FUNDAMENTAL SOLUTIONS OF INVARIANT DIFFERENTIAL OPERATORS ON A SEMISIMPLE LIE GROUP II

GUILLERMO AMES

Let $G$ be a linear connected semisimple Lie group. We denote by $\mathcal{U}(g)^K$ the algebra of left invariant differential operators on $G$ that are also right invariant by $K$, and $Z(\mathcal{U}(g)^K)$ denotes center of $\mathcal{U}(g)^K$.

In this paper we give a sufficient condition for a differential operator $P \in Z(\mathcal{U}(g)^K)$ to have a fundamental solution on $G$. This result extends the same one obtained previously for real rank one Lie groups and groups with only one conjugacy class of Cartan subgroups.

1. Introduction.

Let $G$ be a linear connected semisimple Lie group. The algebra of left invariant differential operators on $G$ is canonically identified with the universal algebra $\mathcal{U}(g)$. The operators of the center $Z(g)$ are the bi-invariant differential operators on $G$. More generally, we consider the algebra $\mathcal{U}(g)^K$ of right $K$-invariant differential operators in $\mathcal{U}(g)$, where $K$ is a maximal compact subgroup of $G$. $Z(\mathcal{U}(g)^K)$ will denote its center.

We denote by $\mathcal{D}(G)$ the space of $C^\infty$ functions with compact support. The dual $\mathcal{D}'(G)$ of continuous linear functionals in $\mathcal{D}(G)$ is the space of distributions of $G$.

An operator $P$ in $\mathcal{U}(g)$ acts on $\mathcal{D}'(G)$ in the following way:

$$PE(f) = E(P^t f),$$

where $P^t \in \mathcal{U}(g)$ is such that, if $dx$ is a Haar measure on $G$,

$$\int_G Pf(x)g(x)dx = \int_G f(x)P^t g(x)dx.$$ 

In addition, if $X \in g$, $X^t = -X$, so the $P \mapsto P^t$ is the anti-automorphism of $\mathcal{U}(g)$ extending $-Id$ of $g$. Also, this map preserves the subalgebras $Z(g)$ and $\mathcal{U}(g)^K$.

**Definition 1.** A distribution $E \in \mathcal{D}'(G)$ is a fundamental solution of a differential operator $P \in \mathcal{U}(g)$ if $PE = \delta$, where $\delta(f) = f(1)$; and $E$ is a parametrix of $P$ if $PE - \delta \in C^\infty(G)$. 

1
In this paper we extend the main result of [1] for a connected semisimple Lie group with a simply connected complexification. We start by recalling this result as it was stated in [1].

Remember that in general $Z(\mathcal{U}(g)^K) \simeq Z(g) \otimes Z(t)$ (Knop’s Theorem [9]). Given $h_0$ and $t_0$ Cartan subalgebras of $g_0$ and $t_0$ respectively, we will denote $\gamma^G_h$ and $\gamma^K_t$ the Harish-Chandra homomorphisms of $Z(g)$ and $Z(t)$ with respect to the subalgebras $h$ and $t$: Then we have

$$Z(g) \times Z(t) \xrightarrow{\gamma^G_h \otimes \gamma^K_t} U(h)^W \times U(t)^W_K \xrightarrow{\downarrow \otimes} U(h)^{iW} \otimes U(t)^{W_K} \xrightarrow{i \otimes} U(h \oplus t).$$

Therefore, by the way of the homomorphisms described above, we can associate to $P \in Z(\mathcal{U}(g)^K)$ a differential operator $(\gamma^G_h \otimes \gamma^K_t)(P)$ in the group $H \times T$, where $H$ and $T$ are the respective Cartan subgroups of $G$ and $K$ with Lie algebras $h_0$ and $t_0$.

We say that a Cartan subgroup $H$ of $G$ is fundamental if it contains a $G$-conjugate of a Cartan subgroup of $K$. All fundamental Cartan subgroups of $G$ are conjugate.

We will denote $\gamma^G = \gamma^G_h$, where $h_0$ is the Lie subalgebra of a fundamental Cartan subgroup $H$. Because in $K$ all Cartan subgroups are conjugate, we put $\gamma^K = \gamma^K_t$.

We can now state our main result:

**Theorem 1.1.** Let $G$ be a connected semisimple Lie group with a simply connected complexification. Let $H$ be a fundamental Cartan subgroup of $G$ and $P \in Z(\mathcal{U}(g)^K)$. If $(\gamma^G \otimes \gamma^K)(P)$ has a fundamental solution in $H \times T$, then $P$ has a fundamental solution in $G$.

When $P$ is a bi-invariant operator, we obtain a complete proof of the Theorem announced in [2] (see [3] for another proof of this result):

**Corollary 1.2 (Benabdallah-Rouvière).** Let $P \in Z(g)$. If $\gamma^G(P)$ has a fundamental solution in $H$, then $P$ has a fundamental solution in $G$.

The proof of Theorem 1.1 goes by explicit construction of the fundamental solution of $P$, using the Plancherel formula as the main tool.

2. Preliminaries.

In this section we fix notation and summarize some known facts about representation theory that will be needed through this paper. We refer to [1] for any unexplained notation.

**2.1. Notation.** Let $G$ be a linear connected semisimple Lie group, $g_0$ its Lie algebra. The complexification of any real Lie algebra will be denoted without the subscript.
We will assume that $G$ has a simply connected complexification, that is, $G$ is the analytic subgroup corresponding to $\mathfrak{g}_0$ of the simply connected complex group with Lie algebra $\mathfrak{g}$.

$\theta$ will denote a Cartan involution in either $\mathfrak{g}_0$, $\mathfrak{g}$ or $G$. Let $\mathfrak{g}_0 = \mathfrak{k}_0 + \mathfrak{p}_0$ be a Cartan decomposition of $\mathfrak{g}_0$ with respect to $\theta$; that is, $\mathfrak{k}_0 = \{ X \in \mathfrak{g}_0 : \theta X = X \}$ and $\mathfrak{p}_0 = \{ X \in \mathfrak{g}_0 : \theta X = -X \}$.

If $K$ is the analytic subgroup of $G$ with Lie algebra $\mathfrak{k}_0$, $K$ is a maximally compact subgroup of $G$. We fix $t_0$ a Cartan subalgebra of $\mathfrak{k}_0$ coming from a maximal torus $T$ of $K$.

On $\mathfrak{g}$ we define the inner product, $(X,Y) = -B(X,J\theta Y)$, where $B$ is the Killing form of $\mathfrak{g}$ and $J$ is conjugation with respect to $\mathfrak{g}_0$.

Let $H$ be any $\theta$-stable Cartan subgroup. Then $\mathfrak{h}_0 = \mathfrak{b}_0^H + \mathfrak{a}_0^H$, with $\mathfrak{b}_0^H \subseteq \mathfrak{t}_0$, $\mathfrak{a}_0^H \subseteq \mathfrak{p}_0$. We can associate to $H$ a cuspidal parabolic subgroup $Q^H$ with Langlands decomposition $M^HA^HN^H$.

The group $M^H$ is a reductive Lie group with compact Cartan subgroup $B^H$.

To avoid overloading the notation, we will not use superscripts for these subgroups when working with a fixed $H$ or whenever is clear from the context.

2.2. Discrete series of $M$ (cf. [8, XII.8]). Let $M_0$ be the identity component of $M$, $Z_M$ the center of $M$, and define $M^\# = M_0Z_M$.

If $\alpha \in \Delta(\mathfrak{g}_0, \mathfrak{h}_0)$ is a real root, let $\gamma_\alpha = \exp 2\pi i H_\alpha$, where $H_\alpha \in \mathfrak{a}$ is the standard co-root. If $F(B)$ is the group generated by all the $\gamma_\alpha$, then $F(B)$ is a finite abelian subgroup of $Z_M$. It also holds $M^\# = M_0F(B)$.

Discrete series representation of $M^\#$ are parametrized by pairs $(\lambda, \chi)$ such that:

1) $\lambda$ is a discrete series parameter of $M_0$; that is, $\lambda \in i\mathfrak{b}_0'$ is nonsingular and $\lambda - \rho_M$ is analytically integral, or equivalently, $\lambda$ is analytically integral, because $\rho_M$ is.

2) $\chi \in F(B)$.

3) $\chi = e^{\lambda - \rho_M}$ on $B \cap F(B)$.

We denote $\pi^\#(\lambda, \chi)$ the respective discrete series representation of $M^\#$, which has the following properties:

1) $\pi^\#(\lambda, \chi)$ has infinitesimal character $\lambda$.

2) $\pi^\#(\lambda, \chi) \simeq \pi^\#(\lambda', \chi')$ if and only if $\chi = \chi'$ and $\lambda' = w\lambda$ for some $w \in W(B_0, M_0)$.

Discrete series representations of $M$ are exactly the representations

$$\pi(\lambda, \chi) = \text{Ind}_{M_0^\#}^{M^\#} \pi^\#(\lambda, \chi).$$

Properties 1 and 2 above also hold for $\pi(\lambda, \chi)$.

As in the connected case, $\lambda$ is called the Harish-Chandra parameter of $\pi(\lambda, \chi)$ (or $\pi^\#(\lambda, \chi)$). We will denote $S_d(M)$ the set of all pairs $(\lambda, \chi)$ that
defines a discrete series representation of \( M \) and \( S_d(M_0) \) the set of Harish-Chandra parameters \( \lambda \).

2.3. \( H \)-series of \( G \). Given \( \pi(\lambda, \chi) \) a discrete series representation of \( M \), \( \nu \in a' \) a complex linear functional on \( a \), the \( H \)-series of \( G \) consists of the representations
\[
\pi(H, \lambda, \chi, \nu) = \text{Ind}_{M_0}^G \pi(\lambda, \chi) \otimes e^{\nu} \otimes 1 = \text{Ind}_{M_0}^{\#} \pi(\lambda, \chi) \otimes e^{\nu} \otimes 1.
\]
\( \pi(H, \lambda, \chi, \nu) \) has infinitesimal character \( \lambda + \nu \) relative to \( h_0 \). We will denote its global character by \( \Theta(H, \lambda, \chi, \nu) \). The unitary \( H \)-series is the subset when \( \nu \) takes pure imaginary values on \( a_0 \).

2.4. Plancherel formula. We can now write down the Plancherel formula for semisimple groups. Let \( \text{Car}(G) \) denote a set of representatives of \( \theta \)-stables Cartan subgroups.

\[\text{Theorem 2.1. There is a nonnegative function } m(H, \lambda, \chi, \nu) \text{ defined in } S_d(M) \times i a'_0 \text{ such that} \]
\[
\delta = \sum_{H \in \text{Car}(G)} \left( \sum_{(\lambda, \chi) \in S_d(M)} \int_{\nu \in i a'_0} \Theta(H, \lambda, \chi, \nu) m(H, \lambda, \chi, \nu) d\nu \right).
\]
The function \( m(H, \lambda, \chi, \nu) \) has the following properties:

(i) For each \( (\lambda, \chi) \in S_d(M) \), \( m(H, \lambda, \chi, \nu) \) is the restriction to \( i a'_0 \) of a meromorphic function on \( a' \) without poles on \( i a'_0 \).

(ii) Exist a positive constant \( C \) and a positive integer \( l \) such that for all \( (\lambda, \chi) \in S_d(M), \nu \in i a'_0 \), we have
\[|m(H, \lambda, \chi, \nu)| \leq C(1 + |\lambda|^2)^l(1 + |\nu|^2)^l.\]

We refer to [5, Theorem 6.17] for a more general and explicit statement of the above formula, and to [6, Lemma 3.3] for the inequality (ii).

3. Action of \( P \) on characters.

If \( \pi \) is an admissible representation with global character \( \Theta_\pi \) and infinitesimal character \( \chi_\pi \), and \( P \in \mathcal{Z}(g) \) is a bi-invariant differential operator, then
\[P \Theta_\pi = \chi_\pi(P) \Theta_\pi \quad ([8, \text{Prop. 10.24}]).\]
In [1] we prove a similar result for an operator \( P \in \mathcal{Z}(U(g)^K) \). We recall it here. If \( E \in \mathcal{D}'(G) \) and \( \tau \in K \) is an irreducible unitary representation of \( K \), then \( E^\tau \) denotes its isotypic component.

\[\text{Proposition 3.1. If } \pi \text{ is an admissible representation with infinitesimal character } \chi_\pi \text{ and global character } \Theta_\pi, \text{ and } P \text{ is a differential operator in } \mathcal{Z}(U(g)^K), \text{ then} \]
\[P \Theta_\pi^\tau = (\chi_\pi \otimes \chi_\tau)(P) \Theta_\pi^\tau.\]
4. Inversion of infinitesimal characters.

Let \( H \in \text{Car}(G) \), \( P \in \mathcal{Z}(U(g)^K) \), \( \lambda \in b' \), \( \mu \in t' \). We will set

\[
P(H, \lambda, \mu, \nu) = (\chi_{\lambda+\nu} \otimes \chi_{\mu})(P) = ((\gamma_{\lambda}^G \otimes \gamma^K)(P))(\lambda + \nu, \mu).
\]

So if \( \pi(H, \lambda, \chi, \nu) \) is an \( H \)-series representation, \( \tau \in \hat{K} \) with infinitesimal character \( \mu_{\tau} \in it'_{0} \), and \( P \in \mathcal{Z}(U(g)^K) \), we denote

\[
P(H, \lambda, \tau, \nu) = (\gamma_{\lambda}^G \otimes \gamma^K)(P)(\lambda + \nu, \mu_{\tau}).
\]

Notice that this equation is independent of \( \chi \). With this notation, Proposition 3.1 can be written

\[
P(\Theta(H, \lambda, \chi, \nu)) = P(H, \lambda, \tau, \nu) \Theta(\tau(H, \lambda, \chi, \nu)).
\]

Observe that \( P(H, \lambda, \tau, \nu) \) is a polynomial function of \( \nu \in a' \), and we will denote it \( P(H, \lambda, \tau) \). Also, if \( P \in \mathcal{Z}(g) \) or in \( \mathcal{Z}(t) \), the expression \( P(H, \lambda, \tau, \nu) \) simplifies to \( P(H, \lambda, \nu) \) or \( P(\tau) \), respectively.

**Proposition 4.1.** Given \( P \in \mathcal{Z}(U(g)^K) \) and \( H \in \text{Car}(G) \), if \( (\gamma_{G}^G \otimes \gamma^K)(P) \) has a fundamental solution on \( H_0 \times T = A \times B_0 \times T \), then there exist a constant \( C \) and a positive integer \( k \) such that

\[
\|P(H, \lambda, \tau)\| \geq \frac{C}{(1 + |\lambda|^2)^k (1 + |\tau|^2)^k} \quad \forall (\lambda, \tau) \in S_d(M_0) \times \hat{K}.
\]

**Proof.** This follows directly from (3) and Theorem 4.1 in [1].

Proposition 4.1 allows us to invert infinitesimal characters of the fundamental series, but we need to do it simultaneously for every \( H \)-series, \( H \in \text{Car}(G) \).

Then we will need the following:

**Proposition 4.2.** Let \( H, J \in \text{Car}(G) \), \( H \) fundamental, \( P \in \mathcal{Z}(U(g)^K) \). If \( (\gamma_{G}^G \otimes \gamma^K)(P) \) has a fundamental solution on \( H \times T \), then \( (\gamma_{J}^G \otimes \gamma^K)(P) \) has a fundamental solution on \( J_0 \times T \).

**Proof.** According to Theorem 4.1 in [1], if \( H \in \text{Car}(G) \), then \( (\gamma_{G}^G \otimes \gamma^K)(P) \) has a fundamental solution on \( H_0 \times T \) if and only if

\[
\|P(H, \lambda, \mu)\| \geq \frac{C}{(1 + |\lambda|^2)^k (1 + |\mu|^2)^k} \quad \forall (\lambda, \mu) \in \hat{B}_0 \times \hat{T}.
\]

So we just need to verify (5) for \( J_0 \times T \) provided that it is satisfied for \( H \times T \) with \( H \) fundamental.

Let \( \Gamma \) be a set of strongly orthogonal noncompact imaginary roots such that \( b^J \) is obtained from \( b \) by Cayley transform \( c_{\Gamma} \). Then \( \gamma_{G} = c_{\Gamma}^{-1} \circ \gamma_{J}^G \).
Given $\lambda^J \in \mathcal{S}_d(M_0^J)$, its extension by 0 to $i b_0'$ defines a character in $\hat{B}$. We write $\lambda$ for this extension, so $c_\Gamma(\lambda) = \lambda^J$. Then, if $\nu \in a'$,

$$P(H, \lambda, \mu)(\nu) = ((\gamma^G \otimes \gamma^K)(P))(\lambda + \nu, \mu) = ((\gamma^G \otimes \gamma^K)(P))(c_\Gamma(\lambda + \nu), \mu) = ((\gamma^G \otimes \gamma^K)(P))(\lambda^J + \nu, \mu) = P(J, \lambda^J, \mu)(\nu).$$

If $P$ has order $m$ in $\mathcal{U}(\mathfrak{g})$, $P(H, \lambda, \mu)$ and $P(J, \lambda^J, \mu)$ are polynomial functions of order $\leq m$ on $a'$ and $(a'^J)'$ respectively, and their norms are the norms of the vectors in $\mathbb{C}^{m+1}$ formed with their coefficients, then, by Schwarz inequality,

$$\|P(H, \lambda, \mu)\| \leq \|P(J, \lambda^J, \mu)\|;$$

so by hypothesis

$$\|P(J, \lambda^J, \mu)\| \geq \frac{\overline{C}}{(1 + |\lambda^J|^2)^k (1 + |\mu|^2)^k} \quad \forall (\lambda^J, \mu) \in \tilde{J}_0 \times \tilde{T}.$$

\[ \square \]

5. Inversion of global characters.

One important step in building the fundamental solution of $P$ is the construction of distributions $R_\pi$ such that $PR_\pi = \Theta_\pi$ for each representation $\pi$ that appears in the Plancherel formula. In this section we will define these distributions $R_\pi$. First we state an analog of Proposition 6.3 in [1] for the $H$-series. We notice that the proof goes exactly the same way.

**Proposition 5.1.** Let $H \in \text{Car}(G)$. There exist $Z \in \mathcal{Z}(\mathfrak{g})$, a positive constant $C$, an $\varepsilon > 0$ and a positive integer $k$ such that

$$|Z(H, \lambda, \nu + z)| \geq C(1 + |\lambda|^2)^k (1 + |\nu|^2)^k \quad \forall \lambda \in \mathcal{S}_d(M_0), \nu \in \mathfrak{a}_0', \quad z \in a', |z| < \varepsilon.$$

1 Although this is a well-known result, we include a sketch of the proof here, since we don’t have a reference. Let $T$ be a torus, and $T_1 \subseteq T$ any subtorus. Then there is a vector space $V$ and a lattice $\Lambda$ such that $T = V/\Lambda$. Also $T_1 = V_1/\Lambda_1$, where $V_1$ is a subspace of $V$ and $\Lambda_1 = \Lambda \cap V_1$.

It suffices to find another subtorus $T_2$ such that $T = T_1 \times T_2$, and for that it suffices to find a sublattice $\Lambda_2$ such that $\Lambda = \Lambda_1 \oplus \Lambda_2$.

Let’s see first that $\Lambda/\Lambda_1$ is torsion free: let $\eta \in \Lambda$ and suppose that $k\eta \in \Lambda_1$ for some positive integer $k$. Then $\eta = (1/k)\eta_1 \in V_1$, so $\eta \in \Lambda_1$.

Now if $\{\xi_1, \ldots, \xi_n\}$ and $\{\eta_1 + \Lambda_1, \ldots, \eta_n + \Lambda_1\}$ are generating sets as free abelian groups of $\Lambda_1$ and $\Lambda/\Lambda_1$ respectively, then it’s easy to see that $\Lambda = \mathbb{Z}[\xi_i, \eta_j]$, and we set $\Lambda_2 = \mathbb{Z}[\eta_1]$. I wish to thank Prof. Joseph Wolf for giving me this sketch.
We are now in position to define the distributions $R(H, \lambda, \chi, \nu)$ for $P \in \mathcal{Z}(\mathcal{U}(g)K)$ satisfying the hypothesis of Theorem 1.1. Let $\varepsilon > 0$ be given by Proposition 5.1, $m$ the order of $P$ and let $\Phi \in C^\infty(\text{Pol}^0(m) \times \mathbb{C})$ be the nonnegative function given by [1, Lemma 6.6].

Given $(\lambda, \chi) \in \mathcal{S}_d(M)$, $\tau \in \hat{K}$, and a fixed $\nu \in \mathfrak{a}'$, we put

$$(6) \quad P^\nu(H, \lambda, \tau)(z) = P(H, \lambda, \tau)(\nu + z) = P(H, \lambda, \tau, \nu + z).$$

Finally, if $dz$ is Lebesgue measure in $\mathfrak{a}'$, $f \in \mathcal{D}(G)$, we define

$$(7) \quad R(H, \lambda, \chi, \nu) = \sum_{\tau \in \hat{K}} \int_{|z| < \varepsilon} \frac{\Theta^\tau(H, \lambda, \chi, \nu + z)}{P^\nu(H, \lambda, \tau)(z)} \Phi(P^\nu(H, \lambda, \tau), z) \, dz.$$ 

This definition makes sense because $P^\nu(H, \lambda, \tau)(z) \neq 0$ if $\Phi(P^\nu(H, \lambda, \tau), z) \neq 0$ ([1, Lemma 6.6 (iv)]).

**Proposition 5.2.** The map defined by (7) is a finite order distribution for all $(\lambda, \chi) \in \mathcal{S}_d(M)$, $\nu \in \mathfrak{a}'$. This map has also the following properties:

(i) $PR(H, \lambda, \chi, \nu) = \Theta(H, \lambda, \chi, \nu)$.

(ii) For every positive integer $k$ and $f \in \mathcal{D}(G)$, exist a constant $C > 0$ which only depends on the support of $f$ and a differential operator $D_k \in \mathcal{Z}(\mathcal{U}(g)K)$ such that $\forall (\lambda, \chi) \in \mathcal{S}_d(M)$, $\nu \in i\mathfrak{a}'_0$

$$|R(H, \lambda, \chi, \nu)(f)| \leq \frac{C}{(1 + |\nu|^2)^k(1 + |\lambda|^2)^k} \|D_k f\|_{L^2(G)}.$$ 

**Proof.** Although the argument is the same as [1, Proposition 6.8] for MAN minimal parabolic, we will sketch it here in order to verify that every step of the argument still applies for MAN any cuspidal parabolic.

(i) is clear, combining (4) with [1, Lemma 6.6 (iii)] and the fact that $P(H, \lambda, \tau, \nu + z)$ is holomorphic in $z$.

Let’s see that (7) defines a distribution: According to [1, Lemma 6.6 (iv)] and [1, Lemma 6.7] together with Proposition 4.2, it holds, for all $|z| < \varepsilon$ and some $k$,

$$(8) \quad \left| \frac{\Phi(P^\nu(H, \lambda, \tau), z)}{P^\nu(H, \lambda, \tau)(z)} \right| \leq C_1 (1 + |\nu|^2)^m (1 + |\lambda|^2)^k (1 + |\tau|^2)^{\tilde{k}}.$$ 

On the other hand, if $Z \in \mathcal{Z}(g)$ and $\Omega \in \mathcal{Z}(\mathfrak{t})$ are given by Proposition 5.1 and [1, Lemma 6.1] respectively, and if $s_1$ and $s_2$ are positive integers, we have, for some positive integer $\tilde{k}$,

$$(9) \quad |\Theta^\tau(H, \lambda, \chi, \nu + z)(f)| \leq \frac{C_2 |\Theta(H, \lambda, \chi, \nu + z)\left((Z')^{s_1} (\Omega')^{s_2} f\right)|}{((1 + |\lambda|^2)(1 + |\nu|^2))^{s_1 - \tilde{k}} (1 + |\tau|^2)^{s_2}}.$$ 

Let $\tilde{K} \subseteq G$ be a compact subset. Note that for $H$-series representations the multiplicity $n_{\tau}$ of any $K$-type satisfies $n_{\tau} \leq \dim \tau$ ([8, p. 207]); so
by [1, Lemma 6.5] exist \( \Omega \in \mathcal{Z}(\mathfrak{g}) \) and a constant \( C_3 \) independent of \( \lambda \in \mathcal{S}_d(M_0) \), \( \nu, z \in \mathfrak{a}' \) such that

\[
|\Theta(H, \lambda, \chi, \nu + z)(f)| \leq C_3 \left( \int_G |\tilde{\Omega}f(g)|^2 \|\pi(H, \lambda, \chi, \nu + z)(g)\|^2 \, dg \right)^{1/2}.
\]

On the other hand, given \( \varphi \) in the space where \( \pi(H, \lambda, \chi, \nu + z) \) acts, if \( a(g) \) is the \( A \)-component of \( g \) in the \( K\)MAN decomposition, then (cf. [8, p. 169]),

\[
(\pi(H, \lambda, \chi, \nu + z)(g)\varphi)(k) = e^{-z \log a(g^{-1}k)} (\pi(H, \lambda, \chi, \nu)(g)\varphi)(k),
\]

and taking \( A = \sup_{g \in \tilde{K}, k \in K, |z| < \varepsilon} |e^{-z \log a(g^{-1}k)}| \) and \( B_{\lambda, \nu} = \sup_{g \in \tilde{K}} \|\pi(H, \lambda, \chi, \nu)(g)\| \),

then \( \|\pi(H, \lambda, \chi, \nu + z)(g)\| \leq AB_{\lambda, \nu} \) uniformly on \( \tilde{K} \), so for all \( f \) supported in \( \tilde{K} \),

\[
|\Theta^T(H, \lambda, \chi, \nu + z)(f)| \leq AB_{\lambda, \nu}\|\tilde{\Omega}f\|_{L^2(G)}.
\]

Now combining (8), (9) and (10) we obtain, for all \( f \) supported in \( \tilde{K} \),

\[
|R(H, \lambda, \chi, \nu)(f)| \leq \left( \sum_{\tau \in \tilde{K}} \frac{1}{(1 + |\tau|^2)^{s_2-k}} \right)^{1/2} \frac{C_1 B_{\lambda, \nu} \|\tilde{\Omega}(\Omega^T)^{s_1}(\Omega^T)^{s_2} f\|_{L^2(G)}}{(1 + |\nu|^2)^{s_1-m-\tilde{k}}(1 + |\sigma|^2)^{s_1-k-\tilde{k}}}
\]

and \( \sum_{\tau \in \tilde{K}} \frac{1}{(1 + |\tau|^2)^{s_2-k}} \) is finite if we choose \( s_2 > \tilde{k} + 1/2 \dim K \). Hence \( R_{\sigma, \nu} \) is a finite order distribution.

To see (ii), just observe that \( B_{\lambda, \nu} = 1 \) if \( \nu \in i\mathfrak{a}'_0 \), so given \( k \) if we take \( D_k = \tilde{\Omega}(\Omega^T)^{s_1}(\Omega^T)^{s_2} \) with the \( s_2 \) chosen above and \( s_1 \geq k + \tilde{k} + \max(k, m) \),

\[
|R(H, \lambda, \chi, \nu)(f)| \leq \frac{C}{(1 + |\nu|^2)^{k}(1 + |\lambda|^2)^k} \|D_k f\|_{L^2(G)}
\]

with \( C \) depending only \( \tilde{K} \). \( \square \)

6. Demonstration of Theorem 1.1.

Now we are ready to complete the proof of Theorem 1.1 with the explicit construction of the fundamental solution of \( P \).

Proposition 6.1. Let \( P \in \mathcal{Z}(\mathcal{U}(\mathfrak{g})^K) \). Suppose that for each \( H \in \text{Car}(G) \) there exist a constant \( C \) and a positive integer \( k \) such that

\[
\|P(H, \lambda, \tau)\| \geq \frac{C}{(1 + |\lambda|^2)^k(1 + |\tau|^2)^k} \quad \forall (\lambda, \tau) \in \mathcal{S}_d(M_0) \times \tilde{K}.
\]
If $R(H, \lambda, \chi, \nu)$ are the distributions defined by (7), then the map $R$ defined by

$$
R = \sum_{H \in \text{Car}(G)} \left( \sum_{(\lambda, \chi) \in S_d(M)} \int_{\nu \in i\mathfrak{a}'_0} R(H, \lambda, \chi, \nu) \cdot m(H, \lambda, \chi, \nu) \, d\nu \right)
$$

is a finite order distribution which is a fundamental solution of $P$.

**Remark.** If $P \in \mathcal{Z}(\mathcal{U}(\mathfrak{g})^K)$ is such that $(\gamma^G \otimes \gamma^K)(P)$ has a fundamental solution in $H \times T$, Propositions 4.1 and 4.2 imply (12) for all $H \in \text{Car}(G)$, so Theorem 1.1 is a direct consequence of Proposition 6.1.

**Proof.** Equality $PR = \delta$ is clear by Plancherel formula (Theorem 2.1) and because $PR(H, \lambda, \chi, \nu) = \Theta(H, \lambda, \chi, \nu)$; it only remains to prove that $R$ is a finite order distribution. For that we will prove that each of the following are finite order distributions:

$$
R_H = \sum_{(\lambda, \chi) \in S_d(M)} \int_{\nu \in i\mathfrak{a}'_0} R(H, \lambda, \chi, \nu) \cdot m(H, \lambda, \chi, \nu) \, d\nu.
$$

Let $\tilde{K}$ be a compact subset and $f \in \mathcal{D}_{\tilde{K}}(G)$; for each positive integer $k$ let $D_k \in \mathcal{Z}(\mathcal{U}(\mathfrak{g})^K)$ be given by Proposition 5.2 (ii); then, using Theorem 2.1 (ii),

$$
|R_H(f)| \leq C_1 C_2 \left( \sum_{\lambda \in S_d(M_0)} \frac{1}{(1 + |\lambda|^2)^{l_2}} \left( \int_{\nu \in i\mathfrak{a}'_0} \frac{1}{(1 + |\nu|^2)^{k-l_1}} \right) \right) \|D_k f\|_{L^2(G)}.
$$

Choosing $k$ large enough so that the sum and the integral are finite, we obtain an operator $D \in \mathcal{Z}(\mathcal{U}(\mathfrak{g})^K)$ and a constant $C$ depending only on $\tilde{K}$ such that

$$
|R_H(f)| \leq C \|D f\|_{L^2(G)},
$$

and this proves that $R_H$ is a distribution of finite order less or equal that the order of $D$. \qed

7. **Final remarks.**

7.1. **P-convexity of $G$.** Suppose that $P \in \mathcal{Z}(\mathcal{U}(\mathfrak{g})^K)$ satisfies the conditions of Proposition 6.1. The existence of fundamental solution of $P$ implies that the differential equation $Pu = f$ has a solution $u \in C^\infty(G)$ for all $f \in \mathcal{D}(G)$. Now, in order to guarantee the solvability of $Pu = f$ when $f \in C^\infty(G)$, it is necessary to analyze the $P$-convexity of $G$.

We have done this already in [1, §8]. We just have to notice that that argument applies for general linear semisimple groups, since Johnson’s injectivity criterion (Theorems 5.1 and 5.2 in [7]) holds for these groups.
7.2. Casimir operator. Let $\Omega \in Z(g)$ be the Casimir operator of $G$. As it was remarked in [2], Corollary 1.2 provides a fundamental solution of $\Omega$ only if $G$ doesn’t have a compact Cartan subgroup. For the sake of completeness, we recall the argument.

If $h$ is a Cartan subalgebra of $g$, and if $\lambda \in h'$, then $\chi_\lambda(\Omega) = B(\lambda,\lambda) - B(\rho,\rho)$; in particular, if $H \in \text{Car}(G)$, $h = b \oplus a$, $\lambda \in \hat{B}_0$, $\nu \in i\hat{a}_0$, then

$$\Omega(H,\lambda,\nu) = \chi_\lambda(\Omega) = |\lambda|^2 - |\nu|^2 - |\rho|^2,$$

so if $a \neq 0$, $\|\Omega(H,\lambda)\| \geq 1$ for all $\lambda \in \hat{B}_0$, and $\gamma_B^G(\Omega)$ has a fundamental solution on $H_0$. Now if $G$ has a compact Cartan subgroup $B$

$$\Omega(B,\lambda) = |\lambda|^2 - |\rho|^2.$$

Since $G^C$ is simply connected, $\lambda = \rho$ is a character of $B$, and consequently we cannot apply Corollary 1.2 in this case.

7.3. Operators of $Z(k)$. Let’s consider $P \in Z(k) \subset Z(U(g)^K)$. Then $P$ can be seen as a differential operator acting on both $G$ and $K$. Now, according to [4, Theorem I], $P$ has a fundamental solution on $K$ if and only if there exists $C > 0$, $k \in \mathbb{N}$ such that

$$|P(\tau)| \geq \frac{C}{(1 + |\tau|^2)^k} \quad \forall \tau \in \hat{K}.$$

We are in position to prove the following:

**Proposition 7.1.** If $P \in Z(k) \subset Z(U(g)^K)$ has a fundamental solution on $K$, then it has one in $G$.

**Proof.** All we have to do at this point is notice that inequality (16) is nothing else than (12) for $P \in Z(t)$, so we can apply Proposition 6.1 directly. $\square$

Let’s finish analyzing the case of $\Omega_K \in Z(t)$ the Casimir of $K$. In this case we have (see [1, §11])

$$\gamma_a \otimes \chi_\tau(\Omega_K) = \chi_\tau(\Omega_K) = |\lambda_\tau|^2 - |\rho_K|^2.$$

So if we choose $\tau$ such that $\lambda_\tau = \rho_K$, by [1, Prop. 11.1], $\Omega_K$ doesn’t have a parametrix. However, if we take $\Omega_K + C$ such that $C > |\rho_K|^2$, then

$$(\Omega_K + C)(H,\lambda,\mu,\nu) = \chi_\mu(\Omega_K + C) \geq \text{Constant}$$

and Theorem 1.1 applies, therefore $\Omega_K + C$ does have a fundamental solution.
Acknowledgments. I wish to thank Prof. Joseph Wolf, who is my advisor as a postdoctoral fellow in Berkeley, for all his support during this period; and particularly for his help with many topics of this paper.

I also want to thank my friend Dr. Leandro Cagliero for his unconditional will to spend his time to read, discuss, and help me with this work.

Finally, I’d like to thank the Mathematics Department of the University of California at Berkeley.

References


Received March 1, 2001 and revised May 30, 2001. This work was supported by CONICET.
FREE STOCHASTIC MEASURES VIA NONCROSSING
PARTITIONS II

MICHAEL ANSHELEVICH

We show that for stochastic processes with freely independent
increments, the partition-dependent stochastic measures
can be expressed purely in terms of the higher stochastic mea-
sures and the higher diagonal measures of the original process.

1. Introduction.

Starting with an operator-valued stochastic process with freely independent
increments $X(t)$, in [A] we defined two families $\{Pr_\pi\}$ and $\{St_\pi\}$ indexed
by set partitions. These objects give a precise meaning to the following
heuristic expressions. For a partition $\pi = (B_1, B_2, \ldots, B_n) \in P(k)$, tem-
porarily denote by $c(i)$ the number of the class $B_{c(i)}$ to which $i$ belongs.
Then, heuristically,

$$Pr_\pi(t) = \int_{[0,t]^n} dX(s_{c(1)})dX(s_{c(2)})\ldots dX(s_{c(k)})$$

and

$$St_\pi(t) = \int_{[0,t]^n} \text{all } s_i \text{'s distinct} dX(s_{c(1)})dX(s_{c(2)})\ldots dX(s_{c(k)}).$$

In particular, denote by $\psi_k$ and $\Delta_k$ the higher stochastic measures and the
higher diagonal measures, defined, respectively, by

$$\psi_k(t) = \int_{[0,t]^k} \text{all } s_i \text{'s distinct} dX(s_1)dX(s_2)\ldots dX(s_k)$$

and

$$\Delta_k(t) = \int_{[0,t)} (dX(s))^k.$$

Rigorous definitions of all these objects in terms of Riemann sums are given
below. These definitions were motivated by [RW], where corresponding
objects were defined for the usual Lévy processes. There is a number of
differences between the classical and the free case. First, the free increments
property implies that $St_\pi = 0$ unless $\pi$ is a noncrossing partition. Second,
the point of the analysis of [RW] was that while we are really interested in
the stochastic measures $St_\pi$, notably $\psi_k$, these are rather hard to define or
to handle. However, by the use of Möbius inversion these can be expressed through the $\Pr_\pi$. It is easy to see that if the increments of the process $X$ commute, in the defining expression for $\Pr_\pi$ all the terms corresponding to the same class can be collected together, and the result is just a product measure over the classes of the partition, $\Pr_\pi = \Delta_{B_1} \Delta_{B_2} \cdots \Delta_{B_n}$. So in this way stochastic measures $\St$ can be defined using ordinary product measures. This fact is a consequence of the commutativity of the increments of the process; in the free probability case the operators do not commute, and unless the classes of the partition $\pi$ are just intervals we cannot expect $\Pr_\pi$ to be a product measure; indeed a counterexample was given in [A].

In this paper we show that while we cannot expect nice factorization properties in the general case of noncommuting variables, the free independence of the increments does imply a product-like property. Namely, by an argument similar to the above one, if the increments of the process commute, then $\St_\pi = \psi_k(\Delta_{B_1}, \Delta_{B_2}, \cdots, \Delta_{B_n})$. This property certainly does not hold either if the increments do not commute, but if the increments are freely independent it can be modified as follows. In a noncrossing partition, one distinguishes classes which are inner, or covered by some other classes, and outer. For example, in the noncrossing partition $((1, 6, 7)(2, 5)(3)(4)(8)(9, 10))$, the classes $\{2, 5\}, \{3\}, \{4\}$ are inner while the classes $\{1, 6, 7\}, \{8\}, \{9, 10\}$ are outer. For a partition with only outer classes, which therefore have to be intervals, the product decomposition of $\Pr_\pi$ and the above decomposition of $\St_\pi$ hold even in the noncommutative case. We show in the main theorem of this paper that the inner classes, while making a complicated contribution to $\Pr_\pi$, make only scalar contributions to $\St_\pi$, and those contributions commute with everything.

We also use the opportunity to make extensions of the definitions of [A] in various directions. The objects $\St_\pi$ and $\Pr_\pi$ were defined by limits of Riemann-like sums with uniform subdivisions. In this paper we extend that definition to arbitrary subdivisions. We also make some preliminary steps towards defining multi-dimensional free stochastic measures.

2. Preliminaries.

This paper is a sequel to [A]; see that paper for all the definitions that are not explicitly provided here.

2.1. Notation. Denote by $[n]$ the set $\{1, 2, \ldots, n\}$.

For two vectors $X = (X_1, X_2, \ldots, X_k)$ and $Z = (Z_1, Z_2, \ldots, Z_{k-1})$ denote

$$X \circ Z = X_1 Z_1 X_2 Z_2 \cdots X_{k-1} Z_{k-1} X_k.$$ 

For a collection of vectors $\{X_i\}_{i=1}^n$, denote by $(X_1, X_2, \ldots, X_n)$ their concatenation.
For a collection of objects \( \{ y_j \} \) and a multi-index \( \vec{v} = (v_1, v_2, \ldots, v_n) \), we will throughout the paper use the notation \( y_{\vec{v}} \) to denote \( \prod_{j=1}^{n} y_{v_j}^{(j)} \).

For a family of functions \( \{ F_j \} \), where \( F_j \) is a function of \( j \) arguments, \( \vec{v} \) a vector with \( k \) components, and \( B \subset [k] \), denote \( F(\vec{v}) = F_k(\vec{v}) \) and

\[
F(B; \vec{v}) = F_{|B|}(v_{i(1)}, v_{i(2)}, \ldots, v_{i(|B|)}),
\]

where \( B = (i(1), i(2), \ldots, i(|B|)) \). In particular, using this notation \( y_{(B; \vec{v})} = \prod_{i \in B} y_{v_i}^{(i)} \).

2.2. Partitions. Denote by \( \mathcal{P}(k) \) and \( NC(k) \) the lattice of all set partitions of the set \([k]\) and its sub-lattice of noncrossing partitions. Let \( \hat{0} \) and \( \hat{1} \) be the smallest and the largest elements in the lattice ordering, namely \( \hat{0} = ((1), (2), \ldots, (k)) \) and \( \hat{1} = (1, 2, \ldots, k) \). Denote by \( \wedge \) the join operation in the lattices. For \( \pi \in NC(k) \), denote by \( K(\pi) \) its Kreweras complement. For \( \pi \in \mathcal{P}(n) \), define \( \pi_{\text{op}} \in \mathcal{P}(n) \) to be \( \pi \) taken in the opposite order, i.e.,

\[
i \sim_{\text{op}} j \iff (n - i + 1) \sim (n - j + 1).
\]

For \( \pi \in \mathcal{P}(n), \sigma \in \mathcal{P}(k) \), define \( \pi + \sigma \in \mathcal{P}(n + k) \) by

\[
i \hat{\sim}_{\pi + \sigma} j \iff ((i, j \leq n, i \hat{\sim} j) \text{ or } (i, j > n, (i - n) \hat{\sim} (j - n))).
\]

2.3. Free cumulants. All the operators involved will live in an ambient noncommutative probability space \( (\mathcal{A}, \varphi) \), where \( \mathcal{A} \) is a finite von Neumann algebra, and \( \varphi \) is a faithful normal tracial state on it. Let \( A = (A_1, A_2, \ldots, A_k) \) be a \( k \)-tuple of self-adjoint operators in \( (\mathcal{A}, \varphi) \). Denote their joint moments by

\[
M(\mathcal{A}) = \varphi[A_1A_2 \ldots A_k].
\]

For a noncrossing partition \( \pi \), denote using the above notation

\[
M_\pi(\mathcal{A}) = \prod_{B \in \pi} M(B; \mathcal{A}).
\]

Also define the combinatorial \( R \)-transform, or the collection of joint free cumulants \( R(\mathcal{A}) \): Denoting

\[
R_\pi(\mathcal{A}) = \prod_{B \in \pi} R(B; \mathcal{A}),
\]

the functional \( R \) is determined inductively by

\[
M(\mathcal{A}) = \sum_{\sigma \in NC(k)} R_\sigma(\mathcal{A}),
\]

or more generally by

\[
M_\pi(\mathcal{A}) = \sum_{\sigma \in NC(k) \atop \sigma \leq \pi} R_\sigma(\mathcal{A}).
\]
Any such relation can be inverted by using Möbius inversion, so we also get
\[ R_{\pi}(A) = \sum_{\sigma \in NC(k) \atop \sigma \leq \pi} \text{Möb}(\sigma, \pi) M_{\sigma}(A), \]
where Möb is the relative Möbius function on the lattice of noncrossing partitions. In particular, since \(|M_{\pi}(A)| \leq \prod_{i=1}^{k} \|A_i\|\) and Möb(\(\pi, \sigma\), \(|NC(k)|\) are products of Catalan numbers and so are bounded in norm by \(4^k\), we conclude that \(|R_{\pi}(A)| \leq 16^k \prod_{i=1}^{k} \|A_i\|\).

Finally, the relation between the free cumulants and free independence is expressed in the “mixed cumulants are zero” condition: \(R(A) = 0\) whenever some \(A_i, A_j\) are freely independent.

**Definition 1.** Let \(X = (X^{(1)}, X^{(2)}, \ldots, X^{(k)})\) be a \(k\)-tuple of free stochastic measures with distributions \(\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(k)}\). Here, \(\mu^{(i)}\) is a freely infinitely divisible distribution with compact support, and \(X^{(i)}\) is an operator-valued measure on \(\mathbb{R}\) that is self-adjoint, additive, stationary, and has freely independent increments. We say that the \(k\)-tuple is **consistent** if the following extra conditions are satisfied.

