MAXIMAL MONOTONICITY OF DENSE TYPE, LOCAL MAXIMAL MONOTONICITY, AND MONOTONICITY OF THE CONJUGATE ARE ALL THE SAME FOR CONTINUOUS LINEAR OPERATORS

Heinz H. Bauschke and Jonathan M. Borwein

The concept of a monotone operator — which covers both linear positive semi-definite operators and subdifferentials of convex functions — is fundamental in various branches of mathematics. Over the last few decades, several stronger notions of monotonicity have been introduced: Gossez’s maximal monotonicity of dense type, Fitzpatrick and Phelps’s local maximal monotonicity, and Simons’s monotonicity of type (NI). While these monotonicities are automatic for maximal monotone operators in reflexive Banach spaces and for subdifferentials of convex functions, their precise relationship is largely unknown.

Here, it is shown — within the beautiful framework of Convex Analysis — that for continuous linear monotone operators, all these notions coincide and are equivalent to the monotonicity of the conjugate operator. This condition is further analyzed and illustrated by examples.

1. Introduction.

Motivation.

A monotone operator is a (possibly set-valued) map from a Banach space to its dual satisfying a certain relation. In the simplest case, when the space is just the real line, this relation corresponds precisely to increasing (possibly set-valued) functions, hence the name.

Monotone operators appear in diverse areas such as Operator Theory, Numerical Analysis, Differentiability Theory of Convex Functions, and Partial Differential Equations, because the notion of a monotone operator is broad enough to cover two fundamental mathematical objects: Linear positive semi-definite operators and subdifferentials of convex functions. Although the former object gave rise to the field, it is the latter that has been receiving much of the recent attention. (For more on monotone operators, the reader is referred to the conference proceedings [11, 4, 31], the books
The urge to extract and study the quite strong monotonicity properties of subdifferentials of convex functions has led to the introduction of several new more powerful notions of monotonicity. While these notions are automatic for maximal monotone operators on reflexive Banach spaces, the situation in nonreflexive Banach spaces is far less well understood. Surprisingly, these notions of monotonicity were largely untested even for the most natural candidates: Continuous linear monotone operators. Thus:

The aim of this paper is to study the various notions of monotonicity for continuous linear monotone operators.

Using elegant and potent tools from Convex Analysis, we show that these notions all coincide with the monotonicity of the conjugate operator.

In contrast to the subdifferential case, this condition is not automatic and we present a new derivation of two classical counter-examples. Using Banach space theory, it can be shown that monotonicity of the conjugate operator is the rule — with the notable exception of spaces containing a complemented copy of $\ell_1$.

Overview.

In Section 2, we introduce the various notions of monotonicity coined by Gossez, by Fitzpatrick and Phelps, and by Simons and then review their basic relationship. From Section 3 on, we focus on the case when the monotone operator is continuous and linear. The main result, whose proof depends crucially on Fenchel's Duality Theorem, is presented in Section 4. It allows us to give an affirmative answer to a question posed by Gossez more than two decades ago. In the last section, we derive and extend classical counter-examples by Gossez and by Fitzpatrick and Phelps systematically from a result that can be viewed as an "instruction manual" for constructing interesting continuous linear monotone operators whose skew parts have nonmonotone conjugates. We conclude by remarking that such strange operators occur only in a few classical Banach spaces like $\ell_1$ and $L_1[0,1]$ and that preliminary results on regularizations demonstrate the close relationship between the various monotonicities even in a nonlinear context.

Notation.

The notation we employ is standard. Throughout, we assume that

$X$ is a real Banach space with norm $\| \cdot \|$ and dual $X^*$.

The evaluation of a functional $x^* \in X^*$ at a point $x \in X$ is written as $\langle x^*, x \rangle$ or as $\langle x, x^* \rangle$. We often view $X$ as a subspace in its bidual $X^{**} := (X^*)^*$. The unit ball $\{ x \in X : \|x\| \leq 1 \}$ is denoted $B_X$. If $(x^*_\alpha)$ is a net in some dual space, then we write $x^*_\alpha \overset{w*}{\to} x^*$ (resp. $x^*_\alpha \to x^*$) to indicate convergence in the weak* (resp. norm) topology with limit point $x^*$. If $x, y \in X$, then $[x, y]$
stands for the line segment \(\{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}\). If \(U\) (resp. \(V\)) is a subset of \(X\) (resp. \(X^*\)), then \(U^\perp\) (resp. \(+V\)) stands for the annihilator \(\{x^* \in X^* : \langle x^*, u \rangle = 0, \forall u \in U\}\) (resp. \(\{x \in X : \langle x, v \rangle = 0, \forall v \in V\} = X \cap V^\perp\)).

If \(T\) is a continuous linear operator from \(X\) to some other Banach space, then the conjugate (or adjoint, transpose) is denoted \(T^*\), the restriction of \(T\) to some subset \(U\) of \(X\) is written as \(T|_U\), and \(\ker T\) is the kernel (or null space) of \(T\): \(\ker T := \{x \in X : Tx = 0\}\). Suppose \(C\) is a subset of \(X\). Then \(\text{span} C\) of \(C\) (i.e. the set of all linear combinations of elements of \(C\)). Also, \(cl C\) (resp. \(\text{int} C\)) stands for the closure (resp. interior) of \(C\); here, the norm topology is the “default topology”. If these operations are meant with respect to some other topology \(T\), then we indicate this by subscripts; for instance, \(cl_{\tau} C\) would be the closure of \(C\) with respect to the topology \(\tau\).

Suppose \(Y\), \(Z\) are sets and \(T\) is a set-valued map from \(Y\) to \(Z\), i.e., \(T\) is a map from \(Y\) to \(2^Z\). Then the graph of \(T\) is denoted \(\text{gra} T\); so \(z \in Ty\) if and only if \((y, z) \in \text{gra} T, \forall y \in Y, z \in Z\). The domain (resp. range) of \(T\) is given by \(\text{dom} T := \{y \in Y : Ty \neq \emptyset\}\) (resp. \(\text{ran} T := \{z \in Z : z \in Ty\\) for some \(y \in Y\)\}). The inverse of \(T\), denoted \(T^{-1}\), is the set-valued map from \(Z\) to \(Y\) defined by \(y \in T^{-1} z\) if and only if \(z \in Ty, \forall y \in Y, z \in Z\). If \(U\) is a subset of \(Y\), then we write \(T(U)\) for \(\bigcup_{u \in U} Tu\).

Notation from Convex Analysis appears throughout the paper. For the reader’s convenience, we review the definitions. The indicator function of a subset \(C\) of \(X\), denoted \(\iota_C\), is given by

\[
\iota_C : X \to \mathbb{R} \cup \{+\infty\} : x \mapsto \begin{cases} 0, & \text{if } x \in C; \\ +\infty, & \text{otherwise.} \end{cases}
\]

Suppose \(f\) is a convex lower semi-continuous function from \(X\) to \(\mathbb{R} \cup \{+\infty\}\).

Then the (essential) domain of \(f\) is the set \(\text{dom} f := \{x \in X : f(x) < +\infty\}\).

The conjugate of \(f\), denoted \(f^*\), is given by

\[
f^* : X^* \to \mathbb{R} \cup \{+\infty\} : x^* \mapsto \sup_{x \in X} \langle x^*, x \rangle - f(x).
\]

Note that \(f^*\) is defined on \(X^*\) and hence \(f^{**} := (f^*)^*\) is defined on \(X^{**}\).

The subdifferential of \(f\) at \(x \in \text{dom} f\) is the set

\[
\{x^* \in X^* : \langle x^*, y - x \rangle \leq f(y) - f(x), \ \forall y \in X\}.
\]

If \(x \in \text{int dom} f\) and \(\partial f(x)\) is singleton, then the element in \(\partial f(x)\) coincides with the (Gâteaux) gradient \(\nabla f(x)\).

Finally, the reals (resp. strictly positive integers \(\{1, 2, 3, \ldots\}\)) are abbreviated \(\mathbb{R}\) (resp. \(\mathbb{N}\)) and we used already \(\forall\) (resp. \(\exists\)) as a short form for “for all” (resp. “there exists”).

As general references on Functional Analysis, we recommend [5, 17, 18, 30]; for more on Convex Analysis, see [25, 1, 7, 15, 16].
The single most important tool from Convex Analysis is the celebrated Fenchel Duality Theorem:

**Fact 1.1.** (Fenchel Duality; see, e.g., [1, Theorem 4.6.1].) Suppose $A$ is a continuous linear operator from a Banach space $X$ to a Banach space $Y$, $f$ is a convex lower semi-continuous function on $X$ and $g$ is a convex lower semi-continuous function on $Y$. Consider the convex programs

$$p := \inf_{x \in X} \left[ f(x) + g(Ax) \right]$$ (P)

and

$$d := -\inf_{y^* \in Y^*} \left[ f^*(-A^*y^*) + g^*(y^*) \right].$$ (D)

Then $p \geq d$. If $A(\text{dom } f) \cap \text{int } \text{dom } g \neq \emptyset$ and $p$ is finite, then $p = d$ and $d$ is attained.

### 2. General tools.

Recall that a set-valued map from $X$ to $X^*$ is a monotone operator if

$$\langle Tx - Ty, x - y \rangle \geq 0, \quad \forall x, y \in X.$$

If $T$ is monotone and $\text{gra } T$ is a maximal subset in $X \times X^*$, then $T$ is called maximal monotone. Zorn’s Lemma guarantees the existence of maximal monotone extensions for any given monotone operator. Analogously, one can speak of (maximal) monotone operators from $X$ to $X^*$ or from $X^*$ to $X$ or of monotone operators whose graphs are maximal monotone with respect to some subsets and so forth.

The following extensions have turned out to be useful when studying the nonreflexive case.

**Definition 2.1.** Suppose $T$ is a set-valued map from $X$ to $X^*$. Define set-valued maps $T_1, T_0, \mathcal{T}$ from $X^{**}$ to $X^*$ via their graphs as follows:

(i) $(x^{**}, x^*) \in \text{gra } T_1$, if there exists a bounded net $(x_\alpha, x_\alpha^*)$ in $\text{gra } T$ with $x_\alpha \rightharpoonup x^{**}$ and $x_\alpha^* \to x^*$.

(ii) $(x^{**}, x^*) \in \text{gra } T_0$, if $\inf_{(y, y^*) \in \text{gra } T} \langle y^* - x^*, y - x^{**} \rangle = 0$.

(iii) $(x^{**}, x^*) \in \text{gra } \mathcal{T}$, if $\inf_{(y, y^*) \in \text{gra } T} \langle y^* - x^*, y - x^{**} \rangle \geq 0$.

**Proposition 2.2.** Suppose $T$ is a monotone operator from $X$ to $X^*$. Then the following inclusions hold in $X^{**} \times X^*$:

$$\text{gra } T \subseteq \text{gra } T_1 \subseteq \text{gra } T_0 \subseteq \text{gra } \mathcal{T} = \text{gra } \mathcal{T}_1 \cap (X^{**} \times X^*).$$

**Proof.** The inclusions $\text{gra } T \subseteq \text{gra } T_1$ and $\text{gra } T_0 \subseteq \text{gra } \mathcal{T} \supseteq \text{gra } \mathcal{T}_1 \cap (X^{**} \times X^*)$ are obvious (even without monotonicity). Fix an arbitrary $(x^{**}, x^*) \in \text{gra } T_1$ and obtain a bounded net $(x_\alpha, x_\alpha^*)$ in $\text{gra } T$ with $x_\alpha \rightharpoonup x^{**}$ and $x_\alpha^* \to x^*$. Then $(x_\alpha - y, x_\alpha^* - y^*) \geq 0$, $\forall \alpha, (y, y^*) \in \text{gra } T$; taking limits yields $(x^{**} -$...
\[ y, x^* - y^* \geq 0. \] On the other hand, \( \langle x^{**} - x_{\alpha}, x^* - x_{\alpha}^* \rangle \to 0 \); altogether, \( 0 = \inf_{(y,y^*) \in \text{gra} T} \langle x^{**} - y, x^* - y^* \rangle \), i.e., \( (x^{**}, x^*) \in \text{gra} T_0 \). Hence \( \text{gra} T_1 \subseteq \text{gra} T_0 \). Finally, pick \((z^{**}, z^*) \in \text{gra} T\). Then \( 0 \leq \inf_{(y,y^*) \in \text{gra} T} \langle y^* - z^*, y - z^{**} \rangle \leq \lim_{\alpha} \langle x^*_{\alpha} - z^*, x_{\alpha} - z^{**} \rangle = \langle x^* - z^*, x^{**} - z^{**} \rangle \); so \((z^{**}, z^*)\) is monotonically related to \( \text{gra} T_1 \), hence \( \text{gra} T \subseteq \text{gra} T_1 \cap (X^{**} \times X^*) \). □

**Definition 2.3.** Suppose \( T \) is a monotone operator from \( X \) to \( X^* \). Then:

(i) (Gossez [14]) \( T \) is of dense type or of type (D), if \( T^1 = T \).

(ii) (Simons [28, Definition 14]) \( T \) is of range-dense type or of type (WD), if for every \( x^* \in \text{ran} T \), there exists a bounded net \((x_\alpha, x_\alpha^*) \in \text{gra} T \) with \( x_\alpha^* \to x^* \).

(iii) (Simons [28, Definition 10]) \( T \) is of type (NI), if \( \inf_{(y,y^*) \in \text{gra} T} \langle y^* - x^*, y - x^{**} \rangle \leq 0 \), for all \((x^{**}, x^*) \in X^{**} \times X^* \). If this holds only on some subset of \( X^{**} \times X^* \), then we say that \( T \) is of type (NI) with respect to this subset.

(iv) (Fitzpatrick and Phelps [8, Section 3]) \( T \) is locally maximal monotone, if \( (\text{gra} T^{-1}) \cap (V \times X) \) is maximal monotone in \( V \times X \), for every convex open set \( V \) in \( X^* \) with \( V \cap \text{ran} T \neq \emptyset \).

(v) \( T \) is unique, if all maximal monotone extensions of \( T \) in \( X^{**} \times X^* \) coincide.

A monotone operator which is either maximal monotone and of dense type or locally maximal monotone is certainly maximal monotone; the converse is true in reflexive spaces:

**Fact 2.4.** (See, e.g., Phelps’s [22, Example 3.2.(b) and Proposition 4.4].) Suppose \( X \) is reflexive and \( T \) is a monotone operator from \( X \) to \( X^* \). Then TFAE: (i) \( T \) is maximal monotone; (ii) \( T \) is maximal monotone and of dense type; (iii) \( T \) is locally maximal monotone.

It is known and very useful that subdifferentials of convex functions are “everything”: Maximal monotone, of dense type, and locally maximal monotone.

**Fact 2.5.** Suppose \( f \) is convex lower semi-continuous proper function on \( X \). Then:

(i) (Rockafellar [26]) \( \partial f \) is maximal monotone.

(ii) (Gossez [12, Théorème 3.1]) \( \partial f \) is of dense type and \( (\partial f)_1 = (\partial f^*)^{-1} \).

(iii) (Simons [27]) \( \partial f \) is locally maximal monotone.

In general, the following is known to be true.

**Fact 2.6.** (Simons’s [28, Lemma 15 and Theorem 19].) For any monotone operator \( T \) from \( X \) to \( X^* \), the following implications hold:

- dense type \( \Rightarrow \) range-dense type \( \Rightarrow \) type (NI) \( \Rightarrow \) unique.

Moreover, TFAE:
Then suppose Corollary 3.2.

Fact 3.1 yields immediately: if ran \( T \) is weakly compact, then \( T \) is maximal monotone.

It is sometimes more handy to work with the following reformulations of the various monotonicities.

**Proposition 2.7.** Suppose \( T \) is a monotone operator from \( X \to X^* \). Then:

(i) \( T \) is of dense type if and only if \( T_1 = ran \ T \).

(ii) \( T \) is of range-dense type if and only if \( ran \ T_1 = ran \ T \).

(iii) \( T \) is of type (NI) if and only if \( T_0 = T \).

(iv) (Phelps’s [22, Proposition 4.3]) \( T \) is locally maximal monotone if and only if for every weak* closed convex bounded subset \( C \) of \( X^* \) with \( ran \ T \cap int C \neq \emptyset \), and for every \( x_0 \in X \), \( x_0 \in (int C) \setminus T x_0 \), there exist \( (z, z^*) \in gra \ T \cap (X \times C) \) with \( (z^* - x_0^*, z - x_0) < 0 \).

**Proof.** (i): “\( \Rightarrow \)”: \( T \) is of dense type \( \iff T_1 = T \). Now \( T_1 \) is monotone (because \( T \) is), hence so is \( T \). By Fact 2.6, \( T = T_1 \) is maximal monotone. “\( \Leftarrow \)”: Pick \( (x^*, x^*) \in gra \ T \). Then (by Proposition 2.2) \( (x^*, x^*) \in gra \ T_1 \), i.e., this point is monotonically related to \( gra \ T_1 \). Now \( T_1 \) is maximal monotone, hence \( (x^*, x^*) \in gra \ T_1 \).

(ii): “\( \Rightarrow \)”: Pick \( x^* \in ran \ T \). By assumption, there exists a bounded net \((x_\alpha, x_\alpha^*) \in gra \ T \) such that \( x_\alpha^* \to x^* \). Without loss, we can assume that \( x_\alpha \weakstar x^* \). Then \( (x^*, x^*) \in gra \ T_1 \) and in particular \( x^* \in ran \ T_1 \). “\( \Leftarrow \)” is even simpler.

(iii): Let us abbreviate \( \inf_{(y, y^*) \in gra \ T} (x^* - y, x^* - y^*) \) by \( I \). “\( \Rightarrow \)”: If \( (x^*, x^*) \in gra \ T \), then \( I \geq 0 \). Now \( T \) is of type (NI), hence \( I \leq 0 \). Thus \( I = 0 \). “\( \Leftarrow \)”: Fix \( (x^*, x^*) \in X^* \times X^* \). If \( (x^*, x^*) \notin gra \ T \), then \( I < 0 \). Otherwise, \( (x^*, x^*) \in T = T_0 \) and hence \( I = 0 \).

\[ \square \]

3. Linear tools.

For a continuous linear operator \( T \) from \( X \to X^* \), the extension \( T_1 \) has the following explicit description:

**Fact 3.1.** (Gossez’s [13, End of Section 2].) Suppose \( T \) is a continuous linear operator from \( X \to X^* \). Then:

\[ gra \ T_1 = cl_{\weakstar \times X^*} gra \ T = (gra \ T^*) \cap (X^* \times X^*). \]

Recall that a continuous linear operator \( T \) from \( X \to X^* \) is weakly compact, if \( ran \ T^* \mid_{X \times X^*} \subseteq X^* \); equivalently, if \( cl \ T(B_X) \) is weakly compact or if \( T^* \) is weakly compact. Fact 3.1 yields immediately:

**Corollary 3.2.** Suppose \( T \) is a continuous linear operator from \( X \to X^* \). Then \( T \) is weakly compact if and only if \( T_1 = T^* \).
Further recall that if \( T \) is a continuous linear operator from \( X \) to \( X^* \) with \( \langle Tx, x \rangle \geq 0, \forall x \in X \), then \( T \) is called positive or positive semi-definite. The following result is part of the folklore.

**Proposition 3.3.** Suppose \( T \) is a continuous linear operator from \( X \) to \( X^* \). Then TFAE: (i) \( T \) is positive; (ii) \( T \) is monotone; (iii) \( T \) is maximal monotone.

**Proof.** By linearity of \( T \), (i) and (ii) are equivalent; also, (iii) implies (ii). For “(ii)⇒(iii)” see, e.g., [22, Proof of Example 1.5.(b)]. \( \square \)

Monotonicity of type (NI) relates to monotonicity of the conjugate operator:

**Proposition 3.4.** Suppose \( T \) is a continuous linear operator from \( X \) to \( X^* \). Then \( T \) is monotone and of type (NI) with respect to \( gra (-T^*) \) if and only if \( T^* \) is monotone.

**Proof.** Clearly, if \( T^* \) is monotone, then so is \( T \). So suppose \( T \) is monotone. Fix \( x^* \in X^* \) and \( x \in X \). Then
\[
\langle Tx + T^*x^*, x - x^* \rangle = \langle Tx, x \rangle - \langle Tx, x^* \rangle + \langle T^*x^*, x \rangle - \langle T^*x^*, x^* \rangle \geq -\langle T^*x^*, x^* \rangle.
\]
Hence
\[
-\langle T^*x^*, x^* \rangle \leq \inf_{x \in X} \langle Tx + T^*x^*, x - x^* \rangle \leq \langle 0 + T^*x^*, 0 - x^* \rangle = -\langle T^*x^*, x^* \rangle
\]
and thus:
\[
\inf_{(y,y^*) \in gra T} \langle y^* - (-T^*x^*), y - x^* \rangle = \inf_{x \in X} \langle Tx + T^*x^*, x - x^* \rangle = -\langle T^*x^*, x^* \rangle.
\]
The result follows. \( \square \)

Recall also that a continuous linear operator from \( X \) to \( X^* \) is symmetric (resp. skew), if \( T^*|_X = T \) (resp. \( T^*|_X = -T \)); equivalently, if \( \langle Tx, y \rangle = \langle x, Ty \rangle \) (resp. \( \langle Tx, y \rangle = -\langle x, Ty \rangle \)), \forall x, y \in X.

Our study of continuous linear monotone operators relies also on the following easy-to-prove yet immensely useful decomposition principle.

**Proposition 3.5.** Suppose \( T \) is a continuous linear operator from \( X \) to \( X^* \). Then \( T \) can be written as the sum of two continuous linear operators, \( T = P + S \), where \( P \) is symmetric and \( S \) is skew. This decomposition is unique; in fact:
\[
P x = \frac{1}{2} Tx + \frac{1}{2} T^* x \quad \text{and} \quad S x = \frac{1}{2} Tx - \frac{1}{2} T^* x, \quad \forall x \in X.
\]
We refer to \( P \) (resp. \( S \)) as the symmetric part (resp. skew part) of \( T \).
Symmetric operators.

**Theorem 3.6.** Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with symmetric part $P$. Let $q(x) := \frac{1}{2} \langle x, Tx \rangle$, $\forall x \in X$. Then:

$q$ is convex $\iff T$ is monotone $\iff P$ is monotone.

Assume in addition that $T$ is monotone. Then:

(i) $\nabla q = P$ and $q^* \circ P = q$.

(ii) $P_1 = P_0 = \overline{P} = P^* = P^{**}$. Hence: $P$ is maximal monotone of dense type, weakly compact, and locally maximal monotone; $P^*$ is monotone and symmetric.

(iii) For every $x^{**} \in X^{**}$, there exists a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \rightharpoonup^{ws} x^{**}$ and $P x_\alpha \to P^* x^{**} = P^{**} x^{**}$.

(iv) $q^*(P^* x^{**}) = \frac{1}{2} \langle x^{**}, P^* x^{**} \rangle = q^{**}(x^{**})$, for every $x^{**} \in X^{**}$ and $\nabla q^{**} = P^{**} = P^*$.

**Proof.** Since $q$ is continuous, it suffices to check midpoint convexity; fixing two arbitrary points $x, y \in X$, we have $q(\frac{1}{2} x + \frac{1}{2} y) \leq \frac{1}{2} q(x) + \frac{1}{2} q(y) \iff 0 \leq \langle x - y, T(x - y) \rangle \iff 0 \leq \langle x - y, P(x - y) \rangle$. The displayed equivalences follow.

(i): $q$ is continuous, convex, and finite on $X$; hence $q$ is subdifferentiable everywhere. So fix an arbitrary $x_0 \in X$ and pick $x^* \in \partial q(x_0)$. Then $t \langle x^*, h \rangle \leq q(x_0 + th) - q(x_0), \forall h \in X, t > 0$; this simplifies to $\langle x^*, h \rangle \leq \langle \frac{1}{2} T x_0 + \frac{1}{2} T^* x_0, h \rangle + \frac{1}{2} t \langle h, Th \rangle$. Letting $t$ tend to $0$ yields $x^* = \frac{1}{2} T x_0 + \frac{1}{2} T^* x_0 = P x_0$. Now

$$q^*(P x_0) = \sup_{x \in X} \langle P x_0, x \rangle - q(x) = -\inf_{x \in X} \{ q(x) + \langle -P x_0, x \rangle \};$$

this last infimum can be viewed as a little optimization problem which is easy to solve: Indeed, after taking gradients, we learn that the set of minimizers equals $x_0 + \ker P$. It follows that $q^*(P x_0) = q(x_0)$.

(ii): By (i), $P$ is the subdifferential of $q$. Consequently (Fact 2.5), $P$ is maximal monotone, of dense type, and locally maximal monotone. In particular (Fact 2.6 and Proposition 2.7.(iii)), $P_1 = P_0 = \overline{P}$ and $P$ is of type (NI). It follows that on the one hand, $P^*$ is a maximal monotone extension of $P$ (Proposition 3.4 and Proposition 3.3). On the other hand, $\overline{P}$ is the unique maximal monotone extension of $P$ in $X^{**} \times X^*$ (Fact 2.6). Altogether, $\overline{P} = P^*$. Now $P^* = P^{**}$, because $P_1 = P^*$ and $\gra P_1 \subseteq \gra P^{**}$ (Fact 3.1). Thus the weak compactness of $P$ follows from Corollary 3.2.

(iii): By (ii), $P_1 = P^* = P^{**}$.

(iv): Fix $x^{**} \in X^{**}$ and define $g(x) := \langle -P^* x^{**}, x \rangle + \frac{1}{2} \langle x^{**}, P^* x^{**} \rangle, \forall x \in X$. Then, by (ii), $(x^{**}, P^* x^{**}) \in \gra P_0$ and hence

$$0 = \frac{1}{2} \inf_{x \in X} \langle P x - P^* x^{**}, x - x^{**} \rangle = \inf_{x \in X} q(x) + g(x).$$
The conjugate of \( g \) is given by \( g^*(x^*) = -\frac{1}{2}(x^{**}, P^*x^{**}) + \iota_{\{-P^*x^{**}\}}(x^*) \), \( \forall x^* \in X^* \). Fact 1.1 yields
\[
0 = - \inf_{x^* \in X^*} \{ q^*(x^*) + g^*(-x^*) \} = \frac{1}{2}(x^{**}, P^*x^{**}) - q^*(P^*x^{**}),
\]
which is the first equality. To prove the second equality, we first note that the first equality implies
\[
q^{**}(x^{**}) \geq (x^{**}, P^*x^{**}) - q^*(P^*x^{**}) = \frac{1}{2}(x^{**}, P^*x^{**}).
\]
On the other hand, by (iii), there is a bounded net \((x_\alpha)\) in \( X \) such that \( x_\alpha \rightharpoonup x^{**} \) and \( P x_\alpha \to P^*x^{**} \). Then for every \( x^* \in X^* \), we estimate
\[
q^*(x^*) \geq \lim_{\alpha} \langle x^*, x_\alpha \rangle - \frac{1}{2} \langle x_\alpha, P x_\alpha \rangle = \langle x^{**}, x^* \rangle - \frac{1}{2} \langle x^{**}, P^*x^{**} \rangle.
\]
This in turn implies \( \frac{1}{2} \langle x^{**}, P^*x^{**} \rangle \geq \sup_{x^* \in X^*} \langle x^{**}, x^* \rangle - q^*(x^*) = q^{**}(x^{**}) \), which yields the second equality. Applying (i) to \( P^{**} \), which is monotone and symmetric, finally yields \( \nabla q^{**} = P^{**} \).

**Skew operators.**

**Theorem 3.7.** Suppose \( S \) is a continuous linear skew operator from \( X \) to \( X^* \). If \((x^{**}, x^*) \in X^{**} \times X^* \), then:
(i) \( x^* \in S_1 x^{**} \iff x^* = S^{**} x^{**} = -S^* x^{**} \).
(ii) \( x^* \in S_0 x^{**} \iff x^* = -S^* x^{**} \) and \( \langle S^* x^{**}, x^{**} \rangle = 0 \).
(iii) \( x^* \in S x^{**} \iff x^* = -S^* x^{**} \) and \( \langle S^* x^{**}, x^{**} \rangle \leq 0 \).

Hence: \( S^* \) is skew if and only if \( S \) is weakly compact.

If \( S^* \) is monotone and \( x^{**} \in X^{**} \) with \( \langle S^* x^{**}, x^{**} \rangle = 0 \), then \( S^{**} x^{**} = -S^* x^{**} \).

**Proof.** First “If” part: Fix an arbitrary \( y \in X \). Then, using the skewness of \( S \), \( \langle x^{**} - y, x^* - S y \rangle = \langle x^{**}, x^* \rangle - \langle y, S^* x^{**} \rangle \). Hence
\[
\inf_{(y^*, y^{**}) \in \text{gra} S} \langle x^{**} - y, x^* - y^* \rangle = \begin{cases} 
\langle x^{**}, x^* \rangle, & \text{if } x^* = -S^* x^{**}; \\
-\infty, & \text{otherwise}.
\end{cases}
\]
(ii) and (iii) follow readily. For (i), observe that: \( x^* \in S_1 x^{**} \iff (x^{**}, x^*) \in \text{gra} S_1 \cap \text{gra} S_0 \) (Proposition 2.2) \( \iff x^* = S^{**} x^{**} = -S^* x^{**} \in X^* \) (ii) and Fact 3.1) \( \iff x^* = S^{**} x^{**} = -S^* x^{**} \) (ran \( S^* \subseteq X^* \)).

“Hence” part: \( S^* \text{ skew} \iff S^{**} = -S^* \iff S_1 = S^{**} \) (use (i)) \( \iff S \text{ weakly compact} \) (Corollary 3.2).

Second “If” part: Fix an arbitrary \( y^{**} \in X^{**} \) and \( \lambda > 0 \). Thus:
\[
0 = \langle S^{**} x^{**}, x^* \rangle
\]
\[
= \langle S^*(x^{**} + \lambda y^{**}), x^{**} + \lambda y^{**} \rangle - \lambda \langle S^*(x^{**} + \lambda y^{**}), y^{**} \rangle - \lambda \langle S^* y^{**}, x^{**} \rangle
\]
\[
\geq -\lambda \langle S^* x^{**}, y^{**} \rangle - \lambda \langle S^* y^{**}, x^{**} \rangle - \lambda^2 \langle S^* y^{**}, y^{**} \rangle.
\]
Now divide by $\lambda$ and then let $\lambda$ tend to 0 to conclude $\langle S^*x^*, y^* \rangle \geq -\langle S^*y^*, x^* \rangle$, $\forall y^* \in X^*$. The result follows. \hfill $\square$


We are now ready for the main result.

**Theorem 4.1.** Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with symmetric part $P$ and skew part $S$. Then TFAE:

(i) $T$ is monotone and of dense type.
(ii) $T$ is monotone and of range-dense type.
(iii) $T$ is monotone and of type (NI).
(iv) $T$ is locally maximal monotone.
(v) $T^*$ is monotone.
(vi) $P$ and $S^*$ are monotone.
(vii) $P$ is monotone and $S$ is of dense type.
(viii) $P$ is monotone and $S$ is of range-dense type.
(ix) $P$ is monotone and $S$ is of type (NI).
(x) $P$ is monotone and $S$ is locally maximal monotone.

**Proof.** Throughout, let $q(x) := \frac{1}{2} \langle Tx, x \rangle = \frac{1}{2} \langle Px, x \rangle$, $\forall x \in X$.

“(i)⇒(ii)⇒(iii)”: Fact 2.6.

“(ii)⇒(v)”: Proposition 3.4.

“(v)⇒(vi)”: $T$ and $P$ are monotone, because $T^*$ is. Fix an arbitrary $x^* \in X^*$. By Theorem 3.6.(iii), obtain a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \rightharpoonup x^*$ and $Px_\alpha \to P^*x^*$. Now

$$0 \leq \langle T^*(x^* - x_\alpha), x^* - x_\alpha \rangle = \langle T^*x^*, x^* - x_\alpha \rangle - \langle T^*x_\alpha, x^* - x_\alpha \rangle = \langle T^*x^* - Px_\alpha, x^* - x_\alpha \rangle - \langle Px_\alpha - Sx_\alpha, x^* - x_\alpha \rangle $$

$$= \langle T^*x^* - Px_\alpha, x^* - x_\alpha \rangle + \langle x_\alpha, S^*x^* \rangle \to \langle x^*, S^*x^* \rangle;$$

consequently, $S^*$ is monotone and (vi) holds.

“(vi)⇒(i)”: We start by noting that if $(x^*, x^*)$ belongs to $X^* \times X^*$, then

$$\frac{1}{2} \inf_{x \in X} \langle Tx - x^*, x - x^* \rangle = \frac{1}{2} \langle x^*, x^* \rangle - q^* \left( \frac{1}{2} x^* + \frac{1}{2} T^*x^* \right).$$

Now fix an arbitrary $(x^*, x^*) \in \text{gra } T$. Then, on the one hand, $q^* \left( \frac{1}{2} x^* + \frac{1}{2} T^*x^* \right) \leq \frac{1}{2} \langle x^*, x^* \rangle$. On the other hand, Theorem 3.6.(iii) yields a bounded net $(x_\alpha)$ in $X$ such that $x_\alpha \rightharpoonup x^*$ and $Px_\alpha \to P^*x^*$. Using the monotonicity of $S^*$, we conclude altogether

$$\frac{1}{2} \langle x^*, x^* \rangle \geq q^* \left( \frac{1}{2} x^* + \frac{1}{2} T^*x^* \right) \geq \lim_{\alpha} \frac{1}{2} \left( \frac{1}{2} x^* + \frac{1}{2} T^*x^*, x_\alpha \right) - \frac{1}{2} \langle x_\alpha, Px_\alpha \rangle$$

$$= \frac{1}{2} \langle x^*, x^* \rangle + \frac{1}{2} \langle S^*x^*, x^* \rangle \geq \frac{1}{2} \langle x^*, x^* \rangle.$$
Hence \( \langle S^*x^*, x^* \rangle = 0 \) and \( q^*(\frac{1}{2}x^* + \frac{1}{2}T^*x^*) = \frac{1}{2}\langle x^*, x^* \rangle \). This has two important consequences: firstly, by Theorem 3.7,

\[ S^*x^* = -S^*x^*. \]

Secondly, using Theorem 3.6.(iv), \( \langle \frac{1}{2}x^* + \frac{1}{2}T^*x^*, x^* \rangle = \frac{1}{2}\langle x^*, x^* \rangle + \frac{1}{2}\langle x^*, P^*x^* \rangle = q^*(\frac{1}{2}x^* + \frac{1}{2}T^*x^*) + q^*(x^*); \) thus: \( x^* \in \partial q^*(\frac{1}{2}x^* + \frac{1}{2}T^*x^*) \Rightarrow \frac{1}{2}x^* + \frac{1}{2}T^*x^* \in \partial q^*(x^*) = \{P^*x^*\} \Leftrightarrow x^* = P^*x^* - S^*x^*. \)

Altogether, \( x^* = P^*x^* - S^*x^* = P^*x^* + S^*x^* = T^*x^* \in X^* \), so that (Fact 3.1) \( (x^*, x^*) \in \text{gra} T_1 \), as desired.

“(iv)\( \Rightarrow \)(v)”: \( T \) is maximal monotone (use \( V = X^* \) in Definition 2.3.(iv)), hence so is \( P \) and the function \( q \) is convex (Theorem 3.6). Fix an arbitrary \( x_0^* \in X^* \). We aim for \( \langle T^*x_0^*, x_0^* \rangle \geq 0 \) and can thus assume WLOG that \( x_0^* := T^*x_0^* \neq 0 \). Select \( x_1 \in X \) with \( \langle x_0^*, x_1 \rangle < 0 \) and let \( x_1^* := Tx_1 \). Let \( x_0^* := 0 \), fix an arbitrary \( \epsilon > 0 \), and define

\[ C_\epsilon := [x_0^*, x_1^*] + \epsilon B_{X^*}. \]

Then \( C_\epsilon \) is weak* closed, convex, bounded with \( x_1^* \in \text{ran} T \cap \text{int} C_\epsilon \). Also, \( x_0^* \in (\text{int} C_\epsilon) \setminus T x_0^* \). Local maximal monotonicity (via Proposition 2.7.(iv)) and Fact 1.1 yield

\[ 0 \geq \inf_{x \in X ; Tx \in C_\epsilon} \langle Tx - x_0^*, x - x_0^* \rangle = \inf_{x \in X} q(x) + \langle -\frac{1}{2}x_0^*, x \rangle + \mu_{C_\epsilon}(Tx) \]

\[ \geq -\inf_{x^* \in X^*} \left\{ q^* \left( -T^*x^* + \frac{1}{2}x_0^* \right) + \mu_{C_\epsilon}^*(x^*) \right\}. \]

Now pick \( x^* := \frac{1}{2}x_0^* \); then, using the fact that \( q^*(0) = 0 \),

\[ 0 < q^* \left( -\frac{1}{2}T^*x_0^* + \frac{1}{2}x_0^* \right) + \mu_{C_\epsilon}^* \left( \frac{1}{2}x_0^* \right) = \epsilon \left\| x_0^* \right\| + \max \left\{ \left\langle \frac{1}{2}x_0^*, x_0^* \right\rangle, \left\langle \frac{1}{2}x_0^*, x_1^* \right\rangle \right\}. \]

Multiply by 2 and let \( \epsilon \) tend to 0 to obtain \( 0 \leq \max \{ \langle x_0^*, T^*x_0^* \rangle, \langle x_0^*, T x_1 \rangle \} \). Since \( \langle x_0^*, T x_1 \rangle = \langle T^*x_0^*, x_1 \rangle = \langle x_0^*, x_1 \rangle < 0 \), we obtain \( \langle x_0^*, T^*x_0^* \rangle \geq 0 \).

“(vi)\( \Rightarrow \)(iv)”: In view of Proposition 2.7.(iv), let us fix a weak* closed convex bounded subset \( C \) of \( X^* \) with \( \text{ran} T \cap \text{int} C \neq \emptyset \), \( x_0 \in X \), \( x_0^* \in (\text{int} C) \setminus T x_0^* \). Let

\[ p := \inf_{x \in X ; T x \in C} \frac{1}{2}\langle Tx - x_0^*, x - x_0^* \rangle. \]

Clearly, \( p < +\infty \) and our aim is \( p < 0 \). We thus can assume WLOG that \( p > -\infty \), hence \( p \) is finite. Let \( f(x) := q(x) + \frac{1}{2}\langle -x_0^* - T^*x_0^* + \frac{1}{2}\langle x_0^*, x_0^* \rangle, \]

\[ \forall x \in X, \text{ and let } g := \iota_C. \text{ Then, using Fact 1.1,} \]
\[ p = \inf_{x \in X} f(x) + g(Tx) = -\inf_{x^{**} \in X^{**}} \{ f^*(-T^*x^{**}) + g(x^{**}) \} \]
\[ = \frac{1}{2} \langle x_0^*, x_0 \rangle - \inf_{x^{**} \in X^{**}} \left\{ q^* \left( -T^*x^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0 \right) + \iota_C^*(x^{**}) \right\}. \]

Moreover: The last infimum is attained (by Fact 1.1 and \( \text{ran } T \cap \text{int } C \neq \emptyset \)), say at some \( x_0^{**} \in X^{**} \). Thus the proof of "(vi)\( \Rightarrow \) (iv)" would be complete after reaching the following

(Aim) \[ \frac{1}{2} \langle x_0^*, x_0 \rangle < q^* \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0 \right) + \iota_C^*(x_0^{**}). \]

By assumption, \( 0 \leq \langle S^*(x_0 - 2x_0^{**}), x_0 - 2x_0^{**} \rangle \), which is equivalent to
\[ \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0, x_0 - 2x_0^{**} \right) - \frac{1}{2} \langle x_0 - 2x_0^{**}, P^*(x_0 - 2x_0^{**}) \rangle \geq \frac{1}{2} \langle x_0^*, x_0 \rangle - \langle x_0^{**}, x_0^{**} \rangle. \]

On the other hand, Theorem 3.6.(iii) gives a bounded net \((x_{\alpha})\) in \( X \) such that \( x_{\alpha} w^{\star} \rightarrow x_0 - 2x_0^{**} \) and \( P x_{\alpha} \rightarrow P^*(x_0 - 2x_0^{**}) \); thus altogether
\[ q^* \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0 \right) \]
\[ \geq \lim_{\alpha} \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0, x_{\alpha} \right) - \frac{1}{2} \langle x_{\alpha}, P x_{\alpha} \rangle \]
\[ = \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0, x_0 - 2x_0^{**} \right) \]
\[ - \frac{1}{2} \langle x_0 - 2x_0^{**}, P^*(x_0 - 2x_0^{**}) \rangle \]
\[ \geq \frac{1}{2} \langle x_0^*, x_0 \rangle - \langle x_0^{**}, x_0^{**} \rangle. \]

Consequently, since \( x_0^{**} \) is in the interior of \( C \),
\[ \frac{1}{2} \langle x_0^*, x_0 \rangle \leq q^* \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0 \right) + \langle x_0^{**}, x_0^{**} \rangle \]
\[ < q^* \left( -T^*x_0^{**} + \frac{1}{2} x_0^* + \frac{1}{2} T^*x_0 \right) + \iota_C^*(x_0^{**}), \]

which is what we aimed for.

We just proved that (i)-(vi) are equivalent for an arbitrary continuous linear operator \( \hat{T} \) from \( X \) to \( X^* \). If we apply this to \( \hat{T} = S \), then the remaining items are readily seen to be equivalent as well. \( \square \)
Remark 4.2. Gossez [13, End of Section 2] found the following question interesting:

Suppose that $T$ is a closed densely defined linear monotone operator from $X$ to $X^*$ and that $T^*$ is monotone. Is $T_1$ maximal monotone?

He then proved that the answer is “yes” if $T$ is continuous and skew. We are now able to give an affirmative answer to this question provided that $T$ is merely continuous: Indeed, this follows from Theorem 4.1. “(v)$\Rightarrow$(i)” and Proposition 2.7.(i).

Theorem 4.3. Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with symmetric part $P$ and skew part $S$. Then TFAE:

(i) $T$ and $T^*|X$ are monotone and of dense type.
(ii) $T$ and $T^*|X$ are monotone and of type (NI).
(iii) $T$ and $T^*|X$ are locally maximal monotone.
(iv) $T^*$ and $(T^*|X)^*$ are monotone.
(v) $T$ is monotone and weakly compact.
(vi) $P$ is monotone and $S$ is weakly compact.
(vii) $P$ is monotone and $S^*$ is skew.
(viii) $P$ is monotone and $S$, $-S$ are of dense type.
(ix) $P$ is monotone and $S$, $-S$ are of type (NI).
(x) $P$ is monotone and $S$, $-S$ are locally maximal monotone.

Proof. Applying Theorem 4.1 to $T = P + S$ and $T^*|X = P - S$ yields the equivalence of (i), (ii), (iii), (iv), (v), (vi), (vii), (ix), and (x). Now (v)$\Leftrightarrow$(vi) (by weak compactness of $P$; see Theorem 3.6.(ii)) $\Leftrightarrow$(vii) (by Theorem 3.7.(v)); so (i)-(x) are all equivalent. $\square$

In hindsight, we can interpret monotonicity of the conjugate of the skew part of a given continuous linear monotone operator as “one half of weak compactness”.

5. Examples and concluding remarks.

Suppose we are given a continuous linear monotone operator $T$ from $X$ to $X^*$ with skew part $S$. In view of Theorem 4.1 and Theorem 4.3, the following three mutually exclusive alternatives are conceivable:

- $T$ is “good”: Both $S^*$ and $-S^*$ are monotone.
- $T$ is “so-so”: Either $S^*$ or $-S^*$ is monotone but not both.
- $T$ is “bad”: Neither $S^*$ nor $-S^*$ is monotone.

A priori, it is not clear that “so-so” or “bad” operators exist. However, this is indeed the case and we will now systematically recover two classical examples.
Theorem 5.1. Suppose $T$ is a continuous linear operator from $X$ to $X^*$ with skew part $S$ and there exists some $e \in X^*$ such that
\[ e \notin \text{cl ran } T \quad \text{and} \quad \langle Tx, x \rangle = \langle e, x \rangle^2, \quad \forall x \in X. \]
Then $T$ is monotone but $S^*$ is not.

If $\text{ran } T^* = \text{ran } T^*|_X$, equivalently, if $(\text{ran } T)^* \subseteq X$, then $-S^*$ is monotone.

Proof. $T$ is obviously monotone.

Let $Px := \langle e, x \rangle e$, $\forall x \in X$; then $\langle P^*x^*, x \rangle = \langle x^*, Px \rangle = \langle x^*, e \rangle \langle e, x \rangle$ and hence $P^*x^* = \langle x^*, e \rangle e$, $\forall x^* \in X^*$. So $P$ is symmetric. Consider now $S := T - P$. Then $\langle Sx, x \rangle = \langle Tx, x \rangle - \langle Px, x \rangle = \langle Tx, x \rangle - \langle e, x \rangle^2 = 0$, $\forall x \in X$, thus $S$ is skew. Since $T = P + S$, the symmetric (resp. skew) part of $T$ is $P$ (resp. $S$) by Proposition 3.5. Because $e \notin \text{cl ran } T = \overline{\text{ker } T^*}$, there exists some $x^*_0 \in \ker T^*$ with $\langle x^*_0, e \rangle \neq 0$. Hence
\[ \langle S^*x^*_0, x^*_0 \rangle = \langle T^*x^*_0, x^*_0 \rangle - \langle P^*x^*_0, x^*_0 \rangle = 0 - \langle x^*_0, e \rangle^2 < 0; \]
so $S^*$ is not monotone.

"If" part: First note that since $\text{ran } T \subseteq X^*$, the Hahn/Banach Theorem allows us to identify $(\text{ran } T)^*$ with $\{x^{**}|_{\text{ran } T} : x^{**} \in X^{**}\}$. We thus derive the "equivalently" part as follows.

$\text{ran } T^* = \text{ran } (T^*|_X) \Leftrightarrow \forall x^{**} \in X^{**} \quad \exists \hat{x} \in X : T^*x^{**} = T^*\hat{x}$
\[ \Leftrightarrow \forall x^{**} \in X^{**} \quad \exists \hat{x} \in X : T^*x^{**} = T^*\hat{x} \quad \forall x \in X : \langle x^{**} - \hat{x}, Tx \rangle = 0 \]
\[ \Leftrightarrow \forall x^{**} \in X^{**} \quad \exists \hat{x} \in X : x^{**}|_{\text{ran } T} = \hat{x}|_{\text{ran } T} \quad \forall x \in X \subseteq X^{**} : z^* = \hat{x}|_{\text{ran } T} \]
\[ \Leftrightarrow (\text{ran } T)^* \subseteq X. \]

Now fix an arbitrary $x^{**} \in X^{**}$. Then there exists $\hat{x} \in X \subseteq X^{**}$ with $T^*x^{**} = T^*\hat{x}$. Thus we have $\langle S^*x^{**}, x \rangle = \langle T^*x^{**}, x \rangle - \langle P^*x^{**}, x \rangle = \langle T^*\hat{x}, x \rangle - \langle P^*x^{**}, x \rangle$, $\forall x \in X$; hence
\[ S^*x^{**} = T^*\hat{x} - P^*x^{**} = T^*\hat{x} - \langle x^{**}, e \rangle e. \]

Because $T^*|_X = P - S = 2P - T$, we further obtain
\[ \langle S^*x^{**}, x^{**} \rangle = \langle T^*\hat{x}, x^{**} \rangle - \langle x^{**}, e \rangle^2 = 2\langle P\hat{x}, x^{**} \rangle - \langle T^*\hat{x}, x^{**} \rangle - \langle x^{**}, e \rangle^2 \]
\[ = 2\langle P\hat{x}, x^{**} \rangle - \langle \hat{x}, T^*x^{**} \rangle - \langle x^{**}, e \rangle^2 \]
\[ = 2\langle e, \hat{x} \rangle \langle x^{**}, e \rangle - \langle e, \hat{x} \rangle^2 - \langle x^{**}, e \rangle^2 = -\langle e, x^{**} - \hat{x} \rangle^2 \leq 0; \]
consequently, $-S^*$ is monotone. \qed

Here is the announced example of a “so-so” operator.
Example 5.2 (Gossez). Define the map $G$ from $\ell_1$ to $\ell_\infty$ by

$$(Gx)_n := -\sum_{k < n} x_k + \sum_{k > n} x_k, \quad \forall x = (x_k) \in \ell_1, n \in \mathbb{N}.$$ 

Then: $G$ and $-G$ are continuous linear skew operators from $\ell_1$ to $\ell_1^* = \ell_\infty$.

$G^*$ is not monotone whereas $-G^*$ is; consequently:

$G$ is neither of type (NI) nor locally maximal monotone;

$-G$ is both of dense type and locally maximal monotone.

**Proof.** Consider the map $T$ from $\ell_1$ to $\ell_\infty$ given by

$$(Tx)_n := x_n + 2 \sum_{k > n} x_k, \quad \forall x = (x_k) \in \ell_1, n \in \mathbb{N}.$$ 

Then $T$ is linear, continuous (in fact, $\|T\| = 2$), and $\text{ran } T \subseteq c_0 \subseteq \ell_\infty$. Let $e := (1, 1, 1, \ldots) \in \ell_\infty = \ell_1^*$. Then $e \notin \text{cl ran } T \subseteq \text{cl } c_0 = c_0$ and for every $x \in \ell_1$,

$$\langle Tx, x \rangle = \sum_n x_n \left(x_n + 2 \sum_{k > n} x_k\right) = \sum_n x_n^2 + \sum_{n > k} 2x_n x_k$$

$$= \sum_n x_n^2 + \sum_{n \neq k} x_n x_k$$

$$= \sum_{n,k} x_n x_k = \left(\sum_n x_n \cdot 1\right) \left(\sum_k x_k \cdot 1\right) = \langle e, x \rangle^2.$$ 

The proof of Theorem 5.1 shows that the symmetric part $P$ of $T$ is given by $Px = \langle e, x \rangle e$, $\forall x \in \ell_1$, and the skew part of $T$ is $S := T - P$. Now for all $x \in \ell_1, n \in \mathbb{N}$:

$$(Sx)_n = (Tx)_n - (Px)_n = x_n + 2 \sum_{k > n} x_k - \sum_k x_k$$

$$= -\sum_{k < n} x_k + \sum_{k > n} x_k = (Gx)_n;$$

hence $S = G$. By Theorem 5.1, $G^*$ is not monotone. Hence (Theorem 4.1) $G$ is neither of type (NI) nor locally maximal monotone. Because $(\text{ran } T)^* \subseteq c_0^* = \ell_1$, Theorem 5.1 yields that $-G^*$ is monotone. By Theorem 4.1, $-G$ is of dense type and locally maximal monotone. □

Somewhat surprisingly, the “continuous” version of the (negative) Gossez operator yields a “bad” operator.

Example 5.3 (Fitzpatrick and Phelps). Define the map $F$ from $L_1[0,1]$ to $L_\infty[0,1]$ by

$$(Fx)(t) := \int_0^t x(s)ds - \int_t^1 x(s)ds, \quad \forall x \in L_1[0,1], t \in [0,1].$$
Then $F, -F$ are continuous linear skew operators from $L_1[0, 1]$ to $L_1^*[0, 1] = L_\infty[0, 1]$.

Neither $F^*$ nor $-F^*$ is monotone; consequently:

$F$ and $-F$ are not of type (NI) nor locally maximal monotone.

**Proof.** *Step 1.* Define the map $T$ from $L_1[0, 1]$ to $L_\infty[0, 1]$ by

$$(Tx)(t) := 2 \int_0^t x(s) \, ds, \quad \forall x \in L_1[0, 1], t \in [0, 1].$$

Then $T$ is linear and continuous (with $\|T\| = 2$). The range of $T$ is contained in the subspace $C_{0,0}$ of $L_\infty[0, 1]$ that consists of all equivalence classes that contain a continuous function vanishing at 0. Let $e$ denote the equivalence class in $L_\infty[0, 1]$ that contains the constant function 1. Then the distance from $e$ to any member in $C_{0,0}$ is at least 1; thus certainly $e \notin \text{cl ran } T$. Also, for every $x \in L_1[0, 1],$

$$
\langle Tx, x \rangle = 2 \int_0^1 \left( \int_0^t x(s) \, ds \right) x(t) \, dt = 2 \int \int_{s \leq t} x(s) x(t) \, ds \, dt
= \int \int_{[0,1] \times [0,1]} x(s) x(t) \, ds \, dt
= \left( \int_0^1 x(s) \cdot 1 \, ds \right) \left( \int_0^1 x(t) \cdot 1 \, dt \right) = \langle x, e \rangle^2.
$$

Then (see again the proof of Theorem 5.1) the positive part $P$ of $T$ is given by $Px := \langle e, x \rangle e, \forall x \in L_1[0, 1]$. The skew part $S$ of $T$ is given by

$$(Sx)(t) := (Tx)(t) - (Px)(t) = 2 \int_0^t x(s) \, ds - \langle e, x \rangle e(t)
= 2 \int_0^t x(s) \, ds - \int_0^1 x(s) \, ds
= \int_0^t x(s) \, ds - \int_t^1 x(s) \, ds,
$$

for every $x \in L_1[0, 1], t \in [0, 1]$; consequently, $S = F$. Now Theorem 5.1 and Theorem 4.1 imply that $F^*$ is not monotone and $F$ is neither of type (NI) nor locally maximal monotone.

*Step 2.* Here we define the map $T$ by $(Tx)(t) := 2 \int_0^1 x(s) \, ds, \forall x \in L_1[0, 1], t \in [0, 1]$. We let $e$ be as in Step 1 and check analogously: $T$ is continuous, linear, $e \notin \text{cl ran } T$, and $\langle Tx, x \rangle = \langle x, e \rangle^2, \forall x \in L_1[0, 1]$. This time, however, the skew part of $T$ is equal to $-F$. We deduce as in Step 1 that $-F^*$ is not monotone and that $-F$ is neither of type (NI) nor locally maximal monotone.

**Remark 5.4.** Let $G$ be the operator from Example 5.2. Gossez [13] proved that $G$ is not of dense type whereas Phelps [22, Example 4.5] showed that $G$...
is not locally maximal monotone. Let $F$ be the operator from Example 5.3. Fitzpatrick and Phelps [9, Example 3.2] showed that $F$ is not locally maximal monotone.

We observe that our discussion of $G$ and $F$ via Theorem 5.1 is much simpler.

We conclude by reporting on two sets of results that are closely connected to the present paper. We omit the proofs as we think the results are not in their final form; nonetheless, the interested reader is able to find the details in [3] or in [2].

**Remark 5.5** (conjugate monotone spaces). We say that $X$ is a conjugate monotone space (cms), if the conjugate of every continuous linear monotone operator from $X$ to $X^*$ is monotone as well. Thus $X$ is a conjugate monotone space precisely when it does not admit “so-so” or “bad” operators; equivalently, when every continuous linear monotone operator from $X$ to $X^*$ is weakly compact.

It is clear that reflexive spaces are (cms). Also, one can use Example 5.2 and Example 5.3 to show that none of $\ell_1$, $L_1[0,1]$, $\ell_1^*$, $L_1^*[0,1]$ is (cms).

A result relying on deeper Banach space theory states that for a Banach lattice $X$, TFAE:

(i) $X$ is (cms).

(ii) $X$ is $(w)$, i.e., every continuous linear operator from $X$ to $X^*$ is weakly compact.

(iii) $X$ does not contain a complemented copy of $\ell_1$.

As a consequence, the classical spaces $c_0$, $c$, $\ell_1$, $L_1[0,1]$, $\ell_\infty$, $L_\infty[0,1]$ are all (cms). It would be interesting to know whether or not (i)–(iii) are still equivalent for Banach spaces. (Note that (i) always implies (iii).)

**Remark 5.6** (some nonlinear results). Recall that the duality map, denoted $J$, is the subdifferential map of the function $\frac{1}{2}\| \cdot \|^2$. Suppose $T$ is a continuous linear monotone operator from $X$ to $X^*$ and $\lambda > 0$. Then the operator $T + \lambda J$ is called a regularization of $T$. Regularizations of monotone operators constitute perhaps the simplest class of nonlinear monotone operators. Then one can show that TFAE:

(i) $T$ is of dense type.

(ii) $\text{cl ran} (T + \lambda J) = X^*$, $\forall \lambda > 0$.

(iii) $T + \lambda J$ is of range-dense type, $\forall \lambda \geq 0$.

If the underlying space $X$ is rugged (i.e., $\text{cl span ran} (J - J) = X^*$), then (i)-(iii) are also equivalent to:

(iv) $\text{cl ran} (T + \lambda J)$ is convex, $\forall \lambda > 0$.

(v) $T + \lambda J$ is locally maximal monotone, $\forall \lambda > 0$.

The full equivalence of (i)–(v) thus holds for the following rugged spaces: $c_0$, $c$, $\ell_1$, $\ell_\infty$, $L_1[0,1]$, $L_\infty[0,1]$, $C[0,1]$.
Conclusion.

Maximal monotonicities of dense type, range-dense type, or type (NI), and local maximal monotonicity all coincide: • For subdifferentials of convex functions (Fact 2.5); • in reflexive spaces; • for continuous linear monotone operators (Theorem 4.1). These monotonicities always hold for subdifferentials of convex functions. They may well be absent for continuous linear monotone operators (Example 5.2 and Example 5.3); however, in reflexive and most of the classical nonreflexive spaces, they are automatic (Remark 5.5).

The question whether or not the monotonicities all coincide for a general maximal monotone operator remains open.

Some preliminary results (Remark 5.6) seem to indicate that this may well be the case.

Acknowledgments.

It is our great pleasure to thank Tamás Erdélyi, Simon Fitzpatrick, and Jon Vanderwerff for various helpful comments and discussions. We also wish to thank Robert Phelps and Stephen Simons for kindly sending us [23, 24, 28, 29]. We would like to refer the reader especially to [24], which contains many results on linear maximal monotone operators that are not necessarily continuous.

References


Received November 25, 1996 and revised August 28, 1998. The first author’s research was supported by an NSERC Postdoctoral Fellowship. The second author’s research was supported by NSERC and by the Shrum Endowment.

Okanagan University College
Kelowna, B.C. V1V 1V7
Canada
E-mail address: bauschke@cecm.sfu.ca

Simon Fraser University
Burnaby, British Columbia V5A 1S6
Canada
E-mail address: jborwein@cecm.sfu.ca
ON THE BOUNDEDNESS OF SINGULAR INTEGRALS

EUGENE FABES, IRINA MITREA, AND MARIUS MITREA

We present a simple, elementary proof of the \( T(1) \) theorem of G. David and J.-L. Journé, as well as related results, which is based on a Krein type lemma.

1. Introduction.

Establishing the \( L^2 \)-boundedness of singular integral operators is a fundamental problem in harmonic analysis which has always enjoyed center stage. For a broader view of this active area of research see, e.g., \([Ca]\), \([CMM]\), \([MM]\), \([CDM]\), \([DJ]\), \([DJS]\), \([CJS]\), \([Me]\), \([Ch]\), \([St]\) and the references therein.

The main aim of this note is to indicate yet another way of proving such \( L^2 \) boundedness results which is inspired by an old lemma of M. G. Krein (\([Kr]\)). Essentially, the latter asserts that if a linear operator \( T \) and its formal adjoint \( T^\ast \) are bounded on a Banach space \( \mathcal{X} \) which is densely and continuously embedded in a Hilbert space \( \mathcal{H} \), then \( T \) extends to a bounded operator on \( \mathcal{H} \). We use this to give a new proof of the celebrated \( T(1) \) theorem of David and Journé (\([DJ]\)) in its full strength.

The actual context in which this lemma applies for a singular integral operator \( T \) satisfying the usual set of hypotheses of the \( T(1) \) theorem is when \( \mathcal{H} = L^2 \) and \( \mathcal{X} = C^\alpha \). Of course, one first needs to localize the problem for the latter space to embed properly in the former but, more importantly, one has to ensure boundedness for \( T \) at the level of Hölder continuous functions. However, this is essentially well known and requires appropriate cancellations for \( T \) which, in turn, are secured by subtracting off paraproducts. Other choices for \( \mathcal{X} \) are possible but the one just described yields perhaps the most elementary proof.

This proof is, in principle, quite flexible; in fact, we shall prove a “real-valued” version of the \( T(b) \) theorem (cf. \([MM]\), \([DJS]\)), i.e. when the constant 1 is replaced by an arbitrary positive, measurable function \( b \) which is bounded away from zero and infinity. The approach can be adapted to spaces of homogeneous type in the sense of \([CoWe]\) and we believe that the same strategy may also be successful in other cases of interest.

Acknowledgments. We thank Y. Brudnyi, A. McIntosh and M. Christ for several discussions and a suggestion which has been incorporated.
2. Statement of theorem.

Recall that a distribution \( k \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n) \) is called a standard kernel if its singular support lies on the diagonal and, for some constant \( C > 0 \), satisfies

\[
|k(x, y)| \leq C|x - y|^{-n}, \quad \forall \ x \neq y, \tag{2.1}
\]

\[
|\nabla_x k(x, y)| + |\nabla_y k(x, y)| \leq C|x - y|^{-(n+1)} \quad \forall \ x \neq y. \tag{2.2}
\]

Let \( \|k\| := \inf \{C; (2.1) - (2.2) \text{ hold}\} \). A linear, continuous operator \( T : \mathcal{D}(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n) \) is called a singular integral operator (of Calderón-Zygmund type) if its Schwartz distribution kernel is standard. Also, \( T \) is said to have the weak boundedness property if there exists a constant \( C \) such that

\[
|\langle T\phi, \psi \rangle| \leq CR^n(|\phi|_{\infty} + R|\nabla \phi|_{\infty})(|\psi|_{\infty} + R|\nabla \psi|_{\infty}) \tag{2.3}
\]

holds uniformly for \( \phi, \psi \in C_0^\infty(B_R(z)) \), \( R > 0 \), \( z \in \mathbb{R}^n \). Here, and elsewhere, \( B_R(z) \) stands for the Euclidean ball of radius \( R \) centered at \( z \in \mathbb{R}^n \). We set

\[
\|T\|_{\text{WBP}} := \inf \{C; \text{ so that (2.3) holds uniformly}
\]

in the natural parameters.

To state the main result, recall that BMO stands for the (John-Nirenberg) space of functions of bounded mean oscillations in \( \mathbb{R}^n \). Also, a measurable function \( b : \mathbb{R}^n \rightarrow \mathbb{C} \) is called accretive if there exists \( \kappa > 0 \) (the accretivity constant) such that \( \text{Re} \ b(x) \geq \kappa > 0 \) for a.e. \( x \in \mathbb{R}^n \). Finally, \( M_b \) is the operator of (pointwise) multiplication by \( b \).

**Theorem 2.1 ([DJ], [MM], [DJS]).** Let \( b_1, b_2 \) be two essentially bounded, accretive, real-valued functions and let \( T \) be a linear continuous operator from \( b_1C_0^\infty(\mathbb{R}^n) \) into \( (b_2C_0^\infty(\mathbb{R}^n))^\prime \). Assume that \( M_{b_2}TM_{b_1} \) satisfies the weak boundedness property and that, if \( K(x, y) \) stands for its distributional kernel, then \( \text{singsupp} K(x, y) \subseteq \text{diag}(\mathbb{R}^n \times \mathbb{R}^n) \) and the kernel \( k(x, y) := b_2(x)^{-1}K(x, y)b_1(y)^{-1} \) satisfies (2.1)-(2.2).

Then, if \( T(b_1), T^t(b_2) \in \text{BMO} \), where \( T^t \) is the formal adjoint of \( T \), it follows that \( T \) can be extended to a bounded operator on \( L^2(\mathbb{R}^n) \) with operator norm controlled by \( \|k\|, \|T(b_1)\|_{\text{BMO}}, \|T^t(b_2)\|_{\text{BMO}} \) and \( \|M_{b_2}TM_{b_1}\|_{\text{WBP}} \).

As usual, \( T(b_1) \) is defined as a “non-standard” distribution in view of the fact that it pairs well with elements from \( b_2C_0^\infty(\mathbb{R}^n) \) having vanishing moment. Something similar applies to \( T^t(b_2) \).

Before commencing the actual proof, we shall record a couple of reduction steps which are going to be important in subsequent arguments. The first one, i.e. the reduction to the case when \( T(b_1) = T^t(b_2) = 0 \), is classical. As usual, this is accomplished by subtracting from \( T \) two paraproduct operators \( \mathcal{L}, \mathcal{M} \) satisfying

\[
\mathcal{L}(b_1) = T(b_1), \quad \mathcal{L}^t(b_2) = 0 \quad \text{and} \quad \mathcal{M}^t(b_2) = T^t(b_2), \quad \mathcal{M}(b_1) = 0. \tag{2.5}
\]
The existence of such operators is well known; cf., e.g., [DJS], [Me], [Da].
In the case when \( b_1 = b_2 = 1 \), the operators \( \mathcal{L} \) and \( \mathcal{M} \) have a particularly simple form; see, e.g., [DJ], [Ch]. Re-denoting by \( T \) the difference \( T - \mathcal{L} - \mathcal{M} \), it follows that \( T \) satisfies the same hypotheses as in Theorem 2.1 and, in addition, \( T(b_1) = T^t(b_2) = 0 \).

Next, note that it suffices to show that \( T \) maps \( L^2(B_R(0)) \) boundedly into itself with norm controlled by the same constitutive constants as in the statement of Theorem 2.1, uniformly in \( R \). That this implies the desired conclusion is easily seen by letting \( R \) go to \( \infty \) and using Fatou’s lemma. Furthermore, it is enough to prove this only for \( R = 1 \), i.e. for

\[
(2.6) \quad T : L^2(B_1(0)) \rightarrow L^2(B_1(0)).
\]

This is a manifestation of the dilation invariant nature of our assumptions. Specifically, if \( (D_\rho f)(x) := f(px) \) is the dilation operator of factor \( \rho > 0 \), then \( D_\rho T(D_\rho^{-1} f) = T_R f \) where \( T_R \) is the operator associated with the kernel \( R^k k(Rx, Ry) \). It is trivial to check that this kernel satisfies (2.1)-(2.2) with the same constant \( \|k\| \). Also, \( M_{D_R b_1} T_R M_{D_R b_2} \) satisfies the weak boundedness property with the same constant which works for \( M_{b_1} T M_{b_2} \), and, since \( D_\rho \) is an isometry of BMO for any \( \rho > 0 \), \( \|T_R(D_R b_1)\|_{\text{BMO}} = \|T(b_1)\|_{\text{BMO}} \), \( \|T_R(D_R b_2)\|_{\text{BMO}} = \|T^t(b_2)\|_{\text{BMO}} \). Of course, \( D_R b_1, D_R b_2 \) remain bounded and accretive with the same bounds as \( b_1, b_2 \).


With an eye toward proving the boundedness of the operator in (2.6) (with appropriate control) which suffices for our purposes, we shall now analyze the action of operators of this type on spaces of Hölder continuous functions. To state the main result in this direction recall that, for \( 0 < \alpha < 1 \) and \( \Omega \subseteq \mathbb{R}^n \),

\[
(3.1) \quad C^\alpha(\Omega) := \left\{ f : \Omega \rightarrow \mathbb{C}; \|f\|_{C^\alpha} := \|f\|_\infty + \sup_{x,y \in \Omega} |x-y|^{-\alpha}|f(x) - f(y)| < \infty \right\}.
\]

**Proposition 3.1.** Let \( T \) be a singular integral operator with kernel \( k \) and let \( b_1, b_2 \) be two essentially bounded, accretive functions such that \( M_{b_2} T M_{b_1} \) has the weak boundedness property and that \( T(b_1) = 0 \). Then for any \( \alpha \in (0,1) \) and any \( \eta \in C^\infty_c(B_2(0)) \) the operator

\[
(3.2) \quad M_\eta T M_\eta : b_1 C^\alpha(B_2(0)) \rightarrow C^\alpha(B_2(0))
\]

is bounded with operator norm controlled in terms of \( \alpha, \|k\|, \eta, \|M_{b_2} T M_{b_1}\|_{\text{WBP}}, \|b_1\|_\infty, \|b_2\|_\infty \) and the accretivity constants.

The proof of this proposition involves only elementary estimates and closely related results have been known for a long time (compare with, e.g.,
However, for the reader’s convenience, below we sketch a simple argument. To facilitate the presentation, we first isolate a technical result (which is precisely where the weak boundedness property is used).

**Lemma 3.2.** With the hypotheses of Proposition 3.1, there holds

\[
\left| \int_{|x-y|<\delta} k(x, y)b_1(y) \, dy \right| \leq C, \quad x \in \mathbb{R}^n
\]

uniformly for \(0 < \delta < \infty\), where \(C\) is controlled by the same parameters as in the statement of Proposition 3.1.

**Proof.** To begin with, let us note that the integral in (3.3) may not converge absolutely and, in fact, must be interpreted in a distributional sense, as explained below.

Denote by \(\chi_E\) the characteristic function of a set \(E \subseteq \mathbb{R}^n\). Also, fix some even function \(\psi \in C^\infty_c(B_3(0))\) with \(\psi \equiv 1\) in \(B_2(0)\) and set \(\psi_t,w(x) := \psi(\frac{x-w}{t})\) for \(t > 0\) and \(w \in \mathbb{R}^n\). Then, since \(T(b_1) = 0\), we may formally write

\[
\int_{|x-y|<\delta} k(x, y)b_1(y) \, dy
= \int_{|x-y|>\delta} (k(x, y) - k(z, y)) \, b_1(y) \, dy
- T(b_1\psi^{\delta,x})(z) - \int_{\mathbb{R}^n} k(z, y)b_1(y)(\chi_{B_3(x)} - \psi^{\delta,x})(y) \, dy
=: I_1(x, z) + I_2(x, z) + I_3(x, z).
\]

Before going any further a comment is in order here. The point is that, generally speaking, the above integrals are to be understood in the sense of distributions. In particular, in order to derive size estimates, they should be integrated against (arbitrary) test functions. It is precisely in this sense that all our subsequent estimates must be interpreted even though, in order to shorten and simplify the exposition, we shall continue to manipulate such expressions in a formal manner. Of course, the main emphasis is to obtain bounds which depend exclusively on the relevant constants (as in the statement of Theorem 2.1).

Next, we shall average the identity (3.4). To this end, let \(\phi \in C^\infty_c(B_{1/2}(0))\) be an even, nonnegative function with \(\int_{\mathbb{R}^n} \phi \, dx = 1\). Also, denote the argument of the complex number \(\int_{|x-y|<\delta} k(x, y)b_1(y) \, dy\) by \(\theta = \theta(x, \delta) \in [0, 2\pi)\).
Then, by virtue of the accretivity of $b_2$,

$$\begin{align*}
C \left| \int_{|x-y|<\delta} k(x,y)b_1(y) \, dy \right| & \leq \Re \int_{\mathbb{R}^n} b_2(z) \delta^{-n} \phi_{\delta,z}(z) e^{i\theta} \left( \int_{|x-y|<\delta} k(x,y)b_1(y) \, dy \right) \, dz \\
& = \sum_{j=1}^{3} \Re \int_{\mathbb{R}^n} b_2(z) \delta^{-n} \phi_{\delta,z}(z) e^{i\theta} I_j(x,z) \, dz \\
& =: II_1(x) + II_2(x) + II_3(x).
\end{align*}$$

(3.5)

For $II_1(x)$, the mean value theorem, (2.2) and the fact that $|x-z| \leq \frac{1}{2}|x-y|$ on the domain of integration give

$$|II_1(x)| \leq C \|k\|_1 \|b_1\|_\infty \|b_2\|_\infty \delta^{-n} \int_{\mathbb{R}^n} \phi_{\delta,z}(z) \int_{\delta<|x-y|<|x-z|} \frac{|x-z|}{|x-y|^{n+1}} \, dydz \leq C.$$  

(3.6)

Also, $II_2(x)$ is clearly controlled by $\|M_{b_2}TM_{b_1}\|_{\text{WBP}}$, whereas

$$|II_3(x)| \leq C \|b_1\|_\infty \|b_2\|_\infty \int_{\mathbb{R}^n} \delta^{-n} \phi_{\delta,z}(z) \left( \int_{\delta/2<|z-y|<7\delta/2} |k(z,y)| \, dy \right) \, dz.$$  

(3.7)

Since $k$ is standard, the last inner integral is trivially bounded independently of $\delta$ and $z$ and the desired conclusion follows.

