{[0,\infty[} \notin \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$, and $z_{[1,\infty[} = x$, then $z_0 \in \mathcal{P}_1(x_1)$, and $\iota_{\underline{\mathsf{X}}}(z_{]-\infty,-1]}) = ([z_0x_n]_n)_{n\in\mathbb{N}_0}$, and if $x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$ and $\sigma_+(x') = x$, then $x'_0 \in \mathcal{P}_1(x_1)$, and $\phi_{\underline{\mathsf{X}}}([x']_\infty) = [x'_0x_n]_{n\in\mathbb{N}_0}$.

Thus (1) holds, and

$$\begin{aligned}
\lambda_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}) &= \sum_{a\in\mathcal{P}_1(x_1)} h([ax_n]_{n\in\mathbb{N}_0}) \\
&= \sum_{\substack{x'\in\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)\\ \sigma_+(x')=x}} h(\phi_{\underline{\mathsf{X}}}([x']_\infty)) + \sum_{\substack{z\in\underline{\mathsf{X}}\\ z_{[0,\infty[}\notin\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)\\ z_{[1,\infty[}=x}} h(\iota_{\underline{\mathsf{X}}}(z_{[-\infty,-1]})) \\
&= \gamma_{\underline{\mathsf{X}}}^{-1}\circ\mathcal{A}_{\underline{\mathsf{X}}}\circ\gamma_{\underline{\mathsf{X}}}(h)(([x_n]_n)_{n\in\mathbb{N}_0}).
\end{aligned}$$

□

The following corollary now follows from Proposition 2.5:

COROLLARY 5.6. *Let $\underline{\mathsf{X}}$ be a shift space which has property* $(**)$. *Then* $K_0(\underline{\mathsf{X}})$ *and*

$$\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id}-\mathcal{A}_{\underline{\mathsf{X}}})\mathcal{G}_{\underline{\mathsf{X}}}$$

are isomorphic as groups.

5.2 THE SPACE $\mathcal{I}_{\underline{\mathsf{X}}}$

In order to get a better understanding of the group $\mathcal{G}_{\underline{\mathsf{X}}}$ and the map $\mathcal{A}_{\underline{\mathsf{X}}}$, we will now try to describe $\mathcal{I}_{\underline{\mathsf{X}}}$ in the case where $\underline{\mathsf{X}}$ has properties $(**)$. For that we will need the concept of right shift tail equivalence (cf. section 2.2).

Denote the set of those right shift tail equivalence classes of $\underline{\mathsf{X}}$ which contains a left special element by $\mathcal{J}_{\underline{\mathsf{X}}}$. Notice that it is finite. Let for every $\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}$, $M_{\mathbf{j}}$ be the set of adjusted left special elements belonging to $\mathbf{j}$. Notice that there only is a finite – but positive – number of elements in $M_{\mathbf{j}}$.

Let us take a closer look at $\pi_+(\mathbf{j})$. It is clear that

$$\pi_+(\mathbf{j}) = \{z_{[n,\infty[} \mid z\in M_{\mathbf{j}}, n\in\mathbb{Z}\},$$

and it follows from the definition of adjusted left special elements that $z_{[n,\infty[}\in\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$ if and only if $n\geq 0$. It follows from the definition of adjusted left special elements and the fact that $\underline{\mathsf{X}}$ contains no periodic left special elements that if $z,z'\in M_{\mathbf{j}}$ and $n,n'<0$, then

$$z_{[n,\infty[} = z'_{[n',\infty[} \iff z=z'\wedge n=n'.$$

Contrary to this, it might happen that $z_{[n,\infty[}=z'_{[n',\infty[}$ for $z\neq z'$ if $n,n'\geq 0$. In fact, it turns out that $\mathbf{j}$ has a "common tail".

DEFINITION 5.7. *Let $\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}$. An $x\in\underline{\mathsf{X}}^+$ such that there for every $z\in\mathbf{j}$ is an $n\in\mathbb{Z}$ such that $z_{[n,\infty[}=x$ is called a common tail of $\mathbf{j}$.*

LEMMA 5.8. *Let z be a left special element and $n\in\mathbb{Z}$. Then $z_{[n,\infty[}$ is a common tail of $\mathbf{z}$ if and only if $\sigma^m(z)$ is not left special for any $m>n$.*

Proof: Assume that $\sigma^m(z)$ is not left special for any $m > n$, and let $z' \in \mathbf{z}$. Then there are $k, k' \in \mathbb{Z}$ such that $z_{[k,\infty[} = z'_{[k',\infty[}$, and since $\sigma^m(z)$ is not left special for any $m > n$, $z_{[n,\infty[} = z'_{[n-k+k',\infty[}$ if $k > n$. If $k \leq n$, then obviously $z_{[n,\infty[} = z'_{[n-k+k',\infty[}$. Thus $z_{[n,\infty[}$ is a common tail of $\mathbf{z}$.
Assume now that there is an $m > n$ such that $\sigma^m(z)$ is left special. Then there is a $z' \in \underline{\mathsf{X}}$ such that $z_{[m,\infty[} = z'_{[m,\infty[}$, but $z_{m-1} \neq z'_{m-1}$. This implies that $z' \in \mathbf{z}$, so if $z_{[n,\infty[}$ is a common tail of $\mathbf{z}$, then there is a $k \in \mathbb{Z}$ such that $z'_{[k,\infty[} = z_{[n,\infty[}$, and since $z_{m-1} \neq z'_{m-1}$, $k \neq n$. But we then have for all $i \geq m$ that

$$z_i = z'_{i+k-n} = z_{i+k-n},$$

which cannot be true, since there are no periodic left special words in $\underline{\mathsf{X}}$. □

The reason for introducing the concept of common tails is illustrated by the following lemma.

LEMMA 5.9. *If x is a common tail of a $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$, then in the notation of Definition 5.4,*

$$\tilde{\alpha}_{[\sigma_+^{n+1}(x)]_\infty} = \alpha_{[\sigma_+^n(x)]_\infty}$$

for every $n \in \mathbb{N}_0$.

Proof: It follows from Lemma 5.8 that $\mathcal{P}_1(\sigma_+^{n+1}(x)) = \{x_n\}$. Thus there is no $z \in \underline{\mathsf{X}}$ such that $z_{[0,\infty[} \notin \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$ and $z_{[1,\infty[} = \sigma_+^{n+1}(x)$, and the only $x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$ such that $\sigma_+(x') = \sigma_+^{n+1}(x)$ is $\sigma_+^n(x)$. Hence $\tilde{\alpha}_{[\sigma_+^{n+1}(x)]_\infty} = \alpha_{[\sigma_+^n(x)]_\infty}$. □

DEFINITION 5.10. *An $x \in \underline{\mathsf{X}}^+$ is called isolated if there is a $k \in \mathbb{N}_0$ such that $[x]_k = \{x\}$.*

LEMMA 5.11. *Every $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ has an isolated common tail.*

Proof: Let z be the cofinal left special element of $\mathbf{j}$. Then $z_{[0,\infty[}$, and thus $z_{[n,\infty[}$ for every $n \in \mathbb{N}_0$, is a common tail by Lemma 5.8. Since there only are finitely many left special words, $[z_{[0,\infty[}]_1$ is finite. Hence there is an $n \in \mathbb{N}$ such that

$$x \in [z_{[0,\infty[}]_1 \wedge x_{[0,n]} = z_{[0,n]} \Rightarrow x = z_{[0,\infty[}.$$

Thus $[z_{[n,\infty[}]_{n+1} = \{z_{[n,\infty[}\}$ and therefore $z_{[n,\infty[}$ is an isolated common tail. □

REMARK 5.12. In [22] Matsumoto introduced the condition (I) for shift spaces, which is a generalization of the condition (I) for topological Markov shifts in the sense of Cuntz and Krieger (cf. [10]).
A shift space $\underline{\mathsf{X}}$ satisfies condition (I) if and only if $\underline{\mathsf{X}}^+$ has no isolated elements (cf. [22, Lemma 5.1]). Thus, it follows from Lemma 5.11 that a shift space which has property (∗∗) does not satisfy condition (I).

Let $\underline{\mathsf{X}}$ be a shift space which has property (∗∗). Choose once and for all, for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ an isolated common tail $x^{\mathbf{j}}$ and a $z^{\mathbf{j}} \in \underline{\mathsf{X}}$ such that $\pi_+(z^{\mathbf{j}}) = x^{\mathbf{j}}$.

Remark 5.13. Notice that $\sigma_+^n(x^{\mathbf{j}})$ is isolated for every $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and every $n \in \mathbb{N}_0$, because if $[x^{\mathbf{j}}]_k = \{x^{\mathbf{j}}\}$, then $[\sigma_+^n(x^{\mathbf{j}})]_{k+n} = \{\sigma_+^n(x^{\mathbf{j}})\}$.