1) **Free increments:** For a family of disjoint intervals \(\{I_i\}_{i=1}^{n}\) and a multi-index \(\bar{u}\) of length \(n\), the family \(\{X^{(u_1)}(I_1)X^{(u_2)}(I_2)\ldots X^{(u_n)}(I_n)\}\) depends only on \(\bar{u}\) and \(|I|\).

2) **Stationarity:** For an interval \(I\) and a multi-index \(\bar{u}\),
\[ \varphi[X^{(u_1)}(I)X^{(u_2)}(I)\ldots X^{(u_n)}(I)] \]
depends only on \(\bar{u}\) and \(|I|\).

3) **Continuity:** For a fixed \(\bar{u}\), the function
\[ |I| \mapsto \varphi[X^{(u_1)}(I)X^{(u_2)}(I)\ldots X^{(u_n)}(I)] \]
is continuous.

Note that these conditions together imply that for an arbitrary collection of intervals \(\{I_i\}_{i=1}^{n}\), \(\varphi[X^{(u_1)}(I_1)X^{(u_2)}(I_2)\ldots X^{(u_n)}(I_n)]\) depends only on \(\bar{u}\) and the sizes of all elements of \(\{\bigcap_{i \in G} I_i : G \subset [n]\}\). Note also that by definition and Möbius inversion, the stationarity and continuity properties apply not just to \(M\), but to \(M_{\pi}, R, R_{\pi}\) as well.

**Remark 1.** It is easy to see that a \(k\)-tuple \(X\) is consistent if \(X^{(i)} = X\) for all \(i\), or if the family \(\{X^{(i)}\}_{i=1}^{k}\) is a freely independent family. More examples are given in Lemma 10 and in Remark 5.

Fix a consistent \(k\)-tuple of free stochastic measures \(X\). For \(t > 0\), denote \(X^{(i)}(t) = X^{(i)}([0, t))\). Throughout most of the paper, we will consider \(t = 1\), in which case we will omit \(t\) from the notation; in particular we will denote \(X^{(i)}([0, 1))\) simply by \(X^{(i)}\). Let \(S = (I_1, I_2, \ldots, I_N)\) be a subdivision of the
interval $[0,t)$ into $N = |S|$ disjoint half-open intervals, listed in increasing order. Denote $\delta(S) = \max_{1 \leq i \leq N} |I_i|$. Let $X^{(i)}(S) = X^{(i)}(I_j)$. In the future we will frequently omit the dependence on $\hat{S}$ and $N$ in the notation.

**Notation 2.** For any set $G$ and a partition $\pi \in P(k)$, denote

$$G^k_{\pi} = \{ v \in G^k : i \sim j \Leftrightarrow v_i = v_j \}$$

and

$$G^k_{\geq \pi} = \{ v \in G^k : i \not\sim j \Rightarrow v_i = v_j \}.$$

Denote

$$St_\pi(X,S) = \sum_{\bar{v} \in [N]^k_{\pi}} X_{\bar{v}}(S)$$

and

$$Pr_\pi(X,S) = \sum_{\bar{v} \in [N]^k_{\geq \pi}} X_{\bar{v}}(S).$$

**Definition 3.** Define the free stochastic and product measures depending on a partition to be the limits along the net of subdivisions of the interval $[0,t)$

$$St_\pi(X,t) = \lim_{\delta(S) \to 0} St_\pi(X,S),$$

$$Pr_\pi(X,t) = \lim_{\delta(S) \to 0} Pr_\pi(X,S).$$

In particular, let the higher diagonal measure be

$$\Delta(X,t) = St_1(X,t) = Pr_1(X,t),$$

and the $k$-dimensional free stochastic measure be $\psi(X,t) = St_0(X,t)$. If $X^{(i)} = X$ for all $i$, we denote $\Delta(X,t)$ by $\Delta_k(t)$ and $\psi(X,t)$ by $\psi_k(t)$.

Here the limits are taken in the operator norm; the proof of their existence is part of the arguments in the next section.

**Lemma 1.** For an arbitrary family of intervals $\{J_i \subset [0,1]\}_{i=1}^n$, a multi-index $\bar{u}$, and $\pi \in NC(k)$,

$$R_\pi(X^{(u_1)}(J_1), X^{(u_2)}(J_2), \ldots, X^{(u_n)}(J_n)) = \left( \prod_{B \in \pi} \left| \bigcap_{i \in B} J_i \right| \right) R_\pi(X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)}).$$

In particular, for a subdivision $S = (I_1, I_2, \ldots, I_N)$ of $[0,1)$,

$$R_\pi \left( X_j^{(u_1)}, X_j^{(u_2)}, \ldots, X_j^{(u_n)} \right) = |I_j|^{|\pi|} R_\pi \left( X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)} \right).$$
Proof. The second statement follows from the first one with all intervals $J_i = I_j$ with a fixed $j$. For the first statement, it suffices to show that

$$R(X^{(u_1)}(J_1), X^{(u_2)}(J_2), \ldots, X^{(u_n)}(J_n))$$

$$= \bigcap_{i=1}^{n} J_i \left| R(X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)}).$$

Moreover, since each $X^{(j)}$ is an additive process with freely independent increments, $R_{\pi}(A_1, A_2, \ldots, A_n)$ is multi-linear in its arguments, and all mixed cumulants are equal to 0, it suffices to show that

$$R \left( X^{(u_1)}(I), X^{(u_2)}(I), \ldots, X^{(u_n)}(I) \right)$$

$$= |I| R \left( X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)} \right)$$

with $I = \bigcap_{i=1}^{n} J_i$. First suppose that $I = I_j$, one of the intervals in a uniform subdivision, with $I_i = \left[ \frac{i-1}{N}, \frac{i}{N} \right)$. Then using the same properties as above,

$$R \left( X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)} \right) = \sum_{\pi \in [N]^n} R \left( X^{(u_1)}_{\pi_1}, X^{(u_2)}_{\pi_2}, \ldots, X^{(u_n)}_{\pi_n} \right)$$

$$= \sum_{i=1}^{N} R \left( X^{(u_1)}_{i}, X^{(u_2)}_{i}, \ldots, X^{(u_n)}_{i} \right)$$

$$= NR \left( X^{(u_1)}_{j}, X^{(u_2)}_{j}, \ldots, X^{(u_n)}_{j} \right),$$

where in the last equality we have used that fact that by stationarity of $X$, $R(X^{(u_1)}_{i}, X^{(u_2)}_{i}, \ldots, X^{(u_n)}_{i})$ does not depend on $i$. Therefore

$$R \left( X^{(u_1)}_{j}, X^{(u_2)}_{j}, \ldots, X^{(u_n)}_{j} \right) = N^{-1} R \left( X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)} \right)$$

$$= |I_j| R \left( X^{(u_1)}, X^{(u_2)}, \ldots, X^{(u_n)} \right).$$

By stationarity, it follows that Equation (1) holds for $|I| = 1/N$ and consequently for any rational $|I|$. The result follows for general $I$ by the continuity assumption on $X$. \qed

This was the main fact used in the proofs of [A], and so the results from that paper carry over to the consistent $k$-tuples of free stochastic measures. Note that all of these results were proven for uniform subdivisions. However, their proofs carry over to the more general definitions of this paper without difficulty; see Lemma 2 for an example of a computation. We list some of those results, with the numbering from [A].

1) $St_\pi(X) = 0$ unless $\pi$ is noncrossing (Theorem 1).
2) $\varphi[St_\pi(X)] = R_\pi(X)$ (Corollary 2).
3) If $\pi$ contains an inner singleton and $X$ is centered, then $\text{St}_\pi(X) = 0$ (Proposition 1).

4) If $Z$ is freely independent from the free stochastic measure $X$, then
\[
\lim_{\delta(S) \to 0} \sum_{i=1}^{N} X_i Z X_i = \varphi[Z] \Delta_2(X)
\]
(Corollary 13).

5) The limit defining $\text{St}_\pi$ exists in the norm topology if the corresponding limit exists for the free Poisson process. In particular, for any consistent $k$-tuple $X$ of free stochastic measures, $\Delta(X)$ is well-defined. In fact, the argument in [A] needs to be modified (Pr should be used in place of St); such a modification is contained in Lemma 7 of this paper.

Results 2 and 3 are consequences of, and result 4 is parallel to, the following theorem.

**Main Theorem.** Let $\pi$ be a noncrossing partition of $[k]$ with $o(\pi)$ outer classes $B_1, B_2, \ldots, B_{o(\pi)}$ and $i(\pi)$ inner classes $C_1, C_2, \ldots, C_{i(\pi)}$. Then
\[
\text{St}_\pi(X) = \prod_{i=1}^{i(\pi)} R(C_i; X) \cdot \psi(\Delta(B_1; X), \Delta(B_2; X), \ldots, \Delta(B_{o(\pi)}; X)).
\]

**Remark 2.** The distinction between the inner and the outer classes of a noncrossing partition was noted in [BLS]. It would be interesting to see what the relation is between the conditionally free cumulants of that paper (which are scalar-valued) and our $\text{St}_\pi$; cf. also [M].

**Example 3.** Let $\pi$ be as in the theorem, and $\{X(t)\}$ be the free Brownian motion. Then the free cumulants of $\mu_t$ are $r_1(t) = 0$, $r_2(t) = t$, $r_n(t) = 0$ for $n > 2$, and the diagonal measures of $X$ are $\Delta_1(t) = X(t)$, $\Delta_2(t) = t$, $\Delta_n(t) = 0$ for $n > 2$. Therefore the theorem states that
\[
\text{St}_\pi(X, t) = \begin{cases} 
0 & \text{if } \pi \text{ contains a class of more than 2 elements,} \\
0 & \text{if } \pi \text{ contains an inner singleton,} \\
t^{i(\pi)} \#\{B_j: |B_j|=2\} \psi_{\#\{B_j: |B_j|=1\}}(t) & \text{otherwise} \\
t^{\#\{V \in \pi: |V|=2\}} \psi_{\#\{V \in \pi: |V|=1\}}(t) & \forall i, j, |C_i| = 2, |B_j| = 1, 2, \\
0 & \text{otherwise.}
\end{cases}
\]

**Example 4.** Let $\pi$ be as in the theorem, and $\{X(t)\}$ be the free Poisson process. Then for $n \geq 1$, $r_n(t) = t$, $\Delta_n(t) = X(t)$. Therefore the theorem states that
\[
\text{St}_\pi(t) = t^{i(\pi)} \psi_{o(\pi)}(t).
\]
3. Proof of the theorem.

We start the analysis with a single free Poisson stochastic measure. It has the following remarkable representation: One can take $I \mapsto X(I)$ to be $sp(I)s$, where $s$ is a variable with a semicircular distribution freely independent from $p(I)$, and $I \mapsto p(I)$ is a projection-valued measure, such that disjoint intervals correspond to orthogonal projections, and $\varphi[p(I)] = |I|$. 

**Lemma 2.** Given a subdivision $\mathcal{S} = (I_1, I_2, \ldots, I_N)$ of $[0, 1)$, let $\{p_i\}_{i=1}^N$ be orthogonal projections adding up to 1 with $\varphi[p_i] = |I_i|$. Let $\{Z_{i,j}\}_{i \in [N], j \in [k]}$ be a family of operators (dependent on $\mathcal{S}$) such that for each $i$, the family $\{Z_{i,j}\}_{j \in [k]}$ is freely independent from $p_i$. In addition, assume that for each $i$, at least one of $Z_{i,j}$ is centered, and that for all $i, j, \mathcal{S}$, $\|Z_{i,j}\| < c/16$. Then

$$\lim_{\delta(\mathcal{S}) \to 0} \sum_{i=1}^N p_i Z_{i,1} p_i Z_{i,2} \ldots Z_{i,k} p_i = 0,$$

where the limit is taken in the operator norm.

**Proof.** Denote

$$A(\mathcal{S}) = \sum_{i=1}^N p_i Z_{i,1} p_i Z_{i,2} \ldots Z_{i,k} p_i = \sum_{i=1}^N (p_i, \ldots, p_i) \circ Z_i,$$

where $Z_i = (Z_{i,1}, \ldots, Z_{i,k})$. Then

$$(A(\mathcal{S}) A(\mathcal{S})^*)^n = \sum_{i=1}^N (p_i, \ldots, p_i) \circ (Z_i, (Z_i)^*, \ldots, Z_i, (Z_i)^*),$$

where $(Z_i)^* = (Z_{i,k}^*, \ldots, Z_{i,1}^*)$. Then as in [A, Theorem 3],

$$\varphi[(A(\mathcal{S}) A(\mathcal{S})^*)^n] = \sum_{i=1}^N \sum_{\pi \in NC(2nk)} R_K(\pi) (Z_i, (Z_i)^*, \ldots, Z_i, (Z_i)^*) \cdot M_\pi(p_i, p_i, \ldots, p_i).$$

At least one $Z_{i,j}$ is centered, so for partitions that contribute to the sum, at least $n$ elements do not belong to singleton classes of $K(\pi)$. Therefore $|K(\pi)| \leq 2nk - n$ and $|\pi| \geq (n + 1)$ (since $|\pi| + |K(\pi)| = 2nk + 1$). Thus

$$\varphi[(A(\mathcal{S}) A(\mathcal{S})^*)^n] \leq \sum_{i=1}^N \sum_{\pi \in NC(2nk) \atop |\pi| \geq n+1} c^{2nk} \varphi[p_i] |\pi| \leq 4^{2nk} c^{2nk} \delta(\mathcal{S})^n.$$

Therefore

$$\|A(\mathcal{S})\|_{2n} < \delta(\mathcal{S})^{1/2k} 4^k c^k,$$

and so $\|A(\mathcal{S})\| < \delta(\mathcal{S})^{1/2k} (4c)^k$, which converges to 0 as $\delta(\mathcal{S}) \to 0$. \qed
Lemma 3. Let $I \mapsto X(I) = \text{sp}(I)s$ be a free Poisson stochastic measure (see [A]), so that $X_i = \text{sp}_i s$ with $p_i$ as in Lemma 2. Let $Z_{i,j}$ be as in Lemma 2. Then

$$\lim_{\delta(S) \to 0} \sum_{i=1}^{N} X_i Z_{i,1} X_i Z_{i,2} \ldots Z_{i,k} X_i = 0.$$ 

Proof. By the free independence assumption on $\{Z_{i,j}\}$, the joint distribution of $\{X_i, Z_{i,1}, Z_{i,2}, \ldots, Z_{i,k}\}$ is, for each $i$, entirely determined by the distribution of $X_i$ and the joint distribution of $\{Z_{i,j}\}_{j=1}^{k}$. These distributions, and hence the conclusion of the lemma, are not changed if we assume in addition that the family $\{Z_{i,j}\}_{j=1}^{k}$ is freely independent from $s$.

$$(X_i, \ldots, X_i) \circ Z_i = (s p_s, \ldots, s p_s) \circ Z_i = s((p_1, \ldots, p_i) \circ (sZ_i s))s$$

where $sZ_i s$ denotes the vector $Z_i$ with each term multiplied by $s$ on both sides. Since $s$ and $Z_i$ are freely independent, if $Z_{i,j}$ is centered then so is $s Z_{i,j} s$. Then Lemma 2 implies the result. □

Lemma 4. Let $\pi \in NC(k)$ have only one outer class $B$ consisting of $n + 1$ elements. That is,

$$\pi = \left\{(u_0 = 1, u_1, u_2, \ldots, u_n = k), \pi(1), \pi(2), \ldots, \pi(n)\right\},$$

where $\pi(j)$ is supported on $C_j = [u_{j-1} + 1, u_j - 1]$. Let $X$ be a free Poisson stochastic measure. Then for $X$ with all $X^{(i)} = X$,

$$\text{St}_\pi(X) = \prod_{j=1}^{n} R_{\pi(j)}(C_j; X) \cdot \Delta(B; X).$$

Proof. Let $Z_{i,j}(N) = \sum_{\pi \in ([N] \setminus \{i\})_{\pi(j)}} X_{(C_j; \pi)}$. Then

$$\sum_{i=1}^{N} X_i^{(u_0)} Z_{i,1} X_i^{(u_1)} Z_{i,2} \ldots Z_{i,n} X_i^{(u_n)} = \sum_{i=1}^{N} \sum_{G \subset \{n\}} \left(X_i^{(u_0)}, \ldots, X_i^{(u_n)}\right) \circ Z(G),$$

where $Z(G)$ is a vector of length $n$ such that

$$Z(G)_j = \begin{cases} Z_{i,j} - \varphi[Z_{i,j}], & j \in G, \\ \varphi[Z_{i,j}], & j \not\in G. \end{cases}$$

For $G \neq \emptyset$, at least one of the $Z(G)_j$ is centered. So Lemma 3 applies and the limit, as $\delta(S)$ goes to 0, of the appropriate term is 0. On the other hand, it follows from [A, Corollary 2] that for any $i$ the limit of $\varphi[Z_{i,j}]$ is
We conclude that, denoting \( t = (i, i, \ldots, i) \),

\[
\lim_{\delta(S) \to 0} \sum_{i=1}^{N} X_{i,1}^{(u_0)} Z_{i,2} \ldots Z_{i,n} X_{i}^{(u_n)} = \prod_{j=1}^{n} \Pr_{\pi(j)}(X) \cdot \lim_{\delta(S) \to 0} \sum_{i=1}^{N} X_{(B; \tilde{I})}
\]

where \( \Delta(B; X) \) is well-defined for the free Poisson stochastic measure by [A, Corollary 4].

On the other hand,

\[
\sum_{i=1}^{N} X_{i,1}^{(u_0)} Z_{i,2} \ldots Z_{i,n} X_{i}^{(u_n)} = \sum_{\sigma} \St_{\sigma}(X, S),
\]

where the sum is taken over all partitions \( \sigma \) of \([k]\) which contain the class \( B \) and such that for all \( j \), the restriction of \( \sigma \) to \( C_j \) is \( \pi(j) \). The only noncrossing partition satisfying these requirements is \( \pi \), so

\[
\lim_{\delta(S) \to 0} \sum_{\sigma} \St_{\sigma}(X, S) = \St_{\pi}(X).
\]

\[\square\]

**Notation 4.** Let \( \pi \in NC(k) \). Then \( \pi \) can be written as

\[
\pi = (B_1(\pi), B_2(\pi), \ldots, B_o(\pi)(\pi), I_1(\pi), I_2(\pi), \ldots, I_o(\pi)(\pi)).
\]

Here, \( \{B_i(\pi)\} \) are outer classes of \( \pi \), listed in increasing order. Denote \( B_i = \{j| \exists a, b \in B_i : a \leq j \leq b\} \), the subset covered by \( B_i \), and let \( I_i(\pi) \) be the restriction of \( \pi \) to the set \( B_i \backslash B_{i} \) strictly covered by \( B_i \). Denote by \( I_i(\pi) \) the noncrossing partition \( (B_i(\pi), I_i(\pi)) \). Finally, denote \( C(\pi) = [k] \backslash \bigcup_{i=1}^{o(\pi)} B_i \), and let \( I(\pi) \) be the partition consisting of all inner classes of \( \pi \), i.e., the restriction of \( \pi \) to \( C(\pi) \).

**Lemma 5.** With the above notation, for a consistent \( k \)-tuple \( X \) of free stochastic measures,

\[
\Pr_{\pi}(X) = \prod_{\sigma \in I(\pi)} \Pr_{\sigma(\pi)}(X).
\]

**Proof.** Such a product decomposition is valid for any subdivision \( S \). \[\square\]

**Lemma 6.** Let \( X \) be a consistent \( k \)-tuple of free stochastic measures. Then:

1) The measures \( \Pr_{\pi}(X) \) and \( \St_{\pi}(X) \) are related as follows: For \( \pi \in NC(k) \),

\[
\Pr_{\pi}(X) = \sum_{\sigma \in NC(k)} \St_{\sigma}(X),
\]

\[
\St_{\pi}(X) = \sum_{\sigma \in NC(k)} \text{Möb}(\pi, \sigma) \Pr_{\sigma}(X).
\]
2) Let \( \pi_1, \pi_2, \ldots, \pi_n \) be noncrossing partitions such that \( \pi = \pi_1 + \pi_2 + \cdots + \pi_n \in NC(k) \). For each \( i \), identify \( \pi_i \) with a sub-partition of \( \pi \), and let \( C_i \) be the support of \( \pi_i \) in \([k]\). Denote \( \tau \in NC(k) \) the partition \((C_1, C_2, \ldots, C_n)\). Then

\[
\prod_{i=1}^{n} \text{St}_{\pi_i}(C_i; X) = \sum_{\sigma \in NC(k)} \text{St}_{\sigma}(X).
\]

**Proof.** Statement 1) is based on a purely combinatorial observation that

\[
\text{Pr}_{\pi}(X, S) = \sum_{\sigma \in P(k)} \text{St}_{\sigma}(X, S)
\]

and the fact that \( \text{St}_{\sigma}(X) = 0 \) for \( \sigma \notin NC(k) \); see Corollary 1 of \([A]\). Statement 2) is based on a purely combinatorial observation that

\[
\prod_{i=1}^{n} \text{St}_{\pi_i}((C_i; X), S) = \sum_{\sigma \in P(k)} \text{St}_{\sigma}(X, S)
\]

and the same fact. \( \square \)

**Lemma 7.** The limit defining \( \text{Pr}_{\pi}(X) \) exists in norm if the corresponding limit exists for the free Poisson stochastic measure.

**Proof.** Let \( S = (I_1, I_2, \ldots, I_N) \) be a subdivision of \([0, 1]\). Let \( T \) be another such subdivision, and let \( S \wedge T = (J_1, J_2, \ldots, J_M) \) be their common refinement. Temporarily denote by \( p(s) \) the index \( i \) such that \( J_s \subset I_i \). Denote \( A(S) = \text{Pr}_{\pi}(X, S) \), and similarly for \( S \wedge T \).

\[
A(S) - A(S \wedge T) = \sum_{v_1} \sum_{p^{-1}(s_1) = v_1} X_{s_1}^{(1)}(S \wedge T) \cdots \sum_{p^{-1}(s_k) = v_k} X_{s_k}^{(k)}(S \wedge T)
\]

\[
= \sum_{\pi \in [M]^k_{\geq \sigma}} X_{\pi}(S \wedge T).
\]

The above expression \( A(S) - A(S \wedge T) \) is a sum with positive coefficients. Hence so is \(((A(S) - A(S \wedge T)) (A(S) - A(S \wedge T))^*)^n\). Therefore its expectation is a sum over a collection of indices, with weights given by products of \(|J_s|\), all of which are independent of the distribution of \( X \), of free cumulants of \( X \) of order \( 2kn \). Each of those free cumulants is bounded in norm by \((16 \|X\|)^{2kn}\), where \( \|X\| = \max_i \|X^{(i)}\| \). Since for the free Poisson process all such cumulants are equal to 1, the result is at most \((16 \|X\|)^{2kn}\) times
the corresponding quantity for the free Poisson process, for which we denote $\Pr_\pi(X,S)$ by $a(S)$. That is,
\begin{align*}
\varphi\left((A(S) - A(S \land T))(A(S) - A(S \land T))^n\right) \\
\leq (16 \|X\|^{2n_k} \varphi\left((a(S) - a(S \land T))(a(S) - a(S \land T))^n\right),
\end{align*}
and so
\begin{equation*}
\|A(S) - A(S \land T)\|_{2n} \leq (16 \|X\|)^k \|a(S) - a(S \land T)\|_{2n},
\end{equation*}
which implies in particular that
\begin{equation*}
\|A(S) - A(S \land T)\| \leq (16 \|X\|)^k \|a(S) - a(S \land T)\|.
\end{equation*}
By assumption, the net $a(S)$ converges in norm, and
\begin{equation*}
\|A(S) - A(T)\| \leq \|A(S) - A(S \land T)\| + \|A(T) - A(S \land T)\|
\leq (16 \|X\|)^k (\|a(S) - a(S \land T)\| + \|a(T) - a(S \land T)\|).
\end{equation*}
Therefore the net $A(S)$ is a Cauchy net, and so converges. \hfill \Box

**Corollary 8.** \(\Pr_\pi(X),\) and hence \(\St_\pi(X),\) is well-defined for all \(\pi, X.\)

**Proof.** Let \(X\) be a free Poisson stochastic measure. By Lemma 4, \(\St_\pi(X)\) is well-defined for \(\pi \in NC(k)\) with a single outer class. Since for such \(\pi\) and \(\sigma \in NC(k), \sigma \geq \pi, \sigma\) also contains only one outer class, by Lemma 6 Part (1) we conclude that for such \(\pi, \Pr_\pi(X)\) is well-defined as well. By Lemma 5, \(\Pr_\pi(X)\) is then well-defined for an arbitrary \(\pi \in NC(k),\) and applying Lemma 6 Part (1) again implies that \(\St_\pi(X)\) is well-defined for an arbitrary \(\pi\) as well. Finally, by Lemma 7 the same is true for an arbitrary consistent \(k\)-tuple \(X\) of free stochastic measures. \hfill \Box

**Corollary 9.** Let \(X\) be a consistent \(k\)-tuple of free stochastic measures. For an interval \(I,\) define \(\Delta(X)(I) = \lim_{\delta(S) \to 0} \St_\delta(X,S),\) where \(S\) is a subdivision of \(I\) in place of \([0,1).\) With this notation, \(\Delta(X)\) is a free stochastic measure.

**Lemma 10.** Let \(X\) be a consistent \(k\)-tuple of free stochastic measures. Let \(G_1, G_2, \ldots, G_m \subset [k],\) and denote \(X_G = (X^{(u_1)}, X^{(u_2)}, \ldots X^{(u_{|G|})})\) for \(G = (u_1 < u_2 < \cdots < u_{|G|}).\) Then the \(m\)-tuple
\begin{equation*}
(\Delta(X_{G_1}), \Delta(X_{G_2}), \ldots, \Delta(X_{G_m}))
\end{equation*}
is also consistent.

**Proof.** The free increments property and stationarity follow immediately from the corresponding properties of \(X.\) For a general \(n\)-tuple \(Y\) of free stochastic measures that has these two properties, by stationarity the continuity property is equivalent to the continuity of the function
\begin{equation*}
t \mapsto \varphi[Y^{(v_1)}(t) \ldots Y^{(v_n)}(t)]
\end{equation*}
for all $t, \mathbf{v}$. By Möbius inversion, this is equivalent to the continuity of $t \mapsto R(Y^{(v_1)}(t), \ldots, Y^{(v_l)}(t))$ for all $t, \mathbf{v}$. By additivity and the free increments property, this is equivalent to the continuity of this function, for all $\mathbf{v}$, at $t = 0$, and so to the same property for $M$.

Thus finally, for the $m$-tuple in the hypothesis, it suffices to prove that

$$\varphi[\Delta(X_{G_1}, t) \Delta(X_{G_2}, t) \ldots \Delta(X_{G_m}, t)] \to 0 \text{ as } t \to 0.$$ 

Note that we do not need to put in a multi-index $\mathbf{v}$ since $\{G_i\}_{i=1}^m$ is already an arbitrary collection of subsets of $[k]$. Denote by $\sigma \in NC(l)$ the partition $(B_1, B_2, \ldots, B_m)$ with interval classes

$$B_j = \left\{ \left( \sum_{s=1}^{j-1} |G_s| \right) + 1, \ldots, \sum_{s=1}^j |G_s| \right\},$$

and let $Y = (X_{G_1}, X_{G_2}, \ldots, X_{G_m})$. Clearly $Y$ is a consistent $l$-tuple. Then

$$\varphi[\Delta(X_{G_1}, t) \Delta(X_{G_2}, t) \ldots \Delta(X_{G_m}, t)] = \varphi[\Pr_{\sigma}(Y(t))] = \sum_{\tau \geq \sigma} \varphi[St_{\tau}(Y(t))]$$

$$= \sum_{\tau \geq \sigma} R_{\tau}(Y(t)) = \sum_{\tau \geq \sigma} |\tau|^l R_{\tau}(Y)$$

by Lemma 1, and so goes to 0 as $t \to 0$. □

Lemma 11. Let $\sigma = (B_1, B_2, \ldots, B_n)$ be an interval partition of $[k]$. Then

$$\Delta(\Delta(B_1; X), \ldots, \Delta(B_n; X)) = \Delta(X).$$

Proof. Let $S = (I_1, I_2, \ldots, I_N)$ be a subdivision of $[0,1)$. For each $i$, let $S_i = (I_{i,1}, I_{i,2}, \ldots, I_{i,M_i})$ be a subdivision of $I_i$, and $T$ be the subdivision of $[0,1)$ obtained by combining $\{S_i\}_{i=1}^N$. Then as $\delta(S_1), \ldots, \delta(S_N) \to 0$, also $\delta(T) \to 0$. Therefore

$$\lim_{\delta(S_1), \ldots, \delta(S_N) \to 0} \Delta(X, T) = \Delta(X),$$

and so $\Delta(X)$ is also the limit of the left-hand-side if in addition $\delta(S) \to 0$. Here

$$\Delta(X, T) = \sum_{i=1}^N \sum_{s=1}^{M_i} \prod_{t=1}^k X_{i,s}^{(t)}(I_{i,s}).$$

On the other hand,

$$\sum_{i=1}^N \prod_{j=1}^n \prod_{s=1}^{M_i} X_{i,s}^{(t)}(I_{i,s}) = \sum_{i=1}^N \prod_{j=1}^n \Delta((B_j; X), S_i)$$
and
\[
\lim_{\delta(S) \to 0} \lim_{\delta(S_1), \ldots, \delta(S_N) \to 0} \sum_{i=1}^{N} \prod_{j=1}^{n} \Delta ((B_j; X), S_i) = \lim_{\delta(S) \to 0} \sum_{i=1}^{N} \prod_{j=1}^{n} \Delta (B_j; X) = \Delta (\Delta (B_1; X), \ldots, \Delta (B_n; X)).
\]

Therefore the difference
\[
\Delta (\Delta (B_1; X), \ldots, \Delta (B_n; X)) - \Delta (X)
\]
is the limit, as \(\delta(S_1), \ldots, \delta(S_N) \to 0\) and then as \(\delta(S) \to 0\), of
\[
\sum_{i=1}^{N} \prod_{j=1}^{n} \prod_{s=1}^{M_i} X(t)(I_{i,s}) - \sum_{i=1}^{N} \prod_{s=1}^{M_i} \prod_{t \in B_j} X(t)(I_{i,s}).
\]
This expression is a sum with positive coefficients. Also, for the free Poisson process,
\[
\Delta (\Delta (B_1; X), \ldots, \Delta (B_n; X)) - \Delta (X) = \Delta (X, \ldots, X) - X = 0.
\]
By the same estimates as in Lemma 7, the result follows. \(\square\)

**Lemma 12.** For \(\pi \in NC(k)\),
\[
\St_\pi (X) = R_{\mathcal{I}(\pi)} (C(\pi); X) \cdot \St_{(B_1(\pi), B_2(\pi), \ldots, B_{o(\pi)}(\pi))} \left( \bigcup_{i=1}^{o(\pi)} B_i(\pi); X \right).
\]

**Proof.** Let \(C\) be an inner class of \(\pi\), and let \(\pi' \in NC(k - |C|)\) be the restriction of \(\pi\) to \([k]\setminus C\). Then it suffices to prove that
\[
\St_\pi (X) = R(C; X) \cdot \St_{\pi'}(([k]\setminus C); X).
\]
Denote \(A = \St_\pi (X) - R(C; X) \cdot \St_{\pi'}(([k]\setminus C); X)\).

\[
\varphi ((AA^*)^n) = \sum_{G \subset [2n]} (-R(C; X))^{|G|} \varphi [\St_{\pi_1} (X_1) \St_{\pi_2} (X_2) \ldots \St_{\pi_{2n}} (X_{2n})],
\]
where:
- If \(j \notin G\), \(j\) odd, then \(\pi_j = \pi, X_j = X\).
- If \(j \notin G\), \(j\) even, then \(\pi_j = \pi^{op}, X_j = X^{op}\).
- If \(j \in G\), \(j\) odd, then \(\pi_j = \pi', X_j = ([k]\setminus C); X\).
- If \(j \in G\), \(j\) even, then \(\pi_j = (\pi')^{op}, X_j = ([k]\setminus C); X)^{op}\).

Denote \(\pi_G = \pi_1 + \pi_2 + \cdots + \pi_{2n}\). Let \(C_i(G)\) be the support of \(\pi_i\) identified as a sub-partition of \(\pi_G\), and let \(\tau_G = (C_1(G), C_2(G), \ldots, C_{2n}(G))\). Then
by Part (2) of Lemma 6,
\[
\|A\|_{2n}^2 = \sum_{G \subseteq [2n]} (-1)^{|G|} R(C; X)^{|G|} \sum_{\sigma \in NC(2nk-|G|\cdot |C|)} \varphi[\text{St}_\sigma(X_1, X_2, \ldots, X_{2n})]
\]
\[
= \sum_{G \subseteq [2n]} (-1)^{|G|} \sum_{\sigma \in NC(2nk-|G|\cdot |C|)} R(C; X)^{|G|} R(\sigma(X_1, X_2, \ldots, X_{2n})).
\]

Fix \( G \subseteq [2n] \). Let \( \sigma \in NC(2nk) \), \( \sigma \land \tau_0 = \pi_0 \), where \( \tau_0 = \hat{1}_k + \cdots + \hat{1}_k \) and \( \pi_0 = \pi + \pi^{\text{op}} + \pi + \cdots + \pi^{\text{op}} \). Denote \( C^{\text{op}} = (k + 1 - C) \) the class of \( \pi^{\text{op}} \) corresponding to \( C \). Since \( C \) is an inner class of \( \pi \), the condition \( \sigma \land \tau_0 = \pi_0 \) implies that \( (2jk + C) \) and \( ((2j + 1)k + C^{\text{op}}) \) are classes of \( \sigma \) for \( 0 \leq j < n \). Let \( g_G \) map such a \( \sigma \) to the partition in \( NC(2nk-|G|\cdot |C|) \) obtained by removing from \( \sigma \) the classes \( (2jk + C) \) for \( (2j + 1) \in G \) and \( ((2j + 1)k + C^{\text{op}}) \) for \( (2j + 2) \in G \). It is easy to see that \( g_G \) is a bijection onto \( \{ \sigma \in NC(2nk-|G|\cdot |C|) | \sigma \land \tau_G = \pi_G \} \), and that
\[
R(C; X)^{|G|} R(g_G(\sigma))(X_1, X_2, \ldots, X_{2n}) = R_{\sigma}(X, X, \ldots, X).
\]
Therefore
\[
\|A\|_{2n}^2 = \sum_{G \subseteq [2n]} (-1)^{|G|} \sum_{\sigma \in NC(2nk) \land \tau_0 = \pi_0} R_{\sigma}(X, X, \ldots, X) = 0
\]
since the first sum equals to 0.

\[\square\]

**Proof of the Main Theorem.** The statement of the theorem holds for \( \pi = \hat{0}_k \). From now on, assume \( \pi > \hat{0}_k \). The proof will proceed by induction on \( k \). The statement of the theorem is vacuous for \( k = 1 \); assume that it holds for all tuples of less than \( k \) elements.

By Lemma 12,
\[
\text{St}_{I_\pi}(\overline{B_i(\pi)}; X) = R_{I_\pi(\pi)}(\overline{B_i(\pi)} \setminus B_i(\pi); X) \cdot \Delta(B_i(\pi); X).
\]
Therefore
\[
\Pr_{I_\pi(\pi)}(\overline{B_i(\pi)}; X) = \sum_{\sigma_1 \geq I_\pi(\pi)} \text{St}_{\sigma_1}(\overline{B_i(\pi)}; X)
\]
\[
= \sum_{\sigma_1 \geq I_\pi(\pi)} (R_{I_\pi(\sigma_1)}(\overline{B(\sigma_1)} \setminus B(\sigma_1); X) \cdot \Delta(B(\sigma_1); X)).
\]
Then by Lemma 5,
\[ \Pr_\pi(X) = \prod_{i=1}^{o(\pi)} \Pr_{I_i(\pi)}(B_i(\pi); X) \]
\[ = \prod_{i=1}^{o(\pi)} \sum_{\sigma \geq \pi_i} (R_{I_i(\sigma)}(\mathcal{B}(\sigma_i) \setminus B(\sigma_i); X) \cdot \Delta(B(\sigma_i); X)) \]
\[ = \sum_{\sigma \geq \pi} R_{I(\sigma)}(C(\sigma); X) \prod_{j=1}^{o(\pi)} \Delta(B_j(\sigma); X) \]
\[ \times \Pr_{\hat{0}_{o(\pi)}}(\Delta(B_1(\sigma); X), \ldots, \Delta(B_{o(\pi)}(\sigma); X)) \]

In its turn,
\[ \Pr_{\hat{0}_{o(\pi)}}(\Delta(B_1(\sigma); X), \ldots, \Delta(B_{o(\pi)}(\sigma); X)) \]
\[ = \sum_{\rho \in NC(o(\pi))} \St_\rho(\Delta(B_1(\sigma); X), \ldots, \Delta(B_{o(\pi)}(\sigma); X)). \]

Since \( \pi > \hat{0}_k, \sigma \) has at most \( k - 1 \) classes, so the induction hypothesis applies to \( Y = (\Delta(B_1(\sigma); X), \ldots, \Delta(B_{o(\pi)}(\sigma); X)) \). Thus
\[ (2) \quad \St_\rho(Y) = R_{I(\rho)}(C(\rho); Y) \cdot \psi(\Delta(B_1(\rho); Y), \ldots, \Delta(B_{o(\rho)}(\rho); Y)). \]

Define the map \( f : NC(o(\pi)) \times \{ \sigma \in NC(k)|\sigma \geq \pi, \forall i : B_i(\sigma) = B_i(\pi) \} \) \( \rightarrow \) NC\( (k) \) by \( i \stackrel{f(\rho, \sigma)}{\sim} j \Leftrightarrow ((i^\sigma \sim j) \text{ or } (i \in B_s(\sigma), j \in B_t(\sigma), s \sim \tau)) \). Note that the outer classes of \( f(\rho, \sigma) \) are in one-to-one correspondence with the outer classes of \( \rho \), and each inner class of \( f(\rho, \sigma) \) corresponds to a unique inner class of either \( \rho \) or \( \sigma \). It is easy to see that \( f \) is in fact a bijection onto \( \{ \tau \in NC(k)|\tau \geq \pi \} \). Combining Equation (2) with Lemma 11, we see that
\[ R_{I(\sigma)}(C(\sigma); X) \cdot \St_\rho(\Delta(B_1(\sigma); X), \ldots, \Delta(B_{o(\pi)}(\sigma); X)) \]
\[ = R_{I(\tau)}(C(\tau); X) \cdot \psi(\Delta(B_1(\tau); X), \ldots, \Delta(B_{o(\tau)}(\tau); X)), \]

with \( \tau = f(\rho, \sigma) \). Therefore
\[ \Pr_\pi(X) = \sum_{\tau \geq \pi} R_{I(\tau)}(C(\tau); X) \]
\[ \times \psi(\Delta(B_1(\tau); X), \Delta(B_2(\tau); X), \ldots, \Delta(B_{o(\tau)}(\tau); X)). \]
On the other hand, for all \( \pi \), \( \Pr_\pi(X) = \sum_{\pi \geq \pi} \text{St}_\pi(X) \). Note that the Möbius inversion formula for \( \pi > \hat{0}_k \) involves only \( \sigma > \hat{0}_k \). Therefore, applying this formula,

\[
\text{St}_\pi(X) = R_{\mathcal{I}(\pi)}(C(\pi); X) \\
\times \psi(\Delta(B_1(\pi), X), \Delta(B_2(\pi); X), \ldots, \Delta(B_{o(\pi)}(\pi); X)) \\
= \prod_{i(\pi)} R(C_i; X) \cdot \psi(\Delta(B_1; X), \Delta(B_2; X), \ldots, \Delta(B_{o(\pi)}; X)).
\]

Remark 5 (Higher-dimensional analogs). The Main Theorem gives a complete description of the higher stochastic measures \( \text{St}_\pi \) as given in Definition 3. However, under the original definitions of [RW] (modified for processes with freely independent increments) these only correspond to values on cubes, hence their dependence on only 1 and not \( k \) parameters. In this remark we briefly describe how one could extend the definition to more general rectangles of the form \( \mathbf{I} = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_k, b_k] \). It is clear that it suffices to give the definition only for the case when for \( 1 \leq i, j \leq k \) the intervals \([a_i, b_i]\) and \([a_j, b_j]\) are either disjoint or the same (one then needs to show that the resulting definition is consistent).

Assume that the rectangle \( \mathbf{I} \) is of this form. Then we can define a partition \( \pi(\mathbf{I}) \in \mathcal{P}(k) \) by \( i \sim^\pi j \Leftrightarrow [a_i, b_i] = [a_j, b_j] \). Let \( \pi(\mathbf{I}) \) have classes \( B_1, B_2, \ldots, B_l \). Let \( c(i) \) be the index such that \( i \in B_{c(i)}, 1 \leq i \leq k \).

Let \( X \) be a free stochastic measure, and \( X \) a \( k \)-tuple of free stochastic measures given by \( X^{(j)}([a, b]) = X([a - a_j, b - a_j]) \). The conditions on \( a_i, b_i \) imply that this \( k \)-tuple is consistent. Let \( S = \{S_j\} \) be subdivisions of \([a_{c-1}(j), b_{c-1}(j)]\), \( 1 \leq j \leq l \) into intervals \( I_{j,s} \). For \( \sigma \in \mathcal{NC}(k) \), denote \( S_\sigma = \left\{ \bar{s} \in \mathbb{N}^k : \left(\prod_{i=1}^k I_{c(i), \nu_i}\right) \cap \mathbb{R}_\sigma^k \neq \emptyset \right\} \). Note that if \( \pi(\mathbf{I}) = \hat{1} \) and \( S \) is a single subdivision with \( N \) classes, \( S_\sigma = [N]_\sigma^k \). Define

\[
\text{St}_\sigma(X, S) = \sum_{\bar{s} \in S_\sigma} \prod_{i=1}^k X(I_{c(i), \nu_i})
\]

and \( \text{St}(\mathbf{I}) = \lim_{\delta(S) \to 0} \text{St}_\sigma(X, S) \). It follows immediately that \( \text{St}(\mathbf{I}) = 0 \) unless \( \sigma \leq \pi(\mathbf{I}) \). Indeed, if \( \sigma \nleq \pi(\mathbf{I}) \) then for any subdivision \( S, S_\sigma = \emptyset \).

Acknowledgments. This paper stems from the work started during the special semester “Free probability and operator spaces” at the Institut Henri Poincaré – Centre Emile Borel, with the support from the Fannie and John Hertz Foundation; it was completed with the support from an NSF postdoctoral fellowship. The author is grateful to the referee for numerous comments and suggestions.
References


Received March 1, 2001 and revised July 13, 2001.

Department of Mathematics
University at California
Berkeley, CA 94720
E-mail address: manshel@math.berkeley.edu
GROUPS ACTING ON CANTOR SETS AND THE END STRUCTURE OF GRAPHS

BRIAN H. BOWDITCH

We describe a variation of the Bergman norm for the algebra of cuts of a connected graph admitting a cofinite group action. By a construction of Dunwoody, this enables us to obtain nested generating sets for invariant subalgebras. We describe a few applications, in particular, to convergence groups acting on Cantor sets. Under certain finiteness assumptions one can deduce that such actions are necessarily geometrically finite, and hence arise as the boundaries of relatively hyperbolic groups. Similar results have already been obtained by Gerasimov by other methods. One can also use these techniques to give an alternative approach to the Almost Stability Theorem of Dicks and Dunwoody.

0. Introduction.

In this paper, we describe some of the interconnections between the end structure of graphs, groups acting on protrees, and convergence actions on Cantor sets. Our work ties in with recent work of Gerasimov and Dunwoody and earlier work of Bergman.