Based on this result it is now easy to tackle the

Proof of Proposition 3.1. Fix some $0 < \alpha < 1$ and take an arbitrary $f \in C^\alpha(B_2(0))$. Then, for $x \in B_2(0)$ we write

$$
T(b_1\eta f)(x) = \int_{|x-y|\leq 4} k(x,y)b_1(y)((\eta f)(y) - (\eta f)(x)) \, dy \\
+ (\eta f)(x) \int_{|x-y|\leq 4} k(x,y)b_1(y) \, dy.
$$

In particular,

$$\|T(M_{\eta}b_1 f)\|_{L^\infty(B_2(0))} \leq C \|f\|_{C^\alpha(B_2(0))}.$$  

(3.8)

Next, for $x, h \in \mathbb{R}^n$ with $|x|, |x+h| \leq 2$, we may write on account of the fact that $T(b_1) = 0$

$$|T(b_1\eta f)(x+h) - T(b_1\eta f)(x)| \leq I + II + III,$$
where

\[ I := \left| \int_{|x-y| \geq 2|h|} \left( k(x+h, y) - k(x, y) \right) b_1(y)((\eta f)(y) - (\eta f)(x)) \, dy \right|, \]

\[ II := \left| \int_{|x-y| \leq 2|h|} k(x, y)b_1(y)((\eta f)(y) - (\eta f)(x)) \, dy \right|, \]

\[ III := \left| \int_{|x-y| \leq 2|h|} k(x+h, y)b_1(y)((\eta f)(y) - (\eta f)(x)) \, dy \right|. \]

We seek a bound of the order of \(|h|^{\alpha} \|f\|_{C^\alpha(B_2(0))}\) for each quantity above. Based solely on (2.1)-(2.2), the first two integrals are handled in a crude, straightforward fashion and we omit the details. It is only the last term which requires one more application of Lemma 3.2. Indeed, we have

\[ |III| \leq \int_{|x-y| \leq 2|h|} |k(x+h, y)b_1(y)((\eta f)(y) - (\eta f)(x+h))| \, dy + |h|^\alpha \|\eta f\|_{C^\alpha(B_2(0))} \int_{|x-y| \leq 2|h|} k(x+h, y)b_1(y) \, dy \]

(3.9)

\[ =: III_1 + |h|^\alpha \|\eta f\|_{C^\alpha(B_2(0))} III_2. \]

The first integral, \(III_1\), is treated essentially as \(II\) above, producing a bound of the same order. Finally, for the second integral, Lemma 3.2 and (2.1) give

\[ |III_2| \leq \left| \int_{|x+h-y| < |h|} \ldots \right| + \left| \int_{|h| < |x+h-y| < 3|h|} \ldots \right| \]

(3.10)

\[ \leq C_0 + C_1 \|b_1\|_\infty \|k\| \int_{|h| < |z-y| < 3|h|} \frac{dy}{|z-y|^n} \, dy \leq C_2. \]

This yields the right estimate for \(III\) also. To sum up, we have proved that

\[ \sup_{x,y \in \mathbb{R}^n} |x-y|^{-\alpha}|T(M_\eta b_1 f)(x) - T(M_\eta b_1 f)(y)| \leq C \|f\|_{C^\alpha(B_2(0))} \]

(3.11)

which, together with (3.8), readily implies the desired conclusion. \(\square\)


What allows us to pass from continuity on spaces of Hölder continuous functions to continuity on spaces of square integrable functions is a certain functional analytic argument which is a variation of a lemma due to M.G. Krein \([Kr]\) (cf. also \([La]\), \([Di]\)).
Lemma 4.1. Let \( \mathcal{X} \) be a Banach space and assume that

\[
(\cdot, \cdot)_i : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}, \quad i = 1, 2,
\]

are two forms such that, for some \( M > 0 \),

\[
|x, y| \leq M\|x\|\|y\|\mathcal{X}, \quad \forall x, y \in \mathcal{X}, \quad i = 1, 2.
\]

We also assume that each \((\cdot, \cdot)_i, i = 1, 2\), is accretive and symmetric in the sense that

\[
\text{Re}(x, x) \geq k\|x\|\mathcal{X}, \quad \text{and} \quad (x, y)_i = (y, x)_i, \quad \forall x, y \in \mathcal{X}.
\]

Consider next a linear normed space \( \mathcal{Y} \) for which the inclusion \( i : \mathcal{Y} \hookrightarrow \mathcal{X} \) is well defined, continuous and with dense range. Finally, let \( A, B : \mathcal{Y} \rightarrow \mathcal{Y} \) be two bounded, linear operators so that

\[
(Ax, y)_1 = (x, By)_2, \quad \forall x, y \in \mathcal{Y}.
\]

Then both \( A \) and \( B \) extend to continuous operators on \( \mathcal{X} \) with

\[
\|A\|\mathcal{X} \rightarrow \mathcal{X}, \|B\|\mathcal{X} \rightarrow \mathcal{X} \leq M\kappa^{-1}\|A\|\mathcal{Y} \rightarrow \mathcal{Y}\|B\|\mathcal{Y} \rightarrow \mathcal{Y}^{1/2}.
\]

Proof. For an arbitrary \( x \in \mathcal{Y} \), using (4.2)-(4.4) it follows that

\[
\kappa\|Ax\|\mathcal{X} \leq \|(Ax, Ax)_1\| = \|(x, BAx)_2\| \leq M\|x\|\mathcal{X}\|BAx\|\mathcal{X}.
\]

Therefore,

\[
\|Ax\|\mathcal{X} \leq \kappa^{-1/2}M^{1/2}\|x\|\mathcal{X}^{1/2}\|BAx\|\mathcal{X}^{1/2}
\]

which, in particular, shows that it suffices to prove that \( BA \) extends to a continuous operator on \( \mathcal{X} \) with a suitable bound for its norm. To this end, parallelising (4.6)-(4.7) but with \( BA \) in place of \( A \) (note that \( BA \) has a bounded transpose in the sense of (4.4)) gives the estimate \( \|(BA)x\|\mathcal{X} \leq \kappa^{-1/2}M^{1/2}\|x\|\mathcal{X}^{1/2}\|BA\|\mathcal{X}^{1/2} \). The key observation is that, in fact, any power \( (BA)^{N} \) will work in place of \( BA \) in this last inequality and, hence, iterating \( j \) times according to powers of 2 gives

\[
\|(BA)x\|\mathcal{X} \\
\leq (M/\kappa)^{2^{-1}+2^{-2}j}\|x\|\mathcal{X}^{2^{-1}+2^{-2}j}\|(BA)^{2j}x\|\mathcal{X}^{2^{-j}} \\
\leq (M/\kappa)^{2^{-1}+2^{-2}j}\|x\|\mathcal{X}^{2^{-1}+2^{-2}j}\|\|\mathcal{Y} \rightarrow \mathcal{X}\|\mathcal{Y} \rightarrow \mathcal{Y}^{2^{-j}}\|\|\mathcal{Y} \rightarrow \mathcal{X}\|\mathcal{Y} \rightarrow \mathcal{Y}^{2^{-j}}\|x\|\mathcal{Y}^{2^{-j}} \\
\leq (M/\kappa)^{2^{-1}+2^{-2}j}\|x\|\mathcal{X}^{2^{-1}+2^{-2}j}\|\|\mathcal{Y} \rightarrow \mathcal{X}\|\mathcal{Y} \rightarrow \mathcal{Y}^{2^{-j}}\|BA\|\mathcal{Y} \rightarrow \mathcal{Y}^{2^{-j}}\|1\|\mathcal{Y}^{2^{-j}}.
\]

Passing to the limit \( j \rightarrow \infty \) yields \( \|(BA)x\|\mathcal{X} \leq M\kappa^{-1}\|x\|\mathcal{X}\|BA\|\mathcal{Y} \rightarrow \mathcal{Y} \) as desired.
Parenthetically, let us note that one can avoid arbitrarily high order iterations in the above reasoning. Indeed, for each \( x \in \mathcal{Y} \) with \( \|x\|_{\mathcal{X}} \leq 1 \), set

\[
K(x) := \sup \{ \|Ax\|_{\mathcal{X}} \mid \forall A, B \text{ which satisfy the hypotheses of the lemma and} \|A\|_{\mathcal{Y} \rightarrow \mathcal{Y}}, \|B\|_{\mathcal{Y} \rightarrow \mathcal{Y}} \leq 1 \}.
\]

Now, by (4.7) it follows that

\[
K(x) \leq \kappa - \frac{1}{2} M_{1/2} K(x)^{1/2}
\]

and, further, \( K(x) \leq \kappa^{-1} M \), since \( K(x) \) is finite. The estimate (4.5) is then obtained by rescaling. \( \square \)

With all ingredients in place, we are now in a position to present the final details in the proof of the Theorem 2.1. To this effect, let us recall that we have reduced matters to showing that the operator \( T \) in (2.6) is bounded with operator norm appropriately controlled. Further, it is trivial to check that this is true if we can show that, for a fixed real valued function \( \eta \in C_c^\infty(\mathbb{B}^2(0)) \) with \( \eta \equiv 1 \) on \( \mathbb{B}^1(0) \),

\[
(4.9) \quad M_\eta TM_\eta : L^2(\mathbb{B}^2(0)) \rightarrow L^2(\mathbb{B}^2(0))
\]

is bounded with the right norm control. The important thing is that \( T(b_1) = T^t(b_2) = 0 \) so that Proposition 3.1 applied both to \( T \) and \( T^t \) gives that

\[
(4.10) \quad M_\eta TM_\eta : b_1 C^\alpha(\mathbb{B}^2(0)) \rightarrow C^\alpha(\mathbb{B}^2(0)),
\]

\[
M_\eta T^t M_\eta : b_2 C^\alpha(\mathbb{B}^2(0)) \rightarrow C^\alpha(\mathbb{B}^2(0))
\]

are bounded with appropriate control.

At this point Lemma 4.1 applies to lift this to the \( L^2 \) setting in the following context. We take \( \mathcal{X} := L^2(\mathbb{B}^2(0)) \), \( \mathcal{Y} := C^\alpha(\mathbb{B}^2(0)) \), \( A := M_\eta TM_\eta M_{b_1} \), \( B := M_\eta T^t M_\eta M_{b_2} \), where, if \( U \) is an operator then \( \overline{Uf} := \overline{U}f \), and

\[
(4.11) \quad (f,g)_1 := \int f b_2 \overline{g} \, dx, \quad (f,g)_2 := \int f b_1 \overline{g} \, dx, \quad f, g \in L^2(\mathbb{B}^2(0)).
\]

Since \( b_1 \) and \( b_2 \) are real-valued we have that \( (f,g)_i = (\overline{g},f)_i \) for \( i = 1, 2 \). Next, the accretivity of \( b_i \) entails the accretivity of the corresponding paring \( (\cdot,\cdot)_i \), \( i = 1, 2 \), whereas

\[
(4.12) \quad (Af, g)_1 = \int M_\eta TM_\eta M_{b_1} f M_{b_2} \overline{g} \, dx = \int f M_{b_1} M_\eta T^t M_\eta M_{b_2} \overline{g} \, dx = \int f M_{b_1} \overline{M_\eta T^t M_\eta M_{b_2}} \overline{g} \, dx = (f, Bg)_2.
\]

Note that the third equality above follows from the fact that \( Uf = \overline{U}f \). Thus, since \( M_{b_i} \) is a isomorphism of \( L^2(\mathbb{B}^2(0)) \) onto itself, Lemma 4.1 gives that \( M_\eta TM_\eta \) is bounded on \( L^2(\mathbb{B}^2(0)) \) with the desired norm control. The proof of the Theorem 2.1 is therefore finished.
References


Received August 15, 1997 and revised July 14, 1998. The first author was partially supported by a UMC Research Board grant and a NSF grant.

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455

UNIVERSITY OF MISSOURI
COLUMBIA, MO 65211

UNIVERSITY OF MISSOURI
COLUMBIA, MO 65211

E-mail address: marius@math.missouri.edu
SMALL EXTENSIONS OF WITT RINGS

Robert W. Fitzgerald

We consider certain Witt ring extensions $S$ of a noetherian Witt ring $R$ obtained by adding one new generator. The conditions on the new generator are those known to hold when $R$ is the Witt ring of a field $F$, $S$ is the Witt ring of a field $K$ and $K/F$ is an odd degree extension. We show that if $R$ is of elementary type then so is $S$.

The elementary type conjecture is a proposed classification of noetherian Witt rings. A potential source of counter-examples is as follows: Start with a field $F$ where $WF$ is known (necessarily of elementary type) then look at noetherian $WK$ for extension fields $K$ of $F$. Jacob and Ware [3] have shown that $WK$ is again of elementary type when $[K : F] = 2$. Here we look at the simplest case of odd degree extensions, again showing $WK$ is of elementary type. We note that $WF$ is noetherian iff $G(F) \equiv F/F^2$ is finite. Also when $K/F$ has odd degree then $G(F) \cong F/K^2/K^2$ embeds into $G(K)$.

We will in fact work with abstract Witt rings $R$ (as defined by Marshall [4]) with associated group of one dimensional forms $G(R)$. The small extensions considered here are as follows. Let $H$ be a subgroup of $G(R)$. We say a Witt ring $S$ is an $H$-extension of $R$ if there exists an $\alpha \in G(S)$ such that:

1. $G(S) = \{1, \alpha\}G(R)$, and
2. For all $x \in G(R)$ we have:

$$DS\langle 1, -x \rangle = \begin{cases} DR\langle 1, -x \rangle, & \text{if } x \notin H \\ \{1, \alpha\}DR\langle 1, -x \rangle, & \text{if } x \in H \end{cases}$$

$$DS\langle 1, -\alpha x \rangle = \{1, -\alpha x\} \left(DR\langle 1, -x \rangle \cap H\right).$$

These conditions hold for $R = WF, S = WK$ when $K/F$ is an odd degree extension and $[G(K) : G(F)] = 2$ by [2, 4.7] (we note that [2, 4.7] should include the condition that $N_{K/F}(a) = 1$). No such field extensions are known. However, there are many examples of $H$-extensions of abstract Witt rings, which we determine inductively. This can be viewed as a first step in classifying extensions of noetherian Witt rings. It also helps the search for odd degree extensions $K/F$ with $[G(K) : G(F)] = 2$, while lessening the motivation for such a search.
The proof of the main result is a long series of technical lemmas, only one of which (3.1) has independent interest. However, most of the results mimic the expected behavior of the field extension case, evidence the elementary type conjecture holds. This is speculative since there may be no field extensions yielding \( H \)-extensions. However, we indulge in this suggestive speculation once, after (1.5).

For any group \( H, H' \) denotes \( H \setminus \{1\} \). The quaternionic mapping associated to \( R \) will be denoted by \( q \). For \( x \in G(R), Q(x) = \{q(x, y) : y \in G(R)\} \) and for a subgroup \( H, Q(H) = \{q(h, y) : h \in H, y \in G(R)\} \). The value set of \( (1, -x) \) is \( D(1, -x) = \{y \in G(R) : q(x, y) = 1\} \). We will often work with several Witt rings at once and write \( q_R, Q_R(x) \) and \( D_R(1, -x) \) to indicate these objects for \( R \).

\( R \) is of local type if \( |q(G(R), G(R))| = 2 \). We let \( E_n \) denote the elementary 2-group of order \( n \). The group ring \( R[E_n] \) is again a Witt ring. An element \( t \in G(R) \) is rigid if \( D(1, t) = \{1, t\} \) and \( t \) is birigid if both \( t \) and \( -t \) are rigid. The basic part of \( R, B(R) \), consists of \( \pm 1 \) and all \( x \in G(R) \) with either \( x \) or \( -x \) not rigid. \( B(R) \) is a subgroup of \( G(R) \) and \( R = R_0[G(R)/B(R)] \), where \( R_0 \) is the Witt ring generated by \( B(R) \). We express this last statement by writing \( R_0 = W(B(R)) \).

The product in the category of Witt rings is:

\[
R_1 \cap R_2 = \{(r_1, r_2) : r_i \in R_i \quad \text{and} \quad \dim r_1 \equiv \dim r_2 \pmod{2}\}.
\]

If \( R = R_1 \cap R_2 \) then \( G(R) = G(R_1) \times G(R_2) \) and:

\[
D_R((1, 1), (x, y)) = D_{R_1}(1, x) \times D_{R_2}(1, y).
\]

The radical of \( R \) is \( \text{rad}(R) = \{x \in G(R) : D(1, -x) = G(R)\} \). We say \( R \) is degenerate if \( \text{rad}(R) \neq 1 \) and totally degenerate if \( \text{rad}(R) = G(R) \). \( D_n \) denotes a totally degenerate Witt ring with square class group of order \( 2^n \).

There are two possibilities for \( D_n \) depending on whether \(-1\) is a square or not. Specifically, \( D_n \) is either a product of \( n \) copies of \( (\mathbb{Z}/2\mathbb{Z})[E_1] \) or \( n \) copies of \( \mathbb{Z}/4\mathbb{Z} \). If \( R \) is degenerate then there exist uniquely determined \( n \) and non-degenerate Witt ring \( R_0 \) such that \( R = D_n \cap R_0 \). \( R_0 \) is the non-degenerate part of \( R \).

\( R \) is of elementary type if it can be built from \( \mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/4\mathbb{Z} \) and Witt rings of local type by a succession of group ring extensions (for some \( E_n \)) and products. The elementary type conjecture is that every noetherian Witt ring is of elementary type.

1. **Group ring extensions.**

To help the reader navigate the following six sections of lemmas, we point out here the highlights. We will prove \( S \) is of elementary type by induction on \( |G(R)| \). If \( R \) is of local type then generally \( H = 1 \) or \( G(R) \) and \( S \) is determined (2.3), (1.1). The exceptions occur when \( R = L_{2,0} \) (there are
three possible $H$'s) and when $R = L_{2,1}$ (four possible $H$'s). The resulting extensions are determined in (1.5).

If $R = R_0[E_n]$ is a group ring either $H$ is one of three special subgroups or $S = S_0[E_n]$ with $S_0$ an $H$-extension of $R_0$ (1.5). If $R = R_1 \cap R_2$ is a product then $S$ is generally a product of one factor with an $H$-extension of the other factor (5.7), (6.1). There are exceptional cases when $R_1 = \mathbb{Z}$ (5.9) or $R_1$ is totally degenerate (4.1).

**Lemma 1.1.** Let $S$ be an $H$-extension of $R$.

(a) If $H = 1$ then $S = R[E_1]$, with $E_1$ generated by $\alpha$.

(b) If $H = G(R)$ then $S = D_1 \cap R$, with $D_1$ generated by $\alpha$.

*Proof.* Suppose first that $H = 1$. Then for all $g \in G(R)$ we have from the definition of an $H$-extension that $D_S(1, -\alpha g) = \{1, -\alpha g\}(D_R(1, -g) \cap H)$ and so $\alpha g$ is birigid. Thus $B(S) \subseteq G(R)$ and $G(S) = \{1, \alpha\}G(R)$. So by [4, 5.19] $S = R[E_1]$, where $E_1$ is generated by $\alpha$.

Next suppose that $H = G(R)$. Then $D_S(1, -\alpha) = \{1, -\alpha\}(D_R(1, -1) \cap H) = G(S)$. Hence $\alpha \in \text{rad}(S)$. Then by [4, pp. 105-106] $S = D_1 \cap R$, where $D_1$ is generated by $\alpha$. \hfill \Box

When $H \neq 1$, which we will often assume in light of (1.1), we use the following notation (recall that $B(R)$ is the basic part of $R$):

$$T = \bigcup_{h \in H} D_R(1, -h)$$

$$T_0 = \bigcup_{h \in H} (D_R(1, -h) \setminus \{-h\})$$

$$B(H) = H \cap B(R)$$

$$BT = \bigcup_{h \in B(H)} D_R(1, -h)$$

$$BT_0 = \bigcup_{h \in B(H)} (D_R(1, -h) \setminus \{-h\}).$$

**Lemma 1.2.** Let $S$ be an $H$-extension of $R$ with $|H| > 1$. Then:

(a) $\pm T_0 \subseteq B(R) \subseteq \pm T$,

(b) $B(S) = \pm \{1, \alpha\} T$.

*Proof.* First note that $D_S(1, -\alpha) = \{1, -\alpha\}H$, and $|H| > 1$ imply $\alpha \in B(S)$. If $x \in G(R) \setminus \pm T$ then $D_R(1, \pm x) \cap H = \{1\}$ so by the definition of $H$-extensions, $D_S(1, \pm \alpha x) = \{1, \pm \alpha x\}$. Hence $\alpha x \notin B(S)$, and as $\alpha \in B(S)$, $x \notin B(S)$. That is, $x$ is birigid in $S$, and so also in $R$. Thus $x \notin B(R)$ and we have:

$$B(R) \subseteq \pm T,$$

$$B(S) \subseteq \pm \{1, \alpha\} T.$$
Let $x \in T_0, x \neq -1$ so that for some $h \in H^*, x \in D_R(1, -h)$ and $x \neq -h$. Then $D_R(1, -x)$ contains $-x, h$ which are distinct and not equal to 1. So $-x \in B(R)$ and also $x \in B(R)$. If $x = -1$ then again $x \in B(R)$. Thus $T_0 \subset B(R)$, and so $\pm T_0 \subset B(R)$ completing the proof of (a).

If $x \in T$ with $x \in D_R(1, -h), h \in H^*$ then

$$DS(1, -x) = \{1, -\alpha x\} (D_R(1, -x) \cap H)$$

contains $\{1, -\alpha x, h, -\alpha x h\}$. Thus $-\alpha x \in B(S)$. Again $\alpha \in B(S)$ so $x \in B(S)$. This shows $\pm \{1, \alpha\}T \subset B(S)$, completing the proof of (b).

Lemma 1.3. If $S$ is an $H$-extension of $R$ with $|H| > 1$ then $B(R) = \pm BT$ and either:

(a) $B(R) = \pm B(H)$ and $B(S) = \pm \{1, \alpha\}H$, or
(b) $B(R) = \pm T$ and $B(S) = \{1, \alpha\}B(R)$.

Proof. If $h \in H \setminus B(H)$ then $D_R(1, -h) = \{1, -h\}$. So:

(1.4) $$\pm T = \pm BT \cup \pm (H \setminus B(H)).$$

Now $B(R) \subset \pm T$ by (1.2)(a) and $B(R) \cap \pm (H \setminus B(H)) = \emptyset$ so $B(R) \subset \pm BT$.

Conversely,

$$\pm BT = \pm (BT \cup B(H)) \subset \pm T_0 \cup \pm B(H) \subset B(R),$$

which proves the first statement.

Now (1.4) gives:

$$\pm T = \pm BT \cup \pm H = B(R) \cup \pm H,$$

and (1.2)(b) gives:

$$B(S) = \pm \{1, \alpha\}T = \{1, \alpha\}B(R) \cup \pm \{1, \alpha\}H.$$ 

This expresses the group $B(S)$ as the union of two subgroups, hence either:

(i) $\{1, \alpha\}B(R) \subset \pm \{1, \alpha\}H$, or
(ii) $\pm \{1, \alpha\}H \subset \{1, \alpha\}B(R)$.

In case (i) $B(S) = \pm \{1, \alpha\}H$ and $B(R) \subset \pm \{1, \alpha\}H \cap G(R) = \pm H$. Hence $B(R) = \pm B(H)$. In case (ii) $B(S) = \{1, \alpha\}B(R)$ and $H \subset B(R)$. Then $H = B(H), BT = T$ and by the first statement $B(R) = \pm T$.

Recall that any Witt ring $R$ can be written as $R_0[G(R)/B(R)]$, where $R_0 = W(B(R))$, the Witt ring generated by $B(R)$. See [4, Chapter 5, Section 7] for details.

Proposition 1.5. Let $R = R_0[E_n]$, with $R_0$ basic. Let $S$ be an $H$-extension of $R$. Then:

(a) If $H = 1$ then $S = R_0[E_{n+1}]$.
(b) If $|H| > 1$ and $H \subset G(R_0)$ then $S = S_0[E_n]$, for some Witt ring $S_0$ that is an $H$-extension of $R_0$ (with the same $\alpha$).
(c) If $H \not\subset G(R_0)$ and $-1 \in H$ then $S = (D_1 \cap R_0[H/B(H)]) [G(R)/H]$, and $G(R_0) \subset H$.

(d) If $H \not\subset G(R_0)$ and $-1 \not\in H$ then $S = (\mathbb{Z} \cap R_0[H/B(H)]) [G(R)/\pm H]$ and $G(R_0) \subset \pm H$.

Proof. If $H = 1$ then $S = R[E_1]$ by (1.1), which gives (a). So assume $|H| > 1$. Further suppose that $H \subset G(R_0)$ so that $B(H) = H \cap B(R) = H \cap G(R_0) = H$. Then $-H \subset T \subset G(R_0)$. Hence $\pm H \subset \pm T \subset G(R_0)$. Thus if (1.3)(a) holds, so that $G(R_0) = B(R) = \pm B(H) = \pm H$, then $B(R) = \pm T$ also. So we are always in case (b) of (1.3). Then, since $G(S)/B(S) \cong G(R)/B(R)$, we have $S = S_0[E_n]$, where $S_0 = W(B(S))$. From $H \subset B(R)$ we have that $S_0$ is an $H$-extension of $R_0$.

Next suppose that $H \not\subset G(R_0)$. We still have that $G(R_0) = B(R)$. If $B(R) = \pm T$ then $H \subset B(R)$, contrary to our assumption. Thus we are in Case (a) of (1.3). First say that $-1 \in H$. Note that $-1 \in H \cap B(R) = B(H)$ also. Then $B(R) = B(H)$ and $B(S) = \{1, \alpha\}H$. Thus $S = S_0[G(R)/H]$, for $S_0 = W(\{1, \alpha\}H)$, since $G(S)/B(S) \cong G(R)/H$.

Now $D_{S_0}(1, -\alpha) = D_S(1, -\alpha) \cap \{1, \alpha\}H = \{1, \alpha\}H$. Hence $\alpha \in \text{rad}(S_0)$. Write $S_0 = D_1 \cap S_1$, for some Witt ring $S_1$ and with $D_1$, generated by $\alpha$, being $\mathbb{Z}_2[E_1]$ or $\mathbb{Z}_4$ using [4, p. 104]. Note that $S = W(H)$.

If $h \in H \setminus B(H)$ then $D_R(1, -h) = \{1, -h\}$ and $D_S(1, -h) = \{1, \alpha, -\alpha h\}$ so that $D_{S_1}(1, -h) = D_S(1, -h) \cap H = \{1, -h\}$. Similarly, $D_{S_1}(h) = \{1, h\}$. And if $h \in B(H)$ then $D_{S_1}(1, -h) = \{1, \alpha\}D_R(1, -h) \cap H = D_R(1, -h) \cap H = D_R(1, -h) \cap B(H) = D_{R_0}(1, -h)$. Thus $S_1 = R_0[H/B(H)]$.

We still suppose $H \not\subset G(R_0)$, so that we are in Case (a) of (1.3), and now say that $-1 \not\in H$. Then $S = S_0[G(R)/\pm H]$ as $G(S)/B(S) = \{1, \alpha\}G(R)/\pm \{1, \alpha\}H \cong G(R)/\pm H$. Here $S_0 = W(\pm\{1, \alpha\}H)$

Now $D_{S_0}(1, -\alpha) = \{1, \alpha\}H$ has index two in $G(S_0) = \pm \{1, \alpha\}H$. Further $\alpha \not\in D_S(1, -\alpha)$ else $-1 \in \{1, -\alpha\}H$ and $-1 \in H$. Thus we have an orthogonal decomposition in the sense of [1]:

$$G(S_0) = \{1, \alpha\} \perp D_S(1, -\alpha).$$

Set $S_1 = W(\{1, \alpha\})$ and $S_2 = W(D_S(1, -\alpha))$. Note that $S_1 = \mathbb{Z}$ as $-1 \not\in D_S(1, -\alpha)$. If $h \in \pm H \setminus \pm B(H)$ then $D_R(1, \pm h) = \{1, \pm h\}$, $D_R(1, h) = \{1, \alpha, -\alpha h\}$ and $D_S(1, h) = \{1, h\}$. Thus we have $D_{S_1}(\pm h) = D_S(\pm h) \cap H = \{1, \pm h\}$. So $S_2 = S_3[H/B(H)]$, for some Witt ring $S_3$. $S_2$ is indeed a group ring as $H \not\subset G(R_0)$ implies $H \not\subset B(H)$.

We wish to apply [1, 3.4] and deduce that the decomposition (1.6) yields a product of Witt rings. First we need to handle the case where $S_2$ is decomposable, that is, $S_2 = \mathbb{Z}[E_1]$. In this case $|G(S_2)| = |D_S(1, -\alpha)| = 4$ so that $H = \{1, t\}$, for some $t \not\in G(R_0)$ and $G(R_0) = \{\pm 1\}$ as $G(R_0) \subset \pm H$. Now $D_S(1, -t) = \{1, \alpha, -t, -\alpha t\}$ so we consider instead the orthogonal decomposition:

$$\{1, t\} \perp D_S(1, -t).$$
Then $Q_S(\{1,t\}) = \{1,q(t,-1)\}$ and $Q_S(D_S(1,-t)) = \{1,q(\alpha,-1), q(-t,-1), q(-\alpha t,-1)\}$, using $q(\alpha,-t) = q(\alpha,-1)$ and $q(t,-\alpha t) = 1$. Then $Q_S(\{1,t\}) \cap Q_S(D_S(1,-t)) = 1$. Thus:

$$S_0 = W(\{1,t\}) \cap W(D_S(1,-t)),$$

Now $W(\{1,t\}) = \mathbb{Z}$ as $-1 \notin D_S(1,-t)$ and $W(D_S(1,-t)) = R_0[\{1,\alpha\}]$ since $\alpha$ is birigid in $q_{fst}$ and $G(R_0) = \{\pm 1\}$. This gives the desired result of (d).

We now apply [1, 3.4] and obtain that either (1.6) yields a product or $Q_S(\{1,\alpha\}) = Q_S(D_S(1,-\alpha))$. But if $t \in H \setminus G(R_0)$ then $q(t,t) = q(t,-1)$ and $q(\alpha,\alpha) = q(\alpha,-1)$ are distinct, since $-1 \notin D_S(1,-\alpha t)$. Hence the decomposition (1.6) in fact yields the product $S_0 = S_1 \cap S_2$. We have already seen that $S_1 = \mathbb{Z}$ and that $S_2 = S_3[H/B(H)]$. If $h \in \pm B(H)$ then

$$D_{S_3}(1,-h) = D_{S_1}(1,-h) = D_R(1,-h) = D_{R_0}(1,-h).$$

So $S_3 = R_0$. \hfill \Box

$H$-extensions are motivated by the behavior of odd degree field extensions. (1.5) and other lemmas do mimic the results expected in the field case, at least for valued fields. Pointing out these parallels is speculation (there may be no field extensions yielding an $H$-extension) but it is instructive.

Suppose then that $K/F$ is an odd degree extension with $WK$ an $H$-extension of $WF$. Further suppose that $K$ has a 2-henselian valuation $B$ with a basic residue field $k_B$. If $\alpha$ is not a unit, modulo squares, then $\alpha$ is birigid and we have case (a) of (1.5). Otherwise, $\alpha$ pushes down to $\overline{\alpha} \in k_B$. Set $A = B \cap F$. If $A$ has a basic residue field $k_A$ then $Wk_A \subset Wk_B$, $\overline{A} = D(1,-\overline{\alpha}) \cap k_B$ and $Wk_B$ is an $\overline{A}$-extension of $Wk_A$. This is (1.5)(b).

If $k_A$ is not basic then $H$ contains birigid elements of $F$ and so $B(k_A) \subset \pm H$. This yields cases (c) and (d) of (1.5).

2. Local type rings.

Notation. For a subset $A \subset G(R)$ set:

$$C_R(A) = \bigcap_{a \in A} D_R(1,-a).$$

Lemma 2.1. Suppose that $S$ is an $H$-extension of $R$. Suppose $k \in C_R(H) \setminus H$. Then $Q_R(H) \cap Q_R(k) = 1$.

Proof. Let $\rho \in Q_R(H) \cap Q_R(k)$ so that $\rho = q(h,x) = q(k,y)$ with $h \in H$, and $x,y \in G(R)$. Since $H \subset D_S(1,-\alpha)$ we have that $q(k,y) = q(\alpha x,h)$. By linkage there exists a $t \in G(S)$ such that:

$$q(k,y) = q(k,t) = q(\alpha x,t) = q(\alpha x,h).$$

The first equality gives $ty \in D_S(1,-k) = D_R(1,-k) \subset G(R)$, since $k \notin H$. Hence $t \in G(R)$. The second equality gives:

$$t \in D_S(1,-\alpha x k) \cap G(R) = D_R(1,-x k) \cap H.$$
This implies $t \in D_R(1, -x)$ since $H \subset D_R(1, -k)$. The third equality gives:

$$ht \in D_S(1, -ax) \cap G(R) = D_R(1, -x) \cap H.$$ 

Hence $h \in D_R(1, -x)$ and $\rho = q(x, h) = 1$. □

The small Witt rings of local type will often be treated separately. The only local type Witt ring with two generators is $\mathbb{Z}$. There are two Witt rings of local type on four generators and both are group rings. If $L$ is local type and $|G(L)| \geq 8$ then $L$ is not a group ring. See [4, Chapter 5, Section 3] for details.

**Lemma 2.2.** Suppose $R = L \cap R_2$, with $L$ a Witt ring of local type and $|G(L)| \geq 8$. Let $\pi_1$ be the projection map of $G(R)$ onto $G(L)$. Let $S$ be an $H$-extension of $R$. Then $\pi_1(H) = 1$ or $G(L)$.

**Proof.** Set $B = \pi_1(H)$ and write $Q(L) = \{1, \rho\}$. Suppose that $B \neq 1$. If $(u, v) \in H$ with $u \neq 1$ then pick $r \in G(L) \setminus D_L(1, -u)$. We get $q((u, v), (r, 1)) = (\rho, 1)$ and so $(\rho, 1) \in Q(H)$.

Now $H \subset B \times G(R_2)$ so that $C_L(B) \times 1 = C_R(B \times G(R_2)) \subset C_R(H)$. If $C_L(B) = 1$ then $B = G(L)$ and we are done. Suppose there exists $1 \neq z \in C_L(B)$. Set $h = (z, 1) \in C_R(H)$. Then $Q(h) = \{(\rho, 1)\} \subset Q(H)$. By (2.1) we must have $h \in H$ and so $C_L(B) \times 1 \subset H$. In particular, $C_L(B) \subset B$.

Continue to let $h = (z, 1)$ where $1 \neq z \in C_L(B)$. We claim there exist an $h_1 \in H$ and an $x \in D_R(1, -h)$ such that $q(x, h_1) = (\rho, 1)$. Suppose not. We consider any $x = (a, 1)$ with $a \in D_L(1, -z)$. Then $q(x, (u, v)) = (\rho, 1)$ unless $a \in D_L(1, -u)$. Thus $D_L(1, -z) \subset D_L(1, -u)$ for all $(u, v) \in H$, that is, $D_L(1, -z) \subset C_L(B)$. Then $C_L(B) = D_L(1, -z)$ and so $B = \{1, z\}$. But $C_L(B) \subset B$ so that $D_L(1, -z) \subset \{1, z\}$ and $|G(L)| = 4$, a case we are excluding.

Thus there does exist an $h_1 \in H$ and an $x \in D_R(1, -h)$ such that $q(x, h_1) = (\rho, 1)$. Then in $S$ we have $q(\alpha x, h_1) = (\rho, 1)$ and so $Q_S(h) \subset Q_S(\alpha x)$. We obtain:

$$|Q_S(\alpha x) \cap Q_S(h)| = 2.$$

On the other hand:

$$D_S(1, -ax) \cap D_S(1, -h) = \{1, -ax\} \cap \{1, \alpha\} \cap D_R(1, -h).$$

Here $D_R(1, -x) \cap H \subset D_R(1, -h)$ as $h \in C_R(H)$. Also, by construction, $x$ is an element of $D_R(1, -h)$ and $h \in C_R(H) \subset D_R(1, -h)$. So $-1, x \in D_R(1, -h)$ and $-ax \in \alpha D_R(1, -h)$. Thus $D_S(1, -ax) \subset D_S(1, -h)$. By [4, 5.2]:

$$|Q_S(\alpha x) \cap Q_S(h)| = \frac{|D_S(1, -ax)|}{|D_S(1, -ax) \cap D_S(1, -h)|} = \frac{2|D_R(1, -x) \cap H|}{2|D_R(1, -x) \cap H|},$$

and $D_R(1, -x) \cap H = D_R(1, -x) \cap H$ as $h \in C_R(H)$ implies $H \subset D_R(1, -h)$. So $|Q_S(\alpha x) \cap Q_S(h)| = 1$, a contradiction. Hence $\pi_1(H) = G(L)$.
**Corollary 2.3.** Suppose $R$ is of local type with $|G(R)| \neq 4$. If $S$ is an $H$-extension of $R$ then either:

(a) $H = 1$ and $S = R[E_1]$, with $E_1$ generated by $\alpha$, or

(b) $H = G(R)$ and $S = D_1 \cap R$, with $D_1$ generated by $\alpha$.

**Proof.** If $|G(R)| = 2$ then it is clear that $H = 1$ or $H = G(R)$. If $|G(R)| \geq 8$ then either:


We start with a lemma that may be of general interest.

**Lemma 3.1.** Let $R$ be a Witt ring of elementary type. Let $K$ be a proper subgroup of $G(R)$ and let $y \in G(R)$. If

$$G(R) = \bigcup_{k \in K} D\langle 1, -yk \rangle,$$

then $y \in \text{rad}(R) \cdot K$.

**Proof.** We first prove the result for non-degenerate $R$ where we must show $y \in K$. It suffices to prove this for subgroups $K$ of index two. Namely, if $K_0$ is any subgroup satisfying the hypothesis let $A$ denote the set of subgroups $K$ of index two that contain $K_0$. Then for any $K \in A$:

$$G(R) = \bigcup_{k \in K_0} D\langle 1, -yk \rangle \subset \bigcup_{k \in K} D\langle 1, -yk \rangle.$$

Assuming the result holds for subgroups of index two, we obtain $y \in K$. Then $y \in \bigcap_{K \in A} K = K_0$, as desired.

So suppose $[G(R) : K] = 2$. We work by induction on $|G(R)|$. We need to prove that $y \in K$ when $R$ is of local type, a group ring or a product. First suppose $R$ is of local type. Then $K = D\langle 1, -a \rangle$, for some $a \in G(R)$. We have:

$$G(R) = \bigcup_{k \in D\langle 1, -a \rangle} D\langle 1, -yk \rangle = D\langle 1, -y, ay \rangle.$$

Multiplying by $-a$ gives $G(R) = D\langle \langle -a, -y \rangle \rangle'$, the pure part of the Pfister form $\langle \langle -a, -y \rangle \rangle$. In particular, $-1 \in D\langle \langle -a, -y \rangle \rangle'$ so $\langle \langle -a, -y \rangle \rangle = 0$ and $y \in D\langle 1, -a \rangle = K$.

Next let $R = R_0[E_1]$, with $E_1 = \{1, t\}$ and $|G(R_0)| \geq 2$ (if $G(R_0) = 1$ then $R$ is degenerate). Suppose $y \notin K$. We claim $G(R_0) \subset K$. Choose any $g \in G(R_0)$. Then there exist $k_1, k_2 \in K$ with $-gt \in D\langle 1, -yk_1 \rangle$ and $-t \in D\langle 1, -yk_2 \rangle$. We get $gt = yk_1$ and $t = yk_2$ since $y \notin K$. Hence $g = k_1k_2 \in K$. This proves the claim. Both $G(R_0)$ and $K$ have index two so $K = G(R_0)$. From $t = yk_2$ we have $y \in tG(R_0)$. Pick $g \in G(R_0)$. Then $g \in D\langle 1, -yk_3 \rangle$ for some $k_3 \in K$. But $yk_3 \in yK = tG(R_0)$, so this is impossible. The contradiction implies $y \in K$. 

□
Lastly, say \( R = R_1 \cap R_2 \). Write \( y = (y_1, y_2) \). Now \( K \cap (G(R_1) \times 1) \) is a subgroup of index at most two in \( G(R_1) \times 1 \). Let \( K_1 \) be its projection into \( G(R_1) \). Similarly, let \( K_2 \) be the projection of \( K \cap (1 \times G(R_2)) \) into \( G(R_2) \). Then \( [G(R_i) : K_i] \leq 2 \), for \( i = 1, 2 \). If \( K_2 = G(R_2) \) then:

\[
\bigcup_{k \in K} D_R(1, -yk) = \bigcup_{k_1 \in K_1} D(1, -y_1 k_1) \times G(R_2)
\]

so that \( G(R_1) = \bigcup_{K_1} D(1, -y_1 k_1) \). By induction then \( y_1 \in K_1 \) and hence \( y = (y_1, y_2) \in K_1 \times G(R_2) = K \). In the same way, if \( K_1 = G(R_1) \) then \( y \in K \) as desired. So we may assume that \( [G(R_i) : K_i] = 2 \) for \( i = 1, 2 \). Write \( K = \{1, \gamma\} K_1 \times K_2 \). We have:

\[
G(R) = \bigcup_{k_1 \in K_1, k_2 \in K_2} (D(1, -y_1 k_1) \times D(1, -y_2 k_2))
\]

\[
\cup D(1, -y_1 \gamma k_1) \times D(1, -y_2 \gamma k_2)，
\]

where \( \gamma = (\gamma_1, \gamma_2) \).

Suppose \( y_1 \in K_1 \). If \( G(R_1) = \cup D(1, -y_1 \gamma_1 k_1) \) then by induction we have \( y_1 \gamma_1 \in K_1 \) and so \( \gamma_1 \in K_1 \). Then \( K = K_1 \times G(R_2) \) and \( K_2 = G(R_2) \) a case we have already covered. We may thus assume there exists a \( g_1 \in G(R_1) \setminus \cup D(1, -y_1 \gamma_1 k_1) \). Then \( g_1 \times G(R_2) \subset \cup (D(1, -y_1 k_1) \times D(1, -y_2 k_2)) \) and so \( G(R_2) = \cup D(1, -y_2 k_2) \). By induction \( y_2 \in K_2 \) and \( y = (y_1, y_2) \in K_1 \times K_2 \subset K \), and we are done.

We may thus assume \( y_1 \notin K_1 \). Similarly, \( y_2 \notin K_2 \). Pick, for \( i = 1, 2 \), a \( g_i \in G(R_i) \setminus \cup D(1, -y_i k_i) \), which is possible by induction. Then \( g_1 \times G(R_2) \subset \cup (D(1, -y_1 \gamma_1 k_1) \times D(1, -y_2 \gamma_2 k_2)) \) and so \( G(R_2) = \cup D(1, -y_2 \gamma_2 k_2) \). By induction once more, we have \( y_2 \gamma_2 \in K_2 \). Similarly, \( y_1 \gamma_1 \in K_1 \). Then \( y \in (\gamma_1, \gamma_2)(K_1 \times K_2) \subset K \) as desired. This proves the result for non-degenerate \( R \).

Now suppose \( R \) is degenerate. Write \( R = D \cap R_2 \), with \( \text{rad}(R) = G(D) \times 1 \) and \( R_2 \) non-degenerate. Let \( \pi_2 \) be the projection of \( G(R) \) onto \( G(R_2) \). Set \( K_2 = \pi_2(K) \) and write \( y = (y_1, y_2) \), with \( y_1 \in G(D) \) and \( y_2 \in G(R_2) \). Our assumption is:

\[
G(R) = G(D) \times G(R_2) = \bigcup_{(k_1, k_2) \in K} D((1, 1), -(y_1 k_1, y_2 k_2))
\]

\[
= G(D) \times \left( \bigcup_{k_2 \in K_2} D_{R_2}(1, -y_2 k_2) \right).
\]

From the non-degenerate case we get \( y_2 \in K_2 = \pi_2(K) \). Thus there exists a \( d \in G(D) \) such that \( (d, y_2) \in K \). Hence \( y = (y_1, y_2) = (dy_1, 1)(d, y_2) \in \text{rad}(R) \cdot K \). □

Our key reduction lemma follows.
Lemma 3.2. Let \( R = R_1 \cap R_2 \) and suppose \( S \) is an \( H \)-extension of \( R \). If \( H = H_1 \times G(R_2) \) then there exists a Witt ring \( T \) that is an \( H_1 \)-extension of \( R_1 \) such that \( S \cong T \cap R_2 \).

Proof. We first construct \( T \). Let \( G(T) \) be a group containing \( G(R_1) \) as a subgroup of index 2; write \( G(T) = \{1, \beta\}G(R_1) \). Let \( \varphi : G(T) \rightarrow \{1, \alpha\}(G(R_1) \times 1) \) be the isomorphism sending \( g_1 \mapsto (g_1, 1) \) and \( \beta g_1 \mapsto \alpha(g_1, 1) \), where \( g_1 \in G(R_1) \). For \( z \in G(T) \) define:

\[
D_T(1, -z) = \varphi^{-1}(D_S(1, -\varphi(z)) \cap \{1, \alpha\}(G(R_1) \times 1)).
\]

We check that \( T \) is an \( H_1 \)-extension of \( R_1 \). If \( z \in G(R_1) \setminus H_1 \) then:

\[
D_T(1, -z) = \varphi^{-1}((D_{R_1}(1, -z) \times G(R_2)) \cap \{1, \alpha\}(G(R_1) \times 1))
= \varphi^{-1}(D_{R_1}(1, -z) \times 1)
= D_{R_1}(1, -z).
\]

If \( z \in H_1 \) then:

\[
D_T(1, -z) = \varphi^{-1}(\{1, \alpha\}(D_{R_1}(1, -z) \times G(R_2)) \cap \{1, \alpha\}(G(R_1) \times 1))
= \varphi^{-1}(\{1, \alpha\}(D_{R_1}(1, -z) \times 1))
= \{1, \beta\}D_{R_1}(1, -z).
\]

Lastly, if \( z \in G(R_1) \) then:

\[
D_T(1, -\beta z) = \varphi^{-1}(D_S(1, -\alpha(z, 1)) \cap \{1, \alpha\}(G(R_1) \times 1)).
\]

Now:

\[
D_S(1, -\alpha(z, 1)) = \{1, -\alpha(z, 1)\}(D_{R_1}(1, -z) \times G(R_2)) \cap (H_1 \times G(R_2))
= \{1, -\alpha(z, 1)\}(D_{R_1}(1, -z) \cap H_1 \times G(R_2)).
\]

Since \( (1, -1) \in D_{R_1}(1, -z) \cap H_1 \times G(R_2) \) we have:

\[
D_S(1, -\alpha(z, 1)) = \{1, \alpha(-z, 1)\}(D_{R_1}(1, -z) \cap H_1 \times G(R_2)).
\]

Hence:

\[
D_T(1, -\beta z) = \varphi^{-1}(\{1, \alpha(-z, 1)\}(D_{R_1}(1, -z) \cap H_1 \times G(R_2))
\cap \{1, \alpha(-z, 1)\}(G(R_1) \times 1)
= \varphi^{-1}(\{1, \alpha(-z, 1)\}(D_{R_1}(1, -z) \cap H_1 \times 1)
= \{1, -\beta z\}(D_{R_1}(1, -z) \cap H_1).
\]

We begin the verification that \( (G(T), D_T) \) is linked, so that \( T \) is indeed a Witt ring. Let \( t = \alpha^{\epsilon_1}(u, v) \in G(S) \) and let \( \beta^{\epsilon_2}x, \beta^{\epsilon_3}y \in G(T) \) with each \( \epsilon_i = 0 \) or 1.

Claim. If \( t \in \varphi(\beta^{\epsilon_3}y)D_S(1, -\varphi(\beta^{\epsilon_2}x)) \) then \( \beta^{\epsilon_1}u \in \beta^{\epsilon_3}yD_T(1, -\beta^{\epsilon_2}x) \).

We first assume that \( \epsilon_1 = 0 \). We have four cases:
Case 1. $\epsilon_2 = 0, \epsilon_3 = 0$. Here $(uy, v) \in D_S\langle 1, -(x, 1) \rangle$, hence $uy \in D_{R_1}\langle 1, -x \rangle \subset D_T\langle 1, -x \rangle$.

Case 2. $\epsilon_2 = 0, \epsilon_3 = 1$. Here $\alpha(uy, v) \in D_S\langle 1, -(x, 1) \rangle$. We must have that $x \in H_1$ so $\alpha(uy, v) \in \{1, \alpha\}(D_{R_1}\langle 1, -x \rangle \times G(R_2))$. Then $uy \in D_{R_1}\langle 1, -x \rangle$. Thus $\beta uy \in D_T\langle 1, -x \rangle = \{1, \beta\}D_{R_1}\langle 1, -x \rangle$.

Case 3. $\epsilon_2 = 1, \epsilon_3 = 0$. Here:

$$
(uy, v) \in \{1, -\alpha(x, 1)\}(D_{R_1}\langle 1, -x \rangle \times G(R_2)) \cap H
= \{1, -\alpha(x, 1)\}(D_{R_1}\langle 1, -x \rangle \cap H_1) \times G(R_2)).
$$

Thus $uy \in D_{R_1}\langle 1, -x \rangle \cap H_1$. We obtain

$$
uy \in D_T\langle 1, -\beta x \rangle = \{1, -\beta x\}(D_{R_1}\langle 1, -x \rangle \cap H_1).
$$

Case 4. $\epsilon_2 = 1, \epsilon_3 = 1$. Here $\alpha(uy, v) \in \{1, -\alpha(x, 1)\}(D_{R_1}\langle 1, -x \rangle \cap H_1) \times G(R_2))$ so that $-(xuy, v) \in (D_{R_1}\langle 1, -x \rangle \cap H_1) \times G(R_2)$. Thus $-xuy \in D_{R_1}\langle 1, -x \rangle \cap H_1$ and $\beta uy \in D_T\langle 1, -\beta x \rangle = \{1, -\beta x\}(D_{R_1}\langle 1, -x \rangle \cap H_1)$.

The four cases with $\epsilon_1 = 1$ are identical to the four above cases. For example, if $\epsilon_1 = 1, \epsilon_2 = 0, \epsilon_3 = 0$ then we have $\alpha(uy, v) \in D_S\langle 1, -(x, 1) \rangle$, which is Case 2 above. Thus the Claim is proven.

We can now check linkage in $T$. Let $x, y, z, w \in G(T)$ and suppose:

$$
xD_T\langle 1, -y \rangle \cap D_T\langle 1, -yz \rangle \cap wD_T\langle 1, -z \rangle \neq \emptyset.
$$

Apply $\varphi$ to get:

$$
\varphi(x)D_S\langle 1, -\varphi(y) \rangle \cap D_S\langle 1, -\varphi(yz) \rangle \cap \varphi(w)D_S\langle 1, -\varphi(z) \rangle \neq \emptyset.
$$

By linkage on $S$, there exists a $t \in G(S)$ in:

$$
\varphi(y)D_S\langle 1, -\varphi(x) \rangle \cap D_S\langle 1, -\varphi(xw) \rangle \cap \varphi(z)D_S\langle 1, -\varphi(w) \rangle.
$$

Now apply the

Claim.

$$
yD_T\langle 1, -x \rangle \cap D_T\langle 1, -xw \rangle \cap zD_T\langle 1, -w \rangle \neq \emptyset,
$$

as desired.

Lastly, set $W = T \cap R_2$. Then $G(W) = (\{1, \beta\}G(R_1)) \times G(R_2)$. Set $\gamma = (\beta, 1)$ so that $G(W) = \{1, \gamma\}(G(R_1) \times G(R_2)) = \{1, \gamma\}G(R)$. We will show $W$ is an $H$-extension of $R$, via $\gamma$, and hence that $S \cong W$.

First let $h = (h_1, g_2) \in H$, where $h_1 \in H_1 \subset G(R_1)$ and $g_2 \in G(R_2)$. Then:

$$
D_W\langle 1, -h \rangle = D_T\langle 1, -h_1 \rangle \times D_{R_2}\langle 1, -g_2 \rangle
= ((1, \beta)D_{R_1}\langle 1, -h_1 \rangle) \times D_{R_2}\langle 1, -g_2 \rangle
= \{1, \gamma\}D_R\langle 1, -h \rangle.
$$
Next let \( g = (g_1, g_2) \in G(R) \setminus H \), with \( g_1 \in G(R_1) \setminus H_1 \) and \( g_2 \in G(R_2) \). Then:

\[
D_W(1, -g) = DT(1, -g_1) \times DR_2(1, -g_2) = DR_1(1, -g_1) \times DR_2(1, -g_2) = DR(1, -g).
\]

Lastly, let \( g = (g_1, g_2) \in G(R) \), with \( g_1 \in G(R_1) \) and \( g_2 \in G(R_2) \). Then:

\[
D_W(1, -\gamma g) = D_W(1, -(\beta g_1, g_2)) = DR_1(1, -\beta g_1) \times DR_2(1, -g_2) = DR_1(1, -g_1) \cap H_1 \times DR_2(1, -g_2)
\]

Using the fact that if \( u,v \in H \) then \( R_u \times R_v \subseteq H \), we have:

\[
DR(1, -\gamma g) \cap H = \{1, -\gamma (g_1, g_2)\} \cap H = \{1, -\gamma (g_1, g_2)\} \cap H.
\]

Thus \( W = T \cap R_2 \) is an \( H \)-extension of \( R \) and so is isomorphic to \( S \). \( \square \)

Our last general lemma is the most technical, but it also does most of the work.

**Lemma 3.3.** Let \( u \in G(R) \) and \( h \in H \). Then:

\[
DR(1, uh, -h) \setminus uH \subset \bigcup_{t \in h(D(1, -u) \cap H)} DR(1, -t).
\]

**Proof.** Let \( w \in DR(1, uh, -h) \setminus uH \). Then \( w \in DR(1, -hv) \) for some \( v \in DR(1, -u) \) and \( uv \notin H \). We have:

\[
q(\alpha u, h) = q(u, h) = q(u, vh) = q(uv, vh).
\]

Thus, by linkage, there exists \( t \in G(S) \) such that:

\[
q(\alpha u, h) = q(\alpha v, t) = q(uv, t) = q(uv, vh).
\]

The third equality gives \( vht \in DS(1, -uv) \). Since \( uv \notin H \) this implies \( t \in G(R) \). Then the first two equalities give:

\[
h \in DS(1, -\alpha u) \cap G(R) = DR(1, -u) \cap H,
\]

\[
t \in DS(1, -\alpha w) \cap G(R) = DR(1, -w) \cap H.
\]

Hence \( w \in DR(1, -t) \) where \( t \in h(DR(1, -u) \cap H) \). \( \square \)

### 4. Products: Degenerate Witt rings.

If \( R \) is a degenerate Witt ring then \( R = DR \cap R_2 \), for some Witt ring \( R_2 \) and where \( G(D) = \{1, d\} \), with \( D_1(1, 1) = D_1(1, d) = \{1, d\} \). We will often use the fact that if \( (u, v) \in G(R) \) then \( DR(1, -(u, v)) = DR(1, -(du, v)) \).
Lemma 4.1. Suppose $R = D \cap R_2$ is degenerate. Let $\pi_1$ be the projection of $G(R)$ onto $G(D)$. Let $S$ be an $H$-extension of $R$. Then either $G(D) \times 1 \subset H$ or $S$ is isomorphic to an $H_0$-extension of $R$, for some subgroup $H_0 \subset G(R)$ with $\pi_1(H_0) = 1$.

Proof. Suppose $(d,1) \notin H$ and that $\pi_1(H) \neq 1$, so that $(d,y) \in H$ for some $y \in G(R_2)$. Let $G$ be the subgroup of $G(R_2)$ such that $1 \times G_2 = H \cap (1 \times G(R_2))$. Then $H = (1 \times G_2) \cup (d \times y G_2)$. Set $H_0 = 1 \times \{1,y\}G_2$, and note that $\pi_1(H_0) = 1$.

Let $\beta^2 = 1$ and set $G(S_0) = \{1, \beta\}G(R_2)$. Define $S_0$-value set $s$ so that $S_0$ is an $H_0$-extension of $R$. We wish to show $S \cong S_0$. Extend $G_2$ to a subgroup $K$ of index two in $G(R_2)$ that does not contain $y$. Define $\varphi : G(S) \to G(S_0)$ by $\alpha \mapsto \beta$ and for $(u,v) \in G(R)$:

$$\varphi(u,v) = \begin{cases} (u,v), & \text{if } v \in K \\ (du,v), & \text{if } v \notin K. \end{cases}$$

It is quickly checked that $\varphi$ is an isomorphism. We will show:

$$\varphi(D_S(1,{-s})) = D_{S_0}(1,{-\varphi(s)}),$$

for all $s \in G(S)$. This shows both that $S_0$ is a Witt ring and that $S \cong S_0$.

Claim. If $(u,v) \in G(R)$ then $\varphi(D_R(1,{-}(u,v))) = D_R(1,{-}(u,v))$.

$$D_R(1,{-}(u,v)) = \{1,d\} \times D_{R_2}(1,{-}v)$$

$$= 1 \times (D_{R_2}(1,{-}v) \cap K) \cup d \times (D_{R_2}(1,{-}v) \cap K)$$

$$\cup 1 \times (D_{R_2}(1,{-}v) \cap y K) \cup d \times (D_{R_2}(1,{-}v) \cap y K).$$

Thus:

$$\varphi(D_R(1,{-}(u,v))) = 1 \times (D_{R_2}(1,{-}v) \cap K) \cup d \times (D_{R_2}(1,{-}v) \cap K)$$

$$\cup 1 \times (D_{R_2}(1,{-}v) \cap y K) \cup 1 \times (D_{R_2}(1,{-}v) \cap y K)$$

$$= D_R(1,{-}(u,v)),$$

proving the Claim.

We now check (4.2) in various cases. First suppose $s = (u,v) \in G(R)$, with $v \in K$. Then $s \in H$ iff $u = 1$ and $v \in G_2$. We have $\varphi(s) = s$ and $\varphi(s) \in H_0$ iff $u = 1$ and $v \in G_2$ iff $s \in H$. $D_S(1,{-}s) = \{1,\alpha\}D_R(1,{-}s)$ or $D_R(1,{-}s)$ depending on whether or not $s \in H$. So by the Claim, $\varphi(D_S(1,{-}s)) = \{1,\beta\}D_R(1,{-}s)$ or $D_R(1,{-}s)$ depending on whether or not $\varphi(s) \in H_0$. Thus $\varphi(D_S(1,{-}s)) = D_{S_0}(1,{-}\varphi(s))$.

Next suppose that $s = (u,v) \in G(R)$ with $v \in yK$. Then $s \in H$ iff $u = d$ and $v \in yG_2$. We have $\varphi(s) = (du,v)$ so that $\varphi(s) \in H_0$ iff $u = d$ and $v \in yG_2$ iff $s \in H$. Again using the Claim:

$$D_S(1,{-}s) = \begin{cases} \{1,\alpha\}D_R(1,{-}s), & \text{if } s \in H \\ D_R(1,{-}s), & \text{if } s \notin H. \end{cases}$$
Thus:

\[
\varphi(D_S(1, -s)) = \begin{cases} 
1, \beta \} D_R(1, -s), & \text{if } \varphi(s) \in H_0 \\
D_R(1, -s) & \text{if } \varphi(s) \notin H_0.
\end{cases}
\]

Now \(D_R(1, -s) = D_R(1, -(u, v)) = D_R(1, -(du, v)) = D_R(1, -\varphi(s))\). Hence we have as desired that \(\varphi(D_S(1, -s)) = D_{S_0}(1, -\varphi(s))\).

Now suppose \(s = \alpha(u, v) \in \alpha G(R)\). Then:

\[
D_S(1, -s) = \{1, -\alpha(u, v)\}(D_R(1, -(u, v)) \cap H)
= \{1, -\alpha(u, v)\}[(\{1, d\} \times D_{R_2}(1, -v)) \cap (1 \times G_2 \cup d \times yG_2)]
= \{1, -\alpha(u, v)\}[1 \times (D_{R_2}(1, -v) \cap G_2) \cup d \times (D_{R_2}(1, -v) \cap yG_2)].
\]

Now:

\[
\varphi((1 \times (D_{R_2}(1, -v) \cap G_2)) \cup (d \times (D_{R_2}(1, -v) \cap yG_2)))
= 1 \times (D_{R_2}(1, -v) \cap G_2) \cup 1 \times (D_{R_2}(1, -v) \cap yG_2)
= 1 \times (D_{R_2}(1, -v) \cap \{1, y\}G_2)
= D_R(1, -(u, v)) \cap H_0.
\]

Thus if \(v \in K\) then:

\[
\varphi(D_S(1, -s)) = \varphi(D_S(1, -\alpha(u, v)))
= \{1, -\beta(u, v)\}(D_R(1, -(u, v)) \cap H_0) = D_{S_0}(1, -\varphi(s)),
\]

verifying (4.2) in this case.

Lastly, if \(v \in G(K)\) then:

\[
\varphi(D_S(1, -s)) = \varphi(D_S(1, -\alpha(u, v)))
= \{1, -\beta(du, v)\}(D_R(1, -(u, v)) \cap H_0)
= \{1, -\beta(du, v)\}(D_R(1, -(du, v)) \cap H_0)
= D_{S_0}(1, -\varphi(s)).
\]

Thus (4.2) holds in all cases. \(\square\)

### 5. Products: Local type factors.

Lemma (3.3) looks simpler when one factor has local type.

**Lemma 5.1.** Suppose \(R = L \cap R_2\), with \(L\) of local type. Suppose \(S\) is an \(H\)-extension of \(R\). Let \(h = (h_1, h_2) \in H\) and \(u = (u_1, u_2) \in G(R)\) such that \(u_1 \notin D_L(1, -h_1)\) while \(u_2 \in D_{R_2}(1, -h_2)\). Then:

\[
G(R) = uH \cup (u_1 \times G(R_2)) \cup \bigcup_{t \in h(D_R(1, -(u_1) \cap H)} D_R(1, -t).
\]
Proof. Write $Q(L) = \{1, \rho\}$. Then $\langle \langle -u, -h \rangle \rangle = (\rho, 1)$. Hence $-D_R(\langle -u, -h \rangle)' = \{(x, y) \in G(R) : x \neq 1\}$. Now $-u \cdot \langle -u, -h, uh \rangle \simeq \langle 1, uh, -h \rangle$. Thus $D_R(1, uh, -h) = \{(x, y) \in G(R) : x \neq u_1\}$. Apply (3.3).

Lemma 5.2. Suppose $R = L \cap R_2$, with $L$ of local type. Let $\pi_1$ be the projection of $G(R)$ onto $G(L)$. Let $S$ be an $H$-extension of $R$ and suppose that $\pi_1(H) = G(L)$. Let $u = (u_1, u_2) \in G(R)$.

(a) If $|G(L)| \geq 4$ then $\pi_1(D_R(1, -u) \cap H) = D_L(1, -u_1)$. 
(b) If $L = \mathbb{Z}$ and $\pi_1(D_R(1, -u) \cap H) \neq D_L(1, -u_1)$ for some $u$, then $H = 1 \times H_2 \cup -1 \times -H_2$, where $1 \times H_2 = H \cap (1 \times G(R_2))$ and $H_2$ is an ordering on $R_2$.

Proof. Suppose that $\pi_1(D_R(1, -u) \cap H) = K < D_L(1, -u_1)$.

Claim. If $v \in G(L) \setminus K$ then $(v, -1) \in H$.

Since $\pi_1(H) = G(L)$ there exists a $w \in G(R_2)$ such that $(v, w) = h \in H$. Now:

$$D_R(1, uh, -h) = -u(D_L(\langle -u, -v \rangle)' \times D_{R_2}(\langle -u_2, -w \rangle)')$$

$$\supseteq u_1G(L)' \times u_2T,$$

where $T = -D_{R_2}(\langle -u_2, -w \rangle)'$. Also:

$$\bigcup_{t \in h(D_R(1, -u) \cap H)} D_R(1, -t) \subset \left( \bigcup_{k \in K} D_L(1, -vk) \right) \times G(R_2).$$

Now $v \notin K$ implies $G(L) \neq \cup D_L(1, -vk)$ by (3.1). Choose a $g \in G(L) \setminus \cup D_L(1, -vk)$.

We check that we may assume $g \neq u_1$. If $v \in D_L(1, -u_1)$ then we have $u_1D_L(1, -v) = D_L(1, -v) \subset \cup D_L(1, -vk)$ and so no $g \in G(L) \setminus \cup D_L(1, -vk)$ is equal to $u_1$. If instead $v \notin D_L(1, -u_1)$ then, as $|K| < |D_L(1, -u_1)|$, there exists a $w \neq u_1$ such that $K \subset D_L(1, -u_1) \cap D_L(1, -w)$. If $v \notin D_L(1, -w)$ then $w$ is not in any $D_L(1, -vk)$, for $k \in K$, and we may take $g = w$. If $v \in D_L(1, -w)$ then $v \notin D_L(1, -uw)$ and we may take $g = uw$.

We thus have $g \in u_1G(L) \setminus \cup D_L(1, -vk)$. So $g \times u_2T \subset uH$ by (3.3).

Hence $u_1g \times T \subset H$. Then $(u_1g, u_2), (u_1g, -u_2w) \in H$ and so $(1, -w) \in H$. We obtain that $(v, -1) = (v, w)(1, -w) \in H$ and the Claim is proven.

Now suppose $u_1 \neq 1$. Let $x \in G(L)$. Since $|K| < |D_L(1, -u_1)|$ we have $|K| \leq \frac{1}{2}|G(L)|$. So we can choose $v \in G(L) \setminus \{1, x\}K$. Then $(v, -1)$ and $(vx, -1)$ are in $H$ by the Claim. Hence $(x, 1) \in H$. This shows that $G(L) \times 1 \subset H$. But then $D_L(1, -u_1) \times 1 \subset D_R(1, -u) \cap H$ and $\pi_1(D_R(1, -u) \cap H) = D_L(1, -u_1)$, as desired.

Next suppose $u_1 = 1$ and $|G(L)| \geq 4$. We show $\pi_1(D_R(1, -u) \cap H) = G(L)$.

Pick any $a \in G(L)$ and pick a $b \in G(L)'$ such that $a \in D_L(1, -b)$. This is possible since $|G(L)| \geq 4$ implies there are at least two $b$'s such that $a \in
\(D_L(1, -b)\). So there is such a \(b\) not equal to 1. Then, by the above paragraph, 
\(\pi_1(D_R(1, -(b, u_2)) \cap H) = D_L(1, -b)\) contains \(a\). Thus there exists a \(k \in D_{R_2}^1, -u_2\) such that \((a, k) \in H\). Thus \((a, k) \in D_R(1, -(u_1, u_2))\), as \(u_1 = 1\), and so \(a \in \pi_1(D_R(1, -u) \cap H)\).

Lastly, suppose \(u_1 = 1\) and \(L = \mathbb{Z}\). Here \(K = \{1\}\) and \(v = -1\). The Claim shows that \((-1, -2) \in H\). Now \(H_2 = H \cap (1 \times G(R_2))\) has index 2 in \(H\) since \(1 \times G(R_2)\) has index 2 in \(G(R)\) and \(H \not\subset 1 \times G(R_2)\). Hence \(H = 1 \times H_2 \cup -1 \times -H_2\). The last paragraph of the proof of the Claim gives \(u_1g \times T \subset H\), where \(g \neq u_1\). Thus \(g = -1\) and after multiplying by \(-1 \in H\) we get:

\[
(5.3) \quad 1 \times D_{R_2}^1((-u_2, -w))' \subset H.
\]

This holds for all \(w \in G(R_2)\) such that \((-1, w) \in H\), that is, for all \(w \in -H_2\). Thus for any \(h_2 \in H_2\):

\[
-u_2D_{R_2}^1(1, h_2) \subset D_{R_2}^1(-u_2, h_2, -u_2h_2) \subset H_2
\]

Thus \(H_2\) is a preordering. Also \((5.3)\) holds for any \(u_2 \in G(R_2)\) with 
\(\pi_1(D_R(1, -(1, u_2))) = 1\). That is, \(D_R(1, -(1, u_2)) \cap (-1 \times -H_2) = \emptyset\) or equivalently, \(u_2\) is not in \(D_{R_2}^1(1, h_2)\) for any \(h_2 \in H_2\). For such a \(u_2\), \((5.3)\) implies \(-u_2 \in H_2\). Hence:

\[
G(R_2) = -H_2 \cup \bigcup_{h_2 \in H_2} D_{R_2}^1(1, h_2).
\]

But \(H_2\) is a preordering so that \(\cup D_{R_2}^1(1, h_2) \subset H_2\). Thus \(G(R_2) = -H_2 \cup H_2\), \(H_2\) has index 2 in \(G(R_2)\) and so \(H_2\) is an ordering. \(\square\)

**Notation.** Suppose \(R = R_1 \cap R_2\) and that \(H\) is a subgroup of \(G(R)\). For \(x \in G(R_1)\), set \(F(x) = \{y \in G(R_2) : (x, y) \in H\}\).

**Lemma 5.4.** Let \(R = R_1 \cap R_2\) and let \(S\) be an \(H\)-extension of \(R\). Let \(\pi_1\) be the projection of \(G(R)\) onto \(G(R_1)\) and suppose \(\pi_1(H) = G(R_1)\). Then for all \(x \in G(R_1)\):

(a) \(F(x)\) is non-empty.

(b) \(F(1)\) is a subgroup of \(G(R_1)\).

(c) \(F(x)\) is a coset of \(F(1)\).

**Proof.** No \(F(x)\) is empty since \(\pi_1(H) = G(R_1)\). Clearly \(F(1)\) is a subgroup. Fix \(y_0 \in F(x)\). If \(y \in F(1)\) then \((1, y), (x, y_0) \in H\) implies \((x, y_0) \in H\) and so \(yy_0 \in F(x)\). This says \(y_0F(1) \subset F(x)\).

Now let \(y \in F(x)\). Then \((x, y_0), (x, y) \in H\) so that \((1, y_0) \in H\). Hence \(yy_0 \in F(1)\) and we have the reverse inclusion \(F(x) \subset y_0F(1)\). \(\square\)
Lemma 5.5. Let \( R = R_1 \cap R_2 \) be of elementary type. Let \( S \) be an \( H \)-extension of \( R \). Let \( \pi_1 \) be the projection of \( G(R) \) onto \( G(R_1) \). Suppose the following:

1. \( \pi_1(H) = G(R_1) \).
2. \( F(a) \cap \text{rad}(R_2) \subset \{1\} \), for all \( a \in G(R_1) \).
3. For all \( u = (u_1, u_2) \in G(R) \) we have \( \pi_1(D_R(1, -u) \cap H) = D_{R_1}(1, -u_1) \).

Then \( H = G(R_1) \times H_2 \), for some subgroup \( H_2 \subset G(R_2) \).

Proof. Let \( a \in G(R_1) \). We will first show that:

\[
G(R_2) = \bigcup_{k \in F(1)} D_{R_2}(1, -k).
\]

Pick any \( b \in D_{R_1}(1, -a) \) and any \( g \in G(R_2) \). Then \( a \in D_{R_1}(1, -b) = \pi_1(D_R(1, -(b, g)) \cap H) \) by assumption (3). Hence there exists a \( k \in D_{R_2}(1, -g) \) with \( (a, k) \in H \). That is, \( g \in D_{R_2}(1, -k) \) for some \( k \in F(a) \), proving (5.6).

Write \( F(a) = yF(1) \) as in (5.4). Then (5.6) becomes:

\[
G(R_2) = \bigcup_{k \in F(1)} D_{R_2}(1, -yk).
\]

Thus \( y \in \text{rad}(R_2) \cdot F(1) \) by (3.1). That is, there exists a \( d \in \text{rad}(R_2) \) such that \( d \in yF(1) = F(a) \). By assumption (2) then \( d = 1 \). Hence \( y \in F(1) \) and so \( F(a) = F(1) \). By assumption (1) we have \( (a, m) \in H \) for some \( m \in G(R_2) \). Then \( m \in F(a) = F(1) \) so that \( (1, m) \in H \) also. So \( (a, 1) = (a, m)(1, m) \in H \). Hence \( G(R_1) \times 1 \subset H \) and \( H = G(R_1) \times F(1) \).

We first complete the case of a local factor \( L \) with \( |G(L)| \geq 8 \).

Corollary 5.7. Let \( R = L \cap R_2 \), with \( R_2 \) of elementary type, \( L \) of local type and \( |G(L)| \geq 8 \). Let \( S \) be an \( H \)-extension of \( R \). Suppose \( F(a) \cap \text{rad}(R_2) \subset \{1\} \) for all \( a \in G(L) \). Then either \( H = 1 \times H_2 \) or \( H = G(L) \times H_2 \) for some subgroup \( H_2 \subset G(R_2) \).

Proof. Again let \( \pi_1 \) denote the projection of \( G(R) \) onto \( G(L) \). We know that \( \pi_1(H) = 1 \) or \( G(L) \), by (2.2). If \( \pi_1(H) = 1 \) then clearly \( H = 1 \times H_2 \) for some subgroup \( H_2 \). So suppose that \( \pi_1(H) = G(L) \), the first hypothesis of (5.5). We are assuming the second hypothesis as well. And (5.2) shows the third hypothesis of (5.5) holds. Hence \( H = G(L) \times H_2 \), for some subgroup \( H_2 \).

The argument for \( R = \mathbb{Z} \cap R_2 \) is different.

Lemma 5.8. Let \( R \) be a real Witt ring of elementary type. Let \( P \subset G(R) \) be an ordering. Suppose that for all \( x \in P \) that:

\[
P = \bigcup_{k \in D(1, x) \cap P} D(1, xk).
\]

Then \( R = \mathbb{Z} \cap R_2 \), for some Witt ring \( R_2 \).
Proof. Suppose \( Z \) is not a factor of \( R \). Then \( R \) has a group ring factor that is real. Thus \( R = R_0[E_1] \cap R_2 \), for some Witt rings \( R_0, R_2 \), and we may assume \( P = P_0 \{1, t\} \times G(R_2) \), where \( P_0 \subset G(R_0) \) is an ordering on \( R_0 \) and \( E_1 = \{1, t\} \). Then take \( x = (t, 1) \). We have that \( D(1, -x) = \{1, -t\} \times G(R_2) \) and \( D(1, -x) \cap P = 1 \times G(R_2) \). Thus:

\[
P = \bigcup_{k \in D(1, -x) \cap P} D(1, xk) = \bigcup_{g_2 \in G(R_2)} D(\langle (1, 1), (t, g_2) \rangle)
\]

Hence \( P_0 = 1 \) and \( R_0 = Z \), giving a contradiction. \( \square \)

Lemma 5.9. Let \( R = Z \cap R_2 \) and suppose \( S \) is an \( H \)-extension of \( R \). Then one of the following occurs.

(a) \( H = 1 \times H_2 \) for some subgroup \( H_2 \subset G(R_2) \).
(b) \( R = Z \cap R_3 \), for some Witt ring \( R_3 \), and \( \{\pm 1\} \times 1 \subset H \).
(c) \( R = Z \cap Z \cap R_3 \), for some Witt ring \( R_3 \), and \( (1, 1) \times G(R_3) \subset H \).