Let z be an adjusted left special element of $\underline{\mathsf{X}}$. Since $x^{\mathbf{z}}$ is a common tail of $\mathbf{z}$, there exists an $n_z \in \mathbb{N}_0$ such that $z_{[n_z,\infty[} = x^{\mathbf{z}}$. We let

$$K_{\underline{\mathsf{X}}} = \{[z_{[n,\infty[}]_\infty \mid z \text{ is an adjusted left special element of } \underline{\mathsf{X}},\ 0 \le n < n_z\},$$

and we let for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$,

$$K_{\mathbf{j}} = \{[z_{[n,\infty[}]_\infty \mid z \in M_{\mathbf{j}},\ 0 \le n \le n_z\}.$$

We notice that

$$K_{\underline{\mathsf{X}}} = \bigcup_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}} \left(K_{\mathbf{j}} \setminus \{x^{\mathbf{j}}\}\right).$$

The following lemma shows that

$$K_{\underline{\mathsf{X}}} \cup \bigcup_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}} \bigcup_{n \in \mathbb{N}_0} \{[\sigma_+^n(x^{\mathbf{j}})]_\infty\}$$

is a partition of $\mathcal{I}_{\underline{\mathsf{X}}}$.

Lemma 5.14.

1. *$K_{\underline{\mathsf{X}}} \cup \{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ n \in \mathbb{N}_0\} = \mathcal{I}_{\underline{\mathsf{X}}}$,*
2. *$K_{\underline{\mathsf{X}}} \cap \{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ n \in \mathbb{N}_0\} = \emptyset$,*
3. *the map $(\mathbf{j}, n) \mapsto [\sigma_+^n(x^{\mathbf{j}})]_\infty$, from $\mathcal{J}_{\underline{\mathsf{X}}} \times \mathbb{N}_0$ to $\mathcal{I}_{\underline{\mathsf{X}}}$ is injective.*

Proof: Let $x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$. Then there is an adjusted left special word z and an $n \in \mathbb{N}_0$ such that $x = z_{[n,\infty[}$. If $n \ge n_z$, then

$$x = z_{[n,\infty[} = z^{\mathbf{z}}_{[n-n_z,\infty[},$$

and if $n < n_z$, then $[x]_\infty = [z_{[n,\infty[}]_\infty \in K_{\underline{\mathsf{X}}}$. Thus

$$K_{\underline{\mathsf{X}}} \cup \{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ n \in \mathbb{N}_0\} = \mathcal{I}_{\underline{\mathsf{X}}}.$$

Assume that $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$, $n \in \mathbb{N}_0$ and $[\sigma_+^n(x^{\mathbf{j}})]_\infty \in K_{\underline{\mathsf{X}}}$. Since $\sigma_+^n(x^{\mathbf{j}})$ is isolated, this implies that there exist an adjusted left special element z and $0 \le m < n_z$ such that $\sigma_+^n(x^{\mathbf{j}}) = z_{[m,\infty[}$. But then

$$z_{[m,\infty[} = \sigma_+^n(x^{\mathbf{j}}) = z_{[n_z+n,\infty[}$$

which cannot be true since there are no periodic left special words in $\underline{\mathsf{X}}$. Thus

$$K_{\underline{\mathsf{X}}} \cap \{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ n \in \mathbb{N}_0\} = \emptyset.$$

Assume that $[\sigma_+^{n_1}(x^{\mathbf{j}_1})]_\infty = [\sigma_+(x^{\mathbf{j}_2})]_\infty$. Since $\sigma_+^{n_1}(x^{\mathbf{j}_1})$ is isolated, $\sigma_+^{n_1}(x^{\mathbf{j}_1})$ must be equal to $\sigma_+^{n_2}(x^{\mathbf{j}_2})$. This implies that $z^{\mathbf{j}_1}$ and $z^{\mathbf{j}_2}$ are right shift tail equivalent, so $\mathbf{j}_1 = \mathbf{j}_2$, and since there are no periodic left special words in $\underline{\mathsf{X}}$, n_1 and n_2 must be equal. $\square$

Remark 5.13 shows that if $[x]_\infty \in \{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ n \in \mathbb{N}_0\}$, then x is isolated. Although it can happen that x is not isolated if $[x]_\infty \in K_{\underline{\mathsf{X}}}$, the following lemma shows that we anyway can separate $K_{\underline{\mathsf{X}}}$ from $\{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ n \in \mathbb{N}_0\}$.

LEMMA 5.15. *There exists an $N_{K_{\underline{\mathsf{X}}}} \in \mathbb{N}_0$ such that if $[x]_\infty \in K_{\underline{\mathsf{X}}}$, then $\#\mathcal{P}_{N_{K_{\underline{\mathsf{X}}}}}(x) > 1$ and*

$$[x]_{N_{K_{\underline{\mathsf{X}}}}} = [x']_{N_{K_{\underline{\mathsf{X}}}}} \Rightarrow [x]_\infty = [x']_\infty$$

for every $x' \in \underline{\mathsf{X}}^+$.

Proof: Since $K_{\underline{\mathsf{X}}}$ is a finite set, it is enough to find for each adjusted left special word $z \in \underline{\mathsf{X}}$ and each $0 \le n < n_z$, an $m \in \mathbb{N}_0$ such that $\#\mathcal{P}_m(z_{[n,\infty[}) > 1$ and $[z_{[n,\infty[}]_m = [x]_m \Rightarrow [z_{[n,\infty[}]_\infty = [x]_\infty$ for every $x \in \underline{\mathsf{X}}^+$.
If z is an adjusted left special element and $0 \le n < n_z$, then $\#\mathcal{P}_{n+1}(z_{[n,\infty[}) > 1$, and since there only is a finite number of left special element in $\underline{\mathsf{X}}$, $[z_{[n,\infty[}]_{n+1}$ is finite, so there exists an $m \in \mathbb{N}_0$ such that $\#\mathcal{P}_m(z_{[n,\infty[}) > 1$ and $[z_{[n,\infty[}]_m = [x]_m \Rightarrow [z_{[n,\infty[}]_\infty = [x]_\infty$ for every $x \in \underline{\mathsf{X}}^+$. $\square$

We have now described the space $\mathcal{I}_{\underline{\mathsf{X}}}$ is such great detail that we are able to rephrase the condition of Lemma 5.3 for when a pair $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})$ belongs to $\mathcal{G}_{\underline{\mathsf{X}}}$ into a condition which is more readily checkable.

LEMMA 5.16. *Let g be a function from $\underline{\mathsf{X}}^-$ to $\mathbb{Z}$ and let for every $\mathsf{i} \in \mathcal{I}_{\underline{\mathsf{X}}}$, α_{i} be an integer. Then $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$ if and only if g is continuous and there exists an $N \in \mathbb{N}_0$ such that $\alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = g(z^{\mathbf{j}}_{]-\infty,n[})$ for all $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and all $n > N$.*

Proof: Assume that $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$. Then there exists by Lemma 5.3 an $N \in \mathbb{N}_0$ such that

1. $\forall y, y' \in \underline{\mathsf{X}}^- : y_{[-N,-1]} = y'_{[-N,-1]} \Rightarrow g(y) = g(y')$,
2. $\forall x, x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) : [x]_N = [x']_N \Rightarrow \alpha_{[x]_\infty} = \alpha_{[x']_\infty}$,
3. $\forall x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+),\ y \in \underline{\mathsf{X}}^- : \mathcal{P}_N(x) = \{y_{[-N,-1]}\} \Rightarrow \alpha_{[x]_\infty} = g(y)$.

It follows from 1. that g is continuous, and since $\mathcal{P}_N(\sigma_+^n(x^{\mathbf{j}})) = \{z^{\mathbf{j}}_{[n-N,n-1]}\}$ for every $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and all $n > N$, it follows from 3. that $\alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = g(z^{\mathbf{j}}_{]-\infty,n[})$.
Assume now that g is continuous and there exists an $N \in \mathbb{N}_0$ such that $\alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = g(z^{\mathbf{j}}_{]-\infty,n[})$ for all $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and all $n > N$. Since g is continuous there is an $M \in \mathbb{N}_0$ such that $y_{[-M,-1]} = y'_{[-M,-1]} \Rightarrow g(y) = g(y')$ for all