It was shown by Hopf in the 1930s that the space of ends of (any Cayley graph of) a finitely generated group, \( \Gamma \), either consists of 0, 1 or 2 points, or else is a Cantor set. In the late 1960s, Stallings used topological methods to show that in the last case \( \Gamma \) splits nontrivially over a finite subgroup. Shortly afterwards, Bergman \( \text{[Ber]} \) gave a different proof of the same result. One can interpret his construction as giving us a \( \Gamma \)-invariant nested subset of the Boolean algebra of clopen sets of the space of ends, and hence a splitting of the group via Bass-Serre theory. In fact, one can think of this algebra combinatorially in terms of cuts of the Cayley graph, i.e., finite sets of edges which separate the graph into more than one infinite subset (cf. \( \text{[Du1]} \)). Bergman’s proof uses a certain “norm” defined on the set of such cuts. Recently Dunwoody and Swenson showed how one can use Bergman’s norm to construct nested generating sets of arbitrary invariant subalgebras of the algebra of cuts of a locally finite graph (see \( \text{[DuS]} \) and Section 8). In this paper we generalise Bergman’s norm, and some its consequences, to certain non-locally finite graphs (see Section 9). One of the main applications we
describe here will be to convergence group actions on Cantor sets. Similar results have already been obtained by Gerasimov [Ger], by other methods. One can also use these ideas to give a simplified proof of the Almost Stability Theorem of [DiD] (see Section 14).

The notion of a convergence group was defined by Gehring and Martin [GehM]. It extracts the essential dynamical features of a kleinian group acting on the boundary of (classical) hyperbolic space. This generalises to boundaries of proper hyperbolic spaces in the sense of Gromov [Gr], see for example, [T2, F, Bo3]. The celebrated result of [T1, Ga, CJ] tells us that every convergence action on the circle arises from a fuchsian group. (In particular, the work of Tukia [T1] deals with the case of cyclically ordered Cantor sets.) There has also been a study of convergence actions on more general continua and their applications to hyperbolic and relatively hyperbolic groups (see for example [Bo2] and the references therein). Here we turn our attention to the opposite extreme, namely convergence actions on Cantor sets. An extensive study of such actions in relation to median algebras can be found in [Ger]. An example of such an action is that of a finitely generated group acting on its space of ends, which is a Cantor set if the group is not virtually cyclic and splits nontrivially over a finite subgroup. We shall see that, under some finiteness assumptions on the group, this type of example is typical.

Via Stone duality, a convergence action of a group, $\Gamma$, on a Cantor set is equivalent to a certain kind of action on a Boolean algebra. To obtain a splitting of $\Gamma$ from this action, there are two issues to be addressed. The first is to find a $\Gamma$-invariant nested generating set for the Boolean algebra, and the second is to arrange that this set is cofinite (in other words to show that the algebra is finitely generated as a $\Gamma$-Boolean algebra). To get a handle on these issues, we assume that $\Gamma$ is (relatively) finitely generated, and work with the algebra of cuts of a connected $\Gamma$-graph given by this hypothesis. The nested generating set is obtained using the Bergman norm (see Sections 8 and 9). In [Ger], a similar result is obtained via the theory of median algebras. For the second part, we need to assume that $\Gamma$ is (relatively) almost finitely presented. We use a result of [DiD] or of [BesF] to obtain a cofinite generating set (Section 11). This is really an issue of accessibility. (We remark that it is shown in [DiD] that a finitely generated group, $\Gamma$, is accessible if and only if the Boolean algebra of almost invariant subsets is finitely generated as a $\Gamma$-Boolean algebra.)

I am indebted to Victor Gerasimov for explaining to me some of his work in this area. I have benefited greatly from discussions with Martin Dunwoody, in particular for introducing me to the work of Bergman, as well as explaining to me how accessibility results could be used to obtain finitely generated algebras (see Sections 10 and 11). Some of the work for the present article was carried out while visiting the Centre de Recerca Matemàtica in
Barcelona. I am grateful for the support and hospitality of this institution, and for helpful discussions with Warren Dicks while there.

1. Summary of results.

We shall introduce some of the results of this paper by giving a series of results, which are, in some sense, increasingly general, but which require greater elaboration in their formulation. They will be proven in Section 12. As mentioned in the introduction, most of these results have been obtained, in some form, by Gerasimov using median algebras [Ger]. The results will be made more precise in later sections. We shall assume throughout that the convergence actions are minimal, i.e., that there is no discontinuity domain. For definitions regarding convergence actions, see Section 4.

The simplest example of a convergence action on a Cantor set is that of a (virtually) free group acting on its space of ends (which is the same as its boundary as a hyperbolic group). Such an action has no parabolic points. One result says that this is typical:

**Theorem 1.1.** If $\Gamma$ is an almost finitely presented group acting as a minimal convergence group on a Cantor set, $M$, without parabolic points, then $\Gamma$ is virtually free, and there is a $\Gamma$-equivariant homeomorphism from $M$ to the space of ends of $\Gamma$.

By *almost finitely presented* we mean that $\Gamma$ admits a cocompact properly discontinuous action on a locally finite connected 2-complex, $\Sigma$, with $H_1(\Sigma; \mathbb{Z}_2) = 0$. (Without loss of generality, one can assume that the action on $\Sigma$ is also free.) Clearly, finite presentability in the usual sense implies almost finite presentability.

More generally, it is well-known that any almost finitely generated group acts as a convergence group on its space of ends (see Section 5). If $\Gamma$ is finitely presented, then Dunwoody’s accessibility theorem [Du2] tells us that we can represent $\Gamma$ as a finite graph of groups with finite edge groups and finite or one-ended vertex groups. The one-ended vertex groups are uniquely determined (up to conjugacy) as the maximal one-ended subgroups of $\Gamma$, and are precisely the maximal parabolic subgroups with respect to this convergence action. We have a converse:

**Theorem 1.2.** Suppose that $\Gamma$ is an almost finitely presented group and acts as a minimal convergence group on a Cantor set, $M$, with all maximal parabolic subgroups finitely generated and one-ended. Then, $M$ is equivariantly homeomorphic to the space of ends of $\Gamma$.

The hypothesis on parabolic subgroups in the above result is somewhat unnatural. Note that the splitting given by Dunwoody’s accessibility result gives us, via Bass-Serre theory, an action on a simplicial tree with finite edge stabilisers and finite quotient. To such a tree, we can associate a
“boundary” as described in Section 6. The group acts as a convergence group on this boundary, with the infinite vertex stabilisers as the maximal parabolic subgroups. If these happen to be one-ended, then we recover the space of ends of $\Gamma$. We also have a converse:

**Theorem 1.3.** Suppose that $\Gamma$ is an almost finitely presented group and acts as a minimal convergence group on a Cantor set, $M$. Then, $\Gamma$ has a representation as a finite graph of groups with finite edge groups such that $M$ is equivariantly homeomorphic to the “boundary” of the associated Bass-Serre tree.

In fact, we only really require that $\Gamma$ be almost finitely presented relative to a set of parabolic subgroups (see Section 2 for a definition). More precisely:

**Theorem 1.4.** Suppose that a group, $\Gamma$, acts as a minimal convergence group on a Cantor set, $M$. Suppose that $\mathcal{G}$ is a finite collection of parabolic subgroups of $\Gamma$ with respect to this action. Suppose that $\Gamma$ is almost finitely presented relative to $\mathcal{G}$. Then, the conclusion of Theorem 1.3 holds.

Note that, from the conclusion of Theorem 1.3, one may deduce that there are only finitely many conjugacy classes of maximal parabolic subgroups. In fact, it follows that the action of $\Gamma$ on $M$ is geometrically finite, and that $\Gamma$ is hyperbolic relative to the collection of maximal parabolic subgroups (see Section 4 for definitions). Moreover, the boundary of $\Gamma$ as a relatively hyperbolic group may be identified with the boundary of the Bass-Serre tree. Note that in Theorem 1.4, the splitting obtained is “relative to” $\mathcal{G}$, i.e., each element of $\mathcal{G}$ is conjugate into a vertex group.

We can further weaken the hypotheses, and assume only that $\Gamma$ is finitely generated (relative to a class of parabolic subgroups). However, in this case we can only be assured of an action of $\Gamma$ on a protree, as opposed to a simplicial tree. From this, one can deduce that the original action on $M$ is an inverse limit of geometrically finite actions of the above type. This is elaborated on in Section 13.

### 2. Finiteness conditions.

In this section, we elaborate on the finiteness conditions featured in Section 1. These are most conveniently expressed in terms of group actions on sets (cf. [Bo5]).

Suppose $\Gamma$ is a group. A $\Gamma$-set, $V$, is a set on which the group $\Gamma$ acts. Given $x \in V$, we write $\Gamma(x) = \{ g \in \Gamma \mid gx = x \}$. We write $V_\infty = \{ x \in V \mid |\Gamma(x)| = \infty \}$. We can interpret a property of $V$ as a property of the group, $\Gamma$ “relative to” the set of infinite point stabilisers, $\{ \Gamma(x) \mid x \in V_\infty \}$. We shall also speak of “$\Gamma$-graphs”, “$\Gamma$-trees” and “$\Gamma$-Boolean algebras” etc. for such objects admitting $\Gamma$-actions.
We shall say that the $\Gamma$-set, $V$, is cofinite if $|V/\Gamma|$ is finite. A pair stabiliser is a subgroup of the form $\Gamma(x) \cap \Gamma(y)$ for $x \neq y$. We shall say that $V$ is “0-connected” if it can be identified with the vertex set, $V(K)$, of a connected cofinite $\Gamma$-graph, $K$ (or equivalently, as a $\Gamma$-invariant subset of $V(K)$ containing $V_\infty(K)$). We say that $V$ is 1-connected if it is the vertex set of a cofinite simply connected CW-complex (or equivalently, a $\Gamma$-invariant subset containing $V_\infty(\Sigma)$ of the vertex set of a cofinite simplicial complex, $\Sigma$). Note that 1-connectedness is equivalent to the assertion that for some (or equivalently any) cofinite connected $\Gamma$-graph with vertex set $V$, there is some $n \geq 2$ such that $\Omega_n(K)$ is simply connected. Here $\Omega_n(K)$ is the 2-complex obtained by attaching a 2-cell along every circuit of length at most $n$ in $K$. More generally, we say that $V$ is $\mathbb{Z}_2$-homologically 1-connected if $\Omega_n(K)$ is $\mathbb{Z}_2$-acyclic, i.e., $H_1(\Omega_n(K); \mathbb{Z}_2) = 0$.

We say that a graph, $K$, is fine if, for each $n$, there are only finitely many circuits of length $n$ containing any given edge. Clearly, this implies that the complex $\Omega_n(K)$ described above is locally finite away from $V = V(K)$ (and can thus be subdivided to give a simplicial complex that is locally finite away from $V$). We say that a $\Gamma$-set is fine if some (hence any) cofinite connected $\Gamma$-graph with vertex set $V$, is fine. Here, the fineness of $K$ is equivalent to saying that there are finitely many circuits of any given length modulo $\Gamma$.

Note that if $V$ is fine and $\mathbb{Z}_2$-homologically 1-connected we can embed $V$ equivariantly in the vertex set, $V(\Sigma)$, of a cocompact simply connected ($\mathbb{Z}_2$-acyclic) 2-dimensional simplicial complex, $\Sigma$, which is locally finite away from $V(\Sigma)$, and such that the stabiliser of each element of $V(\Sigma) \setminus V$ is finite.

Suppose that $\mathcal{G}$ is a nonempty collection of self-normalising subgroups of $\Gamma$, which is a finite union of conjugacy classes, and such that the intersection of any two distinct elements of $\mathcal{G}$ is finite. We may view $\mathcal{G}$ as a $\Gamma$-set with $\Gamma$ acting by conjugation. We say that $\Gamma$ is finitely generated relative to $\mathcal{G}$ if $\mathcal{G}$ is 0-connected. We say that $\Gamma$ is (almost) finitely presented relative to $\mathcal{G}$ if $\mathcal{G}$ is $\mathbb{Z}_2$-homologically 1-connected. In the case of interest to us, $\mathcal{G}$ will always be fine.

This ties in with the usual group theoretical terminology. A group, $\Gamma$, is finitely generated if and only if it admits a cofinite action on a locally finite simplicial 1-complex. It is (almost) finitely presented if and only if it admits a cofinite action on a ($\mathbb{Z}_2$-homologically) 1-connected locally finite simplicial 2-complex. Here the actions can in fact be taken to be free.

If $G \subseteq \Gamma$ is self-normalising with presentation $\langle A; R \rangle$, then $\Gamma$ is finitely generated relative to $G$ if and only if there is a finite set $B \subseteq \Gamma$, such that $\langle A \cup B \rangle$. It is finitely presented relative to $G$, if and only if there is a finite set of words, $S$, such that $\langle A \cup B; R \cup S \rangle$ is a presentation of $\Gamma$. One can generalise this to a finite collection of self-normalising subgroups using the language of groupoids, though we shall not make that explicit here.

Let $B$ be a Boolean ring, i.e., a commutative ring with a unity element, 1, satisfying $x^2 = x$ for all $x \in B$. We write $x^* = 1 + x$, $x \wedge y = xy$ and $x \vee y = x + y + xy$. Thus, $(B, \wedge, \vee, *)$ is a Boolean algebra. We write $x \leq y$ to mean that $xy = x$. Thus $\leq$ is a partial order on $B$, and $[x \mapsto x^*]$ is an order reversing involution. Given any set, $X$, its power set, $\mathcal{P}(X)$ is a Boolean algebra with $\mathcal{P}^* = X \setminus P$, $P \wedge Q = P \cap Q$ and $P \vee Q = P \cup Q$, for $P, Q \in \mathcal{P}(X)$. Suppose $M$ is a compact totally disconnected topological space. The set, $B(M)$ of clopen subsets of $M$ is a Boolean subalgebra of $\mathcal{P}(M)$.

The Stone duality theorem [St] tells us that every Boolean algebra arises in this way (see for example [Si]). Suppose that $B$ is a Boolean algebra. We associate to $B$ a compact totally disconnected space $\Xi = \Xi(B)$, called the Stone dual, such that $B(\Xi) \cong B$. This can be described in a number of equivalent ways. For example we can define $\Xi$ as the set of Boolean ring homomorphisms from $B$ to $\mathbb{Z}_2$. This is a closed subset of the Tychonoff cube $\mathbb{Z}_2^B$, and we topologise $\Xi$ accordingly.

Alternatively, we define $\Xi$ as the maximal ideal spectrum of the ring $B$ with the Zariski topology. Note that the complement of a maximal ideal in $B$ is an ultrafilter. We therefore get the same thing by taking the set of ultrafilters on $B$. From this point of view, we can define a basis for the closed sets by taking a typical basis element to be the set of all ultrafilters that contain a given element of $B$.

Suppose that $B$ is a subalgebra of $\mathcal{P}(X)$ for some set $X$. If $x \in X$, then the ultrafilter $\mathcal{O}(x) = \{P \in B \mid x \in P\}$ determines a point of $\Xi(B)$, so we get natural map from $X$ to $\Xi(B)$. If $M$ is a compact and totally disconnected topological space, and $B = B(M)$, then this map gives us the Stone isomorphism from $M$ to $\Xi(B(M))$.

Note that if $f : B \rightarrow B'$ is a homomorphism, we get a continuous dual map, $f_* : \Xi(B') \rightarrow \Xi(B)$. If $f$ is surjective, then $f_*$ is injective, so we can identify $\Xi(B')$ as a closed subset of $\Xi(B)$.

We say that two nonzero elements $x, y \in B$ are nested if $xy$ is equal to 0, $x$, $y$ or $1 + x + y$. This is equivalent to saying that one of $xy$, $xy^*$, $x^*y$ or $x^*y^*$ equals 0. We say that a subset, $E \subseteq B$ is nested if $0, 1 \notin E$ and every pair of elements of $E$ are nested. Note that if $E$ is nested, then so is the $^*$-invariant set $E \cup E^*$.

4. Convergence groups.

The notion of a convergence group was defined in [GehM]. For further discussion, see [T2, Bo3, T3].

Suppose that $M$ is a compact metrisable space and that $\Gamma$ is a group acting by homeomorphism on $M$. We say that this is a convergence action
(or that $\Gamma$ is a *convergence group*) if the induced action on the space of distinct triples of $M$ (i.e., $M \times M \times M$ minus the large diagonal) is properly discontinuous. This is equivalent to the statement that if $(g_n)_{n \in \mathbb{N}}$ is any infinite sequence of distinct elements of $\Gamma$, then there are points, $a, b \in M$, and a subsequence $(g_n)_i$ such that $g_n|M \setminus \{a\}$ converges locally uniformly to $b$. We refer to the latter statement as the “convergence property” of $\Gamma$.

A subgroup, $G$, of $\Gamma$ is *parabolic* if it is infinite and fixes a unique point. Such a fixed point, $x$, is called a *parabolic point*, and its stabiliser, $\Gamma(x)$, is a maximal parabolic subgroup of $\Gamma$. We say that $x$ is a *bounded parabolic point* if $(M \setminus \{x\})/\Gamma(x)$ is compact. (We allow for the possibility of a parabolic group being an infinite torsion group.)

A point $x \in M$ is a *conical limit point* if there is a sequence of elements $g_n \in \Gamma$, and distinct points, $a, b \in M$ such that $g_n x \to a$ and $g_n y \to b$ for all $y \in M \setminus \{x\}$. It is shown in \cite{T3} that a conical limit point cannot be a parabolic point. We say that the action of $\Gamma$ on $M$ is *geometrically finite* if every point of $M$ is either a conical limit point or a bounded parabolic point. Such actions have been studied by Tukia \cite{T3}.

By the *space of distinct pairs* of $M$, we mean $M \times M$ minus the diagonal.

**Lemma 4.1.** Suppose that $\Gamma$ acts on $M$ as a convergence group, and that $x, y \in M$ are distinct and not conical limit points. Then, $\Gamma(x) \cap \Gamma(y)$ is finite, and the $\Gamma$-orbit of $(x, y)$ is a discrete subset of the space of distinct pairs.

**Proof.** Suppose, for contradiction, that $g_n \in \Gamma$ is a sequence of distinct elements of $\Gamma$ with $g_n x \to a$ and $g_n y \to b$ for all $y \in M \setminus \{x\}$. It is shown in \cite{T3} that a conical limit point cannot be a parabolic point. We say that the action of $\Gamma$ on $M$ is geometrically finite if every point of $M$ is either a conical limit point or a bounded parabolic point. Such actions have been studied by Tukia \cite{T3}.

By the *space of distinct pairs* of $M$, we mean $M \times M$ minus the diagonal.

**Lemma 4.2.** Suppose that $\Gamma$ acts on $M$ as a convergence group, and that $\Pi \subseteq M$ is a $\Gamma$-invariant subset. Suppose that no point of $\Pi$ is a conical limit point. If $\Pi$ is connected (as a $\Gamma$-set) then it is fine.

**Proof.** We show that if $K$ is any cofinite $\Gamma$-graph with vertex set $V(K) = \Pi$, then modulo $\Gamma$, there are only finitely many circuits of length $n$ for any given $n$. Since $\Pi$ is connected, we can take $K$ to be connected, and we see that $K$ is fine. Together with the first part of Lemma 4.1, this implies that $\Pi$ is fine as claimed.

Suppose, for contradiction, that $(\beta^k)_{k \in \mathbb{N}}$ is an infinite sequence of circuits of length $n$ in $K$, each lying in a different $\Gamma$-orbit. We write $\beta^k = x_1^k \ldots x_n^k$, taking subscripts mod $n$. Passing to a subsequence, we can suppose that for all $i$ each of the edges $\{x_i^k x_i^{k+1}\} \in K$ lie in the same $\Gamma$-orbit. Thus, modulo $\Gamma$, we can suppose that $x_0^k = x_0$ and $x_1^k = x_1$ are independent of $k$. Now
by Lemma 4.1, the set of pairs \( \{(x_1, x_2^k) \mid k \in \mathbb{N}\} \) is discrete in the space of distinct pairs. Thus, again after passing to a subsequence, we can suppose that either \( x_2^k = x_2 \) is constant, or that \( x_2^k \to x_1 \). In the latter case, we can suppose that the \( x_2^k \) are all distinct, so since the set of pairs \( (x_2^k, x_3^k)_k \) is discrete, we must also have \( x_3^k \to x_1 \). It follows inductively that \( x_1^k \to x_1 \) for all \( i \), giving the contradiction that \( x_0 = x_n^k \to x_1 \). We can thus assume that \( x_2^k = x_2 \) is constant. But now, the same argument tells us that \( x_3^k \) is constant, so by induction, \( x_i^k \) is constant for all \( i \). We derive the contradiction that \( \beta^k \) is constant. \( \square \)

Note that, by the result of Tukia [T3], Lemma 4.2 applies to a set of parabolic points.

A standard example of a convergence group is the action induced on the boundary, \( \partial X \), by any properly discontinuous action of a group, \( \Gamma \), on a proper (complete locally compact) hyperbolic space, \( X \). If the action on \( \partial X \) is geometrically finite, we say that \( \Gamma \) is hyperbolic relative to the set, \( \mathcal{G} \), of maximal parabolic subgroups of \( \Gamma \). (In [Bo5], we impose the additional requirement that each element of \( \mathcal{G} \) be finitely generated, but this need not concern us here.) It is necessarily the case that \( \mathcal{G} \) is cofinite and connected (and hence fine) as a \( \Gamma \)-set. Relatively hyperbolic groups were introduced by Gromov [Gr]. It turns out that they can be characterised dynamically as geometrically finite convergence groups [Y].

We note that, in the case where \( M \) is totally disconnected, we can express the convergence property in terms of the Boolean algebra, \( B = B(M) \) of clopen subsets of \( M \). By a ternary partition of \( B \), we mean a triple of pairwise disjoint nonzero elements, \( A, B, C \in B \) such that \( A + B + C = 1 \). (In other words, \( M = A \sqcup B \sqcup C \).) Note that \( A \times B \times C \) is a compact subset of the space of distinct triples of \( M \).

Suppose that \( \Gamma \) acts by isomorphism on a Boolean algebra \( B \).

**Definition.** We say that the action is a convergence action if, for any two ternary partitions, \( (a_1, a_2, a_3) \) and \( (b_1, b_2, b_3) \) of \( B \), the set \( \{g \in \Gamma \mid b_1 \land ga_1 \neq 0, b_2 \land ga_2 \neq 0, b_3 \land ga_3 \neq 0\} \) is finite.

To see that this agrees with the notion already defined for \( M \), note that compact subset of the set of distinct triples of \( M \) can be finitely covered by sets of the form \( A \times B \times C \), where \( A, B, C \) is a ternary partition of \( M \).

We finally note that if \( M \) is an inverse limit of compact spaces, \( (M_n)_{n \in \mathbb{N}} \), each admitting a \( \Gamma \)-action that commutes with the inverse limit system, then the action of \( \Gamma \) on \( M \) is a convergence action if and only if the action on each \( M_n \) is a convergence action.
5. Ends of graphs.

In this section, we explain how to associate a “space of ends”, $\Xi_0(K)$, to a connected graph, $K$.

Let $K$ be a connected graph, with vertex set $V = V(K)$, and edge set $E(K)$. Let $V_0 = V_0(K)$ be the set of vertices of finite degree. If $W \subseteq V(K)$ we write $I(W)$ for the set of edges with precisely one endpoint in $W$.

**Definition.** A subset $W \subseteq V(K)$ is a $K$-break if $I(W)$ is finite.

We write $B = B(K)$ for the Boolean algebra of $K$-breaks. We write $\Xi(K) = \Xi(B)$ for the Stone dual of $B$.

There is a natural map, $\xi : V \rightarrow \Xi(K)$, defined by sending $x \in V$ to the ultrafilter of elements of $B$ containing $x$. Note that $\xi|V_0$ is injective, and every point of $\xi(V_0)$ is isolated in $\Xi(K)$. We write $\Xi_0(K) = \Xi(K) \setminus \xi(V_0)$. This is a closed subset of $\Xi(K)$. Note that if $K$ is locally finite, then $\Xi_0(K)$ is the precisely the space of ends of $K$ in the usual sense.

Another way to define $\Xi_0(K)$ is as follows. Let $I$ be the ideal of $B$ consisting of finite subsets of $V$, and let $\mathcal{C}(K)$ be the quotient $B/I$. There is an inclusion of $\Xi(C)$ in $\Xi(B)$ whose image is precisely $\Xi_0(K)$.

We shall say that $K$ is one-ended if $\Xi_0(K)$ consists of a single point, i.e., no finite set of edges separates $K$ into two or more infinite subgraphs.

**Lemma 5.1.** If a group $\Gamma$ acts on a connected graph, $K$, with finite edge stabilisers, then the induced action of $\Gamma$ on $\Xi(K)$ is a convergence action.

**Proof.** Note that if $F \subseteq E(K)$ is any finite set of edges, then $\{g \in \Gamma \mid F \cap gF \neq \emptyset\}$ is finite.

Suppose that $(A_1, A_2, A_3)$ and $(B_1, B_2, B_3)$ are ternary partitions of $B$. Let $I = I(A_1) \cup I(A_2) \cup I(A_3)$ and $J = I(B_1) \cup I(B_2) \cup I(B_3)$. Let $X$ and $Y$ be any connected finite subgraphs of $K$ containing $I$ and $J$ respectively.

Suppose that $E(X) \cap E(Y) = \emptyset$. Now, $I \cap E(Y) = \emptyset$ and so all the vertices of $Y$ must lie in the same element of $\{A_1, A_2, A_3\}$, say $V(Y) \subseteq A_i$. Similarly, $V(X) \subseteq B_j$ for some $j$. We claim that $V(K) \subseteq A_i \cup B_j$. For suppose $x \in V(K) \setminus (A_i \cup B_j)$. Let $\alpha$ be a shortest path connecting $x$ to $A_i \cup B_j$. Let $y, z$ be, respectively, the last and last but one vertices of $\alpha$. Without loss of generality, $y \in A_i$. Now $z \notin A_i$, and so the edge connecting $y$ and $z$ must lie in $I \subseteq E(X)$. It follows that $z \in V(X) \subseteq B_j$, giving a contradiction, and hence proving the claim. Note that if $k \neq i, j$, then $A_k \cap A_i = B_k \cap B_j = \emptyset$, so that $A_k \cap B_k = \emptyset$.

Now if $g \in \Gamma$ with $B_i \cap gA_i \neq \emptyset$ for each $i = 1, 2, 3$, then, by the previous paragraph, we see that $E(Y) \cap gE(X) \neq \emptyset$. But the set of such $g$ for given $(A_1, A_2, A_3)$ and $(B_1, B_2, B_3)$ is finite. This shows that $\Gamma$ is a convergence group as claimed. \qed
In particular, we deduce the well-known fact that any finitely generated group acts a convergence group on its space of ends. (Take $K$ to be any Cayley graph of $\Gamma$.)

Suppose that $f : K \to L$ is a contraction onto a connected graph, $L$ (i.e., a map such that the preimage of each edge of $L$ is an edge of $K$, and the preimage of every vertex of $L$ is a connected subgraph of $K$). There is a natural inclusion of $B(L)$ into $B(K)$. If the preimage of every finite-degree vertex of $L$ is finite, then this descends to an injection from $C(L)$ to $C(K)$. If, in addition, the preimage of every infinite degree vertex of $L$ is one-ended, then this is an isomorphism, so that $\Xi_0(K)$ and $\Xi_0(L)$ are canonically homeomorphic.


Let $T$ be a simplicial tree with vertex set $V = V(T)$, edge set $E(T)$, and directed edge set $\vec{E}(T)$. Given $\vec{e} \in \vec{E}(T)$, we write $e$ for the underlying undirected edge, and $-\vec{e}$ for the same edge pointing in the opposite direction.

Let $V(\vec{e}) \subseteq V$ be the set of vertices, $v$, such that $\vec{e}$ points towards $v$. If $\vec{e}, \vec{f} \in \vec{E}(T)$, we write $\vec{e} < \vec{f}$ to mean that $\vec{f}$ points towards $\vec{e}$ and $\vec{e}$ points away from $\vec{f}$. (In some papers the opposite convention is used.) Note that this is equivalent to saying that $V(\vec{e})$ is strictly contained in $V(\vec{f})$. Clearly $\leq$ is a partial order on $\vec{E}(T)$ with order reversing involution $[\vec{e} \mapsto -\vec{e}]$.

A subset, $F \subseteq \vec{E}(T)$ is a transversal if, for all $\vec{e} \in \vec{E}(T)$, precisely one of $\vec{e}$ or $-\vec{e}$ lies in $F$. A transversal, $F$, is a flow on $T$ if no two elements of $F$ point away from each other (i.e., there do not exist $\vec{e}, \vec{f} \in F$ with $\vec{e} \leq -\vec{f}$). A flow must be of one of two types. Either there is some (unique) $v \in V(T)$ such that each element of $F$ points towards $v$, or else there is some infinite ray, $\alpha \subseteq T$ such that all edges of $F \cap E(\alpha)$ point away from its basepoint, and all other elements of $F$ point towards $\alpha$. We can identify the set of flows of the second kind with the boundary, $\partial T$, of $T$, thought of as a hyperbolic space in the sense of Gromov [Gr].

Recall the definitions of $\mathcal{B} = \mathcal{B}(T)$ and $\Xi(T)$ from Section 5. Note that $\mathcal{E} = \mathcal{E}(T) = \{V(\vec{e}) \mid \vec{e} \in \vec{E}(T)\}$ is a nested set of generators for $\mathcal{B}$. (The partial order, $\leq$, on $\vec{E}(T)$ agrees with that on $\mathcal{E} \subseteq \mathcal{B}$, and the involution $[\vec{e} \mapsto -\vec{e}]$ corresponds to the involution $[W \mapsto W^\ast]$.) We can think of an element of $\Xi(T)$ as an ultrafilter on $\mathcal{B}(T)$. The elements of $\mathcal{E}$ lying in that ultrafilter define a flow on $T$. In this way we may identify $\Xi(T)$ with the set of flows on $T$, and hence with $V(T) \sqcup \partial T$.

Now, $T$ is fine (as defined in Section 2) and hyperbolic (in the sense of Gromov). It thus has associated with it a compact space $\Delta T$, as in [Bo5]. As a set, $\Delta T$ may be defined as $V(T) \sqcup \partial T$. Note that any two elements, $x, y \in \Delta T$ can be connected by a unique arc, $[x, y]$, which may be compact,
a ray, or a bi-infinite geodesic depending on whether or not \(x, y \in V(T)\). Given \(x \in \Delta T\) and \(I \subseteq E(T)\), let \(B(x, I) = \{y \in \Delta T \mid I \cap E([x, y]) = \emptyset\}\). We define a topology on \(\Delta T\) by taking a neighbourhood base of \(x \in \Delta(T)\) to be the collection \(\{B(x, I)\}_I\) as \(I\) runs over all finite subsets \(E(T)\). In this topology, \(\Delta T\) is compact and Hausdorff. We see that, as sets, \(\Delta T\) can be identified with \(\Xi(T)\), and one can readily verify that the two topologies agree.

As with more general graphs (Section 5), the isolated points of \(\Delta T \equiv \Xi(T)\) are precisely the vertices of \(T\) of finite degree. We shall define the ideal boundary \(\Delta_0 T\), of the tree, \(T\), as the space \(\Delta T\) minus the vertices of \(T\) of finite degree. Again, this can be identified with with \(\Xi_0(T)\) as defined in Section 5.

If a group \(\Gamma\) acts on a simplicial tree, \(T\), with finite edge stabilisers, then the induced action on \(\Xi(T) \equiv \Delta T\) is a convergence action. In fact, if the action of \(\Gamma\) on \(T\) is cofinite, then the action on \(\Delta T\) is geometrically finite, with the infinite vertex groups precisely the maximal parabolic subgroups. Thus, \(\Gamma\) is hyperbolic relative to the infinite vertex groups, and its boundary can be identified with \(\Xi_0(T) \equiv \Delta_0 T\) (see [Bo5]). Moreover, if \(\Gamma\) is finitely generated, and each infinite vertex group is finitely generated and one-ended, then from the last paragraph of Section 5, we see that \(\Delta_0 T\) is canonically homeomorphic to the space of ends of \(\Gamma\).

### 7. Protrees and nested sets.

A **protree** is a set \(\Theta\), with an involution \([x \mapsto x^*]\) and a partial order, \(\leq\), with the property that for any \(x, y \in \Theta\), precisely one of the six relations \(x < y, x^* < y, x < y^*, x^* < y^*, x = y, x = y^*\) holds. This notion is due to Dunwoody (see for example [Du4]). We refer to a \(^*\)-invariant subset of \(\Theta\) as a **subprotree**.

An example of a protree is the directed edge set of any simplicial tree, as described in the last section. In fact, any finite protree can be realised as the directed edge set of a finite tree (a property which could serve as an equivalent definition). More generally any “discrete” protree can be realised as the directed edge set of a simplicial tree. A protree is said to be **discrete** if, for all \(x, y \in \Theta\), the set \(\{z \in \Theta \mid x \leq z \leq y\}\) is finite. We can define transversals and flows on a protree, exactly as for simplicial trees. (However, we cannot in general classify flows in the same way.)

Suppose \(\Theta\) is a protree. Let \(F = F(\Theta)\) be the Boolean ring with generating set \(\Theta\), and with relations \(x + x^* = 1\) for all \(x \in \Theta\) and \(xy = 0\) for all \(x, y \in \Theta\) satisfying \(x < y^*\) in \(\Theta\). (It follows that \(x < y^*\) also in \(F(\Theta)\) in the sense defined in Section 3.) Note that it follows that if \(x, y \in \Theta\) are distinct, then \(xy\) must be equal to 0, \(x, y\) or \(1 + x + y\). In particular, it follows that any element of \(F\) can be written as a sum of finitely many elements of \(\Theta\).
Note that if \( x \in \Theta \), then we can define an epimorphism, \( \theta : \mathcal{F} \rightarrow \mathbb{Z}_2 \) by \( \theta(x) = 1 \) and \( \theta(y) = 0 \) for all \( y \in \Theta \setminus \{x\} \). It therefore follows that \( x \neq 0 \) for all \( x \in \Theta \).

Suppose that \( F \subseteq \Theta \) is any transversal. We may identify \( F \) as a subset of \( \mathcal{F}(\Theta) \), and as such, it generates \( \mathcal{F}(\Theta) \) as a ring with unity (indeed as an additive group if we include also 1). If \( x, y \in F \), then precisely one of the relations \( xy = 0 \), \( x^*y = 0 \), \( xy^* = 0 \), \( x^*y^* = 0 \) holds. Thus, \( \Theta \) is a nested set of generators for \( \mathcal{F}(\Theta) \). Any element of \( \mathcal{F}(\Theta) \) can be written in a standard form
\[
\epsilon + \sum_{i=1}^n x_i
\]
where \( \epsilon \in \{0, 1\} \) and \( x_1, \ldots, x_n \) are distinct elements of \( F \).

If \( \Theta \) is a discrete protree, and \( T \) is the corresponding simplicial tree, then there is an epimorphism \( \phi : \mathcal{F}(\Theta) \rightarrow \mathcal{B}(T) \) defined by \( \phi(x) = V(\vec{e}) \), where \( \vec{e} \in \vec{E}(T) \) is the directed edge corresponding to \( x \in \Theta \). Let \( F \subseteq \Theta \) be any transversal. If \( x_1, \ldots, x_n \in F \) are distinct, then the corresponding elements of \( \mathcal{B}(T) \) are distinct. If \( n \neq 0 \), it follows easily that their symmetric difference can be neither \( \emptyset \) nor \( V(T) \). In other words, we see that any standard form of any element in the kernel of \( \phi \) must be trivial. Hence, the kernel is trivial, so \( \phi \) is, in fact, an isomorphism.

We now return to the case of a general protree, \( \Theta \), with transversal \( F \). Suppose that \( \Phi \subseteq \Theta \) is any subprotree. Then, there is a natural epimorphism, \( \theta : \mathcal{F}(\Theta) \rightarrow \mathcal{F}(\Phi) \) defined by \( \theta(x) = x \) if \( x \in \Phi \), and \( \theta(x) = 0 \) if \( x \in \Theta \setminus \Phi \). In particular, suppose \( x_1, \ldots, x_n \) are distinct. We get an epimorphism from \( \mathcal{F}(\Theta) \) to \( \mathcal{F}(\Phi) \), where \( \Phi = \bigcup_{i=1}^n \{x_i, x_i^*\} \). Now, \( \mathcal{F}(\Phi) \) is isomorphic to the Boolean algebra on a finite tree, as above, and so it follows that \( x_1 + \cdots + x_n \notin \{0, 1\} \). We see that the standard form of an element of \( \Theta \) (with respect to a given transversal, \( F \)) is unique. We therefore have an explicit description of the ring \( \mathcal{F}(\Theta) \).

Let \( \Xi(\Theta) = \Xi(\mathcal{F}(\Theta)) \) be the Stone dual. If we think of an element of \( \Xi(\Theta) \) as an ultrafilter on \( \mathcal{F}(\Theta) \), then its intersection with \( \Theta \) is a flow on \( \Theta \). We may therefore identify \( \Xi(\Theta) \) with the set of flows on \( \Theta \). For non-discrete protrees, however, we do not get a clear distinction between vertices and boundary points, as in the simplicial case.

Suppose that \( \mathcal{B} \) is any Boolean algebra with a nested set of generators, \( \mathcal{E} \subseteq \mathcal{B} \). Now, \( \mathcal{E} \) has the structure of a protree, with involution and partial order induced from \( \mathcal{B} \). We can therefore construct the Boolean algebra \( \mathcal{F} = \mathcal{F}(\mathcal{E}) \) as above, and identify \( \mathcal{B} \) as a quotient of \( \mathcal{F} \). Note that \( \Xi(\mathcal{B}) \) can thus be identified as a closed subset of the space \( \Xi(\mathcal{E}) \).
In fact, we can say more than this. We can formally identify \( E \) as a subset of \( F \). When composed with the canonical epimorphism from \( F \) to \( B \), this gives the inclusion of \( E \) into \( B \). Let \( \mathcal{I} \) be the kernel of the canonical epimorphism from \( F \) to \( B \). Note that if \( x, y \in E \), then \( x, y, x + y \notin \mathcal{I} \) (since, by the definition of a nested set, they correspond to distinct nonzero elements of \( B \)). Now a combinatorial argument shows that \( \mathcal{I} \) is generated by elements of the form \( 1 + \sum_{i=1}^{n} x_i \), where \( x_1, \ldots, x_n \in E \) have the property that \( x_i x_j = 0 \) for \( i \neq j \), and if \( y \in E \) then for some \( i \in \{1, \ldots, n\} \), we have \( yx_i \in \{x_i, 1+x_i+y\} \). In other words, we can think of the elements \( x_1, \ldots, x_n \) as a set of edges whose tails all meet at a “vertex” of degree \( n \) of the protree \( E \). (This statement is precise if the protree \( E \) happens to be discrete, and hence the edge set of a simplicial tree.) Suppose that \( a \in \Xi(\mathcal{E}) \setminus \Xi(\mathcal{B}) \). Now \( a \) corresponds to a flow, \( F \), on \( E \). This cannot be identically zero on \( \mathcal{I} \) and so must be nonzero on some generator of \( \mathcal{I} \) of the above form. From this, it is easy to see that \( F \) converges to some vertex of finite degree. But such a point is easily seen to be isolated in \( \Xi(\mathcal{E}) \). We have shown that every point of \( \Xi(\mathcal{E}) \setminus \Xi(\mathcal{B}) \) is isolated.

Finally, suppose that \( \mathcal{B} \) is a subalgebra of \( \mathcal{P}(V) \) for some set, \( V \). There is a natural map from \( V \) to \( \Xi(\mathcal{B}) \) as defined in Section 3. We therefore get a map from \( V \) to \( \Xi(\mathcal{E}) \). The image of a point \( x \in V \) in \( \Xi(\mathcal{E}) \) is defined by the flow \( \{A \in E \mid x \in A\} \) on \( E \).

8. Construction of nested generating sets.

It was shown in [DuS] how the Bergman norm can be used to construct invariant nested subsets of a Boolean algebra. Dunwoody has observed how an elaboration of this argument in fact gives us nested generating sets. The central idea may be formulated in a general fashion as follows.

Let \( \mathcal{B} \) be a Boolean ring. We say that two elements, \( x, y \in \mathcal{B} \) are disjoint if \( xy = 0 \).

Suppose that \( S \) is an ordered abelian group (or cancellative semigroup). Suppose that to each disjoint pair, \( x, y \in \mathcal{B} \), we have associated an element \( \sigma(x, y) \in S \). We suppose that \( \sigma(x, y) = \sigma(y, x) \geq 0 \), and that \( \sigma(x, y) > 0 \) if \( x, y \neq 0 \). Moreover, if \( x, y, z \in \mathcal{B} \) are pairwise disjoint, then \( \sigma(x, y + z) = \sigma(x, y) + \sigma(x, z) \). Given any \( x, y \in \mathcal{B} \), we write \( \mu(x) = \sigma(x, x^*) \) and \( \mu(x|y) = \sigma(xy, x^*y) \). Clearly \( \mu(x^*) = \mu(x) \) and \( \mu(x^*|y) = \mu(x|y) \).

Suppose now that \( x, y \in \mathcal{B} \) are non-nested, i.e., that \( xy, x^*y, xy^*, x^*y^* \) are all nonzero. If \( \mu(y|x) \leq \mu(x|y^*) \), then

\[
\mu(xy) = \sigma(xy, (xy)^*) \\
= \sigma(xy, x^* + xy^*) \\
= \sigma(xy, x^*) + \sigma(xy, xy^*)
\]
Similarly, if \( \mu(y|x) \leq \mu(x|y) \), then \( \mu(xy^*) < \mu(x) \).

Now, if we allow ourselves to permute \( x, y, x^*, y^* \), then we can always arrange that \( \mu(y|x) \) is minimal among \( \{ \mu(x|y), \mu(y|x), \mu(x|y^*), \mu(y|x^*) \} \), so that \( \max\{\mu(xy), \mu(xy^*)\} < \mu(x) \).

**Lemma 8.1.** If \( \mu(B \setminus \{0,1\}) \) is well-ordered (as a subset of \( S \)), then \( B \) has a nested set of generators.

**Proof.** Let \( \mathcal{E} \subseteq B \setminus \{0,1\} \) be the set of \( x \in B \) such that \( x \) does not lie in the ring generated by \( \{ z \in B \setminus \{0,1\} \mid \mu(z) < \mu(x) \} \). Clearly, \( \mathcal{E} \) generates \( B \). Moreover, \( \mathcal{E} \) is nested. For if \( x, y \in \mathcal{E} \) were not nested, then, without loss of generality, \( \max\{\mu(xy), \mu(xy^*)\} < \mu(x) \). But \( x = xy + xy^* \), giving a contradiction.

Note that the generating set we obtain is \( * \)-invariant.

**9. A variation on the Bergman norm.**

In this section, we give a generalisation of Bergman’s result [Ber].

Let \( V \) be a set, and let \( \mathcal{P}(V) \) be its power set. Let \( \mathcal{R} = \mathcal{R}(V) \) of binary partitions of \( V \), i.e., pairs \( \{A, B\} \subseteq \mathcal{P}(V) \) such that \( V = A \sqcup B \). We say that \( \{A, B\} \) is nontrivial if \( A, B \neq \emptyset \). Note that \( \mathcal{R} \) has the structure of an abelian group, with addition defined by \( \{A, A^*\} + \{B, B^*\} = \{(A \cap B) \cup (A^* \cap B^*), (A \cap B^*) \cup (A^* \cap B)\} \). The same structure can be obtained by quotienting the additive group of the Boolean ring \( \mathcal{P}(V) \) by the subgroup \( \{0,1\} \).

Let \( \Psi \) be the set of unordered pairs (i.e., 2-element subsets) of \( V \). If \( \pi \in \Psi \) and \( P = \{A, A^*\} \), we say that \( \pi \) crosses \( P \) if \( \pi \cap A \neq \emptyset \) and \( \pi \cap A^* \neq \emptyset \). Let \( \Psi(P) = \Psi(A) \) be the set of \( \pi \in \Psi \) such that \( \pi \) crosses \( P \).