Proof. Again let \( \pi_1 \) be the projection of \( G(R) \) onto \( G(Z) = \{\pm 1\} \). If \( \pi_1(H) = 1 \) the we are in case (a). Thus we may assume that \( \pi_1(H) = G(Z) \). If for every \( u \in G(R) \) we have that \( \pi_1(D_R(1, -u) \cap H) = D(1, -\pi_1(u)) \) then (5.5) implies we are in case (b). So suppose this fails for some \( u \in G(R) \). Then by (5.2) \( H = 1 \times H_2 \cup (-1 \times -H_2) \), for some ordering \( H_2 \) of \( G(R_2) \). We will first show that for every \( h_2 \in H_2 \) that:

\[
H_2 = \bigcup_{k \in D_{R_2}(1, -h_2) \cap H_2} D_{R_2}(1, h_2 k).
\]

Consider \( \varphi = ((1, 1), (1, h_2), (1, -1)) \in S \). We compute its value set two ways. First:

\[
D_S((1, 1), (1, h_2), (1, -1)) = \bigcup_{\beta \in D_S((1, 1), (1, -h_2))} D_S((1, 1), \beta(1, h_2)).
\]

Now \( (-1, h_2) \notin H \) so \( D_S((1, 1), (1, -h_2)) = 1 \times D_{R_2}(1, -h_2) \). For \( \varphi \) to represent an element of \( \alpha(1 \times G(R_2)) \) we must have \( \beta \in -H = H \). That is, \( \beta = (1, \beta_2) \) with \( \beta_2 \in H_2 \). Thus:

\[
D_S(\varphi) \cap \alpha(1 \times G(R_2)) = \alpha \cdot \bigcup_{\beta_2 \in D_{R_2}(1, -h_2) \cap H_2} (1 \times D_{R_2}(1, \beta_2 h_2)).
\]

Next:

\[
D_S(\varphi) = (1, h_2) \cdot \bigcup_{\gamma \in D_S((1, 1), (1, -1))} D_S((1, 1), \gamma(1, h_2)).
\]
For any $x \in H_2$ take $\gamma = (1, x) \in D_S((1, 1), (1, -1))$. Then since $(-1, -xh_2) \in H$:

$$
\alpha(1, x) \in (1, h_2)D_S((1, 1), (-1, -xh_2)) = (1, h_2) \cdot \{1, \alpha\}(1 \times D_{R_2}(1, xh_2)).
$$

Thus $D_S(\varphi) \cap \alpha(1 \times G(R_2)) = \alpha(1 \times H_2)$. The two computations of $D_S(\varphi)$ thus yield $H_2 = \cup D_{R_2}(1, \beta h_2)$, over $\beta \in D_{R_2}(1, -h_2) \cap H_2$.

We may now apply (5.8) to obtain $R_2 = \mathbb{Z} \cap R_3$, for some Witt ring $R_3$. Let $H_3 \subset G(R_3)$ be the subgroup such that $H_2 \cap (1 \times G(R_3)) = 1 \times H_3$. We note that both $H_2$ and $1 \times G(R_3)$ have index two in $G(R_2)$.

If $H_2 = 1 \times G(R_3)$ then $(1, 1) \times G(R_3) \subset H$ and we are in case (c). So suppose $H_2 \neq 1 \times G(R_3)$. Then $1 \times H_3$ has index two in $H_2$ and $H_3$ has index two in $G(R_3)$. Write $H_2 = 1 \times H_3 \cup (-1 \times zH_3)$, for some $z \in G(R_3)$. Then:

$$
(5.10)
H = [(1, 1) \times H_3] \cup [(1, -1) \times zH_3] \cup [(-1, -1) \times -H_3] \cup [(-1, 1) \times -zH_3].
$$

Now $[G(R_3) : H_3] = 2$ implies at least one of the cosets $zH_3, -H_3, -zH_3$ equals $H_3$. Say $zH_3 = H_3$. Then the second term of (5.10) shows $(1, 1, 1), (1, -1, 1) \in H$. Set $R_4$ equal to the product of the first copy of $\mathbb{Z}$ and $R_3$. Then $R = \mathbb{Z} \cap R_4$ and $\{\pm 1\} \times 1 \subset H$. We are thus in case (b). Next say $-H_3 = H_3$. Then $(1, 1) \times H_3 \subset H_2$. Since $H_2$ is an ordering we have $D_{R_2}((1, 1), (1, -1)) \subset H_2$. But the $1 \times G(R_3) \subset H_2$, a case we have already considered. Lastly, suppose $-zH_3 = H_3$. Then the fourth term of (5.10) shows $(1, 1, 1), (-1, 1, 1) \in H$. This is case (b) again. \hfill \Box


**Lemma 6.1.** Let $R = R_1 \cap R_2$, with $R_1 = R_0[E_1]$ and $E_1$ generated by $t$. Let $S$ be an $H$-extension of $R$. Let $\pi_1$ be the projection of $G(R)$ onto $G(R_1)$ and suppose $\pi_1(H) \not\subset G(R_0)$. Then either $\pi_1(H) = G(R_1)$ or $1 \times G(R_2) \subset H$.

**Proof.** From $\pi_1(H) \not\subset G(R_0)$ we may assume $h = (t, g_2) \in H$, for some $g_2 \in G(R_2)$. Suppose $\pi_1(H) \neq G(R_1)$. Choose $-g_1 \in G(R_1) \setminus \pi_1(H)$. Then $-g_1t \notin \pi_1(H)$. Set $u = (g_1t, 1)$ and note that $\pi_1(D_R(1, -u) \cap H) = 1$.

$$
D_R(1, uh, -h) = D_{R_1}(1, g_1, -t) \times D_{R_2}(1, g_2, -g_2) \supset g_1 \times G(R_2).
$$

Also:

$$
\bigcup_{k \in h(D_R(1, -u) \cap H)} D_R(1, -k) \subset D_{R_1}(1, -t) \times G(R_2) = \{1, -t\} \times G(R_2).
$$

Hence by (3.3), $g_1 \times G(R_2) \subset uH$. Multiplying by $u$ gives $t \times G(R_2) \subset H$. Thus $1 \times G(R_2) \subset H$. \hfill \Box
Lemma 6.2. Let \( R = R_1 \cap R_2 \), with \( R_1 = R_0[E_1] \) and \( E_1 \) generated by \( t \). Let \( S \) be an \( H \)-extension of \( R \). Let \( \pi_1 \) be the projection of \( G(R) \) onto \( G(R_1) \) and suppose \( \pi_1(H) = G(R_1) \). If \( u_1 \in G(R_0) \) and \( u = (u_1, u_2) \) then \( \pi_1(D_R(1, -u) \cap H) = D_{R_1}(1, -u_1) \).

Proof. Set \( K = \pi_1(D_R(1, -u) \cap H) \) and suppose \( K < D_{R_1}(1, -u_1) \). Let \( g \in G(R_0) \). Then \( (gt, g_2) \in H \) for some \( g_2 \in G(R_2) \), since \( \pi_1(H) = G(R_1) \). Now \( D_R(1, uh, -h) \) contains \(-gtD_R(1, -u_1) \times u_2T\), where \( T = -D_R(\{-u_2, -g_2\})' \). Also:
\[
\bigcup_{w \in b(D_R(1, -u) \cap H)} D_R(1, -w) \subset \bigcup_{k \in K} D_{R_1}(1, -kt) \times G(R_2) = (\{1\} \cup -gK) \times G(R_2).
\]
Hence by (3.3), if \( y \in D_{R_1}(1, -u_1) \setminus K \) then:
\[-gtu_1y \times u_2T \subset uH,
-gtu_1y \times T \subset H.
\]
Now \( u_2 \) and \(-u_2g_2 \) are in \( T \) so \((-gtu_1y, u_2)\) and \((-gtu_1y, -u_2g_2)\) are in \( H \). Thus \((1, -g_2) \in H \) and as result \((gt, -1) \in H \).

This holds for all \( g \in G(R_0) \) so we have that \( tG(R_0) \times -1 \subset H \). Thus \( G(R_0) \times 1 \subset H \). But the \( D_{R_1}(1, -u_1) \times 1 \subset D_R(1, -u) \cap H \) and \( \pi_1(D_R(1, -u) \cap H) = D_{R_1}(1, -u_1) \), a contradiction. \( \square \)

Lemma 6.3. Let \( R = R_1 \cap R_2 \), with \( R_1 = R_0[E_1] \) non-degenerate and \( E_1 \) generated by \( t \). Let \( S \) be an \( H \)-extension of \( R \). Let \( \pi_1 \) be the projection of \( G(R) \) onto \( G(R_1) \) and suppose \( \pi_1(H) = G(R_1) \). Then \( G(R_1) \times 1 \subset H \).

Proof. We will first show \( F(g) = F(1) \) for all \( g \in G(R_0) \). Let \( g \in G(R_0) \). Pick \( u_1 \in G(R_0) \) such that \( g \in D_{R_1}(1, -u_1) \). For all \( u_2 \in G(R_2) \), since \( g \in D_{R_1}(1, -u_1) = \pi_1(D_R(1, -\{u_1, u_2\}) \cap H) \), there exists a \( k \in G(R_2) \) with \((g, k) \in H \) and \( k \in D_{R_2}(1, -u_2) \). That is, \( G(R_2) = \cup_{k \in F(g)} D_{R_2}(1, -k) \). By (3.1) and (5.4), \( F(g) = F(1) \).

We next show \( G(R_0) \times 1 \subset H \). Continue to let \( g \in G(R_0) \). Now, as \( g \in \pi_1(H) = G(R_1) \), we have \((g, m) \in H \) for some \( m \in G(R_2) \). Then \( m \in F(g) = F(1) \) so \((1, m) \in H \) and hence \((g, 1) = (g, m)(1, m) \in H \). This shows \( G(R_0) \subset H \).

We will be done if we show \( F(t) = F(1) \). Then, if \((t, k) \in H \) we get \((1, k) \) and hence \((t, 1) \) are in \( H \). Apply the previous paragraph to get \( G(R_1) \times 1 = \{1, t\}G(R_0) \times 1 \subset H \).

So suppose \( F(t) \neq F(1) \). We have by (3.1):
\[
G(R_2) \neq \bigcup_{k \in F(t)} D_{R_2}(1, -k).
\]
Pick \( u_2 \in G(R_2) \setminus \cup D_{R_2}(1, -k) \). Set \( u = (-t, u_2) \). Then, as there is no \( k \) with \((t, k) \in H \) and \( k \in D_{R_2}(1, -k) \), we have \( \pi_1(D_R(1, -u) \cap H) = 1 \). Pick
any $g \in G(R)$ (we note that $|G(R)| > 1$, else $R_1$ is degenerate). Pick any $g_2 \in F(g)$ and set $h = (g, g_2) \in H$. Now:

$$D_R(1, uh, -h) = D_{R_1}(1, -gt, -g) \times D_{R_2}(1, u_2g_2, -g_2)$$

$$\supset -gt \times D_{R_2}(1, u_2g_2, -g_2).$$

Also:

$$\bigcup_{w \in h(D_R(1, -u) \cap H)} D_R(1, -w) \subset D_{R_1}(1, -g) \times G(R_2).$$

Hence, by (3.3),

$$-gt \times D_{R_2}(1, u_2g_2, -g_2) \subset uH$$

$$-t \times D_{R_2}(g_2, u_2, -1) \subset H.$$  

In particular, $(-t, g_2) \in H$. Since $(-1, 1) \in G(R_0) \times 1 \subset H$, we get $(t, g_2) \in H$. But then $g_3 \in F(t) \cap F(g)$, which equals $F(t) \cap F(1)$ by previous work. $F(t)$ is a coset of $F(1)$, by (5.4), so in fact $F(t) = F(1)$ as desired. 

7. The Main Theorem.

**Theorem 7.1.** Let $R$ be a Witt ring of elementary type. If $S$ is an $H$-extension of $R$, for some subgroup $H \subset G(R)$, then $S$ is also of elementary type.

**Proof.** We argue by induction on $|G(R)|$. If $|G(R)| \leq 2$ then either $H = 1$ or $H = G(R)$ and we are done by (1.1). Suppose $|G(R)| > 2$. If $R = A \cap B$ then $\pi_A$ will denote the projection of $G(R)$ onto $G(A)$.

Now suppose $R$ is degenerate. Write $R = D_k \cap R_2$, where $R_2$ is non-degenerate and $G(D_k) \times 1 = \text{rad}(R)$. If some $d \in \text{rad}(R) \cap H$ then write $R = D_1 \cap R_3$, where $d$ generates $D_1$. We have $G(D_1) \times 1 \subset H$ and so $H = G(D_1) \times H_3$, for some subgroup $H_3 \subset G(R_3)$. Then (3.2) implies $S = D_1 \cap S_3$, where $S_3$ is an $H_3$-extension of $R_3$. By induction, $S_3$ is of elementary type and so $S$ is also. We may thus assume $\text{rad}(R) \cap H = 1$. If $\pi_{D_k}(H) \cap G(D_k) \neq 1$ then by (4.1) we may replace $H$, without affecting $S$, by another subgroup $H_0$ such that $\pi_{D_k}(H_0) \cap G(D_k) = 1$. We assume this has already been done so that $\pi_{D_k}(H) \cap G(D_k) = 1$. We note this also holds trivially if $R$ is non-degenerate and $k = 0$.

Next suppose $R$ has a local type factor $L$ with $|G(L)| \geq 8$. Write $R = L \cap R_4$. $D_k$ is a factor of $R_4$. We check the hypothesis of (5.7). Let $a \in G(L)$ and suppose $x \in F(a) \cap \text{rad}(R_4)$. This means $(a, x) \in H$ and so $x \in \pi_{D_k}(H) = 1$. Thus $F(a) \cap \text{rad}(R_4) \subset \{1\}$, as desired. Apply (5.6) to get that either $G(L) \times 1 \subset H$ or $\pi_L(H) = 1$. In the first case, $H = G(L) \times H_4$, for some subgroup $H_4$ of $G(R_4)$. Then (3.2) implies $S = L \cap S_4$, where $S_4$ is an $H_4$-extension of $R_4$. By induction, $S_4$ is of elementary type and so $S$ is also. We may thus assume we are in the second case: $\pi_L(H) = 1$. 

Now suppose the local factor is \( L_1 = \mathbb{Z} \). There are three cases according to (5.9). In case (b) we can write \( R = \mathbb{Z} \cap R_5 \), with \( \{ \pm 1 \} \times 1 \subset H \). Then \( H = G(\mathbb{Z}) \times H_5 \), for some subgroup \( H_5 \subset G(R_5) \). Applying (3.2) again gives \( S = \mathbb{Z} \cap S_5 \), where \( S_5 \) is an \( H_5 \)-extension of \( R_5 \). Induction again shows \( S \) is of elementary type. In case (c) we can write \( R = \mathbb{Z} \cap \mathbb{Z} \cap R_6 \), with \( (1,1) \times G(R_6) \subset H \). Once again (3.2) yields \( S = S_0 \cap R_6 \), where \( S_0 \) is an \( H_0 \)-extension of \( \mathbb{Z} \cap \mathbb{Z} \), for some subgroup \( H_0 \). Since \( \mathbb{Z} \cap \mathbb{Z} \cong \mathbb{Z}[E_1] \), (1.5) shows \( S_0 \) is of elementary type (in (1.5)(b) we have \( R_0 = \mathbb{Z} \) so that its extension is of elementary type as \( |G(R_0)| = 2 \). Thus \( S \) is also of elementary type. We may thus assume we are in case (a) of (5.9), namely, that \( \pi_{L_1}(H) = 1 \). This is the same conclusion as when the local factor has at least 8 square classes.

The only local type factors we have omitted are those with 4 square classes and these Witt rings are group rings. We are thus in the following position: \( R = Y \cap W_1 \cap \ldots \cap W_n \), where \( Y \) is a product of \( D_k \) with \( k \geq 0 \), and local type rings \( L \) with \( |G(L)| \neq 4 \) and each \( W_i \) is a non-degenerate group ring. (It is possible that \( Y \) type rings \( L \) and these Witt rings are group rings. We are thus in the following position: \( S \) is itself a group ring, a case we covered two paragraphs ago. We drop the subscript \( i \) and suppose then that \( R = W = W_0[E_n] \), where \( W_0 \) is basic and \( n \geq 1 \). In cases (a),(c),(d) we have \( S \) is of elementary type. In case (b) \( S = S_0[E_n] \), where \( S_0 \) is an \( H \)-extension of \( W_0 \), and so again \( S \) is of elementary type by induction.

The second possibility in (6.1) is that \( \pi_{W_1}(H) = G(W_1) \). Then by (6.3) we have \( H = G(W_1) \times H_7 \), where \( H_7 \) is a subgroup of \( R_7 \). Apply (3.2) once again to get that \( S = W_7 \cap S_7 \), where \( S_7 \) is an \( H_7 \)-extension of \( R_7 \). Induction gives that \( S \) is of elementary type. This completes the result when (6.1) applies, that is, when \( \pi_{W_1}(H) \not\subset G(V_i) \), for some \( i \).

We may thus assume we have \( R = Y \cap W_1 \cap \ldots \cap W_n \), with \( \pi_Y(H) = 1 \) and every \( \pi_{W_i}(H) \subset V_i \). Choose \( t_i \notin V_i \), for each \( i \) and set \( g = (1,t_1,\ldots,t_n) \in G(R) \). Then:

\[
g \notin \pm \bigcup_{h \in H} D_{R}(1,-h),
\]

as for any \( h \in H \) has a coordinate in \( G(V_i) \). By (1.2), \( g \notin B(R) \). Thus \( R \) is itself a group ring, a case we covered two paragraphs ago.

The previous sections can be used to determine the possible \( H \)-extensions of a given ring \( R \). As an example, consider \( R = (D_1 \cap L_3)[E_2] \). Let \( d \) generate \( D_1 \), \( -1, a, b \) generate \( L_3 \) and \( s, t \) generate \( E_2 \). Thus \( G(R) = \text{gp}(d,-1,a,b,s,t) \), where \( \text{gp}(A) \) denotes the group generated by \( A \). \( G(R) \) has 2825 subgroups,
47 of which will yield $H$-extensions. Up to isomorphism, there are exactly 8 $H$-extensions of $R$. Below we list the 8 extensions $S$ along with one choice for the corresponding subgroup $H$.

1. $(D_1 \cap L_3)[E_3]$  
   $1$
2. $(D_1 \cap (D_1 \cap L_3)[E_1])[E_1]$  
   $\text{gp}(d, -1, a, b, s)$
3. $D_1 \cap (D_1 \cap L_3)[E_2]$  
   $G(R)$
4. $(\mathbb{Z} \cap (D_1 \cap L_3)[E_1])[E_1]$  
   $\text{gp}(d, a, b, s)$
5. $\mathbb{Z} \cap (D_1 \cap L_3)[E_2]$  
   $\text{gp}(d, a, b, s, t)$
6. $(D_1 \cap L_3[E_1])[E_2]$  
   $\text{gp}(d)$
7. $(D_2 \cap L_3)[E_2]$  
   $\text{gp}(d, -1, a, b)$
8. $(D_1[E_1] \cap L_3)[E_2]$  
   $\text{gp}(-1, a, b)$

We give a brief sketch of how this list was derived. Begin by running through the cases of (1.5). In (a) $H = 1$ and $S$ is (1) by (1.1). In (c), $G(R_0)$ is a proper subgroup of $H$, so $S$ is (2) or (3), depending on whether or not $H = G(R)$. In (d) $G(R_0) \subset \pm H, -1 \notin H$ and $H \not\subset G(R_0)$. Thus $H$ looks like a subgroup $K$ of index 2 in $G(R_0)$ that does not contain -1, together with one or more elements from \{t, s, ts\}. $S$ is (4) if $|H \cap E_2| = 2$ and (5) if $|H \cap E_2| = 4$. In (1.5)(b) $S = S_0[E_2]$, where $S_0$ is an $H$-extension of $R_0$. Now $R_0 = D_1 \cap L_3$. By (4.1) we can assume that either $H = G(D_1) \times H_2$, or that $H = 1 \times H_2$, for some $H_2 \subset G(L_3)$. Now $H_2 = 1$ or $H_2 = G(L_3)$ by (2.3) and (5.7). Since we have already done the case $H = 1$ this gives three choices: $G(D_1) \times 1, G(D_1) \times G(L_3)$ and $1 \times G(L_3)$. The corresponding $S$ is (6), (7) and (8), respectively.

References


Received September 29, 1997 and revised March 25, 1998.

Southern Illinois University  
Carbondale, IL 62901-4408  
E-mail address: rfitzg@math.siu.edu
EQUIVARIANT EMBEDDINGS OF STEIN DOMAINS
SITTING INSIDE OF COMPLEX SEMIGROUPS

Bernhard Krötz

In this paper we prove an equivariant version of Hörmanders embedding theorem for Stein manifolds. More concretely, let $G$ be a connected Lie group sitting in its complexification $G_C$ and $D \subseteq G_C$ a $G \times G$-invariant Stein domain. Under slight obstructions on $D$ we construct a Hilbert space $\mathcal{H}$ equipped with a unitary $G \times G$-action and a holomorphic equivariant closed embedding $e : D \to \mathcal{H}^* \setminus \{0\}$.

Introduction.

An interesting problem in the field of equivariant complex analysis is: Given a connected Lie group $G$ sitting in its universal complexification $G_C$, how do the $G \times G$-invariant Stein domains in $G_C$ look like. K.-H. Neeb has shown in [Ne98] that all domains of the form

$$D = G \exp_{G_C}(iD_h),$$

where $D_h \subseteq \mathfrak{g}$ is a $\text{Ad}(G)$-invariant convex domain consisting of elliptic elements, i.e., all operators $i \text{ad} X$, $X \in D_h$, are diagonalizable over the reals, are Stein manifolds. Moreover there is also strong evidence for that these $D$ exhaust up to multiplication with $N_{G_C}(G)$ all proper bi-invariant Stein domains in $G_C$ (cf. [GG77], [Ne98]).

By Hörmander’s Embedding Theorem one knows that every Stein manifold of dimension $n$ can be embedded biholomorphically as a closed submanifold of $\mathbb{C}^{2n+1}$ (cf. [Hör73]). Now the natural question is: Given a bi-invariant Stein domain $D = G \exp_{G_C}(iD_h)$ in $G_C$, does there exist a $G \times G$-equivariant embedding into some complex Hilbert space $\mathcal{H}$ endowed with a unitary $G \times G$-action. In this paper we show that under quite natural assumptions the answer is affirmative. More concretely, if $\text{Ad}(G)$ is closed in $\text{Aut}(\mathfrak{g})$, the center $Z(G)$ is compact and the convex domain $D_h$ is pointed, then there exists a positive definite biinvariant holomorphic kernel $K$ on $D$, such that the map

$$e_K : D \to \mathcal{H}_K^* \setminus \{0\}, \quad z \mapsto K_z$$

defines a $G \times G$-equivariant closed embedding. Here $\mathcal{H}_K$ denotes the reproducing kernel Hilbert space and $K_z : \mathcal{H}_K \to \mathbb{C}$, $f \mapsto f(z)$ the point evaluations corresponding to $K$. 

55
Our method to construct such a kernel $K$ is to sum up kernels $K^\lambda$ associated to unitary highest weight representations $(\pi_\lambda, \mathcal{H}_\lambda)$ of $G$ over a certain lattice $\Gamma \subseteq i\mathfrak{t}^*$, where $\mathfrak{t}$ denotes a compactly embedded Cartan subalgebra of $\mathfrak{g}$. More precisely, we set

$$K = \sum_{\lambda \in \Gamma} \|\lambda\|^N K^\lambda$$

with $\| \cdot \|$ denoting a norm on $i\mathfrak{t}^*$ and $N \in \mathbb{N}$. These kernels $K$ have the important property of tending to infinity at the boundary of $D$, i.e.,

$$\lim_{z \to \partial D} K(z, z) = \infty;$$

a result which is crucial for verifying the closedness of the map $e_K$.

We think that our results are a little bit surprising and we do not really understand what is actually going on. For instance, what is the reason for that one has to exclude zero in $\mathcal{H}_K^\vee$ to achieve the closedness of the map $e_K$, or, is it possible to find an equivariant closed holomorphic embedding $D \to E$ into a complex topological vector space endowed with a continuous $G \times G$-action. We hope that our results give rise to a further discussion leading to a better understanding of these phenomena.

I. The boundary behaviour of bi-invariant kernels.

In this first section we characterize the boundary behaviour of bi-invariant holomorphic positive definite kernels on a bi-invariant domain $D = G \text{Exp}(iD_h)$ by means of the boundary behaviour on the abelian submanifold $D_T := T \text{Exp}(i(D_h \cap \mathfrak{t}))$. If the convex invariant set $D_h \subseteq \mathfrak{g}$ is a pointed cone, we show that $\lim_{z \to \partial D} K(z, z) = \infty$ if and only if $\lim_{z \to \partial D_T} K(z, z) = \infty$. As abelian domains are comparable easily to deal with contrary to the highly non-commutative bi-invariant domains $D$, this result allows us in the sequel to make quite explicit computations.

Definition I.1. Let $V$ be a finite dimensional real vector space and $V^*$ its dual.

(a) For each subset $E \subseteq V$ we define its dual cone by $E^* := \{\alpha \in V^* : (\forall x \in E) \alpha(x) \geq 0\}$. We note that $E^*$ is a convex closed subcone of $V^*$.

(b) For a convex subset $E \subseteq V$ we set $H(E) := \{x \in V : x + E = E\}$, and $\lim E := \{x \in V : x + E \subseteq E\}$. We call $H(E)$ the edge and $\lim E$ the limit cone of $E$. Note that $H(E)$ is a vector space, $H(E) = H(E)$ if $E$ is open and that $\lim E$ is a convex cone in $V$.

(c) A convex set $E$ is called pointed if it contains no affine lines. Note that if $E$ is open or closed then $E$ is pointed if and only if its edge is zero. □
Definition I.2. Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{R}$.

(a) An element $X \in \mathfrak{g}$ is called elliptic if $\text{ad} X$ operates semisimply with purely imaginary spectrum. A convex cone $W \subseteq \mathfrak{g}$ is said to be elliptic if $W^0 \neq \emptyset$ and all $X \in W^0$ are elliptic.

(b) For a subalgebra $\mathfrak{a} \subseteq \mathfrak{g}$ we write $\text{Inn}(\mathfrak{a}) := \langle e^{\text{ad} a} \rangle \subseteq \text{Aut}(\mathfrak{g})$ for the corresponding group of inner automorphisms. A subalgebra $\mathfrak{a} \subseteq \mathfrak{g}$ is said to be compactly embedded if $\text{Inn}(\mathfrak{a})$ is relatively compact in $\text{Aut}(\mathfrak{g})$.

(c) Let $t \subseteq \mathfrak{g}$ be a compactly embedded Cartan subalgebra and recall that there exists a unique maximal compactly embedded subalgebra $\mathfrak{k}$ containing $t$ (cf. [HHL89, A.2.40]).

(d) Associated to the Cartan subalgebra $\mathfrak{t}_C$ in the complexification $\mathfrak{g}_C$ is a root decomposition as follows. For a linear functional $\alpha \in \mathfrak{t}_C^*$ we set

$$\mathfrak{g}_C^\alpha := \{ X \in \mathfrak{g}_C : (\forall Y \in \mathfrak{t}_C) [Y, X] = \alpha(Y) X \}$$

and write $\Delta_\mathfrak{t} := \{ \alpha \in \mathfrak{t}_C^* \setminus \{ 0 \} : \mathfrak{g}_C^\alpha \neq \{ 0 \} \}$ for the set of roots. Then $\mathfrak{g}_C = \mathfrak{t}_C \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_C^\alpha$, $\alpha(t) \subseteq i \mathbb{R}$ for all $\alpha \in \Delta$ and $\overline{\mathfrak{g}_C}^\alpha = \mathfrak{g}_C^{-\alpha}$, where $X \mapsto \overline{X}$ denotes complex conjugation on $\mathfrak{g}_C$ with respect to $\mathfrak{g}$.

(e) A root $\alpha$ is said to be compact if $\mathfrak{g}_C^\alpha \subseteq \mathfrak{t}_C$ and non-compact otherwise. We write $\Delta_k$ for the set of compact roots and $\Delta_n$ for the non-compact ones.

(f) A positive system $\Delta^+$ of roots is a subset of $\Delta$ for which there exists a regular element $X_0 \in i \mathfrak{t}^*$ with $\Delta^+ := \{ \alpha \in \Delta : \alpha(X_0) > 0 \}$. A positive system is said to be $\mathfrak{t}$-adapted if the set $\Delta^+_n := \Delta_n \cap \Delta^+$ is invariant under the Weyl group $W_\mathfrak{t} := N_{\text{Inn}(\mathfrak{t})}(\mathfrak{t})/Z_{\text{Inn}(\mathfrak{t})}(\mathfrak{t})$ acting on $\mathfrak{t}$. We recall from [Ne99, Ch. V] that there exists a $\mathfrak{t}$-adapted positive system if and only if $\mathfrak{g}(\mathfrak{t}(\mathfrak{f})) = \mathfrak{t}$. In this case we call $\mathfrak{g}$ quasimeritian. In this case it is easy to see that $\mathfrak{s}$ is quasimeritian too, and so all simple ideals of $\mathfrak{s}$ are either compact or hermitian.

(g) We associate to a positive system $\Delta^+$ the convex cones

$$C_{\text{min}} := \text{cone} \{ i[X_\alpha, X_\alpha] : X_\alpha \in \mathfrak{g}_C^\alpha, \alpha \in \Delta_n^+ \},$$

and $C_{\text{max}} := (i\Delta_n^+)^* = \{ X \in \mathfrak{t} : (\forall \alpha \in \Delta_n^+) i\alpha(X) \geq 0 \}$. Note that both $C_{\text{min}}$ and $C_{\text{max}}$ are closed convex cones in $\mathfrak{t}$.

(h) Write $p_\mathfrak{t} : \mathfrak{g} \rightarrow \mathfrak{t}$ for the orthogonal projection along $[\mathfrak{t}, \mathfrak{g}]$ and set $O_X := \text{Inn}(\mathfrak{g}).X$ for the adjoint orbit through $X \in \mathfrak{g}$. We define the minimal and maximal cone associated to $\Delta^+$ by

$$W_{\text{min}} := \{ X \in \mathfrak{g} : p_\mathfrak{t}(O_X) \subseteq C_{\text{min}} \} \quad \text{and} \quad W_{\text{max}} := \{ X \in \mathfrak{g} : p_\mathfrak{t}(O_X) \subseteq C_{\text{max}} \}$$

and note that both cones are convex closed and $\text{Inn}(\mathfrak{g})$-invariant. $\square$
From now on we assume that $\mathfrak{g}$ contains a compactly embedded Cartan subalgebra $\mathfrak{t} \subseteq \mathfrak{g}$ and that there exists an elliptic cone $W \subseteq \mathfrak{g}$. Then there exists a $\mathfrak{t}$-adapted positive system $\Delta^+$ such that

$$C_{\min} \subseteq W \cap \mathfrak{t} \subseteq C_{\max}$$

holds and $W_{\max}$ is an elliptic cone (cf. [Ne96b, Th. II.11]). Moreover, we have $W_{\min} \cap \mathfrak{t} = C_{\min}$ and $W_{\max} \cap \mathfrak{t} = C_{\max}$ (cf. [Ne97, Lemma I.1]).

**Definition I.3.** (a) Let $W \subseteq \mathfrak{g}$ be a closed elliptic cone. Let $\tilde{G}$, resp. $\mathcal{G}_C$, be the simply connected Lie groups associated to $\mathfrak{g}$, resp. $\mathfrak{g}_C$, and set $G_1: = \langle \exp \mathfrak{g} \rangle \subseteq \tilde{G}_C$. Then Lawson’s Theorem (cf. [HiNe93, Th. 7.34, 35]) says that the subset $\Gamma_{G_1}(W): = G_1 \exp(iW)$ is a closed subsemigroup of $G_C$ and the polar map

$$G_1 \times W \to \Gamma_{G_1}(W), \quad (g, X) \mapsto g \exp(iX)$$

is a homeomorphism.

Now the universal covering semigroup $\Gamma_{\tilde{G}}(W): = \tilde{\Gamma}_{G_1}(W)$ has a similar structure. We can lift the exponential function $\exp: \mathfrak{g} + iW \to \Gamma_{G_1}(W)$ to an exponential mapping $\Exp: \mathfrak{g} + iW \to \Gamma_{\tilde{G}}(W)$ with $\Exp(0) = 1$ and thus obtain a polar map

$$\tilde{G} \times W \to \Gamma_{\tilde{G}}(W), \quad (g, X) \mapsto g \Exp(iX)$$

which is a homeomorphism.

If $G$ is a connected Lie group associated to $\mathfrak{g}$, then $\pi_1(G)$ is a discrete central subgroup of $\Gamma_{\tilde{G}}(W)$ and we obtain a covering homomorphism $\Gamma_G(W) \to \Gamma_G(W): = \Gamma_{\tilde{G}}(W)/\pi_1(G)$ (cf. [HiNe93, Ch. 3]). It is easy to see that there is also a polar map $G \times W \to \Gamma_G(W), (g, X) \mapsto g \Exp(iX)$ which is a homeomorphism. The semigroups of the type $\Gamma_G(W)$ are called complex Ol’shanskiĭ semigroups.

The subset $\Gamma_G(W^0) \subseteq \Gamma_G(W)$ is an open semigroup carrying a complex manifold structure such that semigroup multiplication is holomorphic. Moreover there is an involution on $\Gamma_G(W)$ given by

$$\ast: \Gamma_G(W) \to \Gamma_G(W), s = g \Exp(iX) \mapsto s^* = \Exp(iX)g^{-1}$$

which is antiholomorphic on $\Gamma_G(W^0)$ (cf. [HiNe93, Th. 9.15] for a proof of all that). Thus $\Gamma_G(W)$ is an involutive semigroup.

(b) A bi-invariant domain $D \subseteq \Gamma_G(W^0_{\max})$ is an open connected $G \times G$ bi-invariant subset of $\Gamma_G(W^0_{\max})$. Note that

$$D = G \Exp(iD_h) = G \Exp(iD)G,$$

where $D_h \subseteq W^0_{\max}$ and $D = D_h \cap \mathfrak{t}$. Recall that $D$ is a Stein manifold if and only if $D_h$ is convex (cf. [Ne98, Th. 6.1]). In this case $D$ is called a bi-invariant Stein domain. We call $D$ pointed if $D_h$ is pointed in $\mathfrak{g}$. The
boundary of a left $G$-invariant subset $E = G \exp(iE_h) \subseteq \Gamma_G(W_{\text{max}})$ is defined as $\partial E = G \exp(i\partial E_h)$. Note that $\partial D = \overline{D}\setminus D$ for every bi-invariant domain $D$, where the closure $\overline{D}$ is taken in $\Gamma_G(W_{\text{max}})$.

**Lemma I.4.** Let $W \subseteq \mathfrak{g}$ be an invariant elliptic pointed convex cone and set $C := W \cap \mathfrak{t}$. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence in $W^0$ converging to $X \in \partial W$. Then there exists a subsequence $(X_{n_k})_{k \in \mathbb{N}}$ and a sequence $(Y_{n_k})_{k \in \mathbb{N}}$ in $C^0$ with $Y_{n_k} \in \text{Inn}(\mathfrak{g}).X_{n_k}$ and $Y_{n_k} \to Y \in \partial C$.

**Proof.** W.l.o.g. we may assume that $W$ is closed. According to [HiNe93, Th. 7.27], we can reconstruct $W^0$ from $C^0$, i.e., we have $W^0 = \text{Inn}(\mathfrak{g}).C^0$. In particular, we find a sequence $(Y_n)_{n \in \mathbb{N}}$ in $C^0$ and a sequence $(g_n)_{n \in \mathbb{N}}$ in $\text{Inn}(\mathfrak{g})$ such that $g_n.X_n = Y_n$. We claim that $(Y_n)_{n \in \mathbb{N}}$ is bounded.

The Convexity Theorem for Adjoint Orbits (cf. [KrNe96, Th. VIII.9]) implies that

$$p_t(X_n) \in \text{conv}(\mathcal{W}_t.Y_n) + C \subseteq C$$

for all $n \in \mathbb{N}$.

As $C$ is pointed, a sequence $(Z_n)_{n \in \mathbb{N}}$ in $C$ is unbounded if and only if $\lim_{n \to \infty} \alpha(Z_n) = \infty$ holds for one $\alpha \in \text{int } C^*$. Thus if $(Y_n)_{n \in \mathbb{N}}$ is unbounded, then (1.1) together with the invariance of $C$ under $\mathcal{W}_t$ implies that $(p_t(X_n))_{n \in \mathbb{N}}$ is unbounded. But this contradicts the fact that $(p_t(X_n))_{n \in \mathbb{N}}$ being a continuous image of a Cauchy sequence is bounded, proving the claim.

Let now $(Y_{n_k})_{k \in \mathbb{N}}$ be a convergent subsequence of $(Y_n)_{n \in \mathbb{N}}$ and $Y = \lim_{k \to \infty} Y_{n_k}$ the corresponding limit in $C$. It remains to show that $Y \in \partial C$. To obtain a contradiction we assume that $Y \in C^0$.

We write $\text{Sl}(W)$ for the special automorphism group of the cone $W$ and note that $\text{Inn}(\mathfrak{g}) \subseteq \text{Sl}(W)$ (cf. [HiNe93, Prop. 7.3(v)]). Then [HiNe93, Prop. 1.11] implies that there exists a convergent subsequence of $(g_{n_k})_{k \in \mathbb{N}}$ in $\text{Sl}(W)$ which we also denote by $(g_{n_k})_{k \in \mathbb{N}}$. Write $g$ for the corresponding limit. Then

$$X = \lim_{k \to \infty} g_{n_k}^{-1}.Y_{n_k} = g^{-1}.Y.$$

Since $\text{Sl}(W).W^0 = W^0$ and $Y \in W^0$, this implies that $X \in W^0$; a contradiction, concluding the proof of the lemma.

**Definition I.5.** Let $M$ be a complex manifold and $\text{Hol}(M)$ denote the space of holomorphic functions on $M$. We write $\overline{M}$ for $M$ equipped with the opposite complex structure.

(a) A function $K \in \text{Hol}(M \times \overline{M})$ is called a holomorphic positive definite kernel if for every sequence $z_1, \ldots, z_n$ in $M$ the matrix $(K(z_i, z_j))_{i,j}$ is positive semi-definite. We write $\mathcal{P}(M^2)$ for the convex cone of all holomorphic positive definite kernels on $M$. Note that every $K \in \mathcal{P}(M^2)$ satisfies the
Lemma I.6. Let \( V \) be a finite dimensional real vector space, \( V^\oplus = V \oplus \mathbb{R} \) and \( E \subseteq V \) a convex subset. Set \( E^\sharp = \mathbb{R}^+ (E \times \{1\}) \). Then \( E^\sharp \) is a convex subcone of \( V^\sharp \) and the following assertions hold:

(i) \( E \) is closed in \( E^\sharp \).
(ii) \( E \) is pointed if and only if \( E^\sharp \) is pointed.
(iii) If \( E \) is open or closed, then \( \partial E^\sharp = ]0, \infty[ \times \{1\} \cup (\text{lim } E \times \{0\}) \).

Proof. (i) This is clear.
Let $E$ be a finite dimensional real vector space and $V \subseteq E$ a convex set. Further let $V_1 := V/H(E)$, denote $q : V \to V_1$ the corresponding quotient homomorphism and set $E_1 := q(E)$. Then we have $q(\partial E) = \partial E_1$.

Proof. As $E + H(E) = E$ it follows that $E^0 + H(E) = E^0$ and $\overline{E} + H(E) = H(E)$. Thus $q(\overline{E}) = \overline{E}_1$, and $q(\overline{E}^0) = \overline{E}_1^0$ since $q$ is an open mapping. This proves the lemma. \hfill \Box

Lemma 1.7. Let $E$ be a finite dimensional real vector space and $V \subseteq E$ a convex set. Further let $V_1 := V/H(E)$, denote $q : V \to V_1$ the corresponding quotient homomorphism.

Let $\gamma^\sharp = \gamma \oplus \mathbb{R}$, $\gamma^\sharp = G \times \mathbb{R}$, and

$$ D^\sharp := \Gamma_G \left( D^\sharp_h \right) \subseteq \Gamma_G \left( W^0_{\max} \oplus \mathbb{R} \right) \cong \Gamma_G(W^0_{\max}) \oplus \mathbb{C}. $$

(i) The map $j : D \to D^\sharp$, $s \mapsto (s, i)$ is a $G \times G$-equivariant holomorphic closed embedding inducing a map

$$ P_{G \times G} \left( D^{\sharp 2} \right) \to P_{G \times G}(D^2), \quad K \mapsto K^\sharp := K^\sharp \circ j. $$

(ii) Let $K \in P_{G \times G}(D^2)$. Then the following statements are equivalent:

(a) $\lim_{z \to \partial D} K(z, z) = \infty$.

(b) $\lim_{x \to \partial D_h} K(\text{Exp}(iX), \text{Exp}(iX)) = \infty$. Moreover, if $K = K^\sharp \circ j$ with some $K^\sharp \in P_{G \times G}(D^\sharp 2)$, then (a)-(b) are implied by

(c) $\lim_{x \to \partial D^\sharp_h} K^\sharp(\text{Exp}(iX), \text{Exp}(iX)) = \infty$.

(iii) Let $a := H(D_h)$ be the edge of $D_h$, $g_1 := g/\mathfrak{a}$ and $q : g \to g_1$ the corresponding quotient morphism. Further let $A := \langle \exp(a) \rangle$, $G_1 := G/A$ and $D_1 := G_1 \text{Exp}(iD_{h,1})$. Then the quotient morphism $\tilde{q} : D \to D_1$ induced by $q$ gives rise to an injection

$$ \tilde{q}^* : P_{G_1 \times G_1}(D_1^2) \to P_{G \times G}(D^2), \quad K_1 \mapsto K := K_1 \circ \tilde{q} $$

and the following statements are equivalent:

(a) $\lim_{x \to \partial D_h} K(\text{Exp}(iX), \text{Exp}(iX)) = \infty$.

(b) $\lim_{x \to \partial D_h} K_1(\text{Exp}(iX), \text{Exp}(iX)) = \infty$.

(c) $\lim_{x \to \partial D_1} K_1(\text{Exp}(iX), \text{Exp}(iX)) = \infty$. 

Proof. (i) Obviously, \( j \) is holomorphic and \( G \times G \)-equivariant. Further the closedness of \( j \) follows from Lemma I.6(i) together with the Polar Decomposition of a complex Ol’shanski˘ı semigroup. The second assertion is clear.

(ii) (a)\( \Rightarrow \) (b) is obvious.

(b) \( \Rightarrow \) (a): This follows from the biinvariance of \( K \).

The remaining statement follows from Lemma I.6(iii).

(iii) (a) \( \iff \) (b): In view of \( q(Dh) = D_{h,1} \) (cf. Lemma I.7), this follows from the equivalence of (a) and (b) in (ii).

(b) \( \Rightarrow \) (c) is clear.

(c) \( \Rightarrow \) (b) follows from the \( \text{Inn}(\mathfrak{g}) \)-invariance of the map

\[
D_{h,1} \to \mathbb{R}^+, \quad X \mapsto K_1(\exp(iX), \exp(iX))
\]
together with Lemma I.4. \( \square \)

II. A kernel tending to infinity at the boundary.

In this section we construct a kernel \( K \in \mathcal{P}_{G \times G}(D^2) \) tending to infinity at the boundary. In view of Proposition I.8, this reduces to the case where \( D = \Gamma_G(W) \) is a pointed complex Ol’shanski˘ı semigroup. But Proposition I.8(iii) tells us even more: We only have to check that \( \lim_{X \to \partial C} K(\exp(iX), \exp(iX)) = \infty \). To start out we need some notation concerning highest weight representations and their associated characters.

Definition II.1. Let \( \Delta^+ \) be a positive system.

(a) For a \( \mathfrak{g}_\mathbb{C} \)-module \( V \) and \( \beta \in (\mathfrak{t}_\mathbb{C})^* \) we write \( V^\beta := \{ v \in V : (\forall X \in \mathfrak{t}_\mathbb{C}) X.v = \beta(X)v \} \) for the weight space of weight \( \beta \) and \( \mathcal{P}_V = \{ \beta : V^\beta \neq \{0\} \} \) for the set of weights of \( V \).

(b) Let \( V \) be a \( \mathfrak{g}_\mathbb{C} \)-module and \( v \in V^\lambda \) a \( \mathfrak{t}_\mathbb{C} \)-weight vector. We say that \( v \) is a primitive element of \( V \) (with respect to \( \Delta^+ \)) if \( g_\mathfrak{a} \cdot v = \{0\} \) holds for all \( \alpha \in \Delta^+ \).

(c) A \( \mathfrak{g}_\mathbb{C} \)-module \( V \) is called a highest weight module with highest weight \( \lambda \) (with respect to \( \Delta^+ \)) if it is generated by a primitive element of weight \( \lambda \).

(d) Let \( \lambda \in \mathfrak{t}_\mathbb{C}^* \) be dominant integral w.r.t. \( \Delta_\mathbb{R}^+ \) and \( F(\lambda) \) the corresponding highest weight module for \( \mathfrak{t}_\mathbb{C} \). Assume that \( \Delta_\mathbb{R}^+ \) is \( \mathfrak{t} \)-adapted and set \( \mathfrak{p}^\pm = \bigoplus_{\alpha \in \Delta_\mathbb{R}^\pm} \mathfrak{g}_\mathbb{C}^\alpha \). We define the generalized Verma module by

\[
N(\lambda) := \mathcal{U}(\mathfrak{g}_\mathbb{C}) \otimes_{\mathcal{U}(\mathfrak{t}_\mathbb{C} + \mathfrak{p}^+)} F(\lambda).
\]

Note that \( N(\lambda) \) is a highest weight module for \( \mathcal{U}(\mathfrak{g}_\mathbb{C}) \) with highest weight \( \lambda \). We denote by \( L(\lambda) \) the unique irreducible quotient of \( N(\lambda) \).

(e) Let \( G \) be a connected Lie group with Lie algebra \( \mathfrak{g} \). We write \( K \) for the analytic subgroup of \( G \) corresponding to \( \mathfrak{k} \). Let \( (\pi, \mathcal{H}) \) be a unitary representation of \( G \). A vector \( v \in \mathcal{H} \) is called \( K \)-finite if it is contained in
a finite dimensional $K$-invariant subspace. We write $\mathcal{H}^{K,\omega}$ for the space of analytic $K$-finite vectors.

(f) An irreducible unitary representation $(\pi, \mathcal{H})$ of $G$ is called a highest weight representation w.r.t. $\Delta^+$ with highest weight $\lambda \in i\mathfrak{t}^*$ if $\mathcal{H}^{K,\omega}$ is a highest weight module for $\mathfrak{g}_C$ w.r.t. $\Delta^+$ and highest weight $\lambda$. We say that the irreducible highest weight module $L(\lambda)$ is unitarizable if there exists a unitary highest weight representation $(\pi_\lambda, \mathcal{H}_\lambda)$ of $G$ with $\mathcal{H}^{K,\omega}_\lambda \cong L(\lambda)$ as $\mathfrak{g}_C$-modules. We write $\text{HW}(G, \Delta^+) \subseteq i\mathfrak{t}^*$ for the set of highest weights corresponding to unitary highest weight representations of $G$ w.r.t. $\Delta^+$ and write $\text{HW}(\Delta^+) := \text{HW}(G, \Delta^+)$ for the set of all highest weights w.r.t. $\Delta^+$ which correspond to a unitarizable $L(\lambda)$.

(g) Let $\lambda \in \text{HW}(\Delta^+)$. We call $\lambda$ singular if the natural map $N(\lambda) \to L(\lambda)$ has a non-trivial kernel and non-singular otherwise. \hfill \square

For each unitary representation $(\pi, \mathcal{H})$ of $G$ we write $(\pi^*, \mathcal{H}^*)$ for the corresponding dual representation. Let $B_2(\mathcal{H})$ be the space of Hilbert Schmidt operators on $\mathcal{H}$. We define a representation of $G \times G$ on $B_2(\mathcal{H})$ by

$$\pi^c : G \times G \to U(B_2(\mathcal{H})), \quad \pi^c(g_1, g_2).A := \pi(g_2)A\pi(g_1)^*.$$

Note that there is a canonical isomorphism between $(\pi^* \otimes \pi, \mathcal{H}^* \otimes \mathcal{H})$ and $(\pi^c, B_2(\mathcal{H}))$.

Now we fix a positive system $\Delta^+$ associated to $W_{\text{max}}$.

Recall from [HiNe96, Th. 3.6, Th. B] that each highest weight representation $(\pi_\lambda, \mathcal{H}_\lambda)$ of $G$ extends to an holomorphic representation of $\Gamma_G(W_{\text{max}})$ denoted by the same symbol. Moreover all operators $\pi_\lambda(s), s \in \Gamma_G(W_{\text{max}}^0)$, are of trace class (cf. [Ne94, Th. III.8]), so that $\Theta_\lambda(s) := \text{tr} \pi_\lambda(s)$ makes sense for all $s \in \Gamma_G(W_{\text{max}}^0)$. We call $\Theta_\lambda$ the character of $(\pi_\lambda, \mathcal{H}_\lambda)$ and note that $\Theta_\lambda$ is holomorphic on $\Gamma_G(W_{\text{max}}^0)$ (cf. [Ne94, Th. IV.11]).

Associated to a $\mathfrak{t}$-adapted positive system, we define the function

$$\phi : \mathfrak{c}^{0}_{\text{max}} \to \mathbb{R}^+, \quad X \mapsto \frac{1}{\prod_{\alpha \in \Delta^+} (1 - e^{-\alpha(X)})^{m_\alpha}}$$

where $m_\alpha := \dim_C \mathfrak{g}_C^\alpha$ for all $\alpha \in \Delta$.

**Lemma II.2.** Let $\lambda \in \text{HW}(G, \Delta^+)$ be non-singular and $(\pi_\lambda, \mathcal{H}_\lambda)$ an associated highest weight representation of $G$. If $\Theta^K_\lambda$ denotes the character of $F(\lambda)$, then

$$\Theta_\lambda(\text{Exp } X) = \phi(X)\Theta^K_\lambda(\text{Exp } X)$$

for all $X \in \mathfrak{c}^{0}_{\text{max}}$.

**Proof.** [Kr97, Lemma IV.8(i)]. \hfill \square

**Proposition II.3.** Let $K$ be a connected Lie group with compact Lie algebra $\mathfrak{k}$. Let $\mathfrak{t}$ be a Cartan subalgebra of $\mathfrak{k}$, $\Delta^+_\mathfrak{c}$ be a positive system of roots and $\lambda \in i\mathfrak{t}^*$ be a dominant analytically integral element. Further let $\Theta^K_\lambda$ denote
the character of the holomorphic representation \((\pi^K, F(\lambda))\) of \(K_C\) and set \(d_\lambda := \dim F(\lambda)\).

(i) \((\forall X \in i t)\) \(e^{\lambda(X)} \leq \Theta^K_\lambda(\exp K_C(X)) \leq d_\lambda \sup_{w \in W_k} e^{\lambda(w.X)}\).

(ii) If \(\| \cdot \|\) denotes a norm on \(i t^*\), then there exists a constant \(c > 0\) and an element \(n \in \mathbb{N}\) such that

\[
d_\lambda \leq c \| \lambda \|^n + c
\]

holds for all integral elements \(\lambda\).

Proof. (i) Let \(\{v_j: 1 \leq j \leq d_\lambda\}\) be an orthonormal basis of weight vectors of \(F(\lambda)\). For each \(j\) let \(\alpha_j\) be the weight corresponding to \(v_j\). Then we have for all \(X \in i t_C\) that

\[
(2.1) \quad \Theta^K_\lambda(\exp K_C(X)) = \sum_{j=1}^{d_\lambda} \langle \pi^K_\lambda(\exp K_C(X)), v_j, v_j \rangle = \sum_{j=1}^{d_\lambda} e^{\alpha_j(X)}.
\]

Since \(\lambda = \alpha_j\) for some \(1 \leq j \leq d_\lambda\), this proves the first inequality.

To prove the second inequality, we first observe that both \(\Theta^K_\lambda(\exp K_C(X))\) and \(d_\lambda \sup_{w \in W_k} e^{\lambda(w.X)}\) considered as functions of \(X \in i t^*\) are invariant under the Weyl group \(W_k\). Thus we may assume that \(X \in - (\Delta^+_k)^* := \{Y \in i t^*: (\forall \alpha \in \Delta^+_k) \alpha(X) \leq 0\}\). Since \(\lambda(X) = \sup_{1 \leq j \leq d_\lambda} \alpha_j(X)\) whenever \(X \in - (\Delta^+_k)^*\), (2.1) implies that \(\Theta^K_\lambda(\exp K_C(X)) \leq d_\lambda e^{\lambda(X)}\), concluding the proof of (i).

(ii) This is a direct consequence of the Weyl Dimension Formula. \(\square\)

Corollary II.4. If \(\lambda \in HW(G, \Delta^+)\) is non-singular, then there exist a constant \(c > 0\) and \(n \in \mathbb{N}\) such that

\[
e^{\lambda(X)} \leq \Theta_\lambda(\exp(X)) \leq (c \| \lambda \|^n + c) \phi(X) \sup_{w \in W_k} e^{\lambda(w.X)}
\]

holds for all \(X \in i C^0_{\text{max}}\).

Proof. As \(1 = \inf_{X \in i C^0_{\text{max}}} \phi(X)\), the corollary follows from Lemma II.2 and Proposition II.3. \(\square\)

Lemma II.5. Let \(V\) be a finite dimensional real vector space, \(C \subseteq V\) a convex cone with non-empty interior, \(\Gamma \subseteq V\) a lattice and \(Q \subseteq V\) a compact subset. Then

\[
\Gamma(C, Q) := \{\gamma \in \Gamma: \gamma \in C, \gamma + Q \subseteq C\}
\]

is an additive subsemigroup of \(\Gamma\), \(R^+ \Gamma(C, Q)\) is a closed convex cone and the following equalities hold:

(i) \(R^+ \Gamma(C, Q) = \overline{C}\).

(ii) \(\Gamma(C, Q)^* = C^*\).
Proof. First we show that $\Gamma(C, Q)$ is an additive semigroup. Let $\gamma_1, \gamma_2 \in \Gamma(C, Q)$. Then $\gamma_1 + \gamma_2 \in C$ since $C$ is a convex cone and further we have for all $x \in Q$

$$\gamma_1 + \gamma_2 + x = \gamma_1 + (\gamma_2 + x) \in C + C = C,$$

proving that $\Gamma(C, Q)$ is an additive semigroup.

It follows in particular that $Q^+ \Gamma(C, Q)$ is an additive semigroup and hence the same holds for $\mathbb{R}^+ \Gamma(C, Q) = Q^+ \Gamma(C, Q)$. This proves that $\mathbb{R}^+ \Gamma(C, Q)$ is a closed convex cone.

(i) Since $\Gamma(C, Q) \subseteq C$ we obtain in particular that $\mathbb{R}^+ \Gamma(C, Q) \subseteq C$.

To prove the converse inclusion, we assume that $\mathbb{R}^+ \Gamma(C, Q) \neq C$. Then we find an open ball $B \subseteq V$ such that $B \subseteq C^0 \setminus \mathbb{R}^+ \Gamma(C, Q)$. Since $C^0 \setminus \mathbb{R}^+ \Gamma(C, Q)$ is an open cone, this implies in particular that

$$\forall \lambda > 0 \, \lambda B \cap \Gamma(C, Q) = \emptyset. \quad (2.2)$$

If $\lambda$ is sufficiently large, then we have $\lambda B \subseteq C - x$ for all $x \in Q$, because $B + \frac{1}{\lambda} x$ is contained in $C$ for sufficiently large $\lambda$ and $Q$ is compact. Let $\lambda_0 > 0$ such that $\lambda B \subseteq C - x$ for all $x \in Q$ and $\lambda > \lambda_0$.

In view of this and $\Gamma(C, Q) = \Gamma \cap C \cap \bigcap_{x \in Q} (C - x)$, (2.2) implies in particular that

$$\forall \lambda > \lambda_0 \, \Gamma \cap C \cap \lambda B = \Gamma(C, Q) \cap \lambda B = \emptyset;$$

a contradiction, concluding the proof of (i).

(ii) This follows from (i) and $\mathbb{R}^+ \Gamma(C, Q)^\star = \Gamma(C, Q)^\star$. \hfill \Box

Let $t_e := \{X \in t : \exp(X) = 1\}$ and note that $t_e$ is a lattice in span $\{t_e\}$. Hence we find a lattice $\Gamma \in it^\star$ which is contained in the set $\{\alpha \in it^\star : (\forall X \in t_e)(\alpha(X) \in 2\pi i \mathbb{Z})\}$. From now on we fix a lattice $\Gamma \subseteq it^\star$ having this property.

Lemma II.6. Let $W \subseteq \mathfrak{g}$ be a pointed closed $\text{Inn}(\mathfrak{g})$-invariant cone with non-empty interior, $\Delta^+$ a positive system satisfying $C_{\min} \subseteq C = W \cap t \subseteq C_{\max}$ and $\rho = \frac{1}{2} \sum_{\alpha \in \Delta^+} m_\alpha \alpha$. Then

$$\Gamma(\mathbb{i} \text{int } C^*, 2\rho) := \{\lambda \in \Gamma : \lambda \in \mathbb{i} \text{int } C^*, \lambda + 2\rho \in \mathbb{i} \text{int } C^*\}$$

consists of non-singular $G$-analytically integral elements.

Proof. First note that $\Gamma$ consists of $G$-analytically integral elements, hence the same holds for the subset $\Gamma(\mathbb{i} \text{int } C^*, 2\rho)$. We claim that $\Gamma(\mathbb{i} \text{int } C^*, 2\rho) \subseteq \Gamma(\mathbb{i} \text{int } C^*, \rho)$. In fact, Lemma II.5 implies that $\Gamma(\mathbb{i} \text{int } C^*, 2\rho)$ is an additive semigroup, so that $\lambda \in \Gamma(\mathbb{i} \text{int } C^*, 2\rho)$ implies that $2\lambda \in \Gamma(\mathbb{i} \text{int } C^*, 2\rho)$ which means that $2\lambda + 2\rho \in \mathbb{i} \text{int } C^*$ or equivalently $\lambda + \rho \in \Gamma(\mathbb{i} \text{int } C^*, \rho)$. This proves the claim.

Further, $C_{\min} \subseteq C$ implies that $\mathbb{i} \text{int } C^* \subseteq \mathbb{i} \text{int } C_{\min}^*$ and we therefore get

$$\Gamma(\mathbb{i} \text{int } C^*, 2\rho) \subseteq \{\lambda \in \Gamma : \lambda + \rho \in \mathbb{i} \text{int } C_{\min}^*\}.$$
In view of [Ne99, Ch. IX], this implies that all elements of $\Gamma(i \text{int } C^*, 2\rho)$ are non-singular.

Lemma II.7. Let $V$ be a finite dimensional real vector space, $C \subseteq V$ a convex pointed cone with non-empty interior, $\Gamma \subseteq V^*$ a lattice and $Q \subseteq V^*$ a compact subset. We fix a norm $\| \cdot \|$ on $V^*$ with $\|\gamma\| \geq 1$ for all $\gamma \in \Gamma \setminus \{0\}$ and consider for each $N \in \mathbb{N}_0$ the mapping

$$F^N : V \to \mathbb{R}^+ \cup \{\infty\}, \quad x \mapsto \sum_{\gamma \in \Gamma(\text{int } C^*, Q)} \|\gamma\|^N e^{-\gamma(x)}.$$

(i) For all $N \in \mathbb{N}_0$ the series defining $F^N$ converges compactly on $C^0$. In particular, $F^N|_{C^0}$ is continuous.

(ii) For all $N \geq 1$ we have $\lim_{x \to \partial C} F^N(x) = \infty$.

Proof. (i) Since $C^*$ is pointed we find for every $x \in C^0$ a constant $C_x > 0$ such that $\|\alpha\| \leq C_x \alpha(x)$ holds for every $\alpha \in C^*$. Thus we find for every compact subset $K \subseteq C^0$ a constant $C_K > 0$ such that

$$(\forall x \in K) (\forall \alpha \in C^*) \|\alpha\| \leq C_K \alpha(x).$$

This in turn implies that

$$\sup_{x \in K} \sum_{\gamma \in \Gamma(\text{int } C^*, Q)} \|\gamma\|^N e^{-\gamma(x)} \leq \sum_{\gamma \in \Gamma(\text{int } C^*, Q)} \|\gamma\|^N e^{-\frac{1}{C_K} \|\gamma\|} < \infty,$$

proving (i).

(ii) As $\|\gamma\| \geq 1$ for all $\gamma \in \Gamma \setminus \{0\}$, we have $F^1 \leq F^N$ for all $N \in \mathbb{N}$ and hence we only have to prove the assertion for $N = 1$. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in $C^0$ converging to $x \in \partial C$. Then there exists an element $0 \neq \alpha \in \partial C^*$ such that $\alpha(x) = 0$. Choose $\alpha_n \in \mathbb{R}^+ \alpha$ with $\alpha_n(x_n) = 1$. We claim that $\lim_{n \to \infty} \|\alpha_n\| = \infty$. Indeed, otherwise we find a number $L > 0$ and a subsequence $(\alpha_{n_k})_{k \in \mathbb{N}}$ such that $\alpha_{n_k} \in [0, L] \alpha$. But then

$$1 = \alpha_{n_k}(x_{n_k}) \leq L \alpha(x_{n_k}) \to 0$$

yields a contradiction and proves the claim.

Next we claim that there exists a constant $c > 0$ and elements $\gamma_n \in \Gamma(\text{int } C^*, Q)$ such that $|\gamma_n - \alpha_n| < c$ for all $n \in \mathbb{N}$. Let $r > 0$ such that $B(\beta, r) = \{\mu \in V^* : \|\beta - \mu\| \leq r\}$ intersects $\Gamma$ for all $\beta \in V^*$. Now choose $\mu \in \text{int } C^*$ such that $B(\mu + t\alpha, r) + Q \subseteq \text{int } C^*$ holds for all $t \geq 0$. Then we find elements $\gamma_n \in \Gamma(\text{int } C^*, Q) \cap B(\mu + \alpha_n, r)$. These elements $\gamma_n$ satisfy $|\gamma_n - \alpha_n| < c$ with $c = r + \|\mu\|$, proving our second claim.
Now we get
\[ F_1(x_n) = \sum_{\gamma \in \Gamma(\text{int} C^\star, Q)} \|\gamma\| e^{-\gamma(x_n)} \geq \|\gamma_n\| e^{-(\gamma_n(x_n) - \alpha_n(x_n))} e^{-\alpha_n(x_n)} \]
and so
\[ \lim_{n \to \infty} F_1(x_n) \geq \lim_{n \to \infty} \frac{1}{e} e^{-c\|x_n\| (\|\alpha_n\| - c)} = \infty. \]
This proves (ii).

\textbf{Theorem II.8.} Let \( S = \Gamma_G(W) \) be a pointed complex Ol’shanskii semi-group. Then for all \( N \in \mathbb{N} \) the prescription
\[ K^N : S^0 \times S^0 \to \mathbb{C}, \quad (z, w) \mapsto \sum_{\lambda \in \Gamma(i \text{int} C^\star, 2\rho)} \|\lambda\|^N \Theta_\lambda(zw^\star) \]
defines an element of \( \mathcal{P}_{G \times G}(S^0^2) \) satisfying
\[ \lim_{z \to \partial S^0} K^N(z, z) = \infty. \]

\textbf{Proof.} First we show that \( K^N \in \mathcal{P}_{G \times G}(S^0^2) \) for all \( N \in \mathbb{N} \). Since all kernels
\[ K^\lambda : S^0 \times S^0 \to \mathbb{C}, \quad (z, w) \mapsto \Theta_\lambda(zw^\star) \]
belong to \( \mathcal{P}_{G \times G}(S^0^2) \) (cf. \[ Ne94, \text{Th. IV.11} \]), we only have to show that the series defining \( K^N \) converges uniformly on compact subsets. In view of (1.2), a series of holomorphic positive definite kernels on a complex manifold converges compactly if and only if it converges uniformly on compact subsets on the diagonal. Therefore the bi-invariance of the kernels \( K^\lambda \) together with the Polar Decomposition of \( S^0 \) imply that it suffices to prove the compact convergence of the series defining the function
\[ \Theta^N : iW^0 \to \mathbb{R}^+, \quad X \mapsto \sum_{\lambda \in \Gamma(i \text{int} C^\star, 2\rho)} \|\lambda\|^N \Theta_\lambda(\text{Exp}(X)). \]
If \( C(W^0)^G \) denotes the \( \text{Ad}(G) \)-invariant continuous functions on \( W^0 \) endowed with the topology of uniform convergence on compact subsets, then \[ Ne96b, \text{Prop. III.6} \] entails that the restriction mapping \( C(W^0)^G \to C(C^0)^{W_0} \) is an isomorphism of Fréchet spaces. Thus we have reduced the problem to showing the compact convergence of the series defining \( \Theta^N |_{C^0} \). In view of Corollary II.4, we therefore have to prove that the series defined by
\[ F^N |_{C^0} : iC^0 \to \mathbb{R}^+, \quad X \mapsto \sum_{\lambda \in \Gamma(i \text{int} C^\star, 2\rho)} \|\lambda\|^N e^{\lambda(X)} \]
converges compactly. But this is exactly the contents of Lemma II.7(i).
In order to prove the second assertion, Proposition I.8(iii) implies that we only have to check that \( \lim_{X \to \partial C} K^N(X, X) = \infty \). According to Corollary II.4, this follows from \( \lim_{X \to \partial C} F^N(X) = \infty \) which is the contents of Lemma II.7(ii). This proves the theorem.

Corollary II.9. Let \( D \subseteq \Gamma_G(W_{\max}^0) \) be a bi-invariant Stein domain. Then there exists a kernel \( K \in \mathcal{P}_{G \times G}(D^2) \) such that
\[
\lim_{z \to \partial D} K(z, z) = \infty.
\]

Proof. In view of Proposition I.8(ii), we may assume that \( D_h = D_h^e \) is a cone, and hence Proposition I.8(iii) implies that we even may assume that \( D_h = D_{h,1}^e \) is a pointed cone. Now the corollary follows from Theorem II.8.

Remark II.10. One can modify \( \Gamma \) a little bit without affecting the contents Theorem II.8 as follows. If one takes \( \Gamma' = \Gamma \cup F \), where \( F \subseteq HW(G, \Delta^+) \) is a finite set of highest weights, then Theorem II.8 remains true with \( \Gamma \) replaced by \( \Gamma' \).

III. The equivariant embedding theorem.

In this section we apply the results of Section II to construct a kernel \( K \in \mathcal{P}_{G \times G}(D^2) \) such that the map
\[
e_K : D \mapsto \mathcal{H}^*_K \setminus \{0\}, \quad z \mapsto K_z
\]
defines a \( G \times G \)-equivariant holomorphic embedding with closed image.

Proposition III.1. Let \( M \) be a complex manifold, \( S \) an involutive semi-group acting on \( M \) by holomorphic maps and \( K \in \mathcal{P}_S(M^2) \). Then
\[
e_K : M \to \mathcal{H}^*_K, \quad z \mapsto K_z
\]
is an \( S \)-equivariant holomorphic map.

Proof. From the \( S \)-invariance of \( K \) it follows that
\[
e_K(s.z) = K_{s.z} = \pi_K(s).K_z = \pi_K(s).e_K(z)
\]
for all \( s \in S \) and \( z \in M \), proving the \( S \)-equivariance of the map \( e_K \). It remains to show that \( e_K \) is holomorphic. As \( \mathcal{H}_K \) is a Hilbert space, \( e_K \) is holomorphic if and only if the following conditions are satisfied:

1. \( e_K \) is locally bounded.
2. There exists a total subset \( T \subseteq \mathcal{H}_K \) such that the mappings \( M \to C, \ z \mapsto K_z(f) = \langle f, K_z \rangle \) are holomorphic for all \( f \in T \).
If \( Q \subseteq D \) is a compact subset, then we have

\[
\sup_{z \in Q} \|K_z\|^2 = \sup_{z \in Q} K(z, z) < \infty,
\]

proving (1). Finally \( T = \{K_w : w \in M\} \) is a total subset in \( \mathcal{H}_K \) and \( z \mapsto K_z(K_w) = \langle K_w, K_z \rangle = K(z, w) \) is holomorphic because \( K \in \text{Hol}(M \times M) \). This proves (2) and concludes the proof of the lemma. \( \square \)

**Corollary III.2.** If \( K \in \mathcal{P}_{G \times G}(D^2) \) is non-zero, then \( K_z \neq 0 \) for all \( z \in D \) and the mapping

\[
e_K : D \to \mathcal{H}_K^* \setminus \{0\}, \quad z \mapsto K_z
\]

is \( G \times G \)-equivariant and holomorphic.

**Proof.** It follows from [Ne97, Lemma III.6] that \( K_z \neq 0 \) for all \( z \in D \) and thus the map \( e_K \) is well defined. Now the assertions follow from Proposition III.1. \( \square \)

**Definition III.3.** A connected Lie group is called a (CA)-Lie group if \( \text{Ad}(G) \subseteq \text{Aut}(g) \) is closed. Note that all connected reductive and nilpotent Lie groups are (CA)-Lie groups. \( \square \)

The next lemma is our key observation. The proof depends heavily on the special choice of the lattice \( \Gamma(i \text{int} C^*, 2\rho) \) and is a little bit tricky.

**Lemma III.4.** Let \( S = \Gamma_G(W) \) be a pointed complex Ol’shanski˘ı semi-group and suppose that \( G \) is a (CA)-Lie group. Let \( K^N, N \in \mathbb{N}, \) be as in Theorem II.8 and \( (K^N_{s_n})_{n \in \mathbb{N}} \) a convergent sequence in \( \mathcal{H}_K^N \) with limit different from zero. Further let \( Z \) denote the center of \( G \).

(i) The set \( \{s_n Z : n \in \mathbb{N}\} \) is relatively compact in \( S^0/Z \).

(ii) If, in addition, \( Z \) is compact, then \( \{s_n : n \in \mathbb{N}\} \) is relatively compact in \( S^0 \).

**Proof.** (i) For simplicity, we write \( K \) instead of \( K^N \). Since \( K = \sum \lambda \|\lambda\|^N K^\lambda \) is a direct sum of positive definite kernels corresponding to inequivalent irreducible unitary representations of \( G \times G \), it follows in particular that \( \mathcal{H}_K = \bigoplus \mathcal{H}_K^\lambda \) (cf. [Ne99, Th. I.11, Rem. I.12(a)]). Thus \( (K_{s_n})_{n \in \mathbb{N}} \) being a convergent sequence with non-zero limit implies in particular that all sequences \( (K^\lambda_{s_n})_{n \in \mathbb{N}} \) are convergent and at least one limit \( f^\lambda = \lim_{n \to \infty} K^\lambda_{s_n} \) is different from zero.

Step 1: The set \( \{s_n Z : n \in \mathbb{N}\} \) is relatively compact in \( S/Z \).

Let \( \lambda \in \Gamma(i \text{int} C^*, 2\rho) \) be such that \( (K^\lambda_{s_n})_{n \in \mathbb{N}} \) converges with limit different from zero. Let

\[
\chi_{\lambda} : Z \to S^1, \quad z = \exp(X) \mapsto e^{-\lambda(X)}
\]

and note that \( \chi_{\lambda} \) is an element of \( \hat{Z} \).
The Bergman space corresponding to the character \( \chi_{\lambda} \) is defined as

\[
B^2(S/Z, \chi_{\lambda}) = \left\{ f \in \text{Hol}(S^0) : (\forall z \in Z, s \in S^0) f(sz) = \chi_{\lambda}(z)^{-1} f(s), \right. \\
\|f\|_2^2 := \int_{S^0/Z} \|f(s)\|^2 \, d\mu_{S^0/Z}(sZ) < \infty \}
\]

where \( \mu_{S^0/Z} \) denotes the canonical left \( S \)-invariant measure on \( S^0/Z \) (cf. \cite[Sect. II]{Kr98}). Recall from \cite[Prop. II.4, Th. IV.5]{Kr98} that \( B^2(S/Z, \chi_{\lambda}) \) is a closed subspace of the Hilbert space \( L^2(S/Z, \chi_{\lambda}) \) and that there exists a positive constant \( c > 0 \) such that the prescription

\[
\mathcal{H}_{K^\lambda} \to B^2(S/Z, \chi_{\lambda}), \quad K^\lambda \mapsto c K^\lambda
\]

defines an \( S \times S \)-equivariant isometric embedding.

We obtain in particular that \( (K^\lambda_n)_{n \in \mathbb{N}} \) is a convergent sequence in \( B^2(S/Z, \chi_{\lambda}) \) with limit \( f^\lambda \neq 0 \). To obtain a contradiction, we now assume that there exists a subsequence of \( (s_n Z)_{n \in \mathbb{N}} \) leaving every compact subset of \( S/Z \). To avoid further notation we denote this subsequence again by \( (s_n Z)_{n \in \mathbb{N}} \). Note that \( (s_n^* Z)_{n \in \mathbb{N}} \) also leaves every compact subset of \( S/Z \), since the involution on \( S \) induces an involution \( * : S/Z \to S/Z \). We write

\[
\rho : S \to B(B^2(S/Z, \chi_{\lambda})), \quad (\rho(s).f)(z) = f(sz)
\]

for the right regular representation of \( S \) on \( B^2(S/Z, \chi_{\lambda}) \) and note that \( (\rho, B^2(S/Z, \chi_{\lambda})) \) is a holomorphic contraction representation of \( S \) (cf. \cite[Prop. II.4]{Kr98}). Let \( s_0 \in S^0 \). Then \( \rho(s_0).K^\lambda_{s_n} \to \rho(s_0).f^\lambda \) and \( \rho(s_0).f^\lambda \neq 0 \) since \( \rho(s_0) \) is an injective operator. It follows in particular that there exists a convergent subsequence of \( (\rho(s_0).K^\lambda_{s_n})_{n \in \mathbb{N}} \) converging to \( \rho(s_0).f^\lambda \) pointwise. Note that \( |K^\lambda_{s_n}| \in C_0(S/Z) \, |_{S^0/Z} \) for all \( s \in S^0 \) (cf. \cite[Prop. II.4]{Kr98}). Thus we obtain from

\[
\rho(s_0).K^\lambda_{s_n}(z) = K^\lambda_{s_n}(zs_0) = K^\lambda(zs_0, s_n) = K^\lambda(z, s_ns_0^*)
\]

for all \( z \in S^0 \) that \( \rho(s_0).K^\lambda_{s_n} \to 0 \) pointwise. This is a contradiction to \( \rho(s_0).f^\lambda \neq 0 \) and proves our first step.

Step 2: Every cluster point of \( (s_n Z)_{n \in \mathbb{N}} \) lies in \( S^0/Z \).