$y, y' \in \underline{\mathsf{X}}^-$, and since $\sigma_+^n(x^{\mathbf{j}})$ is isolated for every $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and every $n \in \mathbb{N}_0$ (cf. Remark 5.13), there is for each $0 \leq n \leq \max\{M, N\}$ a $k_n^{\mathbf{j}} \in \mathbb{N}$ such that $[\sigma_+^n(x^{\mathbf{j}})]_{k_n^{\mathbf{j}}} = \{\sigma_+(x^{\mathbf{j}})\}$, and by increasing $k_n^{\mathbf{j}}$ if necessary, we may (and will) assume that $\#\mathcal{P}_{k_n^{\mathbf{j}}}(\sigma_+(x^{\mathbf{j}})) > 1$. Let

$$N' = \max\Big(\{k_n^{\mathbf{j}} \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, 0 \leq n \leq \max\{M, N\}\} \cup \{N_{K_{\underline{\mathsf{X}}}}, M, N\}\Big),$$

where $N_{K_{\underline{\mathsf{X}}}}$ is as in Lemma 5.15. We claim that

1. $\forall y, y' \in \underline{\mathsf{X}}^- : y_{[-N',-1]} = y'_{[-N',-1]} \Rightarrow g(y) = g(y')$,
2. $\forall x, x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) : [x]_{N'} = [x']_{N'} \Rightarrow \alpha_{[x]_\infty} = \alpha_{[x']_\infty}$,
3. $\forall x \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+), y \in \underline{\mathsf{X}}^- : \mathcal{P}_{N'}(x) = \{y_{[-N',-1]}\} \Rightarrow \alpha_{[x]_\infty} = g(y)$,

which implies that $(g, (\alpha_{\mathsf{i}})_{\mathsf{i} \in \mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$. 1. follows from the fact that $N' \geq M$. Notice that if

$$[x]_\infty \in K_{\underline{\mathsf{X}}} \cup \{[\sigma_+^n(x^{\mathbf{j}})]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ 0 \leq n \leq \max\{M, N\}\},$$

then $[x]_{N'} = [x']_{N'} \Rightarrow [x]_\infty = [x']_\infty$. This takes care of 2. in the case where $[x]_\infty \in K_{\underline{\mathsf{X}}} \cup \{[z^{\mathbf{j}}_{[n,\infty[}]_\infty \mid \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}},\ 0 \leq n \leq \max\{M, N\}\}$.
Since

$$\begin{aligned}[\sigma_+^n(x^{\mathbf{j}})]_{N'} = [\sigma_+^{n'}(x^{\mathbf{j}'})]_{N'} &\Rightarrow z^{\mathbf{j}}_{[n-M,n-1]} = z^{\mathbf{j}'}_{[n'-M,n'-1]} \\ &\Rightarrow \alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = g(z^{\mathbf{j}}_{]-\infty,n[}) = g(z^{\mathbf{j}'}_{]-\infty,n'[}) = \alpha_{[\sigma_+^{n'}(x^{\mathbf{j}'})]_\infty},\end{aligned}$$

for $\mathbf{j}, \mathbf{j}' \in \mathcal{J}_{\underline{\mathsf{X}}}$ and $n, n' > \max\{M, N\}$, 2. and 3. hold, and $(g, (\alpha_{\mathsf{i}})_{\mathsf{i} \in \mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$. □

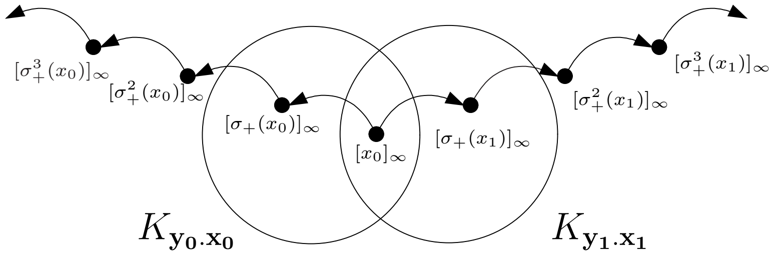
We will now look at $\mathcal{I}_{\underline{\mathsf{X}}}$ for three examples. First let $\underline{\mathsf{X}}$ be the shift space associated with the Morse substitution (see for example [12])

$$0 \mapsto 01, \quad 1 \mapsto 10.$$

The shift space $\underline{\mathsf{X}}$ is minimal and has 4 left special elements:

$$y_0.x_0 \quad y_0.x_1 \quad y_1.x_0 \quad y_1.x_1$$

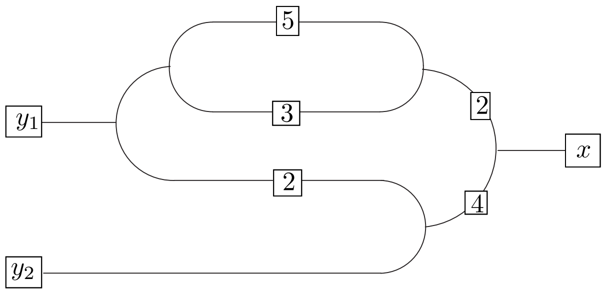
where y_0, y_1 are the fixpoints in $\underline{\mathsf{X}}^-$ of the substitution ending with 0 respectively 1, and x_0, x_1 are the fixpoints in $\underline{\mathsf{X}}^+$ of the substitution beginning with 0 respectively 1. Thus it follows from Example 3.6 that $\underline{\mathsf{X}}$ has property (∗∗). We see that $\mathcal{J}_{\underline{\mathsf{X}}}$ consists of 2 elements: $\mathbf{y_0.x_0}$ and $\mathbf{y_1.x_1}$. Notice that although all of the 4 left special elements are cofinal (and adjusted) neither x_0 nor x_1 are isolated, because $[x_0]_\infty = [x_1]_\infty$, but $\sigma_+(x_0)$ and $\sigma_+(x_1)$ are, so we can choose $\sigma(y_0.x_0)$ and $\sigma(y_1.x_1)$ as $z^{\mathbf{y_0.x_0}}$ and $z^{\mathbf{y_1.x_1}}$ respectively. We then have that $K_{\underline{\mathsf{X}}} = \{[x_0]_\infty\}$, and that the whole of $\mathcal{I}_{\underline{\mathsf{X}}}$ looks like this:



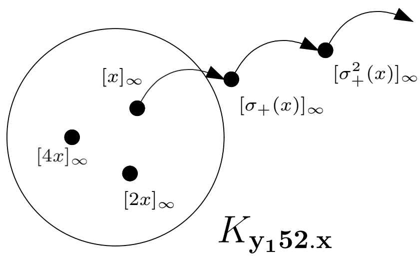
where an arrow from a to b means that in Definition 5.4, $\tilde{\alpha}_b = \alpha_a$. We notice further that $\tilde{\alpha}_{[x_0]_\infty} = g(\sigma_-(y_0)) + g(\sigma_-(y_1))$.
Our second example is the shift space associated to the substitution

$$1 \mapsto 123514, \quad 2 \mapsto 124, \quad 3 \mapsto 13214, \quad 4 \mapsto 14124, \quad 5 \mapsto 15214.$$

The shift space $\underline{\mathsf{X}}$ is minimal and has 8 left special elements (4 adjusted and 4 cofinal) as illustrated on this figure:



where $x \in \underline{\mathsf{X}}^+$ and $y_1, y_2 \in \underline{\mathsf{X}}^-$. Thus it follows from 3.6 that $\underline{\mathsf{X}}$ has property $(**)$. The set $\mathcal{J}_{\underline{\mathsf{X}}}$ consists of one element $\mathbf{y_1 52.x}$, and since x is isolated, we can choose $y_1 52.x$ as $z^{\mathbf{y_1 52.x}}$. We then have that $K_{\underline{\mathsf{X}}} = \{[2x]_\infty, [4x]_\infty\}$, and that the whole of $\mathcal{I}_{\underline{\mathsf{X}}}$ looks like this:

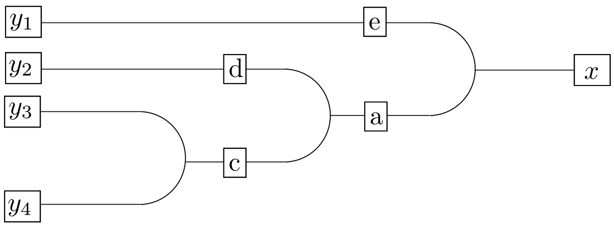


where an arrow from a to b means that in Definition 5.4, $\tilde{\alpha}_b = \alpha_a$. We notice further that $\tilde{\alpha}_{[x]_\infty} = \alpha_{[2x]_\infty} + \alpha_{[4x]_\infty}$, $\tilde{\alpha}_{[2x]_\infty} = 2g(y_1)$ and $\tilde{\alpha}_{[4x]_\infty} = g(y_1) + g(\sigma_-(y_2))$.