Suppose that a group \( \Gamma \) acts on \( V \).

**Definition.** A (directed) slice of \( V \) is a subset \( A \subseteq V \), such that \( \Phi \cap \Psi(A) \) is finite for every \( \Gamma \)-orbit, \( \Phi \), in \( \Pi \).

We refer to \( P = \{A, A^*\} \) as an undirected slice. We write \( \mathcal{S}(V) \) for the Boolean algebra of directed slices, and \( \hat{\mathcal{S}}(V) \) for the additive group of undirected slices.
Suppose \((\Psi_n)_{n \in \mathbb{N}}\) is a collection of cofinite \(\Gamma\)-invariant subsets of \(\Psi\) with \(\Psi = \bigcup \Psi_n\). Given \(n \in \mathbb{N}\) and \(P \in \mathcal{R}\), write \(\Psi_n(P) = \Psi_n \cap \Psi(P)\). Thus, \(P \in \mathcal{S}\) if and only if \(\Psi_n(P)\) is finite for all \(P\). We write \(K_n\) for the graph with vertex set \(V\) and edge set \(\Psi_n\).

Let \(\mathbb{N}^\mathbb{N}\) be the set of infinite sequences of natural numbers. This has the structure of an ordered abelian group with the lexicographic order. We define a map, \(\mu : \hat{\mathcal{S}} \rightarrow \mathbb{N}^\mathbb{N}\) by setting \(\mu(P) = (\mu_n(P))_n\), where \(\mu_n(P) = |\Psi_n(P)|\).

Recall that \(V\) is connected if it is the vertex set of some connected \(\Gamma\)-graph with finite quotient (i.e., some finite union of the \(K_n\) is connected).

We show:

**Theorem 9.1.** Let \(V\) be a connected cofinite \(\Gamma\)-set. Then \(\mu(\hat{\mathcal{S}}) \subseteq \mathbb{N}^\mathbb{N}\) is well-ordered, and the map \(\mu\) is finite-to-one modulo the action of \(\Gamma\).

It will be convenient for the proof to assume that the sets \(\Psi_n\) are increasing, i.e., \(\Psi_m \subseteq \Psi_n\) whenever \(m \leq n\). There is no loss of generality in doing this, for if we set \(\Psi'_n = \bigcup_{m \leq n} \Psi_m\), then if \((P^\alpha)_{\alpha \in \mathbb{N}}\) is an infinite sequence of undirected slices which is non-increasing with respect to the order determined by \((\Psi_n)_{n}\), then some subsequence will be non-increasing with respect to the order determined by \((\Psi'_n)_{n}\). We may as well also assume that \(K_0\), and hence every \(K_n\), is connected. It is easily verified that if \(n \in \mathbb{N}\), then \(\mu_n(P + Q) \leq \mu_n(P) + \mu_n(Q)\), with equality precisely if no element of \(\Psi_n\) crosses both \(P\) and \(Q\).

We now set about the proof of Theorem 9.1.

For the moment, we can forget about the group, \(\Gamma\). Let \(K\) be a connected graph, with vertex set \(V(K) = V\) and edge set \(E(K)\). A finite subset, \(I \subseteq E(K)\) is separating if \(K \setminus I\) is disconnected. (Here, \(K \setminus I\) denotes the graph with vertex set \(V(K)\) and edge set \(E(K) \setminus I\).)

**Definition.** A cut is a finite nonempty subset, \(I \subseteq E(K)\), such that every circuit (or equivalently, cycle) in \(K\) contains an even number of edges of \(I\).

Thus, every cut is separating. A cut is minimal if it contains no proper subcut. We similarly define minimality for separating sets. It is easily verified that if \(I \subseteq E(K)\) is finite, then the following three conditions are equivalent:

1. \(I\) is a minimal cut,
2. \(I\) is a minimal separating set, or
3. \(K \setminus I\) has precisely two connected components.

Note that each cut \(I \subseteq K\) determines a nontrivial partition, \(P(K, I) = \{A, A^*\}\), where each path from \(A\) to \(A^*\) contains an odd number of edges of \(I\). Conversely, each nontrivial partition \(P = \{A, A^*\}\) determines the subset, \(I = I(K, P)\), of edges which cross from \(A\) to \(A^*\). If \(I\) is finite, then it is a
cut, and we say that $P$ is an (undirected) $K$-break. There is thus a natural bijection between cuts and $K$-breaks.

Given two cuts $I$ and $J$, we write $I + J$ for their symmetric difference. This is also a cut, and the operation agrees with that already defined on the set of partitions. If $J \subseteq I$ is a subcut, then $I \setminus J = I + J$ is also a subcut, and $I = J + (I \setminus J)$. We shall measure the “size” of a cut by its cardinality.

We note:

**Lemma 9.2.** Suppose $e \in E(K)$ and $n \in \mathbb{N}$. There are finitely many minimal cuts of size $n$ containing $e$.

**Proof.** Suppose, for contradiction, that the set, $I$, of such minimal cuts is infinite. Choose $I \subseteq E(K)$ maximal such that $I$ is contained in infinitely many elements of $\mathcal{I}$. Let $\mathcal{J} = \{J \in \mathcal{I} \mid I \subseteq J\}$. Now, $I$ cannot separate $K$ (otherwise $I$ would be a minimal cut, and $J = \{I\}$). Let $\gamma$ be a path in $K \setminus I$ connecting the endpoints of $e$. Now each element of $\mathcal{J}$ contains some edge of $\gamma$, and so some infinite subset of elements of $\mathcal{J}$ all contain the same edge, say $f$, of $\gamma$. But now $I \cup \{f\}$ is contained in infinitely many elements of $\mathcal{J}$ and hence of $\mathcal{I}$, contradicting the maximality of $I$. $\square$

**Definition.** We say that a cut, $I$, is blocklike if every pair of elements of $I$ lie in a minimal subcut of $I$.

Clearly, in Lemma 9.2, one can replace “minimal cut” by “blocklike cut”. Every cut can be uniquely decomposed into maximal blocklike cuts. One way to see this is as follows.

Suppose $\Upsilon$ is a finite connected graph. A block of $\Upsilon$ is a maximal 2-vertex-connected subgraph. Two blocks intersect, if at all, in a single vertex. If $e, f \in E(\Upsilon)$ are distinct, then $e, f$ lie in the same block if and only if they lie in some circuit in $\Upsilon$, and if and only if they lie in some minimal separating set. Note that $\Upsilon$ is bipartite (i.e., $E(\Upsilon)$ is a cut) if and only if each of its blocks is bipartite.

Suppose that $K$ is any connected graph, and $I \subseteq E(K)$ is a cut. Let $\hat{\Upsilon} = \hat{\Upsilon}(K, I)$ be the graph obtained by collapsing each connected component of $K \setminus I$ to a point. Thus, $\hat{\Upsilon}$ is a finite connected bipartite graph. There is a canonical surjective map $\phi : K \rightarrow \hat{\Upsilon}$. The preimage of every connected subgraph is connected.

Now it is easily checked that the preimage of every cut in $\hat{\Upsilon}$ is a subcut of $I$. Moreover, every subcut, $J$, arises in this way: $J = \phi^{-1}(\phi J)$. We also see easily that $J$ is minimal if and only if $\phi J$ is minimal (since $K \setminus J$ has the same number of components as $\phi(K \setminus J) = K \setminus \phi J$). Thus if follows that $J$ is blocklike if and only if $\phi J$ is blocklike. We can thus decompose $I$ canonically by taking the preimages of (the edge sets of) blocks of $\hat{\Upsilon}$.

If $P$ is a $K$-break, then we shall write $\Upsilon(K, P) = \Upsilon(K, I(K, P))$. The canonical decomposition of $I(K, P)$ gives us a canonical decomposition of $P$ (depending on $K$).
Now suppose that $K'$ is another graph on the same vertex set, $V$, with $K \subseteq K'$. Let $\phi : K \to \Upsilon = \Upsilon(K, P)$ and $\phi' : K' \to \Upsilon' = \Upsilon(K', P)$ be collapsing maps described above. We can obtain $\Upsilon'$ from $\Upsilon$ by identifying certain vertices of the same colour and/or adding edges between vertices of different colours. There is thus a natural map, $\psi : \Upsilon \to \Upsilon'$, such that $\psi \circ \phi = \phi' \circ \iota$, where $\iota$ is the inclusion of $K$ in $K'$. If we measure the complexity, $c(\Upsilon)$, of a finite graph, $\Upsilon$, by the number of edges in the complementary graph, i.e., $c(\Upsilon) = \frac{1}{2} |V(\Upsilon)|(|V(\Upsilon)| - 1) - E(\Upsilon)$, then we see that the map $\psi$ cannot increase complexity. We have $c(\Upsilon') = c(\Upsilon)$ if and only if $\psi$ is an isomorphism. In this case, the canonical decomposition of $P$ with respect to $K$ is identical to its canonical decomposition with respect to $K'$. Note also that in general, $|V(\Upsilon')| \leq |V(\Upsilon)| \leq |I(K, P)| + 1$.

Suppose now that $\mathcal{T}$ is an infinite set of undirected slices with $\mu_0(P) = |I(K_0, P)|$ bounded, by $\nu$, say for $P \in \mathcal{T}$. Thus, for each $P \in \mathcal{T}$ and $n \in \mathbb{N}$, we have $|V(\Upsilon(K_n, P))| \leq |V(\Upsilon(K_0, P))| \leq \mu_0(P) + 1 \leq \nu + 1$, so there are only finitely many possibilities for the graph $\Upsilon(K_n, P)$. Given a graph $\Upsilon$, and $n \in \mathbb{N}$, let $\mathcal{T}(n, \Upsilon) = \{ P \in \mathcal{T} \mid \Upsilon(K_n, P) = \Upsilon \}$. For any fixed $n$, this partitions $\mathcal{T}$ into finitely many subsets. We can therefore choose $\Upsilon = \Upsilon_0$ of minimal complexity, $c(\Upsilon_0)$, with the property that for some $n = n_0$, say, the set $\mathcal{T}_0 = \mathcal{T}(n_0, \Upsilon_0)$ is infinite. For any $n \geq n_0$, $\Upsilon(K_n, P) = \Upsilon_0$ for all but finitely many $P \in \mathcal{T}_0$. Since $c(\Upsilon(K_n, P)) \leq c(\Upsilon(K_{n_0}, P)) = c(\Upsilon_0)$, and if we had strict inequality for infinitely many $P$, then we would contradict the minimality of $c(\Upsilon_0)$.

We now introduce the action of $\Gamma$. Note that by Lemma 9.2, $K_n$ has only finitely many blocklike cuts of size $s$ modulo $\Gamma$, for any $n, s \in \mathbb{N}$.

Suppose that $(P^\alpha)_{\alpha \in \mathbb{N}}$ is a sequence of undirected slices all lying in distinct $\Gamma$-orbits. Suppose that $\mu(P^\alpha)$ is non-increasing. We want to derive a contradiction.

First note that applying the construction above, with $\mathcal{T} = \{ P^\alpha \mid \alpha \in \mathbb{N} \}$, we can assume that, after passing to a subsequence, $(P^\alpha)_\alpha$ has the following property. There exist $n_0 \in \mathbb{N}$ and a finite bipartite graph, $\Upsilon_0$, such that if $n_0 \leq n \leq \alpha$, then $\Upsilon(K_n, P^\alpha) = \Upsilon_0$. For notational convenience (replacing $\Psi_0$ by $\Psi_{n_0}$), we can assume that $n_0 = 0$.

Now for each $\alpha$, we decompose the cut $I(K_0, P^\alpha)$ into its maximal blocklike subcuts, $I(K_0, P^\alpha) = I^1_1 + \cdots + I^p_1$. Note that $p$ is the number of blocks
Thus, there is some edge in $K$. Let $\alpha$ are not disjoint. Let $K$ be the Boolean algebra of slices of $\Gamma$. Thus, for $n < m$, we have $\mu_n(P_\alpha) = \mu_n(P_\beta) + \cdots + \mu_n(P_p) = \mu_n(Q_1) + \cdots + \mu_n(Q_p)$, which is independent of $\beta$.

Now for some $\alpha$, the cuts $\{I(K_m, P_\alpha_i)\}_{i=1}^p$ are no longer disjoint, so this time we get strict inequality: $\mu_m(P_\alpha) < \mu_m(Q_1) + \cdots + \mu_m(Q_p)$. However, for $\beta \geq m$, we have $\Upsilon(K_m, P_\beta) = \Upsilon_0 = \Upsilon(K_0, P_\beta)$, and the natural map from $\Upsilon(K_0, P_\beta)$ to $\Upsilon(K_m, P_\beta)$ is an isomorphism. Thus, the canonical decompositions of $P_\beta$ with respect to $K_0$ and $K_m$ are identical. In particular, the cuts $\{I(K_m, P_\beta_i)\}_{i=1}^p$ are disjoint, and again, we have equality: $\mu_m(P_\beta) = \mu_m(Q_1) + \cdots + \mu_m(Q_p)$. Thus $\mu_m(P_\beta) > \mu_m(P_\alpha)$. Taking $\beta > \alpha$, we derive a contradiction to the assumption that $\mu$ is non-increasing.

This proves Theorem 9.1.

We note the following corollary of this result.

Suppose that $V$ is countable. Let $\{\Psi_n\}_n$ be an enumeration of the $\Gamma$-orbits of the set of pairs, $\Psi$. Suppose $A, B \in S$ are disjoint. Let $\sigma_n(A, B)$ be the number of pairs in $\Psi_n$ with one element in $A$ and one element in $B$. Let $\sigma(A, B) = (\sigma_n(A, B))_n \in \mathbb{N}^\mathbb{N}$. Thus, $\mu(A) = \sigma(A, A^*)$. We are thus in the situation described in Section 8. Applying Lemma 8.1, we deduce:

**Theorem 9.3.** Suppose $\Gamma$ is a countable group, and $V$ is a connected cofinite $\Gamma$-set. Let $S$ be the Boolean algebra of slices of $V$. Then, any $\Gamma$-invariant subalgebra of $S$ has a $\Gamma$-invariant nested set of generators.

**Proof.** The construction of Section 8 was canonical, and hence $\Gamma$-invariant. \hfill \Box

We finish with the following observation with regard to slices.

**Lemma 9.4.** Suppose that $V$ is a cofinite fine $\Gamma$-set with finite pair stabilisers and $W \subseteq V$ is a $\Gamma$-invariant subset. Suppose that each element of $V \setminus W$ has finite stabiliser. Then the map $[A \to A \cap W] : S(V) \to S(W)$ is an epimorphism of $\Gamma$ Boolean algebras.
Proof. We only really need to note that the map is surjective. To see this, suppose \( B \in \mathcal{S}(W) \). Choose any \( b \in B \) and any orbit transversal, \( \{x_1, \ldots, x_n\} \), of \( V \setminus W \). Let \( A = B \cup \{gx_i \mid 1 \leq i \leq n, gb \in B\} \).

To verify that \( A \in \mathcal{S}(V) \), suppose that \( x, y \in V \) are distinct. We connect \( x \) and \( y \) by an edge \( e \). Let \( E_0 \) be the set of edges connecting \( b \) to each of the \( x_i \). Let \( K \) be the graph with vertex set \( V \) and edge set \( \Gamma e \cup \bigcup \Gamma E_0 \). We see that \( K \) is fine.

Now if some \( \Gamma \)-image, \( e' \), of \( e \) connects \( A \) to \( A^* \), then it lies in an arc of length at most 3 connecting \( B \) to \( B^* \), and whose interior vertices are at \( \Gamma \)-images of \( x \) or \( y \). Since \( x \) and \( y \) have finite degree in \( K \), then modulo \( \Gamma \), there are only finitely many possibilities for such arcs, and hence for its pair of endpoints. But since \( B \in \mathcal{S}(\Gamma) \), there can only be finitely many such arcs in total. Thus, there are only finitely many such edges \( e' \). Since \( x, y \) were arbitrary, we have shown that \( A \in \mathcal{S}(V) \). \( \square \)

10. A finiteness result for Boolean algebras related to simplicial trees.

In this section, using a result of [DiD], we give a proof of the following:

**Lemma 10.1.** Suppose that \( \Gamma \) is a countable group, and that \( T \) is a cofinite simplicial \( \Gamma \)-tree with finite edge stabilisers. Suppose that \( A \) is any \( \Gamma \)-invariant Boolean subalgebra of the Boolean algebra, \( \mathcal{B}(T) \). Then \( A \) is finitely generated as a \( \Gamma \)-Boolean algebra.

Here, \( \mathcal{B}(T) \) is the Boolean algebra of \( T \)-breaks, as defined in Section 5. To say that \( A \) is finitely generated as a \( \Gamma \)-Boolean algebra means that it has a generating set which is a finite union of \( \Gamma \)-orbits.

Now, \( V = V(T) \) has finite pair stabilisers, and so Theorem 9.3 gives us a \( \Gamma \)-invariant nested set, \( \mathcal{E} \), of generators of \( A \). If we know already that \( A \) is finitely generated, then we can assume that \( \mathcal{E} \) is cofinite. It follows that if \( e \in E(T) \), the set \( \{A \in \mathcal{E} \mid e \in I(A)\} \) is finite. We see that if \( x, y \in V \), then \( \{A \in \mathcal{E} \mid x \in A, y \notin A\} \) is finite. In particular, if \( A, B \in \mathcal{E} \), then \( \{C \in \mathcal{E} \mid A \subseteq C \subseteq B\} \) is finite. In other words, \( \mathcal{E} \) satisfies the finite interval condition. We can thus identify \( \mathcal{E} \) with the directed edge set, \( \bar{E}(S) \), of a simplicial tree. Note that \( \Gamma \) acts on \( S \) with finite edge stabilisers.

If \( x \in V(T) \), then \( \{A \in \mathcal{E} \mid x \in A\} \) is a flow on \( \mathcal{E} \), and hence determines a flow on \( S \). Moreover, from the observation of the previous paragraph, there can be no infinite decreasing sequence in the flow (i.e., any strictly decreasing sequence, \( A_1 \supset A_2 \supset A_3 \supset \ldots \) with \( A_i \in \mathcal{E} \) must terminate). The flow thus arises from a unique vertex of \( S \). This therefore defines a \( \Gamma \)-equivariant map \( \phi : V(T) \longrightarrow V(S) \).

Suppose \( y \in V(S) \). Let \( \mathcal{E}(y) \subseteq \mathcal{E} \) be the set of elements of \( \mathcal{E} \) which correspond to directed edges with tail at \( y \). Let \( \mathcal{E}^*(y) = \{A^* \mid A \in \mathcal{E}(y)\} \). If \( x \in \bigcap \mathcal{E}^*(y) \subseteq V(T) \), then each edge corresponding to an element of
\( E^*(y) \) must point towards \( \phi(x) \in V(S) \). It follows that \( \phi(x) = y \). From this observation, we see that if \( y \not\in \phi(V(T)) \), then \( \bigcap E^*(y) = \emptyset \), so that \( \bigcup E(y) = V(S) \). If it also happens that \( E(y) \) is finite, then \( \bigcup E(y) \) is the sum (i.e., symmetric difference) of the elements of \( E(y) \). We deduce:

**Lemma 10.2.** If \( y \in V(S) \setminus \phi(V(T)) \) has finite degree in \( S \), then \( \sum_{i=1}^{n} A_i = 1 \) in \( B(T) \), where \( A_1 \ldots A_n \) are the elements of \( E \) which correspond to those edges with tails at \( y \).

This result will be used in the proof of Lemma 10.1.

To prove Lemma 10.1, one can use an accessibility result of Dicks and Dunwoody. Following \([DiD]\), we say that an edge of a \( \Gamma \)-tree is compressible if its endpoints lie in distinct \( \Gamma \)-orbits and if its stabiliser is equal to an incident vertex stabiliser. (Such an edge can be “compressed” in the corresponding graph of groups to give a smaller graph.) A \( \Gamma \)-tree is incompressible if it has no compressible edges.

The following is shown in \([DiD]\) (III 7.5, p. 92):

**Proposition 10.3.** Let \( \Gamma \) be a group. Suppose that \( S, T \) are cofinite simplicial \( \Gamma \)-trees with finite edge stabilisers, and that \( S \) is incompressible. If there is a \( \Gamma \)-equivariant map from \( V(T) \) to \( V(S) \), then \( |V(S)/\Gamma| \leq |V(T)/\Gamma| + |E(T)/\Gamma| \).

(In \([DiD]\) is assumed also that \( T \) is incompressible. However, it is clear that one can always collapse \( T \) to give another tree, \( T' \), with this property, and with \( V(T') \) isomorphic as a \( \Gamma \)-set to a \( \Gamma \)-invariant subset of \( V(T) \). This process can only decrease \( |V(T)/\Gamma| \) and \( |E(T)/\Gamma| \).

Alternatively, we can use (the argument of) the “elliptic” case of the accessibility result of Bestvina and Feighn \([BesF]\). This gives a slightly different result:

**Proposition 10.4.** Suppose that \( \Gamma \) is a group and that \( S, T \) are cofinite simplicial \( \Gamma \)-trees without edge inversions. Suppose that \( S \) is incompressible, and that every edge stabiliser of \( S \) fixes a vertex of \( T \). If there is a \( \Gamma \)-equivariant map from \( V(T) \) to \( V(S) \), then \( |V(S)/\Gamma| \leq \max\{1,5|E(T)/\Gamma|\} \).

In particular, this applies to the case of finite edge-stabilisers. We shall sketch a proof below, which is condensed out of the relevant part of \([BesF]\). Our direct use of Grushko’s Theorem bypasses the use of folding sequences. We begin with some preliminary remarks.

Let \( t \) be a graph of groups, and let \( \Gamma = \pi_1(t) \) be its fundamental group. Thus \( \Gamma \) acts on the corresponding Bass-Serre tree, \( T \), with quotient the underlying graph \( |t| = T/\Gamma \). Given \( v \in V(t) \) or \( e \in E(t) \), we write \( \Gamma(v) = \Gamma_t(v) \) or \( \Gamma(e) = \Gamma_t(e) \) for the corresponding vertex or edge groups. A subgroup of \( \Gamma \) is elliptic if it is conjugate into a vertex group. Let \( t_0 \) be the graph of groups with the same underlying graph, and with all vertex and edge groups
trivial. Thus, $\pi_1(t_0) = \pi_1(|t|)$ is free of rank $\beta(t) = |E(t)| - |V(t)| + 1$. Moreover, there is a natural epimorphism from $\Gamma$ to $\pi_1(t_0)$ whose kernel contains $\langle \bigcup_{v \in V(t)} \Gamma(v) \rangle$, where $\langle . \rangle$ denotes normal closure. In particular, if $\Gamma$ is the normal closure of some vertex group, $\Gamma(v)$, then $|t|$ is a tree. We see easily that if every other vertex group is $\Gamma$-conjugate into $\Gamma(v)$, then $\Gamma = \Gamma(v)$. Indeed, if every vertex group is conjugate into a subgroup, $G \leq \Gamma(v)$, then $\Gamma = G = \Gamma(v)$.

Proof of Proposition 10.4. After subdividing the edges of $T$, we obtain a $\Gamma$-tree, $T'$, with $V(T) \subseteq V(T')$, and an equivariant morphism $\phi : T' \to S$ (which sends each edge of $T'$ to an edge or vertex of $S$). This descends to a graph-of-groups morphism $\phi : t' \to s$ (where $|t'| = T'/\Gamma$ and $|s| = S/\Gamma$) which induces the identity map on $\gamma = \pi_1(t') = \pi_1(s)$. We make a series of observations.

Claim 1. If $v \in V(s) \setminus \phi(V(t))$ and $G \subseteq \Gamma(v)$ is $T$-elliptic, then $G$ is $\Gamma(v)$-conjugate into an incident edge group. This is easily seen by considering the arc connecting a lift of $v$ to $V(S)$ (fixed by $\Gamma(v)$) to the $\phi$-image of a fixed point of $G$ in $V(T)$.

In fact, the same argument shows that if $v, w \in V(s) \setminus \phi(V(t))$ are the endpoints of an edge $e \in E(s)$, then any $T$-elliptic subgroup, $G$, of $(\Gamma(v) \cup \Gamma(w))$ is conjugate into an incident edge group adjacent to (but different from) $e$. In particular, from the $T$-ellipticity hypothesis of the proposition, this applies to $G = \Gamma(e) = \Gamma(v) \cap \Gamma(w)$.

We say that a vertex, $v \in V(s)$ is dead if $v \notin \phi(V(t))$ and if $\Gamma(v)$ is the normal closure of the incident edge groups. Otherwise it is live. We thus decompose $V(s) = V_D(s) \sqcup V_L(s)$ into dead and live vertices.

Claim 2. $|V_L(s) \setminus \phi(V(t))| \leq \beta(t) - \beta(s)$. To see this, let $\overline{s}$ be the graph of groups with underlying graph $|s|$ obtained by collapsing each edge group in $E(s)$ and each vertex group in $\phi(V(t))$ to the trivial group, and by collapsing each remaining vertex group, $\Gamma_{\overline{s}}(v)$, to the quotient of $\Gamma_s(v)$ by the normal closure of its incident edge groups. Thus if $v \in V(\overline{s})$, then $\Gamma_{\overline{s}}(v)$ is nontrivial if and only if $v \in V_L(s) \setminus \phi(V(t))$. It follows that $\pi_1(\overline{s})$ has at least $|V_L(s) \setminus \phi(V(t))| + \beta(s)$ nontrivial free factors. Now there is a natural epimorphism from $\pi_1(t_0)$ to $\pi_1(\overline{s})$. The former group is free of rank $\beta(t)$, and so Claim 2 follows by Grushko’s Theorem.

Clearly, $|\phi(V(t))| \leq |V(t)|$ and so $|V_L(s)| \leq |E(t)| - \beta(s) + 1$.

Claim 3. Suppose $v \in V_D(s)$ and $G$ is an incident edge group. Suppose that every other incident edge group is $\Gamma(v)$-conjugate into $G$. Then $\Gamma(v) = G$. To see this, consider the action of $\Gamma(v)$ on a minimal $\Gamma(v)$-invariant subtree of $T$. Let $r$ be the corresponding action of $\Gamma(v)$ on a minimal $\Gamma(v)$-invariant subtree of $T$. Let $r$ be the corresponding graph of groups. Now by the $T$-ellipticity hypothesis, $G$ is conjugate into a vertex group, $\Gamma_r(w)$, where $w \in V(r)$.  

Moreover, if $H$ is any other vertex group of $r$, then again $H$ is $T$-elliptic, and hence, by Claim 1, is $\Gamma(v)$-conjugate into an incident edge group to $v$ in $s$. Thus $H$ is $\Gamma(v)$-conjugate into $G$. But now $\Gamma(v) = \langle \langle G \rangle \rangle = \langle \langle \Gamma(w) \rangle \rangle$, and so it follows by the discussion before the proof that $\Gamma(v) = G$. The claim follows.

As an immediate corollary, we see (by the incompressibility of $s$) that any such vertex must have degree at least 3 in $s$. In particular, all terminal vertices of $s$ are live.

**Claim 4.** We cannot have two adjacent dead vertices of degree 2 in $s$. For suppose to the contrary that $e \in E(s)$ has endpoints $v, w \in V_D(s)$, both of degree 2. From the remark after Claim 1, we see that, without loss of generality, $\Gamma(e)$ is conjugate into $\Gamma(f)$, where $f \in E(s)$ is the other edge incident on $v$. But now, from Claim 3, we derive the contradiction that $v$ has degree at least 3.

We now have enough information to bound $|V(s)|$ in terms of the complexity of $|t|$. First note that since every terminal vertex of $s$ is live (Claim 3), the number of such vertices is bounded by Claim 2. Moreover (Claim 2), we have $\beta(s) \leq \beta(t)$. This places a bound on the number of vertices of $s$ of degree at least 3. Finally, Claim 4 together with the bound on live vertices places a bound on the number of vertices of degree 2. More careful bookkeeping shows that if $|s|$ is not a point, then $|V(s)| \leq 5|E(t)| - |V(t)|$. Proposition 10.4 now follows.

Now since, $\beta(s) \leq \beta(t)$, we also get a bound on $|E(s)| = |E(S)/\Gamma|$. In fact, to obtain such a bound, we can weaken the hypotheses slightly:

**Corollary 10.5.** Let $\Gamma$ be a group and that $S, T$ be cofinite $\Gamma$-trees with finite edge stabilisers. Suppose that $\phi : V(T) \rightarrow V(S)$ is a $\Gamma$-equivariant map, and that each compressible edge of $S$ is incident on some element of $\phi(V(T))$. Then there is a bound on $|E(S)/\Gamma|$ in terms of $|E(T)/\Gamma|$.

**Proof.** To see this, note that we can obtain a $\Gamma$-tree, $S'$, by collapsing down a union of trees, $F$, in $S/\Gamma$, consisting of a union of compressible edges. Moreover, we can assume that if $x \in \phi(V(T))$, then the image of $x$ in $S/\Gamma$ is incident to at most one edge of $F$ that is terminal $S/\Gamma$. (Since collapsing such an edge will be sufficient to render all the other edges incident on $x$ incompressible.) Now, Proposition 10.3 or 10.4 gives a bound on the complexity of $S'/\Gamma$. This, in turn, gives a bound on the number of edges of $F$, and hence a bound on the complexity of $S/\Gamma$ as claimed.

We can now set about the proof of Lemma 10.1. Suppose that $\Gamma$, $T$ and $A$ are as in the hypotheses. By Theorem 9.3, there is a $\Gamma$-invariant nested set, $E$, of generators of $A$. Now, if $A$ is not finitely generated as a $\Gamma$-Boolean algebra, then we can find an infinite sequence, $(A_n)_{n \in \mathbb{N}}$, of elements of $E$ such
that $A_n$ does not lie in the $\Gamma$-Boolean algebra generated by $\{A_i \mid i < n\}$. Let $\Gamma(A_n)$ be the stabiliser of $A_n$. After passing to a subsequence, we can assume that $|\Gamma(A_n)|$ is non-decreasing in $n$. Let $E_n$ be the union of the $\Gamma$-orbits of $\{A_i, A_i^* \mid i \leq n\}$. Given $A \in E_n$, we write $m(A) = i$ to mean that $A$ or $A^*$ lies in the $\Gamma$-orbit of $A_i$.

Now fix some $n$. As discussed after the statement of Lemma 10.1, we can identify $E_n$ with the directed edge set of a simplicial tree, $S_n$, and there is an equivariant map, $\phi : V(T) \to V(S_n)$. Note that $|E(S_n)/\Gamma| = n$. We claim that $S_n$ satisfies the weakened hypotheses of Corollary 10.5. In fact, we show that any vertex whose stabiliser fixes an incident edge must lie in $\phi(V(T))$.

Suppose, to the contrary, that $y \in V(S_n) \setminus \phi(V(T))$ is incident on an edge $\vec{e} \in E(S_n)$ with tail at $y$ and with $\Gamma(e) = \Gamma(y)$. Note that $\Gamma(y)$ is finite, so that $y$ has finite degree. Let $A \in E_n$ be the element corresponding to $\vec{e}$, so that $\Gamma(A) = \Gamma(y)$. Let $E_n(y)$ be the set of elements of $E_n$ which correspond to edges with tails at $y$. We can assume that $\vec{e}$ is chosen so that $m(A)$ is maximal among those elements of $E_n(y)$ with stabilisers equal to $\Gamma(y)$. Write $E_n(y) = \{A, B_1, \ldots, B_k\}$. By Lemma 10.2, we have that $A = 1 + \sum_{j=1}^k B_j$ in the Boolean algebra $B(T)$. In particular, $A$ lies in the Boolean algebra generated by $\{B_1, \ldots, B_k\}$.

Now, for each $j$, $\Gamma(B_j) \leq \Gamma(A)$. Either $\Gamma(B_j) = \Gamma(A)$ so that, by the maximality of $m(A)$, we have $m(B_j) < m(A)$, or else $|\Gamma(B_j)| < |\Gamma(A)|$ so that, by the construction of the sequence $(A_i)_i$, we again have $m(B_j) < m(A)$. We therefore deduce that $A$ lies in the $\Gamma$-Boolean algebra generated by $\{A_i \mid i < m(A)\}$, contrary to the construction of $(A_i)_i$. This proves the claim.

Now applying Corollary 10.5, we get a bound on the complexity $n = |E(S_n)/\Gamma|$. But we could have chosen $n$ arbitrarily large, thereby giving a contradiction.

This proves Lemma 10.1.

11. An application to 1-connected $\Gamma$-sets.

In this section, we apply Lemma 10.1 to show:

**Lemma 11.1.** Suppose $\Sigma$ is a countable $\mathbb{Z}_2$-acyclic simplicial 2-complex which is locally finite away from the vertex set $V(\Sigma)$. Suppose a group $\Gamma$ acts on $\Sigma$ with finite quotient and such that $\Gamma(x) \cap \Gamma(y)$ is finite for all distinct $x, y \in V(\Sigma)$. Let $\mathcal{S}(V(\Sigma))$ be the Boolean algebra of slices of $V(\Sigma)$. Then, any $\Gamma$-invariant subalgebra, $A$, of $\mathcal{S}(V)$ has a cofinite nested set of generators which is discrete as a protree.

We can reduce Lemma 11.1 to Lemma 10.1 using the machinery of patterns and tracks as in [Du2]. (The overall strategy of the proof is thus
similar to that of the accessibility result of [BesF]. Let $\Sigma$, $\Gamma$, $V$, $A$ be as in the hypotheses, and let $K$ be the 1-skeleton of $\Sigma$. Recall that a pattern, $t$, on $\Sigma$ is a compact subset of $\Sigma \setminus V(\Sigma)$ which meets each 1-simplex either in the empty set or a single point, and which meets each 2-simplex, $\sigma$, either in the empty set or in a single arc connecting two distinct faces of $\sigma$. It represents a subset, $A \subseteq V(\Sigma)$ if it meets precisely those edges of $\Sigma$ which connect $A$ to $A^*$. Every $\Sigma$-slice is represented by a pattern. A track is a connected pattern. Two disjoint tracks, $s, t$, are parallel if there is a closed subset of $\Sigma \setminus V(\Sigma)$ homeomorphic to $s \times [0, 1] \cong t \times [0, 1]$ with boundary $s \sqcup t$. If $T$ is a $\Gamma$-equivariant set of disjoint pairwise non-parallel tracks, then there is a bound on $|T/\Gamma|$, (see [Du2]).

By Theorem 9.3 there is a $\Gamma$-invariant nested set of generators, $\mathcal{E}$, for $\mathcal{A}$. By a standard construction (cf. [Du2]), we can find a set of patterns $(t(A))_{A \in \mathcal{E}}$ such that $t(A)$ represents $A$, $t(A) = t(A^*)$, $t(A) \cap t(B) = \emptyset$ if $B \not= A, A^*$ and $t(gA) = gt(A)$ for all $g \in \Gamma$. By the observation of the previous paragraph, we can find a cofinite $\Gamma$-invariant set, $T$, of tracks such that if $A \in \mathcal{E}$, then each connected component of the pattern $t(A)$ is parallel to some element of $T$. Now $T$ determines a simplicial tree, $T$, whose edges are in bijective correspondence with $T$, and whose vertices are in bijective correspondence with the connected components of $\Sigma \setminus \bigcup T$. (It is here that we use the fact that $\Sigma$ is $\mathbb{Z}_2$-acyclic, so that every track separates $\Sigma$.) There is a canonical map $\phi : V(\Sigma) \longrightarrow V(T)$, where two vertices of $\Sigma$ get mapped to the same vertex of $T$ if and only if they are not separated by any element of $\mathcal{A}$. Note that $\Gamma$ acts with finite edge stabilisers on $T$.

Suppose that $A \in \mathcal{E}$. Let $I_A \subseteq E(T)$ be the set of edges of $T$ that correspond to the connected components of $t(A)$. Now $I_A$ in turn determines an element, $B(A) \in B(T)$, with the property that $I(B(A)) = I_A$ and such that $\phi(A) \subseteq B(A)$. It follows that $A = \phi^{-1}(B(A))$. Let $\mathcal{F} = \{B(A) \mid A \in \mathcal{E}\}$. Thus, $\mathcal{F}$ is a nested subset of $B(T)$. By Lemma 10.1, some cofinite $\Gamma$-invariant subset of $\mathcal{F}$ is sufficient to generate the Boolean algebra generated by $\mathcal{F}$. The corresponding elements of $\mathcal{E}$ now generate $\mathcal{A}$. (This follows because $A = \phi^{-1}(B(A))$ for all $A \in \mathcal{E}$, so that any relation between the elements of $\mathcal{F}$ also holds between the corresponding elements of $\mathcal{E}$.) We see that $\mathcal{A}$ has a cofinite generating set, as required. This proves Lemma 11.1.

**Corollary 11.2.** Let $V$ be a fine $\mathbb{Z}_2$-homologically 1-connected $\Gamma$-set with finite pair stabilisers. Then any $\Gamma$-invariant subalgebra of $S(V)$ has a cofinite nested set of generators which is discrete as a protree.

**Proof.** As described in Section 2, $V$ can be embedded in a $\mathbb{Z}_2$-acyclic 2-complex which is locally finite away from $V$. Let $\mathcal{A}' = \{A \in S(V(\Sigma)) \mid V \cap A \in \mathcal{A}\}$. By Lemma 9.4, the map $[A \longrightarrow V \cap A]$ is an epimorphism from $\mathcal{A}'$ to $\mathcal{A}$. A nested set of generators for $\mathcal{A}'$ as given by Lemma 11.1 then gives us the required generating set for $V$. $\square$
12. Convergence actions on Cantor sets in the finitely presented case.

In this section, we shall give proofs of the main results stated in Section 1. Suppose the group \( \Gamma \) acts as a convergence group on the Cantor set, \( M \). We write \( \Pi \subseteq M \) for the set of non-conical limit points, and \( \mathcal{B}(M) \) for the Boolean algebra of clopen sets of \( M \). We suppose that there is a cofinite \( \Gamma \)-invariant collection, \( \mathcal{G} \), of parabolic subgroups of \( \Gamma \) such that \( \Gamma \) is almost finitely presented relative to \( \mathcal{G} \). As a \( \Gamma \)-set, \( \mathcal{G} \) is isomorphic to a \( \Gamma \)-invariant subset, \( \Pi_0 \), of \( \Pi \). By Lemma 4.2, \( \Pi_0 \) is fine.

Firstly consider the case where \( \Pi_0 \neq \emptyset \). By hypothesis, \( \Pi_0 \) is \( \mathbb{Z}_2 \)-homologically 1-connected. Let \( A = \{ A \cap \Pi_0 \mid A \in \mathcal{B}(M) \} \). Since \( \Pi_0 \) is dense in \( M \), \( A \) is isomorphic to \( \mathcal{B}(M) \) as a \( \Gamma \)-Boolean algebra, and so \( \Xi(A) \) is \( \Gamma \)-equivariantly homeomorphic to \( \Xi(\mathcal{B}(M)) \cong M \). Moreover, applying Lemma 4.1, we see that \( A \) is a subalgebra of \( S(\Pi_0) \). By Corollary 11.2, \( A \) has a cofinite \( \Gamma \)-invariant nested set of generators, \( \mathcal{E} \), which is discrete as a protree. It can thus be identified with the directed edge set of a cofinite simplicial tree, \( T \), with finite edge stabilisers. We can construct the space \( \Xi(\mathcal{E}) \) as in Section 7, where \( \mathcal{E} \) is viewed as a protree. As in Section 6, we see that \( \Xi(\mathcal{E}) \) be identified with \( \Delta T \), and so \( \Delta_0 T \) is precisely \( \Xi(\mathcal{E}) \) with its isolated points removed. Moreover, in Section 7, we saw that \( \Xi(A) \) can be canonically embedded in \( \Xi(\mathcal{E}) \) so that every point of \( \Xi(\mathcal{E}) \setminus \Xi(A) \) is isolated. Since \( \Xi(A) \equiv M \) is a Cantor set, we see that \( M \) can also be identified with \( \Xi(\mathcal{E}) \) with its isolated points removed, and hence with \( \Delta_0 T \).

Now, \( \Gamma \) is hyperbolic relative to the infinite vertex stabilisers and its boundary as a relatively hyperbolic group is precisely \( \Delta_0 T \) (see Section 6). We have thus shown that the boundary is \( \Gamma \)-equivariantly homeomorphic to \( M \). (In retrospect, we see that \( \Pi \) is precisely the set of parabolic points.)

In the case where \( \Pi_0 = \emptyset \), we can find a \( \mathbb{Z}_2 \)-acyclic complex on which \( \Gamma \) acts freely. If \( \Pi \neq \emptyset \), we take a \( \Gamma \)-equivariant map \( \phi : V(\Sigma) \longrightarrow \Pi \). Again, the image of \( \phi \) is a fine \( \Gamma \)-set. Let \( A \) be the Boolean algebra \( \{ \phi^{-1} A \mid A \in \mathcal{B}(M) \} \). Again, \( A \) is a subalgebra of \( S(V) \) isomorphic to \( \mathcal{B}(M) \) and the argument proceeds as before.

It remains to deal with the case where \( \Pi = \emptyset \), in other words, every point of \( M \) is a conical limit point. By [Bo1] it follows that \( \Gamma \) is hyperbolic with boundary \( M \). It is now an easy consequence of Dunwoody’s accessibility theorem [Du2] that \( \Gamma \) is virtually free and that \( M \equiv \partial \Gamma \) may be identified with the space of ends of \( \Gamma \).

This concludes the proofs of Theorems 1.4 and 1.1. We immediately deduce Theorem 1.3. Theorem 1.2 follows from Theorem 1.4 and the discussion at the end of Section 5.
13. Convergence actions arising from protrees.

Suppose \( \Theta \) is a \( \Gamma \)-protree such that the stabiliser of each element is finite. We say that \( \Theta \) is **locally discrete** if each cofinite subprotree is discrete. Note that \( \Gamma \) acts on \( \Xi(\Theta) \) by homeomorphism.

**Proposition 13.1.** If \( \Theta \) is locally discrete, then \( \Gamma \) acts on \( \Xi(\Theta) \) as a convergence group.

**Proof.** By definition, \( \Xi(\Theta) \) is dual to the Boolean ring \( \mathcal{F}(\Theta) \) defined in Section 7. It is easily verified that the action of \( \Gamma \) on \( \mathcal{F}(\Theta) \) is a convergence action, as defined at the end of Section 4. \( \square \)

If \( \Theta \) is countable, then we can write it as an increasing union, \( \Theta = \bigcup_{n=1}^{\infty} \Theta_n \), of cofinite discrete \( \Gamma \)-protrees, \( \Theta_n \). We can identify \( \Theta_n \) as the directed edge set of a simplicial \( \Gamma \)-tree, \( T_n \). We see that \( \Xi(\Theta) \) is an equivariant inverse limit of the spaces \( \Xi(\Theta_n) \cong \Delta T_n \), and that \( \Xi_0(\Theta) \) is an equivariant inverse limit of the spaces \( \Delta_0 T_n \). (This gives another proof of the fact that \( \Gamma \) acts as a convergence group on \( \Xi(\Theta) \).) We see that the action of \( \Gamma \) on \( \Xi(\Theta) \) is an inverse limit of geometrically finite actions.

Note that an inaccessible group admits a locally discrete action on a non-discrete protree. Dunwoody’s example of a finitely generated inaccessible group \[Du3\] thus gives an example of a non-geometrically finite action of such a group on a Cantor set.

We show that examples of this type are typical of convergence actions of (relatively) finitely generated groups on Cantor sets:

**Theorem 13.2.** Suppose that \( \Gamma \) acts as a minimal convergence group on a Cantor set, \( M \), and that \( \mathcal{G} \) is a finite collection of parabolic subgroups. Suppose that \( \Gamma \) is finitely generated relative to \( \mathcal{G} \). Then, \( \Gamma \) admits a locally discrete action on a countable protree, \( \Theta \), such that \( M \) is equivariantly homeomorphic to \( \Xi(\Theta) \).