As the Polar Decomposition of \( S \) is inherited by \( S/Z \), i.e., the mapping

\[
G/Z \times W \to S/Z, \quad (gZ, X) \mapsto g \text{Exp}(iX)Z
\]

is a homeomorphism, we can write \( s_n = g_n \text{Exp}(X_n) \), where \( g_n \in G \) and \( X_n \in iW^0 \). According to Step 1, we now may assume that both \( (g_n Z)_{n \in \mathbb{N}} \) and \( (X_n)_{n \in \mathbb{N}} \) converge. Let \( X \in iW \) be the limit of \( (X_n)_{n \in \mathbb{N}} \). Note that it suffices to show that \( X \in iW^0 \). As \( (K_{s_n})_{n \in \mathbb{N}} \) is convergent with non-zero
limit, \((\|K_{s_n}\|)^2)_{n \in \mathbb{N}}\) is a convergent sequence in \(\mathbb{R}^+\) with positive limit. The bi-invariance of \(K\) further implies that

\[
\|K_{s_n}\|^2 = K(s_n, s_n) = K(g_n \text{Exp}(X_n), g_n \text{Exp}(X_n)) \\
= K(\text{Exp}(X_n), \text{Exp}(X_n)),
\]

so that Step 2 follows from Theorem II.8.

(ii) This is a direct consequence of (i) and the compactness of \(Z\). □

**Lemma III.5.** If \(S = \Gamma_G(W)\) is a pointed complex Ol’shanski semigroup, then there exists a finite set \(F \subseteq HW(G, \Delta^+)\) such that all representations \(\pi_{K_N}: S \to B(\mathcal{H}_{K_N})\), \(N \in \mathbb{N}\), associated to \(\Gamma' = \Gamma \cup F\) (cf. Remark II.10) are injective.

**Proof.** Since \(S\) is pointed, it follows from [Kr98, Prop. V.7] that \(\pi_{K_N}\) is injective if and only if \(\pi_{K_N}|_T\) is injective, where \(T = \exp t\). Evaluation of the operators \(\pi_{K_N}(t), \ t \in T\), on highest weight vectors \(v_\lambda, \lambda \in \Gamma(i \text{nt } C^*, 2\rho)\), now easily shows how one can choose \(F\) to obtain injective representations. For more details we refer to [Ne96a, Sect. V]. □

**Theorem III.6.** Let \(D \subseteq \Gamma_G(W_{\text{max}}^0)\) be a pointed bi-invariant Stein domain and suppose that \(G\) is a (CA)-Lie group and \(Z\) is compact. Then there exists a kernel \(K \in P_{G \times G}(D^2)\) such that the mapping

\[e_K: D \to \mathcal{H}_K^* \setminus \{0\}, \quad z \mapsto K_z\]

defines a \(G \times G\)-equivariant holomorphic embedding with closed range.

**Proof.** In view of Proposition I.8(i) and Lemma I.6(ii), we may assume that \(D = S^0 = \Gamma_G(W^0)\) is a pointed open complex Ol’shanski semigroup. Now let \(\Gamma'\) as in Lemma III.5 and \(K = K_N\) for some \(N \in \mathbb{N}\). As \(K \neq 0\), the map \(e_K\) is a well defined \(G \times G\)-equivariant holomorphic map (cf. Corollary III.2). Lemma III.5 implies that the representation \(\pi_K: S \to B(\mathcal{H}_K)\) is injective so that \(\mathcal{H}_K\) separates points by [Kr98, Prop. V.10], which in turn means that \(e_K\) is injective.

Next we show that \(\text{im}e_K\) is closed. In fact, if \(K_{s_n} \to f \neq 0\), then Lemma III.4 implies that \((s_n)_{n \in \mathbb{N}}\) is a bounded sequence in \(S^0\) with all accumulation points in \(S^0\). Thus we find a convergent subsequence \((s_{n_k})_{k \in \mathbb{N}}\) with limit \(s \in S^0\). Now we get

\[f = \lim_{n \to \infty} K_{s_n} = \lim_{k \to \infty} K_{s_{n_k}} = \lim_{k \to \infty} e_K(s_{n_k}) = e_K(s) = K_s,
\]

proving the closedness of \(\text{im}e_K\).

Finally another easy application of Lemma III.4 shows that \(e_K\) is homeomorphic onto its image, concluding the proof of the theorem. □
Example III.7 (The Bergman kernel associated to \( \text{Sl}(2, \mathbb{R}) \)). Let \( G = \text{Sl}(2, \mathbb{R}) \) and \( \mathfrak{g} = \mathfrak{sl}(2, \mathbb{R}) \). We choose
\[
U = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\]
as a basis for \( \mathfrak{g} \).

Then \( t := \mathbb{R}U \) is a compactly embedded Cartan subalgebra. Let \( \alpha \in i t^* \) be defined by \( \alpha(U) = -2i \). The root system of \( \mathfrak{g} \) is given by \( \Delta = \{ \pm \alpha \} \) with root spaces \( \mathfrak{g}_\alpha^+ = \mathbb{C}(T + iH) \) and \( \mathfrak{g}_\alpha^- = \mathbb{C}(T - iH) \). We define a positive system by \( \Delta^+ := \{ \alpha \} \) and write \( \kappa \) for the Cartan-Killing form on \( \mathfrak{g} \).

Then the upper light cone
\[
W := \{ X = uU + tT + hH : u \geq 0, \kappa(X, X) \leq 0 \}
\]
is an invariant pointed cone in \( \mathfrak{g} \). Moreover, \( W \) is up to sign the unique invariant elliptic cone in \( \mathfrak{g} \) (cf. [HiNe93, Th. 7.25]). Thus up to isomorphism \( S := \Gamma_G(W) \) is the unique complex Ol’shanski˘ ı semigroup corresponding to \( G \).

In the following we identify \( t_{\mathbb{C}} \) with \( \mathbb{C} \) via the isomorphism \( t_{\mathbb{C}} \to \mathbb{C}, \lambda \mapsto \lambda(iU) \). Then \( HW(G, \Delta^+) = HW(G, W) = \{ \lambda \in \mathbb{Z} : \lambda \leq 0 \}, \lambda + 2\rho \in i \text{int} C^* \) if and only if \( \lambda \leq -3 \), and thus \( \Gamma(i \text{int} C^*, 2\rho) = \{ \lambda \in \mathbb{Z} : \lambda \leq -3 \} \). The main point is that \( \Gamma(i \text{int} C^*, 2\rho) \) coincides with the weights in the decomposition of the Bergman kernel \( B \) of the Bergman space \( B^2(S) := \left\{ f \in \text{Hol}(S^0) : \|f\|_2^2 = \int_{S^0} |f(s)|^2 \, d\mu_{G_{\mathbb{C}}}(s) < \infty \right\} \) (cf. [Kr98, Th. IV.7]). Further one knows that
\[
B = \sum_{\lambda \leq -3} \lambda(1 + \lambda)^2(4 - \lambda^2)K^\lambda
\]
(cf. [Kr98, Th. IV.7, Ex. IV.8]) so that Theorem II.8 and Theorem III.6 imply that
\[
\lim_{z \to \partial S} B(z, z) = \infty
\]
and that the map
\[
e_B : S^0 \to B^2(S)^* \setminus \{0\}, \quad s \mapsto B_s
\]
is a \( G \times G \)-equivariant holomorphic embedding with closed range. \( \square \)

Problems III.8. (a) What is the reason for that one has to exclude zero in \( \mathcal{H}_K^* \) to obtain the closedness of the map \( e_K \)?

(b) Given a pointed biinvariant Stein domain \( D \), does there exist an equivariant closed embedding of \( D \) into a complex topological vector space \( E \) endowed with a continuous \( G \times G \)-action?
References


[HiNe96] ______, Positive definite spherical functions on Ol’shanskiĭ domains, Mathematik-Bericht, 96/7, Technische Universität Clausthal.


Received September 22, 1997 and revised November 20, 1997.

TU Clausthal
Erzstrasse 1
D-38678 Clausthal-Zellerfeld
Germany
E-mail address: mabk@math.tu-clausthal.de
BASE CHANGE PROBLEMS FOR GENERALIZED WALSH SERIES AND MULTIVARIATE NUMERICAL INTEGRATION

Gerhard Larcher and Gottlieb Pirsic

We recall the notion of Walsh functions over a finite abelian group as it was given for example in Larcher, Niederreiter and Schmid, 1996. These function systems play an important role for various “digital lattice rules” in multivariate numerical integration. We consider the following problem:

Assume, that a function \( f \) can be represented by a Walsh-series over a group \( G_1 \) with a certain speed of convergence. Take another group \( G_2 \). What can be said about the speed of convergence of the Walsh-series of \( f \) over \( G_2 \)?

Answers to this question are essential for certain numerical integration error estimates. We are able to give some results, partly best possible ones.

A connection of the above problem to “digital differentiability” of functions and applications to numerical integration are given. Open problems are stated.

1. Introduction.

The classical Walsh function system \( \{ \text{wal}_n | n = 0, 1, 2, \ldots \} \), \( \text{wal}_n : [0, 1) \rightarrow \mathbb{C} \) (in the Paley enumeration) can be defined in the following way:

For a non-negative integer \( n \) and a real \( x \) in \([0, 1)\) let

\[
n = n_v \cdot 2^{v-1} + \cdots + n_1 \quad \text{and} \quad x = \frac{x_1}{2} + \frac{x_2}{2^2} + \cdots
\]

(with \( x_i \neq 1 \) for infinitely many \( i \)) be the digit representation in base 2.

Then

\[
\text{wal}_n (x) := (-1)^{x_1 \cdot n_1 + \cdots + x_v \cdot n_v}.
\]

The set \( \{ \text{wal}_n | n = 0, 1, 2, \ldots \} \) is a complete orthonormal function system in \( L^2 ([0, 1)) \). (See for example [23].)

There exist various generalizations of this concept in the literature. Chrestenson [1] studied Walsh functions in an arbitrary base \( b \). (“Representation in base 2” is replaced by “representation in base \( b \)” and “\(-1\)” is replaced by \( \left( 2 x_{\frac{v}{b}} \right) \).”) Vilenkin [26] introduced the now so called Vilenkin systems,
and Onneweer studied his concept of Rademacher and Walsh functions over groups [20].

Motivated by problems in multivariate quasi-Monte Carlo integration and also by investigations on the distribution of certain number-theoretical point sets (see [7], and also [6] and [8]), in [9] Walsh-function systems $W_{G,\varphi}$ over finite abelian groups $G$ of, say, order $b$, and with respect to certain bijections $\varphi$ between the “digits” $\{0, 1, \ldots, b-1\}$ and the group $G$, were introduced and used for applications. See also [21], [27], [10]. This concept contains the classical Walsh systems in any base $b$ (use $G = \mathbb{Z}_b$ and $\varphi$ the “identity”). In some sense our concept, which is presented in Definition 2 and in Definition 3 in the next section of this paper, is more general than the systems of Onneweer and of Vilenkin. In other aspects, however, their systems are more general than ours. For some details, see the examples in Section 2.

Solutions (or good estimates) concerning the following problem are of great interest for applications:

**Problem.** Let a finite group $G$ of order $b$ and a corresponding bijection $\varphi$ be given. Assume, that the function $f : [0, 1) \to \mathbb{C}$ can be represented by a series of Walsh-functions from $W_{G,\varphi} := \{G,\varphi \text{wal}_n | n = 0, 1, 2, \ldots \}$, say, $f(x) = \sum_{n=0}^{\infty} \hat{f}(n) \cdot G,\varphi \text{wal}_n(x)$. Assume further, that this Walsh series has a certain speed of convergence in the following way:

There are $\alpha > 1$ and $C > 0$, such that $|\hat{f}(n)| \leq C \cdot n^{-\alpha}$ for all $n = 1, 2, \ldots$. We say “$f$ belongs to $G,\varphi E^\alpha(C)$”.

Let now $H$ be another finite abelian group with, say, order $c$ and $\psi$ a corresponding bijection. We now ask: is there a $\beta > 1$ and a $C' > 0$, such that $f \in H,\psi E^\beta(C')$?

That is: We are looking for the following “base change coefficient”:

**Definition 1.** For given finite abelian groups $G, H$ and corresponding bijections $\varphi$ and $\psi$ and for $\alpha > 1$ let the base change coefficient $\beta(G, \varphi, H, \psi, \alpha)$ be defined by

$$
\beta(G, \varphi, H, \psi, \alpha) := \sup \{\beta > 1 \mid \text{for all } C > 0 \text{ there is a } C' > 0 \text{ with } G,\varphi E^\alpha(C) \subseteq H,\psi E^\beta(C')\}.
$$

(We set $\beta(G, \varphi, H, \psi, \alpha) = 1$ if there is no such $\beta$.)

Until now, an exact solution to this question was given only in [12] for the case $G = \mathbb{Z}_2, H = \mathbb{Z}_2^h$, and $\varphi, \psi$ identities. We obtained

$$
\beta(\mathbb{Z}_2, \text{id}, \mathbb{Z}_2^h, \text{id}, \alpha) = \alpha - \beta_h \quad \text{with}
$$

$$
\beta_h = \frac{h - 1}{2h} + \frac{\sum_{k=0}^{h-2} \log \sin \left( \frac{\pi}{4} + \frac{\pi}{2} \left( \frac{4^{h/2} - 1}{3 \cdot 2^k + 1} \right) \right)}{h \cdot \log 2}.
$$
We have $0 \leq \beta_h < 1/2$ for all $h$ and

$$\lim_{h \to \infty} \beta_h = \frac{1}{2} + \frac{\log \sin \frac{5\pi}{12}}{\log 2} = 0.4499 \ldots .$$

Partial results for $G = \mathbb{Z}_{2h}$ and $H = \mathbb{Z}_2$ (the problem is not “commutative”!) were given in [13].

It is the aim of this paper to give in some sense “exact” solutions to the above problem for many cases and good estimates for $\beta$ in all other cases.

Further we demonstrate a connection of our problem to questions concerning the “digital differentiability” of a function. This is done in analogy to the connection between the differentiability of functions in the usual sense and the speed of convergence of their Fourier-series. Finally we give applications of our results to error estimation in quasi-Monte Carlo integration.

In Section 2 we shall recall the concept of Walsh functions over groups and give their basic properties.

In Section 3 we shall recall the notion of the “digital derivative” and give the connection to our problem.

In Sections 4, 5 and 6 we give solutions to our problem, respectively estimates concerning the quantity $\beta$.

In Sections 4 and 5, especially, we give in some sense “exact” answers for the following cases (let $G$ be of order $b$ and let $H$ be of order $c$):

**Case 1:** $b \nmid c^N$ for all positive integers $N$.

**Case 2:** $b^M = c^N$ for some positive integers $M, N$.

In Section 6 we give estimates of $\beta$ for the remaining

**Case 3:** $b^c$ for some positive integer $N$ but $b^M \neq c^N$ for all positive integers $M$ and $N$.

In a short Section 7 for the sake of the reader we will summarize the results on the base change coefficient given in Sections 4, 5 and 6.

In Section 8, as a consequence of the results in Sections 4, 5 and 6 we give the main error estimate in the theory of “digital lattice rules” in its, until now, most general form.

Finally in Section 9 we state some of the most interesting open problems in the field.

**2. Walsh functions over groups.**

We recall the definitions for the concept of Walsh functions over a finite abelian group given in [9] and in [21].

Let $G$ be a finite abelian group of order $b$ . Let

$$G \cong \mathbb{Z}_{b_1} \times \cdots \times \mathbb{Z}_{b_m}$$

and

$$\varphi : \{0, 1, \ldots, b - 1\} \to G \cong \mathbb{Z}_{b_1} \times \cdots \times \mathbb{Z}_{b_m}$$
be a bijection with $\varphi(0) = 0$.

For $g = (g_1, \ldots, g_m) \in G$ let the character $\chi_g \in \hat{G}$ be defined by

$$\chi_g (y) := \prod_{l=1}^{m} e^{2\pi i \frac{g_l y_l}{b_l}}$$

for $y = (y_1, \ldots, y_m) \in G$. Of course $\hat{G} = \{\chi_g | g \in G\}$.

**Definition 2.** For a non-negative integer $n$ with $b$-adic representation $n = n_v \cdot b^{v-1} + \cdots + n_1$ we define the function $G,\varphi_{\text{wal}} n : [0,1) \to \mathbb{C}$ in the following way:

$$G,\varphi_{\text{wal}} n (x) := \prod_{j=1}^{v} \chi_{\varphi(n_j)} (\varphi(x_j)),$$

where $x = \sum_{j=1}^{\infty} x_j \cdot b^{-j}$ is the $b$-adic representation of $x$ (with infinitely many of the $x_j$ different from $b-1$).

**Definition 3.** The set

$$W_{G,\varphi} := \{G,\varphi_{\text{wal}} n | n = 0, 1, \ldots\}$$

is called the Walsh function system over $G$ with respect to $\varphi$.

In most cases we will write $G_{\text{wal}}$ instead of $G,\varphi_{\text{wal}}$, if it is clear which bijection $\varphi$ we use.

Note, that in classical Fourier theory, i.e. representation with respect to the characters $e^{2\pi ikx}$, $k \in \mathbb{Z}$ of the Torus group $\mathbb{T}$, also something similar to the bijections $\varphi$ occurs, as we have to identify the unit interval with the torus.

**Examples:**

a) Let $G = \mathbb{Z}_b$ and $\varphi$ be the “identity” between $\{0, \ldots, b-1\}$ and $\mathbb{Z}_b$. We then obtain the classical systems of Walsh-Paley and of Chrestenson.

b) In [20] Onneweer defines Walsh functions on the infinite product of groups. A “continuation” of his functions to $[0,1)$ is easily possible in an obvious way.

Our Walsh functions are products of characters on one fixed group. Onneweer’s functions are products of characters on possibly different groups. In this sense Onneweer’s concept is more general, however his groups must be of prime order, and he just uses identities for $\varphi$. In this sense our concept is more general.

c) The bijection $\varphi$ indeed strongly influences the structure of $W_{G,\varphi}$. Consider for example $W_{\mathbb{Z}_4,\text{id}}$ and $W_{\mathbb{Z}_4,\varphi}$ with the transposition $\varphi = (1 2)$ on $\mathbb{Z}_4$. The first four functions in each of the systems can be illustrated by
the following diagrams. (Different grey tones represent different function values and the same grey tones correspond to the same function values on each side.)

Figure 1. $\mathbb{Z}_4$, id $\omega_{n}, n = 0, 1, 2, 3$ and $\mathbb{Z}_4$, $\phi_{n}, n = 0, 1, 2, 3$.

d) Concerning the connection between our concept and Vilenkin systems, see [21]. Again we have the situation, that both concepts have a “non-empty intersection” but neither is a “sub-concept” of the other.

Of course one may ask, why we did not extend the investigations of this paper also to Onneweer systems and Vilenkin systems. There are two reasons for this: The problem studied in this paper was motivated by investigations on numerical integration by digital nets, initiated for example in [9]. The methods developed there were restricted to the classes $W_{G, \phi}$. They cannot be extended in a reasonable way to, say, Onneweer systems.

It should without problems be possible to obtain results for Onneweer and Vilenkin systems with the methods and results of this paper (or in an analogous way). However, since these systems are based on sequences of different groups of possibly different orders, the conditions for non-trivial connections between two different systems would be quite technical and quite restrictive. So we have concentrated ourselves in this paper to the more natural (and for our applications more important) classes $W_{G, \phi}$.

In the following, we will give some basic properties of Walsh functions over groups, which will be used later on.

For $G, \phi$ with $|G| = b$ given, we define digital summation $\oplus = \oplus_{G, \phi}$ on $\mathbb{R}_0^+$ in the following way: For $u, v \in \mathbb{R}_0^+$ let

$$u = \sum_{i=0}^{\infty} u_i \cdot b^{-i} \text{ and } v = \sum_{i=0}^{\infty} v_i \cdot b^{-i}$$
be the $b$-adic representations of $u$ and $v$. Then

$$u \oplus v := \sum_{i=w}^{\infty} z_i \cdot b^{-i}$$

with $z_i := \varphi^{-1}(\varphi(u_i) + \varphi(v_i))$ for $i = w, w + 1, \ldots$. Note that $\varphi(0) = 0$.

Further define

$$\ominus u := \sum_{i=w}^{\infty} \varphi^{-1}(-\varphi(u_i)) \cdot b^{-i},$$

which is the additive inverse of $u$, since $u \oplus (\ominus u) = 0$.

The following properties (compare these with the analogous properties of the function class $\{ e^{2\pi ikx}; k \in \mathbb{Z} \}$) are easily checked by insertion of the definitions. (Note, how the proper definition of digital summation enabled us to carry the character properties over to $\mathbb{R}_0^+$.)

**Lemma 1.** For $G, \varphi$ given, and $\oplus := \oplus_{G, \varphi}$ we have:

a) $G_{\text{wal}}(p) \cdot G_{\text{wal}}(q) = G_{\text{wal}}(p \oplus q)$ for all $p, q = 0, 1, 2, \ldots$ and $x \in [0, 1)$.

b) $G_{\text{wal}}(x) \cdot G_{\text{wal}}(y) = G_{\text{wal}}(x \oplus y)$ for all $n = 0, 1, 2, \ldots$ and $x, y \in [0, 1]$.

c) $G_{\text{wal}}(x) = \frac{1}{G_{\text{wal}}(x)} = G_{\text{wal}}(\ominus x)$.

We omit the easy proof.

The next lemma proves, that integrals of Walsh functions over certain intervals can be omitted.

**Lemma 2.** Let $m > 0$ and $n, a$ be integers with $b^{m-1} \leq n < b^m$ and $0 \leq a < b^{m-1}$ . Then

$$\int_{b^{m-1}}^{b^{m-1} \cdot \frac{a+1}{b^m-1}} G_{\text{wal}}(x) \, dx = 0.$$

Proof.

$$\int_{b^{m-1}}^{\frac{a+1}{b^m-1}} G_{\text{wal}}(x) \, dx = G_{\text{wal}}\left(\frac{a}{b^{m-1}}\right) \cdot \int_{b^{m-1}}^{\frac{1}{b^m-1}} G_{\text{wal}}(x) \, dx$$

and (with $n = \sum_{i=1}^{m} n_i \cdot b^i$ and $x = \sum_{i=m}^{\infty} x_i \cdot b^{-i}$)

$$\int_{0}^{b^{m-1}} G_{\text{wal}}(x) \, dx = \int_{0}^{b^{m-1}} \chi_{\varphi(n_m)}(\varphi(x_m)) \, dx$$

$$= \frac{1}{b^m} \cdot \sum_{g \in G} \chi_{\varphi(n_m)}(g) = 0$$

since $\varphi(n_m) \neq 0$. \qed
From Lemma 1 and Lemma 2 we immediately obtain the orthonormality of \( W_{G,\varphi} \):

**Lemma 3.**

\[
\int_0^1 G_{\text{wal}}_n(x) \, dx = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n \geq 1 \end{cases}.
\]

\[
\int_0^1 G_{\text{wal}}_n(x) \cdot G_{\text{wal}}_m(x) \, dx = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}.
\]

Indeed, \( W_{G,\varphi} \) also is a complete orthonormal system in \( L^2([0,1]) \). We do not use the completeness in the following. A proof can be found in [21].

For Walsh functions, some Dirichlet kernels take a very simple form:

**Lemma 4.** For any non-negative integer \( v \) we have

\[
\sum_{j=0}^{b^v-1} G_{\text{wal}}_j(x) = \begin{cases} b^v & \text{for } x \in [0, \frac{1}{b^v}) \\ 0 & \text{otherwise} \end{cases}.
\]

**Proof.** For a non-negative integer \( j \) with \( b \)-adic representation \( j = \sum_{i=1}^{v} j_i \cdot b^{i-1} \) let \( d_i(j) := j_i \) denote the \( i \)-th digit of \( j \).

Let \( x = \sum_{i=1}^{\infty} x_i \cdot b^{-i} \). Then

\[
\sum_{j=0}^{b^v-1} G_{\text{wal}}_j(x) = \sum_{j=0}^{b^v-1} \prod_{i=1}^{v} \chi_{\varphi(d_i(j))}(\varphi(x_i))
\]

\[
= \sum_{j_1, \ldots, j_v = 0}^{b-1} \prod_{i=1}^{v} \chi_{\varphi(j_i)}(\varphi(x_i))
\]

\[
= \prod_{i=1}^{v} \sum_{g \in G} \chi_{g}(\varphi(x_i))
\]

\[
= \begin{cases} b^v & \text{if } x_i = 0 \text{ for } i = 1, \ldots, v \\ 0 & \text{otherwise} \end{cases}.
\]

The result follows. \( \square \)

The next lemma provides a useful method to lower the index of a Walsh function, when it is of a special form:

**Lemma 5.** Let \( l, m, M \) be arbitrary positive integers. Let \( d := b^M \) and let \( j \) and \( n \) be integers with \( 0 \leq j, n < d \). Then

\[
G_{\text{wal}}_{j \cdot d^m-1}(\frac{n}{d^l}) = \begin{cases} 1 & \text{if } l \neq m \\ G_{\text{wal}}_j(\frac{n}{d}) & \text{if } l = m \end{cases}.
\]
Proof. For $M = 1$ we have

$$G_{\text{wal}}^{j,M-1} \left( \frac{n}{b} \right) = \begin{cases} 1 & \text{if } l \neq m \\ \chi_{\varphi(j)}(\varphi(n)) & \text{if } l = m \end{cases}.$$ 

If $m = l$ then moreover

$$\chi_{\varphi(j)}(\varphi(n)) = G_{\text{wal}}^j \left( \frac{n}{b} \right).$$

Therefore for $M \geq 2$ we get: Let $n = \sum_{i=1}^M n_i \cdot b^{i-1}$ and $j = \sum_{i=1}^M j_i \cdot b^{i-1}$, then

$$G_{\text{wal}}^{j,d_{M-1}} \left( \frac{n}{b} \right) = G_{\text{wal}}^{j_1} \left( \frac{n_{M}}{b} \right) \cdot G_{\text{wal}}^{j_2} \left( \frac{n_{M-1}}{b} \right) \cdots G_{\text{wal}}^{j_M} \left( \frac{n_1}{b} \right)$$

if $l = m$, and 1 if $l \neq m$ by Lemma 1 and since the result holds for $M = 1$.

On the other hand

$$G_{\text{wal}}^j \left( \frac{n}{b} \right) = G_{\text{wal}}^{j_1 + \cdots + j_M} \left( \frac{n_1 + \cdots + n_M \cdot b^{M-1}}{b^M} \right)$$

and the result follows. \(\square\)

A function $f : [0, 1) \to \mathbb{C}$ of the form

$$f(x) = \sum_{n=0}^{\infty} \hat{f}_G(n) \cdot G_{\text{wal}}^n(x)$$

with certain $\hat{f}_G(n) \in \mathbb{C}$, shall be called a Walsh series over $G$ with respect to $\varphi$.

In the following we will only deal with absolutely convergent Walsh series. Thus we will always have

$$\hat{f}_G(n) = \int_0^1 f(x) \cdot \overline{G_{\text{wal}}^n(x)} \, dx.$$ 

Recall the definition of the classes $G,\varphi E^\alpha(C)$ given in Section 1:

$$f \in G,\varphi E^\alpha(C) \iff |\hat{f}_G(n)| \leq C \cdot n^{-\alpha} \text{ for } n = 1, 2, \ldots.$$ 

3. Digital derivatives and speed of convergence of Walsh series.

In [20] a derivative for functions defined on groups $G$, that are the direct product of countably many groups of prime order, was introduced. Such functions can be continued to functions on $[0, 1)$ quite naturally in the same manner as it was done for our Walsh functions $W_{G,\varphi}$. In this sense the
following notion of a (digital) derivative with respect to $G$ and $\varphi$ is an analog to Onneweer's concept of a derivative.

**Definition 4.** For $G$ of order $b$, a corresponding bijection $\varphi$, a function $f : [0, 1) \to \mathbb{C}$, for $x \in [0, 1)$ and positive integers $n$ let

$$d_n f (x) := \sum_{j=0}^{n} b^{j-1} \sum_{k=0}^{b-1} \sum_{l=0}^{b-1} g_{\varphi,wal}(\frac{k}{b}) \cdot f \left( x \bigoplus \frac{l}{b^{j+1}} \right).$$

We say, that $f$ is digitally differentiable in $x$ with respect to $G$ and $\varphi$ if

$$f^{[1]} (x) := \lim_{n \to \infty} d_n f (x)$$

exists. $f^{[1]} (x)$ is called the digital derivative of $f$ in $x$.

Higher derivatives $f^{[n]} (x)$; $n = 2, 3, \ldots$ can be defined in the usual way.

A function $f \in L_p ([0, 1))$, $1 \leq p < \infty$ is called strongly differentiable with respect to $G$ and $\varphi$ if there exists $g \in L_p ([0, 1))$ with

$$\lim_{n \to \infty} \|d_n f - g\|_p = 0.$$ 

We then denote $g$ by $df$.

**Example.** As an illustration for the difficulty in combining this concept of a derivative with "geometric intuition", consider for example the function $f := 1_{[0, 1/8)}$, the indicator function of the interval $[0, 1/8)$, and its derivative with respect to $\mathbb{Z}_2$ and identity.

![Figure 2: The indicator function of [0, 1) and its digital derivative with respect to $\mathbb{Z}_2$ and id.](image)

We may state the following question: Assume, that $f : [0, 1) \to \mathbb{C}$ is differentiable $\alpha$ times with respect to $G$ and $\varphi$. How often is $f$ differentiable with respect to another group $H$ with corresponding bijection $\psi$?

Or: Given $G, \varphi, H, \psi$ as above and a positive integer $\alpha$. What can be said about the quantity

$$\gamma (G, \varphi, H, \psi, \alpha) := \sup \{ \gamma \mid \text{if } f \text{ is } \alpha \text{ times differentiable with respect to } G \text{ and } \varphi, \text{ then } f \text{ is at least } \gamma \text{ times differentiable with respect to } H \text{ and } \psi \}?$$
A certain connection between \( \beta(G, \varphi, H, \psi, \alpha) \) as defined in Section 1 and \( \gamma(G, \varphi, H, \psi, \alpha) \) is given by the subsequent Theorem 1.

First we need two further lemmata. They already show, that the digital derivative plays the same role for Walsh functions over groups as the usual derivative does for exponential function, i.e. for Fourier analysis.

**Lemma 6.** For \( G \) of order \( b \) and \( \varphi \) fixed, \( G_{\text{wal}} \) is strongly differentiable with \( d(G_{\text{wal}}) = w \cdot G_{\text{wal}} \).

**Proof.** Let \( 0 \leq l < b \). Let \( w = \sum_{i=1}^{\infty} w_i \cdot b^{i-1} \) and \( j \) be a non-negative integer. Then

\[
G_{\text{wal}} \left( \frac{l}{b^j+1} \right) = \prod_{v=1}^{m} e^{2\pi i \frac{\varphi_v(w_j+1) \varphi_v(0)}{b}} = G_{\text{walt}} \left( \frac{w_{j+1}}{b^j} \right).
\]

(Here we used the representation \( G \cong \mathbb{Z}_{b_1} \times \cdots \times \mathbb{Z}_{b_m} \) for \( G \) again.)

So

\[
\begin{align*}
\hat{d}_n(G_{\text{wal}}(x)) &= \sum_{j=0}^{n} b^{j-1} \sum_{k=0}^{b-1} \sum_{l=0}^{b-1} \hat{G_{\text{walt}}}(k \overline{b}) \cdot G_{\text{wal}} \left( x \oplus \frac{l}{b^j+1} \right) \\
&= G_{\text{wal}}(x) \cdot \sum_{j=0}^{n} b^{j-1} \sum_{k=0}^{b-1} \sum_{l=0}^{b-1} \hat{G_{\text{walt}}}(w_{j+1} \oplus k) = w \cdot G_{\text{wal}}(x)
\end{align*}
\]

by Lemma 4 and for all \( n \) large enough. \( \square \)

**Lemma 7.** If \( f : [0,1) \to \mathbb{R} \) is strongly differentiable, then \( \hat{d}_f G(m) = m \cdot \hat{f}_G(m) \) for all \( m \).

**Proof.** Note that for positive integers \( j, l \) we have

\[
\int_0^1 f \left( x \oplus \frac{l}{b^j+1} \right) \cdot G_{\text{wal}_m}(x) \, dx = \int_0^1 f(x) \cdot G_{\text{wal}_m} \left( x \oplus \frac{l}{b^j+1} \right) \, dx.
\]

So

\[
\begin{align*}
\hat{d}_n(f_G(m)) &= \int_0^1 \sum_{j=0}^{n} b^{j-1} \sum_{k=0}^{b-1} \sum_{l=0}^{b-1} \hat{G_{\text{walt}}}(k \overline{b}) \cdot f \left( x \oplus \frac{l}{b^j+1} \right) \cdot G_{\text{wal}_m}(x) \, dx \\
&= \int_0^1 f(x) \cdot G_{\text{wal}_m}(x) \, dx \cdot \sum_{j=0}^{n} b^{j-1} \sum_{k=0}^{b-1} \sum_{l=0}^{b-1} \hat{G_{\text{walt}}}(w_{j+1} \oplus k)
\end{align*}
\]
\[ \hat{f}_G(m) \cdot \sum_{j=0}^{n} b^{j-1} \sum_{k=0}^{b-1} k \sum_{l=0}^{b-1} G_{\text{wal}} \left( \frac{m_{j+1} \oplus k}{b} \right) = 0 \]

with \( m = \sum_{j=1}^{\infty} m_j \cdot b^{j-1} \), by Lemma 4, for all \( n \) large enough. \( \square \)

As is the aim of this section, a certain connection between the quantities \( \beta \) and \( \gamma \) is now given by the following result. (Compare with analogous results for Fourier series as they appear for example in [5] in a higher dimensional form.)

**Theorem 1.**

- **a)** Let \( \alpha > 1 \) be an integer. If a Walsh series \( f \) over \( G \) and \( \varphi \) is \( \alpha \) times strongly differentiable as a function in \( L^1(\mathbb{R}) \), then \( f \in G, G^E \alpha (C) \) for some \( C > 0 \).

- **b)** If for some integer \( \alpha > 2 \) we have \( f \in G, G^E \alpha (C) \) then \( \sum_{n=0}^{\infty} n^{\alpha-2} \hat{f}_G(n) \cdot G_{\text{wal}}(x) \) is absolutely convergent and therefore, by Lemma 7, the strong \( \alpha \)-th derivative of \( f \) with respect to \( G \) and \( \varphi \).

\( \square \)

Consequently we have, for example, the following relation between the quantities \( \beta \) and \( \gamma \): If \( \alpha > 1 \) and \( \beta(G, \varphi, H, \psi, \alpha) > 1 \), then \( \gamma(G, \varphi, H, \psi, \alpha) \leq \beta(G, \varphi, H, \psi, \alpha) \).

**4. The base change coefficient in the case:**

\[ |G| \not| |H|^N \text{ for all positive } N. \]

In this case there is no “convergence connection” between the Walsh series representations over \( G \) and \( H \). We have the following result.

**Theorem 2.** If \( G, H \) are such, that for all positive integers \( N \) we have \( |G| \not| |H|^N \), then \( \beta(G, \varphi, H, \psi, \alpha) = 1 \) for all \( \alpha > 1 \).

**Proof.** With \( b := |G| \), let

\[ f(x) := \sum_{i=0}^{b-1} G_{\text{wal}}(x) = \begin{cases} b & \text{for } x \in [0, 1/b) \\ 0 & \text{otherwise.} \end{cases} \]
So for all \( \alpha > 1 \) we have \( f \in \mathbb{G} E^\alpha (b^\alpha) \). We show, that for all \( \epsilon > 0 \) and all \( K > 0 \), \( f \notin \mathbb{H} E^{1+\epsilon} (K) \).

Let \( c := |H| \) and \( 1/b = \sum_{i=1}^{\infty} \kappa_i \cdot c^{-i} \) be the representation of \( 1/b \) in base \( c \). This representation is periodic and non-terminating. So for \( k = \sum_{i=1}^{m} \kappa_i \cdot c^{i-1} \) we have

\[
\hat{f}_H (k) = \int_0^{1/b} \overline{H \text{wal}_k (x)} \cdot \overline{H \text{wal}_n (x)} \, dx = \int_0^{1/b} \overline{H \text{wal}_n (x)} \, dx
\]

by Lemma 2. Therefore

\[
\hat{f}_H (k) = \overline{H \text{wal}_k} \left( \sum_{i=1}^{m-1} \frac{\kappa_i}{c^i} \right) \cdot \int_0^{\sum_{i=1}^{\infty} \kappa_i \cdot c^{-i}} \overline{H \text{wal}_n} (x) \, dx
\]

\[
= \overline{H \text{wal}_k} \left( \sum_{i=1}^{m-1} \frac{\kappa_i}{c^i} \right) \cdot \left( \frac{1}{c^m} \sum_{j=0}^{\kappa_m-1} \overline{\chi_{\varphi(k_m)}} (\varphi(j)) + \sum_{i=m+1}^{\infty} \frac{\kappa_i}{c^i} \cdot \overline{\chi_{\varphi(k_m)}} (\varphi(k_m)) \right) =: \frac{1}{c^m} \cdot T (k).
\]

Since the sequence of the \( \kappa_i \) is periodic, \( T (k) \) attains only finitely many different values.

A finite Walsh polynomial over \( H \) is certainly continuous in \( x = 1/b \) (the number \( b \) contains prime factors not dividing \( c! \)) and therefore cannot represent \( f \).

So there is a \( T > 0 \) with \( T = |T (k)| \) for infinitely many \( k \). And for these \( k \) we have

\[
|\hat{f}_H (k)| \geq \frac{1}{c^m} \cdot T \geq \frac{T}{c} \cdot \frac{1}{k}.
\]

The result follows.

\[\square\]

5. The base change coefficient in the case: \( |G|^M = |H|^N \) for some positive \( M \) and \( N \).

In this case the “convergence connection” is non-trivial and we can give the exact form of \( \beta (G, \varphi, H, \psi, \alpha) \). This quantity \( \beta \) shows great similarities to a quantity studied in [27], which measures a certain distance between groups.

Before we state the main results in Theorem 3 and Theorem 4, we give some auxiliary technical results, which will be needed in the proofs. The first lemma in this section establishes the fact, that in the currently considered case the scalar product of Walsh functions is non-zero only for constrained index ranges.

**Lemma 8.** Let \( G, H \) be groups with order \( b \) and \( c \). Let the corresponding bijections \( \varphi \) and \( \psi \) be fixed. Assume, that there exist \( v, w \geq 1 \), such that
for all $j, k$ with $b^v - 1 \leq j < b^v$ we have

$$\gamma(j, k) := \int_0^1 G_{\text{wal}}(x) \cdot H_{\text{wal}}(x) \, dx = 0.$$  

Proof. Note that $G_{\text{wal}}(x)$ is constant on $[b^v l/b^v, l+1/b^v)$ for all $0 \leq l < b^v$. So by Lemma 2 and with $A := c^w/b^v$ we have

$$\gamma(j, k) = \sum_{l=0}^{b^v-1} G_{\text{wal}}(l/b^v) \cdot \int_{A c^w l + A}^{A c^w l + A} H_{\text{wal}}(x) \, dx = 0.$$  

□

In the next lemma we see, how we can estimate $\beta(G, \varphi, H, \psi, \alpha)$ by estimating certain sums involving the $\gamma(j, k)$ defined in the last lemma.

Let $G, H, \varphi, \psi, b, c$ be as above ($b^v | c^w - 1$ for some positive $v, w$) and let $\gamma$ be defined like in Lemma 8.

For non-negative integers $m, k$ let

$$P(m) := \min \{n : b^m | c^n\}$$

$$Q(m) := \min \{n : k < c^{P(n)}\}.$$  

Lemma 9. Let $\alpha, \beta > 1$ and $C_1, C_2 > 0$ be reals and $f \in GE^\alpha(C_1)$, such that

$$\left| \sum_{j=b^Q(k)-1}^{\infty} \hat{f}_G(j) \cdot \gamma(j, k) \right| < \frac{C_2}{k^\beta}$$

for all $k \geq 1$. Then for all $x \in [0, 1)$, we have

$$f(x) = \hat{f}_G(0) + \sum_{k=1}^{\infty} \left( \sum_{j=b^Q(k)-1}^{\infty} \hat{f}_G(j) \cdot \gamma(j, k) \right) \cdot H_{\text{wal}},$$

therefore

$$\hat{f}_H(k) = \sum_{j=b^Q(k)-1}^{\infty} \hat{f}_G(j) \cdot \gamma(j, k)$$

for $k > 0$ and $f \in HE^\beta(C_2)$.

Proof. By Lemma 8 for $j$ with $b^n - 1 \leq j < b^n$ we have

$$G_{\text{wal}}(x) = \sum_{k=0}^{c^{P(m)} - 1} \gamma(j, k) \cdot H_{\text{wal}}(x).$$
So (since \( \gamma(j,0) = 0 \) for \( j > 1 \) )

\[
f(x) = \sum_{j=0}^{\infty} \hat{f}_G(j) \cdot G_{\text{wal}}(x)
\]

\[
= \hat{f}_G(0) + \sum_{m=1}^{\infty} \sum_{b^m-1}^{c^{P(m)}-1} \hat{f}_G(j) \sum_{k=0}^{b^m-1} \gamma(j,k) \cdot H_{\text{wal}}(x)
\]

\[
= \hat{f}_G(0) + \sum_{k=1}^{\infty} \sum_{m, k \geq b^{P(m)}+1}^{k} \left( \sum_{j=b^m-1}^{b^m-1} \hat{f}_G(j) \cdot \gamma(j,k) \right) \cdot H_{\text{wal}}(x)
\]

\[
= \hat{f}_G(0) + \sum_{k=1}^{\infty} \left( \sum_{j=b^{Q(k)}+1}^{\infty} \hat{f}_G(j) \cdot \gamma(j,k) \right) \cdot H_{\text{wal}}(x).
\]

\[\square\]

For the remaining part of this chapter, let \( M, N \) be positive integers, such that \( b^M = c^N =: d \). We show, that in this case, the \( \gamma(j,k) \) can be evaluated in terms of \( \gamma(j_l,k_l) \) with only finitely many indices \( j_l, k_l \).

**Lemma 10.** Let \( j, k \) be positive integers with \( j, k < d^L \) and let \( j = \sum_{l=1}^{L} j_l \cdot d^{l-1}, k = \sum_{l=1}^{L} k_l \cdot d^{l-1} \) be their \( d \)-adic representations. Then

\[
\gamma(j,k) = \prod_{l=1}^{L} \gamma(j_l,k_l).
\]

**Proof.** In the following we use Lemma 5:

\[
\gamma(j,k) = \int_{0}^{1} G_{\text{wal}}(x) \cdot \overline{H_{\text{wal}}(x)} \, dx
\]

\[
= \frac{1}{d^L} \sum_{n=0}^{d^L-1} G_{\text{wal}}(\frac{n}{d^L}) \cdot \overline{\frac{n}{d^L}}
\]

\[
= \frac{1}{d^L} \sum_{n_1, \ldots, n_L=0}^{d-1} \left( \prod_{m=1}^{L} G_{\text{wal},d^{m-1}}(\frac{n_m}{d^m}) \right)
\]

\[
\cdot \left( \prod_{m=1}^{L} \overline{H_{\text{wal},d^{m-1}}(\frac{n_m}{d^m})} \right)
\]
\[= \prod_{m=1}^{L} \left( \frac{1}{d} \sum_{n=0}^{d-1} G_{\text{wal}}_{jm \cdot d^{m-1}} \left( \frac{n}{d^m} \right) \cdot \overline{H_{\text{wal}}}_{km \cdot d^{m-1}} \left( \frac{n}{d^m} \right) \right)\]

\[= \prod_{m=1}^{L} \left( \frac{1}{d} \sum_{n=0}^{d-1} G_{\text{wal}}_{jm} \left( \frac{n}{d} \right) \cdot \overline{H_{\text{wal}}}_{km} \left( \frac{n}{d} \right) \right) = \prod_{l=1}^{L} \gamma (j_l, k_l). \]

We define a quantity, using those finitely many \(\gamma(j, k)\), which will help us in estimating the sum of the above Lemma 9.

**Definition 5.** Let \(G\) and \(H\) be finite abelian groups of order \(b\) and \(c\). Let \(\varphi, \psi\) be corresponding bijections. Assume, there are positive integers \(M, N\), such that \(b^M = c^N \vdash d\). Then

\[\beta_{G, H, \varphi, \psi} := \log_d \left( \max_{k=0, \ldots, d-1} \sum_{j=0}^{d-1} |\gamma(j, k)| \right). \]

(log \(d\) denotes logarithm to base \(d\).)

The quantity \(\beta_{G, H, \varphi, \psi}\) also was studied in [27].

The following theorem is the main result of this section:

**Theorem 3.** Let \(G\) and \(H\) be finite abelian groups of order \(b\) and \(c\). Let \(\varphi, \psi\) be corresponding bijections. Assume, there are positive integers \(M, N\), such that \(b^M = c^N \vdash d\). Then for all \(\alpha > 1 + \beta_{G, H, \varphi, \psi}\) we have

\[\beta(G, H, \varphi, \psi, \alpha) = \alpha - \beta_{G, H, \varphi, \psi}. \]

**Proof.** Let \(f \in G, \varphi E^\alpha (C)\). Let \(d^{L-1} \leq k < d^L, k = \sum_{i=1}^{L} k_i \cdot d^{i-1}\). By Lemma 8 we have \(\gamma(j, k) = 0\) if \(j < d^{L-1}\) or \(j \geq d^L\). By \(l(j)\) we denote the \(l\)-th digit of a non-negative integer \(j\) in base \(d\). Then by Lemma 9 and Lemma 10:

\[\left| \hat{f}_H(k) \right| = \sum_{j=0}^{d^{L-1}} \hat{f}_G(j) \cdot |\gamma(j, k)| \]

\[= \sum_{j=0}^{d^{L-1}} \hat{f}_G(j) \cdot \left( \prod_{l=1}^{L} \gamma(l(j), k_l) \right) \]

\[\leq \frac{C}{d(L-1)\alpha} \sum_{j_1, \ldots, j_L=0}^{d-1} \prod_{l=1}^{L} |\gamma(l(j), k_l)| \]

\[\leq \frac{C}{d(L-1)\alpha} \left( \max_{k=0, \ldots, d-1} \sum_{j=0}^{d-1} |\gamma(j, k)| \right)^L \]
\[ \leq C \cdot d^\alpha \cdot k^{\beta_{G,H,\varphi,\psi} - \alpha} \]

and therefore (again using Lemma 9) \( f \in H,\psi E^{\alpha - \beta_{G,H,\varphi,\psi}} (C \cdot d^\alpha) \) so that
\( \beta (G,H,\varphi,\psi,\alpha) \geq \alpha - \beta_{G,H,\varphi,\psi} \).

On the other hand, consider the function
\[ f (x) = \sum_{j=1}^{\infty} \hat{f}_G (j) \cdot G,\varphi \text{wal}_j (x) \]
with \( \hat{f}_G (j) \) of the following form: Let \( k_0 \) be such, that
\[ \sum_{j=0}^{d-1} |\gamma (j,k_0)| = \max_{k=0,\ldots,d-1} \sum_{j=0}^{d-1} |\gamma (j,k)| . \]

For \( L \geq 1 \) let \( K (L) := k_0 + k_0 \cdot d + \cdots + k_0 \cdot d^{L-1} \) and for \( d^{L-1} \leq j < d^L \) we set
\[ \hat{f}_G (j) := \frac{1}{d^{L \alpha}} \frac{|\gamma (j,K (L))|}{\gamma (j,K (L))} . \]

Then \( f \in G,\varphi E^{\alpha} (1) \).

Further for all \( L \geq 1 \) we have
\[
\begin{align*}
\left| \hat{f}_H (K (L)) \right| &= \left| \sum_{j=0}^{d^L-1} \hat{f}_G (j) \cdot \gamma (j,K (L)) \right| \\
&= \frac{1}{d^{L \alpha}} \sum_{j_{1,\ldots,j_L=0}^{d^L-1}} \prod_{l=1}^{L} |\gamma (j_l,k_0)| \\
&= \frac{1}{d^{L \alpha}} \prod_{l=1}^{L} \sum_{j=0}^{d^L-1} |\gamma (j,k_0)| \geq \frac{1}{d^{\alpha - \beta_{G,H,\varphi,\psi}}} \cdot \frac{1}{K(L)^{\alpha - \beta_{G,H,\varphi,\psi}}} .
\end{align*}
\]

so that \( \beta (G,\varphi,H,\psi,\alpha) \leq \alpha - \beta_{G,H,\varphi,\psi} \), and the result follows. \( \square \)

The constant \( \beta_{G,H,\varphi,\psi} \) is computable in finitely many steps for every \( G,H,\varphi,\psi \), and, as already mentioned in Section 1, \( \beta_{G,H,\varphi,\psi} \) was explicitly computed for \( G = \mathbb{Z}_2, H = \mathbb{Z}_2 \) and \( \varphi,\psi \) identities in [12].

In the following we give an estimate for the quantity \( \beta (G,H,\varphi,\psi,\alpha) \) based on Theorem 3.

Let \( G,H,\varphi,\psi,d \) be as above. We say \( G \) and \( H \) are of comparable order with common multiple \( d = |G|^M = |H|^N \). We define the bijection \( \tau : G^M \rightarrow H^N \) "induced by \( \varphi \) and \( \psi \): Let
\[ \bar{\varphi} : \{0,\ldots,b^M - 1\} \rightarrow G^M, \]
\[ x = a_M \cdot b^M - 1 + \cdots + a_1 \mapsto \bar{\varphi} (x) := (\varphi (a_M),\ldots,\varphi (a_1)) . \]
Let
\[ \bar{\psi} : \{0, \ldots, c^N - 1\} \rightarrow H^N, \]
\[ y = e_N \cdot c^{N-1} + \cdots + e_1 \mapsto \bar{\psi}(y) := (\psi(e_N), \ldots, \psi(e_1)). \]
Then \( \tau := \bar{\psi} \circ \bar{\phi}^{-1}. \)

**Theorem 4.** Let \( G \) and \( H \) be of comparable order. Let \( \varphi \) and \( \psi \) be corresponding bijections and \( \tau \) the bijection induced by \( \varphi \) and \( \psi \). Then for all \( \alpha > 1 \) we have
\[ \alpha - 1/2 < \beta(G, \varphi, H, \psi, \alpha) \leq \alpha \]
with equality on the right side if and only if \( \tau \) is a group isomorphism.

**Proof.** We have \( \beta_{G,H,\varphi,\psi} \geq 0 \) by definition. By Theorem 1 in [27], \( \beta_{G,H,\varphi,\psi} = 0 \) if and only if \( \tau \) is a group isomorphism. From Theorem 3 the right side of the inequality follows.

Now let \( \gamma(l) := (\gamma(0, l), \ldots, \gamma(d - 1, l)) \in \mathbb{C}^d. \) By Cauchy’s inequality we have
\[ \beta_{G,H,\varphi,\psi} = \log_d \left( \max_{l=0, \ldots, d-1} \|\gamma(l)\|_1 \right) \leq \log_d \left( \max_{l=0, \ldots, d-1} \sqrt{d} \cdot \|\gamma(l)\|_2 \right) \]
\[ = \log_d \sqrt{d} = 1/2. \]
Moreover we never have \( \beta_{G,H,\varphi,\psi} = 1/2, \) since \( \gamma(0, l) = 0 \) for \( l > 0 \) and \( \|\gamma(0)\|_1 = 1, \) so that in the above application of Cauchy’s inequality the inequality is strict. Again from Theorem 3 the result follows. \( \square \)

**6. The base change coefficient in the case:**
\( |G| \) divides \( |H|^N \) for some positive \( N, \) but
\( |G|^M \neq |H|^N \) for all positive \( M \) and \( N. \)

In this section we will always assume, that the orders \( b \) and \( c \) of the groups \( G \) and \( H \) satisfy \( b|c^N \) for some positive integer \( N, \) that is: Let \( c \) have canonical prime factorization \( c = \prod_{i=1}^r p_i^{\nu_i} \) with \( \nu_i \geq 1 \) for \( i = 1, \ldots, r \), then \( b = \prod_{i=1}^r p_i^{\mu_i} \) with \( \mu_i \geq 0 \) for \( i = 1, \ldots, r. \)

**Definition 6.** Let \( \rho = \rho(b,c) := \min \{ \nu_i/\mu_i : i = 1, \ldots, r \}. \)

**Definition 7.** Let
\[ \sigma_w := \max_{c^{w-1} \leq k < c^w} \sum_{w=(w-1)\rho+1}^\infty \frac{1}{b^v\alpha} \sum_{j=b^{v-1}} b^v-1 \left| \gamma(j,k) \right| \]
and
\[ \lambda = \lambda(G, \varphi, H, \psi, \alpha) := -\limsup_{w \to \infty} \frac{\log_c \sigma_w}{w}. \]
With this notation, we have:

**Theorem 5.** If \( \lambda(G, \varphi, H, \psi, \alpha) > 1 \) then

\[
\beta(G, \varphi, H, \psi, \alpha) = \lambda(G, \varphi, H, \psi, \alpha).
\]

**Proof.** For \( f \in G, \varphi E^\alpha(C) \) and for \( k \) with \( c^{w-1} \leq k < c^w \) and since \( \lambda > 1 \) we have by Lemma 9:

\[
\left| \hat{f}_H(k) \right| = \left| \sum_{j=b_Q(k)-1}^\infty \hat{f}_G(j) \cdot \gamma(j, k) \right|
\leq C \cdot b^\alpha \cdot \sum_{v=Q(k)}^\infty \frac{1}{b^{v\alpha}} \sum_{j=b^{v-1}}^{b^v-1} |\gamma(j, k)|.
\]

Remember that \( P(m) := \min \{ n : b^m | c^n \} \), \( Q(k) := \min \{ n : k < c^{P(n)} \} \), so

\[
P(m) = \left\lceil m \cdot \max_{i=1, \ldots, r} \nu_i \right\rceil
\]
and therefore

\[
Q(k) = Q(c^{w-1}) = \min \left\{ n : w \leq \left\lceil m \cdot \max_{i=1, \ldots, r} \frac{\mu_i}{\nu_i} \right\rceil \right\}
= \lfloor (w - 1) \cdot \rho \rfloor + 1.
\]
(Here by \( \lfloor y \rfloor \) we denote the smallest integer larger or equal to \( y \).) Consequently

\[
\left| \hat{f}_H(k) \right| \leq C \cdot b^\alpha \cdot \sigma_w \leq C'( \epsilon) \cdot \frac{1}{k^{\lambda - \epsilon}}
\]
for all \( \epsilon > 0 \) and a suitable \( C'( \epsilon) > 0 \), so \( \beta(G, \varphi, H, \psi, \alpha) \geq \lambda(G, \varphi, H, \psi, \alpha) \).

Now let \( \epsilon > 0 \) be given. Let \( 1 \leq \omega_1 < \omega_2 < \ldots \) be any sequence \( \omega \) of positive integers. Then we define a function \( f^{(\omega)} \) in the following way:

Let \( k_i \) with \( c^{\omega_i - 1} \leq k_i < c^{\omega_i} \) be such, that

\[
\sum_{v=\lfloor (\omega_i - 1) \rho \rfloor + 1}^{\lfloor (\omega_i + 1) \rho \rfloor} \frac{1}{b^{v\alpha}} \sum_{j=b^{v-1}}^{b^v-1} |\gamma(j, k_i)|
\]
attains its maximum for \( k = k_i \). Then set

\[
\overline{f}_G^{(\omega)}(j) := \frac{1}{b^{v\alpha}} \cdot \frac{|\gamma(j, k_i)|}{\gamma(j, k_i)}
\]
for all \( j \) with \( b^{\lfloor (\omega_i - 1) \rho \rfloor} \leq b^{v-1} \leq j < b^v \leq b^{\lfloor (\omega_i + 1) \rho \rfloor} \), and \( \overline{f}_G^{(\omega)}(j) = 0 \) for all other \( j \).

Thus \( f^{(\omega)} \in G, \varphi E^\alpha(1) \).
On the other hand
\[ \left| \hat{f}_H (k_i) \right| \geq \sum_{v=[(\omega_i-1)\rho]+1}^{[(\omega_i+1)\rho]} \frac{1}{b^{v\alpha}} \sum_{j=b^{v-1}}^{b^v-1} |\gamma (j, k_i)| - \sum_{j=b^{[(\omega_i+1)\rho]}}^{\infty} \frac{1}{j^{\alpha}}. \]

We have
\[ \left| \sigma_{\omega_i} - \sum_{v=[(\omega_i-1)\rho]+1}^{[(\omega_i+1)\rho]} \frac{1}{b^{v\alpha}} \sum_{j=b^{v-1}}^{b^v-1} |\gamma (j, k_i)| \right| \leq \sum_{j=b^{[(\omega_i+1)\rho]}}^{\infty} \frac{1}{j^{\alpha}}, \]
and therefore
\[ \left| \hat{f}_H^{(\omega)} (k_i) \right| \geq \sigma_{\omega_i} - 2 \sum_{j=b^{[(\omega_i+1)\rho]}}^{\infty} \frac{1}{j^{\alpha}}. \]

Let now the sequence \( \omega \) be such, that \( \lambda \geq -\frac{\log_{\omega_i} \sigma_{\omega_i}}{\omega_i} - \epsilon \) for all \( i \) and such, that
\[ \sum_{j=b^{[(\omega_i+1)\rho]}}^{\infty} \frac{1}{j^{\alpha}} < \frac{1}{4} \cdot \frac{1}{c^{(\lambda-\epsilon)}} \]
for all \( i \). Then
\[ \left| \hat{f}_H^{(\omega)} (k_i) \right| \geq \frac{1}{c^{\omega_i (\lambda-\epsilon)}} - \frac{2}{4} \cdot \frac{1}{c^{\omega_i (\lambda-\epsilon)}} \geq \frac{1}{2} \cdot \frac{1}{c^{\alpha (\lambda-\epsilon)}} \cdot \frac{1}{k_i^{\lambda-\epsilon}} \]
for all \( i \).

Therefore \( \beta (G, \varphi, H, \psi, \alpha) \leq \lambda (G, \varphi, H, \psi, \alpha) \) and the result follows. \( \square \)

Although we now have an exact formula for \( \beta \), it can, however, not be computed in finite time. Consequently, until now we do not even know, for example, the value of \( \beta (\mathbb{Z}_2, \text{id}, \mathbb{Z}_6, \text{id}, \alpha) \). But we can use Theorem 5 to obtain good estimates for \( \beta (G, \varphi, H, \psi, \alpha) \). Using the trivial estimate \( |\gamma (j, k)| \leq 1 \) together with Theorem 5 would lead to \( \beta (G, \varphi, H, \psi, \alpha) \geq (\alpha - 1) \cdot \rho \cdot \log (b) / \log (c) \). An upper bound and a sharper lower bound is given by the following result.

**Theorem 6.** Let \( G \) and \( H \) be of order \( b \) and \( c \) and let \( \varphi \) and \( \psi \) be suitable bijections. Assume, that there is a positive integer \( N \), such that \( b|c^N \). Let \( \rho = \rho (b, c) \) be defined as in Definition 6. Then with \( \theta := \rho \cdot \log (b) / \log (c) \) we have
\[ \alpha \cdot \theta - \min (\theta/2, 2\theta - 1) \leq \beta (G, \varphi, H, \psi, \alpha) \leq \alpha \cdot \theta + (1 - \theta). \]

**Remark.** Note that \( 0 < \theta \leq 1 \) and that the upper and the lower bound differ at most by \( 2/3 \). If \( b \) and \( c \) are as in Section 5 then \( \theta = 1 \) and we obtain the bounds from Theorem 4.
Proof of the theorem. By Cauchy’s inequality we have for \( c^{w-1} \leq k < c^w \) and with

\[
\sigma_w (k) := \sum_{v=\lfloor (w-1)\rho \rfloor+1}^{\infty} \frac{1}{b^{v\alpha}} \sum_{j=b^v}^{b^v-1} |\gamma (j,k)|,
\]

that

\[
|\sigma_w (k)| \leq \sum_{v=\lfloor (w-1)\rho \rfloor+1}^{\infty} \frac{b^{v/2}}{b^{v\alpha}} \left( \sum_{j=b^v}^{b^v-1} |\gamma (j,k)|^2 \right)^{1/2}.
\]

By the Bessel inequality for \( L_2 ([0,1]) \) with the orthonormal basis \( W_{G,\varphi} \)

\[
\sum_{j=b^v}^{b^v-1} |\gamma (j,k)|^2 = \sum_{j=b^v}^{b^v-1} \left| \int_0^1 H_{wal_j} (x) \cdot \overline{G_{wal_k} (x)} \, dx \right|^2 \leq \sum_{j=0}^{b^v-1} \left| \int_0^1 H_{wal_k} (x) \cdot \overline{G_{wal_j} (x)} \, dx \right|^2 \leq \| H_{wal_k} \|_2^2
\]

for all \( v \).

So

\[
|\sigma_w (k)| \leq \sum_{v=\lfloor (w-1)\rho \rfloor+1}^{\infty} \frac{b^{v/2}}{b^{v\alpha}} \leq C_1 (b, \alpha) \cdot \frac{1}{b^{w\rho (\alpha-1/2)}}
\]

for all \( w \) and \( k \) (and a constant \( C_1 (b, \alpha) \) depending only on \( b \) and \( \alpha \)) and we obtain \( \lambda \geq (\alpha - 1/2) \cdot \theta \).

To obtain a further lower bound for \( \lambda \) we estimate the coefficients \( \gamma (j,k) \) “individually”: For \( k \) given with \( c^{w-1} \leq k < c^w \), for \( v \) with

\[
\lfloor (w-1)\rho \rfloor + 1 \leq v \leq (w-1) \cdot \frac{\log c}{\log b}
\]

and for \( j \) with \( b^v-1 \leq j < b^v \) we have \( 1/c^{w-1} \leq 1/b^v \) and therefore by Lemma 2 and since \( G_{wal_j} \) is constant on intervals of the form \([A/b^v, (A+1)/b^v])\):

\[
|\gamma (j,k)| = \left| \int_0^1 G_{wal_j} (x) \cdot \overline{H_{wal_k} (x)} \, dx \right| = \left| \sum_{a=0}^{b^v-1} \int_{a/b^v}^{(a+1)/b^v} G_{wal_j} (x) \cdot \overline{H_{wal_k} (x)} \, dx \right| \leq \frac{b^v}{c^{w-1}}.
\]
So

\[ \sigma_w \leq \sum_{v=\lceil (w-1) \rho \rceil + 1}^{c^{w-1}} \frac{1}{b^v} \cdot \frac{1}{b^{v-2}} + \sum_{v=\lceil (w-1) \log_2 \log_3 \rangle + 1}^{\infty} \frac{1}{b^{v(\alpha-1)}} \]

\[ \leq C_2 (b, c, \alpha) \cdot \left( \frac{1}{c^w} \cdot \frac{1}{b^{w-\rho-2}} + \frac{1}{c^{w-1}} \right) \]

\[ \leq C_2 (b, c, \alpha) \cdot \left( \frac{1}{c^{w-1+\theta-(\alpha-2)}} + \frac{1}{c^{w-1}} \right) \]

with the constant \( C_2 (b, c, \alpha) \) depending only on \( b, c \) and \( \alpha \), and since \( \theta \leq 1 \).

From the definition of \( \lambda (G, \varphi, H, \psi, \alpha) \) we immediately get \( \lambda \geq \theta \cdot \alpha - 2 \theta + 1 \) and by Theorem 5 the lower bound for \( \beta (G, \varphi, H, \psi, \alpha) \) follows.

To show the upper bound it suffices to prove the following: For any positive integer \( w \) let \( v := \lceil (w-1) \rho \rceil + 1 \). Then there is a \( k \) with \( c^{w-1} \leq k < c^w \), such that

\[ \sum_{j=0}^{b^{v-1}-1} |\gamma(j,k)| \geq \frac{b^{v-1}-1}{c^w}. \]

Since then for all \( w \)

\[ \sigma_w \geq C_3 (b, c, \alpha) \cdot \frac{1}{b^{w-\rho-\alpha}} \cdot \frac{b^{w-\rho}}{c^w} \geq C_3 (b, c, \alpha) \cdot \frac{1}{c^{w-(1+\theta-(\alpha-1))}} \]

(here again the constant is depending only on \( b, c \) and \( \alpha \)), and by Theorem 5 the upper bound follows.

Now by Lemma 8 and 4:

\[ \sum_{j=0}^{b^{v-1}-1} |\gamma(j,k)| = \sum_{j=0}^{b^{v-1}-1} |\gamma(j,k)| \]

\[ \geq \left| \int_0^1 H_{\text{walk}}(x) \cdot \sum_{j=0}^{b^{v-1}-1} G_{\text{walk}}(x) \, dx \right| \]

\[ = b^v \cdot \left| \int_0^{b^{v-1}} H_{\text{walk}}(x) \, dx \right|. \]

Because of the special form of \( v \) the fraction \( A := \frac{c^{w-1}}{b^{v-1}} \) is an integer, but \( A/b \) is not. Let the integer \( E \) be such, that

\[ \frac{E}{c^{w-1}} + y = \frac{1}{b^v}, \]
where $0 < y < 1/c^{w-1}$. Then

$$E = \frac{A}{b} - c^{w-1} \cdot y$$

and

$$\frac{1}{b} \cdot \frac{1}{c^{w-1}} \leq y \leq \left(1 - \frac{1}{b}\right) \cdot \frac{1}{c^{w-1}}.$$ 

Now by Lemma 2

$$\left| \int_{b-v}^{E} \text{Hwal}_k (x) \, dx \right| = \left| \int_{c^{w-1} + y}^{E} \text{Hwal}_k (x) \, dx \right|$$

$$= \left| \int_{0}^{y} \text{Hwal}_k (x) \, dx \right|$$

$$= \frac{1}{c^{w-1}} \left| \int_{0}^{y} \text{Hwal}_{kw} (x) \, dx \right|.$$ 

Here

$$\frac{1}{b} \leq z := y \cdot c^{w-1} \leq 1 - \frac{1}{b}$$

and $k_w \neq 0$ is such, that $k = k_w \cdot c^{w-1} + \cdots + k_1$.

By Lemma 4

$$\sum_{l=1}^{c-1} \int_{0}^{z} \text{Hwal}_l (x) \, dx = \int_{0}^{z} \left( \sum_{l=0}^{c} \text{Hwal}_l (x) \right) - 1 \, dx$$

$$= \begin{cases} c \cdot z - z & \text{if } z \leq 1/c \\ 1 - z & \text{if } z > 1/c \end{cases}.$$ 

So there exists an $l \in \{1, \ldots, c - 1\}$, such that

$$\left| \int_{0}^{z} \text{Hwal}_l (x) \, dx \right| \geq \frac{1}{c} \cdot \min (c \cdot z - z, 1 - z) \geq \frac{1}{b \cdot c}.$$ 

Hereby we have shown the existence of a $k$ with $c^{w-1} \leq k < c^{w}$ and with

$$\sum_{j=b^{v-1}}^{b^{v-1}} |\gamma (j, k)| \geq \frac{b^{v-1}}{c^{w-1}}.$$ 

This completes the proof. \qed
7. Summary of the base change results.

In this section, for the sake of the reader, we collect and summarize the results on the base change coefficient given in Sections 4-6.

Given two finite abelian groups, $G, H$, we distinguished three cases, according to the relations between the group orders:

1) If $|G|$ did not divide any (integral) power of $|H|$, we found, that the relation between the convergence classes was quite bad. We obtained the non-improvable result

$$\beta(G, \varphi, H, \psi, \alpha) = 1$$

for all $\alpha > 1$.

2) If some power of $|G|$ equalled some power of $|H|$, the results showed very good relations between the convergence classes:

$$\beta(G, \varphi, H, \psi, \alpha) = \alpha - \beta_{G,H,\varphi,\psi},$$

where $\beta_{G,H,\varphi,\psi}$ was a finitely computable constant (see Definition 5). Estimates of $\beta_{G,H,\varphi,\psi}$ lead to an estimate of $\beta(G, \varphi, H, \psi, \alpha)$:

$$\alpha - 1/2 < \beta(G, \varphi, H, \psi, \alpha) \leq \alpha.$$ 

3) If $|G|$ divided some power of $|H|$, we had to formulate the result using the, in general, not finitely computable constant $\lambda$ (see Definition 7). Then

$$\beta(G, \varphi, H, \psi, \alpha) = \lambda(G, \varphi, H, \psi, \alpha).$$

We were able to give good estimates for this constant, leading to the inequalities

$$\alpha \cdot \theta - \min (\theta/2, 2\theta - 1) \leq \beta(G, \varphi, H, \psi, \alpha) \leq \alpha \cdot \theta + (1 - \theta),$$

where $\theta$ is a constant in $(0, 1]$, depending only on the group orders (see Definition 6 and the formulation of Theorem 6).

Note, that the second case is actually a special case of the third: It corresponds to the value $\theta = 1$. The resulting bounds for $\beta(G, \varphi, H, \psi, \alpha)$ coincide.

8. An application to quasi-Monte Carlo integration.

In a series of papers (see for example [14], [11], [9], [15],...) a so-called “digital lattice rule” for the numerical integration of functions defined on the $s$-dimensional unit cube $[0, 1)^s$ was developed. The essential observation of this method is, that certain classes of functions $f : [0, 1)^s \to \mathbb{C}$ can be approximately integrated with the help of so-called “digital nets” in a much more accurate way than with other methods. The main result of this “digital lattice rule” is an integration error estimate, which was given in improved and generalized form in [14], [9] and [27].
In this section we will give just the necessary preliminaries to state and prove the most general and sharpest form of this error estimate until now, based on Theorems 3 and 6. The error estimate is then given in Theorem 7.

We begin by extending the concept of Walsh systems over groups to arbitrary dimensions.

**Definition 8.** Let $G$ be a finite abelian group and $\varphi$ a corresponding bijection. Let $W_{G,\varphi}$ be the system of Walsh functions over $G$ and $\varphi$. For an integer $s \geq 2$ and non-negative integers $n_1, \ldots, n_s$ let

$$G_{\varphi} \text{wal}_{n_1, \ldots, n_s} : [0, 1)^s \to \mathbb{C}$$

be defined by

$$G_{\varphi} \text{wal}_{n_1, \ldots, n_s}(x_1, \ldots, x_s) := \prod_{i=1}^{s} G_{\varphi} \text{wal}_{n_i}(x_i),$$

and

$$W_{G,\varphi}^s := \{ G_{\varphi} \text{wal}_{n_1, \ldots, n_s} : n_1, \ldots, n_s \geq 0 \}$$

is called the system of $s$-dimensional Walsh functions over $G$ and $\varphi$.

**Definition 9.** For $G$ and $\varphi$ given, for an integer $s \geq 2$ and real numbers $\alpha > 1$ and $C > 0$ let $G_{\varphi} E^s_\alpha(C)$ denote the class of all functions $f : [0, 1)^s \to \mathbb{C}$ which are representable by an $s$-dimensional Walsh series

$$f(x_1, \ldots, x_s) = \sum_{n_1, \ldots, n_s = 0}^{\infty} \hat{f}_{G}(n_1, \ldots, n_s) \cdot G_{\varphi} \text{wal}_{n_1, \ldots, n_s}(x_1, \ldots, x_s)$$

with Walsh coefficients $\hat{f}_{G}(n_1, \ldots, n_s) \in \mathbb{C}$ satisfying

$$|\hat{f}_{G}(n_1, \ldots, n_s)| \leq C \cdot (n_1 \cdots n_s)^{-\alpha}$$

for all $(n_1, \ldots, n_s) \neq (0, \ldots, 0)$. (Here $\overline{n} := \max (1, n)$.)

**Definition 10.** Let $G$ and $H$ be finite abelian groups and $\varphi, \psi$ corresponding bijections. For an arbitrary dimension $s$ and real numbers $\alpha > 1$ we define: $\beta(G, \varphi, H, \psi, \alpha, s) := \sup \left\{ \beta > 1 : \text{for all } C > 0 \text{ there is a } C' > 0 \right\}$

with $G_{\varphi} E^s_\alpha(C) \subseteq H_{\psi} E^s_\beta(C')$.

Again $\beta(G, \varphi, H, \psi, \alpha, s) := 1$ if no such $\beta$ exists.

For this multi-dimensional extension we can show:

**Lemma 11.** $\beta(G, \varphi, H, \psi, \alpha, s) = \beta(G, \varphi, H, \psi, \alpha)$.

**Proof.** The proof is exactly the same as the proof of Lemma 2 in [12]. We omit the obvious adaptions.
The concept of digital \((t, m, s)\)-nets – these are point sets in the \(s\)-dimensional unit cube of a special structure – was introduced by Niederreiter ([9], [16], see also [17]) and was subsequently investigated in detail by various authors. (Special examples of digital \((t, m, s)\)-nets already can be found in Sobol’ [25] and Faure [4].)

**Definition 11.** Let \(b \geq 2\), \(s \geq 1\), and \(0 \leq t \leq m\) be integers. Then a point set \(P = \{x_0 \ldots , x_{N-1}\}\) consisting of \(N = b^m\) points of \([0, 1)^s\) forms a \((t, m, s)\)-net in base \(b\) if the number of \(n\) with \(0 \leq n \leq N - 1\), for which \(x_n\) is in the subinterval \(J\) of \([0, 1)^s\), is \(b^t\) for all \(J = \prod_{i=1}^{s} [a_i b^{-d_i}, (a_i + 1)b^{-d_i})\) with integers \(d_i \geq 0\) and \(0 \leq a_i < b^{d_i}\) for \(1 \leq i \leq s\), and with \(s\)-dimensional volume \(b^{t-m}\).

**Definition 12.** Let \(b \geq 2\), \(s \geq 1\), and \(m \geq 1\) be integers. We consider the following construction principle for point sets \(P\) consisting of \(b^m\) points in \([0, 1)^s\). We choose:

(i) A commutative ring \(R\) with identity and \(|R| = b\);
(ii) a bijection \(\tau : R \rightarrow Z_b = \{0, 1, \ldots , b-1\}\) with \(\tau(0) = 0\);
(iii) elements \(c_{ijr}^{(i)} \in R\) for \(1 \leq i \leq s\), \(1 \leq j \leq m\), and \(0 \leq r \leq m - 1\).

For \(n = 0, 1, \ldots , b^m - 1\) let

\[
n = \sum_{r=0}^{m-1} a_r(n)b^r \quad \text{with all } a_r(n) \in Z_b
\]

be the digit expansion of \(n\) in base \(b\). We put

\[
x_n^{(i)} = \sum_{j=1}^{m} y_{nij}^{(i)} b^{-j} \quad \text{for } 0 \leq n < b^m \text{ and } 1 \leq i \leq s,
\]

with

\[
y_{nij}^{(i)} = \tau \left( \sum_{r=0}^{m-1} c_{ijr}^{(i)} \tau^{-1}(a_r(n)) \right) \in Z_b
\]

for \(0 \leq n < b^m, 1 \leq i \leq s, 1 \leq j \leq m\).