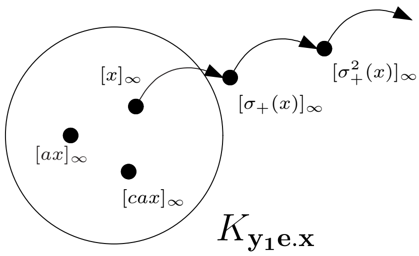
The third example is the shift space associated to the substitution

$$a \mapsto adbac, \quad b \mapsto aedbbc, \quad c \mapsto ac, \quad d \mapsto adac, \quad e \mapsto aecadbac.$$

The shift space $\underline{\mathsf{X}}$ is minimal and has 9 left special elements (1 which is both adjusted and cofinal, 3 which are adjusted but not cofinal, 3 which are cofinal but not adjusted, and 2 which are neither adjusted nor cofinal) as illustrated on this figure:



where $x \in \underline{\mathsf{X}}^+$ and $y_1, y_2, y_3, y_4 \in \underline{\mathsf{X}}^-$. Thus it follows from 3.6 that $\underline{\mathsf{X}}$ has property $(**)$. The set $\mathcal{J}_{\underline{\mathsf{X}}}$ consists of one element $\mathbf{y_1e.x}$, and since x is isolated, we can choose $y_1e.x$ as $z^{\mathbf{y_1e.x}}$. We then have that $K_{\underline{\mathsf{X}}} = \{[cax]_\infty, [ax]_\infty\}$, and that the whole of $\mathcal{I}_{\underline{\mathsf{X}}}$ looks like this:



where an arrow from a to b means that in Definition 5.4, $\tilde{\alpha}_b = \alpha_a$. We notice further that $\tilde{\alpha}_{[x]_\infty} = \alpha_{[ax]_\infty} + g(y_1)$, $\tilde{\alpha}_{[ax]_\infty} = \alpha_{[cax]_\infty} + g(y_2)$ and $\tilde{\alpha}_{[cax]_\infty} = g(\sigma_-(y_3)) + g(\sigma_-(y_4))$.

5.3 $K_0(\underline{\mathsf{X}})$ is a factor of $G_{\underline{\mathsf{X}}}$

We are now ready to define the group $G_{\underline{\mathsf{X}}}$ which has a factor which is isomorphic to $\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}})$.
Loosely speaking, the idea is to simplify $\mathcal{G}_{\underline{\mathsf{X}}}$ in three ways. First we collapse for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$, $K_{\mathbf{j}}$ to one point, which makes it possible to replace $\mathcal{I}_{\underline{\mathsf{X}}}$ by $\mathcal{J}_{\underline{\mathsf{X}}} \times \mathbb{N}_0$, secondly we replace the condition of Lemma 5.16 for when a pair belongs to $\mathcal{G}_{\underline{\mathsf{X}}}$, by the condition that the corresponding sequence in $\mathcal{J}_{\underline{\mathsf{X}}} \times \mathbb{N}_0$ is eventually 0, and thirdly, we replace $\underline{\mathsf{X}}^-$ by $\underline{\mathsf{X}}$. The resulting group $G_{\underline{\mathsf{X}}}$ is of course not necessarily isomorphic to $\mathcal{G}_{\underline{\mathsf{X}}}$, but it turns out that we can still define a map $A_{\underline{\mathsf{X}}} : G_{\underline{\mathsf{X}}} \to G_{\underline{\mathsf{X}}}$ such that $G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\big(G_{\underline{\mathsf{X}}}\big)$ is isomorphic to $\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}})$.

DEFINITION 5.17. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(**)$. Denote by $G_{\underline{\mathsf{X}}}$ the group $C(\underline{\mathsf{X}},\mathbb{Z})\oplus\sum_{n\in\mathbb{N}_0}\mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$, let $A_{\underline{\mathsf{X}}}$ be the map from $G_{\underline{\mathsf{X}}}$ to itself defined by*

$$(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0})\mapsto(f\circ\sigma^{-1},(\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}),$$

where $\tilde{a}_0^{\mathbf{j}}=\sum_{z\in M_{\mathbf{j}}}f(\sigma^{-1}(z))-f(\sigma^{-1}(z^{\mathbf{j}}))$, and $\tilde{a}_n^{\mathbf{j}}=a_{n-1}^{\mathbf{j}}$ for $n>0$, and let ψ be the map from $\mathcal{G}_{\underline{\mathsf{X}}}$ to $G_{\underline{\mathsf{X}}}$ defined by

$$(g,(\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})\mapsto(g\circ\pi_-,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}),$$

where for each $j\in\mathbf{j}$, $a_0^{\mathbf{j}}=\sum_{\mathsf{i}\in K_{\mathbf{j}}}\alpha_{\mathsf{i}}-g(\pi_-(z^{\mathbf{j}}))$ and $a_n^{\mathbf{j}}=\alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty}-g(z^{\mathbf{j}}_{]-\infty,n[})$ for $n>0$.

REMARK 5.18. It directly follows from Lemma 5.16 that ψ in fact maps $\mathcal{G}_{\underline{\mathsf{X}}}$ into $G_{\underline{\mathsf{X}}}$.

PROPOSITION 5.19. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(**)$. Then there is an isomorphism*

$$\bar{\psi}:\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id}-\mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}})\to G_{\underline{\mathsf{X}}}/(\mathrm{Id}-A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)$$

which makes the following diagram commute:

$$\begin{tikzcd}
\mathcal{G}_{\underline{\mathsf{X}}} \arrow[r, "\psi"] \arrow[d, two heads] & G_{\underline{\mathsf{X}}} \arrow[d, two heads] \\
\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id}-\mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[r, "\bar{\psi}"] & G_{\underline{\mathsf{X}}}/(\mathrm{Id}-A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)
\end{tikzcd}$$

We will postpone the proof of proposition 5.19 to section 5.5, and instead state our main theorem which immediately follows from Proposition 5.19 and Corollary 5.6.

THEOREM 5.20. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(**)$. Then $K_0(\underline{\mathsf{X}})$ and*

$$G_{\underline{\mathsf{X}}}/(\mathrm{Id}-A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)$$

are isomorphic as groups.

5.4 EXAMPLES

EXAMPLE 5.21. Let $\underline{\mathsf{X}}$ be a finite shift space. Then $K_0(\underline{\mathsf{X}})$ and

$$C(\underline{\mathsf{X}},\mathbb{Z})/(\mathrm{Id}-(\sigma^{-1})^\star)(C(\underline{\mathsf{X}},\mathbb{Z}))$$

are isomorphic as groups.

Proof: We saw in Example 3.4, that a finite shift space has property $(**)$ and has no left special elements. Thus $\mathcal{J}_{\underline{\mathsf{X}}} = \emptyset$, so $G_{\underline{\mathsf{X}}} = C(\underline{\mathsf{X}}, \mathbb{Z})$ and $A_{\underline{\mathsf{X}}} = (\sigma^{-1})^{\star}$ and it follows from Theorem 5.20, that $K_0(\underline{\mathsf{X}})$ and

$$C(\underline{\mathsf{X}}, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^{\star})(C(\underline{\mathsf{X}}, \mathbb{Z}))$$

are isomorphic as groups. $\square$

Let η be the canonical projection from $G_{\underline{\mathsf{X}}}$ to $C(\underline{\mathsf{X}}, \mathbb{Z})$. We tie things up with the following proposition:

PROPOSITION 5.22. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(**)$. Then there is a surjective group homomorphism*

$$\bar{\eta} : G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right) \to C(\underline{\mathsf{X}}, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^{\star})(C(\underline{\mathsf{X}}, \mathbb{Z}))$$

which makes the following diagram commute:

$$\begin{tikzcd}
C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}) \arrow[rr, "\kappa"] \arrow[dr, "\psi\circ\gamma_{\underline{\mathsf{X}}}"] \arrow[ddd, two heads] & & C(\underline{\mathsf{X}}, \mathbb{Z}) \arrow[ddd, two heads] \\
& G_{\underline{\mathsf{X}}} \arrow[ur, "\eta"] \arrow[d, two heads] & \\
& G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right) \arrow[dr, "\bar{\eta}"] & \\
\frac{C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})}{(\mathrm{Id} - \lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}))} \arrow[ur, "\bar{\psi}\circ\bar{\gamma}_{\underline{\mathsf{X}}}"] \arrow[rr, "\bar{\kappa}"] & & \frac{C(\underline{\mathsf{X}}, \mathbb{Z})}{(\mathrm{Id} - (\sigma^{-1})^{\star})(C(\underline{\mathsf{X}}, \mathbb{Z}))}
\end{tikzcd}$$

where $\bar{\gamma}_{\underline{\mathsf{X}}}$ is the map from $C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z})/(\mathrm{Id} - \lambda_{\underline{\mathsf{X}}})(C(\Omega_{\underline{\mathsf{X}}}, \mathbb{Z}))$ to $\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})\mathcal{G}_{\underline{\mathsf{X}}}$ induced by $\gamma_{\underline{\mathsf{X}}}$.

Proof: Since

$$\eta \circ A_{\underline{\mathsf{X}}} = (\sigma^{-1})^{\star} \circ \eta,$$

η induces a map from $G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)$ to $C(\underline{\mathsf{X}}, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^{\star})(C(\underline{\mathsf{X}}, \mathbb{Z}))$. It is easy to check that this map makes the diagram commute. $\square$

COROLLARY 5.23. *Let $\underline{\mathsf{X}}$ be a shift space which has property $(**)$ and only has two left special words. Then $\bar{\eta}$ is an isomorphism from $G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)$ to $C(\underline{\mathsf{X}}, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^{\star})(C(\underline{\mathsf{X}}, \mathbb{Z}))$. Thus $K_0(\underline{\mathsf{X}})$ and*

$$C(\underline{\mathsf{X}}, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^{\star})(C(\underline{\mathsf{X}}, \mathbb{Z}))$$

are isomorphic as groups.