In particular, the action of \( \Gamma \) on \( M \) is an inverse limit of geometrically finite actions.

The proof of Theorem 13.2 proceeds exactly as with that of Theorem 1.4 (and 1.1) as described in Section 12, except that, in this case, we have to make do with a (fine) connected cofinite \( \Gamma \)-graph, \( K \), instead of the 2-complex, \( \Sigma \). As before, \( M \) is equivariantly homeomorphic to \( \Xi(\mathcal{A}) \) where \( \mathcal{A} \) is a \( \Gamma \)-subalgebra of the Boolean algebra of \( K \)-breaks. Theorem 9.3 gives us an invariant nested generating set, \( \mathcal{E} \) of \( \mathcal{A} \), which has the structure of a protree, \( \Theta \), as described in Section 7. We can canonically identify \( \Xi(\Theta) \) as a closed subset of \( \Xi(\mathcal{A}) \cong M \), whose complement consists of isolated points and is thus empty in this case. We have equivariantly identified \( \Xi(\Theta) \) with \( M \) as required.
14. Other applications.

In this section, we sketch two further applications of Theorem 9.3. One concerns constructions of group splittings, and the other relates to the Almost Stability Theorem of [DiD].

Suppose that $\Gamma$ is a one-ended finitely generated group, and that $G \leq \Gamma$ is any subgroup. Let $X$ be a Cayley graph of $\Gamma$ (or any cofinite locally finite $\Gamma$-graph). As in [Bo4] (cf. [DuS]) we say that $G$ has codimension-one in $\Gamma$ if there is a connected $G$-invariant subset, $Y \subseteq X$, such that $Y/G$ is compact, and such that $X \setminus Y$ has at least two distinct components neither of which is contained in a uniform neighbourhood of $Y$. (This is independent of the choice of $X$.) The following result also follows directly from a result of Niblo [N]:

**Proposition 14.1.** Suppose that $\Gamma$ is finitely generated and that $G \leq \Gamma$ is a codimension-one subgroup such that $\Gamma$ is the commensurator of $G$. Then, $\Gamma$ splits nontrivially as a graph of groups with $G$ conjugate into one of the vertex groups.

The “commensurator” condition means that $G \cap gGg^{-1}$ has finite index in $G$ for all $g \in \Gamma$. If we assume that no vertex group is equal to an incident edge group, then it follows that all the vertex and edge groups will be commensurate with $G$.

To prove Proposition 14.1, let $K$ be a coset graph of $G$ in $\Gamma$. In other words, $K$ is a connected cofinite $\Gamma$-graph, with $V(K)$ isomorphic as a $\Gamma$-set to the set of left translates of $G$, with $\Gamma$ acting by left multiplication. Let $x \in V(K)$ be a vertex stabilised by $G$. Let $Y \subseteq X$ be as in the hypotheses. Thus, we can write $X \setminus Y = P \sqcup Q$ where neither $P$ nor $Q$ is contained in any uniform neighbourhood of $Y$. Now all but finitely many $\Gamma$-images of $Y$ are disjoint from $Y$, and hence contained in either $P$ or $Q$. Each $\Gamma$-image of $Y$ corresponds to a vertex of $K$. This therefore assigns all but finitely many elements of $V(K)$ to one of two disjoint infinite subsets, $A,B \subseteq V(K)$, corresponding to $P$ and $Q$ respectively. Assigning the remaining vertices arbitrarily, we can suppose that $B = A^*$.

If $y \in A$ and $z \in A^*$, then any path connecting the corresponding images of $Y$ in $X$ must intersect $Y$. Since $Y/G$ is finite, it follows that only finitely many $\Gamma$-images of any pair $\{y,z\} \subseteq V(K)$ can meet both $A$ and $A^*$. In other words, $A$ is a slice. We have shown that the algebra of slices, $B(K)$, of $K$ is nontrivial, Theorem 9.3 now applies to give us a nested set of generators for $B(K)$. Since the set of $\Gamma$-images of $Y$ is locally finite in $X$, it follows that this generating set is locally discrete. We thus get an action of $\Gamma$ on a simplicial tree, $T$. Moreover, there is an equivariant map from $V(K)$ to $V(T)$. This proves Proposition 14.1.
This result is, in some sense, “orthogonal” to the constructions of [Bo5]. Note that we cannot expect the splittings we obtain in this case to be canonical.

Another application of Theorem 9.3 (as observed by Dunwoody) is to give an alternative proof of a version of the Almost Stability Theorem of [DiD]. This can be interpreted as giving a criterion for a \( \Gamma \)-set to be embedded in the vertex set of a simplicial \( \Gamma \)-tree with finite edge stabilisers. Let \( \Gamma \) be a group, and \( X \) be a cofinite \( \Gamma \)-set. Let \( \mathcal{P}(X) \) be the power set of \( X \), thought of as a \( \Gamma \)-Boolean algebra, and let \( \mathcal{I} \) be the ideal of finite subsets.

**Proposition 14.2.** Suppose that \( \Gamma \) is a finitely generated group, and that \( X \) is a \( \Gamma \)-set with finite point stabilisers. Suppose that \( V \subseteq \mathcal{P}(X) \) is a \( \Gamma \)-invariant subset of \( \mathcal{P}(X) \) with the property that if \( A, B \in V \) then \( A + B \in \mathcal{I} \). Then \( V \) can be equivariantly embedded in the vertex set of a simplicial \( \Gamma \)-tree.

Suppose we already know that \( V \) embeds in a simplicial \( \Gamma \)-tree, \( T \). We can take \( X \) to be the directed edge set of \( T \). To each \( x \) in \( V \), we associate the set of directed edges which point towards \( x \). This gives a subset of \( \mathcal{P}(X) \) satisfying the hypotheses of Proposition 13.2. Of course, the situation described by the hypotheses may be more general that this.

Proposition 14.2 can be proven as follows. Given any \( x \in X \), let \( A(x) \in \mathcal{P}(V) \) be the subset of \( V \) consisting of those elements of \( V \) that contain \( x \). One verifies that \( A(x) \) is a slice of \( V \). Let \( A \) be the subalgebra of slices of \( V \) generated by \( \{ A(x) \mid x \in V \} \). (A typical element of \( A \) has the form \( B \) or \( B^* \), where \( B \) is either finite or is equal to \( A(x) \) for some \( x \in X \).) We now apply Theorem 9.3. It is easily verified that the resulting generating set is locally discrete.

**References**


[Ger] V. Gerasimov, Manuscript (Russian), Belo Horizonte.


Received February 10, 2001 and revised September 20, 2001.
FACULTY OF MATHEMATICAL STUDIES
UNIVERSITY OF SOUTHAMPTON
HIGHFIELD, SOUTHAMPTON SO17 1BJ
ENGLAND
E-mail address: B.H.Bowditch@maths.soton.ac.uk
ON THE DIOPHANTINE EQUATION $\frac{x^m - 1}{x - 1} = \frac{y^n - 1}{y - 1}$

YANN BUGEAUD AND T.N. SHOREY

We study the Diophantine equation $\frac{x^m - 1}{x - 1} = \frac{y^n - 1}{y - 1}$ in integers $x > 1$, $y > 1$, $m > 1$, $n > 1$ with $x \neq y$. We show that, for given $x$ and $y$, this equation has at most two solutions. Further, we prove that it has finitely many solutions $(x, y, m, n)$ with $m > 2$ and $n > 2$ such that $\gcd(m - 1, n - 1) > 1$ and $(m - 1)/(n - 1)$ is bounded.

1. Introduction.

Goormaghtigh [7] observed in 1917 that

$$31 = \frac{2^5 - 1}{2 - 1} = \frac{5^3 - 1}{5 - 1} \quad \text{and} \quad 8191 = \frac{2^{13} - 1}{2 - 1} = \frac{90^3 - 1}{90 - 1}$$

are two solutions of the Diophantine equation

$$(1) \quad \frac{x^m - 1}{x - 1} = \frac{y^n - 1}{y - 1} \quad \text{in integers } x > 1, y > 1, m > 2, n > 2 \text{ with } x \neq y.$$ 

There is no restriction in assuming that $y > x$ in (1) and thus we have $m > n$. This equation asks for integers having all their digits equal to one with respect to two distinct bases and we still do not know whether or not it has finitely many solutions. Even if we fix one of the four variables, it remains an open question to prove that (1) has finitely many solutions.

However, when either the bases $x$ and $y$, or the base $x$ and the exponent $n$, or the exponents $m$ and $n$ are fixed, then it is proved that (1) has finitely many solutions (see [3] for references). In the first two cases, thanks to Baker’s theory of linear forms in logarithms, we can compute explicit (huge) upper bounds for the size of the solutions. As for the number of solutions, Shorey [14] proved that for two integers $y > x$, the Diophantine equation

$$(2) \quad \frac{x^m - 1}{x - 1} = \frac{y^n - 1}{y - 1} \quad \text{in integers } m > 1, n > 1,$$

has at most 17 solutions, independently of $x$ and $y$. One of the purposes of the present work is to considerably improve this estimate by showing that (2) has at most one solution provided that $y$ is large enough and that otherwise (2) has at most two solutions.
When the exponents $m$ and $n$ are fixed, Davenport, Lewis and Schinzel [6] proved that (1) has finitely many solutions, but their proof rests on a theorem of Siegel and it is ineffective. However, when $\gcd(m - 1, n - 1) > 1$, they are able to replace Siegel’s result by an effective argument due to Runge.

**Theorem DLS.** Equation (1) with $\gcd(m - 1, n - 1) > 1$ implies that $\max(x, y)$ is bounded by an effectively computable number depending only on $m$ and $n$.

Recently, this has been improved by Nesterenko & Shorey [13] as follows.

**Theorem NS.** Let $d \geq 2$, $r \geq 1$ and $s \geq 1$ be integers with $\gcd(r, s) = 1$. Assume that $m - 1 = dr$ and $n - 1 = ds$. If $(x, y, m, n)$ satisfy (1), then $\max\{x, y, m, n\}$ is bounded by an effectively computable number depending only on $r$ and $s$.

This is the first result of the type where there is no restriction on the bases $x$ and $y$ and the exponents $m$ and $n$ extend over an infinite set. In the present work, we show that the assertion of Theorem NS continues to be valid when the ratio $(m - 1)/(n - 1)$ is bounded.

### 2. Statement of the results.

Our first result deals with the number of solutions of Equation (2) and improves a previous estimate of Shorey [14].

**Theorem 1.** Let $y > x > 1$ be integers. If $\gcd(x, y) > 1$ or if $y \geq 10^{11}$, then (2) has at most one solution. Further, if $y \geq 7$, then (2) has at most two solutions. Finally, the only solutions of (2) with $y \leq 6$ are given by $(x, y, m, n) = (2, 5, 5, 3)$ or $(2, 6, 3, 2)$.

M. Mąkowski and A. Schinzel [12] proved that (2) with $y \leq 10$ and $m > n > 2$ has only the solution $(x, y, m, n) = (2, 5, 5, 3)$, however, for sake of completeness, we give a proof of the last statement of Theorem 1. Our proof is based on an idea of Le [10] and it combines the theory of linear forms in logarithms together with a strong gap principle proved by elementary means. As is apparent from the proof, one can derive several other interesting statements including the following.

**Theorem 2.** Let $y > x > 1$ be coprime integers and assume that the smallest integer $s \geq 1$ such that $x^s \equiv 1 \pmod{y}$ satisfies

\[ s > 48^{11}(\log y)^2(\log \log y). \]

Then Equation (2) has no solution.

**Remark 1.** We point out that for given $y$ sufficiently large, Theorem 2 solves (2) for a wide set of integers $x$, indeed for at least $\varphi(y) - \varphi(y)^{2/3}$ integers $x$, with $1 \leq x \leq y$ and $\gcd(x, y) = 1$. Here, $\varphi$ denote the Euler
totient function. A proof of this assertion is given just after the proofs of Theorems 1 & 2.

**Remark 2.** Theorem 1 implies that the equations

\[
\frac{2^m - 1}{2 - 1} = \frac{5^n - 1}{5 - 1} \quad \text{and} \quad \frac{2^m - 1}{2 - 1} = \frac{90^n - 1}{90 - 1}
\]

have exactly one solution in integers \( m > 1 \) and \( n > 1 \), namely \((m,n) = (5,3)\) and \((13,3)\), respectively.

**Remark 3.** M. Makowski and A. Schinzel [11] proved that (2) with \( y \leq 10 \) and \( m > n > 2 \) has only the solution \((x,y,m,n) = (2,5,5,3)\), however, for sake of completeness, we give a proof of the last statement of Theorem 1.

In Theorem NS quoted in the Introduction, the exponents \( m \) and \( n \) are allowed to vary such that the ratio \((m-1)/(n-1)\) is constant. This condition also implies that \( y \) cannot be too large compared with \( x \). We now present two new results under a similar hypothesis. The first one can be seen as an improvement of Theorem NS, although it does not imply the latter. The second one is of a different nature, namely, there is no restriction on the exponents \( m \) and \( n \) and the bases \( x \) and \( y \) extend to an infinite set.

**Theorem 3.** Let \( \alpha > 1 \). Equation (1) with \( \gcd(m-1,n-1) \geq 4\alpha + 6 + \frac{1}{\alpha} \) and \((m-1)/(n-1) \leq \alpha\) implies that \( \max(x,y,m,n) \) is bounded by an effectively computable number depending only on \( \alpha \).

Theorem 3 does not contain Theorem NS because of the condition imposed on \( \gcd(m-1,n-1) \). If \( r \) and \( s \) are fixed, we may remove this condition by Theorem DLS and then Theorem NS follows from Theorem 3. In the course of the proof of Theorem 3, we need an auxiliary result (namely, Lemma 4 below) which enables us to considerably improve Theorem 2 of [13].

**Theorem 4.** Let \((x,y,m,n)\) be a solution of (1) with \( y > x \). Then we have

\[ \gcd(m-1,n-1) \leq 33.4 m^{1/2}. \]

**Remark 4.** Theorem 2 of [13] only asserts that there exist an effectively computable absolute constant \( C \) such that

\[ \gcd(m-1,n-1) \leq C m^{4/5} (\log m)^{3/5}. \]

Further, its proof combines the theory of linear forms in logarithms together with sharp upper bounds for the size of the solutions of (1) obtained by Runge’s method, whereas the proof of Theorem 4 depends only on estimates for linear forms in two logarithms.

**Theorem 5.** Let \( a > 1 \). Let \( y > x > 1 \) be integers such that \( x \) divides \( y - 1 \) and \( y \leq x^a \). If \((x,y,m,n)\) satisfies (1), then

\[ n < m \leq 14000 a^2 (\log 3a)^2 \]
and
\[ x < n, \quad y < n^a. \]

**Remark 5.** It follows from Theorem 5 that for a given \( a > 1 \), Equation (1) has only finitely many solutions \((x, y, m, n)\) with \( y \leq x^a \) and \( x \mid (y - 1) \).

### 3. Auxiliary lemmas.

We begin with the following theorem of Baker and Wüstholz [2] on linear forms in logarithms.

**Lemma 1.** Let \( \alpha_1, \ldots, \alpha_d \) be positive rational numbers of heights not exceeding \( A_1, \ldots, A_d \), respectively, where \( A_j \geq e \) for \( 1 \leq j \leq d \). Put
\[
\Omega = \prod_{j=1}^{d} \log A_j.
\]
Then the inequalities
\[
0 < |b_1 \log \alpha_1 + \cdots + b_d \log \alpha_d| < \exp\left(-16(d+2)^2 \Omega \log B\right)
\]
have no solution in integers \( b_1, \ldots, b_d \) of absolute values not exceeding \( B \), where \( B \geq e \).

In order to prove Theorems 1 & 2, we need an explicit estimate for the size of the solutions of (2).

**Lemma 2.** Let \( y > x > 1 \) be integers and let \((m, n)\) be a solution of (2). Then we have
\[
m \leq 2 \times 48^{10\log y^2 \log m} + 1.
\]

**Proof.** We rewrite (2) as
\[
\frac{x^m}{x - 1} - \frac{y^n}{y - 1} = \frac{1}{x - 1} - \frac{1}{y - 1},
\]
thus we get
\[
0 < 1 - y^n x^{-m} \left( \frac{x - 1}{y - 1} \right) < x^{-m},
\]
which implies that
\[
0 < \left| n \log y - m \log x + \log \left( \frac{x - 1}{y - 1} \right) \right| < 2x^{-m}. \tag{5}
\]
Now we apply Lemma 1 with \( d = 3, A_1 = y, A_2 = x + 1, A_3 = y \) and \( B = m \) to derive
\[
\left| n \log y - m \log x + \log \left( \frac{x - 1}{y - 1} \right) \right| > \exp\left(-48^{10\log y^2 \log(x + 1) \log m}\right)
\]
which, combined with (5), yields the estimate stated in the lemma. \( \square \)
Apart from Lemma 1, we shall also need the following refinement, due to Mignotte [12], of a theorem of Laurent, Mignotte & Nesterenko [9] on linear forms in two logarithms.

**Lemma 3.** Consider the linear form

$$\Lambda = b_2 \log \alpha_2 - b_1 \log \alpha_1,$$

where $b_1$ and $b_2$ are positive integers. Suppose that $\alpha_1$ and $\alpha_2$ are multiplicatively independent. Put

$$D = [Q(\alpha_1, \alpha_2) : Q]/[R(\alpha_1, \alpha_2) : R].$$

Let $a_1, a_2, h, k$ be real positive numbers, and $\rho$ a real number $> 1$. Put $\lambda = \log \rho$, $\chi = h/\lambda$ and suppose that $\chi \geq \chi_0$ for some number $\chi_0 \geq 0$ and that

$$h \geq D \left( \log \left( \frac{b_1}{a_2} + \frac{b_2}{a_1} \right) + \log \lambda + f(\lceil K_0 \rceil) \right) + 0.023,$$

$$a_i \geq \max\{1, \rho | \log | \alpha_i | - \log | \alpha_i | + 2Dh(\alpha_i)\}, \quad (i = 1, 2),$$

$$a_1a_2 \geq \lambda^2$$

where

$$f(x) = \log \left( 1 + \sqrt{x-1} \right) \sqrt{x} + \log \frac{x}{6(x-1)} + \frac{3}{2} + \log 3 + \frac{\log \frac{x}{x-1}}{x-1},$$

and

$$K_0 = \frac{1}{\lambda} \left( \frac{\sqrt{2} + 2\chi_0}{3} + \sqrt{\frac{2(1 + \chi_0)}{9} + \frac{2\lambda}{3} \left( \frac{1}{a_1} + \frac{1}{a_2} \right) + \frac{4\lambda \sqrt{2 + \chi_0}}{3\sqrt{a_1a_2}}} \right)^2 a_1a_2.$$

Put

$$v = 4\chi + 4 + 1/\chi \quad \text{and} \quad m = \max\{2^{5/2}(1 + \chi)^{3/2}, (1 + 2\chi)^{5/2}/\chi\}.$$ 

Then we have the lower bound

$$\log |\Lambda| \geq -\frac{1}{\lambda} \left( \frac{v}{6} + \frac{1}{2} \sqrt{\frac{v^2}{9} + \frac{4\lambda v}{3} \left( \frac{1}{a_1} + \frac{1}{a_2} \right) + \frac{8\lambda m}{3\sqrt{a_1a_2}}} \right)^2 a_1a_2$$

$$- \max\left\{ \lambda(1.5 + 2\chi) + \log \left( (2 + 2\chi)^{3/2} + (2 + 2\chi)^2 \sqrt{k^*} \right) A + (2 + 2\chi), D \log 2 \right\}$$

where

$$A = \max\{a_1, a_2\} \quad \text{and} \quad k^* = \frac{1}{\lambda^2} \left( \frac{1 + 2\chi}{3\chi} \right)^2 + \frac{1}{\lambda} \left( \frac{2}{3\chi} + \frac{2(1 + 2\chi)^{1/2}}{\chi} \right).$$

**Proof.** This is Theorem 2 of [12]. 

We apply Lemma 3 for deriving the following result:
Lemma 4. Let $\alpha > 1$ and $d > 1$ be an integer. Suppose that $(x, y, m, n)$ with $y > x$ is a solution of (1). Assume that

$\gcd(m - 1, n - 1) = d, \quad \frac{m - 1}{n - 1} \leq \alpha.$

Then we have

$$d \leq 743 \left( \alpha + \frac{1}{2} \right).$$

Proof. Let $\alpha > 1$ and $d > 1$ be an integer. We suppose that (1) with $\gcd(m - 1, n - 1) = d$ and $(m - 1)/(n - 1) \leq \alpha$ is satisfied and we put

$$m - 1 = dr, \quad n - 1 = ds$$

where $r$ and $s$ are positive integers. In view of the last assertion of Theorem 1, whose proof is independent of Lemma 3, we may assume that $y \geq 7$. We write (1) as

$$\frac{x}{x - 1} x^r d - \frac{y}{y - 1} y^s d = \frac{1}{x - 1} - \frac{1}{y - 1}$$

which implies that

$$0 < \log \frac{x(y - 1)}{y(x - 1)} - d \log \frac{y^s}{x^r} < y^{-sd}.$$ (6)

Now we apply Lemma 3 with $b_1 = d, \ b_2 = 1, \ \alpha_1 = y^s/x^r$ and $\alpha_2 = (x(y - 1))/(y(x - 1))$ in order to get a lower bound for $\Lambda = b_2 \log \alpha_2 - b_1 \log \alpha_1$. We observe that $h(\alpha_1) \leq s \log y, \ h(\alpha_2) \leq 2 \log y$ and that $\alpha_1$ and $\alpha_2$ are multiplicatively independent. Further, we put

$$\rho = 1 + \frac{3 \log y}{4 \log(1 + \frac{1}{x - 1})}.$$ (7)

Then we may take

$$a_1 = (2s + 3/4) \log y \quad \text{and} \quad a_2 = (19 \log y)/4.$$ (8)

Thus, we check that

$$\log x \leq \log \left( 1 + \frac{3}{4} (x - 1) \log y \right) \leq \lambda := \log \rho \leq \frac{4}{3} \log y$$

and

$$a_1 a_2 \geq 7 \lambda^2, \quad a_1 a_2 / \lambda \geq 19.$$ (9)

By (2), we observe that $y^{n-1} \leq 2x^{m-1}$ which implies that

$$\frac{\log y}{\log x} \leq \alpha + \frac{1}{2}.$$ (10)

We may assume that $d \geq 30$. By (7) and since $[K_0] \geq 16$, in order to apply Lemma 3, we have to choose the parameter $h$ such that

$$h \geq \log d + 0.3.$$
Assume first that $\lambda \geq \log d + 0.3$ and apply Lemma 3 with $h = \lambda$ and $\chi = 1$. Thus $v = 9$, $m = 16$ and, using that $y \geq 7$ in the definition of $\rho$, we infer that $k^* \leq 2.4$. Now, we notice that $\lambda/a_1 + \lambda/a_2 \leq 0.77$ and use (8), to obtain that

$$\log \Lambda > -\frac{19.65}{\lambda} a_1 a_2 - 3.5 \lambda - \log(56.7 a_1 + 4).$$

Finally, using $y \geq 7$, we conclude from (7), (9), (10) and (6) that

$$d \leq 264 \left( \frac{1}{2} + \frac{3}{48} \right) \left( \alpha + \frac{1}{2} \right).$$

We observe that the right hand side of the preceding inequality does not exceed $264(\alpha + 1/2)$. Therefore, we may suppose that $\lambda < \log d + 0.3$, that is, in view of (7)

$$\log \left( 1 + \frac{3 \log y}{4 \log (1 + \frac{1}{x-1})} \right) < \log d + 0.3.$$  

Next, we show that

$$d \leq 135 \log y.$$  

For the proof of (12), we assume that $d > 135 \log y$ and we shall arrive at a contradiction. We apply Lemma 3 once again, but with another choice of the radius $\rho$. Namely, we take $\rho = e^{14}$. Then $\lambda = 4$. We set

$$h = \log \left( \frac{b_1}{a_2} + \frac{b_2}{a_1} \right) + \frac{57}{20}$$

and we see that $h \geq 6$, thus we may choose $\chi_0 = \frac{3}{7}$. We observe that $a_1 a_2 \geq 49$ and $[K_0] \geq 27$. Further, we have $v \leq h + \frac{14}{7}$, $m \leq 2^{5/2}(1 + \chi)^{3/2}$ and we put $H = h + \frac{14}{7}$. Now, Lemma 3 yields

$$\log \Lambda \geq -\frac{3}{40} H^2 a_1 a_2 - 6 - 2 h - \log(2 h^2 a_1 a_2) \geq -\frac{17}{200} H^2 a_1 a_2,$$

since $H \geq \frac{32}{7}$ and $a_1 a_2 \geq 49$. Combining the preceding estimate with (6), we get

$$\frac{d}{\log y} \leq \frac{57}{50} \left( \log \left( \frac{4d}{19 \log y} + \frac{1}{5} \right) + \frac{451}{60} \right)^2.$$  

This is not possible and the proof of (12) is complete.

We combine (7), (11) and (12) to conclude that

$$\frac{3}{4} (x - 1) \log y < 183 \log y.$$
Thus $x \leq 244$. Finally, we apply (9) and (12) to conclude that $d \leq 743(\alpha + 1/2)$.

In addition to Lemma 4, the proof of Theorem 3 uses an irrationality measure [15] of certain algebraic numbers derived from a Theorem of Baker [1].

**Lemma 5.** Let $A, B, K$ and $n$ be positive integers such that $A > B, K < n, n \geq 3$ and $\omega = (B/A)^{1/n}$ is not a rational number. For $0 < \phi < 1$, put

$$\delta = 1 + \frac{2 - \phi}{K}, \quad \sigma = \frac{\delta}{1 - \phi}$$

and

$$u_1 = 40^n((K+1)(\sigma+1)/(K\sigma-1)), \quad u_2^{-1} = K^{2K+\sigma+1}40^n(K+1).$$

Assume that

$$A(A-B)^{-\delta}u_1^{-1} > 1.$$ 

Then

$$\left| \omega - \frac{p}{q} \right| > \frac{u_2}{Aq^K(\sigma+1)}$$

for all integers $p$ and $q$ with $q > 0$.

**Proof.** This is Lemma 1 of Shorey & Nesterenko [15]. We notice that this has been refined by Hirata-Kohno in [8] but the statement of [15] is sufficient for our purpose.

Our last auxiliary result is an estimate from the theory of $p$-adic linear forms in (two) logarithms. For this, we need some notation.

Let $m > 1$ be an integer and write $m = p_1^{u_1} \cdots p_w^{u_w}$, where $p_1 < \cdots < p_w$ are distinct prime numbers and the $u_i$’s are positive integers. Let $x$ be a nonzero integer and let $p$ be a prime. We recall that the $p$-adic valuation of $x$, denoted by $v_p(x)$, is the greatest nonnegative integer $v$ such that $p^v$ divides $x$. Analogously, we define the $m$-adic valuation of $x$, which we denote by $v_m(x)$, to be the greatest nonnegative integer $v$ such that $m^v$ divides $x$.

We observe that

$$v_m(x) = \min_{1 \leq i \leq w} \left[ \frac{v_{p_i}(x)}{u_i} \right],$$

where $[\cdot]$ denotes the integer part. Further, if $a/b$ is a nonzero rational number with $a$ and $b$ coprime, we set $v_m(a/b) = v_m(a) - v_m(b)$.

We let $x_1/y_1$ and $x_2/y_2$ be two nonzero rational numbers with $x_1/y_1 \neq \pm 1$. Lemma 6 provides an upper bound for the $m$-adic valuation of

$$\Lambda = \left( \frac{x_1}{y_1} \right)^{b_1} - \left( \frac{x_2}{y_2} \right)^{b_2},$$
where $b_1$ and $b_2$ are positive integers, assuming that
\[v_{p_i} \left( \frac{x_1}{y_1} - 1 \right) \geq u_i, \quad v_{p_i} \left( \frac{x_2}{y_2} - 1 \right) \geq 1\]
for all prime $p_i$, $1 \leq i \leq w$ and that either $m$ is odd or $4$ divides $m$. Further, let $A_1 > 1, A_2 > 1$ be real numbers such that
\[
\log A_i \geq \max \{\log |x_i|, \log |y_i|, \log m\}, \quad (i = 1, 2),
\]
and put
\[b' = \frac{b_1}{\log A_2} + \frac{b_2}{\log A_1}.\]

**Lemma 6.** With the previous notation and under the above hypothesis, if moreover $m$, $b_1$ and $b_2$ are relatively prime, then we have the upper estimate
\[
v_m(\Lambda) \leq 66.8 \left( \frac{\log b' + \log m + 0.64}{\log m} \right)^2 \log A_1 \log A_2.
\]

**Proof.** This is Theorem 2 of [4]. \qed

### 4. Proofs.

**Proofs of Theorems 1 and 2.** Let $y > x \geq 2$ be integers and denote by $d$ their greatest common divisor. Set $x_0 = x/d$ and $y_0 = y/d$. Assume that (2) has two distinct solutions $(m_2, n_2)$ and $(m_1, n_1)$ with $m_2 > m_1 > 1$. Then we get
\[
(13) \quad (dy_0 - 1)(dx_0)^{m_2} - (dx_0 - 1)(dy_0)^{n_2} = dy_0 - dx_0
\]
and
\[
(14) \quad (dy_0 - 1)(dx_0)^{m_1} - (dx_0 - 1)(dy_0)^{n_1} = dy_0 - dx_0.
\]

We first observe that $d$ and $x_0$ are coprime. Indeed, let $p$ be a prime such that $p^a \parallel d$ and $p|x_0$. Thus $p^{2a}$ must divide the left-hand side of (13) since $m_2 \geq 2$ and $n_2 \geq 2$. But $p^{a+1}$ divides $dx_0$, hence $p$ must divide $y_0$ contradicting $\gcd(x_0, y_0) = 1$. Similarly, we prove that $d$ and $y_0$ are coprime. Using (14) to replace $dy_0 - dx_0$ in (13) and dividing (13) by $(dx_0 - 1)(dy_0)^{n_1}$, we get
\[
(15) \quad 1 = \frac{(dx_0)^{m_1}(dy_0 - 1)(1 - (dx_0)^{m_2 - m_1})}{(dy_0)^{n_1}(dx_0 - 1)} + (dy_0)^{n_2 - n_1}.
\]

Since $d$ and $x_0y_0$ are coprime, this implies that $d$ divides the right-hand side of (15), hence $d = 1$. Consequently, if $x$ and $y$ have a common factor, then (2) has at most one solution.

In the sequel of the Proof of Theorem 1, we always assume that $x$ and $y$ are coprime. Further, we assume that there exist pairs of positive integers $(m_3, n_3), (m_2, n_2)$ and $(m_1, n_1)$, with $m_3 > m_2 > m_1 \geq 1$ and
\[
\frac{x^{m_j} - 1}{x - 1} = \frac{y^{n_j} - 1}{y - 1}, \quad j = 1, 2, 3.
\]
Since $y > x$, we clearly have $m_{j+1} - m_j \geq 2$ for $j = 1, 2$. Further

\begin{equation}
    n_{j+1} - n_j \geq 2, \quad j = 1, 2.
\end{equation}

Indeed, if for example $n_2 = n_1 + 1$, then we get
\[
    \frac{x^{m_2} - 1}{x - 1} = \frac{x^{m_1} - 1}{x - 1} + y^{n_1},
\]
and $x$ divides $y$, which is a contradiction. This proves (16).

Let $\delta = \gcd(x-1, y-1)$ and put $a = (y-1)/\delta$, $b = (x-1)/\delta, c = (y-x)/\delta$.

Let $s$ denote the smallest integer $\geq 1$ such that
\[
    x^s \equiv 1 \pmod{by^{n_1}}.
\]

Then, for $j = 1, 2$, we have
\begin{equation}
    ax^{m_{j+1}} - by^{n_{j+1}} = c
\end{equation}
and
\begin{equation}
    ax^{m_j} - by^{n_j} = c.
\end{equation}

Therefore
\begin{equation}
    x^{m_{j+1} - m_j} \equiv 1 \pmod{by^{n_j}},
\end{equation}
thus $m_3 - m_2$ and $m_2 - m_1$ are multiples of $s$.

Further, (17) and (18) with $j = 1$ yield
\[
    1 = \frac{ax^{m_1}(1 - x^{m_2 - m_1})}{by^{n_1}} + y^{n_2 - n_1},
\]
from which we deduce that $y$ and $(1 - x^{m_2 - m_1})/(by^{n_1})$ are coprime. By the definition of $s$, we deduce that
\begin{equation}
    \gcd \left( y, \frac{x^s - 1}{by^{n_1}} \right) = 1.
\end{equation}

Write $m_3 - m_2 = su$, with $u \geq 1$ integer. By (19), we see that $by^{n_2}$ divides $x^{su} - 1$, thus $y^{n_2 - n_1}$ divides
\begin{equation}
    (1 + x^s + \cdots + x^{s(u-1)})(x^s - 1)by^{n_1}.
\end{equation}

From (20) and (21), we then obtain that
\begin{equation}
    1 + x^s + \cdots + x^{s(u-1)} \equiv 0 \pmod{y^{n_2 - n_1}}.
\end{equation}

Let $p$ be a prime and assume that $p^\alpha || y$. By the definition of $s$, there exists an integer $\lambda \geq 1$ such that $x^s = 1 + \lambda p^\alpha$. We infer from (22) that $p^{\alpha(n_2 - n_1)}$ divides $(1 + \lambda p^\alpha u - 1)/(\lambda p^\alpha) = u + u(u - 1)\lambda p^\alpha/2 + \cdots$. If $p$ is odd or if $\alpha n_1 \geq 2$, we get that $p^{\alpha(n_2 - n_1)}$ divides $u$. If $p = 2$, then $x$ is odd and $b$ must be even, thus $\lambda$ is also even and, arguing as above, we see that $2^{\alpha(n_2 - n_1)}$ divides $u$. Consequently, $y^{n_2 - n_1}$ divides $u$ and we get
\begin{equation}
    m_3 - m_2 \geq y^{n_2 - n_1}.
\end{equation}
By (17) and (18), we have also
\[ y^{n_{j+1}} - y^{n_j} \equiv 1 \pmod{ax^{m_j}} \]

for \( j = 1, 2 \) and we argue as above to conclude that
\[ n_3 - n_2 \geq x^{m_2-m_1}. \]

Assume that (2) has two solutions \((m_3, n_3)\) and \((m_2, n_2)\) with \( m_3 > m_2 > 1 \). We apply our preceding results with \( m_1 = n_1 = 1 \) and (23) yields
\[ m_3 > y^{n_2-1}. \]

Combined with Lemma 2, (25) yields \( y < 10^{11} \), as claimed.

Assume now that (2) has three solutions \((m_4, n_4)\), \((m_3, n_3)\) and \((m_2, n_2)\) with \( m_4 > m_3 > m_2 > 1 \). By (23) we get
\[ m_4 > y^{n_3-n_2}, \]

and by (24), (2) and (16) we see that
\[ n_3 - n_2 > x^{m_2-1} \geq y^{n_2-1}/2 \geq y^2/2. \]

Finally, we have
\[ m_4 > y^{y^2/2}, \]

which, combined with Lemma 2, yields \( y \leq 6 \).

Now, we solve (2) for any pair \((x, y)\) with \( 2 \leq x < y \leq 6 \). For \((2, 3)\) and \((3, 4)\), we observe that the equation \( |3^u - 2^v| = 1 \) has the only solution \((2, 3)\) in integers \( u > 1 \) and \( v > 1 \). For \((x, y) = (2, 4), (2, 6), (3, 6)\) and \((4, 6)\) we conclude by arguing, respectively, modulo 4, 8, 9 and 4 that (2) has no solution other than the one given by \( x = 2, y = 6, m = 3, n = 2 \). For \((x, y) = (4, 5)\), we have to solve the equation \( 4^u - 3 \cdot 5^v = 1 \). Arguing modulo 5, we see that \( u \) is even, hence \((4^u/2 - 1)(4^u/2 + 1) = 3 \cdot 5^v\) and \( u = 2, v = 1 \). We deal with \((x, y) = (5, 6)\) in a similar manner.

Now, for \((x, y) = (2, 5)\), we are left with the equation \( 5^u + 3 = 2^v \). Modulo 8, we see that \( u \) is odd, hence \((5^{(u-1)/2}, v)\) is a solution of \( 5X^2 + 3 = 2^k \). By Theorem 1 of [5], this equation has only two solutions, namely \((X, k) = (1, 3)\) and \((5, 7)\).

Finally, it remains us to treat the pair \((3, 5)\), hence the equation \( 2 \cdot 3^u - 5^v = 1 \). Modulo 3, we see that \( v \) is odd, thus \((5^{(v-1)/2}, u)\) is a solution of \( 1 + 5X^2 = 2 \cdot 3^k \). By Theorem 2 of [5], this equation has only the solution \((X, k) = (1, 1)\). This completes the proof of Theorem 1.

The Proof of Theorem 2 is now easy. Indeed, set \( m_1 = n_1 = 1 \) and let \((m_2, n_2)\) be a solution of (2). As noticed just below (19), \( s \) divides \( m_2 - 1 \). Thus \( s < m_2 \) which, together with (4), contradicts (3). Theorem 2 is then a straightforward consequence of Lemma 1.
Now, we give a proof of Remark 1. Let \( y \) be a given positive integer. For an integer \( 1 \leq x \leq y \) coprime with \( y \), we denote by \( \text{ord}_y(x) \) the smallest integer \( s \geq 1 \) such that \( x^s \equiv 1 \pmod{y} \). Further, we write
\[
\sum_{d|\varphi(y),d<\sqrt{\varphi(y)}} \varphi(d) + \sum_{d|\varphi(y),d\geq\sqrt{\varphi(y)}} \varphi(d) = \varphi(y),
\]
and we denote by \( A_1 \) (resp. \( A_2 \)) the first (resp. the second) summation in the above formula. We observe that
\[
A_1 \leq D(\varphi(y))\sqrt{\varphi(y)},
\]
where \( D(n) \) is the number of divisors of the integer \( n \). Thus, for \( y \) large enough, we have \( A_1 \leq \varphi(y)^{2/3} \).

For any positive integer \( x \leq y \) coprime to \( y \), we have that \( \text{ord}_y(x) \) divides \( \varphi(y) \), thus \( A_2 \) is exactly the number of integers \( 1 \leq x \leq y \), with \( \gcd(x,y) = 1 \), such that there exists an integer \( d \geq \sqrt{\varphi(y)} \) with \( d | \varphi(y) \) and \( \text{ord}_y(x) = d \). Such integers satisfy \( \text{ord}_y(x) \geq \min_{\ell \geq \sqrt{\varphi(y)}} \varphi(\ell) \). But the latter function is at least \( \varphi(y)^{1/3} \) when \( y \) is large enough. Thus, the hypothesis of Theorem 2 is satisfied for large \( y \) by at least \( A_2 \) integers. Since \( A_2 \geq \varphi(y) - \varphi(y)^{2/3} \), the remark following Theorem 2 is proved. \( \square \)

**Proof of Theorem 3.** Let \( 0 < \phi < 1 \) and \( \alpha > 1 \). We denote by \( C_1, C_2 \) and \( C_3 \) effectively computable positive numbers depending only on \( \alpha \). Let \((x, y, m, n)\) be a solution of (1) such that \( \gcd (m - 1, n - 1) = d \geq 3 \) and \((m - 1)/(n - 1) \leq \alpha \). We write
\[
m - 1 = dr, \quad n - 1 = ds
\]
where \( r \) and \( s \) are positive integers. Now we infer from (1) that
\[
x \leq 2y^{s/r}, \quad y \leq 2x^{r/s}.
\]
By Lemma 4, we know that \( d \leq C_1 \). By Theorem 1, we may also suppose that \( y \geq C_2 \) with \( C_2 \) sufficiently large. Further, we rewrite (1) as
\[
x \frac{x^{dr}}{x-1} - y \frac{y^{ds}}{y-1} = \frac{1}{x-1} - \frac{1}{y-1},
\]
which implies that
\[
(26) \quad \left| \left( \frac{y(x-1)}{x(y-1)} \right)^{1/d} - \frac{x^r}{y^s} \right| < \frac{1}{y^{ds}}.
\]
Now we apply Lemma 5 with \( A = x(y - 1), B = y(x - 1), d = n, \sigma = s \) and \( K = \lfloor 2\alpha \rfloor + 1 \). We may suppose that \( K < n \). We choose \( \phi \), depending only on \( \alpha \), suitably such that
\[
1 + \frac{1}{\alpha} + sK \left( 1 + \frac{1}{1-\phi} + \frac{2-\phi}{K(1-\phi)} \right) < \left( 4\alpha + 6 + \frac{1}{\alpha} \right)s.
\]
Finally, setting $\delta = 1 + (2 - \phi)/K$ and $u_1 = 40d(K+1)(\delta+1-\phi)/(K\delta+1-\phi)$, we observe that
\[
A(A - B)^{-\delta}u_1^{-1} > C_3y^{1+\frac{1}{\alpha}}(y - x)^{-\delta} > C_3y^{1+\frac{1}{\alpha}-\delta} > 1
\]
if $C_2$ is sufficiently large. Hence, we conclude from Lemma 5 that the left hand side of (26) exceeds
\[
y^{-s(4\alpha+6+1/\alpha)},
\]
\[
\text{hence}
\]
\[
d < 4\alpha + 6 + \frac{1}{\alpha},
\]
as claimed. \hfill \Box

Proof of Theorem 4. Set $d = \gcd(m-1, n-1)$. We apply Lemma 4 with $\alpha = (m-1)/(n-1)$. We conclude that $d^2 \leq d(n-1) \leq 1114.5m$, whence $d \leq 33.4m^{1/2}$. \hfill \Box

Proof of Theorem 5. Let $(x, y, m, n)$ be a solution of (1) with $x < y \leq x^a$ and $y \equiv 1 \pmod{x}$. Regarding (1) modulo $x$, we deduce that $n \equiv 1 \pmod{x}$, thus $n > x$. We set
\[
\Lambda := (y - 1)x^n = (x - 1)y^n - (x - y)
\]
and we observe that $v_x(\Lambda) \geq m + 1$. Further, if $x \neq 2$, then according as $2\|x$ or not we have
\[
v_x(\Lambda) \leq v_{x/2} \left(y^n - \left(1 - \frac{y - 1}{x - 1}\right)\right)
\]
or
\[
v_x(\Lambda) = v_x \left(y^n - \left(1 - \frac{y - 1}{x - 1}\right)\right).
\]
We check that the hypotheses of Lemma 6 are satisfied, and, thanks to that Lemma, we obtain for $x > 2$ that
\[
m + 1 \leq \frac{66.8}{(\log \frac{y}{x})^4} \left(\max \left\{\log \frac{n+1}{\log y} + \log \log \frac{x}{2} + 0.64, 4\log \frac{x}{2}\right\}\right)^2 (\log y)^2.
\]
Further, $\log y \leq a \log x$ and $\log x/\log \frac{x}{2} \leq \log 6/\log 3 \leq 1.631$, hence we get
\[
m \leq \max\{2843a^2, 148(\log^2 m + 0.64)a^2\}.
\]
It follows that $m \leq 14000a^2(\log 3a)^2$, as claimed. \hfill \Box

Acknowledgements. This work was done when the first author was visiting the Tata Institute of Fundamental Research, Mumbai. He would like to thank the mathematical department for its hospitality and the CNRS for its financial support.
References


Received January 22, 2001.