If for some integer \(t\) with \(0 \leq t \leq m\) the point set

\[
x_n = \left( x_n^{(1)}, \ldots , x_n^{(s)} \right) \in [0, 1)^s \quad \text{for } n = 0, 1, \ldots , b^m - 1
\]

is a \((t, m, s)\)-net in base \(b\), then it is called a digital \((t, m, s)\)-net constructed over \(R\) with respect to the bijection \(\tau\).

The most powerful construction methods for digital \((t, m, s)\)-nets of high quality (i.e. with small \(t\)) are based on methods from algebraic geometry. See for example [18], [19].
Lemma 12. Let $R$ be a finite commutative ring of order $c$ with additive group $H$ and let $\psi : \{0, \ldots, c-1\} \to R$ be a bijection with $\psi(0) = 0$. Let $x_0, \ldots, x_{N-1}$ be a digital $(t, m, s)$-net constructed over $R$ with respect to $\tau = \psi^{-1}$. Then for all $\alpha > 1, C > 0$ we have
\[
\left| \int_{[0,1]^s} f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| \leq K \cdot c^{t\alpha} \cdot \frac{(\log N)^{s-1}}{N^\alpha}
\]
for all $f \in H, \psi E^\alpha_s (C)$, where $K$ is a constant depending only on $s, c, \alpha$ and $C$.

Proof. This is Theorem 1 in [9], stated in a slightly simplified form. □

In this result for the construction of the digital point set, a ring has to be used, which is based on the same additive group $G$ as the considered Walsh system. However, only certain rings are well suited for the construction of digital $(t, m, s)$-nets of highest quality.

Therefore it is sometimes more convenient to use a ring $R$ for the construction of the digital net which is based on another additive group $H$ as the considered Walsh system, which is based on, say, a group $G$. So we need a corresponding, more general integration error estimate.

In Theorem 4 in [27] such an estimate was given for the case $|G| = |H|$. We are now able to give an error estimate for the case, that for some positive integer $L$ we have: $|G|$ divides $|H|^L$. Our result - given in the subsequent Theorem - contains the above mentioned result of Wolf in [27] (Th.4).

Theorem 7. Let $G$ and $H$ be finite abelian groups of orders $b$ and $c$, and $\varphi$ and $\psi$ corresponding bijections. Let $R$ be a commutative ring with additive group $H$. Assume, that there exists a positive integer $L$, such that $b | c^L$. Let $\rho(b, c)$ be defined like in Definition 6 and let
\[
\theta := \rho(b, c) \cdot \frac{\log b}{\log c}.
\]
Let $x_1, \ldots, x_N$ be a digital $(t, m, s)$-net constructed over $R$ with respect to $\tau = \psi^{-1}$.

a) For all $\alpha > 1/\theta + \min(1/2, 2 - 1/\theta)$ and all $C > 0$ we have
\[
\left| \int_{[0,1]^s} f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| \leq K \cdot c^{t(\alpha - \theta - \min(\theta/2, 2\theta - 1))} \cdot \frac{(\log N)^{s-1}}{N^{\alpha - \min(\theta/2, 2\theta - 1)}}
\]
for all $f \in G, \varphi E^\alpha_s (C)$.

b) Assume, that for some positive integers $M$ and $L$ even $b^M = c^L$ holds. Let $\beta_{G,H,\varphi,\psi}$ be defined like in Definition 5.
For all $\alpha > 1 + \beta_{G,H,\varphi,\psi}$ and all $C > 0$ we have
\[
\left| \int_{[0,1]^s} f(x) \, dx - \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) \right| \leq K \cdot c^{s-(\alpha-\beta_{G,H,\varphi,\psi})} \cdot \frac{(\log N)^{s-1}}{N^{\alpha-\beta_{G,H,\varphi,\psi}}}
\]
for all $f \in G,\varphi E^\alpha_s (C)$. (In both cases again the $K$ denote constants depending only on $s, b, c, C$ and $\alpha$.)

Proof. Note, that in the proofs of the lower bound in Theorem 6 and of Theorem 3 we indeed even have proved a little more. From the proofs we even obtain the following:

- For all $\alpha > 1/\theta + \min(1/2, 2 - 1/\theta)$ and all $C > 0$ we have
  
  \[ G,\varphi E^\alpha_s (C) \subseteq H,\psi E^{\alpha-\theta-\min(\theta/2, 2\theta-1)} (C') \]

  for some $C' > 0$;

- respectively if $G$ and $H$ are of comparable order:

  - For all $\alpha > 1 + \beta_{G,H,\varphi,\psi}$ and all $C > 0$ we have
    
    \[ G,\varphi E^\alpha_s (C) \subseteq H,\psi E^{\alpha-\beta_{G,H,\varphi,\psi}} (C') \]

    for some $C' > 0$.

From this and Lemma 12 the assertion of Theorem 7 now immediately follows. □

The results of Theorem 7 (but also the result of Theorem 2, stating that there are no base change connections in Case 1), also are reflected in numerical results. As a small sample, we give in the following some results on the numerical integration of a function $f \in Z_{id} E^3_s (c)$ (i.e. $b = 2$) with digital nets over $Z_2, Z_5, F_8$ and $Z_{10}$, and with the Hammersley-Halton sequence.

As can be seen, we obtain excellent results for nets over $Z_2$ and $F_8$ (Case 2), good results for $Z_{10}$ (Case 3) and the “worst” results for $Z_5$ (Case 1) and the Hammersley-Halton sequence. In the last case (Case 1), the structure of the digital net does not play a role any more in integrating the function $f$. Only the small discrepancy of any digital net of high quality provides an integration error about as small as the error obtained by using the Hammersley-Halton sequence.

In the following table, there are two sections: In the first, the point sets contain between $2^{20}$ and $2^{21}$ points, in the second the range is from $2^{23}$ to $2^{24}$. In each section, the first column describes the point set, in the second the number of integration points $N$ is listed and in the third and fourth column, the integration error of a test function $f$ in 7, resp. 8 variables is shown. The function $f$ is the test function described for example in [12].
We use the digital nets over \( \mathbb{Z}_5 \) and \( \mathbb{Z}_{10} \) described in [9], digital nets over \( \mathbb{F}_8 \) as they are used in [12] and digital nets over \( \mathbb{Z}_2 \) as they are generated in [24].

<table>
<thead>
<tr>
<th>Method</th>
<th>( N )</th>
<th>dim 7</th>
<th>dim 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbb{Z}_2 )</td>
<td>1048576</td>
<td>2.3280821e-12</td>
<td>1.3543555e-11</td>
</tr>
<tr>
<td>( \mathbb{Z}_5 )</td>
<td>1953125</td>
<td>4.3304898e-04</td>
<td>2.5758582e-04</td>
</tr>
<tr>
<td>( \mathbb{F}_8 )</td>
<td>2097152</td>
<td>2.0078383e-13</td>
<td>5.1528231e-10</td>
</tr>
<tr>
<td>( \mathbb{Z}_{10} )</td>
<td>1000000</td>
<td>1.3743740e-05</td>
<td>5.0843230e-06</td>
</tr>
<tr>
<td>Hamm.</td>
<td>1048576</td>
<td>2.0067839e-03</td>
<td>2.2115463e-04</td>
</tr>
<tr>
<td>( \mathbb{Z}_2 )</td>
<td>8388608</td>
<td>2.3425705e-14</td>
<td>1.7175150e-13</td>
</tr>
<tr>
<td>( \mathbb{Z}_5 )</td>
<td>9765625</td>
<td>5.2695372e-05</td>
<td>2.3610550e-04</td>
</tr>
<tr>
<td>( \mathbb{F}_8 )</td>
<td>16777216</td>
<td>9.9364960e-15</td>
<td>1.0041967e-13</td>
</tr>
<tr>
<td>( \mathbb{Z}_{10} )</td>
<td>10000000</td>
<td>1.2159300e-05</td>
<td>5.4922667e-06</td>
</tr>
<tr>
<td>Hamm.</td>
<td>8388608</td>
<td>7.9642499e-05</td>
<td>3.2087933e-05</td>
</tr>
</tbody>
</table>

Table 1. \( N \) denotes the number of integration points, under dim 7 and dim 8 the integration errors of the 7- respective 8-dimensional test function are listed.

Finally we mention, that the above investigations also could be extended to investigations on the connection between Walsh series and Haar series (the problem does not really occur between different classes of Haar series) and the corresponding results could be used to reprove and improve results of Entacher ([2],[3]) on the numerical integration of Haar series. A forthcoming paper concerning this is in preparation. (See also [22].)

9. Some open problems.

Of course many questions remain open. In the following we restate just some of the – in our opinion – most challenging problems once more explicitely.

**Problem 1:** In practice, given \( G \) and \( H \) of comparable order, in the choice of bijections we are quite free. So it is obvious to choose \( \varphi \) and \( \psi \), such that \( \beta_{G,H,\varphi,\psi} \) is minimal. We know (Theorem 4) that

\[
0 \leq \beta_{G,H,\varphi,\psi} < \frac{1}{2}.
\]

Is

\[
\sup_{G,H} \min_{\varphi,\psi} \beta_{G,H,\varphi,\psi} < \frac{1}{2}?
\]

In fact we conjecture

\[
\sup_{G,H} \min_{\varphi,\psi} \beta_{G,H,\varphi,\psi} = \lim_{h \to \infty} \beta_{\mathbb{Z}_2,\mathbb{Z}_2,\text{id},\text{id}}
\]
$= \frac{1}{2} + \frac{\log \sin \frac{5\pi}{12}}{\log 2} = 0.4499...$

(See also [27].)

**Problem 2:** Is

$$\beta_{\mathbb{Z}_2, \mathbb{Z}_2, \text{id}, \text{id}} = \min_{\varphi, \psi} \beta_{\mathbb{Z}_2, \mathbb{Z}_2, \varphi, \psi}$$

We conjecture, yes.

**Problem 3:** We know a closed form for the exact values of $\beta_{\mathbb{Z}_2, \mathbb{Z}_2, \text{id}, \text{id}}$ (see Section 1). However, we do not know a closed form for the exact values of $\beta_{\mathbb{Z}_2, \mathbb{Z}_2, \text{id}, \text{id}}$.

**Problem 4:** Of course also in “Case Three”, $b|c^N$ for some $N$, a solution in the form of Theorem 3 for groups of comparable order would be desirable. However, this seems to be quite difficult.

Until now we do not even know the exact value of $\beta(G, \varphi, H, \psi, \alpha)$ in “easiest” cases, like, for example, $\beta(\mathbb{Z}_2, \text{id}, \mathbb{Z}_6, \text{id}, \alpha)$.

**Problem 5:** Any improvement of the inequality given in Theorem 6 would be of interest.

**Problem 6:** Is it possible to give more exact results than Theorems 5 and 6 in the case that $G$ and $H$ are not of comparable order, but $|G|$ and $|H|$ have the same prime factors?

(For example, it might be easier to compute $\beta(\mathbb{Z}_6, \text{id}, \mathbb{Z}_{12}, \text{id}, \alpha)$ than to compute $\beta(\mathbb{Z}_2, \text{id}, \mathbb{Z}_6, \text{id}, \alpha)$.)

**Problem 7:** Give at least numerical estimates for $\beta(G, \varphi, H, \psi, \alpha)$ for certain examples in “Case Three”, maybe based on Theorem 5.

**Problem 8:** Motivated by Theorem 3 for “Case Two” and Theorem 6 it seems that in “Case Three” we should have

$$\beta(G, \varphi, H, \psi, \alpha) = \alpha \cdot \theta + \delta(G, \varphi, H, \psi)$$

with some quantity $\delta(G, \varphi, H, \psi)$ not depending on $\alpha$. Is this true?

**Problem 9:** Concerning the constant $\gamma$ considered in Section 3, we know that $\gamma(G, \varphi, H, \psi, \alpha) \leq \beta(G, \varphi, H, \psi, \alpha)$. Can $\gamma$ also be estimated from below by some expression depending on $\beta$?

This is motivated by Theorem 1. However, note that $\beta$ is defined just over Walsh series, not over arbitrary functions as $\gamma$ is.

**Acknowledgments.**

We would like to thank the referee for many valuable suggestions that certainly improved the readability of the paper. Thanks also to Wolfgang Ch. Schmid for supporting us with computational work.
References


Received August 29, 1997 and revised July 29, 1998. The work of the second author was supported by the Austrian Research Foundation, Project P 11009 MAT and by the OeNB, Project 6788.

Universität Salzburg
Hellbrunnerstrasse 34
A–5020 Salzburg
Austria
E-mail address: Gerhard.Larcher@sbg.ac.at

Universität Salzburg
Hellbrunnerstrasse 34
A–5020 Salzburg
Austria
E-mail address: pirsic@mat.sbg.ac.at
A NOTE ON THE MOVING SPHERE METHOD

Yūki Naito and Takashi Suzuki

We treat the Dirichlet problem for elliptic equations on annular regions, and show the monotonicity and symmetry properties of positive solutions with respect to the sphere. We generalize the argument of the method of moving spheres to more general partial differential equations.

1. Introduction.

Let $A = \{x \in \mathbb{R}^n : 1/a < |x| < a\}$ be an annulus with $a > 1$ and $n \geq 2$. In [10] Padilla proved the following theorem by employing the method of moving spheres.

**Theorem A.** Let $n > 2$ and let $u$ be a solution to

$$-\Delta u = u^{\frac{n+2}{n-2}}, \quad u > 0 \quad \text{in} \quad A \quad \text{and} \quad u = 0 \quad \text{on} \quad \partial A.$$ 

Then $u$ satisfies

$$u(x) = |x|^{2-n}u\left(\frac{x}{|x|^2}\right) \quad \text{for} \quad x \in A \quad \text{and}$$

$$\left(|x|^{\frac{n-2}{2}}u\right)_r < 0 \quad \text{for} \quad 1 < r = |x| < a.$$ 

The method of moving spheres is a variant of the method of moving planes as presented in Gidas, Ni, and Nirenberg [6] or Berestycki and Nirenberg [1]. Roughly speaking, we make reflection with respect to spheres instead of planes, and then obtain the symmetry of solutions. In the works of Chou and Chu [5], Chen and Li [4], Li and Zhu [9], and Kurata and Matsuda [8], the method of moving spheres is used and is useful for solving various questions about elliptic differential equations.

In this note we generalize the argument of the method of moving spheres to more general partial differential equations. Let us consider the equation

$$(1) \quad -\Delta u = f(x, |x|^{\frac{n-2}{2}}u, (x \cdot \nabla)\left(|x|^{\frac{n-2}{2}}u\right)), \quad u > 0 \quad \text{in} \quad A,$$
where \( x \cdot \nabla = \sum_{i=1}^{n} x_i \partial / \partial x_i \). We assume that \( f = f(x, s, q) \) is continuous on \( \overline{A} \times [0, \infty) \times \mathbb{R} \), \( C^1 \) in \( s \) and \( q \), and even with respect to \( q \):

\[
f(x, s, -q) = f(x, s, q) \quad (x \in A, s \geq 0, q \in \mathbb{R}).
\]

We obtain the following theorems.

**Theorem 1.** Suppose that, for each \( 1 \leq r_0 \leq a \), \( \omega \in S_{n-1} \), \( s \geq 0 \), and \( q \geq 0 \),

\[
r^{\frac{n+2}{2}} f(r\omega, s, q) \geq r^{\frac{n+2}{2}} f(r_0 \omega, s, q) \quad \text{for} \quad \frac{1}{r_0} \leq r \leq r_0.
\]

Then, any \( u \in C^2(A) \cap C(\overline{A}) \) satisfying (1) and \( u = 0 \) on \( |x| = a \) has the properties

\[
|x|^\frac{n-2}{2} u(x) \leq \left( \frac{1}{|x|} \right)^\frac{n-2}{2} u \left( \frac{x}{|x|^2} \right) \quad \text{on} \quad 1 \leq |x| \leq a
\]

and

\[
\left( |x| \frac{n-2}{2} u \right)_r < 0 \quad \text{for} \quad 1 < r = |x| < a.
\]

**Theorem 2.** Suppose that, for each \( 1/a \leq r_0 \leq 1 \), \( \omega \in S_n \), \( s \geq 0 \), and \( q \geq 0 \),

\[
r^{\frac{n+2}{2}} f(r\omega, s, q) \geq r^{\frac{n+2}{2}} f(r_0 \omega, s, q) \quad \text{for} \quad r_0 \leq r \leq \frac{1}{r_0}.
\]

Then, any \( u \in C^2(A) \cap C(\overline{A}) \) satisfying (1) and \( u = 0 \) on \( |x| = 1/a \) has the properties

\[
|x|^\frac{n-2}{2} u(x) \geq \left( \frac{1}{|x|} \right)^\frac{n-2}{2} u \left( \frac{x}{|x|^2} \right) \quad \text{on} \quad 1 \leq |x| \leq a
\]

and

\[
\left( |x| \frac{n-2}{2} u \right)_r > 0 \quad \text{for} \quad 1/a < r = |x| < 1.
\]

**Remark.** It is shown in [6, Theorem 2] by the method of moving planes that the positive solutions \( u \) of the equation \( \Delta u + f(u) = 0 \) in \( A \) with \( u = 0 \) on \( \partial A \) satisfies \( u_r < 0 \) on \((1 <) (a + a^{-1})/2 \leq r < a \).

As a consequence of Theorems 1 and 2 we obtain the following corollary.

**Corollary 1.** Suppose that, for each \( \omega \in S_{n-1} \), \( s \geq 0 \), and \( q \geq 0 \), \( r^{(n+2)/2} f(r\omega, s, q) \) is nonincreasing in \( r \in (1, a) \) and

\[
r^{\frac{n+2}{2}} f(r\omega, s, q) \equiv r^{\frac{n+2}{2}} f(r^{-1} \omega, s, q) \quad \text{for} \quad 1 \leq r \leq a.
\]
Then, any \( u \in C^2(A) \cap C(\overline{A}) \) satisfying (1) and \( u = 0 \) on \( \partial A \) has the properties

\[
|x|^{\frac{n-2}{2}} u(x) \equiv \left( \frac{1}{|x|} \right)^{\frac{n-2}{2}} u \left( \frac{x}{|x|^2} \right) \quad \text{on} \; 1 \leq |x| \leq a
\]

and (3).

**Remark.** If (1) has a solution \( u \) satisfying (4), then we must have

\[
|x|^{\frac{n+2}{2}} f(x, s(x), q(x)) \equiv |x|^{-\frac{n+2}{2}} f(x/|x|^2, s(x), q(x)) \quad \text{for} \; 1 \leq |x| \leq a,
\]

where \( s(x) = |x|^{(n-2)/2} u(x) \) and \( q(x) = x \cdot \nabla s(x) \). In fact, \( v(x) = \frac{n-2}{4} |x|^2 f(x, v(x), (x \cdot \nabla) v(x)) = 0 \)

and

\[
|x|^2 \Delta v - (n-2) x \cdot \nabla v - \frac{(n-2)^2}{4} v + |x|^{\frac{n+2}{2}} f(x, v(x), (x \cdot \nabla) v(x)) = 0,
\]

respectively. (See (7) and (8) below.) Then \( v(x) \equiv w(x) \) implies (5).

We consider the following typical problem

\[
\Delta u + g(x, u) = 0, \quad u > 0 \quad \text{in} \; A \quad u = 0 \quad \text{on} \; \partial A,
\]

where \( g = g(x, s) \) is continuous on \( \overline{A} \times [0, \infty) \) and \( C^1 \) in \( s \). In this case we see that \( f(x, s, q) = g(x, |x|^{(n-2)/2} s) \). Note that the existence of positive nonradial solution \( u \) of the problem (6) has been studied by many authors, see, e.g., Brezis and Nirenberg [2], Suzuki [11], Byeon [3], and the references therein. As a consequence of Corollary 1 we obtain the following corollary, which in the special case \( g(x, u) = u^{(n+2)/(n-2)} \) (and \( f(x, s, q) = |x|^{-(n+2)/2} s^{(n+2)/(n-2)} \) yields Theorem A.

**Corollary 2.** Suppose that, for each \( \omega \in S^{n-1} \) and \( s \geq 0 \), \( r^\frac{n+2}{2} g(r\omega, r^-\frac{n-2}{2} s) \) is nonincreasing in \( r \in (1, a) \) and

\[
r^\frac{n+2}{2} g(r\omega, r^-\frac{n-2}{2} s) \equiv r^-\frac{n+2}{2} g(r^{-1}\omega, r^\frac{n-2}{2} s) \quad \text{for} \; 1 \leq r \leq a.
\]

Let \( u \in C^2(A) \cap C(\overline{A}) \) be a solution of (6). Then \( u \) satisfies the properties (4) and (3).

**Remark.** For example,

\[
g(r\omega, s) = r^{-\frac{(n+2)p(n-2)}{2}} h(\omega) s^p + cs^\frac{n+2}{2}, \quad c, \; p \in \mathbb{R}, \; p \geq 1,
\]
where \( h(\omega) \) is continuous and positive on \( S^{n-1} \), satisfies the conditions in Corollary 2. For the case \( g(r\omega,s) = r^{-2}h(\omega)s + s^{(n+2)/(n-2)} \), the existence of positive solutions for the problem (6) is investigated in [2].

In our proof we use the operator \( \Delta_g v = |x|^2 \Delta v - (n-2)x \cdot \nabla v \). We note that \( \Delta_g \) is the Laplace-Beltrami operator on the Riemannian space \((\mathbb{R}^n, dx^2/|x|^2)\). We find that Equation (1) is written as

\[
\Delta_g v - \frac{(n-2)^2}{4} + |x|^{\frac{n+2}{2}} f(x, v, (x \cdot \nabla) v) = 0 \quad \text{in} \ A
\]

for the function \( v(x) = |x|^{(n-2)/2}u(x) \), and that the operator \( \Delta_g \) is invariant under the transformation \( x \mapsto y = \lambda^2 x/|x|^2 \).

In Section 2 we prove Theorems. In fact, we only present the proof of Theorem 1 since the proof of Theorem 2 is very similar. In Appendix we show that the operator \( \Delta_g \) is invariant under the transformation by using of the property of the Kelvin transformation.

2. Proof of Theorems.

Due to similarity, we only give the proof of Theorem 1. Given \( \lambda \in (1,a) \), we set

\[
T_\lambda = \{|x| = \lambda\} \quad \text{and} \quad \Sigma_\lambda = \{\lambda < |x| < a\}.
\]

For \( x \in \Sigma_\lambda \), let \( x^\lambda = \lambda^2 x/|x|^2 \). Then we have

\[
|x| > |x^\lambda| = \lambda^2 \frac{1}{|x|} > \frac{1}{|x|} \quad \text{for} \quad x \in \Sigma_\lambda.
\]

Define the operator \( \Delta_g \) by \( \Delta_g v = |x|^2 \Delta v - (n-2)x \cdot \nabla v \). We note that \( \Delta_g \) is the Laplace-Beltrami operator on the Riemannian space \((\mathbb{R}^n, dx^2/|x|^2)\). For a solution \( u \) of (1), the function \( v(x) = |x|^{\frac{n-2}{2}}u(x) \) satisfies

\[
|x|^2 \Delta v - (n-2)x \cdot \nabla v - \frac{(n-2)^2}{4} v + |x|^{\frac{n+2}{2}} f(x, v, (x \cdot \nabla) v) = 0
\]

in \( A \), which is written as

\[
\Delta_g v - \frac{(n-2)^2}{4} + |x|^{\frac{n+2}{2}} f(x, v, (x \cdot \nabla) v) = 0 \quad \text{in} \ A.
\]

Let \( v^\lambda(x) = v(x^\lambda) \) and \( y = x^\lambda \). By Lemma A in Appendix, we find that \( \Delta_g v^\lambda(x) = \Delta_g v(y) \). We have \( x \cdot \nabla = r \partial_r \) for \( r = |x| \) (see, e.g., [7]) and hence

\[
x \cdot \nabla x^\lambda = r \partial_r v^\lambda = -s \partial_s v = -y \cdot \nabla y, \n\]

where \( r = |x| \) and \( s = |y| = \lambda^2/r \). Therefore, the property \( f(x, s, -q) = f(x, s, q) \) implies the relation

\[
\Delta_g v^\lambda - \frac{(n-2)^2}{4} v^\lambda + |x^\lambda|^{\frac{n+2}{2}} f\left(x^\lambda, v^\lambda, (x \cdot \nabla) v^\lambda\right) = 0 \quad \text{in} \ A.
\]
It follows that \(|x| > |x^\lambda| > 1/|x|\) and \(1 < \lambda < |x| < a\) for \(x \in \Sigma_\lambda\). Then the assumption on \(f\) in Theorem 1 guarantees

\[
|x^\lambda|^\frac{n+2}{4} f(x^\lambda, s, q) \geq |x|^\frac{n+2}{4} f(x, s, q)
\]

for \(x \in \Sigma_\lambda, s \geq 0,\) and \(q \geq 0\). Therefore, the function \(w_\lambda = v^\lambda - v\) satisfies

\[
\Delta_g w_\lambda - \frac{(n-2)^2}{4} w_\lambda + |x|^\frac{n+2}{2} \left( f \left( x, v^\lambda, (x \cdot \nabla) v^\lambda \right) - f \left( x, v, (x \cdot \nabla) v \right) \right) \leq 0
\]

on \(\Sigma_\lambda\). Writing

\[
b_\lambda(x) = \int_0^1 f_s \left( x, t v^\lambda(x) + (1-t) v(x), (x \cdot \nabla) v^\lambda(x) \right) dt \quad \text{and} \quad c_\lambda(x) = \int_0^1 f_q \left( x, v(x), t(x \cdot \nabla) v^\lambda(x) + (1-t)(x \cdot \nabla) v(x) \right) dt
\]

we obtain

\[
\Delta_g w_\lambda - \frac{(n-2)^2}{4} w_\lambda + |x|^\frac{n+2}{2} \left( b_\lambda(x) + c_\lambda(x) x \cdot \nabla \right) w_\lambda \leq 0 \quad \text{on} \quad \Sigma_\lambda.
\]

Let \(z_\lambda(x) = |x|^{-\frac{n-2}{2}} w_\lambda(x)\). Then we have

\[
|x|^\frac{n+2}{2} \Delta z_\lambda = \Delta_g w_\lambda - \frac{(n-2)^2}{4} w_\lambda \quad \text{and} \quad |x|^\frac{n-2}{2} x \cdot \nabla z_\lambda = -\frac{n-2}{2} w_\lambda + x \cdot \nabla w_\lambda.
\]

Define

\[
\tilde{b}_\lambda(x) = |x|^\frac{n-2}{2} \left( b_\lambda(x) + \frac{n-2}{2} c_\lambda(x) \right) \quad \text{and} \quad \tilde{c}_\lambda(x) = |x|^\frac{n-2}{2} c_\lambda(x).
\]

We have shown the following lemma.

**Lemma 1.** Under the assumptions of Theorem 1, each \(\lambda \in (1, a)\) admits the inequality

\[
(9) \quad \Delta z_\lambda + \tilde{b}_\lambda(x) z_\lambda + \tilde{c}_\lambda(x) x \cdot \nabla z_\lambda \leq 0 \quad \text{on} \quad \Sigma_\lambda,
\]

where \(z_\lambda(x) = |x|^{-\frac{n-2}{2}} (v^\lambda - v)\).

Once Lemma 1 is proven, Theorem 1 follows from the standard argument ([1]). Putting

\[
\Lambda \equiv \{ \lambda \in (1, a) : z_\lambda > 0 \quad \text{in} \quad \Sigma_\lambda \},
\]

we see that the desired consequence follows from \(\Lambda = (1, a)\). We show \(\Lambda = (1, a)\) by three steps.

**Step 1.** We have \([r_0, a) \subset \Lambda\) for \(r_0\) close to \(a\), that is, \(\Lambda \neq \emptyset\).

**Proof.** We see that the coefficients \(\tilde{b}_\lambda(x)\) and \(\tilde{c}_\lambda(x)\) in (9) are uniformly bounded. Then for \(r_0\) close to \(a\), the maximum principle holds for the
Equation (9) on any subdomain of \( A \setminus \overline{B}_{r_0} \) and for any \( \lambda \), where \( B_{r_0} = \{ x \in \mathbb{R}^n : |x| < r_0 \} \). (See e.g. [1].) This implies \([r_0, 1) \subset \Lambda. \)

We prepare the following lemma.

**Lemma 2.** (i) If \( \lambda \in \Lambda \), then

\[
\frac{\partial z_\lambda}{\partial \nu} < 0 \quad \text{on} \quad T_\lambda,
\]

where \( \nu \) denotes the outer unit normal vector on \( T_\lambda \) from \( \Sigma_\lambda \);

(ii) If \( \lambda \not\in \Lambda \), then there exists some \( x_0 \in \Sigma_\lambda \cap \overline{B}_{r_0} \) such that \( z_\lambda(x_0) \leq 0 \).

**Proof.** (i) Let \( \lambda \in \Lambda \). Then we have \( z_\lambda = 0 \) on \( T_\lambda \), and \( z_\lambda > 0 \) in \( \Sigma_\lambda \). Therefore, Hopf’s boundary lemma can be applied by (9) so that (10) holds.

(ii) As we have proven in Step 1, \( \lambda < r_0 \) and hence \( \Sigma_\lambda \cap \overline{B}_{r_0} \neq \emptyset \). Suppose to the contrary that

\[
z_\lambda(x) > 0 \quad \text{on} \quad \Sigma_\lambda \cap \overline{B}_{r_0}.
\]

Then we get

\[
\Delta z_\lambda + b_\lambda(x)z_\lambda + c_\lambda(x)x \cdot \nabla z_\lambda \leq 0 \quad \text{in} \quad \Sigma_\lambda \setminus \overline{B}_{r_0},
\]

and

\[
z_\lambda \geq 0 \quad \text{on} \quad \partial \left( \Sigma_\lambda \setminus \overline{B}_{r_0} \right).
\]

Now the maximum principle guarantees \( z_\lambda > 0 \) in \( \Sigma_\lambda \setminus \overline{B}_{r_0} \). However, we have \( z_\lambda > 0 \) in \( \Sigma_\lambda \cap \overline{B}_{r_0} \) and hence \( z_\lambda > 0 \) in \( \Sigma_\lambda \). This means \( \lambda \in \Lambda \), a contradiction.

**Step 2.** \( \Lambda \) is left-open.

**Proof.** If \( \Lambda \) is not left-open, there exist \( \lambda_0 \in \Lambda \) and a sequence \( \{\lambda_n\} \) satisfying

\[
\lambda_0 - \frac{1}{n} < \lambda_n < \lambda_0 \quad \text{and} \quad \lambda_n \not\in \Lambda.
\]

Lemma 2 (ii) guarantees the existence of \( x_n \in \Sigma_{\lambda_n} \cap \overline{B}_{r_0} \) satisfying

\[
z_{\lambda_n}(x_n) \leq 0.
\]

Note that \( z_{\lambda_n} = 0 \) on \( T_{\lambda_n} \). Then we have a point \( y_n \) on the segment connecting \( x_n \) and \( \lambda_n^2 x_n/|x_n|^2 \) satisfying

\[
\frac{\partial z_{\lambda_n}}{\partial r}(y_n) \leq 0.
\]

Taking a subsequence if necessary, we may suppose the existence of some \( x_0 \in \Sigma_{\lambda_0} \cap \overline{B}_{r_0} \) satisfying \( x_n \to x_0 \). By (11) we obtain \( z_{\lambda_0}(x_0) \leq 0 \). Since
\[ \lambda_0 \in \Lambda, \text{ we must have } x_0 \in T_{\lambda_0}. \] In particular, \( y_n \to x_0 \) and \( \partial z_{\lambda_0} / \partial r(x_0) \leq 0 \) follows from (12). However, this is equivalent to
\[
\frac{\partial z_{\lambda_0}}{\partial \nu}(x_0) \geq 0,
\]
which contradicts to (10) valid for \( \lambda = \lambda_0 \in \Lambda \).

**Step 3.** \( \Lambda \) is left-closed.

Proof. In fact, let \( \{ \lambda_n \} \subset \Lambda \) be a sequence satisfying \( \lambda_n \downarrow \lambda_1 > 1 \). Then, we have
\[
\Delta z_{\lambda_1} + \left( \tilde{b}_{\lambda_1}(x) + \tilde{c}_{\lambda_1}(x) x \cdot \nabla \right) z_{\lambda_1} \leq 0 \quad \text{and} \quad z_{\lambda_1} \geq 0 \quad \text{in } \Sigma_{\lambda_1}.
\]
Since \( z_{\lambda_1} > 0 \) on \( |x| = a \), we have \( z_{\lambda_1} \not\equiv 0 \) in \( \Sigma_{\lambda_1} \). Therefore, the maximum principle implies \( z_{\lambda_1} > 0 \) in \( \Sigma_{\lambda_1} \), or equivalently, \( \lambda_1 \in \Lambda \).

As a consequence of Steps 1-3, we obtain \( \Lambda = (1,a) \). This implies \( v^1(x) \geq v(x) \) on \( 1 \leq |x| \leq a \), and then (2) holds. The property (3) follows from Lemma 2 (i). This completes the proof.

**Appendix.**

Let \( \Delta_g v = |x|^2 \Delta v - (n-2)x \cdot \nabla v \). We show that the operator \( \Delta_g \) is invariant under the transformation \( x \mapsto y = \lambda^2 x/|x|^2 \), that is, \( \Delta_g v = \Delta_g V \) for \( v(x) = V(y) \). Here we use the well-known property of the Kelvin transformation \( \eta = \xi/|\xi|^2 \) expressed as
\[
\Delta_\eta U = |\xi|^{n+2} \Delta_\xi u \quad \text{for} \quad U(\eta) = |\xi|^{n-2} u(\xi).
\]

**Lemma A.** Let \( v(x) = V(y) \) and \( y = \lambda^2 x/|x|^2 \) with \( \lambda > 0 \). Then we have
\[
|x|^2 \Delta_x v - (n-2)x \cdot \nabla_x v = |y|^2 \Delta_y V - (n-2) y \cdot \nabla_y V,
\]
where \( \Delta_x = \sum_{i=1}^n \partial^2 / \partial x_i^2 \) and \( x \cdot \nabla x = \sum_{i=1}^n x_i \partial / \partial x_i \).

Proof. Writing \( w(x) = |x|^{-\frac{n+2}{2}} v(x) \) we have
\[
|x|^{\frac{n+2}{2}} \Delta_x w = |x|^2 \Delta_x v - (n-2)x \cdot \nabla_x v - \frac{(n-2)^2}{4} v.
\]
Similarly, writing \( W(y) = |y|^{-\frac{n+2}{2}} V(y) \) we have
\[
|y|^{\frac{n+2}{2}} \Delta_y W = |y|^2 \Delta_y V - (n-2) y \cdot \nabla_y V - \frac{(n-2)^2}{4} V.
\]
By \(|y| = \lambda^2/|x|\) it follows that
\[
|x|^\frac{n-2}{2} w(x) = v(x) = V(y) = |y|^\frac{n-2}{2} W(y) = \left(\frac{\lambda^2}{|x|}\right)^\frac{n-2}{2} W(y).
\]
Then we obtain
\[
W(y) = \left(\frac{|x|}{\lambda}\right)^{n-2} w(x).
\]
By the property of the Kelvin transformation, we obtain
\[
\Delta_y W = \left(\frac{|x|}{\lambda}\right)^{n+2} \Delta_x w.
\]
Then we have
\[
|x|^\frac{n+2}{2} \Delta_x w = \left(\frac{\lambda^2}{|x|}\right)^\frac{n+2}{2} \left(\frac{|x|}{\lambda}\right)^{n+2} \Delta_x w = |y|^\frac{n+2}{2} \Delta_y W.
\]
Therefore, by (14) and (15), we obtain the property (13). This completes the proof. \(\square\)

**Acknowledgement.** The authors wish to express their thanks to the referee for helpful comments and valuable suggestions which improve the original version of this paper.

**References**


Received August 27, 1997 and revised May 27, 1998.

Kobe University
Kobe 657-8501
JAPAN
E-mail address: naito@pascal.seg.kobe-u.ac.jp

Osaka University
Toyonaka 560-0043
JAPAN
E-mail address: takashi@math.sci.osaka-u.ac.jp
PSEUDODIFFERENTIAL OPERATORS ON DIFFERENTIAL GROUPOIDS

VICTOR NISTOR, ALAN WEINSTEIN, AND PING XU

We construct an algebra of pseudodifferential operators on each groupoid in a class that generalizes differentiable groupoids to allow manifolds with corners. We show that this construction encompasses many examples. The subalgebra of regularizing operators is identified with the smooth algebra of the groupoid, in the sense of non-commutative geometry. Symbol calculus for our algebra lies in the Poisson algebra of functions on the dual of the Lie algebroid of the groupoid. As applications, we give a new proof of the Poincaré-Birkhoff-Witt theorem for Lie algebroids and a concrete quantization of the Lie-Poisson structure on the dual $A^*$ of a Lie algebroid.

Introduction.

Certain important applications of pseudodifferential operators require variants of the original definition. Among the many examples one can find in the literature are regular or adiabatic families of pseudodifferential operators [2, 41], pseudodifferential operators along the leaves of foliations [6, 8, 28, 29], on coverings [9, 30] or on certain singular spaces [21, 22, 25, 26].

Since these classes of operators share many common features, it is natural to ask whether they can be treated in a unified way. In this paper we shall suggest a positive answer to this question. For any “almost differential” groupoid (a class which allows manifolds with corners), we construct an algebra of pseudodifferential operators. We then show that our construction recovers (almost) all the classes described above (for operators on manifolds with boundary, our algebra is slightly smaller than the one defined in [21]). We expect our results to have applications to analysis on singular spaces.

Our construction and results owe a great deal to the previous work of several authors, especially Connes [6] and Melrose [20, 21, 23]. A hint of the direction we take was given at the end of [38]. The basic idea of our construction is to consider families of pseudodifferential operators along the fibers of the domain (or source) map of the groupoid. More precisely, for any almost differentiable groupoid (see Definition 3) we consider the fibers $G_x = d^{-1}(x)$ of the domain map $d$, which consist of all arrows with domain $x$. It follows from the definition of an almost differentiable groupoid that
these fibers are smooth manifolds (without corners). The calculus of pseudo-differential operators on smooth manifolds is well understood and by now a classical subject, see for example [14]. We shall consider differentiable families of pseudodifferential operators $P_x$ on the smooth manifolds $G_x$. Right translation by $g \in G$ defines an isomorphism $G_x \equiv G_y$ where $x$ is the domain of $g$ and $y$ is the range of $g$. We say that the family $P_x$ is invariant if $P_x$ transforms to $P_y$ under the diffeomorphisms above (for all $g$). The algebra $\Psi^\infty(G)$ of pseudodifferential operators on $G$ that we shall consider will consist of invariant differentiable families of operators $P_x$ as explained above (the actual definition also involves a technical condition on the support of these operators). See Definition 7 for details. The relation with the work of Melrose relies on an alternative description of our algebra as an algebra of distributions on $G$ with suitable properties (compactly supported, conormal, and with singular support contained in the set of units). This is contained in Theorem 7. The difference between our theory and Melrose’s lies in the fact that he considers a compactification of $G$ as a manifold with corners, and his distributions are allowed to extend to the compactification, with precise behavior at the boundary. This is useful for the analysis of these operators. In contrast, our work is purely algebraic (or geometric, depending on whether one considers Lie algebroids as part of geometry or algebra).

We now review the contents of the sections of this paper. In the first section we recall the definitions of a groupoid, of a Lie algebroid, and of the less known concept of local Lie groupoid. We extend the definition of a Lie groupoid to include manifolds with corners. These groupoids are called almost differentiable groupoids. The second section contains the definition of a pseudodifferential operator on a groupoid (really a family of pseudodifferential operators, as explained above) and the proof that they form an algebra, if a support condition is included. We also extend this definition to include local Lie groupoids. This is useful in the third section where we use this to give a new proof of the Poincaré-Birkhoff-Witt theorem for Lie algebroids. In the process of proving this theorem we also exemplify our definition of pseudodifferential operators on an almost differentiable groupoid by describing the differential operators in this class. As an application we give an explicit construction of a deformation quantization of the Lie-Poisson structure on $A^*$, the dual of Lie algebroid $A$. The section entitled “Examples” contains just what the title suggests: for many particular examples of groupoids $G$, we explicitly describe the algebra $\Psi^\infty(G)$ of pseudodifferential operators on $G$. This recovers classes of operators that were previously defined using ad hoc constructions. Our definition is often not only more general, but also simpler. This is the case for operators along the leaves of foliations [8, 28] or adiabatic families of operators. Since one of our main themes is that the Lie algebras of vector fields that are central in [24] are in fact the spaces of sections of Lie algebroids, we describe these Lie algebroids explicitly in
each of our examples. In the sixth section of the paper, we describe the
convolution kernels (called reduced kernels) of operators in $\Psi^\infty(\mathcal{G})$. Then
we extend to our setting some fundamental results on principal symbols, by
reducing to the classical results. This makes our proofs short (and easy).
Finally, the last section treats the action of $\Psi^\infty(\mathcal{G})$ on functions on the units
of $\mathcal{G}$, and a few related topics.

Recently, we have learned of certain related results by Monthubert, see
[27] and the references therein. Our paper was circulated as preprint funct-
an/9702004. The first author would like to thank Richard Melrose for several
useful conversations.

1. Preliminaries.

In the following we allow manifolds to have corners. Thus by “manifold”
we shall mean a $C^\infty$ manifold, possibly with corners, and by a “smooth
manifold” we shall mean a manifold without corners. By definition, if $M$ is
a manifold with corners, then every point $p \in M$ has coordinate neighbor-
hoods diffeomorphic to $[0, \infty)^k \times \mathbb{R}^{n-k}$. The transition functions between
such coordinate neighborhoods must be smooth everywhere (including on
the boundary). We shall use the following definition of submersions between
manifolds (with corners).

**Definition 1.** A submersion between two manifolds with corners $M$ and $N$
is a differentiable map $f : M \to N$ such that $df_x : T_x M \to T_{f(x)} N$ is onto
for any $x \in M$ and such that if $df_x(v)$ is an inward pointing tangent vector
to $N$, then $v$ is an inward pointing tangent vector to $M$.

The reason for introducing the definition above is that for any submersion
$f : M \to N$, the set $M_y = f^{-1}(y), y \in N$ is a smooth manifold, just as for
submersions of smooth manifolds.

We shall study groupoids endowed with various structures. ([33] is a
general reference for some of what follows.) We recall first that a small
category is a category whose class of morphisms is a set. The class of objects
of a small category is then a set as well.

**Definition 2.** A groupoid is a small category $\mathcal{G}$ in which every morphism
is invertible.

This is the shortest but least explicit definition. We are going to make
this definition more explicit in cases of interest. The set of objects, or units,
of $\mathcal{G}$ will be denoted by

$$M = \mathcal{G}^{(0)} = \text{Ob} (\mathcal{G}).$$

The set of morphisms, or arrows, of $\mathcal{G}$ will be denoted by

$$\mathcal{G}^{(1)} = \text{Mor} (\mathcal{G}).$$
We shall sometimes write $G$ instead of $G^{(1)}$ by abuse of notation. For example, when we consider a space of functions on $G$, we actually mean a space of functions on $G^{(1)}$. We will denote by $d(g)$ [respectively $r(g)$] the domain [respectively, the range] of the morphism $g : d(g) \to r(g)$. We thus obtain functions

\[ d, r : G^{(1)} \rightarrow G^{(0)} \]

that will play an important role below. The multiplication operator $\mu : (g, h) \mapsto \mu(g, h) = gh$ is defined on the set of composable pairs of arrows $G^{(2)}$:

\[ \mu : G^{(2)} = G^{(1)} \times_M G^{(1)} := \{(g, h) : d(g) = r(h)\} \rightarrow G^{(1)}. \]

The inversion operation is a bijection $\iota : g \mapsto g^{-1}$ of $G^{(1)}$. Denoting by $u(x)$ the identity morphism of the object $x \in M = G^{(0)}$, we obtain an inclusion of $G^{(0)}$ into $G^{(1)}$. We sometimes write $G = (G^{(0)}, G^{(1)}, d, r, \mu, u, \iota)$. The structural maps satisfy the following properties:

(i) $r(gh) = r(g), d(gh) = d(h)$ for any pair $(g, h) \in G^{(2)}$, and the partially defined multiplication $\mu$ is associative.

(ii) $d(u(x)) = r(u(x)) = x, \forall x \in G^{(0)}, u(r(g))g = g$ and $gu(d(g)) = g, \forall g \in G^{(1)}$ and $u : G^{(0)} \rightarrow G^{(1)}$ is one-to-one.

(iii) $r(g^{-1}) = d(g), d(g^{-1}) = r(g), gg^{-1} = u(r(g))$ and $g^{-1}g = u(d(g))$.

**Definition 3.** An almost differentiable groupoid $G = (G^{(0)}, G^{(1)}, d, r, \mu, u, \iota)$ is a groupoid such that $G^{(0)}$ and $G^{(1)}$ are manifolds with corners, the structural maps $d, r, \mu, u, \iota$ are differentiable, and the domain map $d$ is a submersion.

We observe that $\iota$ is a diffeomorphism and hence $d$ is a submersion if and only if $r = d \circ \iota$ is a submersion. Also, it follows from the definition that each fiber $G_x = d^{-1}(x) \subset G^{(1)}$ is a smooth manifold whose dimension $n$ is constant on each connected component of $G^{(0)}$. The étale groupoids considered in [5] are extreme examples of differentiable groupoids (corresponding to $\dim G_x = 0$). If $G^{(0)}$ is smooth (i.e. if it has no corners), then $G^{(1)}$ is also smooth and $G$ becomes what is known as a differentiable, or Lie groupoid.\(^1\)

We now introduce a few important geometric objects associated to an almost differentiable groupoid.

\(^1\)Earlier terminology, such as in [19], used the term Lie groupoid only for differentiable groupoids in which every pair of objects is connected by a morphism.
The vertical tangent bundle (along the fibers of $d$) of an almost differentiable groupoid $\mathcal{G}$ is

\begin{equation}
T_d\mathcal{G} = \ker d = \bigcup_{x \in \mathcal{G}(0)} T\mathcal{G}_x \subset T\mathcal{G}^{(1)}.
\end{equation}

Its restriction $A(\mathcal{G}) = T_d\mathcal{G}|_{\mathcal{G}(0)}$ to the set of units is the Lie algebroid of $\mathcal{G}$ [19, 31]. We denote by $T^*_d\mathcal{G}$ the dual of $T_d\mathcal{G}$ and by $A^*(\mathcal{G})$ the dual of $A(\mathcal{G})$. In addition to these bundles we shall also consider the bundle $\Omega^*_\lambda d$ of $\lambda$-densities along the fibers of $d$. If the fibers of $d$ have dimension $n$, then $\Omega^*_\lambda d = |\Lambda^n T^*_d\mathcal{G}|_\lambda = \bigcup x \Omega^*_\lambda (\mathcal{G}_x)$. By invariance these bundles can be obtained as pull-backs of bundles on $\mathcal{G}(0)$. For example $T_d\mathcal{G} = r^* (A(\mathcal{G}))$ and $\Omega^*_d \lambda = r^*(D^\lambda)$, where $D^\lambda$ denotes $\Omega^*_\lambda|_{\mathcal{G}(0)}$. If $E$ is a (smooth complex) vector bundle on the set of units $\mathcal{G}(0)$, then the pull-back bundle $r^* (E)$ on $\mathcal{G}$ will have right invariant connections obtained as follows. A connection $\nabla$ on $E$ lifts to a connection on $r^* (E)$. Its restriction to any fiber $\mathcal{G}_x$ defines a linear connection in the usual sense, which is denoted by $\nabla_x$. It is easy to see that these connections are right invariant in the sense that

\begin{equation}
R_g^* \nabla_x = \nabla_y, \quad \forall g \in \mathcal{G} \text{ such that } r(g) = x \text{ and } d(g) = y.
\end{equation}

The bundles considered above will thus have invariant connections.

The bundle $A(\mathcal{G})$, called the Lie algebroid of $\mathcal{G}$, plays in the theory of almost differentiable groupoids the rôle Lie algebras play in the theory of Lie groups. We recall for the benefit of the reader the definition of a Lie algebroid [31].

**Definition 4.** A Lie algebroid $A$ over a manifold $M$ is a vector bundle $A$ over $M$ together with a Lie algebra structure on the space $\Gamma(A)$ of smooth sections of $A$, and a bundle map $\rho : A \to TP$, extended to a map between sections of these bundles, such that

(i) $\rho([X,Y]) = [\rho(X), \rho(Y)];$ and

(ii) $[X,fY] = f[X,Y] + (\rho(X)f)Y$

for any smooth sections $X$ and $Y$ of $A$ and any smooth function $f$ on $M$.

Note that we allow the base $M$ in the definition above to be a manifold with corners.

If $\mathcal{G}$ is an almost differentiable groupoid, then $A(\mathcal{G})$ will naturally have the structure of a Lie algebroid [19]. Let us recall how this structure is defined (the original definition easily extends to include manifolds with corners). Clearly $A(\mathcal{G})$ is a vector bundle. The right translation by an arrow $g \in \mathcal{G}$ defines a diffeomorphism $R_g : \mathcal{G}_{r(g)} \ni g' \rightarrow g' g \in \mathcal{G}_{d(g)}$. This allows us to talk about right invariant differential geometric quantities as long as they are completely determined by their restriction to all submanifolds $\mathcal{G}_x$. This is true of functions and $d$–vertical vector fields, and this is all that is needed to...
define the Lie algebroid structure on \( A(G) \). The sections of \( A(G) \) are in one-to-one correspondence with vector fields \( X \) on \( G \) that are \( d \)-vertical, in the sense that \( d_x(X(g)) = 0 \), and right invariant. The condition \( d_x(X(g)) = 0 \) means that \( X \) is tangent to the submanifolds \( \mathcal{G}_x \), the fibers of \( d \). The Lie bracket \([X,Y]\) of two \( d \)-vertical right-invariant vector fields \( X \) and \( Y \) will also be \( d \)-vertical and right-invariant, and hence the Lie bracket induces a Lie algebra structure on the sections of \( A(G) \). To define the action of the sections of \( A(G) \) on functions on \( G^{(0)} \), observe that the right invariance property makes sense also for functions on \( G \) and that \( C^\infty(G^{(0)}) \) may be identified with the subspace of right–invariant functions on \( G \). If \( X \) is a right–invariant vector field on \( G \) and \( f \) is a right-invariant function on \( G \), then \( X(f) \) will still be a right invariant function. This identifies the action of \( \Gamma(A(G)) \) on functions on \( G^{(0)} \).

Not every Lie algebroid is the Lie algebroid of a Lie groupoid (see [1] for an example). However, every Lie algebroid is associated to a local Lie groupoid [32]. The definition of a local Lie (or more generally, almost differentiable) groupoid [10] is obtained by relaxing the condition that the multiplication \( \mu \) be everywhere defined on \( G^{(2)} \) (see Equation (2)), and replacing it by the condition that \( \mu \) be defined in a neighborhood \( U \) of the set of units.

**Definition 5** (van Est). An almost differentiable local groupoid \( L = (L^{(0)}, L^{(1)}) \) is a pair of manifolds with corners together with structural morphisms \( d, r : L^{(1)} \to L^{(0)}, \iota : L^{(1)} \to L^{(1)}, u : L^{(0)} \to L^{(1)} \) and \( \mu : U \to L^{(1)} \), where \( U \) is a neighborhood of \( (u \times u)(L^{(0)}) = \{(u(x), u(x))\} \) in \( L^{(2)} = \{(g, h), d(g) = r(h)\} \subset L^{(1)} \times L^{(1)} \). The structural morphisms are required to be differentiable maps such that \( d \) is a submersion, \( u \) is an embedding, and to satisfy the following properties:

(i) The products \( u(d(g))g, gu(r(g)), g^{-1}g \) and \( g^{-1}g \) are defined and coincide with, respectively, \( g, g, u(r(g)) \) and \( u(d(g)) \), where we denoted \( g^{-1} = \iota(g) \) as usual.

(ii) If \( gh \) is defined, then \( h^{-1}g^{-1} \) is defined and equal to \( (gh)^{-1} \).

(iii) (Local associativity) If \( gg', g'g'' \), and \( (gg')g'' \) are defined, then \( g(g'g'') \) is also defined and equal to \( (gg')g''. \)

The set \( U \) is the set of arrows for which the product \( gh = \mu(g, h) \) is defined.

We see that the only difference between a groupoid and a local groupoid \( \mathcal{L} \) is the fact that the condition \( d(g) = r(h) \) is necessary for the product \( gh = \mu(g, h) \) to be defined, but not sufficient in general. The product is defined as soon as the arrows \( g \) and \( h \) are “small enough.” A consequence of this definition is that the right multiplication by an arrow \( g \in L^{(1)} \) defines only a diffeomorphism

\[
U_{g^{-1}} : g' \to g'g \in U_g
\]
of an open (and possibly empty) subset $U_{g-1}$ of $L_y$, $y = r(g)$ to an open subset $U_g \subset L_x$, $x = d(g)$. This will not affect the considerations above, however, so we can associate a Lie algebroid $A(L)$ to any almost differentiable local groupoid $L$.

In the following, when considering groupoids, we shall sometimes refer to them as *global* groupoids, in order to stress the difference between groupoids and local groupoids.

### 2. Main definition.

Consider a complex vector bundle $E$ on the space of units $G^{(0)}$ of an almost differentiable groupoid $G$. Denote by $r^*(E)$ its pull-back to $G^{(1)}$. Right translations on $G$ define linear isomorphisms

\[
U_g : C^\infty(G_{d(g)}, r^*(E)) \to C^\infty(G_{r(g)}, r^*(E))
\]

\[
(U_g f)(g') = f(g'g) \in (r^*E)_{g'}
\]

which makes sense because $(r^*E)_{g'} = (r^*E)g'g = E_{r(g')}$.

If $G$ is merely a local groupoid, then (6) is replaced by the isomorphisms

\[
U_g : C^\infty(U_g, r^*(E)) \to C^\infty(U_{g-1}, r^*(E))
\]

defined for the open subsets $U_g \subset G_{d(g)}$ and $U_{g-1} \subset G_{r(g)}$ defined in (5).

Let $B \subset \mathbb{R}^n$ be an open subset. Define the space $S^m_n(B \times \mathbb{R}^n)$ of symbols on the bundle $B \times \mathbb{R}^n \to B$ as in [14] to be the set of smooth functions $a : B \times \mathbb{R}^n \to \mathbb{C}$ such that

\[
|\partial_y^\alpha \partial_{\xi}^\beta a(y, \xi)| \leq C_{K,\alpha,\beta} (1 + |\xi|)^{m-|\beta|}
\]

for any compact set $K \subset B$ and any multi-indices $\alpha$ and $\beta$. An element of one of our spaces $S^m_n$ should properly be said to have “order less than or equal to $m$”; however, by abuse of language we will say that it has “order $m$”.

A symbol $a \in S^m_n(B \times \mathbb{R}^n)$ is called *classical* if it has an asymptotic expansion as an infinite sum of homogeneous symbols $a \sim \sum_{k=0}^{\infty} a_{m-k}$, $a_l$ homogeneous of degree $l$: $a_l(y, t\xi) = t^l a_l(y, \xi)$ if $\|\xi\| \geq 1$ and $t \geq 1$. (“Asymptotic expansion” is used here in the sense that $a - \sum_{k=0}^{N-1} a_{m-k}$ belongs to $S^{m-N}_n(B \times \mathbb{R}^n)$.) The space of classical symbols will be denoted by $S^m_{cl}(B \times \mathbb{R}^n)$. We shall be working exclusively with classical symbols in this paper.

This definition immediately extends to give spaces $S^m_{cl}(E; F)$ of symbols on $E$ with values in $F$, where $\pi : E \to B$ and $F \to B$ are smooth euclidian vector bundles. These spaces, which are independent of the metrics used in their definition, are sometimes denoted $S^m_{cl}(E; \pi^*(F))$. Taking $E = B \times \mathbb{R}^n$ and $F = \mathbb{C}$ one recovers $S^m_{cl}(B \times \mathbb{R}^n) = S^m_{cl}(B \times \mathbb{R}^n; \mathbb{C})$. 

\[\text{PSEUDODIFFERENTIAL OPERATORS ON GROUPOIDS 123}\]
A pseudodifferential operator $P$ on $B$ is a linear map $P : C_c^\infty(B) \to C^\infty(B)$ that is locally of the form $P = a(y, D_y)$ plus a regularizing operator, where for any complex valued symbol $a$ on $T^*W = W \times \mathbb{R}^n$, $W$ an open subset of $\mathbb{R}^n$, one defines $a(y, D_y) : C_c^\infty(W) \to C^\infty(W)$ by

\begin{equation}
(9) \quad a(y, D_y)u(y) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{iy\cdot\xi}a(y, \xi)\hat{u}(\xi) d\xi.
\end{equation}

Recall that an operator $T : C_c^\infty(U) \to C^\infty(V)$ is called regularizing if and only if it has a smooth distribution (or Schwartz) kernel. This happens if and only if $T$ is pseudodifferential of order $-\infty$.

The class of $a$ in $S^m_\cl(T^*W)/S^{m-1}_\cl(T^*W)$ does not depend on any choices; the collection of all these classes, for all coordinate neighborhoods $W$, patch together to define a class $\sigma_m(P) \in S^m_\cl(T^*W)/S^{m-1}_\cl(T^*W)$, which is called the principal symbol of $P$. If the operator $P$ acts on sections of a vector bundle $E$, then the principal symbol $\sigma_m(P)$ will belong to $S^m_\cl(T^*B; \text{End}(E))/S^{m-1}_\cl(T^*B; \text{End}(E))$. See [14] for more details on all these constructions.

We shall sometimes refer to pseudodifferential operators acting on a smooth manifold as ordinary pseudodifferential operators, in order to distinguish them from pseudodifferential operators on groupoids, a class of operators, which we now define (and which are really families of ordinary pseudodifferential operators).

Throughout this paper, we shall denote by $(P_x, x \in G^{(0)})$ a family of order $m$ pseudodifferential operators $P_x$, acting on the spaces $C_c^\infty(G_x, r^*(E))$ for some vector bundle $E$ over $G^{(0)}$. Operators between sections of two different vector bundles $E_1$ and $E_2$ are obtained by considering $E = E_1 \oplus E_2$.

**Definition 6.** A family $(P_x, x \in G^{(0)})$ as above is called differentiable if for any open set $V \subset G$, diffeomorphic through a fiber preserving diffeomorphism to $d(V) \times W$, for some open subset $W \subset \mathbb{R}^n$, and for any $\phi \in C_c^\infty(V)$, we can find $a \in S^m_\cl(d(V) \times T^*W; \text{End}(E))$ such that $\phi P_x \phi$ corresponds to $a(x, y, D_y)$ under the diffeomorphism $G_x \cap V \simeq W$, for each $x \in d(V)$.

A fiber preserving diffeomorphism is a diffeomorphism $\psi : d(V) \times W \to V$ satisfying $\psi(x, w) = x$. Thus we require that the operators $P_x$ be given in local coordinates by symbols $a_x$ that depend smoothly on all variables, in particular, on $x \in G^{(0)}$.

**Definition 7.** An order $m$ invariant pseudodifferential operator $P$ on an almost differentiable groupoid $G$, acting on sections of the vector bundle $E$, is a differentiable family $(P_x, x \in G^{(0)})$ of order $m$ classical pseudodifferential operators $P_x$ acting on $C_c^\infty(G_x, r^*(E))$ and satisfying

\begin{equation}
P_{r(g)}U_g = U_gP_{d(g)} \quad \text{(invariance)}
\end{equation}

for any $g \in G^{(1)}$, where $U_g$ is as in (6).
Replacing the coefficient bundle $E$ by $E \otimes \mathcal{D}^\lambda$ and using the isomorphism
\[ \Omega^\lambda_d \simeq r^*(\mathcal{D}^\lambda), \]
we obtain operators acting on sections of density bundles. Note that $P$ can generally not be considered as a single pseudodifferential operator on $\mathcal{G}^{(1)}$. This is because a family of pseudodifferential operators on a smooth manifold $M$, parametrized by a smooth manifold $B$, is not a pseudodifferential operator on the product $M \times B$, although it acts naturally on $C^\infty_c(M \times B)$. (See [2] or [14], page 94.)

Recall [13] that distributions on a manifold $Y$ with coefficients in the bundle $E_0$ are continuous linear maps $C^\infty_c(Y; E_0 \otimes \Omega) \to \mathbb{C}$, where $E_0^*$ is the dual bundle to $E_0$ and $\Omega = \Omega(Y)$ is the space of 1-densities on $Y$. The collection of all distributions on $Y$ with coefficients in the (finite dimensional complex vector) bundle $E_0$ is denoted $C^\infty(Y; E_0)$.

If $P = (P_x, x \in \mathcal{G}^{(0)})$ is a family of pseudodifferential operators acting on $\mathcal{G}_x$, we denote by $k_x$ the distribution kernel of $P_x$

\[ k_x \in C^\infty_c(\mathcal{G}_x \times \mathcal{G}_x; r^*_1(E) \otimes r^*_2(E)' \otimes \Omega_2). \tag{11} \]

Here $\Omega_2$ is the pull-back of the bundle of vertical densities $\Omega_1$ on $\mathcal{G}_x$ to $\mathcal{G}_x \times \mathcal{G}_x$ via the second projection. These distribution kernels are obtained using Schwartz’ kernel theorem. We define the support of the operator $P$ to be

\[ \text{supp}(P) = \bigcup_x \text{supp}(k_x). \tag{12} \]

The support of $P$ is contained in the closed subset \( \{(g, g'), d(g) = d(g')\} \) of the product $\mathcal{G}^{(1)} \times \mathcal{G}^{(1)}$. In particular, \( (id \times \iota)(\text{supp}(P)) \subset \mathcal{G}^{(2)} \). If all operators $P_x$ are of order $-\infty$, then each kernel $k_x$ is a smooth section. Actually we have more:

**Lemma 1.** The collection of all distribution kernels $k_x$ of a differentiable family $P = (P_x, x \in \mathcal{G}^{(0)})$ of order $-\infty$ operators defines a smooth section $k$ of $r^*_1(E) \otimes r^*_2(E)' \otimes \Omega_2$ on \( \{(g, g'), d(g) = d(g')\} \).

**Proof.** Indeed if $\psi : d(V) \times W \to V$ is a fiber preserving diffeomorphism as in Definition 6, then it follows from the definition that $k$ is smooth on $d(V) \times W \times W \subset \{(g, g'), d(g) = d(g')\}$. Since in this way we obtain an atlas of $\{(g, g'), d(g) = d(g')\}$, we obtain that $k$ is smooth as claimed. \( \square \)

**Definition 8.** The family $P = (P_x, x \in \mathcal{G}^{(0)})$ is properly supported if $p_1^{-1}(K) \cap \text{supp}(P)$ is a compact set for any compact subset $K \subset \mathcal{G}$, where $p_1, p_2 : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ are the two projections. The family $P$ is called compactly supported if its support $\text{supp}(P)$ is compact; and, finally, $P$ is called uniformly supported if its reduced support $\text{supp}_\mu(P) = \mu_1(\text{supp}(P))$ is a compact subset of $\mathcal{G}^{(1)}$, where $\mu_1(g', g) = g'g^{-1}$. 

It immediately follows from the definition that a uniformly supported operator is also properly supported, and that a compactly supported operator is uniformly supported. If the family $P = (P_x, x \in \mathcal{G}^{(0)})$ is properly supported, then each $P_x$ is properly supported, but the converse is not true.

Recall that the composition of two ordinary pseudodifferential operators is defined if one of them is properly supported. It follows that we can define the composition $PQ$ of two properly supported families of operators $P = (P_x, x \in \mathcal{G}^{(0)})$ and $Q = (Q_x, x \in \mathcal{G}^{(0)})$ on $\mathcal{G}^{(1)}$ by pointwise composition $PQ = (P_xQ_x, x \in \mathcal{G}^{(0)})$. The action on sections of $r^*(E)$ is also defined pointwise as follows. For any smooth section $f \in C^\infty(G, r^*(E))$ denote by $f_x$ the restriction $f |_{\mathcal{G}_x}$. If each $f_x$ has compact support and $P = (P_x, x \in \mathcal{G}^{(0)})$ is a family of ordinary pseudodifferential operators, then we define $Pf$ by $(Pf)_x = P_x(f_x)$.

**Lemma 2.** (i) If $f \in C^\infty_c(G, r^*(E))$ and $P = (P_x, x \in \mathcal{G}^{(0)})$ is a differentiable family of ordinary pseudodifferential operators, then $Pf \in C^\infty_c(G, r^*(E))$. If $P$ is also properly supported, then $Pf \in C^\infty_c(G, r^*(E))$.

(ii) The composition $PQ = (P_xQ_x, x \in \mathcal{G}^{(0)})$ of two properly supported differentiable families of operators $P = (P_x, x \in \mathcal{G}^{(0)})$ and $Q = (Q_x, x \in \mathcal{G}^{(0)})$ is a properly supported differentiable family.

**Proof.** If $P$ consists of regularizing operators, then

$$Pf(g) = \int_{\mathcal{G}_x} k_x(g, h)f(h), \text{ where } x = d(g).$$

Lemma 1 implies that the formula above for $Pf$ involves only the integration of smooth (uniformly in $g$) compactly supported sections, and hence we can exchange integration and differentiation to obtain the smoothness of $Pf$. This proves (i) in case $P$ consists of regularizing operators. The proof of (ii) if both $P$ and $Q$ consist of regularizing operators follows the same reasoning.

We prove now (i) for $P$ arbitrary. Fix $g \in \mathcal{G}_x$ and $V$ a neighborhood of $g$ fiber preserving diffeomorphic to $d(V) \times W$ for some open convex subset $W$ in $\mathbb{R}^n$, $0 \in W$, such that $(x, 0)$ maps to $g$. Replacing $P_x$ by $P_x - R_x$ for a smooth regularizing family $R_x$ we can assume that the distribution kernels $k_x$ of $P_x$ satisfy

$$p_1^{-1}(d(V) \times W/4) \cap \overline{\text{supp}(k_x)} \subset (d(V) \times W/4) \times (d(V) \times W/2).$$

The smoothness of $Pf$, respectively of $PQ$ if $Q$ consists of regularizing operators, reduces in this way to a computation in local coordinates. This completes the proof of (i) in general, and of (ii) if $Q$ is regularizing.

For arbitrary $Q$ we can replace $Q$, in view of what has already been proved, with $Q - R$, where $R$ is a regularizing family. In this way we may assume that

$$p_1^{-1}(d(V) \times W/2) \cap \overline{\text{supp}(k_x)} \subset (d(V) \times W/2) \times (d(V) \times 3W/4),$$
where \( k'_y \) are the distribution kernels of \( Q_x \). The support estimates above for \( P \) and \( Q \) show that the \( P_yQ_y \) for \( y \in d(V) \) are the compositions of smooth families of pseudodifferential operators acting on \( W \subset \mathbb{R}^n \). The result is then known. \( \square \)

The smaller class of uniformly supported operators is also closed under composition.

**Lemma 3.** The composition \( PQ = (P_xQ_x, x \in G^{(0)}) \) of two uniformly supported families of operators \( P = (P_x, x \in G^{(0)}) \) and \( Q = (Q_x, x \in G^{(0)}) \) is uniformly supported.

**Proof.** The reduced support \( \text{supp}_\mu(PQ) \) (see (12)) of the composition \( PQ \) satisfies

\[
\text{supp}_\mu(PQ) \subset \mu\left( \text{supp}_\mu(P) \times \text{supp}_\mu(Q) \right),
\]

where \( \mu \) is the composition of arrows. Since \( \text{supp}_\mu(P) \) and \( \text{supp}_\mu(Q) \) are compact, the equation above completes the proof of the lemma. \( \square \)

Let \( G \) be an almost differentiable groupoid. The space of order \( m \), invariant, uniformly supported pseudodifferential operators on \( G \), acting on sections of the vector bundle \( E \) will be denoted by \( \Psi^m(G;E) \). We denote \( \Psi^\infty(G;E) = \bigcup_{m \in \mathbb{Z}} \Psi^m(G;E) \) and \( \Psi^{-\infty}(G;E) = \bigcap_{m \in \mathbb{Z}} \Psi^m(G;E) \). Thus an operator \( P \in \Psi^m(G;E) \) is actually a differentiable family \( P = (P_x, x \in G^{(0)}) \) of ordinary pseudodifferential operators.

**Theorem 1.** The set \( \Psi^\infty(G;E) \) of uniformly supported invariant pseudodifferential operators on an almost differentiable groupoid \( G \) is a filtered algebra, i.e.

\[
\Psi^m(G;E)\Psi^{m'}(G;E) \subset \Psi^{m+m'}(G;E).
\]

In particular, \( \Psi^{-\infty}(G;E) \) is a two-sided ideal.

**Proof.** Let \( P = (P_x, x \in G^{(0)}) \) and \( Q = (Q_x, x \in G^{(0)}) \) be two invariant uniformly supported pseudodifferential operators on \( G \), of order \( m \) and \( m' \) respectively. Their composition \( PQ = (P_xQ_x) \), is a uniformly supported operator of order \( m + m' \), in view of Lemma 3. It is also a differentiable family due to Lemma 2. We now check the invariance condition. Let \( g \) be an arbitrary arrow and \( U_g : C^\infty_c(G_x, r^*(E)) \to C^\infty_c(G_y, r^*(E)), x = d(g) \) and \( y = r(g) \), be as in the definition above. Then

\[
(PQ)_y U_g = P_yQ_y U_g = P_yU_yQ_x = U_gP_xQ_x = U_g(PQ)_x.
\]

This proves the theorem. \( \square \)

Properly supported invariant differentiable families of pseudodifferential operators also form a filtered algebra, denoted \( \Psi^\infty_{\text{prop}}(G;E) \). While it is clear that in order for our class of pseudodifferential operators to form an algebra we need some condition on the support of their distribution kernels, exactly
what support condition to impose is a matter of choice. We prefer the uniform support condition because it leads to a better control at infinity of the family of operators $P = (P_x, x \in \mathcal{G}^{(0)})$ and allows us to identify the regularizing ideal (i.e. the ideal of order $-\infty$ operators) with the groupoid convolution algebra of $\mathcal{G}$. The choice of uniform support will also ensure that $\Psi^m(\mathcal{G}; E)$ behaves functorially with respect to open embeddings. The compact support condition enjoys the same properties but is usually too restrictive. The issue of support will be discussed again in examples.

The definition of the principal symbol extends easily to $\Psi^m(\mathcal{G}; E)$. Denote by $\pi : A^* (\mathcal{G}) \to M$, ($M = \mathcal{G}^{(0)}$) the projection. If $P = (P_x, x \in \mathcal{G}^{(0)}) \in \Psi^m(\mathcal{G}; E)$ is an order $m$ pseudodifferential differential operator on $\mathcal{G}$, then the principal symbol $\sigma_m(P)$ of $P$ will be represented by sections of the bundle $\text{End}(\pi^* E)$ and will be defined to satisfy

\begin{equation}
\sigma_m(P)(\xi) = \sigma_m(P_x)(\xi) \in \text{End}(E_x) \text{ if } \xi \in A^*_x(\mathcal{G}) = T^*_x \mathcal{G}_x
\end{equation}

(the equation above is mod $S^m_{cl}(A^*_x(\mathcal{G}); \text{End}(E)))$. This equation will obviously uniquely determine a linear map

$$
\sigma_m : \Psi^m(\mathcal{G}) \to S^m_{cl}(A^*(\mathcal{G}); \text{End}(E))/S^{m-1}_{cl}(A^*(\mathcal{G}); \text{End}(E))
$$

provided we can show that for any $P = (P_x, x \in \mathcal{G}^{(0)})$ there exists a symbol $a \in S^m_{cl}(A^*(\mathcal{G}); \text{End}(E))$ whose restriction to $A^*_x(\mathcal{G})$ is a representative of the principal symbol of $P_x$ in that fiber for each $x$. We thus need to choose for each $P_x$ a representative $a_x \in S^m_{cl}(A^*_x(\mathcal{G}); \text{End}(E))$ of $\sigma_m(P_x)$ such that the family $a_x$ is smooth and invariant. Assume first that $E$ is the trivial line bundle and proceed as in [14] Section 18.1, especially Equation (18.1.27) and below.