Proof: If $\underline{\mathsf{X}}$ only has two left special words, z_1 and z_2, then they must necessarily be right shift tail equivalent, so $\mathcal{J}_{\underline{\mathsf{X}}} = \{\mathbf{j}\}$, where $\mathbf{j} = \mathbf{z_1} = \mathbf{z_2}$. We also have that $z_{1[0,\infty[} = z_{2[0,\infty[}$ is an isolated common tail of $\mathbf{j}$, so we can choose z_2 to be $z^{\mathbf{j}}$. The set $M_{\mathbf{j}}$ is equal to $\{z_1, z_2\}$, so for any $(h, (b^{\mathbf{j}}_n)_{n\in\mathbb{N}_0}) \in G_{\underline{\mathsf{X}}}$ is

$$A_{\underline{\mathsf{X}}}((h, (b^{\mathbf{j}}_n)_{n\in\mathbb{N}_0})) = (h \circ \sigma^{-1}, (\tilde{b}^{\mathbf{j}}_n)_{n\in\mathbb{N}_0}),$$

where $\tilde{b}^{\mathbf{j}}_0 = h(\sigma^{-1}(z_1))$, and $\tilde{b}^{\mathbf{j}}_n = b^{\mathbf{j}}_{n-1}$ for $n > 0$.
Suppose that $(f, (a^{\mathbf{j}}_n)_{n\in\mathbb{N}_0}) \in G_{\underline{\mathsf{X}}}$ and that

$$\eta((f, (a^{\mathbf{j}}_n)_{n\in\mathbb{N}_0})) \in (\mathrm{Id} - (\sigma^{-1})^\star)(C(\underline{\mathsf{X}}, \mathbb{Z})).$$

Then there is a $\tilde{f} \in C(\underline{\mathsf{X}}, \mathbb{Z})$ such that $f = \tilde{f} - \tilde{f} \circ \sigma^{-1}$. Since $(a^{\mathbf{j}}_n)_{n\in\mathbb{N}_0} \in \sum_{n\in\mathbb{N}_0} \mathbb{Z}$, there is an $N \in \mathbb{N}_0$ such that $a^{\mathbf{j}}_n = 0$ for $n > N$. Let

$$c = -\tilde{f}(\sigma^{-1}(z_1)) - \sum_{n=0}^{N} a^{\mathbf{j}}_n$$

and $h \in C(\underline{\mathsf{X}}, \mathbb{Z})$ the function $\tilde{f}$ plus the constant c, and let $b^{\mathbf{j}}_n = \sum_{i=0}^n a^{\mathbf{j}}_i + h(\sigma^{-1}(z_1))$ for $n \in \mathbb{N}_0$. Then $b^{\mathbf{j}}_n = 0$ for $n > N$, so $(h, (b^{\mathbf{j}}_n)_{n\in\mathbb{N}_0}) \in G_{\underline{\mathsf{X}}}$, and

$$(f, (a^{\mathbf{j}}_n)_{n\in\mathbb{N}_0}) = (\mathrm{Id} - A_{\underline{\mathsf{X}}})((h, (b^{\mathbf{j}}_n)_{n\in\mathbb{N}_0})) \in (\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right),$$

which prove that $\bar{\eta}$ is injective and thus an isomorphism. □

Example 5.24. As noted in [12], a Sturmian shift space $\underline{\mathsf{X}}_\alpha$, $\alpha \in [0,1]\backslash\mathbb{Q}$ is minimal and has two special words. Thus it follows from Example 3.6 and Corollary 5.23 that $K_0(\underline{\mathsf{X}}_\alpha)$ and

$$C(\underline{\mathsf{X}}_\alpha, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^\star)(C(\underline{\mathsf{X}}_\alpha, \mathbb{Z}))$$

are isomorphic as groups.
In [31] it is shown that

$$C(\underline{\mathsf{X}}_\alpha, \mathbb{Z})/(\mathrm{Id} - (\sigma^{-1})^\star)(C(\underline{\mathsf{X}}_\alpha, \mathbb{Z}))$$

is isomorphic to $\mathbb{Z} + \mathbb{Z}\alpha$ as an *ordered* group. Thus it follows that $K_0(\underline{\mathsf{X}}_\alpha)$ and $\mathbb{Z} + \mathbb{Z}\alpha$ are isomorphic as groups.
In [9, Corollary 5.2] we prove that $K_0(\underline{\mathsf{X}}_\alpha)$ with the order structure mentioned in the Introduction is isomorphic to $\mathbb{Z} + \mathbb{Z}\alpha$.

Example 5.25. It is proved in [30, pp. 90 and 107] that if τ is an aperiodic and primitive substitution, then the associated shift space X_τ is minimal and only has a finite number of left special words. Thus by Example 3.6, X_τ has property $(**)$. It follows from [6, Proposition 3.5] that if τ furthermore is proper and elementary, then $\pi_+(z)$ is isolated for every left special word z. Thus $K_0(\mathsf{X}_\tau)$ is isomorphic to the cokernel of the map

$$A_\tau(f,[(a_0^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\mathsf{X}_\tau}},(a_1^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\mathsf{X}_\tau}},\dots]) =$$
$$\left(f\circ\sigma^{-1},\left[\left(\left(\sum_{z\in M_{\mathbf{j}}} f(\sigma^{-1}(z))\right)-f(\sigma^{-1}(z^{\mathbf{j}}))\right)_{\mathbf{j}\in\mathcal{J}_{\mathsf{X}_\tau}},(a_0^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\mathsf{X}_\tau}},(a_1^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\mathsf{X}_\tau}},\dots\right]\right)$$

defined on

$$G_\tau = C(\underline{\mathsf{X}}_\tau,\mathbb{Z})\oplus\sum_{i=0}^{\infty}\mathbb{Z}^{\mathcal{J}_{\mathsf{X}_\tau}},$$

where $\mathcal{J}_{\mathsf{X}_\tau}$ and $M_{\mathbf{j}}$ are as defined in section 5.2, and $z^{\mathbf{j}}$ is a cofinal special element belonging to the right shift tail equivalence class $\mathbf{j}$.
In the notation of [8],

$$\mathcal{J}_{\mathsf{X}_\tau}=\{\widetilde{\mathsf{y}}^1,\widetilde{\mathsf{y}}^2,\dots,\widetilde{\mathsf{y}}^{\mathsf{n}_\tau}\},\ M_{\widetilde{\mathsf{y}}^{\mathsf{j}}}=\{\mathsf{y}_1^j,\mathsf{y}_2^j,\dots,\mathsf{y}_{p_j+1}^j\}\ \text{ and } z^{\widetilde{\mathsf{y}}^{\mathsf{j}}}=\widetilde{\mathsf{y}}^j.$$

In [8], this is used for every aperiodic and primitive (but not necessarily proper or elementary) substitution τ, to present $K_0(\mathsf{X}_\tau)$ as a stationary inductive limit of a system associated to an integer matrix defined from combinatorial data which can be computed in an algorithmic way (cf. [6] and [7]).

5.5 The proof of Proposition 5.19

In order to prove Proposition 5.19, we will define maps and groups as indicated on the diagram:

$$\begin{tikzcd}
\mathcal{G}_{\underline{\mathsf{X}}} \arrow[r, two heads] \arrow[d, "\psi_1"] & \mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id}-\mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[d, "\overline{\psi}_1"] \\
\psi_1(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[r, two heads] \arrow[d, "\psi_2"] & \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id}-A_1)(\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})) \arrow[d, "\overline{\psi}_2"] \\
C(\underline{\mathsf{X}}^-,\mathbb{Z})\oplus\sum_{n\in\mathbb{N}_0}\mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \arrow[r, two heads] \arrow[d, "\psi_3"] & \frac{C(\underline{\mathsf{X}}^-,\mathbb{Z})\oplus\sum_{n\in\mathbb{N}_0}\mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}}{(\mathrm{Id}-A_2)\left(C(\underline{\mathsf{X}}^-,\mathbb{Z})\oplus\sum_{n\in\mathbb{N}_0}\mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}\right)} \arrow[d, "\overline{\psi}_3"] \\
G_{\underline{\mathsf{X}}} \arrow[r, two heads] & G_{\underline{\mathsf{X}}}/(\mathrm{Id}-A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)
\end{tikzcd}$$

such that the diagram commutes, $\psi_3 \circ \psi_2 \circ \psi_1 = \psi$, and $\overline{\psi}_1$, $\overline{\psi}_2$ and $\overline{\psi}_3$ are isomorphisms.
Let $\psi_1 : \mathcal{G}_{\underline{\mathsf{X}}} \to C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \prod_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$ be the map defined by

$$(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \mapsto (g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}),$$

where for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$, $a_0^{\mathbf{j}} = \sum_{\mathsf{i}\in K_{\mathbf{j}}} \alpha_{\mathsf{i}}$, and $a_n^{\mathbf{j}} = \alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty}$ for $n \in \mathbb{N}$.