Université Louis Pasteur
U. F. R. de mathématiques
7, rue René Descartes
67084 Strasbourg
France
E-mail address: bugeaud@math.u-strasbg.fr

School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road
Mumbai 400 005
India
E-mail address: shorey@math.tifr.res.in
TODA LATTICE AND TORIC VARIETIES FOR REAL SPLIT SEMISIMPLE LIE ALGEBRAS

Luis G. Casian and Yuji Kodama

The paper concerns the topology of an isospectral real smooth manifold for certain Jacobi element associated with real split semisimple Lie algebra. The manifold is identified as a compact, connected completion of the disconnected Cartan subgroup of the corresponding Lie group $\tilde{G}$ which is a disjoint union of the split Cartan subgroups associated to semisimple portions of Levi factors of all standard parabolic subgroups of $G$. The manifold is also related to the compactified level sets of a generalized Toda lattice equation defined on the semisimple Lie algebra, which is diffeomorphic to a toric variety in the flag manifold $\tilde{G}/B$ with Borel subgroup $B$ of $\tilde{G}$. We then give a cellular decomposition and the associated chain complex of the manifold by introducing colored-signed Dynkin diagrams which parametrize the cells in the decomposition.

1. Introduction.

In this paper, we study the topological structure of certain manifolds that are interesting in two different ways. First they are isospectral manifolds for a signed Toda lattice flow [14]; an integrable system that arises in several physical contexts and has been studied extensively. Secondly they are shown in §8 to be the closures of generic orbits of a split Cartan subgroup on a real flag manifold. These are certain smooth toric varieties that glue together the disconnected pieces of a Cartan subgroup of a semisimple Lie group of the form $\tilde{G}$. In this paper we start from the Toda lattice aspect of this object and end the paper inside a real flag manifold. We thus start by motivating and describing our main constructions from the point of view of the Toda lattice; then we trace a path that starts with this Toda lattice and that naturally leads to the (disconnected) Cartan subgroup of a split semisimple Lie group and a toric orbit. Another path that also leads to a Cartan subgroup starts with Kostant’s paper [16]; this approach is described in §8.

From the Toda lattice end of this story, these manifolds are related to the compactified level set of a generalized (nonperiodic) Toda lattice equation defined on the semisimple Lie algebra (see for example [16]) and, although they share some features with the Tomei manifolds in [21], they are different.
from those (e.g., nonorientable). As background information, we start with
a definition of the generalized Toda lattice equation which led us to our
present study of the manifolds.

Let \( g \) denote a real split semisimple Lie algebra of rank \( l \). We fix a
split Cartan subalgebra \( h \) with root system \( \Delta = \Delta(g, h) \), real root vectors
\( e_{\alpha_i} \) associated with simple roots \( \{\alpha_i : i = 1, \ldots, l\} = \Pi \). We also denote
\( \{h_{\alpha_i}, e_{\pm\alpha_i}\} \) the Cartan-Chevalley basis of \( g \) which satisfies the relations,
\[
[h_{\alpha_i}, h_{\alpha_j}] = 0, \quad [h_{\alpha_i}, e_{\pm\alpha_j}] = \pm C_{i,j} e_{\pm\alpha_j}, \quad [e_{\alpha_i}, e_{-\alpha_j}] = \delta_{i,j} h_{\alpha_j}
\]
where the \( l \times l \) matrix \((C_{i,j})\) is the Cartan matrix corresponding to
\( g \), and \( C_{i,j} = \alpha_i(h_{\alpha_j}) \).

Then the generalized Toda lattice equation related to real split semisim-
ple Lie algebra is defined by the following system of 2nd order differential
equations for the real variables \( f_i(t) \) for \( i = 1, \ldots, l \),
\[
\frac{d^2 f_i}{dt^2} = \epsilon_i \exp \left( - \sum_{j=1}^{l} C_{i,j} f_j \right)
\]
where \( \epsilon_i \in \{\pm 1\} \) which correspond to the signs in the indefinite Toda lattices
introduced in [5, 14]. The main feature of the indefinite Toda equation
having at least one of \( \epsilon_i \) being \(-1\) is that the solution blows up to infinity in
finite time [14, 10]. Having introduced the signs, the group corresponding to
the Toda lattice is a real split Lie group \( \tilde{G} \) with Lie algebra \( g \) which is defined
in \( \S 3 \). For example, in the case of \( g = \mathfrak{sl}(n, \mathbb{R}) \), if \( n \) is odd, \( \tilde{G} = SL(n, \mathbb{R}) \),
and if \( n \) is even, \( \tilde{G} = Ad(SL(n, \mathbb{R})^\pm) \).

The original Toda lattice equation in [20] describing a system of \( l \) particles
on a line interacting pairwise with exponential forces corresponds to the case
with \( g = \mathfrak{sl}(l + 1, \mathbb{R}) \) and \( \epsilon_i = 1 \) for all \( i \), and it is given by
\[
\frac{d^2 q_i}{dt^2} = \exp(q_{i-1} - q_i) - \exp(q_i - q_{i+1}), \quad i = 1, \ldots, l,
\]
where the physical variable \( q_i \), the position of the \( i \)-th particle, is given by
\[
q_i = f_i - f_{i+1}, \quad i = 1, \ldots, l,
\]
with \( f_{l+1} = 0 \) and \( f_0 = f_{l+2} = -\infty \) indicating \( q_0 = -\infty \) and \( q_{l+1} = \infty \).

The system (1) can be written in the so-called Lax equation which de-
scribes an iso-spectral deformation of a Jacobi element of the algebra \( g \). This
is formulated by defining the set of real functions \( \{(a_i(t), b_i(t)) : i = 1, \ldots, l\} \)
with
\[
a_i(t) := \frac{df_i(t)}{dt}, \quad b_i(t) := \epsilon_i \exp \left( - \sum_{j=1}^{l} C_{i,j} f_j(t) \right)
\]
from which the system (1) reads
\[
\frac{da_i}{dt} = b_i, \quad \frac{db_i}{dt} = -b_i \left( \sum_{j=1}^{l} C_{i,j} a_j \right).
\]
(3)

This is then equivalent to the Lax equation defined on \( g \) (see [7] for a nice review of the Toda equation).
\[
\frac{dX(t)}{dt} = [P(t), X(t)],
\]
(4)

where the Lax pair \((X(t), P(t))\) in \( g \) is defined by
\[
\begin{cases}
X(t) &= \sum_{i=1}^{l} a_i(t) h_{\alpha_i} + \sum_{i=1}^{l} (b_i(t) e_{-\alpha_i} + e_{\alpha_i}) \\
P(t) &= -\sum_{i=1}^{l} b_i(t) e_{-\alpha_i}.
\end{cases}
\]
(5)

Although a case with some \( b_i = 0 \) is not defined in (2) the corresponding system (3) is well-defined and is reduced to several noninteracting subsystems separated by \( b_i = 0 \). The constant solution \( a_i \) for \( b_i = 0 \) corresponds to an eigenvalue of the Jacobi (tridiagonal) matrix for \( X(t) \) in the adjoint representation of \( g \). We denote by \( S(F) \) the ad diagonalizable elements in \( g \otimes \mathbb{R} F \) with eigenvalues in \( F \), where \( F \) is \( \mathbb{R} \) or \( \mathbb{C} \).

Then the purpose of this paper is to study the disconnected manifold \( Z_{\mathbb{R}} \) of the set of Jacobi elements in \( g \) associated to the generalized Toda lattices,
\[
Z_{\mathbb{R}} = \left\{ X = x + \sum_{i=1}^{l} (b_i e_{-\alpha_i} + e_{\alpha_i}) : x \in \mathfrak{h}, b_i \in \mathbb{R} \setminus \{0\}, X \in S(\mathbb{R}) \right\},
\]
its iso-spectral leaves \( Z(\gamma)_{\mathbb{R}}, \gamma \in \mathbb{R}^{l} \) and the construction of a smooth connected compactification, \( \hat{Z}(\gamma)_{\mathbb{R}} \) of each \( Z(\gamma)_{\mathbb{R}} \). The construction of \( \hat{Z}(\gamma)_{\mathbb{R}} \) generalizes the construction of such a smooth compact manifold which was carried out in [15] in the important case of \( g = \mathfrak{sl}(l+1, \mathbb{R}) \). The construction there is based on the explicit solution structure in terms of the so-called \( \tau \)-functions, which provide a local coordinate system for the blow-up points. Then by tracing the solution orbit of the indefinite Toda equation, the disconnected components in \( Z(\gamma)_{\mathbb{R}} \) are all glued together to make a smooth compact manifold. The result is maybe well explained in Figure 1 for the case of \( \mathfrak{sl}(3, \mathbb{R}) \). In the figure, the Toda orbits are shown as the dotted lines, and each region labeled by the same signs in \( (\epsilon_1, \epsilon_2) \) with \( \epsilon_i \in \{\pm\} \) are glued together through the boundary (the wavy-lines) of the hexagon. At a point of the boundary the Toda orbit blows up in finite time, but the orbit can be uniquely traced to the one in the next region (marked by the same letter
The compact smooth manifold $\tilde{Z}(\gamma)_\mathbb{R}$ in this case is shown to be isomorphic to the connected sum of two Klein bottles. In the case of $\mathfrak{sl}(l+1, \mathbb{R})$ for $l \geq 2$, $\tilde{Z}(\gamma)_\mathbb{R}$ is shown to be nonorientable and the symmetry group is the semi-direct product of $(\mathbb{Z}_2)^l$ and the Weyl group $W = S_{l+1}$, the permutation group. One should also compare this with the result in [21] where the compact manifolds are associated with the definite (original) Toda lattice equation and the compactification is done by adding only the subsystems. (Also see [4] for some topological aspects of the manifolds.)

The study of $Z(\gamma)_\mathbb{R}$ and of the compact manifolds $\tilde{Z}(\gamma)_\mathbb{R}$, can be physically motivated by the appearance of the indefinite Toda lattices in the context of symmetry reduction of the Wess-Zumino-Novikov-Witten (WZNW) model. For example, the reduced system is shown in [6] to contain the indefinite Toda lattices. The compactification $\tilde{Z}(\gamma)_\mathbb{R}$ can then be viewed as a concrete description of (an expected) regularization of the integral manifolds of these indefinite Toda lattices, where infinities (i.e., blow up points) of the solutions of these Toda systems glue everything into a smooth compact manifold.

In addition our work is mathematically motivated by:

a) The work of Kostant in [16] where he considered the real case with all $b_i > 0$,
b) the construction of the Toda lattice in [17].

In [17], the solution \( \{ b_i(t) : i = 1, \ldots, l \} \) in (2) of the generalized Toda lattice equation is shown to be expressed as an orbit on a connected component \( H_\epsilon \) labeled by \( \epsilon = (\epsilon_1, \ldots, \epsilon_l) \) with \( \epsilon_i = \pm 1 \), of the Cartan subgroup \( H_\mathbb{R} \) defined in §3,

\[
H_\mathbb{R} = \bigcup_{\epsilon \in \{\pm 1\}^l} H_\epsilon.
\]

This can be seen as follows: Let \( g_\epsilon \) be an element of \( H_\epsilon \) given by

\[
g_\epsilon = h_\epsilon \exp \left( \sum_{i=1}^{l} f_i h_{\alpha_i} \right),
\]

which gives a map from \( Z_\mathbb{R} \) into \( H_\mathbb{R} \). Here the element \( h_\epsilon \in H_\epsilon \) satisfies \( \chi_{\alpha_i}(h_\epsilon) = \epsilon_i \) for the group character \( \chi_\phi \) determined by \( \phi \in \Delta \). The solution \( \{ b_i(t) : i = 1, \ldots, l \} \) in (2) is then directly connected to the group character \( \chi_{-\alpha_i} \) evaluated at \( g_\epsilon \), i.e.,

\[
b_i(t) = \chi_{-\alpha_i}(g_\epsilon).
\]

The Toda lattice equation is now written as an evolution of \( g_\epsilon \),

\[
\frac{d}{dt} g_\epsilon^{-1} \frac{d}{dt} g_\epsilon = [g_\epsilon^{-1} e_+, e_-],
\]

where \( e_\pm \) are fixed elements in the simple root spaces \( \mathfrak{g}_\pm \Pi \) so that all the elements in \( \mathfrak{g}_\pm \Pi \) can be generated by \( e_\pm \), i.e., \( \mathfrak{g}_\pm \Pi = \{ \text{Ad}_h(e_\pm) : h \in H \} \). In particular, we take

\[
e_\pm = \sum_{i=1}^{l} e_{\pm \alpha_i}.
\]

Thus the Cartan subgroup \( H_\mathbb{R} \) can be identified as the position space (e.g., \( f_i = q_i + \cdots + q_l \) for \( \mathfrak{sl}(l+1, \mathbb{R}) \)-Toda lattice) of the generalized Toda lattice equation, whose phase space is given by the tangent bundle of \( H_\mathbb{R} \).

One should also note that the boundary of each connected component \( H_\epsilon \) is given by either \( b_i = 0 \) (corresponding to a subsystem) or \( |b_i| = \infty \) (to a blow-up). We are then led to our main construction of the compact smooth manifold of \( H_\mathbb{R} \) in attempting to generalize [16] Theorem 2.4 to our indefinite Toda case including \( b_i < 0 \).

We define a set \( \hat{H}_\mathbb{R} \) containing the Cartan subgroup \( H_\mathbb{R} \) by adding pieces corresponding to the blow-up points and the subsystems (Definition 6.1).

Thus the set \( \hat{H}_\mathbb{R} \) is defined as a disjoint union of split Cartan subgroups \( H^A_\mathbb{R} \) associated to semisimple portions of Levi factors of all standard parabolic subgroups determined by \( A \subset \Pi \), (Definition 3.10),

\[
\hat{H}_\mathbb{R} = \bigcup_{A \subset \Pi} \bigcup_{w \in W/W_{0\Pi \setminus A}} w (H^A_\mathbb{R}) \times \{[w]^A\}.
\]
The space $\hat{H}_R$ then constitutes a kind of compact, connected completion of the disconnected Cartan subgroup $H_R$. Figure 1 also describes how to connect the connected components of $H_R$ to produce the connected manifold $\hat{H}_R$ in the case of $\mathfrak{sl}(3, \mathbb{R})$. The signs must now be interpreted as the signs of simple root characters on the various connected components. The pair of signs $(+, +)$ corresponds to the connected component of the identity $H$ and regions with a particular sign represent one single connected component of $H_R$. Boundaries between regions with a fixed sign correspond to connected components of Cartan subgroups arising from Levi factors of parabolic subgroups. In addition the Weyl group $W$ acts on $\hat{H}_R$.

Most of this paper is then devoted to describing the detailed structure of the manifold $\hat{H}_R$, and in §8 we conclude with a diffeomorphism defined between $\hat{H}_R$ and the isospectral manifold $\hat{Z}(\gamma)_R$ as identified with a toric variety $(\mathbb{H}_{\mathbb{R}} g B)$ in the flag manifold $\tilde{G}/B$.

1.1. Main Theorems. In connection with the construction $\hat{H}_R$, we introduce in Definition 4.1 the set of colored Dynkin diagrams. The colored Dynkin diagrams are simply Dynkin diagrams $D$ where some of the vertices have been colored red or blue. For example in the case $\mathfrak{g} = \mathfrak{sl}(3, \mathbb{R})$: $\circ_R - \circ$ (the sub-index $R$ indicates that $\circ$ is colored red). The full set of colored Dynkin diagrams consists of pairs: $(D, [w]_{\Pi \setminus S})$ where $D$ is a colored Dynkin diagram, $S \subset \Pi$ denotes the set of vertices that are colored in $D$ and $[w]_{\Pi \setminus S}$ is the coset of $w$ in $W/W_S$. To each pair $(D, [w]_{\Pi \setminus S})$ one can associate a set which is actually a cell. First Notation 6.2 associates a subset of $\hat{H}_R$ also denoted $(D, [w]_{\Pi \setminus S})$.

This turns out to be a cell of codimension $k$ with $k = |S|$. We illustrate the example of $\mathfrak{sl}(2, \mathbb{R})$ in Figure 2.

We consider the set $\mathbb{D}^k$ of colored Dynkin diagrams $(D, [w]_{\Pi \setminus S})$ with $|S| = k$. In Figure 2, we have $\mathbb{D}^0 = \{(\circ, e), (\circ, s_{\alpha_1})\}$, $\mathbb{D}^1 = \{(\circ_B, [e]), (\circ_R, [e])\}$. We then obtain the following theorem:
Theorem 1.1. The collection of the sets \( \{D^k : k = 0, 1, \ldots, l\} \) gives a cell decomposition of \( \hat{H}_\mathbb{R} \).

The chain complex \( \mathcal{M}_\ast \) is introduced in §4. The construction, in the case of \( \mathfrak{sl}(2; \mathbb{R}) \), is as follows:

\[
\mathcal{M}_1 = \mathbb{Z} \left[ D_0 \right] \xrightarrow{\partial_1} \mathcal{M}_0 = \mathbb{Z} \left[ D_1 \right].
\]

Here the boundary map \( \partial_1 \) is given by

\[
\partial_1(\circ, e) = (\circ_B, [e]) - (\circ_R, [e]),
\]

\[
\partial_1(\circ, s_{\alpha_1}) = (\circ_B, [s_{\alpha_1}]) - (\circ_R, [s_{\alpha_1}]).
\]

In particular, \( (\circ) := \sum_{w \in W} (-1)^{\ell(w)} (\circ, w) = (\circ, e) - (\circ, s_{\alpha_1}) \) is a cycle, where \( \ell(w) \) is the length of \( w \), and \( (\circ) \) represents the \( \hat{H}_\mathbb{R} \). The following theorem gives a topological description of \( \hat{H}_\mathbb{R} \):

Theorem 1.2. The manifold \( \hat{H}_\mathbb{R} \) is compact, nonorientable (except if \( g \) is of type \( A_1 \)), and it has an action of the Weyl group \( W \). The integral homology of \( \hat{H}_\mathbb{R} \) can be computed as a \( \mathbb{Z}[W] \)-module as the homology of the chain complex \( \mathcal{M}_\ast \) in (13).

Theorem 1.2 is completed in Proposition 7.7. The \( W \)-action is (abstractly) introduced in Definition 4.4 in terms of a representation-theoretic induction process from smaller parabolic subgroups of \( W \). The proof that the \( W \)-action is well-defined is given in Proposition 4.5 and Proposition 5.3. The chain complex \( \mathcal{M}_\ast^{CW} \) in Definition 7.6 is defined so that it computes integral homology. Since each \( X_r \backslash X_{r-1} \) in Definition 7.6 is a union of cells \( (D, [w]^{P \backslash S}) \in \mathbb{D}^{l-r} \), we obtain an identification between \( \mathcal{M}_\ast^{CW} \) and the chain complex \( \mathcal{M}_\ast \).

Then using the Kostant map which can be described as a map from \( \hat{H}_\mathbb{R} \) into the flag manifold \( \tilde{G}/B \) (in Definition 8.8), that is, a torus imbedding, we obtain the following theorem:

Theorem 1.3. The toric variety \( \hat{Z}(\gamma)_\mathbb{R} \) is a smooth compact manifold which is diffeomorphic to \( \hat{H}_\mathbb{R} \).

The complex version of this theorem has been proven in [9], and our proof is essentially given in the same manner.

1.2. Outline of the paper. The paper is organized as follows:

In §2, we begin with two fundamental examples, \( \mathfrak{g} = \mathfrak{sl}(l + 1, \mathbb{R}) \) for \( l = 1, 2 \), which summarize the main results in the paper.

In §3, we present the basic notations necessary for our discussions. We then define a real group \( \tilde{G} \) of rank \( l \) whose split Cartan subgroup \( H_\mathbb{R} \) contains \( 2^l \) connected components. We also define the Lie subgroups of \( \tilde{G} \) corresponding to the subsystems and blow-ups of the generalized Toda lattice.
equations. The reason for the introduction of $\tilde{G}$ and the Cartan subgroup $H_\mathbb{R}$ can be appreciated in Remark 8.10 and Corollary 8.11.

In §4, we introduce colored Dynkin diagrams which will be shown to parametrize the cells in a cellular decomposition of the manifold $\hat{H}_\mathbb{R}$. We then construct a chain complex $M^\ast$ of the $\mathbb{Z}[W]$-modules $M_{l-k}$ (Definition 4.9). The parameters involved in the statement of Theorem 1.1 are given here.

In §5 we show that Weyl group representations introduced in §4 are well-defined. In addition we define $H_\circ$ in §5 by adding some Cartan subgroups associated to semisimple Levi factors of parabolic subgroups to $H_\mathbb{R}$.

In §6, we define $\hat{H}_\mathbb{R}$ as a union of the Cartan subgroup associated to semisimple Levi factors of certain parabolic subgroups (Definition 6.1) using translation by Weyl group elements. We also associate subsets of $\hat{H}_\mathbb{R}$ to colored Dynkin diagrams.

In §7, we discuss the topological structure of $\hat{H}_\mathbb{R}$ expressing $\hat{H}_\mathbb{R}$ as the union of the subsets determined by the colored Dynkin diagrams. We then show that $\hat{H}_\mathbb{R}$ is a smooth compact manifold and those subsets naturally determine a cell decomposition (Theorem 1.2).

In §8, we consider a Kostant map between the isotropy subgroup $G_z\subset G_\mathbb{C}$ of $G_\mathbb{C}$ with $\text{Ad}(g)z = z$ and the isospectral manifold $J(\gamma)_\mathbb{C}$ for some $\gamma \in \mathbb{R}^l$, which can be also described as a map into the flag manifold $\tilde{G}/B$. Then we show that the toric variety $(\hat{H}_\mathbb{R} n B)$ is a smooth manifold and obtain Theorem 1.3.

2. Examples of $\mathfrak{sl}(l+1, \mathbb{R})$, $l = 1, 2$.

This section contains two examples which are the source of insight for the main theorems in this paper. Most of the notation and constructions used later on in the paper can be anticipated through these examples.

Our main object of study in this paper, $\hat{H}_\mathbb{R}$, has nothing to do, in its construction, with the moment map and the Convexity Theorem of [1]. However, because of [1] it is expected to contain a polytope with vertices given by the action of the Weyl group which, gives rise to it through some gluings along its faces. We will identify a convenient polytope of this kind in $\mathfrak{h}'$ (dual of Cartan subalgebra) and show how all the pieces of $\hat{H}_\mathbb{R}$ would fit inside it. The polytope in the example of $\mathfrak{sl}(3, \mathbb{R})$ is a hexagon. Since we are not dealing with the moment map in this paper, this is just done for the purposes of motivation and illustration. The polytope of the Convexity Theorem in [1], strictly speaking, is the convex hull of the orbit of $\rho$. This sits inside the convex hull of the orbit of $2\rho$ which is all a part of $\hat{H}_\mathbb{R}$.

The terms dominant and antidominant being relative to a choice of a Borel subalgebra, we refer to the chamber in $\mathfrak{h}'$ containing $\rho$ as antidominant for no other reason than the fact that we will make it correspond to what we call the antidominant chamber of the Cartan subgroup $H_\mathbb{R}$. This odd
1) The bulk — interior — of $\hat{H}_R$ is made up of a split Cartan subgroup $H_R$ which has $2^l$ connected components $H_{\epsilon}$ parametrized by signs $\epsilon = (\pm, \ldots, \pm)$ for $g$ with rank $l$, i.e., $H_R = \bigcup_{\epsilon \in \{\pm\}^l} H_{\epsilon}$. These disconnected pieces are glued together into a connected manifold by using Cartan subgroups $H_R^A$ associated to Levi factors of parabolic subgroups, which are determined by the set of simple roots $\Pi \setminus A$ for each $A \subset \Pi$.

2) The language of colored Dynkin diagrams and signed colored Dynkin diagrams is introduced in this paper in order to parametrize the pieces of $\hat{H}_R$. However the motivation for their introduction is that these diagrams parametrize the pieces of a convex polytope (hexagon) which lives in $h'$, including its external faces and internal walls. The parameters for the pieces are colored and signed-colored Dynkin diagrams.

3) If one just looks at the antidominant chamber intersected with the $2\rho$ polytope, then it is easy to see that it forms a box. This box is the closure of the antidominant chamber $H_R^{<}$ of the Cartan subgroup $H_R$ (Definition 3.10) inside $\hat{H}_R$, i.e., $\hat{H}_R = \bigcup_{w \in W} w(\overline{H_R^{<}})$. The box $\overline{H_R^{<}}$ has internal chamber walls corresponding to simple roots (thus a Dynkin diagram appears) and it has external faces which are also parametrized by simple roots. Hence one needs not just a Dynkin diagram but also two colors to indicate internal walls and external faces. The color blue indicates internal chamber walls and the color red indicates external faces. The $S$ is usually reserved for the set of colored vertices of a colored Dynkin diagram.

4) To have a correspondence with the $2^l$ connected components of the Cartan subgroup $H_R$ of diagonal matrices, the box $\overline{H_R^{<}}$ must be further subdivided into $2^l$ boxes $H_{\epsilon}^{<}$ with signs $\epsilon$, i.e., $H_R^{<} = \bigcup_{\epsilon \in \{\pm\}^l} H_{\epsilon}^{<}$. These boxes are parametrized by a Dynkin diagram where each simple root has a sign attached to it. The boundaries of these boxes are portions of the internal and external walls. Hence we use Dynkin diagrams with both signs and colors. The boundaries between two signs require to be labeled as 0. The $A$ is usually reserved for the set of vertices of a signed-colored Dynkin diagram assigned 0 (subsystems).

5) To translate the notation of colored and signed-colored Dynkin diagrams to other chamber one needs to consider pairs $(D, w)$ where $w$ is a Weyl group element. However since the Weyl group action has non-trivial isotropy groups in portions of the polytope, it is necessary to
consider Weyl group cosets \([w]^{\Pi \setminus S}\) for \(S \subset \Pi\) the set of colored simple roots (giving reflections generating an isotropy group).

6) To translate the Levi subgroups around, note that the Weyl group of the Levi factor stabilizes the Cartan subgroup of the Levi factor corresponding to the simple roots in \(\Pi \setminus A\). Because of that we consider products \(w(H^A_R) \times \{[w]^A\}\) so that \([w]^A\) is a coset in \(W/\Pi \setminus A\).

\[
\hat{H}_R = \bigcup_{A \subset \Pi} \bigcup_{w \in W/\Pi \setminus A} w(H^A_R) \times \{[w]^A\}.
\]

2.1. The example of \(\mathfrak{sl}(2, \mathbb{R})\). The corresponding group in this example is given by \(G = \text{Ad}(SL(2, \mathbb{R})^\pm)\). The geometric picture that corresponds to \(\hat{H}_R\) is a circle. Consider the interval \([-2, 2]\) where \(-2\) and \(2\) are identified. Here \(2\) represents \(2\pi\). Inside this interval \((-2, 2)\) we consider the subset \([-1, 1]\). The points \(-1, 0, 1\) divide \([-2, 2]\) into four open intervals. These open intervals will correspond to the connected components of a Cartan subgroup \(H_R\) of \(\text{Ad}(SL(2, \mathbb{R})^\pm)\) when the walls corresponding to the points 0 and 2 are deleted. Below we list each cell \(w(H^A_{\leq \xi}) \times \{[w]^A\}\) in \(\hat{H}_R\), where \(H^A_{\leq \xi}\) is the intersection of \(H^A_{\xi}\) with the strictly antidominant chamber (the superscript \(\leq\) means that the walls are included):

Let us take \(h_{\alpha_1} = \text{diag}(1, -1) \in \mathfrak{h}\) and \(h_e = \text{Ad}(\text{diag}(e, 1)) \in H_e\). Then any element in \(\hat{H}_R\) can be expressed as \(h_t \exp(th_{\alpha_1})\) with some parameter \(t \in \mathbb{R}\). We denote \(\exp(th_{\alpha_1}) = \text{diag}(a, a^{-1})\).

We first consider \(A = \emptyset\). Then the cell \(H^\leq_{\mathfrak{so}} \times \{[e]\}\) is given by

\[
\{ \text{Ad}(\text{diag}(a, a^{-1})) : 0 < a < 1 \} \times \{[e]\} = (\mathfrak{o}_{+}, e) \leftrightarrow (0, 1).
\]

Here \(\chi_{\alpha_1}(h) = a^2\) for \(h \in H_{+}\). Also we have the set \(H^\leq_{\mathfrak{so}} \times \{[e]\}\) as

\[
\{ \text{Ad}(\text{diag}(-a, a^{-1})) : 0 < a < 1 \} \times \{[e]\} = (\mathfrak{o}_{-}, e) \leftrightarrow (1, 2)
\]

with \(\chi_{\alpha_1}(h) = -a^2\) for \(h \in H_{-}\).

We now consider the case of \(A = \{\alpha_1\} = \Pi\), which corresponds to a subsystem of Toda lattice with \(b_1 = 0\) in (5). Then we have \(H^{\Pi \setminus \leq}_{\mathfrak{so}} = \{e\}\). This is the degenerate case of \(A = \Pi\) which gives rise to the Levi factor of a Borel subgroup. Since the Levi factor does not contain a semisimple Lie subgroup, \(H^A_{\mathfrak{so}}\) is defined to be \(\{e\}\) (Definition 3.9). Here \([w]^A\) is just the element \(w\). We have

\[
\{ \text{Ad}(h_{+}) \} \times \{e\} = (\mathfrak{o}_0, e) \leftrightarrow \{1\}.
\]

We now describe the box containing the strictly antidominant chamber \(H^\leq_{\mathfrak{so}}\) of \(H_R\). Since \(H^\leq_{\mathfrak{so}}\) is disconnected, \((\mathfrak{o}_0, e)\) has been used to glue the pieces together. We then have a box given by

\[
(\mathfrak{o}, e) = (\mathfrak{o}_{+}, e) \cup (\mathfrak{o}_{-}, e) \cup (\mathfrak{o}_0, e) \leftrightarrow (0, 2).
\]
The bijection which gives rise to local coordinates $\phi_e$ in Subsection 7.2 is given by either $\pm a^2$ or 0. The set $(\circ, e)$ is sent by $\phi_e$ to the interval $(-1, 1)$.

We now apply $s_{\alpha_1}$ on $H^< \times \{[e]\}$ to obtain the cell in the $s_{\alpha_1}$-chamber which corresponds to the negative intervals:

$$\{ \text{Ad}(\text{diag}(a^{-1}, a)) : 0 < a < 1 \} \times \{[e]\} = (\circ_+, s_{\alpha_1}) \leftrightarrow (-1, 0)$$

with $\chi_{\alpha_1}(h) = a^{-2}$. However the local coordinate $\phi_{s_{\alpha_1}}$ is $\chi_{-\alpha_1}$ which equals $a^2$. We also have

$$\{ \text{Ad}(\text{diag}(a^{-1}, -a)) : 0 < a < 1 \} \times \{[e]\} = (\circ_-, s_{\alpha_1}) \leftrightarrow (-2, -1).$$

Also for the case $A = \{\alpha_1\}$, and the set $s_{\alpha_1}(H^<_{\mathbb{R}}) \times \{s_{\alpha_1}\}$ is given by

$$\{ \text{Ad}(h_+) \} \times \{s_{\alpha_1}\} = (\circ_0, s_{\alpha_1}) \leftrightarrow \{-1\}.$$ 

Once again $\phi_{s_{\alpha_1}}$ is given as 0, and we have the set $s_{\alpha_1}(H^<_{\mathbb{R}})$, giving rise to an open box by gluing two disconnected pieces as before,

$$(\circ, s_{\alpha_1}) = (\circ_+, s_{\alpha_1}) \cup (\circ_-, s_{\alpha_1}) \cup (\circ_0, s_{\alpha_1}) \leftrightarrow (-2, 0).$$

These are already associated to colored Dynkin diagrams $\circ_B$ and $\circ_R$.

We can write down $\{-2\}$ by applying $s_{\alpha_1}$ as we did above. However noting $\text{Ad}(\text{diag}(-1, 1)) = \text{Ad}(\text{diag}(1, -1))$, we obtain the same set that defines $\{2\}$. Thus $\{2\}$ and $\{-2\}$ are identified. We then obtain the interval $[-2, 2]$ with $-2$ identified with 2, which is $\hat{H}_{\mathbb{R}}$ diffeomorphic to a circle. We illustrate this example in Figure 3.

**Figure 3.** The manifold $\hat{H}_{\mathbb{R}}$ parametrized by signed-colored Dynkin diagrams for $\mathfrak{sl}(2, \mathbb{R})$. The endpoints in the interval are identified giving rise to a circle.
Figure 4. The square formed by the intersection of the antidominant chamber with the $2\rho$ hexagon for $\mathfrak{sl}(3, \mathbb{R})$.

Maps can be easily found between the intervals on the left and the sets on the right above. With a suitable topology associated to the $(D, [w])$, topology defined in terms of the coordinate functions $\phi_e$ and $\phi_{s_{\alpha_1}}$ in Section 7.3, each interval or point on the left side is homeomorphic to the interval on the right. This is what is indicated with $\leftrightarrow$.

2.2. The example of $\mathfrak{sl}(3, \mathbb{R})$. We here consider the dual of its Cartan subalgebra $\mathfrak{h}'$ and, inside it, a convex region bounded by a hexagon which is determined by the $W$ orbit of $2\rho$. We will later describe how to identify some of the faces of this hexagon. In Figure 4, we illustrate the parametrization of the faces. The two walls of the antidominant chamber intersected with this convex region are denoted by the colored Dynkin diagrams $\circ_B - \circ$ (the $s_{\alpha_1}$-wall) and $\circ - \circ_B$ (the $s_{\alpha_2}$-wall). The intersection of two of the sides or faces of the $2\rho$ hexagon with the antidominant chamber are each denoted by a colored Dynkin diagram $\circ_R - \circ$ or $\circ - \circ_R$. The four colored Dynkin diagrams $\circ_B - \circ, \circ - \circ_B, \circ_R - \circ, \circ - \circ_R$ form a square. This square is the intersection of the antidominant chamber and the $2\rho$ hexagon.

We now further subdivide this square into four subsquares (see Figure 1 and Figure 4). Inside the $2\rho$ hexagon is the orbit of $\rho$ which gives rise to a new smaller hexagon. Consider the two faces in the smaller hexagon intersecting the antidominant chamber. Denote these (intersected with the antidominant chamber) by signed-colored diagrams. The $\circ_0 - \circ_+$ corresponds to the unique face intersecting the wall $\circ - \circ_B$, and $\circ_+ - \circ_0$ corresponds to the unique face intersecting the wall $\circ_B - \circ$. Denote by $\circ_0 - \circ_0$ the vertex of the $\rho$ hexagon which is the intersection of these two faces. We now add a
segment joining the vertex \( o_0 - o_0 \), to a point in the interior of \( o_R - o \), say the midpoint. Denote this segment by the signed-colored diagram \( o_- - o_0 \).

We add another segment joining \( o_0 - o_0 \) to a point in the interior of \( o - o_R \) and denote this second segment by the signed-colored diagram \( o_0 - o_- \). Now the square is divided into four “square” regions denoted by \( o_\pm - o_\pm \).

Note that both \( o_R - o \) and \( o - o_R \) segments parametrized by colored Dynkin diagrams, are now subdivided into two segments parametrized with signed-colored Dynkin diagrams. For instance \( o_- - o_- \) Dynkin diagrams, are now subdivided into two segments parametrized with signed-colored diagrams \( o_- o_- \). We add another segment joining \( o_- o_- \) segment joining the vertex \( o_- o_- \), for example, has a boundary which consists of the segments parametrized by \( o_- o_- \), \( o_- o_- \), \( o_- o_- \) and \( o_- o_- \). The square \( o_- o_- \) has boundary \( o_- o_- \), \( o_- o_- \), \( o_- o_- \), \( o_- o_- \). Here the \( \alpha_2 \)-wall \( o_- o_- \) has also been subdivided into two pieces \( o_- o_- \) and \( o_- o_- \) by the intersection with the \( \rho \) hexagon.

If we now consider the full set of signed-colored Dynkin diagrams by translating with \( W \), we can fill the interior of the \( 2 \rho \) hexagon with a total of 12 regions. The four squares \( (o_- o_- , e) \) form the intersection of the antidominant chamber with the inside of the \( 2 \rho \) hexagon.

2.2.1. The sets in \( \hat{H}_R \) parametrized by the colored and signed-colored Dynkin diagrams. We now proceed to describe explicitly some of the pieces of \( \hat{H}_R \) corresponding to the signed-colored Dynkin diagrams \( (D, [w]^{\Pi \setminus S}) \) with \( +, - \) or 0 on the vertices in \( \Pi \setminus S \) of the diagram \( D \) and \( [w]^{\Pi \setminus S} \in W/W_S \).

When \( A = \emptyset \) implying no 0’s in the vertices, we have \( H_R^A = H_R \) which has 4 connected components,

\[
H_R = \bigcup_{\varepsilon \in \{\pm 1\}^2} \{ h_\varepsilon \text{diag}(a, b, c) : a > 0, b > 0, abc = 1 \},
\]

where \( h_\varepsilon = h_{(\varepsilon_1, \varepsilon_2)} = \text{diag}(\varepsilon_2, \varepsilon_1 \varepsilon_2, \varepsilon_1) \) satisfying \( \chi_{\alpha_i}(h_\varepsilon) = \varepsilon_i \). Since \( A = \emptyset \), \( W_{\Pi \setminus A} = W \), and there is only one coset \([\varepsilon]^A = [\varepsilon]\) in the \( \hat{H}_R \) construction.

We now consider the signed-colored Dynkin diagrams setting \( S = \emptyset \), that is, all the vertices are uncolored and they give a subdivision of the antidominant chamber inside the \( 2 \rho \) hexagon. In order to move around this chamber and its subdivisions using the Weyl group we must also consider six elements \([w]^{\Pi} = w \in W \). We then have, for \( S = \emptyset \) and \( A = \emptyset \), and \((\varepsilon_1, \varepsilon_2) = (\pm 1, \pm 1)\),

\[
(\varepsilon_1 o_- \varepsilon_2, e) = \{ h_\varepsilon \text{diag}(a, b, (ab)^{-1}) : ab^{-1} < 1, \ ab^2 < 1 \} \times \{ [\varepsilon] \}
\]

Note here that the inequalities \(|\chi_{\alpha_i}| < 1\) guarantee that the set in question is contained in the chamber associated to \( e \). Here the local coordinate \( \phi_\varepsilon \) is given by \((\chi_{\alpha_1}, \chi_{\alpha_2})\) which equals \((ab^{-1}, ab^2)\).

For \( A = \{\alpha_1\} \), we have \( h_\varepsilon = \text{diag}(\varepsilon_2, \varepsilon_2, 1) \in H_R^A \) with \( \varepsilon = (1, \varepsilon_2) \); and we multiply this element with \( \text{diag}(1, a, a^{-1}) \). This is a typical element in
the connected component of the Cartan associated to the Levi factor, in accordance with the definition of $H^A_R$ in Definition 3.10. This gives:

$$\phi_0 - \phi_2, e = \{ \text{diag}(e_2, e_2a, a^{-1}) : 0 < a < 1 \} \times \{ [e]^A \}$$

where the local coordinate $\phi_e$ is given by $(0, e_2a^2)$.

For $A = \{ \alpha_2 \}$, we have in a similar way with $h_{(\epsilon_1,1)} = \text{diag}(1, \epsilon_1, \epsilon_1)$:

$$\phi_{\epsilon_1} - \phi_0, e = \{ \text{diag}(e_1a^{-1}, \epsilon_1) : 0 < a < 1 \} \times \{ [e]^A \}$$

where $\phi_e$ now equals $(\epsilon_1a^2, 0)$.

We have an open square associated with the interior of the antidominant chamber,

$$(\circ - \circ, e) = \bigcup_{(\nu_1, \nu_2)\in\{\pm 1, 0\}^2} (\circ_{\nu_1} - \circ_{\nu_2}, e).$$

The image of this set under the map $\phi_e$ is an open square $(-1, 1) \times (-1, 1)$.

We now write down the boundary of this square: We here give an explicit form of $(\circ_R - \circ, [e]^H, [S])$, and the others can be obtained in the similar way. We first have, for $S = \{ \alpha_1 \}, A = \emptyset$, so that $[e]^A = [e] = [s_{\alpha_1}]$ for $i = 1, 2$,

$$[\circ_R - \circ_{e_2}, [e]^{\{\alpha_2\}}] = \{ \text{diag}(e_2a, -e_2a, -a^{-2}) : 0 < a < 1 \} \times \{ [e] \}.$$ 

Here $\phi_e$ equals $(\chi_{\alpha_1}, \chi_{\alpha_2})$, and is given by $(-1, \epsilon_2a^3)$.

With $A = \{ \alpha_2 \}$, we have:

$$[\circ_R - \circ_{e_2}, [e]^{\{\alpha_2\}}] = \{ \text{diag}(1, -1, -1) \} \times \{ [e]^{\{\alpha_2\}} \}.$$ 

The map $\phi_e$ is $(\chi_{\alpha_1}^\Delta, 0)$ and equals $(-1, 0)$. We then have

$$[\circ_R - \circ, [e]^{\{\alpha_2\}}] = \bigcup_{\nu\in\{\pm 1, 0\}} ([\circ_R - \circ_{\nu}, [e]^{\{\alpha_2\}}]).$$

The image of this set under $\phi_e$ is thus $\{-1\} \times (-1, 1)$.

We now consider the parts of the Cartan subgroup of Levi factors corresponding to other chambers inside the hexagon. Note that if we apply $s_{\alpha_1} = w$ to $(\circ_{\epsilon_1} - \circ_{e_2}, e)$, we obtain

$$(\circ_{\epsilon_1} - \circ_{e_1\epsilon_2}, s_{\alpha_1}) = \{ \text{diag}(\epsilon_1e_2b, e_2a, \epsilon_1(ab)^{-1}) : ab^{-1} < 1, ab^2 < 1 \} \times \{ [e] \}.$$ 

Since $\text{sign}(\chi_{\alpha_1}) = \epsilon_1$ and $\text{sign}(\chi_{\alpha_2}) = \epsilon_1\epsilon_2$, this set is no longer contained in $H_{(\epsilon_1, \epsilon_2)}$ but rather in $H_{(\epsilon_1, \epsilon_1\epsilon_2)}$. This justifies the notation $(\circ_{\epsilon_1} - \circ_{e_1\epsilon_2})$. One should note that $e$ for the component $H_e$ in $H_R$ did not change when one uses the simple roots associated to the new positive system $s_{\alpha_1}\Delta_+$. 

Also notice that for \( S = \{ \alpha_1 \} \) we have
\[
\left( \circ_R - \circ_+, [s_{\alpha_1}]^{\Pi \setminus S} \right) = \{ \text{diag}(-a, a, a^{-2}) : a > 0 \} \times \{ [e] \}
\]
and we have an identification of \( \left( \circ_R - \circ_+, [s_{\alpha_1}]^{\Pi \setminus S} \right) \) and \( \left( \circ_R - \circ_-, [e]^{\Pi \setminus S} \right) \). They are the same set. In fact, now \( S = \{ \alpha_1 \} \) and \( e \) and \( s_{\alpha_1} \) give the same coset in \( W/W_\Sigma \) so the corresponding signed-colored Dynkin diagrams agree too. Similarly \( \left( \circ_R - \circ_-, [s_{\alpha_1}]^{\Pi \setminus S} \right) \) is the same as \( \left( \circ_R - \circ_+, [e]^{\Pi \setminus S} \right) \). In our hexagon this means that two of the outer walls must be glued. The fact that a segment with a sign + is glued to one with − corresponds to the fact that the two contiguous chambers will form a Mobius band after the gluing (see Figure 5).

What this means is that in our geometric picture consisting of the inside of the 2ρ hexagon, some portions of the boundary need to be identified. Such identifications take place on all the chambers. This identification provides the gluing rule given in Lemma 4.2 in [15] for the case of \( \mathfrak{sl}(n, \mathbb{R}) \).