Choose a connection $\nabla$ on the vector bundle $A(\mathcal{G}) \to \mathcal{G}^{(0)}$ and consider the pull-back vector bundle $r^*(A) \to \mathcal{G}$ of $A(\mathcal{G}) \to \mathcal{G}^{(0)}$ endowed with the pull-back connection $\tilde{\nabla} = r^* \nabla$. Its restriction on any fiber $\mathcal{G}_x$ defines a linear connection in the usual sense, which is denoted by $\nabla_x$. These connections are right invariant in the sense that

\begin{equation}
R^*_y \nabla_x = \nabla_y, \quad \forall g \in \mathcal{G} \text{ such that } r(g) = x \text{ and } d(g) = y.
\end{equation}

Using such an invariant connection, we may define the exponential map of a Lie algebroid, which generalizes the usual exponential map of a manifold with a connection and the exponential map of a Lie algebra as follows. For any $x \in \mathcal{G}^{(0)}$, define a map $\text{exp}_x : A_x \to \mathcal{G}$ as the composition of the maps:

$$
A_x \xrightarrow{i} T_x \mathcal{G}_x \xrightarrow{\text{exp}_x} \mathcal{G},
$$

where $i$ is the natural inclusion and $\text{exp}_x = \text{exp}_{\nabla_x}$ is the usual exponential map at $x \in \mathcal{G}_x$ on the manifold $\mathcal{G}_x$. By varying the point $x$, we obtain a map
exp_{\nabla} defined in a neighborhood of the zero section, called the exponential map of the Lie algebroid\(^2\). Clearly, exp_{\nabla} is a local diffeomorphism

\begin{equation}
A(\mathcal{G}) \ni V_0 \ni v \mapsto \exp_{\nabla}(v) = y \in V \subset \mathcal{G}
\end{equation}

mapping an open neighborhood \(V_0\) of the zero section in \(A(\mathcal{G})\) diffeomorphically to a neighborhood \(V\) of \(\mathcal{G}^{(0)}\) in \(\mathcal{G}\), and sending the zero section onto the set of units. Choose a cut-off function \(\phi \in C^\infty(\mathcal{G})\) with support in \(V\) and equal to 1 in a smaller neighborhood of \(\mathcal{G}^{(0)}\) in \(\mathcal{G}\). If \(y \in V\), \(x = d(y)\) and \(\xi \in A_x^*(\mathcal{G})\) let \(v \in V_0\) be the unique vector \(v \in A_x(\mathcal{G})\) such that \(y = \exp_{\nabla}(v)\) and denote \(e_\xi(y) = \phi(y)e^{iv \xi}\), which extends then to all \(y \in \mathcal{G}\) due to the cut-off function \(\phi\). Define the \((\nabla, \phi)\)-complete symbol \(\sigma_{\nabla, \phi}(P)\) by

\begin{equation}
\sigma_{\nabla, \phi}(P)(\xi) = (P_x e_\xi)(x), \quad \forall \xi \in T^*_x \mathcal{G}_x = A_x^*(\mathcal{G}).
\end{equation}

**Lemma 4.** If \(P = (P_x, x \in \mathcal{G}^{(0)})\) is an operator in \(\Psi^m(\mathcal{G})\), then the function \(\sigma_{\nabla, \phi}(P)\) defined above is differentiable and defines a symbol in \(S^m_{cl}(A^*(\mathcal{G}))\). Moreover, if \((\nabla_1, \phi_1)\) is another pair consisting of an invariant connection \(\nabla_1\) and a cut-off function \(\phi_1\), then \(\sigma_{\nabla_1, \phi_1}(P) - \sigma_{\nabla, \phi}(P)\) is in \(S^{-\infty}_{cl}(A^*(\mathcal{G}))\) and \(\sigma_{\nabla, \phi}(P) - \sigma_{\nabla_1, \phi_1}(P)\) is in \(S^{-m}_{cl}(A^*(\mathcal{G}))\).

**Proof.** For each \(\xi \in A_x^*\) the function \(e_\xi\) is smooth with compact support on \(\mathcal{G}_x\), so \(P_x e_\xi\) is defined. Equation (18.1.27) of [14] shows that \(a(\xi) = \sigma_{\nabla, \phi}(P)(\xi)\) is the restriction of the complete symbol of \(P_x \phi\) to \(T^*_x \mathcal{G}_x\) if the complete symbol is defined in the normal coordinate system at \(x \in \mathcal{G}_x\) (given by the exponential map). The normal coordinate system defines, using a local trivialization of \(A(\mathcal{G})\), a fiber preserving diffeomorphism \(\psi : d(V) \times W \to V\) for some open subset \(W\) of \(\mathbb{R}^n\) (i.e. satisfying \(d(\psi(x, w)) = x\)). From the definition of the smoothness of the family \(P_x\) (Definition 6) it follows that the complete symbol of \(P \phi\) is in \(S^m_{cl}(d(V) \times T^*W)\) if the support of \(\phi\) is chosen to be in \(V\). This proves that \(\sigma_{\nabla, \phi}(P)\) is in \(S^m_{cl}(A^*(\mathcal{G}))\).

The rest follows in exactly the same way. \(\square\)

The lemma above justifies the following definition of the principal symbol as the class of \(\sigma_{\nabla, \phi}(P)\) modulo terms of lower order (for the trivial line bundle \(E = \mathbb{C}\)). This definition will be, in view of the same lemma, independent on the choice of \(\nabla\) or \(\phi\) and will satisfy Equation (13). If \(E\) is not trivial one can still define a complete symbol \(\sigma_{\nabla', \phi}(P)\), depending also on a second connection \(\nabla'\) on the bundle \(E\), which is used to trivialize \(r^*(E)\) on \(V \subset \mathcal{G}\) (assuming also that \(V_0\) is convex). Alternatively, we can use Proposition 3 below.

\(^2\)See [17] for an alternative definition of the exponential map. One should not confuse this map with the exponential map from \(\Gamma(A)\) to the bisections of the groupoid as defined in [16].
Proposition 1. Let $\nabla$ and $\phi$ be as above. The choice of a connection $\nabla'$ on $E$ defines a complete symbol map $\sigma_{\nabla, \nabla', \phi}: \Psi^m(G; E) \to S^m_{\cl}(A^*(G))$. The principal symbol $\sigma_m: \Psi^m(G; E) \to S^m_{\cl}(A^*(G))/S^{m-1}_{\cl}(A^*(G))$, defined by

\begin{equation}
\sigma_m(P) = \sigma_{\nabla, \nabla', \phi}(P) + S^{m-1}_{\cl}(A^*(G)),
\end{equation}

does not depend on the choice of the connections $\nabla$, $\nabla'$ or the cut-off function $\phi$.

Proof. The $(\nabla, \nabla', \phi)$–complete symbol $\sigma_{\nabla, \nabla', \phi}(P)$ is defined as follows. Let $w$ be a vector in $E_x$. Using the connection $\nabla'$ we can define a section $\bar{w}$ of $\rho^*(E)$ on $G_x \cap V$ by parallel transport along the geodesics of $\nabla$ starting at $x$, and which coincides with $w$ at $x$. Then denote $e_{\xi, w} = e_{\xi} \bar{w}$ and let

\begin{equation}
\sigma_{\nabla, \nabla', \phi}(P)(\xi) w = (P_x e_{\xi, w})(x) \in E_x, \quad \forall \xi \in T^*_x G_x = A^*_x(G).
\end{equation}

The rest of the proof proceeds along the lines of the proof of Lemma 4. □

Note that the principal symbol of $P$ determines the principal symbols of the individual operators $P_x$ by the invariance with respect to right translations. Precisely, we have $\sigma_m(P_x) = \rho^*(\sigma(P))|_{T^*_x G_x}$.

The following result extends some very well known properties of the calculus of pseudodifferential operators on smooth manifolds. We shall prove the surjectivity of the principal symbol in Section 5.

Proposition 2. (i) The principal symbol map

$$\sigma_m: \Psi^m(G; E) \to S^m_{\cl}(G; \End(E))/S^{m-1}_{\cl}(G; \End(E))$$

has kernel $\Psi^{m-1}(G; E)$ and satisfies Equation (13).

(ii) The composition $PQ$ of two operators $P, Q \in \Psi^\infty(G; E)$, of orders $m$ and, respectively, $m'$, satisfies $\sigma_{m+m'}(PQ) = \sigma_m(P)\sigma_{m'}(Q)$.

Proof. (i) The operator $P = (P_x, x \in G^{(0)}) \in \Psi^m(G; E)$ is in the kernel of $\sigma_m$ if and only if all symbols $\sigma_m(P_x)$ vanish. This implies $P_x \in \Psi^{m-1}(G; E)$ for all $x$ and hence $P = (P_x, x \in G^{(0)}) \in \Psi^m(G; E)$. As already observed for $E$ a trivial line bundle, the fact that Equation (13) is satisfied was contained in the proof of Lemma 4. The general case is similar or can be proved using Proposition 3.

The second statement is known for pseudodifferential operators on smooth manifolds [14]; this accounts for the second equality sign in the next equation. We obtain using Equation (13) that

$$\sigma_{m+m'}(PQ)(v) = \sigma_{m+m'}(P_x Q_x)(v) = \sigma_m(P_x)\sigma_{m'}(Q_x)(v) = \sigma_m(P_x)(v)\sigma_{m'}(Q_x)(v),$$

where $v \in A^*_x(G)$. □
Although for the most of this paper we shall be concerned with groupoids, the definition of $\Psi^m(G; E)$ easily extends to local groupoids. Indeed it suffices to modify the invariance condition in Definition 7, using the notation in Equation (7), as follows. We assume that for any $g \in G^{(1)}$ and any smooth compactly supported function $\phi$ on $U_g$ there exists a regularizing operator $R_{g,\phi}$ such that

$$U_g(\phi)P_{r(g)}U_g f - U_g(\phi P_{d(g)} f) = R_{g,\phi} f$$

for any function $f \in \mathcal{C}_c^\infty(U_g)$. We thus replace the strict invariance of the original definition by ‘invariance up to regularizing operators’.

We denote by $\Psi^m_{\text{loc}}(G; E)$ the set of differentiable properly supported families $P = (P_x, x \in G^{(0)})$ of order $m$ pseudodifferential operators satisfying the condition (19) above. Note that, if we regard an almost differentiable groupoid $G$ as a local groupoid, then $\Psi^\infty(G; E) \subset \Psi^m_{\text{loc}}(G; E)$. The inclusion is generally a strict one, though, because Equation (19) gives no condition for order $-\infty$ operators, and so $\Psi^\infty_{\text{loc}}(G; E)$ consists of arbitrary smooth families $P = (P_x, x \in G^{(0)})$ of regularizing operators. This ideal is too big to reflect the structure of $G$. The “symbolic” part remains however the same:

$$\Psi^m_{\text{prop}}(G; E)/\Psi^\infty_{\text{prop}}(G; E) \simeq \Psi^m_{\text{loc}}(G; E)/\Psi^\infty_{\text{loc}}(G; E).$$

If the sets $U_g$ are all connected (in which case the local groupoid $G$ is said to be $d$–connected) an easier condition to use than (19) is

$$[X, P] \in \Psi^\infty_{\text{loc}}(G; E)$$

for all $r$–vertical left–invariant vector fields $X$ on $G^{(1)}$. With this, the following analog of Theorem 1, becomes straightforward.

**Theorem 2.** Assume that $G$ is a $d$–connected almost differentiable groupoid. Then the space $\Psi^\infty_{\text{loc}}(G; E)$ is a filtered algebra, with $\Psi^\infty_{\text{loc}}(G; E)$ as residual ideal.

**Proof.** The only thing to check is that $\Psi^\infty_{\text{loc}}(G; E)$ is closed under composition. The composition of two differentiable, properly supported families $P, Q \in \Psi^\infty_{\text{loc}}(G; E)$ is again differentiable and properly supported, as has already been proved. The infinitesimal invariance condition $[X, PQ] = [X, P]Q + P[X, Q] \in \Psi^\infty_{\text{loc}}(G; E)$ (20) follows from the fact that $\Psi^\infty_{\text{loc}}(G; E)$ is an ideal of $\Psi^\infty_{\text{loc}}(G; E)$. \qed

### 3. Differential operators and quantization.

In this section, we examine the differential operators in $\Psi^\infty(G; E)$, if $G$ is a global groupoid, or in $\Psi^\infty_{\text{loc}}(G; E)$ if $G$ is a local groupoid. We also show how a simple algebraic construction applied to $G$ and to the algebras $\Psi^\infty_{\text{loc}}(G; E)$ leads to a concrete construction of a deformation quantization of the Lie-Poisson structure on the dual of a Lie algebroid.
In this section, \( \mathcal{G} \) will be an almost differentiable local groupoid. This generality is necessary in order to integrate arbitrary Lie algebroids. Nevertheless, when \( A \) is the Lie algebroid of an almost differentiable global groupoid \( \mathcal{G} \) (that is not just a local groupoid), then all results we shall prove for \( \Psi^\infty(\mathcal{G};E) \) in this section extend immediately to \( \Psi^\infty(\mathcal{G};E) \), although we shall not mention this each time.

**Lemma 5.** Let \( P = (P_x, x \in \mathcal{G}^{(0)}) \) be an operator in \( \Psi^\infty_{\text{loc}}(\mathcal{G};E) \). If \( P_x \) is a multiplication operator for all \( x \) in \( \mathcal{G} \), then there exists a smooth endomorphism \( s \) of \( E \) such that \( P_x(g) = s(r(g)) \) for all \( g \in \mathcal{G}_x \). Conversely, every smooth section \( s \) of \( \text{End}(E) \) defines a multiplication operator in \( \Psi^0_{\text{loc}}(\mathcal{G};E) \).

**Proof.** By assumption \( P_x(g) \) is in \( \text{End}(E_{r(g)}) \). The invariance relation shows that \( P_x(g) \) depends only on \( r(g) \). This defines the section \( s \) of \( \text{End}(E) \) such that \( P_x(g) = s(r(g)) \). To show that \( s \) is smooth, we let \( \phi \) be a smooth section of \( E \) over \( \mathcal{G}^{(0)} \) and let \( \tilde{\phi}(g) = \phi(r(g)) \). By assumption \( P \tilde{\phi} \) is smooth and hence \( s\phi = P \tilde{\phi}|_{\mathcal{G}^{(0)}} \) is also smooth. Since \( \phi \) is arbitrary this implies the smoothness of \( s \).

Conversely, if \( s \) is a smooth endomorphism of \( E \), then if we let \( P_x(g) = s(r(x)) \) we obtain a multiplication operator in \( \Psi^0_{\text{loc}}(\mathcal{G};E) \). \( \square \)

The following proposition will allow us to assume that \( E \) is a trivial bundle, which is sometimes useful in applications.

**Proposition 3.** Let \( E \) be a vector bundle on \( \mathcal{G}^{(0)} \) embedded into a trivial hermitian bundle, \( E \subset \mathbb{C}^N \). Denote by \( e_0 \) the projection onto \( E \) regarded as a matrix of multiplication operators in \( M_N(\Psi^0_{\text{loc}}(\mathcal{G})) \), the algebra of \( N \times N \) matrices with values in \( \Psi^0_{\text{loc}}(\mathcal{G}) \). Then \( \Psi^\infty_{\text{loc}}(\mathcal{G};E) \simeq e_0 M_N(\Psi^\infty_{\text{loc}}(\mathcal{G})) e_0 \) as filtered algebras.

**Proof.** The multiplication operator \( e_0 \) defines an element of \( \Psi^0_{\text{loc}}(\mathcal{G};E) \) by the lemma above; hence it acts on all spaces \( C^\infty_c(\mathcal{G}_x, \mathbb{C}^N) \). Then

\[
C^\infty_c(\mathcal{G}_x, r^*(E)) = e_0 C^\infty_c(\mathcal{G}_x, \mathbb{C}^N)
\]

and every pseudodifferential operator \( P_x \) on \( C^\infty_c(\mathcal{G}_x, r^*(E)) \) extends in this way to an operator on \( C^\infty_c(\mathcal{G}_x, \mathbb{C}^N) \). This gives an inclusion \( \Psi^\infty_{\text{loc}}(\mathcal{G};E) \subset e_0 M_N(\Psi^\infty_{\text{loc}}(\mathcal{G})) e_0 \). Conversely if \( P_x \) is a pseudodifferential operator on \( C^\infty_c(\mathcal{G}_x, \mathbb{C}^N) \), then \( e_0 P_x e_0 \) is a pseudodifferential operator on \( C^\infty_c(\mathcal{G}_x, \mathbb{C}^N) \). This gives the opposite inclusion. \( \square \)

The following proposition shows the intimate connection between \( A(\mathcal{G}) \), the Lie algebroid of \( \mathcal{G} \), and \( \Psi^\infty_{\text{loc}}(\mathcal{G}) \) (\( \Psi^\infty(\mathcal{G}) \) if \( \mathcal{G} \) is global). It is morally an equivalent definition of the Lie algebroid associated to an almost differentiable local groupoid.

**Proposition 4.** Let \( \mathcal{G} \) be an almost differentiable local groupoid.
(i) The algebra $C^\infty(G^{(0)})$ is the algebra of multiplication operators in $\Psi^0_{\text{loc}}(G)$.
(ii) The space of sections of the Lie algebroid $A(G)$ can be identified with the space of order 1 differential operators in $\Psi^1_{\text{loc}}(G)$ without constant term.
(iii) The Lie algebroid structure of $A(G)$ is induced by the commutator operations $[,] : \Psi^1_{\text{loc}}(G) \times \Psi^1_{\text{loc}}(G) \to \Psi^1_{\text{loc}}(G)$ and $[,] : \Psi^1_{\text{loc}}(G) \times \Psi^0_{\text{loc}}(G) \to \Psi^0_{\text{loc}}(G)$.

Proof. The first part is a particular case of Lemma 5, only easier. Order 1 differential operators without constant term are vector fields, right invariant by the definition of $\Psi^1_{\text{loc}}(G)$, so they can be identified with the sections of the Lie algebroid $A(G)$ of $G$. This proves (ii). In order to check (iii) recall that, if we regard vector fields on $G$ as linear maps $C^\infty(G) \to C^\infty(G)$, then the Lie bracket coincides with the commutator of linear maps. Moreover the commutator $[X,f]$ of a vector field $X$ and of a multiplication map $f$ is $[X,f] = X(f)$, again regarded as a linear map. Then (iii) follows in view of the discussion above.

The Lie algebroid $A = A(G)$ turns out to determine the structure of the algebra of invariant tangential differential operators on $G$, denoted $\text{Diff}(G)$. We shall see that the subalgebra $\text{Diff}(G) \subset \Psi^\infty_{\text{loc}}(G)$ is a concrete model of the universal enveloping algebra of the Lie algebroid $A$ [15, 36], a concept whose definition we now recall.

Given a Lie algebroid $A \to M$ with anchor $\rho$, we can make the $C^\infty(M)$-module direct sum $C^\infty(M) \oplus \Gamma(A)$ into a Lie algebra over $\mathbb{C}$ by defining

$$[f + X, g + Y] = (\rho(X)g - \rho(Y)f) + [X,Y].$$

Let $U = U(C^\infty(M) \oplus \Gamma(A))$ be its universal enveloping algebra. For any $f \in C^\infty(M)$ and $X \in \Gamma(A)$, denote by $f'$ and $X'$ their canonical image in $U$. Denote by $I$ the two-sided ideal of $U$ generated by all elements of the form $(fg)' - f'g'$ and $(fX)' - f'X'$. Define

$$(21) \quad U(A) = U/I.$$

$U(A)$ is called the universal enveloping algebra of the Lie algebroid $A$. When $A$ is a Lie algebra, this definition reduces to the usual universal enveloping algebra. We shall see, for example, that for the tangent bundle $TM$ this is the algebra of differential operators on $M$.

The maps $f \to f'$ and $X \to X'$ considered above descend to linear embeddings $i_1 : C^\infty(M) \to U(A)$, and $i_2 : \Gamma(A) \to U(A)$; the first map $i_1$ is an algebra morphism. These maps have the following properties:

$$(22) \quad i_1(f)i_2(X) = i_2(fX), \quad [i_2(X), i_1(f)] = i_1(\rho(X)f),$$
$$[i_2(X), i_2(Y)] = i_2([X,Y]).$$
In fact, $U(A)$ is universal among triples $(B, \phi_1, \phi_2)$ having these properties (see [15] for a proof of this easy fact).

In particular, if $M$ is the space of units of an almost differentiable groupoid $G$ with Lie algebroid $A(G)$, the natural morphisms $\phi_1 : C^\infty(M) \to \text{Diff}(G)$ and $\phi_2 : \Gamma(A) \to \text{Diff}(G)$ obtained from Proposition 4 extend to a unique algebra morphism $\tau : U(A) \to \text{Diff}(G)$. (Recall that we denoted by $\text{Diff}(G)$ the algebra of right invariant tangential differential operators on $G$.) Denote by $U_n(A) \subset U(A)$ the space generated by

$$C^\infty(M) \quad \text{and} \quad \text{images of } X_1 \otimes X_2 \otimes \ldots \otimes X_k \in U = U(C^\infty(M) \oplus \Gamma(A)), \text{ for } k \leq n,$$

under the canonical projection $U \to U(A) = U/I$. Then

$$U_0(A) \subset U_1(A) \subset \ldots \subset U_n(A) \subset \ldots$$

is a filtration of $U(A)$. The relations (22) show that, as in the Lie algebra case, the graded algebra $\oplus U_n(A)/U_{n-1}(A)$ is commutative. Similarly, $\text{Diff}(G)$ is naturally filtered by degree.

**Lemma 6.** The map $\tau : U(A) \to \text{Diff}(G)$ maps $U_n(A)$ onto the space $\text{Diff}_n(G)$ of operators of order $\leq n$.

**Proof.** Let $D \in \text{Diff}(G)$ be an invariant tangential differential operator of order $\leq n$. By right invariance $D$ is completely determined by the restrictions $(Du)|_{G(0)}$, $u \in C^\infty_c(G)$. Since $D$ acts on the fibers of $d$ we can write

$$(Du)|_{G(0)} = \sum_{i=1}^{n} D_i u,$$

where $D_i$ is a superposition of derivations $D_i u = X_1^{(i)} X_2^{(i)} \ldots X_k^{(i)} u$ defined using the tangential derivations $X_j^{(i)} \in \Gamma(A)$. By definition it follows that $D$ is the sum of $\tau \left( X_1^{(i)} X_2^{(i)} \ldots X_k^{(i)} \right)$.

Denote by $\text{Symm}(A)$ the symmetric tensor product of the bundle $A$, that is

$$\text{Symm}(A) = \bigoplus_{n=0}^{\infty} S_n(A),$$

where $S_n(A)$ is the symmetric quotient of the bundle $A^\otimes n$, and is isomorphic to the subspace of symmetric tensors, if $S_0(A)$ is the trivial $\mathbb{R}$ bundle by convention. The space $\Gamma(\text{Symm}(A))$ of smooth sections of $\text{Symm}(A)$ identifies with the space of smooth functions on $A^*$ polynomial in each fiber. The complete symbol map $\sigma_{\nabla, \phi}(D)$ of an invariant differential operator $\text{Diff}(G) \subset \Psi^\infty_{\text{loc}}(G)$ (defined in Equation (16)) does not depend on the cut-off function $\phi$ and will be a polynomial in $\xi$, denoted simply by $\sigma_{\nabla}(D)$.

Using the algebra morphism $\tau : U(A) \to \text{Diff}(G)$ obtained from the universality property of $U(A)$, we have the following Poincaré-Birkhoff-Witt
Theorem 3 (Poincaré-Birkhoff-Witt). The composite map,
\[
U(A) \ni D \mapsto \sigma \nabla(\tau(D)) \in \Gamma(\text{Symm}(A)),
\]

is an isomorphism of filtered vector spaces. In particular, \( \tau : U(A) \to \text{Diff}(G) \) is an algebra isomorphism.

Proof. It follows from definitions that the map \( \sigma = \sigma \nabla \circ \tau \) considered in the statement maps \( U_n(A) \) to \( \oplus_{k=0}^n \Gamma(S_k(A)) \) and hence it preserves the filtration. By abuse of notation we shall still denote by \( \sigma \) the induced map \( U_n(A)/U_{n-1}(A) \to \Gamma(S_n(A)) \). It is enough to prove that the map of graded spaces
\[
\sigma : \bigoplus U_n(A)/U_{n-1}(A) \to \bigoplus \Gamma(S_n(A)) = \Gamma(\text{Symm}(A))
\]
is an isomorphism. By Lemma 6 this map is onto. We now prove that it is one-to-one.

The inclusion of \( C^\infty(M) \) in \( U(A) \) makes \( U(A) \) a \( C^\infty(M) \)-bimodule. The filtration \( U_n(A) \) of \( U(A) \) consists of \( C^\infty(M) \)-bimodules. Moreover, since the graded algebra \( \oplus U_n(A)/U_{n-1}(A) \) is commutative, the quotient \( U_n(A)/U_{n-1}(A) \) consists of central elements for this action (i.e. the left and right \( C^\infty(M) \)-module structure coincide). It follows from the definition that the subspace \( \Gamma(A) \otimes C^\infty(M) \) of the universal enveloping algebra \( U = U(C^\infty(M) \oplus \Gamma(A)) \) maps onto \( U_n(A)/U_{n-1}(A) \). The previous discussion shows that this map descends to a map from the tensor product \( \Gamma(A) \otimes C^\infty(M) \cdot \cdot \cdot \otimes C^\infty(M) \) of \( C^\infty(M) \)-modules. By the commutativity of the graded algebra of \( U(A) \) this further descends to a \( C^\infty(M) \)-linear surjective map \( q : \Gamma(S_n(A)) \to U_n(A)/U_{n-1}(A) \).

The composition \( \sigma \circ q : \Gamma(\text{Symm}(A)) \to \Gamma(\text{Symm}(A)) \) is multiplicative since both \( q \) and \( \sigma \) are multiplicative. Moreover \( \sigma \circ q \) is the identity when restricted to \( C^\infty(M) \) (the order 0 elements) and \( \Gamma(A) \) (the elements of order 1). Since these form a system of generators of the commutative algebra \( \Gamma(\text{Symm}(A)) \) it follows that \( \sigma \circ q \) is the identity. This completes the proof. \( \square \)

Remark. The Poincaré-Birkhoff-Witt theorem was proved in the algebraic context by Rinehart [36] for \((L,R)\)-algebras (an algebraic version of Lie algebroids). It essentially stated that the associated graded algebra \( grU(A) = \oplus_n U_{n+1}(A)/U_n(A) \) is isomorphic to the symmetric algebra \( S(\Gamma(A)) = \Gamma(\text{Symm}(A)) \). The role of the connection \( \nabla \) on \( A \to M \) is to establish an explicit isomorphism \( \sigma \nabla \circ \tau \) between \( U(A) \) and \( \Gamma(\text{Symm}(A)) \).

We will now use the results of this and the previous section to obtain an explicit deformation quantization of \( A^* \). In order to do that we need to
establish the relation between commutators and the Poisson bracket in our calculus.

For any \( x \in \mathcal{G}^{(0)} \), \( T^*_d \mathcal{G}_x \) is a symplectic manifold, so \( T^*_d \mathcal{G} \overset{\text{def}}{=} \bigcup_{x \in \mathcal{G}^{(0)}} T^*_d \mathcal{G}_x \) is a regular Poisson manifold with the leafwise symplectic structures. Now the Poisson structure on \( A^* \) can be considered as being induced from that on \( T^*_d \mathcal{G} \). More precisely, let \( \Phi : T^*_d \mathcal{G} \to A^* \) be the natural projection induced by the right translation, used to define a map \( \Phi^* : \mathcal{C}^\infty(A^*(\mathcal{G})) \to \mathcal{C}^\infty(T^*_d \mathcal{G}) \). We then have:

**Lemma 7.** The map \( \Phi \) is a Poisson map.

Of course this lemma is really the definition of the Poisson structure on \( A^* \). The point is to show that the subspace \( \Phi^*(\mathcal{C}^\infty(A^*(\mathcal{G}))) \) of \( \mathcal{C}^\infty(T^*_d \mathcal{G}) \) is closed under the Poisson bracket.

**Proof.** It is enough to check that

\[
\Phi^*(\{f, g\}) = \{\Phi^*(f), \Phi^*(g)\},
\]

where \( f \) and \( g \) are two smooth function on \( A^* \) with polynomial restrictions on each fiber of \( A^* \), that is for \( f \) and \( g \) in \( \Gamma(\text{Sym}(A)) \). Since the Poisson bracket is a derivation in each variable it is further enough to check this for \( f \) constant or \( f \) linear in each fiber. If both \( f \) and \( g \) are constant in each fiber, then both sides of Equation (24) vanish. If \( f \) and \( g \) are of degree one in each fiber, then they correspond to sections \( X \) and \( Y \) of \( A^* \), and their Poisson bracket will identify to \([X, Y]\) (so in particular will also be of degree one in each fiber and this justifies the name of Lie-Poisson structure for this Poisson structure). For this situation the relation (24) follows from the identification of \( \Gamma(A) \) with \( d \)-vertical right invariant vector fields on \( \mathcal{G} \) and the fact that \( T^*_d \mathcal{G} \) is a Lie-Poisson manifold itself. The remaining case is treated similarly. \( \square \)

We shall use the following general fact about the principal symbols of commutators.

**Proposition 5.** When \( E \) is the trivial line bundle, the commutator \([P, Q]\) satisfies \( \sigma_{m+m'-1}([P, Q]) = \{\sigma_m(P), \sigma_{m'}(Q)\} \), where \( \{\ , \ \} \) is the Poisson structure on \( A^*(\mathcal{G}) \).

**Proof.** The map \( \Phi^* : \mathcal{C}^\infty(A^*(\mathcal{G})) \to \mathcal{C}^\infty(T^*_d \mathcal{G}) \) is a Poisson map according to Lemma 7. Since \( \Phi^*(\sigma_m(P)) = \sigma_m(P_x) \) on \( T^*_d \mathcal{G}_x \) the result follows from

\[
\Phi^*(\sigma_{m+m'-1}([P, Q])) = \sigma_{m+m'-1}([P_x, Q_x]) = \{\sigma_m(P_x), \sigma_{m'}(Q_x)\}
\]

\[
= \{\Phi^*(\sigma_m(P)), \Phi^*(\sigma_{m'}(Q))\}
\]

\[
= \Phi^*(\{\sigma_m(P), \sigma_{m'}(Q)\}).
\]

Since \( \Phi^* \) is one-to-one this proves the last statement. \( \square \)
We now use the results above to construct deformation quantizations. Deform the Lie bracket structure on the Lie algebroid $A$ on $M$ to obtain a new algebroid, the adiabatic algebroid $A_t$ associated to $A$, defined over $M \times [0, \infty)$ as follows. As a bundle $A_t$ is the lift of the bundle $A$ to $M \times [0, \infty)$. Regard the sections $X$ of $A_t$ as functions $X : [0, \infty) \to \Gamma(A)$, $t \to X_t$. Then the algebroid structure is obtained by letting 

$$[X, Y]_t = t[X_t, Y_t]$$

and

$$\rho(X)_t = t\rho(X_t)$$

so that $\rho(X)f$ is the function whose restriction to $\{t\} \times M$ is $t\rho(X_t)(f)$, where for any $f \in C^\infty(M \times [0, \infty))$ we denote by $f_t \in C^\infty(M)$ the restriction of $f$ to $\{t\} \times M \equiv M$.

Observe that $C^\infty([0, \infty)) \subset C^\infty(M \times [0, \infty))$ is acted upon trivially by $\Gamma(A_t)$ and hence will define a central subalgebra of the universal enveloping algebra $U(A_t)$ of the adiabatic Lie algebroid $A_t$. Denote by $t \in C^\infty([0, \infty))$ the identity function.

**Theorem 4.** The inverse limit $\proj \lim U(A_t)/t^n U(A_t)$ is a deformation quantization of $\Gamma(\text{Symm}(A))$, the algebra of polynomial functions on $A^*$. Therefore, it induces a $\ast$-product on the Lie-Poisson space $A^*$ in the sense of [3].

**Proof.** It follows from the PBW theorem for Lie algebroids (Theorem 3) that the inverse limit $\proj \lim U(A_t)/t^n U(A_t)$ is isomorphic to $\Gamma(\text{Symm}(A))[\![t]\!]$ as a $\mathbb{C}[\![t]\!]$ module via the complete symbol map $\sigma_\nabla = \sigma_{\nabla, \phi}$ defined in Equation (16). Denote by $\{ , \}'$ the Poisson bracket on $A^*_t$ and identify $C^\infty(A^*)$ with the subset of functions on $A^*_t$ that do not depend on $t$. Then $\{f, g\}' = t\{f, g\}$ if $f, g$ are smooth functions on $A^*_t$ and $\{ , \}$ is the Poisson bracket on $A^*$.

For any polynomial function $f$ on $A^*$ denote by $q(f) \in U(A_t)$ the element with complete symbol $\sigma_\nabla(q(f))(\xi, t) = f(\xi)$ obtained, as an application of the isomorphism in the PBW theorem for $A_t$. (We treat $\tau$ as the identity, which justifies replacing $\sigma_\nabla \circ \tau$ with $\sigma_\nabla$.) The proof will be complete if we check the following quantization relation

$$q(f)q(g) - q(g)q(f) = tq(\{f, g\}) + t^2 h, \quad h \in U(A_t). \quad (25)$$

It is actually enough to do so for $f$ and $g$ among a set of generators of the algebra $\Gamma(\text{Symm}(A))$. Choose the set of generators to be the union of $C^\infty(M)$ and $\Gamma(A)$. Then Equation (25) will obviously be satisfied for $f$ and $g$ in this generating set (with no $t^2$–term) in view of the definition of the Lie bracket on $A_t$. \qed
Remark. When \( M \) is a Lie group \( G \) and \( \nabla \) is the right invariant trivial connection making all right invariant vector fields parallel, this construction, restricted to right invariant differential operators, reduces to the symmetrization correspondence between \( U(g) \) and \( S(g) \) studied by Berezin \cite{4} and Gutt \cite{12}. See also Rieffel’s paper \cite{35}. On the other hand, when the Lie algebroid \( A \) is the tangent bundle Lie algebroid \( TP \), this construction gives rise to a quantization for the canonical symplectic structure on cotangent bundle \( T^*M \).

The quantization of the Lie-Poisson structure on \( A^* \) was investigated by Landsman in terms of Jordan-Lie algebras \cite{17}. His quantization axioms are closer to those of Rieffel’s strict deformation quantization. It was conjectured in \cite{17} that the quantization of \( A^* \) is related to the groupoid \( C^* \)-algebra of the corresponding groupoid \( \mathcal{G} \), and the transitive case was proved in \cite{18}.

4. Examples.

As anticipated in the introduction, we recover many previously defined classes of operators as pseudodifferential operators on groupoids. We begin by showing that pseudodifferential operators on a manifold, in the classical sense, are obtained as a particular case of our construction. In this section we will consider only operators with coefficients in the trivial line bundle \( E = \mathbb{C} \). We include the description of the Lie algebroids associated to each example.

Denote by \( \Psi^m_{\text{prop}}(M) \) the space of properly supported pseudodifferential operators on a smooth manifold \( M \), and by \( \Psi^m_{\text{comp}}(M) \) the subspace of operators with compactly supported Schwartz kernel, regarded as a distribution on \( M \times M \).

**Example 1.** Let \( M \) be a smooth manifold and \( \mathcal{G} = M \times M \) be the pair groupoid: \( \mathcal{G}^{(1)} = M \times M \), \( \mathcal{G}^{(0)} = M \), \( d(x, y) = y \), \( r(x, y) = x \), \( (x, y)(y, z) = (x, z) \). According to the definition, a pseudodifferential operator \( P \in \Psi^m(\mathcal{G}) \) is a uniformly supported invariant family of pseudodifferential operators \( P = (P_x, x \in M) \) on \( M \times \{x\} \). The action by right translation with \( g = (x, y) \) identifies \( M \times \{x\} \) with \( M \times \{y\} \). After we identify all fibers with \( M \), the invariance condition reads \( P_x = P_y \) for all \( x, y \) in \( M \). This shows that the family \( P = (P_x)_{x \in M} \) is constant, and hence reduces to one operator \( P_0 \) on \( M \). The family \( P = (P_x)_{x \in M} \) is uniformly supported if and only if the distribution kernel of \( P_0 \) is compactly supported. The family \( P \) is properly supported if and only if \( P_0 \) is properly supported. If \( M \) is not compact, then \( P = (P_x)_{x \in M} \) will not be compactly supported unless it vanishes. We obtain \( \Psi^m(\mathcal{G}) = \Psi^m_{\text{comp}}(M) \).

In this case, the Lie algebroid \( A(\mathcal{G}) \) is the tangent bundle \( TM \).
Example 2. If \( \mathcal{G} \) has only one unit, i.e. if \( \mathcal{G} = G \), a Lie group, then \( \Psi^m(\mathcal{G}) \simeq \Psi^m_{\text{prop}}(G)^G \), the algebra of properly supported pseudodifferential operators on \( G \), invariant with respect to right translations. In this example, every invariant properly supported operator is also uniformly supported. Again, there are no nontrivial compactly supported operators unless \( G \) is compact.

In this example \( A(\mathcal{G}) \) is the Lie algebra of \( G \).

We continue with some more elaborate examples.

Example 3. If \( \mathcal{G} \) is the holonomy groupoid of a foliation \( \mathcal{F} \) on a smooth manifold \( M \), then \( \Psi^\infty(\mathcal{G}) \) is the algebra of pseudodifferential operators along the leaves of the foliation. Suppose for simplicity that the foliation is given by a (right) locally free action of a Lie group \( G \) on a manifold \( M \), and that the isotropy representation of \( G_x \), the stabilizer of \( x \), on \( N_x \), the normal space to the orbit through \( x \), is faithful. This is equivalent to the condition that the holonomy of the leaf passing through \( x \) be isomorphic to the discrete group \( G_x \).

The Lie algebroid \( A = A(\mathcal{G}) \) is the integrable subbundle of \( TM \) corresponding to the foliation \( \mathcal{F} \).

Example 4. Let \( \mathcal{G} \) be the fundamental groupoid of a compact smooth manifold \( M \) with fundamental group \( \pi_1(M) = \Gamma \). Recall that if we denote by \( \tilde{M} \) a universal covering of \( M \) and let \( \Gamma \) act by covering transformations, then \( \mathcal{G}(0) = M/\Gamma = M \), \( \mathcal{G}(1) = (M \times M)/\Gamma \) and \( d \) and \( r \) are the two projections. Each fiber \( \mathcal{G}_x \) can be identified with \( \tilde{M} \), uniquely up to the action of an element in \( \Gamma \). Let \( P = (P_x, x \in M) \) be an invariant, uniformly supported, pseudodifferential operator on \( \mathcal{G} \). Then each \( P_x, x \in M \) is a pseudodifferential operator on \( \tilde{M} \). The invariance condition applied to the elements \( g \) such that \( x = d(g) = r(g) \) implies that each operator \( P_x \) is invariant with respect to the action of \( \Gamma \). This means that we can identify \( P_x \) with an operator on \( \tilde{M} \) and that the resulting operator does not depend on the identification of \( \mathcal{G}_x \) with \( \tilde{M} \). Then the invariance condition applied to an arbitrary arrow \( g \in \mathcal{G}(1) \) gives that all operators \( P_x \) acting on \( \tilde{M} \)

---

3The holonomy groupoids of some foliations are non-Hausdorff manifolds. We believe that our constructions will extend to this case with the use of the technique in [7] (page 564), where the groupoid algebra is generated by continuous functions supported on Hausdorff open sets.
coincide. We obtain $\Psi^m(\mathcal{G}) \simeq \Psi^m_{\text{prop}}(\tilde{M})^\Gamma$, the algebra of properly supported $\Gamma$-invariant pseudodifferential operators on the universal covering $\tilde{M}$ of $M$.

An alternative definition of this algebra using crossed products is given in [29]. See also [6].

The Lie algebroid is $TM$, as in the first example.

**Example 5.** Let $\Gamma$ be a discrete group acting from the right by diffeomorphisms on a smooth compact manifold $M$. Define $\mathcal{G}$ as follows, $\mathcal{G}^{(0)} = M$, $\mathcal{G}^{(1)} = M \times M \times \Gamma$ with $d(x, y, \gamma) = y\gamma$, $r(x, y, \gamma) = x$ and $(x, y, \gamma)(y\gamma, y', \gamma') = (x, y'\gamma^{-1}, \gamma\gamma')$. Then $\Psi^\infty(\mathcal{G})$ is the algebra generated by $\Gamma$ and $\Psi_{\text{prop}}^\infty(M)$ acting on $C^\infty(M) \otimes \mathbb{C}[\Gamma]$, where $\Gamma$ acts diagonally, $\Psi_{\text{prop}}^\infty(M)$ acts on the first variable, and $\mathbb{C}[\Gamma]$ denotes the set of finite sums of elements in $\Gamma$ with complex coefficients. This algebra coincides with the crossed product algebra $\Psi_{\text{prop}}^\infty(M) \rtimes \Gamma = \{ \sum_{i=0}^n P_i g_i, P_i \in \Psi_{\text{prop}}^\infty(M), g_i \in \Gamma \}$. The regularizing algebra $\Psi_{\text{prop}}^{-\infty}(\mathcal{G})$ is isomorphic to $\Psi_{\text{prop}}^{-\infty}(M) \rtimes \Gamma \simeq \Psi_{\text{comp}}^{-\infty}(M) \otimes \mathbb{C}[\Gamma]$. If we drop the condition that $M$ be compact we obtain $\Psi^\infty(\mathcal{G}) \simeq \Psi_{\text{comp}}^\infty(M) \rtimes \Gamma$.

In general, if a discrete group $\Gamma$ acts on a groupoid $\mathcal{G}_0$, then

$$\Psi_{\text{prop}}^\infty(\mathcal{G}_0 \rtimes \Gamma) \simeq \Psi_{\text{prop}}^\infty(\mathcal{G}_0) \rtimes \Gamma.$$ 

This construction does not change the Lie algebroid.

In the following example we realize the algebra of families of operators in $\Psi^m(\mathcal{G})$ parametrized by a compact space $B$ as the algebra of pseudodifferential operators on the product groupoid $\mathcal{G} \times B$. This example shows that our class of operators on groupoids is closed under formation of families of operators.

**Example 6.** If $B$ is a compact manifold with corners, define $\mathcal{G} \times B$ by $(\mathcal{G} \times B)^{(0)} = \mathcal{G}^{(0)} \times B$, $(\mathcal{G} \times B)^{(1)} = \mathcal{G}^{(1)} \times B$ with the structural maps preserving the $B$-component. Then $\Psi^m(\mathcal{G} \times B)$ contains $\Psi^m(\mathcal{G}) \otimes C^\infty(B)$ as a dense subset in the sense that $\Psi^{m-1}(\mathcal{G}) \otimes C^\infty(B) = \Psi^m(\mathcal{G}) \otimes C^\infty(B) \cap \Psi^{m-1}(\mathcal{G} \times B)$ and $\Psi^m(\mathcal{G}) \otimes C^\infty(B) / \Psi^{m-1}(\mathcal{G}) \otimes C^\infty(B)$ is dense in $\Psi^{m}(\mathcal{G} \times B) / \Psi^{m-1}(\mathcal{G} \times B)$ in the corresponding Frechet topology (defined by the isomorphism of Theorem 8). It follows that $\Psi^m(\mathcal{G} \times B)$ consists of smooth families of operators in $\Psi^m(\mathcal{G})$ parametrized by $B$, see [2], page 122 and after, where families of pseudodifferential operators are discussed.

We obtain that $A(\mathcal{G} \times B)$ is the pull back of $A$ to $\mathcal{G}^{(0)} \times B$.

The following example generalizes the tangent groupoid of Connes; here we closely follow [6, II,5]. The groupoid defined below also appears in [37] and is related to the notion of explosion of manifolds.

**Example 7.** The adiabatic groupoid $\mathcal{G}_{\text{adb}}$ associated to $\mathcal{G}$ is defined as follows. The space of units is $\mathcal{G}^{(0)}_{\text{adb}} = [0, \infty) \times \mathcal{G}^{(0)}$ with the product manifold structure. The set of arrows $\mathcal{G}^{(1)}_{\text{adb}}$ is defined to be the disjoint union $A(\mathcal{G}) \cup$
The smooth structure on the set of arrows is the product structure for \( t > 0 \). In order to define a coordinate chart at a point \( v \in T_x \mathcal{G}_x \) choose first a coordinate system \( \psi : U = U_1 \times U_2 \to \mathcal{G}^{(1)}_1, U_1 \subset \mathbb{R}^p \) and \( U_2 \subset \mathbb{R}^n \) being open sets containing the origin, \( U_2 \) convex, with the following properties: \( \psi(0, 0) = x \in \mathcal{G}^{(0)}_0 \subset \mathcal{G}^{(1)}_1, d(\psi(s, y_1)) = d(\psi(s, y_2)) = \phi(s) \) and \( \psi(U) \cap \mathcal{G}^{(0)}_0 = \psi(U_1 \times \{0\}) \). Here \( \phi : U_1 \to \mathcal{G}^{(0)}_0 \) is a coordinate chart of \( x \) in \( \mathcal{G}^{(0)}_0 \). We identify, using the differential \( D_2 \psi \) of the map \( \psi \), the vector space \( \{s\} \times \mathbb{R}^n \) and the tangent space \( T_{\phi(s)} \mathcal{G}^{(s)}_0 = A_{\phi(s)}(\mathcal{G}) \). We obtain then coordinate charts \( \psi_t : (0, \epsilon) \times U_1 \times \epsilon^{-1} U_2 \to \mathcal{G}^{(1)}_1, \psi_t(0, s, y) = (0, (D_2 \psi)(s, y)) \in T_{\phi(s)} \mathcal{G}^{(s)}_0 = A_{\phi(s)}(\mathcal{G}) \) and \( \psi_t(t, s, y) = (t, \psi(s, ty)) \in (0, 1) \times \mathcal{G}^{(1)}_1 \). For \( \epsilon \) very small the range of \( \psi_t \) will contain \( v \).

For \( \mathcal{G} = M \times M \) as in the first example the groupoid \( \mathcal{G}_{\text{adb}} \) is the tangent groupoid defined by Connes, and the algebra of pseudodifferential operators is the algebra of asymptotic pseudodifferential operators \([39]\). In general an operator \( P \) in \( \Psi^m(\mathcal{G}_{\text{adb}}) \) will restrict to an adiabatic family \( P = (P_{t,x}, t > 0, x \in \mathcal{G}^{(0)}_0) \), which will have an “adiabatic limit” at \( t = 0 \) given by the operator \( P \) at \( t = 0 \).

The Lie algebroid of \( \mathcal{G}_{\text{adb}} \) is the adiabatic Lie algebroid associated to \( A(\mathcal{G}) \), \( A(\mathcal{G}_{\text{adb}}) = A(\mathcal{G})_t \) (using the notation of Theorem 4). This gives a procedure for integrating adiabatic Lie algebroids. Using pseudodifferential operators on the adiabatic groupoid we obtain an explicit quantization of symbols on \( A^* \) generalizing Theorem 4. The proof proceeds exactly in the same way.

**Theorem 5.** The inverse limit \( \text{proj lim} \Psi^\infty(\mathcal{G}_{\text{adb}})/t^n \Psi^\infty(\mathcal{G}_{\text{adb}}) \) is a deformation quantization of the commutative algebra \( S_{\text{cl}}^\infty(A^*(\mathcal{G})) \) of classical symbols.

The space \( S_{\text{cl}}^\infty(A^*(\mathcal{G})) \) appearing in the statement of the theorem above is the union of all symbol spaces \( S_{\text{cl}}^m(A^*(\mathcal{G})) \) and is a commutative algebra under pointwise multiplication.

Of course Theorem 4 provides us with a \( * \)-product whose multiplication is given by differential operators, and hence this \( * \)-product extends to all smooth functions (even to functions defined on open subsets). The usefulness of the theorem above is that it gives in principle a nonperturbative (i.e. not just formal) deformation quantization, close in spirit to that of strict deformation quantization introduced by Rieffel \([34]\).

**Example 8.** This example provides a treatment in our settings of the \( b \)- and \( c \)-calculi defined by Melrose \([21, 22, 25, 26]\) on a manifold with boundary \( M \).
Define first a groupoid $G_\phi(M)$ associated to $M$ and an increasing diffeomorphism $\phi : \mathbb{R} \to (0, \infty)$ as follows. If $M = [0, \infty)$ the action by translation of $\mathbb{R}$ on itself extends to an action on $M$ fixing $0$, not smooth in general, defined using the isomorphism $\phi$. Define $G_\phi(M)$ to be the transformation groupoid associated this action of $\mathbb{R}$ on $M$. If $M = [0, 1)$, then $G_\phi(M)$ is defined to be the reduction $G_\phi(M) = G_\phi([0, \infty)) \cap d^{-1}(M) \cap r^{-1}(M)$ of $G_\phi([0, \infty))$ to $[0, 1)$.

Suppose next that $M = \partial M \times [0, 1)$. We then define $G_\phi(M) = G_\phi([0, 1)) \times (\partial M \times \partial M)$, where $\partial M \times \partial M$ is the pair groupoid of $\partial M$ considered in Example 1. For an arbitrary manifold with boundary $M$ write $M = M_0 \cup U$, where $U = M \setminus \partial M$ and $M_0$ is diffeomorphic to $\partial M \times [0, 1)$. (Our construction will depend on this diffeomorphism.) Then we define $G_\phi^{(0)}(M) = M$ and $G_\phi^{(1)}(M) = G_\phi^{(1)}(M_0) \cup (U \times U)$ with the induced operations.

If $\phi(t) = e^t$ or $\phi(t) = -t^{-1}$ (for $t << 0$), then $G = G_\phi$ will be an almost differentiable groupoid and we obtain $\Psi^m(G) \subset \Psi_b(M)$ in the first case and $\Psi^m(G) \subset \Psi_c(M)$ in the second case. The first groupoid does not depend on any choices.

5. Distribution kernels.

In this section we characterize the reduced (or convolution) distribution kernels of operators in $\Psi^m(G; E)$ following [21] (see also [14]) as compactly supported distributions on $G$, conormal to the set of units $G^{(0)}$.

Denote by $\text{END}_G(E)$ the bundle $\text{Hom}(d^*(E), r^*(E)) = r^*(E) \otimes d^*(E)'$ on $G^{(1)}$, where $V'$ denotes as usual the dual of the vector bundle $V$. Using the relations $d \circ \iota = r$ and $r \circ \iota = d$ we see that $\text{END}_G(E)$ satisfies
\begin{equation}
\iota^*(\text{END}_G(E)) \simeq d^*(E) \otimes r^*(E)' \simeq \text{END}_G(E)'.
\end{equation}

We define a convolution product on the space $\mathcal{C}_c^\infty(G^{(1)}, \text{END}_G(E) \otimes d^*(\mathcal{D}))$ of compactly supported smooth sections of the bundle $\text{END}_G(E) \otimes d^*(\mathcal{D})$ by the formula
\begin{equation}
f_1 \ast f_2(g) = \int_{\{(h_1, h_2), h_1 h_2 = g\}} f_1(h_1) f_2(h_2).\]
\end{equation}
The multiplication on the right hand side is the composition of homomorphisms giving a linear map
\[
\text{Hom}(E_{d(h_1)}, E_{r(h_1)}) \otimes \text{Hom}(E_{d(h_2)}, E_{r(h_2)}) \otimes \mathcal{D}_{d(h_1)} \otimes \mathcal{D}_{d(h_2)} \longrightarrow \text{Hom}(E_{d(g)}, E_{r(g)}) \otimes \mathcal{D}_{d(h_1)} \otimes \mathcal{D}_{d(h_2)} \]
\[
\quad f_1(h_1) \otimes f_2(h_2) \longrightarrow f_1(h_1) f_2(h_2),
\]
defined since $d(h_1) = r(h_2)$. To see that the integration is defined we parametrize the set $\{(h_1, h_2), h_1 h_2 = g\}$ as $\{(g h^{-1}, h), h \in G_{d(g)}\}$, which shows that this set is a smooth manifold, and notice that we can invariantly define
the integration with respect to $h$ taking advantage of the 1-density factor $\mathcal{D}_{d(h)^1} = \mathcal{D}_{r(h)} = (\Omega_d)_h$. If we choose a hermitian metric on $\mathcal{D}^{-1/2} \otimes E$, we obtain a conjugate-linear involution (making the desired isomorphism. □

From this, taking into account the definitions of $k$, where $k$ is a given smooth section $k \in C^\infty(\mathcal{G}_x \times \mathcal{G}_x; r_1^*(E) \otimes r_2^*(E)' \otimes \Omega_2)$, using the notation $\Omega_2 = p_2^*(\Omega_d) = r_2^*(\mathcal{D})$ of (11). We define the reduced distribution kernel $k_P$ of the smoothing operator $P$ by

$$
(28) 
k_P(g) = k_{d(g)}(g, d(g)) \in E_{r(g)} \otimes E'_{d(g)} \otimes \mathcal{D}_{d(g)}.
$$

This definition will be later extended to all of $\Psi^\infty(\mathcal{G}; E)$.

The following theorem is one of the main reasons we consider uniformly supported operators.

**Theorem 6.** The reduced kernel map $P \to k_P$ (28) defines an isomorphism of the residual ideal $\Psi^{-\infty}(\mathcal{G}; E)$ with the convolution algebra $C^\infty_c(\mathcal{G}^{(1)}, END_\mathcal{G}(E) \otimes d^*(\mathcal{D}))$.

**Proof.** Let $P$ and $k$ be as above. We know from Lemma 1 that the collection of all sections $k$ defines a smooth section of $r_1^*(E) \times r_2^*(E)' \otimes \Omega_2$ over the manifold $\{(g_1, g_2), d(g_1) = d(g_2)\}$. The relation $P_{r(g)} U_g = U_g P_{d(g)}$ gives the invariance relation $k_{r(g)}(h', h) = k_{d(g)}(h'g, hg) \in E_{r(h)} \otimes E'_{r(g)} \otimes \mathcal{D}_{r(h)}$ for all arrows $g \in \mathcal{G}^{(1)}$, and $h, h' \in \mathcal{G}_{r(g)}$. It follows that $k_{d(h)}(h', h) = k_{r(h)}(h'h^{-1}, r(h)) = k_P(h'h^{-1})$. The section $k_P$ is well defined, smooth and completely determines all kernels $k_x$ and hence also the operator $P$. Moreover, the section $k_P$ has compact support because $\mathrm{supp}(k_P) = \mathrm{supp}_{\mu}(P) = \mu \circ (id \times \iota)(\cup_{x} \mathrm{supp}(k_x^P))$ and the reduced support $\mathrm{supp}_{\mu}(P)$ of $P$ is compact since $P$ is uniformly supported. The distribution kernel $k_{PQ}$ of the product $P_x Q_x$ of two operators $P_x, Q_x \in \Psi^{-\infty}(\mathcal{G}_x)$ is

$$
k_{PQ}^P(g, g'') = \int_{\mathcal{G}_x} k_{P}^P(g, g') k_{Q}^Q(g', g'') dg',
$$

where $k_{P}^P$ and $k_{Q}^Q$ are the distribution kernels of $P_x$ and, respectively, $Q_x$. From this, taking into account the definitions of $k_{PQ}$, $k_P$ and $k_Q$, we obtain

$$
k_{PQ}(g) = \int_{\mathcal{G}_x} k_P(gg'^{-1}) k_Q(g') dg'.
$$

This means that $k_{PQ} = k_P \ast k_Q$ and hence the map $P \to k_P$ establishes the desired isomorphism. □

We will now use duality to extend the definition of the reduced distribution kernel to any operator $P \in \Psi^\infty(\mathcal{G}; E)$. Let $\mathcal{L} = \Omega_{\mathcal{G}^{(0)}}$ be the line bundle
of 1-densities on $G^{(0)}$ and $\mathcal{D} = \Omega_d[G^{(0)}]$ be the bundle of vertical 1-densities as above. Define

$$\Psi^{-\infty}(G; E)_\mathcal{L} = \Psi^{-\infty}(G; E) \otimes \mathcal{C}^\infty(G^{(0)}, \mathcal{D}^{-1} \otimes \mathcal{L}) \simeq$$

$$\Psi^{-\infty}(G; E) \otimes \mathcal{C}^\infty(G^{(0)}, \mathcal{D}^{-1}) \otimes \mathcal{C}^\infty(G^{(0)}, \mathcal{L}),$$

where the tensor products are defined using the inclusion $\mathcal{C}^\infty(G^{(0)}) \subset \Psi^{-\infty}(G; E)$. We note that the bundle $\mathcal{L}$ plays an important role in connection with the modular class of a groupoid [11], since it carries a natural representation of the groupoid.

The relation $k_P f(g) = k_P(g) f(d(g))$ for $f \in \mathcal{C}^\infty(G^{(0)})$ and $P \in \Psi^{-\infty}(G; E)$ give using Theorem 6 the isomorphism

$$(29) \quad \Psi^{-\infty}(G; E)_\mathcal{L} \simeq \mathcal{C}^\infty_c(G^{(1)}, \text{END}_G(E) \otimes d^*(\mathcal{L})).$$

The space $\Psi^{-\infty}(G; E)_\mathcal{L}$ comes equipped with a natural linear functional $T$ such that, if $P_0 \in \Psi^{-\infty}(G; E)$, $\xi \in \mathcal{C}^\infty(G^{(0)}, \mathcal{D}^{-1})$ and $\nu \in \mathcal{C}^\infty(G^{(0)}, \mathcal{L})$, then

$$T(P_0 \otimes \xi \otimes \nu) = \int_{G^{(0)}} \text{tr}(k_{P_0}(x)\xi(x))d\nu(x)$$

defined by integrating the function $\text{tr}(k_{P_0}(x)\xi(x))$ with respect to the 1-density (i.e. measure) $\nu$. An operator $P \in \Psi^{\text{vis}}(G; E)$ defines a continuous linear functional (i.e. distribution) $k_P^* : \Psi^{-\infty}(G; E)_\mathcal{L} \to \mathbb{C}$ by the formula

$$k_P^*(P_0 \otimes \xi \otimes \nu) = T(P P_0 \otimes \xi \otimes \nu).$$

It is easy to see using Equation (26) that the map $f \to \tilde{f} = f \circ \iota$, $\iota(g) = g^{-1}$, establishes isomorphisms

$$(30) \quad \Phi : \mathcal{C}^\infty_c(G^{(1)}, \text{END}_G(E) \otimes d^*(\mathcal{L})) \xrightarrow{\iota^*} \mathcal{C}^\infty_c(G^{(1)}, \text{END}_G(E)^{\prime} \otimes r^*(\mathcal{L}))$$

$$\simeq \mathcal{C}^\infty_c(G^{(1)}, (\text{END}_G(E) \otimes d^*(\mathcal{L}))^{\prime} \otimes d^*(\mathcal{L}) \otimes r^*(\mathcal{L}))$$

$$\simeq \mathcal{C}^\infty_c(G^{(1)}, \text{END}_G(E) \otimes d^*(\mathcal{D})) \otimes \Omega_G)$$

whose composition we denote by $\Phi$, so that $\Phi(P_0 \otimes \xi \otimes \nu) = (k_{P_0}\xi\nu) \circ \iota = \iota^*(k_{P_0}\xi\nu)$. We obtain in this way from $k_P^*$, a distribution $k_P \in \mathcal{C}^{-\infty}(G^{(1)}; \text{END}_G(E) \otimes d^*(\mathcal{D}))$ defined by the formula

$$(31) \quad (k_P, f) = k_P^*(\Phi^{-1}(f)).$$

An other way of writing the formula above is

$$(32) \quad (k_P, \iota^*(k_{P_0}\xi\nu)) = T(P P_0 \otimes \xi \otimes \nu) = \int_{G^{(0)}} \text{tr}(k_{P P_0}(x)\xi(x))d\nu(x).$$

**Proposition 6.** If $P \in \Psi^{-\infty}(G; E)$ is a regularizing operator, then the kernels $k_P$ defined in Equations (28) and (31) coincide.

**Proof.** To make a distinction for the purpose of this proof, denote by $k_P^{\text{dist}}$ the distribution defined by (31). Let $\nu$ be a smooth section of $\mathcal{L}$, $\xi$ a smooth
Definition 9. The distribution $k_P \in C^{-\infty}(\mathcal{G}; E)$, for any operator $P \in \Psi^{-\infty}(\mathcal{G}; E)$ by Equation (31) will be called the reduced (or convolution) distribution kernel of $P$, or simply the reduced kernel of $P$, and will be denoted $k_P$.

We now relate the action of $\Psi^{-\infty}(\mathcal{G}; E)$ by multiplication on $\Psi^{-\infty}(\mathcal{G}; E)$, respectively on $\Psi^{-\infty}(\mathcal{G}; E)_L$, to that on $C_c^\infty(\mathcal{G}, r(E))$.

\begin{align}
\Psi^{-\infty}(\mathcal{G}; E) &\simeq C_c^\infty(\mathcal{G}^{(1)}; r^*(E)) \otimes_{C^\infty(\mathcal{G}^{(0)})} \Gamma(E^\prime) \otimes_{C^\infty(\mathcal{G}^{(0)})} \Gamma(D) \\
\Psi^{-\infty}(\mathcal{G}; E)_L &\simeq C_c^\infty(\mathcal{G}^{(1)}; r^*(E)) \otimes_{C^\infty(\mathcal{G}^{(0)})} \Gamma(E^\prime) \otimes_{C^\infty(\mathcal{G}^{(0)})} \Gamma(L)
\end{align}

such that the left action by multiplication of $\Psi^{-\infty}(\mathcal{G}; E)$ on $\Psi^{-\infty}(\mathcal{G}; E)$ becomes $P(f \otimes \eta \otimes \xi) = Pf \otimes \eta \otimes \xi$, where $\eta$ is a smooth section of $E^\prime$ and $\xi$ is a smooth section of $\mathcal{D}$ or $\mathcal{L}$. Moreover the kernel of $P_0 = f \otimes \eta \otimes \xi$ is $k_{P_0}(g) = f(g) \otimes \eta(d(g))\xi(d(g))$. Thus in order to define the distribution $k_P$, for arbitrary $P$, it is enough to compute $T(Pf_0 \otimes \eta \otimes \nu)$, where $\nu$ is a density.

Fix a unit $x$ and choose a coordinate chart $\phi : U_0 \to U \subset \mathcal{G}^{(0)}$, where $U_0$ is an open subset of $\mathbb{R}^k$ containing $0$, $k = \dim \mathcal{G}^{(0)}$ and $\phi(0) = x$. By decreasing $U_0$ if necessary we can assume that the tangent space $T\mathcal{G}^{(0)}$ is trivialized over $U$. Consider the diffeomorphism $\exp_{\nabla} : V_0 \to V \subset \mathcal{G}^{(1)}$ associated to a right invariant connection $\nabla$ as in (14) and (15), where $V_0 \subset A(\mathcal{G})$ is an open neighborhood of the zero section. It maps the zero section of $A(\mathcal{G})$ to $\mathcal{G}^{(0)}$. Choose a connection on $E$, which lifts to an invariant connection on $r^*(E)$. By decreasing $V$ if necessary and using the invariant connection on $r^*(E)$ we obtain canonical trivializations of $r^*(E)$ on each fiber $V \cap G_x$. Denote by $\theta_h : E_{r(h)} \otimes E^\prime_{d(h)} \to \text{End}(E_x)$ the isomorphism induced by the connection $\nabla'$ (defined using parallel transport along the geodesics of $\nabla$), where $x = d(h)$ and $h$ is in $V$. Decreasing further $V$ and $U$ we can assume that $E$ is trivialized over $d^{-1}(U) \cap V$ and that $\phi$ and $\exp_{\nabla}$ give a fiber preserving diffeomorphism $\psi : U_0 \times W \to d^{-1}(U) \cap V$, where $W \subset \mathbb{R}^n$ is an open set, identified with an open neighborhood of the zero section in $T_x\mathcal{G}^{(0)}$. 
The diffeomorphism \( \psi \) we have just constructed satisfies \( d(\psi(s,y)) = \phi(s) \).

The maps \( \psi \) and \( \theta_h \) yield isomorphisms

\[
\begin{align*}
\mathcal{C}^{-\infty}(d^{-1}(U) \cap V, \text{END}_{\mathcal{G}}(E) \otimes d^*(\mathcal{D})) \\
\cong \mathcal{C}^{-\infty}(U_0 \times W, \psi^*(\text{END}_{\mathcal{G}}(E) \otimes d^*(\mathcal{D}))) \\
\cong \mathcal{C}^{-\infty}(U_0 \times W, E_x \otimes E'_x)
\end{align*}
\]

whose composition is denoted \( \Theta_\psi \).

Next theorem describes the reduced distribution kernels \( k_P \) of operators \( P \) in \( \Psi^m(\mathcal{G}; E) \). We use the notation introduced above.

**Theorem 7.** For any operator \( P = (P_x, x \in \mathcal{G}^{(0)}) \in \Psi^m(\mathcal{G}; E) \) the reduced distribution kernel \( k_P \) satisfies:

(i) If \( \psi : U_0 \times W \to V_1 \subset V, W \subset \mathbb{R}^n \) open, is a diffeomorphism satisfying \( \psi(s,0) = d(\psi(s,y)) \), then there exists a symbol \( a_P \in S^m_{\text{cl}}(U_0 \times \mathbb{R}^n; \text{End}(E_x)) \), such that \( k_P \circ \iota = \Theta_\psi^{-1}(k) \) on \( V_1 \), where \( k \) is the distribution

\[
k(s, y) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{-iy \cdot \zeta} a_P(s, \zeta) d\zeta \in \text{End}(E_x),
\]

the integral being an oscillatory integral. Moreover, after suitable identifications, \( a_P \) is a representative of the principal symbol of \( P \).

(ii) The singular support of \( k_P \) is contained in \( \mathcal{G}^{(0)} \).

(iii) The support of \( k_P \) is compact, more precisely \( \text{supp}(k_P) = \text{supp}_\mu(P) \).

(iv) For every distribution \( k \in \mathcal{C}^{-\infty}(\mathcal{G}; E_0) \), satisfying the three conditions above, there exists \( P \in \Psi^m(\mathcal{G}; E) \) such that \( k = k_P \).

Note that \( k_P \circ \iota, \iota^*(k_P) \) and \( k'_P \) all denote the same distribution.

**Proof.** Write \( \phi(s) \) for \( \psi(s,0) = d(\psi(s,y)) \). According to Definitions 6 and 7, there exists a classical symbol \( a \in S^m_{\text{cl}}(U_0 \times T^*W; \text{End}(E_x)) \) such that \( P_{\phi(s)} = a(s, y, D_y) \) (modulo regularizing operators) on \( \mathcal{G}_{\phi(s)} \cap V_1 \cong W \).

Let \( P_0 \in \Psi^{-\infty}(\mathcal{G}; E), \xi \in \mathcal{C}^\infty(\mathcal{G}^{(0)}, \mathcal{D}^{-1}) \) and \( \nu \in \mathcal{C}^\infty(\mathcal{G}^{(0)}, \mathcal{L}) \). Assume, using the isomorphisms (33) and (34), that \( P_0 = f_0 \otimes \eta \), where \( \eta \in \Gamma(E') = \mathcal{C}^\infty(\mathcal{G}^{(0)}, E') \) and \( f_0 \in \mathcal{C}^\infty(\mathcal{G}, r^*(E) \otimes \Omega_d) \), so that \( f_0 \xi \) is a section of \( \mathcal{C}_\mathcal{E}^\infty(\mathcal{G}, r^*(E)) \). Then we have

\[
\text{tr}(k_{P_{f_0}}(x) \xi(x)) = \eta(P_x(f_0 \xi|_{\mathcal{G}_x})(x)).
\]

Suppose \( f_0 \) is supported in \( V_1 \), and denote by \( f_s \) the section of \( E_x \) that corresponds to \( f_0 |_{\mathcal{G}_{\phi(s)}} \) under the diffeomorphism \( \mathcal{G}_{\phi(s)} \cap V_1 \cong \{s\} \times W = W \).
induced by \( \psi \). We then have
\[
\langle \iota^*(k_P), k_P \xi \nu \rangle = \langle k_P, \iota^*(k_P \xi \nu) \rangle = T(P P_0 \otimes \xi \otimes \nu)
\]
\[
= \int_{G^{(0)}} \text{tr}(k_P P_0(x) \xi(x)) d\nu(x)
\]
\[
= \int_{G^{(0)} \cap U} \eta(P_x(f_0 \xi|_{G_x})(x)) d\nu(x)
\]
\[
= \int_{U_0} \eta \left( \int_{\mathbb{R}^n} \int_{W} e^{-iy\cdot\zeta} a(s, 0, \zeta) f_s(y) dy d\zeta \right) d\nu(s)
\]
\[
= \int_{U_0} \int_{\mathbb{R}^n} \int_{W} e^{-iy\cdot\zeta} \text{tr}(a(s, 0, \zeta) f_s(y) \otimes \eta) dy d\zeta d\nu(s)
\]
\[
= \int_{U_0 \times W} \text{tr} \left( f_s(y) \otimes \eta \int_{\mathbb{R}^n} e^{-iy\cdot\zeta} a(s, 0, \zeta) d\zeta \right) dy d\nu(s),
\]
where the first integral is really a pairing between the distribution \( k \) obtained from (i) for \( a_P(s, \zeta) = a(s, 0, \zeta) \), and the smooth section \( f_s \otimes \eta \otimes \nu \). Since \( \text{End}(E_x) \) is canonically its own dual this shows that the distribution \( k_P \) is the conormal distribution to \( G^{(0)} \) given by (i).

To prove (iii) and (iv) observe that \( k_x \) is the restriction to \( G_x \times G_x \) of the distribution \( \mu_1^1(k_P) \), where \( \mu_1(h', h) = h'h^{-1} \) and the distribution \( \mu_1^1(k_P) \) is defined by \( \langle \mu_1^1(k_P), f \rangle = \langle k_P(g), \int_{h_1 h_2 = g} f(h_1, h_2) \rangle \). Then we can define \( P_x \) by its distribution kernel \( k_x \). From (i) it follows that \( k_x \) is conormal to the diagonal and hence \( P_x \) is a pseudodifferential operator.

In order to check (ii) fix \( g \notin G^{(0)} \) and let \( \varphi \) be a smooth cut-off function, \( \varphi = 1 \) in a neighborhood of \( G^{(0)} \), \( \varphi = 0 \) in a neighborhood of \( g \). Consider again the distribution \( \mu_1^1(1 - \varphi) k_P) = (1 - \varphi \circ \mu_1) \mu_1^1(k_P) \). Its restriction to \( G_x \) is \( (1 - \varphi \circ \mu_1) k_x \), which is smooth since the singular support of \( k_x \) (= the distribution kernel of \( P_x \)) is contained in the diagonal of \( G_x \times G_x \), and \( 1 - \varphi \circ \mu_1 \) vanishes there. It follows that \( \mu_1^1(1 - \varphi) k_P) \) is smooth and hence \( (1 - \varphi) k_P \) is also smooth. \( \square \)

**Corollary 1.** The distribution \( k_P \) is conormal at \( G^{(0)} \) and smooth everywhere else. In particular, the wave-front set of \( k_P \) is a subset of the annihilator of \( TG^{(0)} \): \( WF(k_P) \subset (TG/TG^{(0)})^* \subset T^*G^{(0)} \).

**Proof.** This is a standard consequence of (i) and (ii) above, see [14], Section 12.2. \( \square \)

We remark that \( (TG/TG^{(0)})^* \) is naturally identified with \( A^*(G) \). Denote by \( S_0^m(A^*(G) \cap \text{End}(E)) \subset S_0^m(A^*(G) \cap \text{End}(E)) \) the space of classical symbols with support in a set of the form \( \pi^{-1}(K) \), where \( \pi : A^*(G) \rightarrow G^{(0)} \) is the projection and \( K \subset G^{(0)} \) is a compact subset.
Corollary 2. Let $V$ be a neighborhood of $\mathcal{G}^{(0)}$ in $\mathcal{G}$. Then any $P \in \Psi^m(\mathcal{G}; E)$ can be written as $P = P_1 + P_2$, where $P_1$ has reduced support $\text{supp}_\mu(P_1)$ contained in $V$ and $P_2 \in \Psi^{-\infty}(\mathcal{G}; E)$.

Proof. Let $\phi$ be a smooth cut-off function, equal to 1 in a neighborhood of $\mathcal{G}^{(0)}$ and with support in $V$. Define $P_2 \in \Psi^{-\infty}(\mathcal{G}; E)$ by $k_{P_2} = k_P(1 - \phi)$. This is possible using Theorem 6 since by (ii) of the theorem above $k_P(1 - \phi)$ is a smooth compactly supported section of an appropriate bundle. Then $P_1 = P - P_2$ and $P_2$ satisfy the requirements of the statement. \hfill \Box

Theorem 8. The principal symbol map $\sigma_m$ in Equation (17) is onto; hence it establishes an isomorphism

$$\Psi^m(\mathcal{G}; E)/\Psi^{m-1}(\mathcal{G}; E) \simeq S^m_c(A^*(\mathcal{G}); \text{End}(E))/S^{m-1}_c(A^*(\mathcal{G}); \text{End}(E))$$

for any $m$.

Proof. We only need to prove that $\sigma_m$ is onto. If follows from the proof of Theorem 7 that $\sigma_m(P)$ is the class of the symbol $a_P$ appearing in the equation in (i). Given a symbol $a \in S^m_c(A^*(\mathcal{G}); \text{End}(E))$ the equation in (i) defines a distribution $k_0$ in a small neighborhood of $\mathcal{G}^{(0)}$ in $\mathcal{G}$. Using a smooth cut-off function we obtain a distribution $k$ on $\mathcal{G}$ that coincides with $k_0$ in a neighborhood of $\mathcal{G}^{(0)}$ and is smooth outside $\mathcal{G}^{(0)}$. From (iv) we conclude that there exists an operator $P$ with $k_P = k$, which will then necessarily satisfy $\sigma_m(P) = a + S^{m-1}_c(A^*(\mathcal{G}); \text{End}(E))$. \hfill \Box

6. The action on sections of $E$.

In this section we define a natural action of $\Psi^m(\mathcal{G}; E)$ on sections of $E$ over $\mathcal{G}^{(0)}$, thus generalizing the action of classical pseudodifferential operators on functions.

Let $\phi$ be a smooth section of $E$ over $\mathcal{G}^{(0)}$. Define

$$\tilde{\phi} \in C^\infty(\mathcal{G}^{(1)}, r^*(E)), \tilde{\phi}(g) = \phi(r(g)).$$

Lemma 8. If $P = (P_x, x \in \mathcal{G}^{(0)})$ belongs to $\Psi^\infty(\mathcal{G}; E)$, then for any section $\phi$ in $C^\infty(\mathcal{G}^{(0)}, E)$ there exists a unique section $\psi \in C^\infty(\mathcal{G}^{(0)}, E)$ such that $P\tilde{\phi} = \tilde{\psi}$.

Proof. Observe first that given a section $\gamma$ of $r^*(E)$ over $\mathcal{G}^{(1)}$ we can find a section $\tilde{\phi}$ of $E$ over $\mathcal{G}^{(0)}$ such that $f = \tilde{\phi}$ if and only if $f(g'g) = f(g')$ for all $g$ and $g'$, i.e. if and only if

$$U_g f_x = f_y, \text{ for all } g, x, y \text{ such that } x = d(g) \text{ and } y = r(g).$$

We have then

$$U_g \tilde{\phi}_x = \tilde{\phi}_y \Rightarrow P_y U_g \tilde{\phi}_x = P_y \tilde{\phi}_y \Rightarrow U_g P_x \tilde{\phi}_x = P_y \tilde{\phi}_y \Rightarrow U_g (P\tilde{\phi})_x = (P\tilde{\phi})_y$$
and hence $P\tilde{\phi}$ satisfies (37). Thus we can find a section $\psi$ of $E$ over $G^{(0)}$ such that $P\tilde{\phi} = \tilde{\psi}$. Note that $P_x\tilde{\phi}_x$ is defined since $P_x$ is properly supported.