LEMMA 5.26. *Let* $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \prod_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$. *Then* $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ *if and only if*

$$\exists N \in \mathbb{N}_0 \forall \mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}} \forall n > N : a_n^{\mathbf{j}} = g(z^{\mathbf{j}}_{]-\infty,n[}).$$

Proof: The forward implication directly follows from Lemma 5.16.
Assume that $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \prod_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$ and there exists an $N \in \mathbb{N}_0$ such that $a_n^{\mathbf{j}} = g(z^{\mathbf{j}}_{]-\infty,n[})$ for all $n > N$. We let $\alpha_{\mathsf{i}} = 0$ for each $\mathsf{i} \in K_{\underline{\mathsf{X}}}$, and we let for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and each $n \in \mathbb{N}_0$, $\alpha_{[z^{\mathbf{j}}_{[n,\infty[}]_\infty} = a_n^{\mathbf{j}}$. It then follows from Lemma 5.16 that $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$, and since $\psi_1(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) = (g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0})$, we have that $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$.

□

Let $A_1 : C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \prod_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \to C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \prod_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$ be the map defined by

$$(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \mapsto (g \circ \sigma_-, (\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}),$$

where for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$, $\tilde{a}_0^{\mathbf{j}} = \sum_{z\in M_{\mathbf{j}}} g(\sigma_-(\pi_-(z)))$, and $\tilde{a}_n^{\mathbf{j}} = a_{n-1}^{\mathbf{j}}$ for $n \in \mathbb{N}$.
It follows from Lemma 5.26 that A_1 maps $\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ into itself. Thus $(\mathrm{Id} - A_1)\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ is a subgroup of $\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$, and we can form the quotient $\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id} - A_1)\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$. Let

$$q : \psi_1(\mathcal{G}_{\underline{\mathsf{X}}}) \mapsto \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id} - A_1)\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$$

be the quotient map. We then have:

LEMMA 5.27. $\ker(q \circ \psi_1) = (\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}})$.

Proof: Assume $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \ker(q \circ \psi_1)$. That means that

$$(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) = \psi_1(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in (\mathrm{Id} - A_1)\psi_1(\mathcal{G}_{\underline{\mathsf{X}}}).$$

Thus there exists $(\tilde{g}, (\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ such that $g = \tilde{g} - \tilde{g} \circ \sigma_-$, and such that for every $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$,

$$a_0^{\mathbf{j}} = \sum_{\mathsf{i}\in K_{\mathbf{j}}} \alpha_{\mathsf{i}} = \tilde{a}_0^{\mathbf{j}} - \sum_{z\in M_{\mathbf{j}}} \tilde{g}(\sigma_-(\pi_-(z))),$$

and $a_n^{\mathbf{j}} = \alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = \tilde{a}_n^{\mathbf{j}} - \tilde{a}_{n-1}^{\mathbf{j}}$ for all $n \in \mathbb{N}$.

Now let $\mathsf{i} \in \bigcup_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}} K_{\mathbf{j}}$. Choose $x_{\mathsf{i}} \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+)$ such that $[x_{\mathsf{i}}]_\infty = \mathsf{i}$. Then there is, for each $z \in \underline{\mathsf{X}}$ which satisfies $\pi_+(z) = x_{\mathsf{i}}$, a unique $m_z \in \mathbb{N}_0$ such that $\sigma^{-m_z}(z)$ is an adjusted left special word. We let

$$L_{\mathsf{i}} = \{[z_{[-m,\infty[}]_\infty \mid \pi_+(z) = x_{\mathsf{i}},\ 0 \le m \le m_z\} \subseteq \mathcal{I}_{\underline{\mathsf{X}}},$$
$$B_{\mathsf{i}} = \{\sigma^{-m_z}(z) \mid \pi_+(z) = x_{\mathsf{i}}\} \subseteq \underline{\mathsf{X}},$$

and

$$\tilde{\alpha}_{\mathsf{i}} = \sum_{\mathsf{i}'\in L_{\mathsf{i}}} \alpha_{\mathsf{i}'} + \sum_{z\in B_{\mathsf{i}}} \tilde{g}(\sigma_-(\pi_-(z))),$$

and we let for $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and $n \in \mathbb{N}$, $\tilde{\alpha}_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = \tilde{a}_n^{\mathbf{j}}$.

Since $(\tilde{g}, (\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$, $\tilde{g}$ is continuous and there exists by Lemma 5.26, an $N \in \mathbb{N}_0$ such that for all $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and all $n > N$, $\tilde{\alpha}_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = \tilde{a}_n^{\mathbf{j}} = \tilde{g}(z^{\mathbf{j}}_{]-\infty,n[})$, so $(\tilde{g}, (\tilde{\alpha}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$ by Lemma 5.16.

Let $(\tilde{\tilde{g}}, (\tilde{\tilde{\alpha}}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) = \mathcal{A}_{\underline{\mathsf{X}}}(\tilde{g}, (\tilde{\alpha}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})$. Then $\tilde{\tilde{g}} = \tilde{g} \circ \sigma_-$, and by lemma 5.9,

$$\tilde{\tilde{\alpha}}_{[\sigma_+^{n+1}(x^{\mathbf{j}})]_\infty} = \tilde{\alpha}_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = \tilde{a}_n^{\mathbf{j}},$$

for $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and $n \in \mathbb{N}$.

Now let $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$. Then $L_{[x^{\mathbf{j}}]_\infty} = K_{\mathbf{j}}$ and $B_{[x^{\mathbf{j}}]_\infty} = M_{\mathbf{j}}$, so

$$\begin{aligned}
\tilde{\tilde{\alpha}}_{[\sigma_+(x^{\mathbf{j}})]_\infty} &= \tilde{\alpha}_{[x^{\mathbf{j}}]_\infty} \\
&= \sum_{\mathsf{i}\in K_{\mathbf{j}}} \alpha_{\mathsf{i}} + \sum_{z\in M_{\mathbf{j}}} \tilde{g}(\sigma_-(\pi_-(z))) \\
&= a_0^{\mathbf{j}} + \sum_{z\in M_{\mathbf{j}}} \tilde{g}(\sigma_-(\pi_-(z))) \\
&= \tilde{a}_0^{\mathbf{j}}.
\end{aligned}$$

If $[x]_\infty \in K_{\mathbf{j}}$, then $L_{[x]_\infty}$ is the disjoint union of $L_{[x']_\infty}$, where $[x']_\infty \in \mathcal{I}_{\underline{\mathsf{X}}}$ and $\sigma_+(x') = x$, and $\{[x]_\infty\}$, and $B_{[x]_\infty}$ is the disjoint union of $B_{[x']_\infty}$, where $[x']_\infty \in \mathcal{I}_{\underline{\mathsf{X}}}$ and $\sigma_+(x') = x$, and $\{\sigma(z) \mid z \in \underline{\mathsf{X}}, z_{[0,\infty[} \notin \mathcal{ND}_\infty(\underline{\mathsf{X}}^+), z_{[1,\infty[} = x\}$.

Hence

$$\begin{aligned}
\tilde{\tilde{\alpha}}_{[x]_\infty} &= \sum_{\substack{[x']_\infty\in\mathcal{I}_{\underline{\mathsf{X}}}\\ \sigma_+(x')=x}} \tilde{\alpha}_{[x']_\infty} + \sum_{\substack{z\in\underline{\mathsf{X}}\\ z_{[0,\infty[}\notin\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)\\ z_{[1,\infty[}=x}} \tilde{g}(z_{]-\infty,-1]})\\
&= \sum_{\substack{[x']_\infty\in\mathcal{I}_{\underline{\mathsf{X}}}\\ \sigma_+(x')=x}} \left(\sum_{\mathsf{i}\in L_{[x']_\infty}} \alpha_{\mathsf{i}} + \sum_{z\in B_{[x']_\infty}} \tilde{g}(\sigma_-(\pi_-(z)))\right) + \sum_{\substack{z\in\underline{\mathsf{X}}\\ z_{[0,\infty[}\notin\mathcal{ND}_\infty(\underline{\mathsf{X}}^+)\\ z_{[1,\infty[}=x}} \tilde{g}(z_{]-\infty,-1]})\\
&= \sum_{\mathsf{i}\in L_{[x]_\infty}} \alpha_{\mathsf{i}} - \alpha_{[x]_\infty} + \sum_{z\in B_{[x]_\infty}} \tilde{g}(\sigma_-(\pi_-(z)))\\
&= \tilde{\alpha}_{[x]_\infty} - \alpha_{[x]_\infty}.
\end{aligned}$$