2.3. The chain complex \( \mathcal{M}_* \). We describe \( \mathcal{M}_* \) in terms of colored Dynkin diagrams. The \( \mathbb{Z} \) modules of chains \( \mathcal{M}_k \) are then given by (see Figure 6):

- \( \mathcal{M}_2 = \mathbb{Z}[[\circ - \circ, w] : w \in W] \). The cells are the chambers of the 2ρ hexagon, and \( \dim \mathcal{M}_2 = 6 \).
- \( \mathcal{M}_1 \) consists of the cells parametrized by the colored Dynkin diagrams \( \left( \circ_R - \circ_+, [w]^{\{\alpha_2\}} \right), \left( \circ_B - \circ_-, [w]^{\{\alpha_2\}} \right) \) with \( w \in \{ e, s_{\alpha_2}, s_{\alpha_1} s_{\alpha_2} \} \) and \( \left( \circ - \circ_R, [w]^{\{\alpha_1\}} \right), \left( \circ - \circ_B, [w]^{\{\alpha_1\}} \right) \) with \( w \in \{ e, s_{\alpha_1}, s_{\alpha_2} s_{\alpha_1} \} \). These are the sides of the different chambers of the hexagon. The blue (B) stands for internal chamber wall and the red (R) for external face of the hexagon. The dimension is then given by \( \dim \mathcal{M}_1 = 12 \).
Figure 6. The manifold $\hat{H}_R$ parametrized by colored Dynkin diagrams for $\mathfrak{sl}(3, \mathbb{R})$.

- $\mathcal{M}_0 = \mathbb{Z}[(\circ_s - \circ_t, [e]) : s, t \in \{R, B\}]$. These are the four vertices of a chamber. Because of identifications, there are only four different points coming from all the six chambers, that is, $\dim \mathcal{M}_0 = 4$.

The chain complex $\mathcal{M}_*: \mathcal{M}_2 \xrightarrow{\partial_2} \mathcal{M}_1 \xrightarrow{\partial_1} \mathcal{M}_0$ leads to the following integral homology: $H_2 = 0$, $H_1 = \mathbb{Z}^3 \oplus \mathbb{Z}/2\mathbb{Z}$, $H_0 = \mathbb{Z}$. This implies that $\hat{H}_R$ is nonorientable and is equivalent to the connected sum of two Klein bottles. Also note that the Euler character is $6 - 12 + 4 = -2$.

According to Proposition 7.7 this computes the homology $H_*(\hat{H}_R, \mathbb{Z})$ of the compact smooth manifold $\hat{H}_R$. The torsion $\mathbb{Z}/2\mathbb{Z}$ in $H_1$ has the following representative:

$$c_1 = \sum_{w \in W/W(\{s_{02}\})} (-1)^{\ell(w)} \left( (\circ_R - \circ, [w]^{\{s_{02}\}}) \right)$$

$$- \sum_{w \in W/W(\{s_{01}\})} (-1)^{\ell(w)} \left( (\circ_R - \circ, [w]^{\{s_{01}\}}) \right).$$

Here $\ell(w)$ denotes the length of $w$. If we let

$$c_2 = \sum_{w \in W} (-1)^{\ell(w)} (\circ - \circ, w),$$

then $\partial_2(c_2) = 2c_1$.

From the chain complex one can compute the three dimensional vector space $H_1(\hat{H}_R, \mathbb{Q})$ as a $W$-module. This is a direct sum of a one dimensional nontrivial (sign) representation and the two dimensional reflection representation. The representation $H_0(\hat{H}_R, \mathbb{Q})$ is trivial.
3. Notation, the group $\tilde{G}$.

3.1. Basic Notation. The important Lie group for the purposes of this paper is a group $\tilde{G}$ that will be technically introduced in Subsection 3.2. This group is $\mathbb{R}$ split and has a split Cartan subgroup with $2^l$ components with $l = \text{rank}(G)$. For example $SL(3, \mathbb{R})$ is of this kind but $SL(4, \mathbb{R})$ is not. We need to introduce the following standard objects before it is possible to define and study $\tilde{G}$.

**Notation 3.1** (Standard Lie theoretic notation). We adhere to standard notation and to the following conventions: $\ldots \mathbb{C}$ denotes a complexification; for any set $S \subset \Pi$, $\ldots_S$ is an object related to a parabolic subgroup or subalgebra determined by $S$; for any subset $A \subset \Pi$, $\ldots_A$ is associated to the parabolic determined by $\Pi \setminus A$. However we will not simultaneously employ $\ldots_A$ and $\ldots_S$ for the same object $\ldots$. As in the introduction $\mathfrak{g}$ denotes a real split semisimple Lie algebra of rank $l$ with complexification $\mathfrak{g}_{\mathbb{C}} = \mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$. We also have $G_{\mathbb{C}}$ the connected adjoint Lie group with Lie algebra $\mathfrak{g}_{\mathbb{C}}$ ($G_{\mathbb{C}} = \text{Ad}(G_{\mathbb{C}}^s)$) and $G$ the connected real semisimple Lie subgroup of $G_{\mathbb{C}}$ with Lie algebra $\mathfrak{g}$. Denote $G_{\mathbb{C}}^s$ the simply connected complex Lie group associated to $\mathfrak{g}_{\mathbb{C}}$. We list some additional very standard Lie theoretic notation:

- $\mathfrak{g}' = \text{Hom}_{\mathbb{R}}(\mathfrak{g}, \mathbb{R})$ and $\mathfrak{g}'_{\mathbb{C}} = \text{Hom}_{\mathbb{C}}(\mathfrak{g}_{\mathbb{C}}, \mathbb{C})$.
- Given $\lambda \in \mathfrak{g}'$ and $x \in \mathfrak{g}$, $\langle \lambda, x \rangle$ is $\lambda$ evaluated in $x$, $\langle \lambda, x \rangle = \lambda(x)$.
- $(,)$ the bilinear form on $\mathfrak{g}$ or $\mathfrak{g}_{\mathbb{C}}$ given by the Killing form (the same notation applies to the Killing form on $\mathfrak{g}'$ and $\mathfrak{g}'_{\mathbb{C}}$).
- $\theta$ a Cartan involution on $\mathfrak{g}$.
- $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$ the Cartan decomposition of $\mathfrak{g}$ associated to $\theta$, where $\mathfrak{k}$ is the Lie algebra of a maximal compact subgroup $K$ of the adjoint group $G$ and $\mathfrak{p}$ is the orthogonal complement to $\mathfrak{k}$ with respect to the Killing form.
- $e_\phi, h_\phi \in \mathfrak{g}, h$ root vectors chosen so that $(e_\phi, e_{-\phi}) = 1$.
- $\Delta_+ \subset \Delta$ be a fixed system of positive roots.
- $\mathfrak{b} = h + \sum_{\phi \in \Delta_+} \mathbb{R}e_\phi$ (Borel subalgebra).
- $\mathfrak{n} = [\mathfrak{b}, \mathfrak{b}] = \sum_{\phi \in \Delta_+} \mathbb{R}e_\phi$ and $\mathfrak{n} = \sum_{\phi \in \Delta_+} \mathbb{R}e_{-\phi}$.
- $\mathfrak{n}_{\mathbb{C}} = \mathfrak{n} \otimes_{\mathbb{R}} \mathbb{C}$, $\mathfrak{n}_{\mathbb{R}} = \mathfrak{n} \otimes_{\mathbb{R}} \mathbb{R}$.
- $H_{\mathbb{C}}$ Cartan subgroup of $G_{\mathbb{C}}$ with Lie algebra $\mathfrak{h}_{\mathbb{C}}$.
- $H = H(\Delta)$ the connected Lie subgroup of $G$ with Lie algebra $\mathfrak{h}$, $H = \exp(\mathfrak{h})$.
- $H_{\mathbb{R}}^1$ Cartan subgroup of $G$ with Lie algebra $\mathfrak{h}_{\mathbb{R}}$.
- $W$ the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{h}$ or the Weyl group of $H_{\mathbb{R}}^1$. 

• $W_S \subset W$, the group generated by the simple reflections corresponding to the elements in $S \subset \Pi$.

**Remark 3.2.** The group $H^1_R$ [22] p. 59, 2.3.6 consists of all $g \in G$ such that $\text{Ad}(g)$ restricted to $\mathfrak{h}$ is the identity. This Cartan subgroup will be usually disconnected and $H$ is the connected component of the identity $e$. The Weyl group $W$ of the Cartan subgroup $H^1_R$ is isomorphic to the group which is generated by the simple reflections $s_{\alpha_i}$ with $\alpha_i \in \Pi$ and which agrees with the Weyl group associated to the pair $(g_C, h_C)$. This is because our group is assumed to be $\mathbb{R}$ split.

**Example 3.3.** The reader may wish to read the whole paper with the following well-known example in mind. Let $G_C = SL(n, \mathbb{C})$ and $G_C = \text{Ad}(SL(n, \mathbb{C}))$. The second group is obtained by dividing $SL(n, \mathbb{C})$ by the finite abelian group consisting of the $n$ roots of unity times the identity matrix. We can set $G = \text{Ad}(SL(n, \mathbb{R}))$. Note that if $n$ is odd then $SL(n, \mathbb{R}) = \text{Ad}(SL(n, \mathbb{R})) = G$. If $n$ is even then $\text{Ad}(SL(n, \mathbb{R}))$ is obtained by dividing by $\{ \pm I \} (I$ the identity matrix). In this example, $\mathfrak{g}$ consists of traceless $n \times n$ real matrices. The Cartan subalgebra $\mathfrak{h}$ in the $\text{Ad}(SL(n, \mathbb{R}))$ case can be taken to be the space of traceless $n \times n$ real diagonal matrices. The root vectors $e_{\alpha}$ are the various matrices with all entries $a_{i,j} = 0$ if $(i, j) \neq (i_o, j_o)$ and $a_{i_o,j_o} = 1$ where $i_o \neq j_o$ are fixed integers. In this case of $G = \text{Ad}(SL(n, \mathbb{R}))$, with $\mathfrak{h}$ chosen as above, the Cartan subgroup $H^1_R$ of $G$ consists of $\text{Ad}$ applied to the group of real diagonal matrices of determinant one. The group $H$ consists of $\text{Ad}$ applied to all diagonal matrices $\text{diag}(r_1, \ldots, r_n)$ with $r_i > 0$ the connected component of the identity of $H^1_R$.

**Notation 3.4.** The following elements in $\mathfrak{h}$ and its dual $\mathfrak{h}'$ will appear often in this paper:

- $\tilde{\alpha}_i = \frac{2\alpha_i}{(\alpha_i, \alpha_i)}$ the coroots.
- $m_{\alpha_i}$, $i = 1, \ldots, l$, the fundamental weights, $m_{\alpha_i} = (m_{\alpha_i}, \tilde{\alpha}_j) = \delta_{i,j}$.
- $m_{\alpha_i}^{\circ}$ the unique element in $\mathfrak{h}$ defined by $\langle m_{\alpha_i}, x \rangle = \langle m_{\alpha_i}^{\circ}, x \rangle$.
- $h_{\alpha_i}$ is the unique element in $\mathfrak{h}$ such that $\langle h_{\alpha_i}, x \rangle = \langle \tilde{\alpha}_i, x \rangle$.
- $y_i = \frac{2\pi m_{\alpha_i}^{\circ}}{(\alpha_i, \alpha_i)}$.
- $\mathfrak{h}^\circ = \{ x \in \mathfrak{h} : \langle \alpha_i, x \rangle < 0 \text{ for all } \alpha_i \in \Pi \}$.

**Example 3.5.** Consider $G_C = \text{Ad}(SL(n, \mathbb{C})) \subset \text{Ad}(GL(n, \mathbb{C}))$ and $n$ even. We view $g_C$ inside the Lie algebra of $GL(n, \mathbb{C})$ (as $n \times n$ complex traceless matrices). Then $m_{\alpha_i}^{\circ} = \text{diag}(t_1, \ldots, t_n) + z$ where $t_j = 1$ for $j \leq i$ and $t_j = 0$ for $j > i$. The element $z$ is in the center of the Lie algebra of $GL(n, \mathbb{C})$ and thus $\text{ad}(z) = 0$. For instance if $n = 2$ then $m_{\alpha_1}^{\circ} = \text{diag}(\frac{1}{2}, -\frac{1}{2}) = \text{diag}(1, 0) - \text{diag}(\frac{1}{2}, \frac{1}{2})$.

**3.2. The group $\tilde{G}$.** We now define, following [16], 3.4.4 in p. 241 an enlargement $\tilde{G}$ of the split group $G$. The purpose of this is to “complete”
the Cartan subgroup $H_{\mathbb{R}}^l$ forcing it to have $2^l$ connected components where $l$ is the rank of $G$. In the case $G = SL(l + 1, \mathbb{R})$ where $l$ is even, $\tilde{G} = G$ already. We thus consider another real split group, slightly bigger than $G$.

Let $x, y \in \mathfrak{h}$ so that $x + \sqrt{-1}y \in \mathfrak{h}_C$. We recall the conjugate linear automorphism of $\mathfrak{g}_C$ given by $z^c = x - \sqrt{-1}y$ where $z = x + \sqrt{-1}y$ and $x, y \in \mathfrak{g}$. We recall that this automorphism induces an automorphism $G_C \to G_C$, $g \mapsto g^c$.

We thus let

$$\tilde{G} = \{g \in G_C : g^c = g\}.$$

By [16] Proposition 3.4, we also have:

$$\tilde{G} = \{g \in G_C : \text{Ad}(g) \mathfrak{g} \subset \mathfrak{g}\}.$$

Then we have the following Proposition whose proof will be given later (right after Proposition 3.15):

**Proposition 3.6.** Let $G = \text{Ad}(SL(n, \mathbb{R}))$, then $\tilde{G} \cong SL(n, \mathbb{R})$ for $n$ odd and $\tilde{G} \cong \text{Ad}(SL(n, \mathbb{R})^\pm)$ if $n$ is even. Thus $\tilde{G}$ is disconnected whenever $n$ is even.

We now describe an element $h_i$ in the Cartan subgroup of $\tilde{G}$. First we have:

**Lemma 3.7.** Let $x, y \in \mathfrak{h}$ and $\alpha_i \in \Pi$. Then the numbers $\exp(\langle \alpha_i, x + \sqrt{-1}y \rangle)$ are real if and only if $y$ has the form: $y = \sum_{i=1}^l k_i y_i$ where $k_i$ is an integer and $y_i$ is as in Notation 3.4. The elements $h_i = \exp(\sqrt{-1}k_i y_i)$ with $k_i$ odd are in $\tilde{G}$ and satisfy $h_i^2 = e$ with $h_i \neq e$.

**Proof.** It is enough to consider the case when $x = 0$. Suppose first that $y = \sum_{i=1}^l k_i y_i$ where each $k_i$ is an integer. Then $e^{\sqrt{-1}(\alpha_i, y)} = e^{\sqrt{-1}\pi k_i}$ takes either +1 or −1. Conversely, suppose that all the $e^{\langle \alpha_i, \sqrt{-1}y \rangle}$ with $i = 1, \ldots, l$ are real. Then $e^{\langle \alpha_i, \sqrt{-1}y \rangle}$ equals ±1 and $\langle \alpha_i, y \rangle = k_i \pi$ for all $i = 1, \ldots, l$. This implies that, for each $i$, $\langle \alpha_i, y \rangle = \frac{2k_i \pi}{(\alpha_i, \alpha_i)}$. Since $\mathfrak{g}$ is semisimple and the $\{m^\circ_{\alpha_i} : i = 1, \ldots, l\}$ forms a basis of $\mathfrak{h}$, necessarily $y = \sum_{i=1}^l c_i m^\circ_{\alpha_i}$ with $\langle \alpha_i, y \rangle = c_i$ and $\langle \alpha_i, y \rangle = \frac{2k_i \pi}{(\alpha_i, \alpha_i)}$. This proves the first part of the statement in Lemma 3.7. We note that since all the $e^{\langle \alpha_j, y_i \rangle}$ are real then $e^{\langle \phi, \sqrt{-1}y_i \rangle}$ is also real for any root $\phi$ which is not necessarily simple. Therefore each $\text{Ad}(h_i)$ stabilizes all the root spaces of $\mathfrak{h}$ in $\mathfrak{g}$. Since $\mathfrak{h}$ is also stabilized, we obtain that $\text{Ad}(h_i)(\mathfrak{g}) \subset \mathfrak{g}$ and thus $h_i \in \tilde{G}$. Clearly $h_i = \exp(\sqrt{-1}y_i)$ satisfies $h_i^2 = 1$ where $h_i \neq e$. This is because $h_i$ is representable by a diagonal matrix with entries of the form ±1. Moreover, at least one diagonal entry must be equal to −1. \(\square\)

**Example 3.8.** In the case of $\text{Ad}(SL(n, \mathbb{R})^\pm)$ with $n$ even, $h_i = \exp(\sqrt{-1}y_i)$ is just $\text{Ad}(\text{diag}(r_1, \ldots, r_n))$ where $r_j = 1$ if $j \leq i$ and $r_j = -1$ if $j > i$. When
$n$ is odd then we have $h_i = (-1)^{n-i} \text{diag}(r_1, \ldots, r_n)$ with the same notation as above for the $r_i$. In this case $SL(n, \mathbb{R}) = \text{Ad}(SL(n, \mathbb{R}))$.

We now describe a split Cartan subgroup $H_R$ of $\tilde{G}$ and other items related to it. The $H_R$ is the real part of $H_C$ on $\tilde{G}$,

$$H_R = H_C \cap \tilde{G},$$

which has $2^l$ components (see Proposition 3.15 below). We denote by $B$ a Borel subgroup with Lie algebra $\mathfrak{h} + \mathfrak{n}$ contained in $\tilde{G}$ or $G$ as will be clear from the context. Thus in $G$ this is $H_R N$, $N = \exp(n)$. From the Bruhat decomposition applied to $\tilde{G}$, we have

$$\tilde{G} = \bigsqcup w \in W \tilde{N} \dot{w} H_R N,$$

with $\tilde{N} = \exp(\tilde{n})$. Here $\dot{w}$ stands for any representative of the Weyl group element $w \in W$ in the normalizer of the Cartan subgroup. Keeping this in mind we will harmlessly drop the $\dot{}$ from the notation.

In addition let $\chi_{\phi}$ denote the group character determined by $\phi \in \Delta$; on $H_R$ each $\chi_{\phi}$ is real and cannot take the value zero. Thus a group character has a fixed sign on each connected component. We denote by $\text{sign}(\chi_{\alpha_i}(h))$ the sign of this character on a specified element $h$ of $H_R$. From the indefinite Toda lattice in p. 323 of [14]. In connection with the connected components let $h_\epsilon = \prod_{\epsilon_i \neq 1} h_i$ for $\epsilon \in \mathcal{E}$ and $h_i = \exp(\sqrt{-1} y_i)$ in Lemma 3.7. Now $H_\epsilon = h_\epsilon H = \{h \in H_R : \text{sign}(\chi_{\alpha_i}(h)) = \epsilon_i \text{ for all } \alpha_i \in \Pi\}$, and we have:

$$H_R = \bigcup_{\epsilon \in \mathcal{E}} H_\epsilon.$$

**Notation 3.9.** We need notation to parametrize connected the components of a split Cartan subgroup, roots and root characters. Unfortunately we need such notation for all the parabolic subgroups associated to arbitrary subsets of $\Pi$. Recall (Notation 3.1) that notation associated to a parabolic subgroup determined by a subset $\Pi \setminus A$, $A \subset \Pi$ is usually indicated by changing the standard notation with the use of a superscript $A$. Thus we have: $\Delta_A \subset \Delta$, root system giving rise to a semisimple Lie algebra $\mathfrak{l}^A \subset \mathfrak{g}$. Also there are corresponding connected semisimple Lie subgroups $L^A_\mathbb{C} \subset G_\mathbb{C}$, $L^A \subset L^A_\mathbb{C}$. The adjoint group is denoted by $L_C(\Delta_A) = \text{Ad}(L^A_\mathbb{C})$ and it has a real connected Lie subgroup $L(\Delta_A)$. Let $\tilde{L}(\Delta_A)$ be defined in the same way as $\tilde{G}$ but relative to the root system $\Delta_A$. We let $\mathfrak{h}^A$ be the real span of the $h_{\alpha_i}$ with $\alpha_i \notin A$. This is a (split) Cartan subalgebra of $\mathfrak{l}^A$ denoted as $H^A_R$. The
corresponding connected Lie subgroup is \( H^A = \exp(\mathfrak{h}^A) \) (exponentiation taking place inside \( G_\mathbb{C} \)).

We also consider Lie subgroups of \( \tilde{G} \) corresponding to the subsystems of the Toda lattice. In accordance to our convention for Levi subgroups associated with \( A \in \Pi \), we have

\[
\mathcal{E}^A = \{ \epsilon = (\epsilon_1, \ldots, \epsilon_i) : \epsilon_i = 1 \text{ if } \alpha_i \in A \}.
\]

Thus \( H^A \) is by definition a subgroup of \( H \). This Lie subgroups of \( H \) is isomorphic to the connected component of the identity of a Cartan subgroup of a real semisimple Lie group that corresponds to \( \mathfrak{t}^A \). There is a bijection

\[
(10) \quad \mathcal{E}^A \cong \mathcal{E}(\Delta^A) = \{ (\epsilon_1, \ldots, \epsilon_m) : \alpha_j, \not\in A \text{ for } i = 1, \ldots, m = |\Pi \setminus A| \}
\]

which is given in the obvious way by restricting a function \( \epsilon : \Pi \to \{\pm1\} \) such that \( \epsilon_i = \epsilon(\alpha_i) = 1 \) for \( \alpha_i \in A \) to a new function \( \epsilon(\Delta^A) \) with domain \( \Pi \setminus A \).

We now consider the Cartan subalgebras and the Cartan subgroups for Levi factors of parabolic subalgebras and subgroups determined by \( A \subset \Pi \):

**Definition 3.10.** For \( A \subset \Pi \), we denote:

- \( H^A_\mathbb{R} = \bigcup_{\epsilon \in \mathcal{E}^A} h_\epsilon H^A \) (if \( \Delta^A = \emptyset \) (i.e., \( A = \Pi \)) then \( H^A_\mathbb{R} = \{e\} \)).
- \( H^A_\mathbb{R}(\Delta^A) \) a Cartan subgroup of \( \tilde{L}(\Delta^A) \), defined in the same way as \( H_\mathbb{R} \) and having Lie algebra \( \mathfrak{h}^A \) (\( H_\mathbb{R}(\Delta^A) = \{e\} \) if \( \Delta^A = \emptyset \)).
- \( H^A_{\mathbb{R},\le} = \{ h \in h_\epsilon H^A : \text{for all } \alpha_i \in \Pi \setminus A : |\chi_\alpha(h)| \le 1 \} \) the antidominant chamber of \( H^A_\mathbb{R} = h_\epsilon H^A \). Similarly we consider \( H^A_{\mathbb{R},<} \) using strict inequalities.
- \( H(\Delta^A)_{\le} = \{ h \in H(\Delta^A)_\mathbb{R} : \text{for all } \alpha_i \in \Pi \setminus A : |\chi_{\Delta^A}(h)| \le 1 \} \). Similarly we consider a version with strict inequalities.
- \( \chi_{\Delta^A} \) the root character associated to \( \alpha_i \) on the Cartan \( H_\mathbb{R}(\Delta^A) \).

We use notation \( \ldots \le \) to indicate the antidominant chamber on Cartan subalgebras and subgroups and \( \ldots < \) for strictly antidominant chambers.

**Example 3.11.** In the case of \( \tilde{G} = SL(3, \mathbb{R}) \), \( H_\mathbb{R} \) is the group

\[
H_\mathbb{R} = \{ \text{diag}(a, b, c) : a \neq 0, b \neq 0, abc = 1 \}.
\]

For \( A = \{\alpha_1\} \), \( H^A \) is the group \( \{ \text{diag}(1,a,a^{-1}) : a > 0 \} \). The set \( \mathcal{E}^A \) consists of \( (1,1) \) and \( (1,-1) \). The element \( h_{(1,-1)} = \text{diag}(-1,-1,1) \) and thus \( h_{(1,-1)} H^A = \{ \text{diag}(-1,-a,a^{-1}) : a > 0 \} \). These two components form \( H^A_\mathbb{R} \). Note that \( H_\mathbb{R} L^A \) is the Lie subgroup of \( SL(3, \mathbb{R}) \) consisting of all real matrices of the form,

\[
H_\mathbb{R} L^{\{\alpha_1\}} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & d & e \end{pmatrix}
\]
having determinant one. The group \( L^A \) is obtained by setting \( a = 1 \) and the determinant equal to one. The group \( L(\Delta^A) \) is isomorphic to \( \text{Ad}(SL(2,\mathbb{R})) \) the adjoint group obtained from \( L^A \) and \( \tilde{L}(\Delta^A) \) is isomorphic to \( \text{Ad}(SL(2,\mathbb{R})^\pm) \). Thus these three groups \( L^A, L(\Delta^A) \) and \( \tilde{L}(\Delta^A) \) are all different in this case.

**Definition 3.12.** Recall that the \( S \) is a subset of \( \Pi \) indicating colored vertices in a Dynkin diagram. Let \( \eta : S \to \{\pm 1\} \) be any function. We let \( \epsilon_\eta \in \mathcal{E}_{\Pi \setminus S} \) defined by \( \epsilon_\eta(\alpha_i) = \eta(\alpha_i) \) if \( \alpha_i \in S \), \( \epsilon_\eta(\alpha_i) = 1 \) if \( \alpha_i \notin S \). Thus there is a bijective correspondence between the set of all functions \( \eta \) and the set \( \mathcal{E}_{\Pi \setminus S} \) and (by Equation (10)) a second bijection with the set \( \mathcal{E}(\Delta^{\Pi \setminus S}) \):

\[
\eta \mapsto \epsilon_\eta \in \mathcal{E}_{\Pi \setminus S},
\]

and

\[
\eta \mapsto \epsilon_\eta(\Delta^{\Pi \setminus S}) \in \mathcal{E}(\Delta^{\Pi \setminus S}).
\]

The exponential map \( h_\epsilon \exp : \mathfrak{h} \to h_\epsilon H \) allows us to define chamber walls in \( H \) and therefore in any of the connected components of \( H_\mathbb{R} \). For any root \( \phi \in \Delta \) the set \( \{ h \in h_\epsilon H : |\chi_\phi(h)| = 1 \} \) defines the \( \phi \)-wall of \( H_\epsilon = h_\epsilon H \). The intersection of all the \( \alpha_i \)-walls of \( H_\epsilon \) is the set \( \{ h_\epsilon \} \). This is also the intersection of all the \( \phi \)-walls of \( H_\epsilon \) with \( \phi \in \Delta \).

We also define the following subsets of \( H_\epsilon \):

**Definition 3.13.** We denote:

- \( \mathcal{D} = \mathcal{D}(\Delta) = \{ h_\epsilon : \epsilon \in \mathcal{E} \} \),
- \( \mathcal{D}(\Delta^{\Pi \setminus S}) = \{ h_\epsilon : \epsilon \in \mathcal{E}(\Delta^{\Pi \setminus S}) \} \).

The set \( \mathcal{D} \) has two structures: It is a finite group and also a set with an action of \( W \). In Proposition 3.15 it is the first structure that is emphasized but in the Proof of Proposition 4.5 it is the second structure which is relevant. We now look at the \( W \)-action.

Since \( W \) acts on \( H_\mathbb{R} \) and \( w \in W \) sends a \( \phi \)-wall of \( H_\epsilon \) to the \( w(\phi) \)-wall of some other \( H_\epsilon \), this set \( \mathcal{D} \) is preserved by \( W \) and thus acquires a \( W \)-action. Given \( S \subset \Pi \) we similarly obtain that \( \mathcal{D}(\Delta^{\Pi \setminus S}) \) has a \( W_S \)-action.

The map

\[
\mathcal{D}(\Delta^{\Pi \setminus S}) \to \mathcal{E}(\Delta^{\Pi \setminus S})
\]

sending \( h_\epsilon \mapsto \epsilon \) also defines an action of \( W_S \) on the set of signs \( \mathcal{E}(\Delta^{\Pi \setminus S}) \). Recall (Notation 3.9) that \( \mathcal{E}^{\Pi \setminus S} \subset \mathcal{E} \) denotes those \( \epsilon \) for which \( \epsilon_i = 1 \) whenever \( \alpha_i \notin S \). We have a bijection (by (11) together with (10))

\[
\mathcal{E}^{\Pi \setminus S} \cong \mathcal{D}(\Delta^{\Pi \setminus S}).
\]

The \( W_S \)-action on the set \( \mathcal{D}(\Delta^{\Pi \setminus S}) \) thus gives a \( W_S \)-action on \( \mathcal{E}^{\Pi \setminus S} \) such that for any \( \epsilon \in \mathcal{E}^{\Pi \setminus S} \) only the \( \epsilon_j \) with \( \alpha_j \in S \) may change in sign under
Proof. Note that this construction requires looking at the $h_\epsilon$ with $\epsilon \in \mathcal{E}(\Delta^{\Pi}/\mathcal{S})$ in the adjoint representation of $t^A$.

Remark 3.14. The root characters can be expressed as a product of the simple root characters raised to certain integral powers $\phi = \sum_{i=1}^l c_i \alpha_i$ with $c_i \in \mathbb{Z}$, $e^\phi = \prod_{i=1}^l \chi_{\alpha_i}^{c_i}$. Therefore if $h \in H_{\mathbb{C}}$ the scalars $\chi_{\alpha_i}(h)$ determine all the scalars $e^{\phi}(h) = \chi_{\phi}(h)$ and thus $h$ is uniquely determined. Moreover $e^{\langle \phi, x + \sqrt{-1}y \rangle}$ is real for all $\phi \in \Delta$ if and only if $e^{\langle \alpha_i, x + \sqrt{-1}y \rangle}$ is real for all $\alpha_i \in \Pi$.

Proposition 3.15. The Cartan subgroup $H_{\mathbb{R}}$ of $\tilde{G}$ has $2^l$ components. We have $H_{\mathbb{R}} = DH$ where $D$ (Definition 3.13) is the finite group of all the $h_\epsilon$, $\epsilon \in \mathcal{E}$.

Proof. It was shown in Lemma 3.7 that the elements $\exp(x + \sqrt{-1}y)$ with $x, y \in \mathfrak{h}$ such that $e^{\langle \alpha_i, x + \sqrt{-1}y \rangle}$ is real for all $\alpha_i \in \Pi$ are those for which $y$ has the form $y = \sum_{i=1}^l k_i y_i$ with $y_i$ as in Notation 3.4, and $k_i$ are integers. Moreover as in Remark 3.14 we also have that $e^{\langle \phi, x + \sqrt{-1}y \rangle}$ is real for any $\phi \in \Delta$ exactly when $y = \sum_{i=1}^l k_i y_i$ for some integers $k_i$. Therefore all the root spaces of $\mathfrak{g}$ are stabilized under the adjoint action of $\exp(x + \sqrt{-1}y)$. Since clearly $\mathfrak{h}$ is also stabilized, then $\exp(x + \sqrt{-1}y)$ stabilizes all of $\mathfrak{g}$ and this implies that $\exp(x + \sqrt{-1}y) \in \tilde{G}$. In fact this shows that $\exp(x + \sqrt{-1}y) \in \tilde{G}$ if and only if $y$ has the form $y = \sum_{i=1}^l k_i y_i$ for certain $k_i$ integers. Thus Lemma 3.7 and these remarks compute the intersection $\tilde{G} \cap H_{\mathbb{C}}$. From here it is easy to conclude.

Note that Proposition 3.15 implies that $D(\Delta)$ in Definition 3.13 is isomorphic to $H_{\mathbb{R}}/H$ as a set with a $W$-action. That the $W$-actions agree is verified in Corollary 4.6.

We now give the Proof of Proposition 3.6:

Proof. Let $\text{Ad}$ denote the representation of $GL(n, \mathbb{C})$ on $\mathfrak{sl}(n, \mathbb{C})$. Then we have $\text{Ad}(GL(n, \mathbb{C})) = \text{Ad}(SL(n, \mathbb{C}))$. If $n$ is odd, $\text{Ad}(SL(n, \mathbb{C}))$ is isomorphic to $SL(n, \mathbb{C})$. Denote $D_i = \text{diag}(r_1, \ldots, r_n)$ with $r_i$ as in Example 3.8. We set $\tilde{f}_i = D_i$ when $n$ is even and $\tilde{f}_i = (-1)^{n-1}D_i$ when $n$ is odd. Let $h_i = \text{Ad}(\tilde{f}_i)$. We have that for each $h_i \in \tilde{G}$, $\chi_{\alpha_i}(h_i) = e^{\pi \sqrt{-1}\delta_{ij}}$ for $\alpha_i \in \Pi$.

The $h_i$ generate the group $D$ with $2^l$ elements where $l = n-1$ and now the group $G_1 = \langle \text{Ad}(SL(n, \mathbb{R})), h_i, i = 1, \ldots, l \rangle$ is a subgroup of $\tilde{G}$ and it is isomorphic to $\text{Ad}(SL(n, \mathbb{R}))$ for $n$ even and to $\text{Ad}(SL(n, \mathbb{R})) \cong SL(n, \mathbb{R})$ for $n$ odd. What remains is to verify that $\tilde{G} \subset G_1$.

Let $\tilde{K} = \{ \text{Ad}(g) : g \in U(n) \} \cap \tilde{G}$. By the Iwasawa decomposition of $\tilde{G}$ and $G_1$, it suffices to show that $\tilde{K} = KD$, $K = SO(n)$. The right side, $KD$ is either $O(n)$ or $SO(n)$ according to the parity of $n$. Recall that the
maximal compact Lie subgroup $\overline{K}$ of $\overline{G}$ acts transitively on the set $X$ of all maximal abelian Lie subalgebras $\mathfrak{a}$ which are contained in the vector space $\mathfrak{p}$. The action of $K = \text{Ad}(SO(n))$ is also transitive on $X$ (by (2.1.9) of [22]) and thus for any $g \in \overline{K}$ there is $k \in K$ such that $g = kD$ where $D$ is the isotropy group (in $\overline{K}$) of an element in $X$, for instance of the element $\mathfrak{h} \in X$. However this isotropy group $D$ has been computed implicitly in the Proof of Proposition 3.15 and $D = \mathcal{D}$. Therefore $\overline{K} = K\mathcal{D}$.

**Proposition 3.16.** Let $\alpha_i \in \Pi$ and assume that $\epsilon = (\epsilon_1, \ldots, \epsilon_l) \in \mathcal{E}$. Then $s_{\alpha_i}h_\epsilon = h_{\epsilon'}$ where $\epsilon'_j = \epsilon_j \epsilon_i^{-C_{j,i}}$ with $(C_{i,j})$ the Cartan matrix.

**Proof.** This follows from the expression of the Weyl group action on elements in $\mathfrak{h}_C^\epsilon$ given by: $s_{\alpha_i}x = x - (\tilde{\alpha}_i, x)\alpha_i$. If this expression is applied to $x = \alpha_j$ it gives $s_{\alpha_i}\alpha_j = \alpha_j - C_{j,i}\alpha_i$. On the level of root characters this becomes, by exponentiation of the previous identity,

$$s_{\alpha_i}\chi_{\alpha_j} = \chi_{\alpha_j}\chi_{\alpha_i}^{-C_{j,i}}.$$  

Now recall that $\epsilon_j$ is just $\chi_{\alpha_j}$ evaluated at $h_\epsilon$. Also $\epsilon'_j$ will be $\chi_{\alpha_j}$ evaluated at $s_{\alpha_i}h_\epsilon$. When we evaluate $\chi_{\alpha_j}$ on $s_{\alpha_i}h_\epsilon$ in order to compute the corresponding $j$-th sign, we obtain $\chi_{s_{\alpha_i}\alpha_j}(h_\epsilon)$. Therefore the sign of $\chi_{\alpha_j}$ on the $s_{\alpha_i}h_\epsilon$ is given by the product $\epsilon_j \epsilon_i^{-C_{j,i}}$. Finally we use the fact that the set of all scalars $\chi_{\alpha_i}(h)$ determines $h$. Thus $\epsilon'$ determines the element $h_{\epsilon'}$ giving rise to the equation $s_{\alpha_i}h_\epsilon = h_{\epsilon'}$. 

The sign change $\epsilon_j \rightarrow \epsilon'_j$ in Proposition 3.16 is precisely the gluing rule in Lemma 4.2 for the indefinite Toda lattice in [15]. Then the gluing pattern using the Toda dynamics is just to identify each piece of the connected component $H_\epsilon$ (see Figure 1). The sign change on subsystem corresponding to $H_\epsilon^A$ with $A \subset \Pi$ can be also formulated as:

**Proposition 3.17.** Let $\alpha_i \in \Pi \setminus A$ and assume that $\epsilon = (\epsilon_1, \ldots, \epsilon_l) \in \mathcal{E}_A$. Then:

a) If $\epsilon_i = 1$, $s_{\alpha_i}h_\epsilon = h_\epsilon$. If $\epsilon_i = -1$ then $s_{\alpha_i}h_\epsilon = h_{\epsilon'}$ where $\epsilon'_j = \epsilon_j \epsilon_i^{-C_{j,i}}$. In addition $h_{\epsilon'}$ factors as a product $(\prod_{\alpha_j \in A,C_{j,i} \text{ is odd}} h_j)h_{\epsilon_A}$ where $\epsilon_A \in \mathcal{E}_A$.

b) If $\epsilon_i = 1$, $s_{\alpha_i}H_\epsilon^A \subset H_\epsilon^A$. If $\epsilon_i = -1$ then $s_{\alpha_i}H_\epsilon^A \subset h_{\epsilon'}H_\epsilon^A$ where $\epsilon'_j = \epsilon_j \epsilon_i^{-C_{j,i}}$. In addition $h_{\epsilon'}$ factors as a product $(\prod_{\alpha_j \in A,C_{j,i} \text{ is odd}} h_j)h_{\epsilon_A}$ where $\epsilon_A \in \mathcal{E}_A$.

c) The sign $\epsilon_k = \text{sign}(\chi_{\alpha_k}(h))$ for any $h \in H_\epsilon$ agrees with $\text{sign}(\chi_{s_{\alpha_i}\alpha_k}(h'))$ for any $h' \in s_{\alpha_i}(H_\epsilon)$.

**Proof.** Part a) follows from Proposition 3.16 but with the observation that in this case, $\epsilon'$ may fail to be in $\mathcal{E}_A$ even if $\epsilon \in \mathcal{E}_A$. This happens exactly
when $\epsilon_i = -1$ and $\alpha_j \in A$ with $C_{j,i}$ odd (either $-3$ or $-1$). Under these circumstances $\epsilon'_j = -1$ (rather than one as required in the definition of $E^A$).

We can fix this problem by factoring $h_{\epsilon'}$ as a product $(\prod_{\alpha_j \in A, C_{j,i} \text{ is odd}} h_j)^{\epsilon_A}$ where $\epsilon_A \in E^A$.

Part b) follows from Part a) and the fact that each $h_{\epsilon} H$ is a connected component of $H_{\mathbb{R}}$ in Proposition 3.15.

Part c) follows easily from $\chi_{s_{\alpha_i}}(s_{\alpha_k}) = \chi_{\alpha_k}(h)$ and that each $s_{\alpha_i} H$. The two desired signs have thus been computed in the two connected components and they agree. □

Remark 3.18. From Proposition 3.17 it follows that any connected component $H_{\epsilon}^A$, is the union of chambers of the form $w \left( H_{\epsilon(w)}^{A_{\leq}} \right)$, with $w \in W_{\Pi \setminus A}$, $\epsilon(w) \in E$,

$$H_{\epsilon}^A = \bigcup_{w \in W_{\Pi \setminus A}} w \left( H_{\epsilon(w)}^{A_{\leq}} \right).$$


We now introduce some notation that will ultimately parametrize the cells in a cellular decomposition of the smooth compact manifold $\hat{H}_{\mathbb{R}}$ to be defined in §6.

4.1. Colored Dynkin Diagrams. Let us first define:

**Definition 4.1** (Colored Dynkin diagrams $\mathbb{D}(S)$). A colored Dynkin diagram is a Dynkin diagram where all the vertices in a set $S \subset \Pi$ have been colored either red $R$ or blue $B$. For example, in $\mathfrak{sl}(4, \mathbb{R})$, $\alpha_R - \alpha_R - \alpha_B$ is a colored Dynkin diagram with $S = \{\alpha_1, \alpha_3\}$. Thus a colored Dynkin diagram where $S \neq \emptyset$, corresponds to a pair $(S, \epsilon_{\eta})$ with $S \subset \Pi$ and $\eta : S \to \{\pm 1\}$ any function. Here $\eta(\alpha_i) = -1$ if $\alpha_i$ is colored $R$ and $\eta(\alpha_i) = 1$ if $\alpha_i$ is colored $B$. If $S = \emptyset$ then $\epsilon_0$ with $\epsilon_0(\alpha_i) = 1$ for all $\alpha_i \in \Pi$ replaces $\epsilon_{\eta}$. We denote:

- $D = (S, \epsilon_{\eta})$ or $(S, \epsilon_{\eta}(\Delta_{\Pi \setminus S}))$ with $\epsilon_{\eta} \in E^{\Pi \setminus S}$;
- $\mathbb{D}(S) = \{D = (S, \epsilon_{\eta}) : \text{the vertices in } S \text{ are colored}\}$.

We also introduce an oriented colored Dynkin diagram which is defined as a pair $(D, o)$ with $o \in \{\pm 1\}$ and $D$ a colored Dynkin diagram.

4.2. Boundary of a colored Dynkin diagram. We now define the boundaries of a cell parametrized by a colored Dynkin diagram $D$.

**Definition 4.2** (The boundary $\partial_{j,c}D$). For each $(j, c)$ with $c = 1, 2$ and $j = 1, \ldots, m$ we define a new colored Dynkin diagram $\partial_{j,c}D$, the $(j, c)$-boundary of the $D$ by considering $\{\alpha_{i_j} : 1 \leq i_1 < \cdots < i_m \leq l\}$ the set...
Figure 7. The boundary \( \partial_{j,c} \) in the case of \( \mathfrak{sl}(3, \mathbb{R}) \).

\( \Pi \setminus S \) of uncolored vertices and \( m = |\Pi \setminus S| \). The \( \partial_{j,c}D \) is then a new colored Dynkin diagram obtained by coloring the \( i_j \)-th vertex with \( R \) if \( c = 1 \) and with \( B \) if \( c = 2 \). The boundary of an oriented colored Dynkin diagram \( (D, o), o \in \{ \pm 1 \} \) is, in addition, given an orientation defined to be the sign \( (−1)^{j+c+1}o \). Recall that a colored Dynkin diagram \( D \) corresponds to a pair \( (S, \epsilon) \) with \( S \subset \Pi \) and \( \eta : S \rightarrow \{ \pm 1 \} \) (Definition 3.12). Thus the boundary \( \partial_{j,c} \) determines a new pair \( (S \cup \{ \alpha_i \}, \epsilon') \) associated to \( \partial_{j,c}D \). We can then define the following boundary maps,

\[
(-1)^{j+c+1}\partial_{j,c} : \mathbb{Z}[\mathcal{D}(S)] \rightarrow \mathbb{Z}[\mathcal{D}(S \cup \{ \alpha_i \})].
\]

**Example 4.3.** The boundary of \( \circ - \circ \) (which we can picture as a box) consists of segments (one dimensional boxes) given by

\[
\partial_{1,1}(\circ - \circ) = \circ_R - \circ, \quad \partial_{2,1}(\circ - \circ) = \circ - \circ_R, \quad \partial_{1,2}(\circ - \circ) = \circ_B - \circ \quad \text{and} \quad \partial_{2,2}(\circ - \circ) = \circ - \circ_B.
\]

The orientation sign associated to \( \circ_R - \circ \) is the following: With \( c = 1 \) and \( j = 1 \) one has \( (−1)^{j+c+1} = (−1)^3 = −1 \). We illustrate the example in Figure 7.