The uniqueness of the section $\psi$ follows from the fact that the map $\phi \to \tilde{\phi}$ is one-to-one, and the smoothness of $\psi$ follows from Lemma 2.

The representation given by the following theorem reduces to the trivial representation in the case of a group (see also comments below).

**Theorem 9.** There exists a canonical representation $\pi_0$ of the algebra $\Psi^\infty(G; E)$ on $C^\infty(G^{(0)}, E)$ given by $\pi_0(P)\phi = \psi$, where, using the notation of the previous lemma, $\psi$ is the unique section satisfying $\tilde{\psi} = P\tilde{\phi}$. Moreover $\pi_0(P)$ maps compactly supported sections to compactly supported sections.

**Proof.** The fact that $\pi_0$ is well defined follows from the uniqueness part of the previous lemma. It is clearly a representation. We only need to check that $\pi_0(P)$ maps compactly supported sections to compactly supported sections. Let $L_1 \subset G^{(0)}$ be the support of $\phi$, $L_2 = \text{supp}(P)$. Then the support of $\pi_0(P)$ is contained in $L_2L_1$. □

Assume that $E$ is a trivial line bundle. $C^\infty(M)$ and $\Gamma(A)$ act naturally on $C^\infty(M)$ and this action satisfies the relations (22), which means that it gives rise to a representation of $U(A) = \text{Diff}(G)$ on $C^\infty(M)$. Then $\pi_0$ is an extension of this representation. If $G = G$ is a group, then $\pi_0$ extends the trivial representation. In order to generalize this fact to arbitrary representations of $G$ we need the following definition.

**Definition 10.** An equivariant bundle $(V, \rho)$ on $G^{(0)}$ is a differentiable vector bundle $E$ together with a bundle isomorphism $\rho : d^*(V) \to r^*(V)$ satisfying $\rho(gh) = \rho(g)\rho(h)$.

An equivariant bundle is also called a representation of $G$. Given an equivariant bundle $(V, \rho)$, we can define a representation $\pi_\rho$ of the groupoid algebra $C^\infty_c(G, d^*(D))$ on $C^\infty_c(G^{(0)}, V)$ by the formula

$$
(\pi_\rho(f)\phi)(x) = \int_{G_x} f(h^{-1})\rho(h^{-1})\phi(r(h)).
$$

Note that the integration is defined and gives an element of $V_x$ since $f(h^{-1})\phi(r(h))$ is in $C^\infty_c(r^*(V) \otimes \Omega_d)$ and hence that $f(h^{-1})\rho(h^{-1})\phi(r(h))$ is a smooth compactly supported section of $d^*(V) \otimes \Omega_d$.

The following proposition has no obvious analog in the classical theory because the pair groupoid has no nontrivial representations. If one moves one step up and considers the fundamental groupoid, nontrivial representations exist, and the following lemma says that geometric operators (i.e. the ones that lift to the universal covering space) act on sections of flat bundles. A representation of a groupoid thus resembles a flat bundle.
Proposition 7. Let $(V, \rho)$ be an equivariant bundle and $E$ an arbitrary bundle on $G^{(0)}$. There exists a natural morphism $T_\rho : \Psi^\infty(G; E) \to \Psi^\infty(G; V \otimes E)$ and hence there exist canonical actions $\pi_\rho = \pi_0 \circ T_\rho$ of $\Psi^\infty(G)$ on $C^\infty(G^{(0)}, E \otimes V)$ and $C^\infty_c(G^{(0)}, E \otimes V)$, which extend the representation defined in (38).

Proof. Let

$$W_{\rho,x} : C^\infty_c(G_x; r^*(E)) \otimes V_x = C^\infty_c(G_x; r^*(E) \otimes d^*(V)) \to C^\infty_c(G_x; r^*(E \otimes V))$$

be the isomorphism defined by $\rho$ as in Definition 10. It is easy to see that this gives an isomorphism $W_\rho : C^\infty_c(G; r^*(E) \otimes d^*(V)) \to C^\infty_c(G; r^*(E \otimes V))$. Define an operator on $C^\infty_c(G_x; r^*(E \otimes V))$ by the formula

$$(T_\rho(P))_x = W_{\rho,x}(P_x \otimes \text{id}_{V_x})W_{\rho,x}^{-1}.$$ 

The relation $W_{\rho,x}(U_g \otimes \rho(g)) = U_g W_{\rho,x}$ shows that the family $(T_\rho(P))_x$, $x \in G^{(0)}$ satisfies the invariance condition $(T_\rho(P))_x U_g = U_g (T_\rho(P))_y$, for $d(g) = x$ and $r(g) = y$. The uniform support condition is satisfied since $\text{supp}(T_\rho(P)) = \text{supp}_0(P)$. It follows that the family $(T_\rho(P))_x$ defines an operator $T_\rho(P)$ in $\Psi^\infty(G; V \otimes E)$. The multiplicativity condition $T_\rho(PQ) = T_\rho(P)T_\rho(Q)$ follows from definition and hence $T_\rho$ is a morphism. \hfill \Box

References


Received April 25, 1997. Nistor is partially supported by NSF Young Investigator Award DMS-9457859 and a Sloan research fellowship, preprints available from http://www.math.psu.edu/nistor. Weinstein is partially supported by NSF Grants DMS-9309653 and DMS-9625122. Xu is partially supported by NSF Grant DMS 95-04913.
RESTRICTIONS OF RANK-2 SEMISTABLE VECTOR BUNDLES ON SURFACES IN POSITIVE CHARACTERISTIC

Atsushi Noma

We work over an algebraically closed field of positive characteristic. Let $E$ be a semistable rank-2 vector bundle with respect to a very ample line bundle $O(1)$ on a smooth projective surface. The purpose here is to give an effective bound $d_0$ such that if $d \geq d_0$ the restriction of $E$ to a general member $C \in |O(d)|$ is semistable.

1. Introduction.

Let $E$ be a rank-$r$ torsion free sheaf on a normal projective variety of dimension $n \geq 2$ defined over an algebraically closed field $k$. Assume that $E$ is semistable with respect to a very ample line bundle $O(1)$: Namely, if we set $\mu(F) = (c_1(F) \cdot O(1)^{n-1})$ for a subsheaf $F$ of $E$, $\mu(E) \geq \mu(F)$ holds for all subsheaf $F$ of $E$.

A problem of finding a condition when the restriction $E|C$ to a member $C \in |O(d)|$ is semistable on $C$ has been considered by several authors ([1], [3], [6], [7], [8]): Maruyama [6] proved that if $r < n$ then $E|C$ is semistable for general $C \in |O(d)|$ for every $d \geq 1$; Mehta and Ramanathan [7] proved that there exists an integer $d_0$ such that if $d \geq d_0$ then $E|C$ is semistable for general $C \in |O(d)|$; Flenner [3] proved that if $k$ is of characteristic 0 and $d$ satisfies $(\frac{d+n}{d})-d^{-1} > (O(1)^n) \cdot \max(\frac{t^2-1}{4}, 1)$ then $E|C$ is semistable for general $C \in |O(d)|$. In other direction, in characteristic 0, Bogomolov [1] and Moriwaki [8] obtained an effective bound $d_0$ for some special restriction $E|C$ to be semistable.

The purpose here is to give an effective bound $d_0$ in positive characteristic when $E$ is a rank-2 vector bundle on a surface: If $d \geq d_0$ the restriction $E|C$ of $E$ to a general member $C \in |O(d)|$ is semistable.

Our result is the following.

Theorem. Let $S$ be a smooth projective surface over an algebraically closed field $k$ of characteristic $\text{char}(k) = p > 0$ and $O_S(1)$ a very ample line bundle on $S$. Let $E$ be a rank-2 semistable vector bundle with respect to $O_S(1)$ on $S$. Set $\deg S = (O_S(1)^2)$, $\Delta(E) = c_2(E) - (1/4)c_1^2(E)$, and $\nu = \min\{(M \cdot \deg S)^{1/2} : $
$\mathcal{O}_S(1)) > 0 : \mathcal{M} \in \text{Pic } S \}$. Let $d$ be an integer with
\[
d > \begin{cases} \frac{\Delta(E)}{\nu} + \frac{\sqrt{\Delta(E)}}{2\sqrt{3\deg S}}, & \text{if } \Delta(E) > 0, \\ 0, & \text{if } \Delta(E) \leq 0. \end{cases}
\]
Then the restriction $E|C$ to a general $C \in |\mathcal{O}_S(d)|$ is semistable on $C$.
The theorem is proved based on ideas of Ein [2] and Flenner [3].

2. Proof of Theorem.

Set $L = \mathcal{O}_S(d)$. Let $\mathbb{P}^{2*}$ be the projective space of lines in $\mathbb{P}^2$ and $F$ the incidence correspondence $\{(x, \ell) \in \mathbb{P}^2 \times \mathbb{P}^{2*} : x \in \ell\}$, namely
\[F = \mathbb{P}_{\mathbb{P}^2}(\Omega_{\mathbb{P}^2}^1(2)) \subseteq \mathbb{P}_{\mathbb{P}^2}(\wedge^2 H^0(\mathcal{O}_{\mathbb{P}^2}(1)) \otimes \mathcal{O}_{\mathbb{P}^2}) = \mathbb{P}^2 \times \mathbb{P}^{2*}.
\]
Let $\phi: S \rightarrow \mathbb{P}^2$ be a (separable) finite morphism defined by a 2-dimensional, base-point-free, linear subsystem $d$ of $L$ containing a general curve $C \in |L|$. Pulling-back the correspondence $F$ by $\phi$, we have the following diagram:
\[
\begin{array}{ccc}
X & \xrightarrow{\phi} & \mathbb{P}^{2*} \\
\pi \downarrow & & \downarrow \pi_0 \\
S & \rightarrow & \mathbb{P}^2
\end{array}
\]
We denote the composite $X \rightarrow \mathbb{F} \rightarrow \mathbb{P}^{2*}$ by $\rho$.

Assume that the restriction $E|C$ to a general curve $C \in d \subset |L|$ is not semistable. In other words, the restriction $\pi^*E|\rho^{-1}(\ell)$ to $\rho^{-1}(\ell)$ for a general $\ell \in \mathbb{P}^{2*}$ is not semistable, since $\rho^{-1}(\ell)$ is isomorphic to a divisor $C \in d$ and $\pi^*E|\rho^{-1}(\ell) \cong E|C$ under this isomorphism. Consider a relative Harder-Narasimhan filtration (HN-filtration) of $\pi^*E$ over $\rho$, which has a property that its restriction to $\rho^{-1}(\ell)$ for a general $\ell \in \mathbb{P}^{2*}$ is the HN-filtration of $\pi^*E|\rho^{-1}(\ell)$ (see [4, (3.2)]). By assumption, the relative HN-filtration is $0 = \mathcal{E}_0 \subset \mathcal{E}_1 \subset \mathcal{E}_2 = \pi^*E$ and we may assume that $\mathcal{E}_1$ is locally free of rank 1 on $X$. Hence if $W$ denotes the class of the tautological bundle of $X = \mathbb{P}_{S}(\phi^*(\Omega_{\mathbb{P}^2}^1(2)))$, we have $\mathcal{E}_1 \cong \mathcal{O}_X(aW) \otimes \pi^*\mathcal{M}$ for some $a \in \mathbb{Z}$ and $\mathcal{M} \in \text{Pic } S$, since $\text{Pic } X \cong \mathbb{Z}W \oplus \pi^*\text{Pic } S$ (see [5, Ch. III Ex. 12.5]). Since $\mathcal{M}|\pi(\rho^{-1}(\ell)) \subset E|\pi(\rho^{-1}(\ell))$ is the HN-filtration of $E|\pi(\rho^{-1}(\ell))$ for a general $\ell \in \mathbb{P}^{2*}$, we have
\[
(1) \quad (c_1(E) - 2\mathcal{M} \cdot L) < 0.
\]
Consequently $H^0(\rho^{-1}(\ell), \pi^*(E \otimes \mathcal{M}^\vee)|\rho^{-1}(\ell)) \cong k$ for a general $\ell \in \mathbb{P}^{2*}$, and hence $\rho_\ast\pi^*(E \otimes \mathcal{M}^\vee)$ is of rank 1 and reflexive. Therefore we have $\rho_\ast\pi^*(E \otimes \mathcal{M}^\vee) = \mathcal{O}_{\mathbb{P}^{2*}}(-t)$ for some $t \in \mathbb{Z}$. The semistability of $E$ implies
\[
(2) \quad t > 0
\]
since $H^0(\mathbb{P}^2, \rho_*\pi^*(E \otimes \mathcal{M}^\vee)) = H^0(X, \pi^*(E \otimes \mathcal{M}^\vee)) = H^0(S, E \otimes \mathcal{M}^\vee)$ and (1) holds. The natural map $\mathcal{O}_X(-tW) = \rho^*\rho_*\pi^*(E \otimes \mathcal{M}^\vee) \rightarrow \pi^*(E \otimes \mathcal{M}^\vee)$ induces an exact sequence

(3) \hspace{1cm} 0 \rightarrow \mathcal{O}_X(-tW) \otimes \pi^*\mathcal{M} \rightarrow \pi^*E \rightarrow \mathcal{O}_X(tW) \otimes \pi^*\mathcal{O}_S(c_1(E) - \mathcal{M}) \otimes \mathcal{I}_Z \rightarrow 0

with a closed subscheme $Z$ of codimension 2 in $X$.

The surjection in (3) induces a unique morphism $\sigma: X \setminus Z \rightarrow \mathbb{P}_S(E)$ with

$\sigma^*\mathcal{O}_{\mathbb{P}(E)}(1) = \mathcal{O}_X(tW) \otimes \pi^*\mathcal{O}_S(c_1(E) - \mathcal{M})|(X \setminus Z)$

$\sigma^*\mathcal{O}_{\mathbb{P}(E)/S}^1 = \mathcal{O}_X(-2tW) \otimes \pi^*\mathcal{O}_S(2\mathcal{M} - c_1(E))|(X \setminus Z)$,

by the universal property of projective bundle $\tau: \mathbb{P}_S(E) \rightarrow S$. If the differential

$$d\sigma: \sigma^*\mathcal{O}_{\mathbb{P}(E)/S}^1 \rightarrow \Omega_{X/S}^1|(X \setminus Z)$$

is zero, then $S$-morphism $\sigma$ factors through the relative Frobenius $F_{(X \setminus Z)/S}: X \setminus Z \rightarrow (X \setminus Z)^{(1)}$ of $X \setminus Z$ over $S$ (see [2, (1.4)]). Namely there exists an $S$-morphism $\sigma_1: (X \setminus Z)^{(1)} \rightarrow \mathbb{P}_S(E)$ such that $\sigma = \sigma_1 \circ F_{(X \setminus Z)/S}$. Here for an $S$-scheme $Y$, by $Y^{(r)}$ we denote the base change of the structure morphism $\eta: Y \rightarrow S$ by the $r$th (absolute) Frobenius $F^r_X: S \rightarrow S$; $F^r_{Y/S}: Y \rightarrow Y^{(r)}$ is the $S$-morphism induced by the (absolute) Frobenius $F^r_Y: Y \rightarrow Y$ of $Y$ and the structure morphism $\eta$ by the property of products. Furthermore, if $d\sigma_1 = 0$, then there exists a morphism $\sigma_2: (X \setminus Z)^{(2)} \rightarrow \mathbb{P}_S(E)$ such that $\sigma = \sigma_2 \circ F^2_{(X \setminus Z)/S}$. Proceeding in this way with $\mathbb{P}_S(E)$, we claim that there exists a morphism $\sigma_r: (X \setminus Z)^{(r)} \rightarrow \mathbb{P}_S(E)$ such that $\sigma = \sigma_r \circ F^r_{(X \setminus Z)/S}$ and the relative differential

$$d\sigma_r: \sigma^*\mathcal{O}_{\mathbb{P}(E)/S}^1 \rightarrow \Omega_{X^{(r)}/S}^1|(X \setminus Z)^{(r)}$$

is nonzero for some $r \geq 0$. In fact, suppose that we have a morphism $\sigma_r: (X \setminus Z)^{(r)} \rightarrow \mathbb{P}_S(E)$ such that $\sigma = \sigma_r \circ F^r_{(X \setminus Z)/S}$ for some $r \geq 0$. Here we set $\sigma_0 = \sigma$ if $r = 0$. Then we have the following diagram:

\[
\begin{array}{ccc}
X \setminus Z & \overset{F^r_{(X/Z)/S}}{\longrightarrow} & (X \setminus Z)^{(r)} \\
\downarrow & & \downarrow \sigma_r \\
X & \overset{F^r_X}{\longrightarrow} & X^{(r)} \\
\downarrow \pi & & \downarrow \pi_r \\
S & = & S.
\end{array}
\]
Since $X \cong \mathbb{P}_S(\mathcal{F}_S^*(\Omega_{\mathbb{P}^2}^1(2)))$, we have $X^{(r)} \cong \mathbb{P}_S(F_{X/S}^r \mathcal{O}_S^*(\Omega_{\mathbb{P}^2}^1(2)))$. If $W'$ is the class of the tautological line bundle of $X^{(r)}$ over $S$, then $F_{X/S}^r \mathcal{O}_S^*(W') \cong \mathcal{O}_X(p'^r W)$ and $\Omega_{X^{(r)}/S}^1 \cong \pi^*(p'^r L) \otimes \mathcal{O}_X(-2W')$. On the other hand, $F_{X/S}^r \pi^*_r \mathcal{A} \cong \pi^*_r \mathcal{A}$ for every $\mathcal{A} \in \text{Pic } S$. Since $\sigma = \sigma_r \circ F_{X/S}^r$ and Pic $X^{(r)} \cong \mathbb{Z}W' \oplus \pi^*_r \text{Pic } S$, the morphism $\sigma_r$ induces an exact sequence

$$0 \to \mathcal{O}_X^{(r)} \left( -\frac{t}{p'^r} W' \right) \otimes \pi^*_r \mathcal{M} \to \pi^*_r \mathcal{E}$$

and we have $t/p'^r \in \mathbb{Z}$, where $Z'$ is a codimension 2 closed subscheme of $X^{(r)}$. Since $t > 0$, the latter implies that $\sigma$ factors through the relative Frobenius over $S$ only in finite times. Therefore for some $r \geq 0$, the morphism $\sigma_r$ must have the nonzero relative differential $d\sigma_r$, as required.

We take such $r \geq 0$ and $\sigma_r \colon (X \setminus Z)^{(r)} \to \mathbb{P}_S(\mathcal{E})$. If $C \in \mathfrak{d} \subset |L|$, since $C \cong \rho^{-1}(\ell) \cong F_{X/S}^r \rho^{-1}(\ell)$ for some $\ell \in \mathbb{P}^{2r}$, we can consider $C \subset X^{(r)}$. Then we have $\mathcal{O}_X^{(r)}(W')|C \cong \mathcal{O}_C$ and $\pi^*_r \mathcal{A}|C \cong \mathcal{A}|C$ for every $\mathcal{A} \in \text{Pic } S$. The restriction $d\sigma_r|C$ to general $C \in \mathfrak{d}$ is nonzero by the choice of $r$. This implies that

$$(L \cdot 2\mathcal{M} - c_1(\mathcal{E})) \leq p'^r(L^2) \leq t(L^2),$$

since

$$\sigma_r^* \Omega_{\mathcal{F}_S(\mathcal{E})/S}|C = \mathcal{O}_C(2\mathcal{M} - c_1(\mathcal{E}))$$

$$\Omega_{X^{(r)}/S}^1|C = \mathcal{O}_C(p'^r L),$$

and since $t/p'^r \in \mathbb{Z}$.

Restricting the exact sequence (3) to a general member $W \in |\mathcal{O}_X(W)|$ not containing any associate points of $Z$, we have an exact sequence

$$0 \to \mathcal{O}_W(-tw) \otimes \pi^* \mathcal{M}|W \to \pi^* \mathcal{E}|W$$

$$\to \mathcal{O}_W(tw) \otimes \pi^* \mathcal{O}_S(c_1(\mathcal{E}) - \mathcal{M})|W \otimes \mathcal{I}_{Z\cap W} \to 0.$$
$W \to S$ is birational via $\pi$, we have
\[
c_2(E) = c_2(\pi^*E|W)
= -t^2(O_W(W)^2) + t(O_W(W) \cdot \pi^*O_S(2\mathcal{M} - c_1(E))|W)
+ (\pi^*\mathcal{M}|W \cdot \pi^*O_S(c_1(E) - \mathcal{M})|W) + \deg(Z \cap W)
= t(L \cdot 2\mathcal{M} - c_1(E)) - (\mathcal{M} \cdot \mathcal{M} - c_1(E)) + \deg(Z \cap W)
\geq t(L \cdot 2\mathcal{M} - c_1(E)) - (\mathcal{M} \cdot \mathcal{M} - c_1(E)),
\]
and hence
\[
\Delta(E) \geq 3t(L \cdot \mathcal{M} - (1/2)c_1(E)) - ((\mathcal{M} - (1/2)c_1(E))^2).
\]
By the Hodge index theorem for $L$ and $\mathcal{M} - (1/2)c_1(E)$, we have
\[
(5) \quad \Delta(E) \geq 2t(L \cdot \mathcal{M} - (1/2)c_1(E)) - \frac{(L \cdot \mathcal{M} - (1/2)c_1(E))^2}{L^2}.
\]

From (4) and (5), it follows that
\[
\Delta(E) \geq 3 \frac{(L \cdot \mathcal{M} - (1/2)c_1(E))^2}{L^2}.
\]
When $\Delta(E) \leq 0$, this contradicts (1). When $\Delta(E) > 0$, we have
\[
(6) \quad (L \cdot \mathcal{M} - (1/2)c_1(E)) \leq \frac{\sqrt{\Delta(E)L^2}}{3}.
\]
On the other hand, from (5), it follows
\[
\frac{\Delta(E)}{(L \cdot \mathcal{M} - (1/2)c_1(E))} + \frac{(L \cdot \mathcal{M} - (1/2)c_1(E))}{L^2} \geq 2t.
\]

Since $L = O_S(d)$, by using the assumption $(O_S(1) \cdot \mathcal{M} - (1/2)c_1(E)) \geq \nu/2$ to the first term and (6) to the second term, we have
\[
\frac{1}{d} \left( \frac{\Delta(E)}{\nu} + \frac{\sqrt{\Delta(E)}}{2\sqrt{3} \deg S} \right) \geq t.
\]
By assumption of $d$, we have $t < 1$ hence $t \leq 0$, which contradicts (2). This completes the proof. \hfill \Box

**Remark.** Let $S$, $O_S(1)$, $E$ and $d$ be as in Theorem.

(1) Let $d$ be a 2-dimensional linear subsystem of $|O_S(d)|$ defining a separable, finite morphism from $S$ to $\mathbb{P}^2$. The proof of theorem implies that the restriction $E|C$ is semistable for a general member $C \in d$.

(2) Assume that $\Delta(E) > 0$ and that the restriction $E|C$ to be a general member of $C \in |O_S(1)|$ is not semistable with HN-filtration $0 = \mathcal{E}_0 \subset \mathcal{E}_1 \subset \mathcal{E}_2 = E|C$. Then it follows from (6) that
\[
\deg \mathcal{E}_1 - (1/2) \deg(E|C) \leq \sqrt{\deg S \cdot \Delta(E)/3}
\]
holds. This inequality is exactly that of Ein in [2, (4.1)] when $S = \mathbb{P}^2$ and $\mathcal{O}_S(1) = \mathcal{O}_{\mathbb{P}^2}(1)$.

(3) I do not know the bound in Theorem is optimal or not. For example, let $E$ be the $m$th Frobenius pull-back $F^m_\ast(\Omega_{\mathbb{P}^2}(2))$ of the twisted cotangent bundle on $\mathbb{P}^2$, which plays an important role in the proof of Theorem. We know that $E$ is semistable (see for example [2]) and $\Delta(E) = p^{2m}/4$. Thus Theorem implies that $E|C$ is semistable on a general curve $C$ of degree $d$ if $d > p^{2m}/4 + p^{m}/(4\sqrt{3})$. On the other hand, from a calculation of $H^0(C, E(-p^m + 1)/2)|C$ by using the Euler sequence, it follows that $E|C$ is semistable for general $C$ of degree $d$ if $d > (3p^m + 5)/4$ for $p \neq 2$.

References


Received July 21, 1997 and revised March 10, 1998.

Yokohama National University
79-2 Tokiwadai, Hodogaya-ku
Yokohama 240-8501
JAPAN
E-mail address: noma@ed.ynu.ac.jp
TRACE FORMULA FOR A SYSTEM OF PARTICLES WITH ELLIPTIC POTENTIAL

K.L. VANINSKY

We consider classical particles on the line with the Weierstrass ζ function as potential. This system parameterizes special solutions of the KP equation. We derive the trace formula which relates the Hamiltonian of the particle system to the residues of some Abelian differential (meromorphic one-form) on the spectral curve. Such formula is important for the construction action-angle variables and study invariant Gibbs' states.

1. Introduction.

The subject of this note is a system of $N$ classical particles on the line interacting with the Hamiltonian

$$H_N = \sum_{n=1}^{N} \frac{p_n^2}{2} - 2\sigma^2 \sum_{n,m=1}^{N} \varphi(q_n - q_m).$$

The parameter $\sigma = 1$ corresponds to attractive particles and $\sigma = i = \sqrt{-1}$ corresponds to repulsive particles. The potential is the Weierstrass $\varphi$ function with real period $2\omega$ and pure imaginary period $2\omega'$. The system includes the well known integrable in the sense of Liouville potentials $\sin^{-2} x$, $\sinh^{-2} x$ and $x^{-2}$; which correspond to various degeneration of $\varphi$.

In a remarkable article Airault, McKean, Moser, [AMM], discovered a connection between particles with rational or elliptic potential and the Korteweg-de Vries equation

$$u_t + \frac{3}{2} uu_x - \frac{1}{4} u_{xxx} = 0.$$  

Few years later Krichever, [K], found an isomorphism between particles with elliptic potentials and special solutions of the Kadomtzev-Petviashvili (KP) equation

$$\frac{3}{4} \sigma^2 u_{yy} = \left(u_t + \frac{3}{2} uu_x - \frac{1}{4} u_{xxx}\right)_x.$$  

1Should not be confused with the Weierstrass $\sigma$, see Section 6.
The case $\sigma = 1$ corresponds to the KP-2 equation and $\sigma = i$ to KP-1. The KP equation arises as a compatibility condition for the zero curvature representation

$$[\sigma \partial_y - L_2, \partial_t - L_3] = 0,$$

where

$$L_2 = \partial_x^2 - u,$$
$$L_3 = \partial_x^3 - \frac{3}{2} u \partial_x - w.$$

Any such solution is associated with a spectral curve $\Gamma_N$ of genus $N$, defined by

$$R_N(k, z) = \sum_{n=0}^{N} r_n(z) k^n = \det(\tilde{L} + 2k),$$

where $\tilde{L}$ is a $N \times N$ matrix which depends on $q, p$ and $z$. The functions $r_n(z)$ are elliptic, so the curve $\Gamma_N$ is an $N$-sheeted covering of the elliptic curve. The matrix $\tilde{L}$ has a simple pole above $z = 0$ and can be expanded in powers of $z$

$$\tilde{L} = \frac{1}{z} L^{(-1)} + O(1),$$

where $L_{nm} = -2(1 - \delta_{nm})$ is a constant matrix. This “zero order” approximation provides all the information needed to solve the direct spectral problem and obtain a formula for the solution in terms of Riemann theta functions, see [K].

In this note we address a different question. Is there a formula of the type

$$H_N = \sum_{\alpha=2}^{N} I'_\alpha,$$

where $I'_\alpha$ are parameters of the Riemann surface associated with the system? We give an affirmative answer to this question here. The formula is needed, see [MCV1, MCV2], to express the canonical measure as

$$e^{-H} d\text{vol} = e^{-H} \prod dI \, d\phi = e^{-\sum I'} \prod dI \, d\phi,$$

where $d\text{vol}$ is produced from the basic symplectic structure $\Omega = dp \wedge dq$ and $I'$s and $\phi$'s are classical action-angle variables constructed from $\Omega$. After this is done one can try to compute the partition function, [V1].

Now analysis of the direct spectral problem requires a “second order” approximation

$$\tilde{L} = \frac{1}{z} L^{(-1)} + L^{(0)} + zL^{(1)} + O(z^2),$$
Such an approximation provides the coefficients $k_1^{(0)}$ and $k_1^{(1)}$ (Theorem 9) for the expansion of the function $k$

$$k_1(z) = \frac{N - 1}{z} + k_1^{(0)} + zk_1^{(1)} + O(z^2)$$

on the “upper” sheet of the curve. The desired formula can be easily obtained by a simple application of Cauchy’s theorem. Moreover, in the repulsive case the parameters $I'_\alpha$ are real for all configurations of particles. In order to prove this we show that in the expansion

$$k_\alpha(z) = -\frac{1}{z} + k_\alpha^{(0)} + O(z), \quad \alpha = 2, \ldots, N,$$

the coefficients $k_\alpha^{(0)}$ are distinct for all $\alpha = 2, \ldots, N$ and a generic configuration of particles. That much information can be obtained by perturbation techniques on “lower” sheets.

Presumably $I'_\alpha$ are moduli of the corresponding $N$-sheeted covers and the actions relative to some symplectic structure $\Omega'$ on the phase space, but this is not proved. Compare [V2] for the case of the cubic Schrödinger curves. We will return to this issue elsewhere.

2. Elliptic solutions of the KP hierarchy.

Consider the $N$ particle Hamiltonian on the line

$$H_N = \sum_{n=1}^{N} \frac{p_n^2}{2} - 2\sigma^2 \sum_{n,m=1}^{N} \varphi(q_n - q_m).$$

The Hamiltonian produces a system of first order equations of motion

$$\begin{align*}
q_n &= \frac{\partial H}{\partial p_n}, \quad n = 1, \ldots, N, \\
p_n &= -\frac{\partial H}{\partial q_n} = 4\sigma^2 \sum_{m\neq n} \varphi'(q_n - q_m), \quad n = 1, \ldots, N.
\end{align*}$$

The system can be written in the form

$$(1) \quad \ddot{q}_n = 4\sigma^2 \sum_{m\neq n} \varphi'(q_n - q_m), \quad n = 1, \ldots, N.$$
**Theorem 1 ([K]).** The equations

\[
\sigma \partial_y - \partial_x^2 + 2 \sum_{n=1}^{N} \varphi(x - q_n(y)) \psi = 0,
\]

\[
\psi^\dagger \left[ \sigma \partial_y - \partial_x^2 + 2 \sum_{n=1}^{N} \varphi(x - q_n(y)) \right] = 0
\]

have solutions of the form

\[
\psi(x, y, k, z) = \sum_{n=1}^{N} a_n(y, k, z) \Phi(x - q_n, z)e^{kx+\sigma^{-1}k^2y}
\]

\[
\psi^\dagger(x, y, k, z) = \sum_{n=1}^{N} a_n^\dagger(y, k, z) \Phi(-x + q_n, z)e^{-kx-\sigma^{-1}k^2y},
\]

where \(2 \Phi(x, z) = \frac{\sigma(z-x)}{\sigma(z)} e^{\xi(z)x}, \) if and only if \(q_n(y)\) satisfy the system of equations (1).

The proof is obtained by requiring that singularities of the form \((x - q_n)^{-2}\) and \((x - q_n)^{-1}\) vanish. This condition can be written in a compact form with the aid of \(N \times N\) matrices \(L\) and \(M\)

\[
L_{nm} = \sigma p_n \delta_{nm} + 2 \Phi(q_n - q_m, z)(1 - \delta_{nm}),
\]

\[
M_{nm} = \left( -\varphi(z) + 2 \sum_{s \neq n} \varphi(q_n - q_s) \right) \delta_{nm} + 2 \Phi'(q_n - q_m, z)(1 - \delta_{nm}).
\]

**Lemma 2 ([K]).** The vectors \(a(y, k, z)\) and \(a^\dagger(y, k, z)\) satisfy the equations

\[
(L + 2k)a = 0 \quad \sigma \partial_y + M)a = 0
\]

and

\[
a^\dagger(L + 2k) = 0 \quad a^\dagger(\sigma \partial_y + M) = 0.
\]

These equations determine the curve \(\Gamma_N\) which is the subject of the next section.

### 3. Riemann surface.

The matrix \(L\) can be simplified using a gauge transformation

\[
L = G\tilde{L}G^{-1},
\]

where \(G_{nm} = e^{\xi(z)q_n} \delta_{nm}\). Then

\[
\tilde{L}_{nm} = \sigma p_n \delta_{nm} + 2 \Phi_0(q_n - q_m, z)(1 - \delta_{nm})\]

\(2\sigma(z)\) denotes the Weierstrass function.
and
\[ \Phi_0(x, z) = \frac{\sigma(z - x)}{\sigma(z)\sigma(x)}. \]
The existence of a nontrivial vector \( a : (\tilde{L} + 2k)a = 0 \) implies that \( R_N(k, z) = \det(\tilde{L} + 2k) \) vanishes and this condition determines the curve. We denote by \( P, Q, \) etc., points \((k, z)\) on the curve.

**Lemma 3 ([K]).** The determinant \( R_N(k, z) \) can be written in the form
\[ R_N(k, z) = \sum_{n=0}^N r_n(z)k^n, \]
where \( r_n(z) \) are elliptic functions of \( z \).

The curve \( \Gamma_N \) is an \( N \)-sheeted covering of the elliptic curve. The next lemma describes symmetries of the curve.

**Lemma 4.** (i) \( \sigma = 1 \). The curve \( \Gamma_N \) admits the antiholomorphic involution
\[ \tau_1 : (k, z) \mapsto (\bar{k}, \bar{z}). \]
(ii) \( \sigma = i \). The curve \( \Gamma_N \) admits the antiholomorphic involution
\[ \tau_i : (k, z) \mapsto (-\bar{k}, -\bar{z}). \]

**Proof.** (i) For rectangular lattice \( \sigma(z) = \sigma(\bar{z}) \) and \( \Phi_0(x, z) = \Phi_0(\bar{x}, \bar{z}) \).
Therefore \( R_N(k, \bar{z}) = R_N(\bar{k}, \bar{z}) \).
(ii) Note first \( \Phi_0(x, -z) = -\Phi_0(-x, z) \) and \( ip + 2k = -(ip + 2(\bar{k})) \).
Therefore \( R_N(k, z) = (-1)^N R_N(-\bar{k}, -\bar{z}) \). The proof is finished. \( \square \)

The function \( \Phi_0(x, z) \) has a simple pole at \( z = 0 \) and can be expanded in powers of \( z \)
\[ \Phi_0(x, z) = -\frac{1}{z} + \zeta(x) + \frac{1}{2}(\wp(x) - \zeta^2(x))z + O(z^2). \]
Therefore\(^3\),
\[ \tilde{L}(p, q, z) = \frac{1}{z}L^{(-1)}(p, q) + L^{(0)}(p, q) + zL^{(1)}(p, q) + O(z^2), \]
where
\[
\begin{align*}
L^{(-1)}_{nm} &= -2(1 - \delta_{nm}), \\
L^{(0)}_{nm} &= \sigma p_n \delta_{nm} + 2\zeta(q_n - q_m)(1 - \delta_{nm}), \\
L^{(1)}_{nm} &= (\wp - \zeta^2)(q_n - q_m)(1 - \delta_{nm}).
\end{align*}
\]
The matrix \( L^{(-1)} \) is a constant matrix. Its spectrum and eigenvectors can be easily computed.

\(^3\)We omit \( \tilde{L} \) above \( L \) and \( a \) to simplify the notations.
Let

\[
a_\alpha = \begin{pmatrix} 1 \\
1 e^{i\beta_\alpha} \\
\vdots \\
e^{i\beta_\alpha(N-1)}
\end{pmatrix} \quad \text{and} \quad \beta_\alpha = \frac{2\pi}{N}(\alpha - 1), \quad \alpha = 1, \ldots, N.
\]

Then

\[
\left(L^{(-1)} + 2k_\alpha\right) a_\alpha = 0
\]

with \(k_\alpha = -1\) for \(\alpha = 2, \ldots, N\) and \(k_1 = N - 1\).

The necessary information about the curve is obtained by perturbation of this trivial case. The eigenvectors \(a_\alpha(z)\) and eigenvalues \(k_\alpha(z)\) can be expanded in power series in \(z\)

\[
a_\alpha(z) = a^{(0)}_\alpha + a^{(1)}_\alpha z + a^{(2)}_\alpha z^2 + \cdots, \quad a^{(0)}_\alpha = a_\alpha,
\]

\[
k_\alpha(z) = \frac{1}{z}k^{(-1)}_\alpha + k^{(0)}_\alpha + k^{(1)}_\alpha z + \cdots, \quad k^{(-1)}_\alpha = k_\alpha.
\]

The index “\(\alpha\)” labels the sheets of the curve \(\Gamma_N\). We call the sheet \(\alpha = 1\) the “upper” sheet. In fact, the upper sheet is distinguished by \(k^{(-1)}\), the leading term of the asymptotics. The “lower” sheets \((\alpha \geq 2)\) are distinguished for generic configuration of particles by different values of \(k^{(0)}\) since all \(k^{(-1)} = -1\). This is proved in Lemma 5. The proof of Lemma 6 shows that all leading terms \(a^{(0)}_\alpha\) are distinct. This implies that all “lower” sheets can be indexed according to these asymptotics and \(a^{(0)}_\alpha = a_\alpha\).

**Lemma 5.** For generic configuration of particles \(k^{(0)}\) are distinct.

**Proof.** To prove the statement we need first order perturbation theory for multiple eigenvalues.

We choose \(N - 1\) vectors \(e^{(0)}\) in the subspace generated by \(a^{(0)}_\alpha, \alpha = 2, \ldots, N;\)

\[
e^{(0)} = \sum_{\alpha=2}^{N} \eta_\alpha a^{(0)}_\alpha,
\]

where the \(\eta\)'s depend on \(e^{(0)}\) and are such that

\[
\left(\tilde{L}(z) + 2k(z)\right) e_\gamma(z) = 0.
\]

\(L(z), k_\gamma(z), e_\gamma(z)\) can be expanded in integer powers of \(z\)

\[
k_\gamma(z) = \frac{1}{z}k^{(-1)}_\gamma + k^{(0)}_\gamma + zk^{(1)}_\gamma + \cdots,
\]

\[
e_\gamma(z) = e^{(0)}_\gamma + ze^{(1)}_\gamma + \cdots.
\]
Now we collect terms in the identity
\[
\left[ \frac{1}{z}L^{(-1)} + L^{(0)} + \cdots + \frac{2}{z}k^{(-1)} + 2k^{(0)} + \cdots \right] \left[ e^{(0)} + ze^{(1)} + \cdots \right] = 0.
\]
Terms in \( z^0 \) produce
\[
L^{(-1)}e^{(1)} + L^{(0)}e^{(0)} + 2k^{(-1)}e^{(1)} + 2k^{(0)}e^{(0)} = 0,
\]
and
\[
\left( L^{(-1)}e^{(1)}, a^{(0)}_{\alpha} \right) + \left( L^{(0)}e^{(0)}, a^{(0)}_{\alpha} \right) + 2k^{(-1)} \left( e^{(1)}, a^{(0)}_{\alpha} \right) + 2k^{(0)} \left( e^{(0)}, a^{(0)}_{\alpha} \right) = 0.
\]
Using the selfadjointness of \( L^{(-1)} \)
\[
\left( L^{(-1)}e^{(1)}, a^{(0)}_{\alpha} \right) = \left( e^{(1)}, L^{(-1)}a^{(0)}_{\alpha} \right) = -2k^{(-1)} \left( e^{(1)}, a^{(0)}_{\alpha} \right).
\]
Therefore
\[
\left( L^{(0)}e^{(0)}, a^{(0)}_{\alpha} \right) = -2k^{(0)} \left( e^{(0)}, a^{(0)}_{\alpha} \right)
\]
or
\[
\sum_{\alpha=2}^{N} \eta_{\alpha} \left( L^{(0)}a^{(0)}_{\alpha}, a^{(0)}_{\alpha'} \right) = -2k^{(0)} \eta_{\alpha'}.
\]
The eigenvalues \(-2k^{(0)}\) of the matrix \( (L^{(0)}a^{(0)}_{\alpha}, a^{(0)}_{\alpha'}) \), \( \alpha, \alpha' = 2, \ldots, N \); are distinct for generic configuration of particles. \( \square \)

**Lemma 6.** For all configurations of particles the “zero” order approximations
\[
\tilde{a}^{(0)} = \begin{pmatrix} \tilde{a}^{(0)}(1) \\ \vdots \\ \tilde{a}^{(0)}(N) \end{pmatrix}
\]
of the eigenvectors \( \tilde{a}_{\alpha}(z) = \tilde{a}^{(0)}_{\alpha} + \tilde{a}^{(1)}_{\alpha} z + \cdots \) of the spectral problem \( (\tilde{L} + 2k)\tilde{a} = 0 \) normalized by the condition \( \tilde{a}^{(0)}(1) = 1 \) are given by the formula
\[
\tilde{a}^{(0)}_{\alpha} = \begin{pmatrix} 1 \\ e^{i\beta_{\alpha}} \\ \vdots \\ e^{i\beta_{\alpha}(N-1)} \end{pmatrix} \quad \text{and} \quad \beta_{\alpha} = \frac{2\pi}{N} (\alpha - 1), \quad \alpha = 1, \ldots, N.
\]

**Proof.** If \( A \) as \( n \times n \) matrix then \( AA^\wedge = \det A^{}I \), where \( A^\wedge \) is the matrix which consists of auxiliary minors of \( A \), see also [KNS]. For any column \( r = 1, \ldots, N \)
\[
\tilde{a}(p) = \left[ \tilde{L} + 2k \right]^\wedge_{pr} \left[ \tilde{L} + 2k \right]^\wedge_{1r}
\]
in the vicinity of $z = 0^4$.

We know that \( k(z) = -\frac{1}{z} + k^{(0)} + \cdots \) and
\[
\tilde{L} = -\frac{2}{z}[E - I] + \tilde{L}^{(0)} + \cdots, \quad E_{nm} = 1.
\]

Then
\[
\left[ \tilde{L} + 2k \right]_{pr}^\wedge = \left[ -\frac{2E}{z} + [\tilde{L}^{(0)} + 2k^{(0)}] + \cdots \right]_{pr}^\wedge.
\]

Since \( \text{rank} E = 1 \) we have
\[
\left[ \tilde{L} + 2k \right]_{pr}^\wedge = -\frac{2}{z} \sum_{s=1}^{N} [\tilde{L}^{(0)} + 2k^{(0)}]_{spr} + \cdots,
\]
where the subscript \( s \) means that the \( s \)-th column is replaced by \( \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \).

For a generic configuration \( k^{(0)} \) are distinct on all sheets of the curve. This and the formula for \( \tilde{a}^{(p)} \) imply that all \( \tilde{a}^{(0)} \) are also distinct and therefore match \( \tilde{a}^{(0)}_\alpha \). Since all \( \tilde{a}^{(0)}_\alpha \) are fixed the statement is true for all configurations of particles.

The following two lemmas are simple consequences of the discussion above:

**Lemma 7 ([K]).** The determinant \( R_N(k, z) \) can be written in the form
\[
R_N(k, z) = 2^N \left( k - \left( \frac{N - 1}{z} + k_1^{(0)} + \cdots \right) \right) \prod_{\alpha=2}^{N} \left( k - \left( -\frac{1}{z} + k^{(0)}_\alpha + \cdots \right) \right).
\]

The genus of the curve \( \Gamma_N \) can be easily computed using the Riemann-Hurwitz formula.

**Lemma 8 ([K, KBBT]).** (i) The elliptic case: \( 2\omega, 2\omega' < \infty \). For generic configuration of particles the genus of the curve \( \Gamma_N \) is \( N \).

(ii) The rational case: \( 2\omega, 2\omega' = \infty \). The genus of the curve \( \Gamma_N \) is 0.

4. **Asymptotics for \( k(z) \).**

The main result of this section is the following:

**Theorem 9.** On the “upper” sheet for \( k_1(z) \) the following asymptotics hold
\[
k_1(z) = \frac{N - 1}{z} - \frac{\sigma P_N}{2N} + z \left( \frac{\sigma^2 H_N}{2N^2} - \frac{\sigma^2 P_N^2}{4N^3} \right) + O(z^2),
\]

\(^4\)Note that \( \tilde{a} \) is not a function on the curve, since the entries of the matrix \( \tilde{L} + 2k \) are not elliptic functions.
where
\[ H_N = K_N - \sigma^2 V_N = \sum_{n=1}^{N} \frac{p_n^2}{2} - 2\sigma^2 \sum_{n,m=1}^{N} \varphi(q_n - q_m), \]

and
\[ P_N = \sum_{n=1}^{N} p_n. \]

To prove the theorem we need the following lemma which is the second order perturbation theory of simple eigenvalues adapted to our considerations.

**Lemma 10.** The following identities hold
\[ 2k^{(0)} = -\left(L^{(0)}a^{(0)}, a^{(0)}\right), \]
\[ 2k^{(1)} = \frac{1}{2N} \sum_{\alpha=2}^{N} \left( L^{(0)}a^{(0)}, a^{(0)} \right) \times \left( L^{(0)}a^{(0)}, a^{(0)} \right) - \left( L^{(1)}a^{(0)}, a^{(0)} \right), \]

where
\[ (f,g) = \frac{1}{N} \sum_{n=1}^{N} f_n \bar{g}_n, \]
for any \( f, g \in \mathbb{C}^N. \)

**Proof.** We start with the identity
\[
\left( \frac{1}{z} L^{(-1)} + L^{(-0)} + zL^{(1)} + \cdots \right) \\
+ \frac{2k^{(-1)}}{z} + 2k^{(0)} + 2k^{(1)} + \cdots \right) \left( a^{(0)} + za^{(1)} + \cdots \right) = 0.
\]
Collecting the terms in \( z^{-1} \)
\[
\left( L^{(-1)} + 2k^{(-1)} \right) a^{(0)} = 0.
\]

Note,
\[ a^{(0)} = a^{(0)}_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \]

Therefore \( k^{(-1)} = N - 1 \), as it has to be.

Now we collect terms with \( z^0 \)
\[ L^{(-1)}a^{(1)} + L^{(0)}a^{(0)} + 2k^{(-1)}a^{(1)} + 2k^{(0)}a^{(0)} = 0 \]
and
\[(3)\]
\(\left(L^{(-1)}a^{(1)}, a^{(0)}\right) + \left(L^{(0)}a^{(0)}, a^{(0)}\right) + \left(2k^{(-1)}a^{(1)}, a^{(0)}\right) + \left(2k^{(0)}a^{(0)}, a^{(0)}\right) = 0.\)

Note
\[(4)\]
\(\left(L^{(-1)}a^{(1)}, a^{(0)}\right) = \left(a^{(1)}, L^{(-1)}a^{(0)}\right) = \left(a^{(1)}, -2k^{(-1)}a^{(0)}\right),\)

which implies that the two terms in (3) cancel each other. From the definition \((a^{(0)}, a^{(0)}) = 1\) we obtain the first statement of the lemma.

Now we derive the formulas for \(a^{(1)}\) on the “upper” sheet, which will be useful later on. Multiply (2) by \(a^{(0)}\), \(\alpha = 2, \ldots, N,\)
\[(5)\]
\(\left(L^{(-1)}a^{(1)}, a^{(0)}\right) + \left(L^{(0)}a^{(0)}, a^{(0)}\right) + \left(2k^{(-1)}a^{(1)}, a^{(0)}\right) + \left(2k^{(0)}a^{(0)}, a^{(0)}\right) = 0.\)

The last term vanishes due to the orthogonality of eigenvectors corresponding different eigenvalues. Similar to (4)
\(\left(L^{(-1)}a^{(1)}, a^{(0)}\right) = \left(a^{(1)}, L^{(-1)}a^{(0)}\right) = 2 \left(a^{(1)}, a^{(0)}\right)\)
due to
\(\left(\frac{1}{z}L^{(-1)} + 2 - \frac{1}{z}\right) a^{(0)} = 0, \quad \alpha = 2, \ldots, N.\)

Therefore
\(2 \left(a^{(1)}, a^{(0)}\right) + \left(L^{(0)}a^{(0)}, a^{(0)}\right) + 2(N-1) \left(a^{(1)}, a^{(0)}\right) = 0\)
and
\(\left(a^{(1)}, a^{(0)}\right) = -\frac{1}{2N} \left(L^{(0)}a^{(0)}, a^{(0)}\right), \quad \alpha = 2, \ldots, N.\)

The condition \(||a(z)||^2 = 1 + O(z^2)\) implies
\(1 = (a(z), a(z)) = (a^{(0)} + za^{(1)} + \cdots, a^{(0)} + za^{(1)} + \cdots)\)
\(= (a^{(0)}, a^{(0)}) + z \left[(a^{(0)}, a^{(1)}) + (a^{(1)}, a^{(0)})\right] + \cdots.\)

and \((a^{(1)}, a^{(0)}) = 0.\) Therefore
\(a^{(1)} = \sum_{\alpha=2}^{N} a^{(0)}\left(a^{(1)}, a^{(0)}\right) = -\frac{1}{2N} \sum_{\alpha=2}^{N} a^{(0)} \left(L^{(0)}a^{(0)}, a^{(0)}\right).\)

In order to prove the second formula of the lemma we collect terms with \(z^1\)
\(L^{(-1)}a^{(2)} + L^{(0)}a^{(1)} + L^{(1)}a^{(0)} + 2k^{(-1)}a^{(2)} + 2k^{(0)}a^{(1)} + 2k^{(1)}a^{(0)} = 0\)
and

$$\begin{align*}
(6) & \quad \left( L^{(-1)} a^{(2)}, a^{(0)} \right) + \left( L^{(0)} a^{(1)}, a^{(0)} \right) + \left( L^{(1)} a^{(0)}, a^{(0)} \right) \\
& \quad + 2k^{(-1)} \left( a^{(2)}, a^{(0)} \right) + 2k^{(0)} \left( a^{(1)}, a^{(0)} \right) + 2k^{(1)} \left( a^{(0)}, a^{(0)} \right) = 0.
\end{align*}$$

Note

$$\left( L^{(-1)} a^{(2)}, a^{(0)} \right) = \left( a^{(2)}, L^{(-1)} a^{(0)} \right) = -2k^{(-1)} \left( a^{(2)}, a^{(0)} \right)$$

and two of the terms in the formula (6) vanish. Using the normalization conditions \( (a^{(0)}, a^{(0)}) = 1 \) and \( (a^{(1)}, a^{(0)}) = 0 \) we obtain

$$2k^{(1)} = - \left( L^{(0)} a^{(1)}, a^{(0)} \right) - \left( L^{(1)} a^{(0)}, a^{(0)} \right).$$

Using the formula for \( a^{(1)} \), we finally have

$$2k^{(1)} = \frac{1}{2N} \sum_{\alpha=2}^{N} \left[ \sigma \left( Aa^{(0)}_{\alpha}, a^{(0)} \right) \times \left( L^{(0)} a^{(0)}, a^{(0)} \right) - \left( L^{(1)} a^{(0)}, a^{(0)} \right) \right].$$

The lemma is proved.

Now we use the lemma to compute the coefficients \( k^{(0)} \) and \( k^{(1)} \). Note, first, that \( L^{(0)} \) can be split into two parts \( L^{(0)} = \sigma A + B \), where

$$A_{mn} = p_n \delta_{nm} \quad \text{and} \quad B_{mn} = 2 \zeta(q_n - q_m)(1 - \delta_{nm}).$$

The matrix \( A \) is symmetric and \( B \) is skew-symmetric

$$2k^{(0)} = - \left( L^{(0)} a^{(0)}, a^{(0)} \right) = -\sigma \left( Aa^{(0)}(0), a^{(0)} \right) - \left( Ba^{(0)}, a^{(0)} \right)$$

$$= -\frac{\sigma P_N}{N}.$$

The term with \( B \) vanishes due to skew-symmetry.

The computation of the next term \( k^{(1)} \) is much more involved. Using the decomposition for \( L^{(0)} \) we have

$$2k^{(1)} = \frac{1}{2N} \sum_{\alpha=2}^{N} \left[ \sigma \left( Aa^{(0)}_{\alpha}, a^{(0)} \right) + \left( Ba^{(0)}_{\alpha}, a^{(0)} \right) \right]$$

$$\times \left[ \sigma \left( Aa^{(0)}, a^{(0)} \right) + \left( Ba^{(0)}, a^{(0)} \right) \right] - \left( L^{(1)} a^{(0)}, a^{(0)} \right) = I + II.$$

**Step 1.** Our goal is to evaluate the first term \( I \). Arguments for the attractive and the repulsive case are different and we consider these two cases separately.
Repulsive case. $\sigma = i$. Using the symmetry of $A$ and skew-symmetry of $B$,

$$
\frac{1}{2N} \sum_{\alpha=2}^{N} \left[ i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right] \times \left[ i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right]
$$

$$
= -\frac{1}{2N} \sum_{\alpha=2}^{N} \left| i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2
$$

$$
= -\frac{1}{2N} \sum_{\alpha=1}^{N} \left| i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2 + \frac{1}{2N} \left| \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2.
$$

The last term can be easily estimated

$$
\frac{1}{2N} \left| \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2 = \frac{1}{2N^3} P_N^2.
$$

To estimate the first term, we introduce $\alpha'$ such that

$$
\alpha' = N - \alpha + 2.
$$

Then

$$
\beta_{\alpha'} = \frac{2\pi}{N} \left( \alpha' - 1 \right) = \frac{2\pi}{N} \left( N - \alpha + 1 \right)
$$

and

$$
\beta_\alpha + \beta_{\alpha'} \equiv 0 \pmod{2\pi}.
$$

Now

$$
-\frac{1}{2N} \sum_{\alpha=1}^{N} \left| i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2
$$

$$
= -\frac{1}{4N} \sum_{\alpha=1}^{N} \left| i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2
$$

$$
+ \left| i \left( A a_\alpha^{(0)}, a_{\alpha'} \right) + \left( B a_\alpha^{(0)}, a_{\alpha'} \right) \right|^2
$$

$$
= -\frac{1}{4N} \sum_{\alpha=1}^{N} \left| i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2
$$

$$
+ \left| -i \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) + \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2.
$$

Using the identity $|a + b|^2 + |a - b|^2 = 2|a|^2 + 2|b|^2$ we obtain

$$
I = -\frac{1}{2N} \sum_{\alpha=1}^{N} \left| \left( A a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2 + \left| \left( B a_\alpha^{(0)}, a_\alpha^{(0)} \right) \right|^2 + \frac{1}{2N^3} P_N^2.
$$
Attractive case. \( \sigma = 1 \). Again using properties of \( A \) and \( B \) we have

\[
\frac{1}{2N} \sum_{\alpha=2}^{N} \left[ (Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)}) + (Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right] \times \left[ (Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)}) + (Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right]
\]

\[
= \frac{1}{2N} \sum_{\alpha=2}^{N} \left[ \left( Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) - \left( Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \right] \times \left[ \left( Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) + \left( Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \right]
\]

\[
= \frac{1}{2N} \sum_{\alpha=2}^{N} \left| (Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 - \left| (Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2
\]

\[
+ 2i \frac{1}{2N} \Im \sum_{\alpha=2}^{N} \left( Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \left( Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right).
\]

The second sum vanishes. Indeed,

\[
2 \Im \sum_{\alpha=2}^{N} \left( Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \left( Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right)
\]

\[
= \Im \sum_{\alpha=1}^{N} \left( Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \left( Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) + \left( Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \left( Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right)
\]

\[
= 0.
\]

Therefore,

\[
I = -\frac{1}{2N} \sum_{\alpha=1}^{N} \left| (Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 + \left| (Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 - \frac{1}{2N^3} P_N^2.
\]

Combining the results for two cases \( \sigma^2 = \pm 1 \), we obtain

\[
I = -\frac{1}{2N} \sum_{\alpha=1}^{N} -\sigma^2 \left| (Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 + \left| (Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 - \frac{\sigma^2}{2N^3} P_N^2.
\]

**Step 2.** Now we estimate \( \sum_{\alpha=1}^{N} \left| (Aa_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 \) and \( \left| (Ba_{\alpha}^{(0)}, a_{\alpha}^{(0)}) \right|^2 \). To do this we introduce for two polynomials

\[
P(z) = \sum_{k=1}^{N} p_k z^k, \quad Q(z) = \sum_{k=1}^{N} q_k z^k,
\]
the inner product

$$\langle P, Q \rangle \equiv \sum_{\alpha=1}^{N} P(e^{i\beta_{\alpha}})Q(e^{i\beta_{\alpha}}).$$

It is easy to see that

$$\langle P, Q \rangle = N \sum_{k=1}^{N} p_{k} q_{k}.$$

Therefore

$$\sum_{\alpha=1}^{N} \left| \left( A_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \right| = \sum_{\alpha=1}^{N} \left| \frac{1}{N} \sum_{n=1}^{N} p_{n} e^{-i\beta_{n}(n-1)} \right|^{2} = \frac{1}{N^{2}} \left( \sum_{n=1}^{N} p_{n} z_{n}, \sum_{n=1}^{N} p_{n} z_{n}^{*} \right) = \frac{1}{N} \sum_{n=1}^{N} p_{n}^{2}.$$

Similarly,

$$\sum_{\alpha=1}^{N} \left| \left( B_{\alpha}^{(0)}, a_{\alpha}^{(0)} \right) \right|^{2} = \frac{1}{N} \sum_{n=1}^{N} (b_{n})^{2}, \quad b_{n} = \sum_{s=1}^{N} b_{ns}.$$

Finally,

$$I = \frac{\sigma^{2}}{2N^{2}} \sum_{n=1}^{N} p_{n}^{2} - \frac{1}{2N^{2}} \sum_{n=1}^{N} b_{n}^{2} - \frac{\sigma^{2}}{2N^{3}} P_{N}^{2} = \frac{\sigma^{2}}{N^{2}} K_{N} - \frac{1}{2N^{2}} \sum_{n=1}^{N} b_{n}^{2} - \frac{\sigma^{2}}{2N^{3}} P_{N}^{2}.$$

**Step 3.** Now using the expression for $II^5$

$$II = -\frac{1}{N} \sum_{s \neq s'} (\phi - \zeta^{2})(q_{s} - q_{s'}),$$

we will estimate

$$I + II = \frac{\sigma^{2}}{N^{2}} K_{N} - \frac{1}{2N^{2}} \sum_{n=1}^{N} \left( \sum_{s=1}^{N} 2\zeta(q_{n} - q_{s}) \right)^{2} - \frac{\sigma^{2}}{2N^{3}} P_{N}^{2}
\quad - \frac{1}{N} \sum_{s \neq s'} \phi(q_{s} - q_{s'}) + \frac{1}{N} \sum_{s \neq s'} \zeta^{2}(q_{s} - q_{s'}).$$

---

5 In the rational case $\phi - \zeta^{2} \equiv 0$; $L^{(1)}$ and $II$ vanish identically.
We will prove that

\[- \frac{1}{2N^2} \sum_{n=1}^{N} \left( \sum_{s=1}^{N} 2\zeta(q_n - q_s) \right)^2 + \frac{1}{N} \sum_{s \neq s'} \zeta^2(q_s - q_{s'}) \]

(7) \[= \frac{N - 2}{N^2} \sum_{s \neq s'} \varphi(q_s - q_{s'}).\]

This would imply

\[I + II = \frac{\sigma^2}{N^2} K_N - \frac{1}{N^2} V_N - \frac{\sigma^2}{2N^3} P_N^2\]

and complete the proof of the theorem.

**Step 4.** To prove the identity (7) we make some manipulations with the sums

\[ - \frac{2}{N^2} \sum_{n=1}^{N} \left( \sum_{s=1}^{N} \zeta(q_n - q_s) \right)^2 + \frac{1}{N} \sum_{s \neq s'} \zeta^2(q_s - q_{s'}) \]

\[ = - \frac{2}{N^2} \sum_{s \neq s', s' \neq n} \zeta(q_n - q_s)\zeta(q_n - q_{s'}) + \frac{1}{N} \sum_{s \neq s'} \zeta^2(q_s - q_{s'}) \]

\[ = \frac{1}{N^2} \left[ (N - 2) \sum_{s \neq s'} \zeta^2(q_s - q_{s'}) - 2 \sum_{s \neq s' \neq n} \zeta(q_n - q_s)\zeta(q_n - q_{s'}) \right] \]

\[ = \frac{1}{N^2} \left[ \sum_{s \neq s', s' \neq n} \zeta^2(q_s - q_{s'}) + 2\zeta(q_n - q_s)\zeta(q_n - q_{s'}) \right] .\]

Fix \( k_1 < k_2 < k_3 \) and consider all their permutations. Collect the terms with such indices

\[ \frac{2}{N^2} \sum_{k_1 < k_2 < k_3} \zeta^2(q_{k_1} - q_{k_2}) + \zeta^2(q_{k_2} - q_{k_3}) + \zeta^2(q_{k_3} - q_{k_1}) \]

\[ + 2\zeta(q_{k_1} - q_{k_3})\zeta(q_{k_3} - q_{k_2}) + 2\zeta(q_{k_1} - q_{k_2})\zeta(q_{k_2} - q_{k_3}) \]

\[ + 2\zeta(q_{k_2} - q_{k_1})\zeta(q_{k_1} - q_{k_3}) \]

\[ = \frac{2}{N^2} \sum_{k_1 < k_2 < k_3} (\zeta(q_{k_1} - q_{k_2}) + \zeta(q_{k_2} - q_{k_3}) + \zeta(q_{k_3} - q_{k_1}))^2 .\]

Using the identity

\[ (\zeta(u) + \zeta(v) + \zeta(s))^2 = \varphi(u) + \varphi(v) + \varphi(s) \]

for \( u + v + s = 0 \), we obtain

\[ \frac{2}{N^2} \sum_{k_1 < k_2 < k_3} \varphi(q_{k_1} - q_{k_2}) + \varphi(q_{k_2} - q_{k_3}) + \varphi(q_{k_3} - q_{k_1}) .\]
The last expression is invariant under permutations and therefore it is equal
\[
\frac{2}{6N^2} \sum_{s \neq s' \neq n} \varphi(q_s - q_{s'}) + \varphi(q_s - q_n) + \varphi(q_{s'} - q_n)
= \frac{N - 2}{N^2} \sum_{s \neq s'} \varphi(q_s - q_{s'}).
\]
This completes the proof of the identity (7) and the theorem.

Example. For \( N = 2 \) the equation defining the curve \( \Gamma_2 \) has the form
\[
k^2 + k \frac{\sigma P}{2} + \left( \frac{\sigma^2 P^2}{8} - \frac{\sigma^2 H}{4} - \varphi(z) \right) = 0.
\]
Solving the quadratic equation
\[
k_1(z) = -\frac{\sigma P}{4} + \frac{1}{2} \sqrt{\frac{\sigma^2 P^2}{4} - 4 \left( \frac{\sigma^2 P^2}{8} - \frac{\sigma^2 H}{4} - \varphi(z) \right)}.
\]
Expanding \( k_1(z) \) at \( z = 0 \) we obtain
\[
k_1(z) = \frac{1}{z} - \frac{\sigma P}{4} + z \left( \frac{\sigma^2 H}{8} - \frac{\sigma^2 P^2}{32} \right) + O(z^2).
\]

5. Trace formula.

If the total momentum vanishes, \( P_N = 0 \), then the result of Theorem 9 becomes
\[
(8) \quad k_1(z) = \frac{N - 1}{z} + z \frac{\sigma^2 H_N}{2N^2} + O(z^2).
\]

Theorem 11. The following identity holds
\[
\frac{\sigma^2}{2N^2} H_N = \sum_{\alpha=2}^{N} I'_{\alpha},
\]
where \( I'_{\alpha} = -k_{\alpha}^{(1)} \).

\( \sigma = i \). The antiinvolution \( \tau_i \) does not permute the sheets of the curve. The variables \( I'_{\alpha} \) are real for all configurations of particles.

\( \sigma = 1 \). For some configurations of particles the antiinvolution \( \tau_1 \) permutes some lower sheets \( \alpha \) and \( \tau_1 \alpha \). It also leaves the other sheets invariant. Variables \( I_{\alpha} \) and \( I'_{\tau_1 \alpha} \) corresponding permuting sheets form complex conjugate pairs. All other \( I' \)s corresponding invariant sheets are real.

Proof. By Cauchy’s theorem
\[
\frac{1}{2\pi i} \int_{\gamma_i} k(P) \varphi(z(P)) dz(P) = -\sum_{\alpha=2}^{N} \frac{1}{2\pi i} \int_{\gamma_\alpha} k(P) \varphi(z(P)) dz(P),
\]
where $\gamma_\alpha$ is a small contour surrounding $P_\alpha$, the point on the $\alpha$'th sheet above $z = 0$. The asymptotics (8) implies the result.

The second part of the proof is different for repulsive and attractive cases. We treat them separately.

**Repulsive case.** $\sigma = i$. If the pair $(k, z)$ satisfies
\[ k(z) = \frac{1}{z}k^{(-1)}_\alpha + k^{(0)}_\alpha + k^{(1)}_\alpha z + \cdots \]
in the vicinity of $P_\alpha$, then the pair $(-\bar{k}, -\bar{z})$ also satisfies similar expression in the vicinity of $P_{\tau_\alpha}$
\[ -\bar{k}(z) = \frac{1}{-\bar{z}}k^{(-1)}_{\tau_\alpha} + k^{(0)}_{\tau_\alpha} + k^{(1)}_{\tau_\alpha}(-\bar{z}) + \cdots. \]

Therefore
\[ k(z) = \frac{1}{z}k^{(-1)}_{\tau_\alpha} - \bar{k}^{(0)}_{\tau_\alpha} + \bar{k}^{(1)}_{\tau_\alpha}z - \cdots. \]

Comparision shows that $k^{(n)}_\alpha = \bar{k}^{(n)}_{\tau_\alpha}$ for $n$ odd and $k^{(n)}_\alpha = -\bar{k}^{(n)}_{\tau_\alpha}$ for $n$ even.

We know that $k^{(-1)}_1 = N - 1$ and $k^{(-1)}_\alpha = -1$ for $\alpha = 2, \ldots, N$ so that $\tau_i$ leaves the upper sheet invariant.

The Lemma 5 states that the values of $k^{(0)}_\alpha$ are distinct on different sheets of the curve for a generic configuration of particles. The matrix $(L^{(0)}a^{(0)}_a, a^{(0)}_\alpha)$ (see the proof of Lemma 5) is skew symmetric and therefore all $k^{(0)}_\alpha$ are pure imaginary: $\bar{k}^{(0)}_\alpha = -k^{(0)}_\alpha$. These together with the identity $\bar{k}^{(1)}_\alpha = -\bar{k}^{(1)}_{\tau_\alpha}$ imply that $\tau_\alpha = \alpha$ and antiinvolution $\tau_1$ does not permute sheets. Therefore $k^{(1)}_\alpha = \bar{k}^{(1)}_{\tau_\alpha}$ and approximation arguments complete the proof.

**Attractive case.** $\sigma = 1$. Arguments as before lead to $k^{(n)}_\alpha = \bar{k}^{(n)}_{\tau_\alpha}$ for all $n$. Skew-symmetry of the matrix $L^{(0)}$ is lost but for generic configuration of particles $k^{(0)}_\alpha$ are distinct. The antiinvolution $\tau_1$ leaves the upper sheet invariant, but is can permute the lower sheets of the curve. An example after the proof of the Theorem shows how it happens for $N = 3$. Therefore, some lower sheets $\alpha$ and $\tau_1\alpha$ are interchanged by antiinvolution $\tau_1$ and some are invariant. These and $k^{(-1)}_\alpha = \bar{k}^{(-1)}_{\tau_\alpha}$, together with approximation arguments complete the proof of Theorem.  

**Example.** For $N = 3$ we can explicitly compute expansion for $k_\alpha(z)$ at $z = 0$. In this case
\[ R_3(k, z) = 8k^3 + 4\sigma Pk^2 + [\sigma^2(P^2 - 2H) - 24\psi(z)]k + 8\sigma J + 8\psi'(z), \]
where
\[ 8J = \sigma^2p_1p_2p_3 + 4p_1\psi(q_2 - q_3) + 4p_2\psi(q_1 - q_3) + 4p_3\psi(q_1 - q_2). \]
If \( P = 0 \) the equation of the curve \( \Gamma_3 \) becomes
\[
k^3 - k \left[ \sigma^2 H \frac{4}{4} + 3\varphi(z) \right] + \sigma J + \varphi'(z) = 0.
\]
Cardano’s formula for the roots \( k^3 + pk - q = 0 \) has the form
\[
k_1(z) = \frac{3}{z} \sqrt{\frac{q}{2} + \sqrt{R}} + \frac{3}{z} \sqrt{\frac{q}{2} - \sqrt{R}},
\]
\[
k_2(z) = \omega^2 \frac{3}{z} \sqrt{\frac{q}{2} + \sqrt{R}} + \omega \left( \frac{3}{z} \sqrt{\frac{q}{2} - \sqrt{R}} \right),
\]
\[
k_3(z) = \omega^3 \frac{3}{z} \sqrt{\frac{q}{2} + \sqrt{R}} + \omega^2 \left( \frac{3}{z} \sqrt{\frac{q}{2} - \sqrt{R}} \right),
\]
where \( \omega = e^{\frac{2\pi i}{3}} \), \( R = \frac{q^2}{4} + \frac{p^3}{27} \). After elementary but tiresome calculations we obtain
\[
k_1(z) = 2 \frac{z}{z} + z \frac{\sigma^2 H}{18} + O(z^2),
\]
\[
k_2(z) = -1 \frac{1}{z} + \frac{\sigma\sqrt{3H}}{6} + z \left( \frac{J}{\sqrt{3H}} - \frac{\sigma^2 H}{36} \right) + O(z^2),
\]
\[
k_3(z) = -1 \frac{1}{z} - \frac{\sigma\sqrt{3H}}{6} + z \left( -\frac{J}{\sqrt{3H}} - \frac{\sigma^2 H}{36} \right) + O(z^2).
\]

Consider the case \( \sigma = 1 \). If \( H < 0 \), then \( \tau_1 \) permutes lower sheets of the curve \( \Gamma_3 \); \( I_2 \) and \( I_3 \) are complex conjugate. If \( H > 0 \), then \( \tau_1 \) leaves lower sheets invariant; \( I_2 \) and \( I_3 \) are real.

**Lemma 12.** Let \( p_n = 0, n = 1, \ldots, N \). Then the curve \( \Gamma_N \) admits involution
\[
\tau_- : (k, z) \rightarrow (-k, -z).
\]
The involution \( \tau_- \) leaves the upper sheet invariant and \( k_1^{(n)} = 0 \) for \( N \) even.
It can permute lower sheets and in this case \( k_\alpha^{(n)} = k_{\tau \alpha}^{(n)} \) for \( n \) odd and \( k_\alpha^{(n)} = -k_{\tau \alpha}^{(n)} \) for \( n \) even.

\( \sigma = 1 \). All \( k_\alpha^{(0)}, \alpha = 2, \ldots, N \) are pure imaginary.

**Proof.** Using the identity \( \Phi(q, -z) = -\Phi(-q, z) \) we have
\[
L(q, p = 0, -z) + 2(-k) = |L(q, p = 0, z) + 2k|^T (-1)^N.
\]
Therefore \( R(-k, -z) = (-1)^N R(k, z) \) and the existence of \( \tau_- \) is proved.

If a pair \( (k, z) \) satisfies
\[
k = \frac{1}{z} k_\alpha^{(-1)} + k_\alpha^{(0)} + k_\alpha^{(1)} z + \cdots.
\]
Then
\[-k = \frac{1}{z} k_{\tau-\alpha}^{(-1)} + k_{\tau-\alpha}^{(0)} + k_{\tau-\alpha}^{(1)}(-z) + \cdots.\]
Therefore $k_{\alpha}^{(n)} = k_{\tau-\alpha}^{(n)}$ for $n$ odd and $k_{\alpha}^{(n)} = -k_{\tau-\alpha}^{(n)}$ for $n$ even. Since $k_1^{(-1)} = N - 1$ and $k_{\alpha}^{(0)} = -1$, $\alpha = 2, \ldots, N$; $\tau_-$ leaves the upper sheet invariant. The involution $\tau_-$ can permute the lower sheets. It is demonstrated by the example for $N = 3$.

$\sigma = 1$. Skew-selfadjointness of the matrix $L^{(0)}$ is restored under the conditions of the Lemma. \[\square\]

6. Appendix.

The Weierstrass $\sigma(z)$ has periods $2\omega$ and $2\omega'$ and is defined as
\[\sigma(z) = z \prod \left\{ \left( 1 - \frac{z}{\omega} \right) \exp \left( \frac{z}{\omega} + \frac{z^2}{2\omega^2} \right) \right\},\]
where $\omega = 2\omega n + 2\omega' n'$ and $\prod$ is taken with $n, n' \in \mathbb{Z}$; $n_2 + n'^2 > 0$. $\zeta(z)$ and $\varphi(z)$ are defined similarly:
\[\zeta(z) = \frac{d}{dz} \log \sigma(z), \quad \varphi(z) = -\frac{d}{dz} \zeta(z).\]
For more information see [HC].

Acknowledgments. Finally, I would like to thank I.M. Krichever and H.P. McKean for numerous stimulating discussions.

References


Received September 25, 1997 and revised May 1, 1998. The work is supported by NSF grant DMS-9501002.

KANSAS STATE UNIVERSITY
MANHATTAN, KS 66502
E-mail address: vaninsky@math.ksu.edu
A COMMUTATION FORMULA FOR ROOT VECTORS
IN QUANTIZED ENVELOPING ALGEBRAS

Nanhua Xi

Root vectors are important to understand quantized enveloping algebras. In this paper we establish a commutation formula for root vectors. By means of the formula we show that particular orders on root system are not necessary in constructing some integral bases of a quantized enveloping algebra (Theorem 2.4). Moreover using the formula we can show that certain PBW bases are orthogonal bases of the bilinear form considered by Kashiwara in his work on crystal bases, see 3.9.

In [CK] there is a commutation formula for root vectors, our formula here is stronger. For the bilinear form obtained through Drinfeld dual (see [L5, LS]) Lusztig and Levendorski-Soibelmana showed that certain PBW bases are orthogonal, see loc. cit. However the proofs in [L5, LS] essentially rely on the property [L5, 38.2.1] which does not hold for the bilinear form in [K], so it is not easy to use the methods of [L5, LS] to prove Theorem 3.9.