So

$$\tilde{g} - \tilde{\tilde{g}} = \tilde{g} - \tilde{g}\circ\sigma_- = g,$$

and for $\mathsf{i} \in \bigcup_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}} K_{\mathbf{j}}$,

$$\tilde{\alpha}_{\mathsf{i}} - \tilde{\tilde{\alpha}}_{\mathsf{i}} = \tilde{\alpha}_{\mathsf{i}} - \tilde{\alpha}_{\mathsf{i}} + \alpha_{\mathsf{i}} = \alpha_{\mathsf{i}},$$

and for $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and $n \in \mathbb{N}$

$$\tilde{\alpha}_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} - \tilde{\tilde{\alpha}}_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} = \tilde{a}_n^{\mathbf{j}} - \tilde{a}_{n-1}^{\mathbf{j}} = a_n^{\mathbf{j}} = \alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty}.$$

Thus

$$\begin{aligned}
&(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) = (\tilde{g}, (\tilde{\alpha}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) - (\tilde{\tilde{g}}, (\tilde{\tilde{\alpha}}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})\\
&= (\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\tilde{g}, (\tilde{\alpha}_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in (\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}}),
\end{aligned}$$

which shows that $\ker(q\circ\psi_1) \subseteq (\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}})$.
Now let $(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}}) \in \mathcal{G}_{\underline{\mathsf{X}}}$. We will find an element $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ such that

$$\psi_1((\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})) = (\mathrm{Id} - A_1)(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}).$$

This will show that $(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}}) \subseteq \ker(\rho\circ\psi_1)$.
Let for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and every $n \in \mathbb{N}_0$, $a_n^{\mathbf{j}} = \alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty}$. It then follows from Lemma 5.16 and 5.26 that $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$.
Now,

$$(\mathrm{Id} - A_1)(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) = (g - g\circ\sigma_-, (a_n^{\mathbf{j}} - \tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}),$$

where for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$, $\tilde{a}_0^{\mathbf{j}} = \sum_{z\in M_{\mathbf{j}}} g(\sigma_-(\pi_-(z)))$, and $\tilde{a}_n^{\mathbf{j}} = a_{n-1}^{\mathbf{j}}$ for $n \in \mathbb{N}$, and

$$\psi_1((\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(g, (\alpha_{\mathsf{i}})_{\mathsf{i}\in\mathcal{I}_{\underline{\mathsf{X}}}})) = (g - g\circ\sigma_-, (b_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}),$$

where by Lemma 5.9

$$b^{\mathbf{j}}_n = \alpha_{[\sigma_+^n(x^{\mathbf{j}})]_\infty} - \alpha_{[\sigma_+^{n-1}(x^{\mathbf{j}})]_\infty} = a^{\mathbf{j}}_n - a^{\mathbf{j}}_{n-1} = a^{\mathbf{j}}_n - \tilde{a}^{\mathbf{j}}_n$$

for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$ and every $n \in \mathbb{N}$, and

$$\begin{aligned} b^{\mathbf{j}}_0 = & \sum_{\mathsf{i} \in K_{\mathbf{j}}} \alpha_{\mathsf{i}} - \sum_{[x]_\infty \in K_{\mathbf{j}}} \left(\sum_{\substack{x' \in \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) \\ \sigma_+(x') = x}} \alpha_{[x']_\infty} + \sum_{\substack{z \in \underline{\mathsf{X}} \\ z_{[0,\infty[} \notin \mathcal{ND}_\infty(\underline{\mathsf{X}}^+) \\ z_{[1,\infty[} = x}} g(\pi_-(z)) \right) \\ = & \alpha_{x^{\mathbf{j}}} - \sum_{z \in M_{\mathbf{j}}} g(\sigma_-(\pi_-(z))) \\ = & a^{\mathbf{j}}_0 - \tilde{a}^{\mathbf{j}}_0 \end{aligned}$$

for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$. Thus $\psi_1((\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(g, (\alpha_{\mathsf{i}})_{\mathsf{i} \in \mathcal{I}_{\underline{\mathsf{X}}}})) = (\mathrm{Id} - A_1)(g, (a^{\mathbf{j}}_n)_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, n \in \mathbb{N}_0})$. □

It follows from the previous lemma, that $\psi_1 : \mathcal{G}_{\underline{\mathsf{X}}} \to \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ induces an isomorphism $\overline{\psi}_1$ from $\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}})$ to $\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id} - A_1)(\psi_1(\mathcal{G}_{\underline{\mathsf{X}}}))$ such that the diagram

$$\begin{tikzcd} \mathcal{G}_{\underline{\mathsf{X}}} \arrow[r, two heads] \arrow[d, "\psi_1"] & \mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[d, "\overline{\psi}_1"] \\ \psi_1(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[r, two heads] & \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id} - A_1)(\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})) \end{tikzcd}$$

commutes.
Let for every $(g, (a^{\mathbf{j}}_n)_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, n \in \mathbb{N}_0}) \in \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$,

$$\psi_2(g, (a^{\mathbf{j}}_n)_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, n \in \mathbb{N}_0}) = (g, (a^{\mathbf{j}}_n - g(z^{\mathbf{j}}_{]-\infty, n[}))_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, n \in \mathbb{N}_0}).$$

It follows from Lemma 5.26 that ψ_2 is an isomorphism from $\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})$ to $C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n \in \mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$.
Let $A_2 : C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n \in \mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \to C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n \in \mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$ be the map given by

$$(g, (a^{\mathbf{j}}_n)_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, n \in \mathbb{N}_0}) \mapsto (g \circ \sigma_-, (\hat{a}^{\mathbf{j}}_n)_{\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}, n \in \mathbb{N}_0}),$$

where for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$,

$$\hat{a}^{\mathbf{j}}_0 = \sum_{z \in M_{\mathbf{j}}} g(\pi_-(\sigma^{-1}(z))) - g(\pi_-(\sigma^{-1}(z^{\mathbf{j}}))),$$

and $\hat{a}^{\mathbf{j}}_n = a^{\mathbf{j}}_{n-1}$ for $n \in \mathbb{N}$.

Then $\psi_2 \circ A_1 = A_2 \circ \psi_2$, so ψ_2 induces an isomorphism

$$\overline{\psi}_2 : \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id} - A_1)(\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})) \to \frac{C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}}{(\mathrm{Id} - A_2)\left(C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}\right)}$$

such that the diagram

$$\begin{tikzcd}
\psi_1(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[r, two heads] \arrow[d, "\psi_2"] & \psi_1(\mathcal{G}_{\underline{\mathsf{X}}})/(\mathrm{Id} - A_1)(\psi_1(\mathcal{G}_{\underline{\mathsf{X}}})) \arrow[d, "\overline{\psi}_2"] \\
C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \arrow[r, two heads] & \frac{C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}}{(\mathrm{Id} - A_2)\left(C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}\right)}
\end{tikzcd}$$

commutes.
Let $\psi_3 : C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \to G_{\underline{\mathsf{X}}}$ be the map defined by

$$(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \mapsto (g \circ \pi_-, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}).$$

We then have:

Lemma 5.28. *$\psi_3 \circ A_2 = A \circ \psi_3$.*

Proof: Let $(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \in C(\underline{\mathsf{X}}^-, \mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$. Then

$$\begin{aligned} A \circ \psi_3(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) &= A(g \circ \pi_-, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \\ &= (g \circ \pi_- \circ \sigma^{-1}, (\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}), \end{aligned}$$

where for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$,

$$\tilde{a}_0^{\mathbf{j}} = \sum_{z\in M_{\mathbf{j}}} g \circ \pi_-(\sigma^{-1}(z)) - g \circ \pi_-(\sigma^{-1}(z^{\mathbf{j}})),$$

and $\tilde{a}_n^{\mathbf{j}} = a_{n-1}^{\mathbf{j}}$ for $n \in \mathbb{N}$, and

$$\begin{aligned} \psi_3 \circ A_2(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) &= \psi_3(g \circ \sigma_-, (\hat{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) \\ &= (g \circ \sigma_- \circ \pi_-, (\hat{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}), \end{aligned}$$

where for each $\mathbf{j} \in \mathcal{J}_{\underline{\mathsf{X}}}$,

$$\hat{a}_0^{\mathbf{j}} = \sum_{z\in M_{\mathbf{j}}} g(\pi_-(\sigma^{-1}(z))) - g(\pi_-(\sigma^{-1}(z^{\mathbf{j}}))) = \tilde{a}_0^{\mathbf{j}},$$

and $\hat{a}_n^{\mathbf{j}} = a_{n-1}^{\mathbf{j}} = \tilde{a}_n^{\mathbf{j}}$ for $n \in \mathbb{N}$, and since $\pi_- \circ \sigma^{-1} = \sigma_- \circ \pi_-$, we have that $\psi_3 \circ A_2(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0}) = A \circ \psi_3(g, (a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}, n\in\mathbb{N}_0})$. □

It follows from the previous lemma that $\psi_3 : C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \to G_{\underline{\mathsf{X}}}$ induces an injective map

$$\overline{\psi}_3 : \frac{C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}}{(\mathrm{Id} - A_2)\left(C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}\right)} \to G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)$$

such that the diagram

$$\begin{tikzcd}
C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}} \arrow[r, two heads] \arrow[d, "\psi_3"] & \frac{C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}}{(\mathrm{Id} - A_2)\left(C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}\right)} \arrow[d, "\overline{\psi}_3"] \\
G_{\underline{\mathsf{X}}} \arrow[r, two heads] & G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)
\end{tikzcd}$$

commutes. We will now show that $\overline{\psi}_3$ in fact is an isomorphism.