The paper is organized as follows. In Section 1 we fix some notation. In Section 2 we establish the commutation formula, then prove Theorem 2.4 and state two conjectures. In Section 3 we show that certain PBW bases are orthogonal bases of the bilinear form considered in [K].

1. Preliminaries.

1.1. Let $U$ be the quantized enveloping algebra over $\mathbb{Q}(v)$ ($v$ an indeterminate) corresponding to a Cartan matrix $(a_{ij})$ of rank $n$. Then $U$ is an associative $\mathbb{Q}(v)$-algebra with generators $E_i$, $F_i$, $K_i, K_i^{-1}$ ($i = 1, 2, ..., n$) which satisfy the quantized Serre relations. The algebra $U$ has a Hopf algebra structure. Let $U_\mathcal{A}$ be the $\mathcal{A} = \mathbb{Z}[v, v^{-1}]$-subalgebra of $U$ generated by all divided powers $E_i^{(a)}$, $F_i^{(a)}$ and $K_i$, $K_i^{-1}$. We refer to [L2] for the definitions, noting that for defining the divided powers we need to choose integers $d_i \in \{1, 2, 3\}$ such that $(d_ia_{ij})$ is symmetric. As usual we denote the positive parts and negative parts by $U^+, U_\mathcal{A}^+$, $U^-$, $U_\mathcal{A}^-$ respectively.

1.2. Let $R \subset \mathbb{Z}^n$ be the root system with simple roots $\alpha_i = (a_{i1}, a_{i2}, ..., a_{in})$. For $\mu = (\mu_1, ..., \mu_n) \in \mathbb{Z}^n$, we also write $\langle \mu, \alpha_i^\vee \rangle$ for $\mu_i$. Define $s_i : \mathbb{Z}^n \to \mathbb{Z}^n$
by \( s_i \mu = \mu - \langle \mu, \alpha_i^\vee \rangle \alpha_i \). The reflections \( s_1, s_2, \ldots, s_n \) generate the Weyl group \( W \) of the root system \( R \). Denote by \( R^+ \) the set of positive roots. For \( \lambda = \sum_{i=1}^{n} a_i \alpha_i, \mu = \sum_{j=1}^{n} b_j \alpha_j \), we define \((\lambda | \mu) = \left\langle \sum_{i=1}^{n} a_i \alpha_i, \sum_{j=1}^{n} b_j \alpha_j \right\rangle \). We have \((\lambda | \mu) = (\mu | \lambda)\). The form \( (\cdot | \cdot) \) is non-degenerate and is \( W \)-invariant.

Let \( T_i \) be the automorphisms \( T_i^{-1} \) of \( U \) in \([L5, 37.1.3]\). For each \( w \in W \) we define \( T_w \) as in \([L5]\). We shall write \( \Omega, \Psi : U \to U^{\text{opp}} \) the \( \mathbb{Q} \)-algebra homomorphisms defined by

\[
\Omega E_i = F_i, \quad \Omega F_i = E_i, \quad \Omega K_i = K_i^{-1}, \quad -v \equiv v^{-1};
\]

\[
\Psi E_i = E_i, \quad \Psi F_i = F_i, \quad \Psi K_i = K_i^{-1}, \quad \Psi v = v.
\]

We have \( \Psi T_i \Psi = T_i^{-1}, \quad \Omega T_i = T_i \Omega \). Let \( \Omega' : U \to U \) be the \( \mathbb{Q}(v) \)-algebra automorphism defined by

\[
\Omega' E_i = F_i, \quad \Omega' F_i = E_i, \quad \Omega' K_i = K_i^{-1}.
\]

2. The commutation formula.

2.1. Let \( s_{i_1} s_{i_2} \cdots s_{i_{\nu}} \) be a reduced expression of the longest element \( w_0 \) of \( W \), thus \( \nu = |R^+| \). We have a bijection \([1, \nu] \to R^+ \) defined by

\[
j \to s_{i_1} s_{i_2} \cdots s_{i_{\nu}}(\alpha_{i_j}).
\]

This gives rise to a total order on \( R^+ \). If \( \beta \in R^+ \) corresponds to \( j \), we set \( w_\beta = s_{i_1} s_{i_2} \cdots s_{i_{\nu-1}} \). Then define

\[
E_\beta^{(a)} = T_{\nu}^{(a)}(E_{i_j}^{(a)}) \in U^+, \quad F_\beta^{(a)} = T_{\nu}^{(a)}(F_{i_j}^{(a)}) \in U^-.
\]

We have \( E_\beta^{(a)} \in U_A^+ \) and \( F_\beta^{(a)} \in U_A^- \) \((a \in \mathbb{N})\).

(a) Let \( i = (i_1, \ldots, i_{\nu}) \). It is known the following elements

\[
E_i^A = E_i^{(a_1)} T_{i_1}^{(a_2)} \cdots T_{i_2}^{(a_3)} \cdots T_{i_\nu}^{(a_{\nu-1})}, \quad A = (a_1, \ldots, a_{\nu}) \in \mathbb{N}^\nu,
\]

form an \( \mathcal{A} \)-base of \( U_A \), see \([DL]\).

(b) Let \( w, u \in W \) such that \( l(wu) = l(w) + l(u) \) and let \( s_{i_1} \cdots s_{i_k} \) be a reduced expression of \( u \). Let \( U_{A,u}^+ \) be the \( \mathcal{A} \)-submodule of \( U_A^+ \) generated by the elements \( E_i^{(a_1)} T_{i_1}^{(a_2)} \cdots T_{i_2}^{(a_3)} \cdots T_{i_{k-1}}^{(a_{k-1})} (E_{i_k}^{(a_k)}) (a_1, \ldots, a_k) \in \mathbb{N} \). Then \( T_w(U_A^+) \) is contained in \( U_A^+ \) and \( U_{A,u}^+ \) is independent of the choice of the reduced expression. See \([DL, L2]\).

2.2. We have seen that for each reduced expression \( s_{i_1} \cdots s_{i_{\nu}} \) of \( w_0 \), one can construct an \( \mathcal{A} \)-basis \( \{E_i^A\}_{A \in \mathbb{N}^\nu} \) of \( U_A^+ \). Note that the element \( E_i^A \) is a product of some divided powers of root vectors and the order of the factors in the product is determined by the reduced expression. We will show that we can arrange the product in any fixed total order on \( R^+ \) (see Theorem
2.4). For the purpose we need the following result, which is stronger than [CK, Lemma 1.7].

**Theorem 2.3.** Let $s_{i_1} \cdots s_{i_k}$ be a reduced expression and $1 \leq j < k$. Let $\beta_m = s_{i_1} \cdots s_{i_{m-1}} (\alpha_{i_m})$ ($1 \leq m \leq k$) and define $E^{(a)}_{\beta_m} = T_{i_1} \cdots T_{i_{m-1}}(E^{(a)}_{i_m})$ $(a \in \mathbb{N})$. Then we have

\[
E^{(a)}_{\beta_k} E^{(b)}_{\beta_j} - v^{ab(\beta_j|\beta_k)} E^{(b)}_{\beta_j} E^{(a)}_{\beta_k} = \sum_{a_j, \ldots, a_k \in \mathbb{N}, a_j < b} \rho(a_j, a_{j+1}, \ldots, a_k) E^{(a_j)}_{\beta_j} E^{(a_{j+1})}_{\beta_{j+1}} \cdots E^{(a_k)}_{\beta_k},
\]

where $\rho(a_j, a_{j+1}, \ldots, a_k) \in A$. Note that we have $a_j \beta_j + \cdots + a_k \beta_k = a \beta_k + b \beta_j$ if $\rho(a_j, a_{j+1}, \ldots, a_k) \neq 0$.

**Proof.** We use induction on $k - j$ to prove the theorem. We may assume that $j = 1$. To see this, apply $T_{i_{j-1}}^{-1} \cdots T_{i_1}^{-1}$ to the wanted identity (*).

Let $u$ be the shortest element of the coset $s_{i_1} \cdots s_{i_{k-1}} \langle s_{i_{k-1}}, s_{i_k} \rangle$ (we use $\langle s_{i_{k-1}}, s_{i_k} \rangle$ for the subgroup of $W$ generated by $s_{i_{k-1}}, s_{i_k}$). Then $s_{i_1} \cdots s_{i_{k-1}} = uu'$, where $u' \in \langle s_{i_{k-1}}, s_{i_k} \rangle$ and $l(uu') = l(u) + l(u')$. Note that $u(\alpha_{i_k})$ and $u(\alpha_{i_{k-1}})$ are contained in $R^+$.

When $u = e$ is the neutral element of $W$, the required identity (*) follows from the formulas in [L2].

From now on, we suppose that $l(u) \geq 1$. Assume that (*) is true if $j, k$ are replaced by $j', k'$ respectively with $1 \leq j' < k' \leq k$ and $k' - j' < k - j$, and assume that the Cartan matrix includes no factors of type $G_2$.

**Case A.** $u = s_{i_1} \cdots s_{i_m}$ and $u' = s_{i_{m+1}} \cdots s_{i_{k-1}}$ for some $m \in [1, k - 2]$.

When $u'(\alpha_{i_k})$ is a simple root $\alpha_i$, then $\beta_k = u(\alpha_i)$. Moreover, we have $E^{(a)}_{\beta_k} = T_u(E^{(a)}_{i_k})$. Note that $l(u) \leq k - 2$. By induction hypothesis we see

\[
E^{(a)}_{\beta_k} E^{(b)}_{\beta_j} - v^{ab(\beta_j|\beta_k)} E^{(b)}_{\beta_j} E^{(a)}_{\beta_k} = \sum_{a_1, \ldots, a_m, a_k \in \mathbb{N}, a_i < b} \rho(a_1, \ldots, a_m, a_k) E^{(a_1)}_{\beta_1} \cdots E^{(a_m)}_{\beta_m} E^{(a_k)}_{\beta_k},
\]

where $\rho(a_1, \ldots, a_m, a_k) \in A$. Thus the desired identity (*) is true in this case.

Now assume that $u'(\alpha_{i_k})$ is not a simple root. We have the following cases.

1. $u' = s_{i_{k-1}}$ and $\langle \alpha_{i_k}, \alpha_{i_{k-1}}^{\vee} \rangle = -1$, then $u = s_{i_1} \cdots s_{i_{k-2}}$. 

---

[CK, Lemma 1.7]

[L2]
(2) \( u' = s_{i_k} s_{i_{k-1}} \) and \( \langle \alpha_{i_k}, \alpha_{i_{k-1}}^\vee \rangle = -1 \), \( \langle \alpha_{i_{k-1}}, \alpha_{i_k}^\vee \rangle = -2 \), that is, \( \alpha_{i_{k-1}} \) is a long root and \( \alpha_{i_k} \) is a short root. We have \( d_{i_{k-1}} = 2 \), \( d_{i_k} = 1 \), and \( u = s_{i_1} \cdots s_{i_{k-1}} \).

(3) \( u' = s_{i_{k-1}} \) and \( \langle \alpha_{i_k}, \alpha_{i_{k-1}}^\vee \rangle = -2 \), \( \langle \alpha_{i_{k-1}}, \alpha_{i_k}^\vee \rangle = -1 \). Then \( d_{i_{k-1}} = 1 \), \( d_{i_k} = 2 \), and \( u = s_{i_1} \cdots s_{i_{k-2}} \).

(4) \( u' = s_{i_k} s_{i_{k-1}} \) and \( \langle \alpha_{i_k}, \alpha_{i_{k-1}}^\vee \rangle = -2 \), \( \langle \alpha_{i_{k-1}}, \alpha_{i_k}^\vee \rangle = -1 \). Then \( d_{i_{k-1}} = 1 \), \( d_{i_k} = 2 \), and \( u = s_{i_1} \cdots s_{i_{k-3}} \).

Define \( \alpha = u(a_{i_{k-1}}) \) and \( \gamma = u(a_{i_k}) \), they are positive roots. Set \( E_\alpha = Tu(E_{i_{k-1}}) \) and \( E_\gamma = Tu(E_{i_k}) \). We have \( E_\alpha, E_\gamma \in U_A^+. \) In cases (1) and (3), we have \( \alpha = \beta_{k-1} \) and \( E_\alpha = E_{\beta_{k-1}} \). In cases (2) and (4), we have \( \gamma = \beta_{k-2} \) and \( E_\gamma = E_{\beta_{k-2}} \).

By induction hypothesis we get

\[ (b) \quad E_\alpha E_{\beta_1} - v^{(\beta_1|\alpha)} E_{\beta_1} E_\alpha = \sum_{\alpha = 0, \cdots, a_{k-2} \in \mathbb{Z}} \rho'(a_2, \ldots, a_{k-2}) E_{\beta_2}^{(a_2)} \cdots E_{\beta_{k-2}}^{(a_{k-2})}, \]

where \( \rho'(a_2, \ldots, a_{k-2}) \in \mathbb{A} \). We shall simply write \( X \) for the right hand side of the above identity. Then \( E_\alpha E_{\beta_1} - v^{(\beta_1|\alpha)} E_{\beta_1} E_\alpha = X \). Note that \( a_2 \beta_2 + \cdots + a_{k-2} \beta_{k-2} = \beta_1 + \alpha \) if \( \rho'(a_2, \ldots, a_{k-2}) \neq 0 \). Moreover, for cases (2) and (4), \( a_{k-2} = 0 \) if \( \rho'(a_2, \ldots, a_{k-2}) \neq 0 \).

\[ (c) \quad E_\gamma E_{\beta_1} - v^{(\beta_1|\gamma)} E_{\gamma} E_{\beta_1} = \sum_{\alpha = 0, \cdots, a_{k-2} \in \mathbb{Z}} \rho''(a_2, \ldots, a_{k-2}) E_{\beta_2}^{(a_2)} \cdots E_{\beta_{k-2}}^{(a_{k-2})}, \]

where \( \rho''(a_2, \ldots, a_{k-2}) \in \mathbb{A} \). We shall simply write \( Y \) for the right hand side of the above identity. Then \( E_\gamma E_{\beta_1} - v^{(\beta_1|\gamma)} E_{\beta_1} E_\gamma = Y \). Note that \( a_2 \beta_2 + \cdots + a_{k-2} \beta_{k-2} = \beta_1 + \gamma \) if \( \rho'(a_2, \ldots, a_{k-2}) \neq 0 \). Moreover, for cases (2) and (4), \( a_{k-2} = 0 \) if \( \rho''(a_2, \ldots, a_{k-2}) \neq 0 \).

Now assume that we are in case (1), then

\[ (d) \quad E_{\beta_k} = Tu(T_{i_{k-1}}(E_{i_k}) = Tu(T_{i_k} E_{i_{k-1}} - v^{-d} E_{i_{k-1}} E_{i_k}) \]

\[ = E_\gamma E_{\beta_{k-1}} - v^{-d} E_{\beta_{k-1}} E_\gamma, \]

where \( d = d_{i_{k-1}} \).

Therefore we have

\[ (e) \quad E_{\beta_k} E_{\beta_1} = E_\gamma E_{\beta_{k-1}} E_{\beta_1} - v^{-d} E_{\beta_{k-1}} E_\gamma, \]

\[ = E_\gamma (v^{(\beta_1|\beta_{k-1})} E_{\beta_1} E_{\beta_{k-1}} + X) - v^{-d} E_{\beta_{k-1}} (v^{(\beta_1|\gamma)} E_{\beta_1} E_\gamma + Y) \]

\[ = v^{(\beta_1|\beta_{k-1})} (v^{(\beta_1|\gamma)} E_{\beta_1} E_\gamma + Y) E_{\beta_{k-1}} + E_\gamma X \]

\[ - v^{-d} v^{(\beta_1|\gamma)} (v^{(\beta_1|\beta_{k-1})} E_{\beta_1} E_{\beta_{k-1}} + X) E_\gamma - v^{-d} E_{\beta_{k-1}} Y. \]
Repeatedly using induction hypothesis we get

\[(f)\]  
\[E_\gamma X = v^{(\gamma|\beta_1 + \beta_{k-1})}XE_\gamma + \sum_{b_2, \ldots, b_{k-2} \in \mathbb{N}} \xi'(b_2, \ldots, b_{k-2})E_{\beta_2}^{(b_2)} \cdots E_{\beta_{k-2}}^{(b_{k-2})},\]

where \(\xi'(b_2, \ldots, b_{k-2}) \in \mathcal{A}\).

\[(g)\]  
\[E_{\beta_{k-1}} Y = \sum_{b_2, \ldots, b_{k-1} \in \mathbb{N}} \xi''(b_2, \ldots, b_{k-1})E_{\beta_2}^{(b_2)} \cdots E_{\beta_{k-1}}^{(b_{k-1})},\]

where \(\xi''(b_2, \ldots, b_{k-1}) \in \mathcal{A}\).

We have

\[(h)\]  
\[(\gamma|\beta_1 + \beta_{k-1}) = (\beta_1|\gamma) + (\beta_{k-1}|\gamma) = (\beta_1|\gamma) + (\alpha_{i_{k-1}}|\alpha_{i_k}) = (\beta_1|\gamma) - d.\]

Moreover \(\beta_k = \beta_{k-1} + \gamma\).

Combining (e)-(h) we get

\[(i)\]  
\[E_{\beta_k}E_{\beta_1} - v^{(\beta_1|\beta_k)}E_{\beta_1}E_{\beta_k} = \sum_{a_2, \ldots, a_{k-1} \in \mathbb{N}} \eta(a_2, \ldots, a_{k-1})E_{\beta_2}^{(a_2)} \cdots E_{\beta_{k-1}}^{(a_{k-1})},\]

where \(\eta(a_2, \ldots, a_{k-1}) \in \mathcal{A}\).

Using induction on \(a, b\), and using (i) and induction hypothesis repeatedly, we see

\[(j)\]  
\[E_{\beta_k}^{(a)}E_{\beta_1}^{(b)} - v^{ab(\beta_1|\beta_k)}E_{\beta_1}^{(b)}E_{\beta_k}^{(a)} = \sum_{a_1, \ldots, a_k \in \mathbb{N}} \rho(a_1, \ldots, a_k)E_{\beta_1}^{(a_1)} \cdots E_{\beta_k}^{(a_k)},\]

where \(\rho(a_1, \ldots, a_k) \in \mathcal{A}\) (here we need 2.1 (a)). Thus in case (1) the identity (*) is true.

Now assume that we are in case (2). Then

\[(k)\]  
\[E_{\beta_k} = T_uT_{i_k}T_{i_{k-1}}(E_{i_k}) = T_u(E_{i_{k-1}}E_{i_k} - v^{-2}E_{i_k}E_{i_{k-1}}) = E_\alpha E_{\beta_{k-2}} - v^{-2}E_{\beta_{k-2}}E_\alpha.\]

As a similar argument for case (1) we see that the identity (*) is true in this case.

Now assume that we are in case (3). Then

\[(l)\]  
\[E_{\beta_k} = T_uT_{i_{k-1}}(E_{i_k}) = T_u(E_{i_k}E_{i_{k-1}}^{(2)} - v^{-1}E_{i_{k-1}}E_{i_k}E_{i_{k-1}} + v^{-2}E_{i_{k-1}}^{(2)}E_{i_k}) = E_\gamma E_{\beta_{k-1}}^{(2)} - v^{-1}E_{\beta_{k-1}}E_{\beta_{k-1}} + v^{-2}E_{\beta_{k-1}}^{(2)}E_{\gamma}.\]

We have
(m) \[ E_\gamma E_{\beta_{k-1}}^{(2)} E_{\beta_1} \]
\[ = \frac{1}{2} \gamma \left( v(\beta_1|\beta_{k-1}) E_{\beta_1} E_{\beta_{k-1}} + X \right) \]
\[ = \frac{1}{2} v(\beta_1|\beta_{k-1}) E_{\gamma} \left( v(\beta_1|\beta_{k-1}) E_{\beta_1} E_{\beta_{k-1}} + X \right) E_{\beta_{k-1}} + \frac{1}{2} E_\gamma E_{\beta_{k-1}} X \]
\[ = v^2(\beta_1|\beta_{k-1}) \left( v(\beta_1|\gamma) \gamma E_{\beta_1} E_{\gamma} + Y \right) E_{\beta_{k-1}}^{(2)} + \frac{1}{2} v(\beta_1|\beta_{k-1}) E_{\gamma} X E_{\beta_{k-1}} \]
\[ + \frac{1}{2} E_\gamma E_{\beta_{k-1}} X, \]

(n) \[ E_{\beta_{k-1}} E_\gamma E_{\beta_{k-1}} E_{\beta_1} \]
\[ = E_{\beta_{k-1}} E_{\gamma} \left( v(\beta_1|\beta_{k-1}) E_{\beta_1} E_{\beta_{k-1}} + X \right) \]
\[ = v(\beta_1|\beta_{k-1}) E_{\beta_{k-1}} \left( v(\beta_1|\gamma) \gamma E_{\beta_1} E_{\gamma} + Y \right) E_{\beta_{k-1}} + E_{\beta_{k-1}} E_{\gamma} X \]
\[ = v(\beta_1|\beta_{k-1} + \gamma) \left( v(\beta_1|\beta_{k-1}) E_{\beta_1} E_{\beta_{k-1}} + X \right) E_{\gamma} E_{\beta_{k-1}} \]
\[ + v(\beta_1|\beta_{k-1}) E_{\beta_{k-1}} Y E_{\beta_{k-1}} + E_{\beta_{k-1}} E_{\gamma} X, \]

(o) \[ E_{\beta_{k-1}}^{(2)} E_\gamma E_{\beta_1} \]
\[ = E_{\beta_{k-1}}^{(2)} \left( v(\beta_1|\gamma) \gamma E_{\beta_1} + Y \right) \]
\[ = \frac{1}{2} \left( v(\beta_1|\gamma) \gamma E_{\beta_{k-1}} \left( v(\beta_1|\beta_{k-1}) E_{\beta_1} E_{\beta_{k-1}} + X \right) E_{\gamma} + E_{\beta_{k-1}}^{(2)} Y \right) \]
\[ = \frac{1}{2} v(\beta_1|\beta_{k-1} + \gamma) \left( v(\beta_1|\beta_{k-1}) E_{\beta_1} E_{\beta_{k-1}} + X \right) E_{\beta_{k-1}} E_{\gamma} \]
\[ + \frac{1}{2} v(\beta_1|\gamma) E_{\beta_{k-1}} X E_{\gamma} + E_{\beta_{k-1}}^{(2)} Y. \]

Using induction hypothesis repeatedly we see

(p) \[ E_\gamma X = v(\gamma|\beta_1+\beta_{k-1}) X E_{\gamma} \]
\[ + \sum_{b_2, \ldots, b_{k-2} \in \mathbb{N}} \xi'(b_2, \ldots, b_{k-2}) E_{\beta_2}^{(b_2)} \cdots E_{\beta_{k-2}}^{(b_{k-2})}, \]
\[ E_{\beta_{k-1}} X = v(\beta_{k-1}|\beta_1+\beta_{k-1}) X E_{\beta_{k-1}} \]
\[ + \sum_{b_2, \ldots, b_{k-2} \in \mathbb{N}} \xi''(b_2, \ldots, b_{k-2}) E_{\beta_2}^{(b_2)} \cdots E_{\beta_{k-2}}^{(b_{k-2})}, \]
where \( \xi'(b_2, \ldots, b_{k-2}) \), \( \xi''(b_2, \ldots, b_{k-2}) \)  \in A. We shall simply write \( X', X'' \) for the second terms of the right hand sides of the above two identities respectively. Then \( E_\gamma X = \nu(\gamma|\beta_1 + \beta_{k-1}) X E_\gamma + X' \),  \( E_{\beta_{k-1}} X = \nu(\gamma|\beta_1 + \beta_{k-1}) X E_{\beta_{k-1}} + X'' \).

Using (p) and induction hypothesis repeatedly, we get

\[
\begin{align*}
(q) \quad E_\gamma E_{\beta_{k-1}} X &= E_\gamma(\nu(\beta_{k-1}|\beta_1 + \beta_{k-1}) X E_{\beta_{k-1}} + X'') \\
&= \nu(\beta_{k-1}+\gamma|\beta_1 + \beta_{k-1}) X E_\gamma E_{\beta_{k-1}} + \nu(\gamma|\beta_1 + \beta_{k-1}) X'' E_\gamma \\
&\quad + \sum_{c_2, \ldots, c_{k-1} \in \mathbb{N}} \eta'(c_2, \ldots, c_{k-1}) E_{\beta_2} \cdots E_{\beta_{k-1}},
\end{align*}
\]

\[
E_{\beta_{k-1}} E_\gamma X = E_{\beta_{k-1}}(\nu(\gamma|\beta_1 + \beta_{k-1}) X E_\gamma + X') \\
= \nu(\gamma+\beta_{k-1}|\beta_1 + \beta_{k-1}) X E_{\beta_{k-1}} E_\gamma + \nu(\gamma|\beta_1 + \beta_{k-1}) X'' E_\gamma \\
\quad + \sum_{c_2, \ldots, c_{k-1} \in \mathbb{N}} \eta''(c_2, \ldots, c_{k-1}) E_{\beta_2} \cdots E_{\beta_{k-1}},
\]

\[- \nu^{-1}(\gamma|\beta_{k-1}) E_{\beta_{k-1}} Y E_{\beta_{k-1}} + E_{\beta_{k-1}}^{(2)} Y = \sum_{c_2, \ldots, c_{k-1} \in \mathbb{N}} \eta'''(c_2, \ldots, c_{k-1}) E_{\beta_2} \cdots E_{\beta_{k-1}},
\]

where \( \eta'(c_2, \ldots, c_{k-1}), \eta''(c_2, \ldots, c_{k-1}), \eta'''(c_2, \ldots, c_{k-1}) \in A. \)

Moreover we have

\[(r) \quad \beta_1 = \gamma + 2\beta_{k-1} \text{ and } (\beta_{k-1}|\beta_1) = 2, \quad (\gamma|\beta_{k-1}) = (\alpha_{i_k}|\alpha_{i_{k-1}}) = -2. \]

Combining (l)-(r) we see

\[
(s) \quad E_{\beta_1} E_{\beta_k} - \nu(\beta_1|\beta_k) E_{\beta_1} E_{\beta_k} = \sum_{a_2, \ldots, a_{k-1} \in \mathbb{N}} \eta(a_2, \ldots, a_{k-1}) E_{\beta_2} \cdots E_{\beta_{k-1}},
\]

where \( \eta(a_2, \ldots, a_{k-1}) \in A. \)

Using induction on \( a, b, \) and using (s) and induction hypothesis repeatedly, we see

\[
(t) \quad E_{\beta_k}^{(a)} E_{\beta_1}^{(b)} - \nu^{ab}(\beta_1|\beta_k) E_{\beta_1}^{(b)} E_{\beta_k}^{(a)} = \sum_{a_1, \ldots, a_{k} \in \mathbb{N}} \rho(a_1, \ldots, a_k) E_{\beta_1}^{(a_1)} \cdots E_{\beta_k}^{(a_k)},
\]

where \( \rho(a_1, \ldots, a_k) \in A \) (here we need 2.1 (a)). Thus in case (3) the identity (*) is true.

Now assume that we are in case (4), then
As a similar argument for case (3) we see that the identity (*) is true in this case.

Thus we proved the theorem for Case A.

**Case B.** $u = s_{j_1}s_{j_2} \cdots s_{j_m}$, $u' = s_{j_{m+1}} \cdots s_{j_{k-1}}$, and $j_1 = i_1$, $j_{k-1} = i_{k-1}$.

Define $\gamma_p = s_{j_1} \cdots s_{j_{p-1}} (\alpha_{j_p})$ $(2 \leq p \leq k - 1)$ and $E_{\gamma_p}^{(a)} = T_{j_1} \cdots T_{j_{p-1}} (E_{j_p}^{(a)})$ $(a \in \mathbb{N})$.

According to the arguments in Case A we get

\[
E_{\beta_k}^{(a)} = T_u T_{i_k} T_{i_{k-1}} (E_{i_k}) = T_u \left( E_{i_{k-1}}^{(a)} E_{i_k} - v^{-1} E_{i_{k-1}} E_{i_k} E_{i_{k-1}} + v^{-2} E_{i_k} E_{i_{k-1}}^{(a)} \right) = E_{\beta_{k-2}}^{(a)} - v^{-1} E_{\alpha} E_{\beta_{k-2}} E_{\alpha} + v^{-2} E_{\beta_{k-2}} E_{\alpha}^{(a)} \]

\[
(E_{\alpha}^{(a)} = T_u \left( E_{i_{k-1}}^{(a)} \right) ).
\]

Thus we proved the theorem for Case A.

**Case C.** $u = s_{j_1}s_{j_2} \cdots s_{j_m}$, $u' = s_{j_{m+1}} \cdots s_{j_{k-1}}$, and $j_1 \neq i_1$, $j_{k-1} = i_{k-1}$.

In this case $uu'$ has a reduced expression of the form $s_{p_1}s_{p_2} \cdots s_{p_{k-1}}$ such that $p_1 = i_1$, $p_{k-1} = i_{k-1}$, and one of the following three cases happens.

1. $\langle \alpha_{p_1}, \alpha_{p_2}^\vee \rangle = 0,$
2. $p_1 = p_3$ and $\langle \alpha_{p_1}, \alpha_{p_2}^\vee \rangle \langle \alpha_{p_2}, \alpha_{p_1}^\vee \rangle = 1,$
3. $p_1 = p_3$, $p_2 = p_4$, and $\langle \alpha_{p_1}, \alpha_{p_2}^\vee \rangle \langle \alpha_{p_2}, \alpha_{p_1}^\vee \rangle = 2.$
Define $p_k = i_k$. We set, for case (5),

$$
\gamma_1 = \alpha_{p_1}, \quad \gamma_3 = s_{p_1}(\alpha_{p_3}), \quad \gamma_h = s_{p_1} s_{p_3} \cdots s_{p_{h-1}}(\alpha_{p_h}) \quad (4 \leq h \leq k),
$$

$$
E^{(a)}_{\gamma_1} = E^{(a)}_{p_1}, \quad E^{(a)}_{\gamma_3} = T_{p_1}(E^{(a)}_{p_3}),
$$

$$
E^{(a)}_{\gamma_h} = T_{p_1} T_{p_3} \cdots T_{p_{h-1}}(E^{(a)}_{p_h}) \quad (4 \leq h \leq k), \quad a \in \mathbb{N};
$$

for case (6),

$$
\gamma_1 = \alpha_{p_2}, \quad \gamma_4 = s_{p_2}(\alpha_{p_4}), \quad \gamma_h = s_{p_2} s_{p_4} \cdots s_{p_{h-1}}(\alpha_{p_h}) \quad (5 \leq h \leq k),
$$

$$
E^{(a)}_{\gamma_1} = E^{(a)}_{p_2}, \quad E^{(a)}_{\gamma_4} = T_{p_2}(E^{(a)}_{p_4}),
$$

$$
E^{(a)}_{\gamma_h} = T_{p_2} T_{p_4} \cdots T_{p_{h-1}}(E^{(a)}_{p_h}) \quad (5 \leq h \leq k), \quad a \in \mathbb{N};
$$

for case (7),

$$
\gamma_1 = \alpha_{p_1}, \quad \gamma_5 = s_{p_1}(\alpha_{p_5}), \quad \gamma_h = s_{p_1} s_{p_3} \cdots s_{p_{h-1}}(\alpha_{p_h}) \quad (6 \leq h \leq k),
$$

$$
E^{(a)}_{\gamma_1} = E^{(a)}_{p_1}, \quad E^{(a)}_{\gamma_5} = T_{p_1}(E^{(a)}_{p_5}),
$$

$$
E^{(a)}_{\gamma_h} = T_{p_1} T_{p_5} \cdots T_{p_{h-1}}(E^{(a)}_{p_h}) \quad (6 \leq h \leq k), \quad a \in \mathbb{N}.
$$

By induction hypothesis we get:

(x1) For case (5), (since $s_{p_1} s_{p_3} \cdots s_{p_k}$ is a reduced expression),

$$
E^{(a)}_{\gamma_k} E^{(b)}_{\gamma_1} - v^{ab(\gamma_1 \gamma_k)} E^{(b)}_{\gamma_1} E^{(a)}_{\gamma_k} = \sum_{a_1, a_3, \ldots, a_k \in \mathbb{N}} \rho'(a_1, a_3, \ldots, a_k) E^{(a_1)}_{\gamma_1} E^{(a_3)}_{\gamma_3} \cdots E^{(a_k)}_{\gamma_k},
$$

where $\rho'(a_1, a_3, \ldots, a_k) \in A$.

(x2) For case (6), (since $s_{p_1} s_{p_4} \cdots s_{p_k}$ is a reduced expression),

$$
E^{(a)}_{\gamma_k} E^{(b)}_{\gamma_1} - v^{ab(\gamma_1 \gamma_k)} E^{(b)}_{\gamma_1} E^{(a)}_{\gamma_k} = \sum_{a_1, a_4, \ldots, a_k \in \mathbb{N}} \rho'(a_1, a_4, \ldots, a_k) E^{(a_1)}_{\gamma_1} E^{(a_4)}_{\gamma_4} \cdots E^{(a_k)}_{\gamma_k},
$$

where $\rho'(a_1, a_4, \ldots, a_k) \in A$.

(x3) For case (7), (since $s_{p_1} s_{p_5} \cdots s_{p_k}$ is a reduced expression),

$$
E^{(a)}_{\gamma_k} E^{(b)}_{\gamma_1} - v^{ab(\gamma_1 \gamma_k)} E^{(b)}_{\gamma_1} E^{(a)}_{\gamma_k} = \sum_{a_1, a_5, \ldots, a_k \in \mathbb{N}} \rho'(a_1, a_5, \ldots, a_k) E^{(a_1)}_{\gamma_1} E^{(a_5)}_{\gamma_5} \cdots E^{(a_k)}_{\gamma_k},
$$
where $\rho'(a_1, a_5, \ldots, a_k) \in A$.

Note that we always have $(\gamma_1 | \gamma_k) = (\beta_1 | \beta_k)$ since $(\cdot | \cdot)$ is $W$-invariant and $p_1 = i_1$. Recall that $T_{p_1}(E_i) = E_j$ if $w(\alpha_i) = \alpha_j$ (see [L5]). Applying $T_{p_2}$ (resp. $T_{p_2}T_{p_1}$; $T_{p_2}T_{p_1}T_{p_2}$) to the identity in (x1) (resp. (x2); (x3)) and using 2.1 (b) (see the argument for Case B) we get

\[(y) \ E_{\beta_k}^{(a)} E_{\beta_1}^{(b)} - v^{\alpha(\beta_1 | \beta_k)} E_{\beta_1}^{(b)} E_{\beta_k}^{(a)} = \sum_{a_1 < a_k, a_1 \in N} \rho(a_1, \ldots, a_k) E_{\beta_1}^{(a_1)} \cdots E_{\beta_k}^{(a_k)}, \]

where $\rho(a_1, \ldots, a_k) \in A$.

Thus the identity $(\ast)$ is true for Case C.

The theorem is proved. □

**Theorem 2.4.** Keep the notation in 2.1. Then:

(i) The elements

\[
\prod_{\beta \in R^+} E_{\beta}^{(a_\beta)} \quad (a_\beta \in \mathbb{N})
\]

form an $A$-basis of $U_+^A$. Where the factors in the product are written in a given total order on $R^+$.

(ii) The elements

\[
\prod_{\beta \in R^+} F_{\beta}^{(a_\beta)} \quad (a_\beta \in \mathbb{N})
\]

form an $A$-basis of $U_-^A$. Where the factors in the product are written in a given total order on $R^+$.

**Proof.** We only need to prove (i) since $\Omega(E_{\beta}^{(a_\beta)}) = F_{\beta}^{(a_\beta)}$ and $\Omega U_+^A = U_-^A$.

Define the lexicographical order $>_{\mathbb{N}}$ on $\mathbb{N}^{\mathbb{N}}$ such that \((1, 0, \ldots, 0) > (0, 1, \ldots, 0) > \cdots > (0, \ldots, 0, 1)\). Using Theorem 2.3 repeatedly we see

\[
\prod_{\beta \in R^+} E_{\beta}^{(a_\beta)} = v^p E_1^A + \sum_{B \in \mathbb{N}^{\mathbb{N}}} \rho_B E_1^B, \quad \rho_B \in A,
\]

where $p \in \mathbb{Z}$ and $A = (a_{\beta_1}, a_{\beta_2}, \ldots, a_{\beta_\nu})$ (we define $\beta_j = s_{i_1} \cdots s_{i_{j-1}}(\alpha_{i_j})$ and $\nu = |R^+|$). Noting that $\rho_B = 0$ if $\sum_{i=1}^{\nu} b_i \beta_i \neq \sum_{\beta \in R^+} a_\beta \beta$ (here $B = (b_1, \ldots, b_\nu)$), we see $E_1^A$ is an $A$-linear combination of the elements

\[\prod_{\beta \in R^+} E_{\beta}^{(c_\beta)} \quad (c_\beta \in \mathbb{N}).\]

Since for any $\lambda$ in $\mathbb{N} R^+$, the number

\[\#\{E_1^A \mid A = (a_1, \ldots, a_\nu) \in \mathbb{N}^\nu \text{ such that } a_1 \beta_1 + \cdots + a_\nu \beta_\nu = \lambda\}\]
is equal to
\[ \# \left\{ \prod_{\beta \in \mathbb{R}^+} E^{(a_\beta)}_\beta \mid \sum_{\beta \in \mathbb{R}^+} a_\beta \beta = \lambda \right\} , \]
by 2.1 (a), the elements
\[ \prod_{\beta \in \mathbb{R}^+} E^{(a_\beta)}_\beta \quad (a_\beta \in \mathbb{N}) \]
form an \( \mathcal{A} \)-basis of \( U^+_\mathcal{A} \).

The theorem is proved.

From the above proof we see the following:

**Corollary 2.5.** Keep the notation in Theorem 2.4 and its proof. If \( A = (a_{\beta_1}, \ldots, a_{\beta_\nu}) \) is minimal in the set\]
\[ \left\{ (b_1, \ldots, b_\nu) \in \mathbb{N}^\nu \mid \sum_{i=1}^\nu b_\beta = \sum_{\beta \in \mathbb{R}^+} a_\beta \beta \right\} , \]
then \( \prod_{\beta \in \mathbb{R}^+} E^{(a_\beta)}_\beta = v^q E^A_1 \). That is, for all \( \beta, \gamma \in \mathbb{R}^+ \) we have \( E^{(a_\beta)}_\beta E^{(a_\gamma)}_\gamma = v^q E^{(a_\gamma)}_\beta E^{(a_\beta)}_\beta \) for some \( q \in \mathbb{Z} \). (Of course, many \( a_\beta \) are 0 in this case.)

**2.6.** We would like to state two conjectures, one describes the root vectors intrinsically. The conjectures might be helpful for constructing an \( \mathcal{A} \)-basis of the \( \mathcal{A} \)-form of the quantized enveloping algebra of a symmetrizable Kac-Moody algebra. For \( \lambda \in \mathbb{N} \mathbb{R}^+ \), we denote by \( U^+_\lambda \) the set \( \{ x \in U^+ \mid K_i x K_i^{-1} = v(\lambda | \alpha_i) x \} \) and let \( U^+_\mathcal{A},\lambda = U^+_\lambda \cap U_\mathcal{A} \). We also write \( U^-_\lambda \) for \( \Omega(U^+_\lambda) \).

**Conjecture A.** Let \( \alpha \in \mathbb{R}^+ \) and set \( d_\alpha = d_i \) if \( w(\alpha_i) = \alpha \) for some \( w \in W \).

Let \( E \in U^+_{\mathcal{A},\alpha} \). If \( E^{(a)} = E^a /[a]_{d_\alpha}^1 \in U^+_\mathcal{A} \) for all \( a \in \mathbb{N} \), then there exists a simple root \( \alpha_j \) and \( u \in W, \ f \in \mathcal{A} \), such that \( u(\alpha_j) = \alpha \) and \( E = f T_u(\alpha_j) \).

(We refer to [L2] for the definition of \([a]_{d_\alpha}^1\).) For type \( A_2 \), the conjecture is true.

**Conjecture B.** For any \( \beta \in \mathbb{R}^+ \), choose \( w_\beta \in W \) and \( i_\beta \in [1, n] \) such that \( w_\beta(\alpha_{i_\beta}) = \beta \). Define \( E^{(a)}_\beta = T_{w_\beta}(E^{(a)}_i) \). Then the elements
\[ \prod_{\beta \in \mathbb{R}^+} E^{(a_\beta)}_\beta \quad (a_\beta \in \mathbb{N}) \]
form an \( \mathcal{A} \)-basis of \( U^+_\mathcal{A} \). Where the factors in the product are written according to a given total order on \( \mathbb{R}^+ \).
3. Some orthogonal bases of the bilinear form in [K].

In this section we show that certain PBW bases are orthogonal bases of the bilinear from considered in [K], see Theorem 3.9. For the bilinear form obtained from the Drinfeld dual, a similar result was established in [L5, LS]. Although the difference between the two bilinear forms are small, it is difficult to apply the methods in [L5, LS] for proving Theorem 3.9, since the methods rely on a property ([L5, 38.2.1]) which does not hold for the bilinear form in [K].

3.1. Following Kashiwara [K, Prop. 3.4.4] we define a bilinear form on $U^+$. (a) For each $P \in U^+$ and $F_i$, there exist unique $P', P'' \in U^+$ such that

$$PF_i - F_iP = \frac{K_i P' - K_i^{-1} P''}{v_i - v_i^{-1}}.$$ 

(We set $v_i = v^{d_i}$.)

Define $\varphi_i(P) = P''$ and $\psi_i(P) = P'$. We have (cf. [K, Prop. 3.4.4]).

(b) There is a unique symmetric bilinear form $(\ , \ )$ on $U^+$ such that $(1, 1) = 1$,

$$(E_i x, y) = (x, \varphi_i(y)) \quad \text{for all } i \in [1, n] \text{ and } x, y \in U^+,$$

$$(x, E_i y) = (\varphi_i(x), y) \quad \text{for all } i \in [1, n] \text{ and } x, y \in U^+.$$ 

We need some preparation for proving Theorem 3.9. Let $\mathcal{X}$ be the set of all sequences $i = (i_1, \ldots, i_\nu)$ in $[1, n]$ such that $s_{i_1} \cdots s_{i_\nu}$ is a reduced expression of the longest element $w_0 \in W$. For $i = (i_1, \ldots, i_\nu) \in \mathcal{X}$, $A = (a_1, \ldots, a_\nu) \in \mathbb{N}^\nu$, we shall write

$$E_i^A = E_{i_1}^{(a_1)} T_{i_1} (E_{i_2}^{(a_2)}) \cdots T_{i_1} \cdots T_{i_\nu-1} (E_{i_\nu}^{(a_\nu)}),$$

$$F_i^A = T_{i_1} \cdots T_{i_\nu-1} (E_{i_\nu}^{(a_\nu)}) \cdots T_{i_1} (E_{i_2}^{(a_2)}) F_{i_1}^{(a_1)} = \Omega(E_i^A),$$

$$\hat{E}_i^A = E_{i_1}^{(a_1)} T_{i_1}^{-1} (E_{i_2}^{(a_2)}) \cdots T_{i_1}^{-1} \cdots T_{i_\nu-1}^{-1} (E_{i_\nu}^{(a_\nu)}).$$

For $\lambda = a_1 \alpha_1 + a_2 \alpha_2 + \cdots + a_\nu \alpha_\nu \in \mathbb{Z}R$, we define $K_\lambda = K^{a_1} K^{a_2} \cdots K^{a_\nu}$.

The following result plays a key role in the proof, which is essentially a variation of Theorem 2.3.

**Lemma 3.2.** Let $i = (i_1, \ldots, i_\nu) \in \mathcal{X}$ and let $s, k \in [0, \nu - 1]$ such that $s > k$. Set $E = T_{i_1} \cdots T_{i_k} (E_{i_{k+1}})$, $F = T_{i_1} \cdots T_{i_k} (F_{i_{k+1}})$. We have

$$EF - FE = \sum \sigma(A, \lambda, B) F_i^A K_\lambda E_i^B, \quad \sigma(A, \lambda, B) \in A,$$
where $A = (a_1, \ldots, a_\nu)$ and $B = (b_1, \ldots, b_\nu)$ run through a finite subset of $\mathbb{N}^\nu$, $\lambda$ runs through a finite subset of $\mathbb{N}R^+$, and $a_{k+1} = \cdots = a_\nu = 0$, $b_1 = \cdots = b_{s+1} = 0$ if $\sigma(A, \lambda, B) \neq 0$.

**Proof.** Set $j_1 = i_s$, $j_2 = i_{s-1}$, \ldots, $j_s = i_1$. Choose $j_{s+1}$, \ldots, $j_\nu \in [1, n]$ such that $(j_1, \ldots, j_\nu) \in X$ and $s_{j_1} \cdots s_{j_{\nu-1}}(\alpha_{j_\nu}) = \alpha_{i_{s+1}}$.

For $m \in [1, \nu]$, $\alpha \in \mathbb{N}$, define

$$X_m^{(a)} = T_{j_1} \cdots T_{j_{m-1}}(E_{j_{m-1}}^{(a)}),$$

$$X_m^{(a)} = \Psi(X_m^{(a)}) = T_{j_1} \cdots T_{j_{m-1}}(E_{j_{m-1}}^{(a)}).$$

Then $T_{i_s} \cdots T_{i_{s+2}}(E_{i_{k+1}}) = X_{s-k}$ and $E_{i_{s+1}} = X_\nu$.

Set $\beta = s_{i_s} \cdots s_{i_{s+2}}(\alpha_{i_{s+1}})$, $\beta' = \alpha_{i_{s+1}}$. Using Theorem 2.3 repeatedly we see

$$E_{i_{s+1}}X_{s-k} - v^{(\beta|\beta')}X_{s-k}E_{i_{s+1}}$$

$$= \sum \sigma(a_{\nu-1}, \ldots, a_{s-k+1})X_{\nu}^{(a_{\nu-1})} \cdots X_{s-k+1}^{(a_{s-k+1})},$$

where $\sigma(a_{\nu-1}, \ldots, a_{s-k+1}) \in A$, and $a_{\nu-1}, \ldots, a_{s-k+1}$ run through a finite subset of $\mathbb{N}$.

Applying $\Psi$ to the identity (a) we get

$$X_{s-k}E_{i_{s+1}} - v^{(\beta'|\beta')}E_{i_{s+1}}X_{s-k}$$

$$= \sum \sigma(a_{\nu-1}, \ldots, a_{s-k+1})X_{s-k+1}^{(a_{s-k+1})} \cdots X_{\nu}^{(a_{\nu-1})}. $$

If $\nu > m \geq s + 1$, then we may find $k_{m+1}, \ldots, k_\nu \in [1, n]$ such that $(k_{\nu}, \ldots, k_{m+1}, j_1, \ldots, j_m) \in X$. Noting that $s_{k_{\nu}} \cdots s_{k_{m+1}}s_{j_1} \cdots s_{j_{m-1}}(\alpha_{j_m})$ is a simple root $\alpha_j$ for some $j \in [1, n]$, we see

$$T_{k_{\nu}} \cdots T_{k_{m+1}}T_{j_1} \cdots T_{j_{m-1}}(E_{j_m}) = E_j.$$

Since $s_{j_1} \cdots s_{j_{m-1}}(\alpha_j) = s_{j_1} \cdots s_{j_1}s_{k_{m+1}} \cdots s_{k_{\nu}}(\alpha_j) \in R^+$, we have

$$Y_m = T_{j_{s+1}} \cdots T_{j_{m-1}}(E_{j_m}) = T_{j_1} \cdots T_{j_1}T_{k_{m+1}} \cdots T_{k_{\nu}}(E_j)$$

$$= T_{i_s} \cdots T_{i_s}T_{k_{m+1}} \cdots T_{k_{\nu}}(E_j) \in U^+, $$

for $s + 1 \leq m \leq \nu - 1$.

By our choice on $j_1, \ldots, j_\nu$ we may require that $k_{m+1} = i_{s+1}$. By (d) and 2.1 (b) we see

$$Y_m = T_{j_{s+1}} \cdots T_{j_{m-1}}(E_{j_m}) = \sum \sigma(A)E_{i_s}^A, \quad \sigma(A) \in A,$$
where $A = (a_1, \ldots, a_\nu)$ runs through a finite subset of $\mathbb{N}^\nu$ and $a_1 = \cdots = a_{s+1} = 0$ if $\sigma'(A) \neq 0$.

Define $\beta_m = s_{i_1} \cdots s_{i_{m-1}}(\alpha_{i_m})$ for $1 \leq m \leq k + 1$, and set $Z_m^{(a)} = T_i \cdots T_{i_{m-1}}(F_i^{(a)})$ for $1 \leq m \leq k + 1$, $a \in \mathbb{N}$. For $m \geq s + 1$, $a \in \mathbb{N}$, we set $Y_m^{(a)} = T_{j_{s+1}}^{-1} \cdots T_{j_{m-1}}^{-1}(E_{j_m}^{(a)}) \in U^+$.

Applying $T_i \cdots T_{i_s}$ to the identity (b) we get

\[(f) \quad - K_{\beta_{k+1}}^{-1} F E + v^{(\beta|\beta')} EK_{\beta_{k+1}}^{-1} F \]
\[= \sum \sigma(a_{\nu-1}, \ldots, a_{s-k+1})v^c K_{\gamma} Z_k^{(a_{s-k+1})} \cdots Z_1^{(a_s)} Y_{s+1}^{(a_1)} \cdots Y_{\nu-1}^{(a_{\nu-1})}, \]

where $c$ is a suitable integer depending on $i_1, \ldots, i_k, a_{s-k+1}, \ldots, a_s$, and $\lambda = a_{s-k+1}\beta_k + a_{s-k+2}\beta_{k-1} + \cdots + a_s\beta_1 \in \mathbb{N}R^+$.

Since $(\beta_{k+1}|s_{i_1} \cdots s_{i_s}(\alpha_{i_{s+1}})) = -(s_{i_1} \cdots s_{i_{k+2}}(\alpha_{i_{k+1}})(\alpha_{i_{s+1}}) = -(\beta|\beta')$, we see

\[(g) \quad EK_{\beta_{k+1}}^{-1} = v^{-\gamma(\beta|\beta')} K_{\beta_{k+1}}^{-1} E. \]

Obviously we have

\[(h) \quad \beta_{k+1} \geq \lambda \quad \text{if} \quad \sigma(a_{\nu-1}, \ldots, a_{s-k+1}) \neq 0. \]

Combining (e)-(h) and Theorem 2.3 we see the lemma is true.

### 3.3.

Let $\beta = \sum a_i\alpha_i \in \mathbb{N}R^+$. We define $\sigma(\beta) = \prod_i (v_i - v_i^{-1}) a_i$, $d_j^\beta = \sum a_i d_i - d_j$ if $\beta \in R^+$. Let $a$ be an integer and $b, d$ positive integers, set

\[\{a\}_d = \frac{1 - v^{-2ad}}{1 - v^{-2d}}, \quad \{b\}_d = \{b\}_d \{b - 1\}_d \cdots \{1\}_d, \quad \{0\}_d = 1, \]
\[\{-b\}_d^\gamma := (-1)^b \{b\}_d^\gamma, \quad \{a\}_d := \prod_{h=1}^{a+b} \frac{1 - v^{-2(a-h+1)d}}{1 - v^{-2hd}}, \quad \{a\}_d := 1. \]

We have

\[(a) \quad \left\{ \begin{array}{l} a + b \\ b \end{array} \right\}_d = \frac{\{a + b\}_d^\gamma}{\{a\}_d^\gamma (b)_d^\gamma} \quad \text{for} \quad a, b \in \mathbb{N}. \]

We shall omit the subscript $d$ if $d = 1$.

\[(b) \quad \sigma(\beta + \gamma) = \sigma(\beta) \sigma(\gamma) \quad \text{for} \quad \beta, \gamma \in \mathbb{N}R^+. \]
Recall for \( \lambda = a_1\alpha_1 + \cdots + a_n\alpha_n \), we write \( K\lambda \) for \( K^{a_1}_\lambda \cdots K^{a_n}_\lambda \). For \( \beta \in R^+ \), we shall write \( \Sigma_\beta \) (resp. \( \Sigma'_\beta \)) for any element of \( U \) of the form
\[
\sum uK\lambda x \quad \text{(resp.} \sum uK\lambda x) ,
\]
where \( u \) runs through a finite subset of \( U^- \), \( x \) runs through a finite subset of \( U^+ \), and \( \lambda \) runs through the set \( \{ \sum b_i\alpha_i \in \mathbb{Z}R \mid |b_i| \leq a_i \text{ for all } i \text{ and } \sum b_i\alpha_i \neq \pm \beta \} \) (resp. \( \{ \sum b_i\alpha_i \in \mathbb{Z}R \mid |b_i| \leq a_i \text{ for all } i \} \)). The following assertions (c) and (d) are obvious.

(c) \( \Sigma_\beta + \Sigma_\beta = \Sigma_\beta \) and \( \Sigma'_\beta + \Sigma'_\beta = \Sigma'_\beta \) for \( \beta \in NR^+ \)

(d) \( \Sigma_\beta \Sigma_\gamma = \Sigma_{\beta+\gamma}, \quad \Sigma'_\beta \Sigma_\gamma = \Sigma_{\beta+\gamma}, \quad \Sigma_\beta \Sigma'_\gamma = \Sigma_{\beta+\gamma} \), for \( \beta, \gamma \in NR^+ \).

**Lemma 3.4.** Let \( \beta \in NR^+ \) and let \( u \in U^-_\beta \) be a monomial of \( F_1, \ldots, F_n \). Then for any \( x \in U^+ \), there exist unique \( x_1, x_2 \in U^+ \) such that
\[
xu - ux = \frac{K_\beta x_1 + ( -1)^{ht(\beta)} K^{-1}_\beta x_2}{\sigma(\beta)} + \Sigma_\beta.
\]

**Proof.** We use induction on \( \text{ht}(\beta) \). When \( \text{ht}(\beta) = 0, 1 \), the lemma is just 3.1 (a). Assume that \( \text{ht}(\beta) \geq 2 \) and \( u = F_iu' \). By induction hypothesis we get
\[
\begin{align*}
xu - ux &= (xF_i - F_i x)u' + F_i(xu' - u'x) \\
&= \frac{K_iy_1 - K_i^{-1}y_2}{\sigma(\alpha_i)} u' + F_i \left( \frac{K_{\beta-\alpha_1}x_1' + ( -1)^{\text{ht}(\beta-\alpha_1)} K^{-1}_{\beta-\alpha_1} x_2'}{\sigma(\beta - \alpha_i)} + \Sigma_{\beta-\alpha_1} \right) \\
&= \frac{1}{\sigma(\alpha_i)} \left\{ K_i \left( u'y_1 + \frac{K_{\beta-\alpha_1}z_1 + ( -1)^{\text{ht}(\beta-\alpha_1)} K^{-1}_{\beta-\alpha_1} z_2}{\sigma(\beta - \alpha_1)} + \Sigma_{\beta-\alpha_1} \right) \\
&\quad - K_i^{-1} \left(u'y_2 + \frac{K_{\beta-\alpha_1}z_1' + ( -1)^{\text{ht}(\beta-\alpha_1)} K^{-1}_{\beta-\alpha_1} z_2'}{\sigma(\beta - \alpha_1)} + \Sigma_{\beta-\alpha_1} \right) \right\} + \Sigma_\beta,
\end{align*}
\]
where \( y_1, y_2, x_1', x_2', z_1, z_2, z_1', z_2' \) are elements of \( U^+ \).

We have \( K_i K^{-1}_{\beta-\alpha_1} = \Sigma_\beta, \ K_i^{-1} K_{\beta-\alpha_1} = \Sigma_\beta, \) and \( K_i \Sigma_{\beta-\alpha_1} = \Sigma_\beta, \ K_i^{-1} \Sigma_{\beta-\alpha_1} = \Sigma_\beta \) (cf. 3.3 (d)). By (b), (c) and (d) in 3.3 we get
\[
xu - ux = \frac{K_\beta z_1 + ( -1)^{\text{ht}(\beta)} K^{-1}_\beta z_2'}{\sigma(\beta)} + \Sigma_\beta.
\]

The uniqueness of \( x_1 = z_1, x_2 = z_2' \) follows from PBW theorem (see [L2]). The lemma is proved.
Proposition 3.5. Let $\beta, \gamma \in \mathbb{NR}^+$ such that $\beta - \gamma \in \mathbb{NR}^+$, and let $x \in U_\beta^+$, $y \in U_\gamma^+$, $z \in U_{\beta - \gamma}^+$. Let $\xi_1, \xi_2 \in U^+$ be such that (see Lemma 3.4)

$$x\Omega'(y) - \Omega'(y)x = \frac{K_{\gamma}\xi_1 + (-1)^{ht(\gamma)}K_{\gamma}^{-1}\xi_2}{\sigma(\gamma)} + \Sigma_{\gamma}.$$  

(See 1.2 for the definition of $\Omega'$.) Then $(x, yz) = (\xi_2, z)$. In particular, if $\beta = \gamma$ and $z = 1$, then $(x, y) = \xi_2$.

Proof. We may assume that $y$ is a monomial $E_{i_1}\cdots E_{i_k}$. Repeatedly using the properties in the definition of the bilinear form we get the proposition.

Corollary 3.6. Let $\beta \in R^+$ and $F$ a root vector corresponding to $-\beta$. Then for any $x \in U^+$ there exist unique $x_1, x_2 \in U^+$ such that

$$xF - Fx = \frac{K_{\beta}x_1 + (-1)^{ht(\beta)}K_{\beta}^{-1}x_2}{\sigma(\beta)} + \Sigma_{\beta}.$$  

We shall write $\varphi_F(x) = x_2$ and $\psi_F(x) = x_1$.

Proof. Since $F$ is a $\mathbb{Q}(v)$-linear combination of monomials of $F_1, \ldots, F_n$ with degree $-\beta$, the corollary follows from Lemma 3.4.

Proposition 3.7. Let $\beta \in R^+$ and $F$ a root vector corresponding to $-\beta$. Then for any $x, y \in U^+$ we have (see 3.3 for the definition of $d''_{\beta}$)

$$(x, Ey) = (-1)^{ht(\beta) - 1}v^{-d''_{\beta}}(\varphi_F(x), y),$$

where $E = \Omega(F) \in U^+$.

Proof. Let $s_{i_1}\cdots s_{i_k}$ be a reduced expression of $w \in W$ such that

$$F = T_{i_1}\cdots T_{i_{k-1}}(F_{i_k}).$$

We use induction on $k = l(w)$ to prove the proposition. When $k = 1$, then $F = F_{i_1}$, the proposition is just a property of the bilinear form $(\ , \ )$ since $d''_{\beta} = 0$ in this case. Assume the proposition is true when $l(w) \leq k - 1$.

Now assume that $k = l(w) \geq 2$. Let $u$ be the shortest element of the coset $w(s_{i_{k-1}}s_{i_k})$, then $w = uu'$ for some $u' \in \langle s_{i_{k-1}}, s_{i_k} \rangle$ and $l(w) = l(u) + l(u')$. Moreover $l(us_{i_{k-1}}) = l(us_{i_k}) = l(u) + 1 \leq k - 1$. If $u'(\alpha_{i_k})$ is a simple root $\alpha_j$, then $j = i_k$ or $i_{k-1}$ and $F = T_u(F_j)$. By induction hypothesis, the proposition is true in this case.

Suppose that $\gamma = u'(\alpha_{i_k})$ is not a simple root, then we have the following cases.

1. $u' = s_{i_{k-1}}$ and $\gamma = \alpha_{i_{k-1}} + \alpha_{i_k}$,
2. $u' = s_{i_{k-1}}$ and $\gamma = 2\alpha_{i_{k-1}} + \alpha_{i_k}$,
3. $u' = s_{i_{k-1}}$ and $\gamma = 3\alpha_{i_{k-1}} + \alpha_{i_k}$,
4. $u' = s_{i_{k-1}}s_{i_{k-1}}$ and $\gamma = \alpha_{i_{k-1}} + \alpha_{i_k}$. 

By induction hypothesis, we have

(5) $u' = s_{i_k} s_{i_{k-1}}$ and $\gamma = 2\alpha_{i_{k-1}} + \alpha_{i_k}$,

(6) $u' = s_{i_k} s_{i_{k-1}}$ and $\gamma = \alpha_{i_{k-1}} + 2\alpha_{i_k}$, (type $G_2$)

(7) $u' = s_{i_k} s_{i_{k-1}}$ and $\gamma = 3\alpha_{i_{k-1}} + 2\alpha_{i_k}$,

(8) $u' = s_{i_{k-1}} s_{i_k} s_{i_{k-1}}$ and $\gamma = \alpha_{i_{k-1}} + 2\alpha_{i_k}$,

(9) $u' = s_{i_k} s_{i_{k-1}} s_{i_k} s_{i_{k-1}}$ and $\gamma = 3\alpha_{i_{k-1}} + 2\alpha_{i_k}$,

(10) $u' = s_{i_k} s_{i_{k-1}} s_{i_k} s_{i_{k-1}}$ and $\gamma = \alpha_{i_{k-1}} + \alpha_{i_k}$,

(11) $u' = s_{i_k} s_{i_{k-1}} s_{i_k} s_{i_{k-1}}$ and $\gamma = 3\alpha_{i_{k-1}} + \alpha_{i_k}$.

Case (1). Let $\beta_1 = u(\alpha_{i_k})$, $\beta_2 = u(\alpha_{i_{k-1}})$, then $\beta_1$, $\beta_2 \in R^+$ and $\beta = \beta_1 + \beta_2$. We have

$$T_{i_{k-1}}(E_{i_k}) = E_{i_k} E_{i_{k-1}} - v^{-d} E_{i_{k-1}} E_{i_k}$$

where $d = -d_{i_{k-1}} a_{i_{k-1}, i_k} = d_{i_{k-1}}$.

Let $E' = T_u(E_{i_k})$, $E'' = T_u(E_{i_{k-1}})$, $F' = \Omega(E') = T_u(F_{i_k})$, $F'' = \Omega(E'') = T_u(F_{i_{k-1}})$. Then $E = E'E'' - v^{-d} E'' E'$ and $F = F'' F' - v^d F' F''$.

By induction hypothesis, we have

(a) $(x, E'E'' y) = (-1)^{ht(\beta_1)-1 + ht(\beta_2)-1} v^{-d_{\beta_1} - d_{\beta_2}} (\varphi_{F''}(\varphi_{F'}(x)), y)$,

(b) $(x, E'' E' y) = (-1)^{ht(\beta_2)-1 + ht(\beta_1)-1} v^{-d_{\beta_1} - d_{\beta_2}} (\varphi_{F'}(\varphi_{F''}(x)), y)$.

Recall that we have

$$xF' - F' x = \frac{K_{\beta_1} \psi_{F'}(x) + (-1)^{ht(\beta_1)} K_{\beta_1}^{-1} \varphi_{F'}(x)}{\sigma(\beta_1)} + \Sigma_{\beta_1},$$

$$\varphi_{F'}(x) F'' - F'\varphi_{F'}(x) = \frac{K_{\beta_2} \psi_{F''}(\varphi_{F'}(x)) + (-1)^{ht(\beta_2)} K_{\beta_2}^{-1} \varphi_{F''}(\varphi_{F'}(x))}{\sigma(\beta_2)} + \Sigma_{\beta_2},$$

$$xF'' - F'' x = \frac{K_{\beta_2} \psi_{F''}(x) + (-1)^{ht(\beta_2)} K_{\beta_2}^{-1} \varphi_{F''}(x)}{\sigma(\beta_2)} + \Sigma_{\beta_2},$$

$$\varphi_{F''}(x) F' - F'\varphi_{F''}(x) = \frac{K_{\beta_1} \psi_{F'}(\varphi_{F''}(x)) + (-1)^{ht(\beta_1)} K_{\beta_1}^{-1} \varphi_{F'}(\varphi_{F''}(x))}{\sigma(\beta_1)} + \Sigma_{\beta_1}.$$
Theorem 3.9. Let
\[ E_i = \ldots \]
Using 3.3 (b)-3.3 (d) and Corollary 3.6 repeatedly, we get
\[ 196 \text{ NANHUA XI} \]
\[ \phi(x) \]
\[ \Omega(\phi) \]
\[ \phi(\psi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
\[ \phi(\phi) \]
Therefore we have
\[ \phi(x) = \phi(x) - \phi(x) \phi(x). \]
Since \( d_\beta = d_{i_k} = d_{\beta_1} \), so, \( d_\beta' = d_{\beta_1}' + d_{\beta_2}' \). Note that \( d_{\beta_2} = d_{i_{k-1}} = d \).
Hence
\[ (x, E_i y) = (x, E_i E_i y - v^{-d} E_i E_i y) \]
\[ = (-1)^{ht(\beta)} v^{-d_\beta} (\phi(x), y) \]
\[ = (-1)^{ht(\beta)} v^{-d_\beta} (\phi(x), y) \]
\[ = (-1)^{ht(\beta) - 1} v^{-d_\beta} (\phi(x), y) \]
We may deal with other cases similarly. The proposition is proved.

Corollary 3.8. Let \( i = (i_1, \ldots, i_{i_\nu}) \in \mathcal{X} \) and let \( F = T_{i_1} \cdots T_{i_{k-1}} (E_{i_k}), E_1 = \Omega(F), A = (a_1, \ldots, a_{i_\nu}) \in \mathbb{N}^{i_\nu}. \) Then
\[ \phi_F(E_i^A) = 0 \text{ if } a_1 = \cdots = a_k = 0, \]
\[ \phi_F(E_i^A) = (-1)^{ht(\beta) - 1} v_{i_k}^{a_k - 1} \frac{\sigma(\beta)}{\sigma(\alpha_{i_k})} E_i^{A'}, \text{ if } a_1 = \cdots = a_{k-1} = 0, \]
where \( A' = (0, \ldots, 0, a_k - 1, a_{k+1}, \ldots, a_{i_\nu}), \beta = s_{i_1} \cdots s_{i_{k-1}}(\alpha_{i_k}). \)

Proof. (a) follows from Lemma 3.2 and the definition of \( \phi_F \) (see Corollary 3.6). (b) follows from the definition of \( \phi_F \) and the following identity
\[ E^{(a)} = F E^{(a)} + K_\beta v_{i_k}^{1-a} - K_\beta v_{i_k}^{a-1} E^{(a-1)}, \text{ } a \in \mathbb{N}. \]

Theorem 3.9. Let \( i = (i_1, \ldots, i_{i_\nu}) \in \mathcal{X} \) and \( A = (a_1, \ldots, a_{i_\nu}), B = (b_1, \ldots, b_{i_\nu}) \) be elements of \( \mathbb{N}^{i_\nu}. \) Then
\[ (E_i^A, E_i^B) = 0 \text{ if } A \neq B. \]
(b) \((E_1^A, E_1^A) = \prod_{k=1}^{\nu} \frac{\xi(\beta_k)^{a_k}}{(a_k!)^d_i},\)

where \(\beta_k = s_{i_k} \cdots s_{i_{k-1}}(\alpha_i),\) \(d'_k = d_k\) and \(\xi(\beta_k) = \frac{\prod (1 - v_i^{-2})^{c_i}}{(1 - v_i^{-2})}\) if \(\beta_k = \sum_i c_i \alpha_i.\) In particular, \(E_i^A \in L_B\) (see 3.11 for definition).

Proof. Repeatedly use Prop. 3.7 and Corollary 3.8.

Corollary 3.10. Let \(x \in U^+.\) Then

(a) \((x, x) \neq 0\) if \(x \neq 0.\) In particular, \(( , )\) is non-degenerate \([K, Corollary 3.4.8]\).

(b) For any \(i \in X\) we have \(x = \sum_{A \in \mathbb{N}^\nu} \frac{(x, E_i^A)}{(E_i^A, E_i^A)} E_i^A.\)

3.11. Let \(B\) be the subring of \(\mathbb{Q}(v)\) consisting of all rational functions which are regular at \(v^{-1} = 0\) (i.e. \(v = \infty\)). Define \(L_B = \{x \in U^+ | (x, x) \in B\}.\) The \(B\)-submodule \(L_B\) of \(U^+\) is crucial for discussing canonical bases.

Corollary 3.12. For any \(i \in X\), the elements \(E_i^A (A \in \mathbb{N}^\nu)\) form a \(B\)-basis of \(L_B.\)

Proof. Let \(\xi \in \mathbb{Q}(v),\) then \(\xi \in B\) if and only if \(\xi^2 \in B.\) The corollary then follows from Theorem 3.9 and 2.1 (a).

Corollary 3.13. Let \(i, j \in X\) and let \(A \in \mathbb{N}^\nu.\) Write

\[E_i^A = \sum_{B \in \mathbb{N}^\nu} \xi_B E_j^B, \quad \xi_B \in A,\]

then there exists a unique \(B_0 \in \mathbb{N}^\nu\) such that \(\xi_{B_0} \in \pm 1 + v^{-1}\mathbb{Z}[v^{-1}],\) and \(\xi_B \in v^{-1}\mathbb{Z}[v^{-1}]\) if \(B \neq B_0\) (see \([L_3, Prop. 2.3]\)).

Proof. By Corollary 3.12 we see that \(\xi_B \in A \cap B = \mathbb{Z}[v^{-1}].\) By Theorem 3.9 (a) we know

\[(E_i^A, E_i^A) = \sum_{B \in \mathbb{N}^\nu} \xi_B^2 (E_j^B, E_j^B).\]

By Theorem 3.9 (b), the values of \((E_i^A, E_i^A), (E_j^B, E_j^B)\) at \(v^{-1} = 0\) are 1. So there is a unique \(B_0 \in \mathbb{N}^\nu\) such that \(\xi_{B_0}^2|_{v^{-1} = 1} = 1,\) and \(\xi_B^2|_{v^{-1} = 0} = 0\) if \(B \neq B_0.\)

The corollary is proved.

Corollary 3.14. (a) Let \(L\) be the \(\mathbb{Z}[v^{-1}]-\)submodule of \(L_B\) spanned by the elements \(E_i^A, i \in X, A \in \mathbb{N}^\nu.\) Then \(L\) is a free \(\mathbb{Z}[v^{-1}]-\)module and for any \(i \in X,\) the elements \(E_i^A (A \in \mathbb{N}^\nu)\) form a \(\mathbb{Z}[v^{-1}]-basis\) of \(L.\)
(b) Let $x \in U^+_A$. Then $x \in \mathcal{L}$ if and only if $x \in \mathcal{L}_B$, i.e. $(x,x) \in B$. See [L3].

3.15. The $\mathbb{Z}[v^{-1}]$-module $\mathcal{L}$ can be defined through Kashiwara’s operators $\tilde{e}_i, \tilde{f}_i : U^+ \to U^+$, which are defined as follows

\[ \tilde{e}_i : \sum_{A \in \mathbb{N}^\nu} \xi_A E_i^A \mapsto \sum_{A \in \mathbb{N}^\nu} \xi_A E_i^{A+A_1}, \quad \xi_A \in \mathbb{Q}(v), \]

\[ \tilde{f}_i : \sum_{A \in \mathbb{N}^\nu} \xi_A E_i^A \mapsto \sum_{A \in \mathbb{N}^\nu} \xi_A E_i^{A-A_1}, \quad \xi_A \in \mathbb{Q}(v), \]

where $i = (i_1, \ldots, i_n) \in \mathcal{X}$ such that $i_1 = i$ and $A_1 = (1,0,\cdots,0)$.

Obviously we have

(a) $\tilde{e}_i$ and $\tilde{f}_i$ map $\mathcal{L}$ to $\mathcal{L}$ for all $i \in [1,n]$,

(b) $\tilde{f}_i \tilde{e}_i = \text{id}$ for all $i \in [1,n]$,

(c) $U^+ = \text{ker} \tilde{f}_i \oplus \text{im} \tilde{e}_i$ for each $i$ in $[1,n]$,

(d) $\ker \tilde{f}_i = \ker \varphi_i$. In particular $\bigcap_{i=1}^n \ker \tilde{f}_i = \mathbb{Q}(v).1$ (cf. [K, Lemma 3.4.7]).

Proposition 3.16. Let $\mathcal{L}'$ be the $\mathbb{Z}[v^{-1}]$-submodule of $U^+$ generated by the elements $\tilde{e}_i \cdots \tilde{e}_{i_k}(1) \ (i_1, \cdots, i_k \in [1,n] \text{ and } k \in \mathbb{N})$. Then we have $\mathcal{L}' \subseteq \mathcal{L}$. (See [L4, Theorem 2.3 (a)].)

Proof. Using Corollary 3.13 and the definition of $\tilde{e}_i$ we see $\mathcal{L}' \subseteq \mathcal{L}$.

It is not difficult to prove that $\mathcal{L}' = \mathcal{L}$, see [L4] or [X3].

Acknowledgement. The author was supported by the National Natural Science Foundation of China (No. 19425003).

References


[LS] Levendorski-Soibelman, Some applications of quantum Weyl groups, J. Geom. and Physics, 7 (1990), 241-254.


Received September 26, 1997.

ACADEMIA SINICA
BEIJING 100080
CHINA
E-mail address: nanhua@math08.math.ac.cn

UNIVERSITY OF CALIFORNIA
RIVERSIDE, CA 92521
Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use $\LaTeX$, but papers in other varieties of $\TeX$, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as $\LaTeX$ sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of Bib$\LaTeX$ is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this.”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.
Maximal monotonicity of dense type, local maximal monotonicity, and monotonicity of the conjugate are all the same for continuous linear operators

Heinz H. Bauschke and Jonathan M. Borwein

On the boundedness of singular integrals

Eugene Fabes, Irina Mitrea and Marius Mitrea

Small extensions of Witt rings

Robert W. FitzGerald

Equivariant embeddings of Stein domains sitting inside of complex semigroups

Bernhard Krotz

Base change problems for generalized Walsh series and multivariate numerical integration

Gerhard Larcher and Gottlieb Pirsic

A note on the moving sphere method

Yuki Naito and Takashi Suzuki

Pseudodifferential operators on differential groupoids

Victor Nistor, Alan Weinstein and Ping Xu

Restrictions of rank-2 semistable vector bundles on surfaces in positive characteristic

Atsushi Noma

Trace formula for a system of particles with elliptic potential

K.L. Vaninsky

A commutation formula for root vectors in quantized enveloping algebras

Nanhua Xi