Lemma 5.29. *The map $\overline{\psi}_3$ is surjective.*

Proof: In order to prove that $\overline{\psi}_3$ is surjective, it is enough to show that for every $(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) \in G_{\underline{\mathsf{X}}}$, there is a $(g,(\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) \in C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$ such that

$$(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) - \psi_3(g,(\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) \in (\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right).$$

Since f is continuous, there are $k,m\in\mathbb{N}$ such that $z_{[-k,m]} = z'_{[-k,m]} \Rightarrow f(z) = f(z')$. Thus

$$z_{[-k-m-1,-1]} = z'_{[-k-m-1,-1]} \Rightarrow f\circ\sigma^{-(m+1)}(z) = f\circ\sigma^{-(m+1)}(z').$$

Hence there is an $g\in C(\underline{\mathsf{X}}^-,\mathbb{Z})$ such that $g\circ\pi_- = f\circ\sigma^{-(m+1)}$. We let for each $\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}}$,

$$\tilde{a}_n^{\mathbf{j}} = \sum_{z\in M_{\mathbf{j}}} f\circ\sigma^{n-m-1}(z) - f\circ\sigma^{n-m-1}(z^{\mathbf{j}})$$

for $0\le n\le m$, and $\tilde{a}_n^{\mathbf{j}} = a_{n-(m+1)}^{\mathbf{j}}$ for $n>m$. Since $(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) \in G_{\underline{\mathsf{X}}}$, $(g,(\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) \in C(\underline{\mathsf{X}}^-,\mathbb{Z}) \oplus \sum_{n\in\mathbb{N}_0} \mathbb{Z}^{\mathcal{J}_{\underline{\mathsf{X}}}}$, and it is easy to check that

$$\psi_3(g,(\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) = A_{\underline{\mathsf{X}}}^{m+1}(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}).$$

Thus

$$\begin{aligned}
(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) - \psi_3(g,(\tilde{a}_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) &= \\
(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) - A_{\underline{\mathsf{X}}}^{m+1}(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0}) &= \\
\sum_{k=0}^{m}(\mathrm{Id} - A_{\underline{\mathsf{X}}})(A_{\underline{\mathsf{X}}}^k(f,(a_n^{\mathbf{j}})_{\mathbf{j}\in\mathcal{J}_{\underline{\mathsf{X}}},n\in\mathbb{N}_0})) &\in (\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right).
\end{aligned}$$

□

We now have that $\bar{\psi} = \overline{\psi}_3 \circ \overline{\psi}_2 \circ \overline{\psi}_1$ is an isomorphism and since $\psi = \psi_3 \circ \psi_2 \circ \psi_1$, the diagram

$$\begin{tikzcd}
\mathcal{G}_{\underline{\mathsf{X}}} \arrow[r, "\psi"] \arrow[d] & G_{\underline{\mathsf{X}}} \arrow[d] \\
\mathcal{G}_{\underline{\mathsf{X}}}/(\mathrm{Id} - \mathcal{A}_{\underline{\mathsf{X}}})(\mathcal{G}_{\underline{\mathsf{X}}}) \arrow[r, "\bar{\psi}"] & G_{\underline{\mathsf{X}}}/(\mathrm{Id} - A_{\underline{\mathsf{X}}})\left(G_{\underline{\mathsf{X}}}\right)
\end{tikzcd}$$

commutes, and we are done with the proof of Proposition 5.19.

References

[1] R. Bowen and J. Franks, *Homology for zero-dimensional nonwandering sets*, Ann. Math. (2) 106 (1977), no. 1, 73–92.

[2] M. Boyle and D. Handelman, *Orbit equivalence, flow equivalence and ordered cohomology*, Israel J. Math. 95 (1996), 169–210.

[3] M. Boyle and D. Lind, *Exponential subdynamics*, Trans. Amer. Math. Soc. 349 (1997), 55–102.

[4] M. Brin and G. Stuck, *Introduction to dynamical systems*, Cambridge University Press, Cambridge, 2002.

[5] T.M. Carlsen, *C^*-algebras associated to general shift spaces*, Master's thesis, Copenhagen University, 2001, available on `www.math.ku.dk/~toke`.

[6] T.M. Carlsen and S. Eilers, *A graph approach to computing nondeterminacy in substitutional dynamical systems*, submitted for publication. `www.math.ku.dk/~eilers/papers/cei`, 2002.

[7] ______, *Java applet*, `www.math.ku.dk/~eilers/papers/cei`, 2002.

[8] ______, *Augmenting dimension group invariants for substitution dynamics*, Ergodic Theory Dynam. Systems 24 (2004), no. 4, 1015-1039.

[9] ______, *Ordered K-groups associated to substitutional dynamics*, submitted for publication. `www.math.ku.dk/~eilers/papers/ceiv`, 2004.

[10] J. Cuntz and W. Krieger, *A class of C^*-algebras and topological markov chains*, Invent. Math. 56 (1980), 251–268.

[11] F. Durand, B. Host and C. Skau, *Substitutional dynamical systems, Bratteli diagrams and dimension groups*, Ergodic Theory Dynam. Systems 19 (1999), no. 4, 953–993.

[12] N.P. Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lecture Notes in Mathematics 1794, Springer-Verlag, Berlin, 2002.

[13] A.H. Forrest, *K-groups associated with substitution minimal systems*, Israel J. Math. 98 (1997), 101–139.

[14] J. Franks, *Flow equivalence of subshifts of finite type*, Ergodic Theory Dynam. Systems 4 (1984), no. 1, 53–66.

[15] T. Giordano, I.F. Putnam, and C.F. Skau, *Topological orbit equivalence and C^*-crossed products*, J. Reine Angew. Math. 469 (1995), 51–111.

[16] R.H. Herman, I.F. Putnam, and C.F. Skau, *Ordered Bratteli diagrams, dimension groups and topological dynamics*, Internat. J. Math. 3 (1992), no. 6, 827–864.

[17] C. Holton and L.Q. Zamboni, *Directed graphs and substitutions*, Theory Comput. Syst. 34 (2001), no. 6, 545–564.

[18] B.P. Kitchens, *Symbolic dynamics*, Springer-Verlag, Berlin, 1998.

[19] W. Krieger, *On dimension functions and topological Markov chains*, Invent. Math. 56 (1980), no. 3, 239–250.

[20] ———, *On a dimension for a class of homeomorphism groups*, Math. Ann. 252 (1980), no. 2, 87–95.

[21] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995.

[22] K. Matsumoto, *K-theory for C^*-algebras associated with subshifts*, Math. Scand. 82 (1998), no. 2, 237–255.

[23] ———, *Dimension groups for subshifts and simplicity of the associated C^*-algebras*, J. Math. Soc. Japan 51 (1999), no. 3, 679–698.

[24] ———, *Presentations of subshifts and their topological conjugacy invariants*, Doc. Math. 4 (1999), 285–340 (electronic).

[25] ———, *Bowen-Franks groups as an invariant for flow equivalence of subshifts*, Ergodic Theory Dynam. Systems 21 (2001), no. 6, 1831–1842.

[26] B. Parry and D. Sullivan, *A topological invariant of flows on 1-dimensional spaces*, Topology 14 (1975), no. 4, 297–299.

[27] W. Parry and S. Tuncel, *Classification problems in ergodic theory*, London Mathematical Society Lecture Note Series 67 Cambridge University Press, Cambridge, 1982.

[28] Y.T. Poon, *A K-theoretic invariant for dynamical systems*, Trans. Amer. Math. Soc., 311 (1989), no. 2, 515–533.

[29] I.F. Putnam, *On the topological stable rank of certain transformation group C^*-algebras*, Ergodic Theory Dynam. Systems 10 (1990), 197–207.

[30] M. Queffélec, *Substitution dynamical systems—spectral analysis*, Springer-Verlag, Berlin, 1987.

[31] C. Skau, *Ordered K-theory and minimal symbolic dynamical systems*, Colloq. Math., 84/85 (2000), part 1, 203–227.

[32] S. Williams, *Toeplitz minimal flows which are not uniquely ergodic*, Z. Wahrsch. Verw. Gebiete, 67 (1984), no. 1, 95–107.

Toke Meier Carlsen
Department of Mathematics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark
toke@math.ku.dk

Søren Eilers
Department of Mathematics
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark
eilers@math.ku.dk