### 4.3. \( W_S \)-action on colored Dynkin diagrams.

We now move these colored Dynkin diagrams around with elements in \( W \). A \( W_S \)-action on the diagram \( D \in \mathcal{D}(S) \), \( W_S : \mathcal{D}(S) \rightarrow \mathcal{D}(S) \), is defined as follows:

**Definition 4.4.** For any \( \alpha_i \in S \) (the \( \alpha_i \) vertex is colored), \( s_{\alpha_i}D = D' \) is a new colored Dynkin diagram having the colors according to the sign change \( \epsilon_j' = \epsilon_j \epsilon_i^{-C_{j,i}} \) in Proposition 3.17 with the identification that \( R \) if the sign is \( −1 \), and \( B \) if it is \( +1 \). For example, in the case of \( \mathfrak{sl}(3, \mathbb{R}) \), \( s_{\alpha_1}(\circ_R - \circ_B) = \circ_R - \circ_R, s_{\alpha_1}(\circ_B - \circ_R) = \circ_B - \circ_R \).
We also define a $W_S$-action on the set $D(S) \times \{\pm 1\}$ of oriented colored Dynkin diagrams. If $\alpha_i \in S$, the action of $s_{\alpha_i}$ on the pair $(D, o)$ is given by $s_{\alpha_i}(D, o) = (s_{\alpha_i}D, (\epsilon_i)^{r_{\alpha_i}} o)$ where $r_{\alpha_i}$ is the number of elements in the set \{\alpha_j \in \Pi \setminus S : C_{j,i} \text{ is odd}\} and $\epsilon_i = \pm 1$ depending on the color of $\alpha_i$.

We confirm the Definition:

**Proposition 4.5.** Definition 4.4 above gives a well-defined action of $W_S$ on the set $D(S)$ of colored Dynkin diagrams with set of colored vertices $S$.

**Proof.** This follows from Proposition 3.16 and the correspondence $(S, \epsilon_\eta) \rightarrow h_\epsilon \eta$ giving a bijection between $D(S)$ and $D(\Delta^{\Pi \setminus S})$. □

**Corollary 4.6.** There is a bijection of $D(\Delta) \rightarrow H_R/H$ intertwining the $W$-actions on both sets.

**Proof.** The two sets $D(\Delta)$ and $H_R/H$ are clearly in bijective correspondence and an element $h_\epsilon$ corresponds to the coset $h_\epsilon H$. By Proposition 3.17, Notation 4.4 and Proposition 4.5, the actions on these two sets agree. They are both given by Part a) in Proposition 3.17. □

**Notation 4.7.** Motivated by Example 2.2 we consider the set $W \times W_S D(S)$ the $W$ translations of the set $D(S)$. We write the elements in this set as pairs $(w, D)$ and introduce an equivalence relation $\sim$ on these pairs, where for any $x \in W_S$, $(wx, D) \sim (w, XD)$. Denote the equivalence classes $[(w, D)]$. This set of equivalence classes is then in bijective correspondence with the set $D(S) \times W/W_S$ of pairs $(D, [w])$ with $[w] = [w]^{\Pi \setminus S} \in W/W_S$ and $D$ in $D(S)$. The correspondence is such that $[(w^*, D)]$ corresponds to $(D, [w^*])$ with $w^* \in [w]^{\Pi \setminus S}$ a minimal length representative of a coset in $W/W_S$.

We denote

$$D^k = \{(D, [w]^{\Pi \setminus S}) : S \subset \Pi, |S| = k, w \in W\},$$

which also parametrizes all the connected components of the Cartan subgroups of the form $H_R(w(\Delta^{\Pi \setminus S}))$.

**Remark 4.8.** For the reader who is not interested in the $W$-action or torsion, the object defined in Definition 4.9 becomes over $\mathbb{Q}$ the vector space with basis given by the full set of colored Dynkin diagrams having a fixed number of uncolored vertices. The boundary maps are obtained by *translating* around the $\partial_{j,c}$ defined for the antidominant chamber (see (14) below). Perhaps this boundary construction can be appreciated in Example 2.2 in Section 2 where portions of the boundary of a chamber are translated and some identifications take place. The tensor product notation below accomplishes the required boundary identifications algebraically.

**Definition 4.9** (The $\mathbb{Z}[W]$-modules $M(S)$). The full set of colored Dynkin diagrams is the set $D(S) \times W/W_S$ of all pairs $(D, [w]^{\Pi \setminus S})$. 
We can also define the full set of oriented colored Dynkin diagrams by considering $D(S) \times \{\pm 1\} \times W/W_S$ for different subsets $S \subseteq \Pi$. As $W$ sets these correspond to $W \times D(S) \times \{\pm 1\}$.

If we consider $D(S) \times \{\pm 1\}$ as imbedded in $Z[D(S)]$ by sending $(D,o)$ to $oD$, $o \in \{\pm 1\}$ we may consider a $W_S$-action on $\pm D(S)$, namely the action on oriented colored Dynkin diagrams. This produces a $Z[W]$-module $Z[W] \otimes Z[D(S)]$, and we denote this module by
\[ \mathcal{M}(S) = Z[W] \otimes_{Z[W_S]} Z[D(S)]. \]
Also denote by $\mathcal{M}_{l-k}$ the direct sum of all these modules over all sets $S$ with exactly $k$ elements,
\[ \mathcal{M}_{l-k} = \bigoplus_{|S|=k} \mathcal{M}(S). \]

**Remark 4.10.** Assume that $A$ is a $Z[W_S]$-module and $B$ is a $Z[W_{S'}]$-module with $S \subseteq S'$. In particular $B$ can be regarded as a $Z[W_S]$-module by restriction. Let $f : A \to B$ be a map intertwining these two $Z[W_S]$-module structures involved. Then, tensoring with $Z[W]$ we obtain a map of $Z[W]$-modules:
\[ F(f) : Z[W] \otimes_{Z[W_S]} A \to Z[W] \otimes_{Z[W_S]} B. \]
Since, in addition, $B$ is a $Z[W_{S'}]$-module where $S \subseteq S'$ then there is a second map:
\[ g : Z[W] \otimes_{Z[W_S]} B \to Z[W] \otimes_{Z[W_{S'}]} B. \]
Let $g \circ F(f) = T(f)$. This is then a map of $Z[W]$-modules
\[ T(f) : Z[W] \otimes_{Z[W_S]} A \to Z[W] \otimes_{Z[W_{S'}]} B. \]

We now give the boundary maps of $\mathcal{M}_{l-k}$. We regard $D(S) \times \{\pm 1\}$ as a $Z[W_S]$-module using Definition 4.4 the oriented case. We can view a pair $(D,o)$ instead as $\pm D \in Z[D(S)]$. This gives a $Z[W_S]$-module structure to each of the $Z$ modules involved on the domain of the map in (12) and a $W_{SU\{\alpha_i\}}$ to those on the co-domain. The map in (12) now has the form described in Remark 4.10. We apply the construction $T(\partial_{j,c})$ in Remark 4.10 to (12) and add over all possible subsets $S \subseteq \Pi$ with $|S| = k$ on the left side and with $|S| = k + 1$ on the right side. We then obtain the boundary maps,
\[ \partial_{l-k} : \mathcal{M}_{l-k} \to \mathcal{M}_{l-(k+1)}, \]
which are all given by

$$\partial_{l-k}(w^* \otimes X) = \sum_{j=1}^{l-k} \sum_{c=1}^{2} (-1)^j c + 1 w^* \otimes T(\partial_{j,c})X,$$

where $X \in \mathbb{D}(S)$ and $w^* \in [w]^\Pi \cap S$ is the minimum length representative in $W/W_S$ (see Notation 4.7). We thus have:

**Proposition 4.11.** The maps $\partial_{l-k}$ of (14) define a chain complex $M_\ast$ of $\mathbb{Z}[W]$-modules.

5. Cartan subgroups and Weyl group actions.

We here discuss relations between some Cartan subgroups of Levi factors, and verify the $W_S$-action on oriented colored Dynkin diagrams.

5.1. Cartan subgroups of Levi factors. Let $\text{Ad}^{\Delta A}$ denote the adjoint representation of the Lie subgroup $H_{\mathbb{R}}L_A$ of $G$ on the Lie algebra $\mathfrak{l}^A$. Note that we are deviating slightly from the standard convention and denoting by $\text{Ad}^{\Delta A}$ the representation of $H_{\mathbb{R}}L_A$ acting on the semisimple part of the Levi factor $\mathfrak{l}^A$, seen as a *quotient* of the group action on $\mathfrak{h}^A + \mathfrak{l}^A$ (thus dividing by the center of this Lie algebra which corresponds to a trivial representation summand). We will use this same notation $\text{Ad}^{\Delta A}$ when restricting to various Lie subgroups of $H_{\mathbb{R}}L_A$ containing $L_A$. The notation $\text{Ad}^{w(\Delta A)}$ with $w \in W$ refers to the similar construction with respect to $w(\Delta A)$. When $A = \Pi$ and $\Delta A = \emptyset$ then $\text{Ad}^{w(\Delta A)}$ will refer to the (trivial) one dimensional representation of $\{ e \}$.

**Definition 5.1.** Let $H_{\text{fund}}^A$ be defined by:

$$\exp \left( \left\{ \sum_{\alpha_i \notin A} c_i m_i^\alpha : c_i \in \mathbb{R} \right\} \right) = H_{\text{fund}}^A.$$

Note that usually $H^A \neq H_{\text{fund}}^A$.

Then we have:

**Proposition 5.2.** The images of $\text{Ad}^{\Delta A}$ on subsets of Cartan subgroups satisfy

1) $\text{Ad}^{\Delta A}(\mathcal{D}(\Delta)) = \mathcal{D}(\Delta A)$,
2) $\text{Ad}^{\Delta A}(H_{\mathbb{R}}^A) = H_{\mathbb{R}}(\Delta A)$,
3) $\text{Ad}^{\Delta A}(H_{\text{fund}}^A) = \text{Ad}^{\Delta A}(H^A)$.

**Proof.** We first point out that the exponential map in the Lie groups $L_{\mathbb{R}}^A$, $\text{Ad}^{\Delta A}$ gives a diffeomorphism between $\mathfrak{h}^A$ and the corresponding connected
Lie group. Thus the Lie groups $H^A$, $\text{Ad}^A(H^A)$ are isomorphic. This takes care of Part 2) on the level of the connected component of the identity.

Also the image under $\text{ad}^A$ of the set $\{m^\alpha_{\alpha_i} : \alpha_i \notin A\}$ gives rise to a basis of the Cartan subalgebra of $\text{ad}^A(t^A)$. Thus exponentiating we obtain Part 3).

Recall that $\chi_{\alpha_i}(h_j) = \exp(\pi \sqrt{-1} \delta_{i,j})$ and $\text{Ad}^A(h) = I$ if and only if $\chi_{\alpha_i}(h) = 1$ for all $\alpha_i \in \Pi \setminus A$. We have thus $\text{Ad}^A(h_i) = I$ for $\alpha_i \in A$ and $\{\text{Ad}^A(h_e) : e \in \mathcal{E}\} = \{\text{Ad}^A(h_e) : e \in \mathcal{E}^A\} = \mathcal{D}(\Delta^A)$. The elements $h_i$ with $\alpha_i \in A$ are in the center of $H^A$. This proves Part 1).

For Part 2), we proceed by noting that by definition of $H^A$ it must be generated by the Lie group $\text{Ad}^A(H^A)$ with Lie algebra $h^A$ and the elements $\text{Ad}(h_i)$ with $\alpha_i \in \Pi \setminus A$ (playing the role that the $h_i$ play in the definition of $H^A$). This is the same as $\text{Ad}^A(H^A)$. $\square$

### 5.2. Action of the Weyl group on oriented colored Dynkin diagrams.

**Proposition 5.3.** The action of $W_S$ on $\mathbb{D}(S) \times \{\pm 1\}$ in Definition 4.4 is well-defined.

**Proof.** Recall that $W$ is a Coxeter group, (Proposition 3.13 [12]) and it thus has defining relations $s_{\alpha_i}^2 = e$ and $(s_{\alpha_i}, s_{\alpha_j})^{m_{ij}} = e$ where $m_{ij} = 2, 3, 4, 6$ depending on the number of lines joining $\alpha_i$ and $\alpha_j$ in the Dynkin diagram. The case $m_{ij} = 2$ occurring when $\alpha_i$ and $\alpha_j$ are not connected in the Dynkin diagram. The only relevant cases then are when $\alpha_i, \alpha_j$ are both colored and connected in the Dynkin diagram. The only relevant vertices in the Dynkin diagram are those connected with these two and which are uncolored. We are thus reduced to very few nontrivial possibilities: $D_5, A_4, F_4, B_4, C_4$; smaller rank cases being very easy cases. We verify only one of these cases, the others being almost identical with no additional difficulties. Consider the case of $D_5$ where $\alpha_i = \alpha, \alpha_j = \beta$ are the two simple roots which are not “endpoints” in the Dynkin diagram and all the others are uncolored. Then $r_\alpha = 1$ but $r_\beta = 2$. For book-keeping purposes it is convenient to temporarily represent red as $-1$ and blue as $1$ and simply follow what happens to these two roots and an orientation $o = 1$. Hence all the information can be encoded in a triple $(\epsilon_1, \epsilon_2, o)$ representing the colors of these two roots and the orientation $o$. Now we apply $(s_{\alpha}, s_{\beta})^3$ to $(\epsilon_1, \epsilon_2, o)$:

\[
(\epsilon_1, \epsilon_2, o = 1) \xrightarrow{s_{\beta}} (\epsilon_1 \epsilon_2, \epsilon_2, (\epsilon_2)^2 o = 1) \xrightarrow{s_{\alpha}}
(\epsilon_1 \epsilon_2, \epsilon_1, (\epsilon_1 \epsilon_2)^1 o) \xrightarrow{s_{\beta}} (\epsilon_2, \epsilon_1, (\epsilon_1)^2 \epsilon_1 \epsilon_2 o = \epsilon_1 \epsilon_2 o) \xrightarrow{s_{\alpha}}
(\epsilon_2, \epsilon_1 \epsilon_2, (\epsilon_2)^1 \epsilon_1 \epsilon_2 o = \epsilon_1 o) \xrightarrow{s_{\beta}} (\epsilon_1, \epsilon_1 \epsilon_2, (\epsilon_1 \epsilon_2)^2 \epsilon_1 o = \epsilon_1 o) \xrightarrow{s_{\alpha}}
(\epsilon_1, \epsilon_2, o).
\]

$\square$
5.3. The set $H^o_R$. By Proposition 5.2, if $\chi_{\alpha_i}^{\Delta^A}, \alpha_i \in \Pi \setminus A$ is a root character of $H_R(\Delta^A)$ acting on $l^A$, then it is related to $\chi_{\alpha_i}$ by $\chi_{\alpha_i}^{\Delta^A} \circ \text{Ad}^{\Delta^A} = \chi_{\alpha_i}$. The map $\chi_{\alpha_i}^{\Delta^A}, \alpha_i \in \Pi \setminus A$ provides the local coordinates on $H^o_R$ which consists of the Cartan subgroups $H_R(\Delta^A)$.

**Definition 5.4.** Let $H^o_R$ be defined as

$$H^o_R = \bigcup_{A \subset \Pi} H_R(\Delta^A) \times \{[e]^A\}.$$

We then define a map $\phi_e : H^o_R \to \mathbb{R}^l$ as follows:

$$\phi_e(h, [e]^A) = (\phi_{e,1}(h), \ldots, \phi_{e,l}(h)),$$

where $\phi_{e,i}(h) = \chi_{\alpha_i}^{\Delta^A}(h)$ whenever $\alpha_i \notin A$ and $\phi_{e,i}(h) = 0$ if $\alpha_i \in A$. Denote $\phi_e^A$ the restriction of $\phi_e$ to $H_R(\Delta^A) \times \{[e]^A\}$.

By Proposition 5.2 Part 2) we can compose with $\text{Ad}^{\Delta^A} \times 1$ and re-write the domain of $\phi_e \circ (\text{Ad}^{\Delta^A} \times 1)$ as $\bigcup_{A \subset \Pi} H^A_R \times \{[e]^A\}$. We also define

$$\phi_w = (\phi_{w,1}, \ldots, \phi_{w,l}) : w(H^o_R) \to \mathbb{R}^l,$$

with

$$w(H^o_R) := \bigcup_{A \subset \Pi} H_R(w(\Delta^A)) \times \{[w]^A\},$$

by setting

$$\phi_{w,i}(x, [w]^A) = \chi_{w\alpha_i}^{w(\Delta^A)}(\text{Ad}^{w(\Delta^A)}(wh)) = \chi_{w\alpha_i}(wh) = \chi_{\alpha_i}(h),$$

where $x = \text{Ad}^{w(\Delta^A)}(wh) \in H_R(w(\Delta^A))$ with $h \in H_R$, if $\alpha_i \notin A$. For the case when $\alpha_i \in A$, we set

$$\phi_{w,i}(x, [w]^A) = 0.$$

5.4. The sets $H^A_R$. We here note an isomorphism between several presentations of a split Cartan subgroup of a Levi factor.

**Proposition 5.5.** All the following are isomorphic as Lie groups:

1) $H^A_R$,
2) $\text{Ad}^{\Delta^A}(H_R)$,
3) $\text{Ad}^{\Delta^A}(H^A_R)$,
4) $H_R(\Delta^A)$.

**Proof.** The groups in 3) and 4) are isomorphic by Proposition 5.2. We can write $h_R = z + h^A_R$ where $z$ is defined by $z \in z$ if and only if $\langle \alpha_i, z \rangle = 0$ for all $\alpha_i \in \Pi \setminus A$. 
Now the group $\exp(z)$ is the center of $H^A$. Intersecting with $\tilde{G}$ we obtain the center of $H^A$. To find this intersection with $\tilde{G}$, we must find all $x + \sqrt{-1}y \in z$ such that $e^{(\alpha_i,x+\sqrt{-1}y)}$ is real for all $\alpha_i \in \Pi$ (see the Proof of Lemma 3.7). We find that $x \in z \cap h$ and that $y$ is an integral linear combination of the $y_i$. As in the Proof of Proposition 5.2 the elements $\exp(\sqrt{-1}y_i)$ which are in the center of $H^A$ are those for which $\alpha_i \in A$. The center of $H^A$ is then the group generated by $h_i$ with $\alpha_i \in A$ and $\exp(z \cap h)$. Since $\exp(z \cap h)$ is isomorphic to $D^A$, the isomorphism between 2), 3) and 4) follows.

6. The set $\hat{H}_R$.

We here define our main object $\hat{H}_R$ as a union of Cartan subgroups of Levi factors and their $W$-translations.

Let $\hat{W}$ be the disjoint union of all the quotients $W/W_{\Pi \setminus A}$ over $A \subset \Pi$. Each of the elements $[w]^A \in W/W_{\Pi \setminus A}$ parametrizes a parabolic subgroup. First $[e]^A$ corresponds to a standard parabolic subgroup of $G_C$ (Proposition 7.76 or Proposition 5.90 of [13]). This is just the parabolic subgroup determined by the subset $\Pi \setminus A$ of $\Pi$. Then we translate such a parabolic subgroup with $w \in W$. The resulting parabolic subgroup corresponds to $[w]^A$. Thus $\hat{W}$ is the set of all the $W$-translations of standard parabolic subgroups.

**Definition 6.1.** We define

$$\hat{H}_R = \bigcup_{A \subset \Pi} \bigcup_{w \in W} H_R(w(\Delta^A)) \times \{[w]^A\}. \quad (16)$$

From Proposition 5.5, $H^A_R$ can be replaced by $H_R(\Delta^A)$ or $\text{Ad}^{w(\Delta^A)}(H_R)$. Thus we can alternatively write $\hat{H}_R$ as a subset of $H_R \times \hat{W}$. We recall $w^{\bullet,A}$ or just $w^{\bullet}$ as in Notation 4.7. We then have:

$$\hat{H}_R = \bigcup_{A \subset \Pi} \bigcup_{w^{\bullet,A} \in W/W_{\Pi \setminus A}} w^{\bullet,A}(H^A_R) \times \{[w]^A\}. \quad (17)$$

Also fixing an isomorphism $\xi_w : H_R(\Delta^A) \rightarrow H_R(w(\Delta^A))$ inducing a set bijection (by composition) $\xi_w^* : w(\Pi) \rightarrow \Pi$, we have

$$\hat{H}_R \cong \bigcup_{A \subset \Pi} \bigcup_{w^{\bullet,A} \in W/W_{\Pi \setminus A}} H_R(\Delta^A),$$

where an element $(w^{\bullet,A},h)$ on the right-hand side is sent to $(\xi_w(h),[w]^A)$. This endows $\hat{H}_R$ with a $W$-action.
In what follows it is useful to think of a colored Dynkin diagram (e.g., \( R-\circ R \)) as parametrizing a “box” (e.g., \([-1, 1]\)). We need to further subdivide this box into \(2^l\) smaller boxes by dividing it into regions according to the sign of each of the coordinates (e.g., \([-1, 0] [0, 1]\)). We also need to consider the boundary between these \(2^l\) regions (e.g., \([-1, 0] [0, 1]\)). We will then do the same thing with a colored Dynkin diagram of the form \((D, [w]_{\Pi, S})\) by assigning labels in \(\{\pm 1, 0\}\) to the vertices in \(\Pi \setminus S\).

The main purpose of the following is to associate certain sets in \(\hat{H}_R\) to the colored Dynkin diagrams, the signed-colored Dynkin diagrams and the sets associated to them only play an auxiliary role.

We introduce the following notation. This notation is illustrated in Example in Section 2 and Figure 3.

**Notation 6.2.** A signed-colored Dynkin diagram \(\tilde{D}\) is a Dynkin diagram with some vertices colored \(R\) or \(B\) and the remaining vertices labeled +, −, or 0. The followings are auxiliary objects to keep track of signs, zeros and colors.

- \(\tilde{\eta} : \Pi \to \{\pm 1, 0\}\) function which agrees with \(\eta\) on \(S\) and determines the sign labels in \(\tilde{D}\).
- \(\tilde{A} = A(\tilde{D}) = A(\tilde{\eta}) = \{\alpha_i \in \Pi : \tilde{\eta}(\alpha_i) = 0\}\).
- \(K(\eta)\) the set of all \(\tilde{\eta} : \Pi \to \{\pm 1, 0\}\) which agree with \(\eta\) in \(S\).
- \(\epsilon_{\tilde{\eta}} \in \mathcal{E}^A\) the element which agrees with \(\tilde{\eta}\) on \(\Pi \setminus A\) \((A = A(\tilde{D}))\).
- \(\tilde{D}(S, A, \epsilon_{\tilde{\eta}}, [w]_{\Pi, S})\) the unique signed-colored Dynkin diagram attached to a colored Dynkin diagram \((D, [w]_{\Pi, S})\) or to \((S, \eta, [w]_{\Pi, S})\), where the vertices in \(\Pi \setminus S\) are assigned a label in \(\{\pm 1, 0\}\).

We now associate a subset of \(\hat{H}_R\) to a signed-colored Dynkin diagram with \(S = \emptyset\). Notice that in this case \(\epsilon_{\tilde{\eta}}\) can be any element of \(\mathcal{E}^A\). Recall that \(\epsilon_{\tilde{\eta}}(\Delta^A)\) is then the element of \(\mathcal{E}(\Delta^A)\) that corresponds.

We associate to a signed colored Dynkin diagram \((\emptyset, A, \epsilon_{\tilde{\eta}}, [e]_{\Pi, S})\) two sets, one includes walls and the other doesn’t (in order to avoid duplicate notation we denote the signed-colored Dynkin diagram and the set with the same notation):

\[
(\emptyset, A, \epsilon_{\tilde{\eta}}, [e]_{\Pi, S})^\leq = H(\Delta^A)^{\leq}_{\epsilon_{\tilde{\eta}}(\Delta^A)} \times \{[e]^A\}.
\]

When \(A = \emptyset\) (no zeros), these are the \(2^l\) boxes in the antidominant chamber. We define the second related set as:

\[
(\emptyset, A, \epsilon_{\tilde{\eta}}, [e]_{\Pi, S})^< = H(\Delta^A)^{<}_{\epsilon_{\tilde{\eta}}(\Delta^A)} \times \{[e]^A\}.
\]

The chamber walls of the antidominant chamber of the Cartan subgroup are defined as:
Proposition 6.4. We have

- \( D(\alpha_i, A, e)^\leq = \{ h \in H(\Delta^A)^\leq_{\epsilon(\Delta^A)} : |\chi_{\alpha_i}(h)| = 1 \} \) (the \( \alpha_i \)-wall),
- \( D(\alpha_i, A, e)^< = D(\alpha_i, A, e)^\leq \cap \{ h \in H(\Delta^A)^\leq_{\epsilon(\Delta^A)} : |\chi_{\alpha_j}(h)| < 1 \text{ if } j \neq i, \alpha_j \in \Pi \setminus A \} \).

We next consider the case of \( S = \{ \alpha_i \} \not\subset A \) and then the general case of any \( S \subset \Pi \) with \( A \subset \Pi \setminus S \) and any \( \epsilon_\eta \in \mathcal{E}^A \). This defines the walls for Levi factor pieces corresponding to subsystems of the Toda lattice (we here list open walls):

- \( \{ \{\alpha_i\}, A, \epsilon_\eta \} = D(\alpha_i, A, \epsilon_\eta)^\leq \times \{[e]^A\} \).
- \( (S, A, \epsilon_\eta) = \bigcap_{\alpha_i \in S} (\{\alpha_i\}, A, \epsilon_\eta) \).

We now associate a set in \( \hat{H}_R \) to a colored Dynkin diagram. We define a set denoted \( D = (S, \epsilon_\eta) \) as follows: For \( S \neq \emptyset \) (so that there is an \( \eta : S \to \{\pm 1\} \)),

- \( (S, \epsilon_\eta) = \bigcup_{\eta \in K(\eta)} (S, A(\eta), \epsilon_\eta), \text{ and if } S = \emptyset, (\emptyset, \epsilon_\eta) = \bigcup_{\eta \in \{\pm 1, 0\}} (\emptyset, \epsilon_\eta) \).

Here \( \epsilon_\eta = (1, \ldots, 1) \), and see (7) and (9) in Section 2.

We consider the \( w \)-translations of the colored Dynkin diagrams: For this write \( w \) uniquely as \( w = w^* \cdot w_\bullet \) with \( w_\bullet \in W_S \).

- \( (w^* \cdot D, [w]^\Pi \setminus S) = w((D, [e]^\Pi \setminus S)) \).

In order to define the \( W \)-translations of sets associated to signed-colored Dynkin diagrams we have to extend the definition of the \( W_S \)-action to the signed-colored Dynkin diagrams. The definition is exactly the same if we treat the label – as if it were an \( R \) and the label + as if it were a \( B \) as in Definition 4.4. Then consider \( (w^* \cdot D, [w]^\Pi \setminus S) \) and let:

- \( (w^* \cdot \bar{D}, [w]^\Pi \setminus S) = w((\bar{D}, [e]^\Pi \setminus S)) \).

Remark 6.3. We refer to the set \( (\emptyset, \epsilon_\eta) \subset \hat{H}_R \) as the antidominant chamber of \( \hat{H}_R \). Recall Notation 4.7 \( w^* \cdot A \in [w]^A \). The following justifies our definition of the “antidominant chamber” of \( \hat{H}_R \) using (17) in the definition of \( \hat{H}_R \).

Proposition 6.4. We have

\begin{equation}
\hat{H}_R = \bigcup_{A \subset \Pi} \bigcup_{w^* \cdot A \in W} \bigcup_{\sigma \in \mathcal{W}_{\Pi \setminus A}} w^* \cdot A \left( \sigma \left( H^A_{\epsilon(\sigma)} \right) \right) \times \{[w]^A\}.
\end{equation}

Proof. We set \( w^* = w^* \cdot A \) for simplicity. By Proposition 3.17 Part b), we have that each \( H^A_\sigma \) can be written as a union over \( \sigma \in \mathcal{W}_{\Pi \setminus A} \) of sets of the form \( \sigma \left( H^A_{\epsilon(\sigma)} \right) \) and thus by the definition (17) of \( \hat{H}_R \), we conclude the statement. \( \square \)
7. Colored Dynkin diagrams and the corresponding cells.

Here we consider the manifold structure and the topology of \( \hat{H}_R \) as the union of cells parametrized by colored Dynkin diagrams.

7.1. Action of the Weyl group on the sets \((S, \epsilon_\eta)\). Let \( \mathcal{M}^{\text{geo}} \) be the complex with the sets \((D, [w]_{\Pi \backslash S})\). We then consider the action of \( W_S \) on the union of all the sets \( D = (S, \epsilon_\eta) \) having a fixed nonempty set \( S \) of colored vertices and endowed with an orientation \( o \). Similarly we can endow each of the terms in the chain complex \( \mathcal{M}^{\text{geo}} \) with a action of \( W \) by translating the sets corresponding to colored Dynkin diagrams with the \( W \)-action and taking into account changes of orientation induced on the oriented boxes. This new action of \( W \) on \( \mathcal{M}^{\text{geo}} \) could in principle be different from the \( W \)-action on the chain complex \( \mathcal{M}_* \) (see Proposition 4.11).

We now become more explicit about the \( W \)-action that was just introduced on \( \mathcal{M}^{\text{geo}} \): Note that since \( \chi_{\alpha_{ji}}(h) = \pm 1 \) for any \( \alpha_{ji} \in S \) and \( h \in (S, \epsilon_\eta) \) (Notation 6.2), if \( h \in (S, \epsilon_\eta) \) then \( s_{\alpha_{ji}} \chi_{\alpha_{ji}}(h) = \chi_{\alpha_{ji}}(h) \chi_{\alpha_{ji}}(h) ^{-C_{j,ji}} = \pm \chi_{\alpha_{ji}}(h) \). Hence, in terms of the coordinates given by \( \phi_e \), the action of \( s_{\alpha_{ji}} \) is given by a diagonal matrix whose nonzero entries are \( \pm 1 \). This matrix has some entries corresponding to the set \( S \) and other entries corresponding to \( \Pi \backslash S \). The entries corresponding to the set \( S \) change by a sign as described in Proposition 3.16. The statement in Proposition 3.16 just means that the set \((S, \epsilon_\eta)\) is sent to \((S, \epsilon_\eta')\) with \( \eta' = (\eta_{j_1}, \ldots, \eta_{j_s}) \), and \( \eta'_{ji} = \eta_{ji} (-1)^{C_{j,ji}} \).

The determinant of the diagonal submatrix corresponding to elements in \( \Pi \backslash S \) is the sign \( (\eta_i)^r \) with (see Definition 4.4)

\[
r = \left| \{ \alpha_j \in \Pi \backslash S : C_{j,ji} \text{ is odd} \} \right|.
\]

Consider the set \((S, \epsilon_\eta)\) endowed with a fixed orientation \( \omega \) corresponding to \( o = 1 \). Then \( s_{\alpha_i} \) with \( \alpha_i \in S \) sends \((S, \epsilon_\eta)\) to \((S, \epsilon_{\eta'})\) endowed with the orientation \((\eta_i)^r \omega\). We now consider the \( \mathbb{Z} \)-module of formal integral combinations of the sets \((S, \epsilon_\eta)\),

\[
\mathbb{Z} \left[ (S, \epsilon_\eta) : S \subset \Pi, \ |S| = k, \ \epsilon_\eta \in E^{\Pi \backslash S} \right].
\]

We keep track of the orientation by putting a sign \( \pm \) in front of \((S, \epsilon_\eta)\). This \( \mathbb{Z} \)-module acquires a \( \mathbb{Z}[W_S] \)-action that corresponds to the abstract construction given in Definition 4.9 with colored Dynkin diagrams. By considering all the \( W \)-translations of the \((S, \epsilon_\eta)\) we generate the module denoted above by \( \mathcal{M}(S) \). The direct sum of all these \( \mathcal{M}(S) \) over \( |S| = k \) is denoted \( \mathcal{M}_{l-k} \) (see Definition 4.9). Hence there is no difference as \( W \)-modules between \( \mathcal{M}^{\text{geo}} \) and \( \mathcal{M}_* \).

We can now summarize this discussion in the following:

**Proposition 7.1.** The action of \( W \) on \( \mathcal{M}_* \) (Subsection 4.3) and the action of \( W \) on \( \mathcal{M}^{\text{geo}} \) are isomorphic.
Proof. We start with the last statement, that we can regard \( H \) in the adjoint group (Remark 3.14). From Proposition 5.2 Part 2) it follows that we have defined \( \chi \in \Delta^A \) between the images \( H \) and \( R \). We will use these maps to give \( \hat{H}_R \) coordinate charts leading to a manifold structure. We then have the following three Propositions:

**Proposition 7.2.** The image \( \phi_e(H_R(\Delta^A)) \times \{ [e]^A \} \) consists of all \((t_1, \ldots, t_l) \in \mathbb{R}^l\) such that \( t_i \neq 0 \) if and only if \( \alpha_i \notin A \). The map \( \phi_e \) is a bijection between \( H_R = \bigcup_{A \in \Pi} (H_R(\Delta^A)) \times \{ [e]^A \} \) and \( \mathbb{R}^l \).

Proof. We start with the last statement, that \( \phi_e \) is a bijection; we have that \( \phi_e \) is injective because the scalars \( \chi^{\Delta^A}_{\alpha_1}(h), \ldots, \chi^{\Delta^A}_{\alpha_{m}}(h) \) determine all the root characters \( \chi^\phi(\hat{A}_e(h), \phi) \in \Delta^A \) for \( h \in H_R(\Delta^A) \) and these scalars determine \( h \) in the adjoint group (Remark 3.14). From Proposition 5.2 Part 2) it follows that we can regard \( H_R(\Delta^A) \) as \( \text{Ad}^A(H_R) \) (see (15)). We prove surjectivity by proving first the statement concerning the image \( \phi_e \circ (\text{Ad}^A \times 1)(H_A \times \{ [e]^A \}) \). When all the sets \( A \) are considered then all of \( \mathbb{R}^l \) will be seen to be in the image of \( \phi_e \). First consider \( h_e h \) with \( h = \exp \left( \sum_{\alpha_i \in A} c_i m^{\alpha_i}_0 \right) \) and \( \epsilon \in \mathbb{E}^A \).

We now apply \( \phi_e \circ (\text{Ad}^A \times 1) \). Since \( \left( \alpha_i, \sum_{\alpha_j \notin A} c_j m^{\alpha_j}_0 \right) = c_i \left( \frac{\alpha_i, \alpha_i}{2} \right) \) we obtain, by exponentiating, \( \chi_{\alpha_i}(h_e h) = e^{\epsilon c_i m^{\alpha_i}_0} \). The set \( \phi_e(H_R(\Delta^A) \times \{ [e]^A \}) \) becomes the image of the map: \( \mathbb{R}^l \to \mathbb{R}^l \) given by first defining a map that sends \((\epsilon t_1, \ldots, \epsilon t_l) \to (f_1, \ldots, f_l) \) with \( f_i = \epsilon t_i \left( \frac{\alpha_i, \alpha_i}{2} \right) \) for \( t_i > 0 \). This map is modified so that whenever \( \alpha_i \in A \) then the \( i \)-th coordinate is replaced with 0. We denote this modified map by \( F^A \). The domain and the image of \( F^A \) therefore consists of the set \( \{(s_1, \ldots, s_l) : s_i = 0 \text{ if } \alpha_i \notin A \} \).

Together all these \( F^A \) give rise to one single map \( F : \mathbb{R}^l \to \mathbb{R}^l \) which is surjective. \( \square \)

**Proposition 7.3.** The image \( \phi_w(H_R(w(\Delta^A)) \times \{ [w]^A \}) \) consists of all \((t_1, \ldots, t_l) \in \mathbb{R}^l\) such that \( t_i \neq 0 \) if and only if \( \alpha_i \notin A \).

Proof. This follows from Proposition 6.4 and the fact that \( \phi_{w_i}(w(h), A) = \chi_{w\alpha_i}(w(h)) = \chi_{\alpha_i}(h) \) for \( \alpha_i \notin A \). \( \square \)

**Proposition 7.4.** The image \( \phi_w((\emptyset, \epsilon_o)) \) consists of all \((t_1, \ldots, t_l) \in \mathbb{R}^l\) such that \(-1 < t_i < 1\). The sets \( \phi_w((S, \epsilon, [w]^{[\Pi,S]})) \) as \( S \subset \Pi, S \neq \emptyset \) varies,
give a cell decomposition of the boundary of the box $[-1,1]^l$. In particular, the sets $(S,\epsilon, [w]^{\Pi\backslash S})$ give a cell decomposition of the smooth manifold $\hat{H}_R$.

Proof. This follows from Proposition 7.3 but is better understood in Example 2.2. We omit details. □

Remark 7.5. There is a more convenient cell decomposition of $\hat{H}_R$ for the purpose of calculating homology explicitly. The only change is that the $l$ dimensional cell becomes the union of all the $l$-cells together with all the (internal) boundaries corresponding to colored Dynkin diagrams where all the colored vertices are colored $B$. This is the set:

$$\hat{H}_R \setminus \bigcup_{S \subset \Pi, w \in W, \eta \text{ such that } \eta(\alpha_i) = -1 \text{ for some } \alpha_i \in S} (S,\epsilon, [w]^{\Pi\backslash S}).$$

This set can be seen to be homeomorphic to $\mathbb{R}^l$. With this cell decomposition there is exactly one $l$ cell; and the other lower dimensional cells correspond to colored Dynkin diagrams which are parametrized by pairs $(D, [w]^{\Pi\backslash S})$, such that, at least one vertex of $D$ has been colored $R$. In terms of the $(S,\epsilon, [w]^{\Pi\backslash S})$, the top cell would instead be defined to consist of $\hat{H}_R$. (See Figure 6 for this remark.)

The big cell in this decomposition with a fixed orientation corresponds to the element $c_l = \sum_{w \in W} (-1)^{l(w)}(D,\epsilon, w)$ $(S = \emptyset)$. This element satisfies $\partial_l(c_l) = 2(c_{l-1})$ for some $c_{l-1} \neq 0$ (except in the case of type $A_1$). Note that $(D,w)$ and $(D,ws\alpha_i)$, with $S = \emptyset$ and $l(ws\alpha_i) = l(w) + 1$, appear with opposite signs in $c_l$. When $\partial_l$ is applied and the $\alpha_i$ is colored $R$ the sign $(-1)^{r_{\alpha_i}}$ in Definition 4.4 makes these two terms contribute as $2(D', [w]^{\Pi\backslash \{\alpha_i\}})$ with $D'$ the new colored Dynkin diagram obtained. The terms from the boundary of $\partial_l$ obtained by coloring $\alpha_i$ with $B$ will cancel since the action of $s_{\alpha_i}$ on colored Dynkin diagrams is trivial when $\alpha_i$ is colored $B$. The case of $A_1$ is an exception because, in that case, once $\alpha_1$ is colored $R$ no more uncolored vertices remain. The set $\{\alpha_j \in \Pi \setminus \alpha_i : C_{j,i} \text{ is odd}\}$ is empty and $r_{\alpha_i} = 0$. Thus there is cancellation in this case.

7.3. Topology on $\hat{H}_R$, coordinate charts, integral homology. We define a topology on $\hat{H}_R$ in which $U \subset \bigcup_{A \subset \Pi} H(w(\Delta^A)) \times \{[w]^A\}$ is open if and only if $\phi_w(U)$ is open in the usual topology of $\mathbb{R}^l$. The maps $\phi_w$ become coordinate charts and since the compositions $\phi_w \circ \phi^{-1}$ are $C^\infty$ on their domain, then $\hat{H}_R$ acquires the structure of a smooth manifold. The $W$-action becomes a smooth action.

Definition 7.6 (Filtration of $\hat{H}_R$ and the chain complex $\mathcal{M}^{CW}_*$). We construct a filtration of the topological space $\hat{H}_R$ in the sense of [18] p. 222.
Let $X_{l-k}$ denote the union of all the sets of the form $(D, [w]^{\Pi \setminus S})$ over all $w \in W$ and $S \in P(\Pi)$ such that $|S| \geq k$. This is a closed set and $X_r \setminus X_{r-1}$ is a union of sets of the form $(D, [w]^{\Pi \setminus S})$ with $|S| = l - r$. The filtration $X_r = 0$, $1, \ldots, l$ satisfies the conditions of Theorem 39.4 in [18].

We define a chain complex $\mathcal{M}_s^{CW}$ with boundary operators as in [18]

$$\partial_r : H_r(X_r, X_{r-1}, \mathbb{Z}) \rightarrow H_{r-1}(X_{r-1}, X_{r-2}, \mathbb{Z}).$$

**Proposition 7.7.** The smooth manifold $\tilde{H}_R$ is compact, nonorientable (except if $g$ is of type $A_1$). The homology of the chain complex $\mathcal{M}_s, H^k(\mathcal{M}_s)$ is isomorphic as a $\mathbb{Z}[W]$ module to $H^k(\tilde{H}_R, \mathbb{Z})$.

**Proof.** The manifold $\tilde{H}_R$ is the finite union of the chambers as in Proposition 6.4. Since the $W$-action on $\tilde{H}_R$ is by continuous transformations, it then suffices to observe that the antidominant chamber is compact. The antidominant chamber $(\emptyset, e_o)$ of Definition 6.2 can be seen to be compact by describing explicitly its image under $\phi_e$. This image is a “box” inside $\mathbb{R}^l$, as can be seen in Propositions 7.2, 7.4 namely the set \{$(t_1, \ldots, t_l) : -1 \leq t_i \leq 1$\}. The space $\tilde{H}_R$ is now the finite union of the $W$-translates of this compact set. That the boundary operators of $\mathcal{M}_s^{CW}$ agree with the boundary operators of $\mathcal{M}_s$ will follow from the fact that in the $\phi_w$ coordinates the $(D, [w])$ is a “box” which is itself part of the boundary of a bigger “box” (Proposition 7.4). We start with the set \{$(t_1, \ldots, t_l) : -1 \leq t_i \leq 1$\} and note that its boundary is combinatorially described by (12) or (13). Note that $(D, w)$ (with $S = \emptyset$, $[w]^{\Pi} = w$) represents the open box. The faces are parametrized by coloring each of the $l$ vertices $R$ or $B$ which then represent opposite faces in the boundary. The signs are just chosen so that $\partial_{k-1} \circ \partial_k = 0$ for $k = 1, \ldots, l$. This description may be best understood by working out Example 2.2.

All the cells thus appear by taking the faces of a box $[-1, 1]^l$ and then faces of faces etc. By the same process of coloring uncolored vertices $R$ or $B$ which give rise, each time, to a pair of opposite faces in a box. In each case (12) or (13) correctly describe the process of taking the boundary of a box. Note that it is enough to study what happens when $w = e$ and then consider the $W$-translates.

We now use Theorem 39.4 of [18] to conclude that $\mathcal{M}_s^{CW}$ computes integral homology. However each $\mathbb{Z}[W]$-module appearing in $\mathcal{M}_s^{CW}$ in a fixed degree, can easily be seen to be identical with the corresponding term in $\mathcal{M}_s$. By Proposition 7.1 and the agreement of the boundary operators, we obtain that $\mathcal{M}_s$ computes integral homology. The nonorientability follows if we use the second cell decomposition described in Remark 7.5. The unique top cell then has a nonzero boundary (except in the case of $g = sl(2, \mathbb{R}))$. □
8. Toda lattice and the manifold $\hat{H}_\mathbb{R}$.

We now associate the manifold $\hat{H}_\mathbb{R}$ with the Toda lattice by extending the results of Kostant in [16]. We start with the definition of the variety $Z_\mathbb{R}$ of Jacobi elements on $\mathfrak{g}$ where the Toda lattice is defined.

8.1. The variety $Z_\mathbb{R}$ and isotropy group $\tilde{G}^\mathbb{R}$.

Definition 8.1 (Varieties of Jacobi elements). Let $S(\mathfrak{g})$ be the symmetric algebra of $\mathfrak{g}$. We may regard $S(\mathfrak{g})$ as the algebra of polynomial functions on the dual $\mathfrak{g}'$.

If we consider the algebra of $G$-invariants of $S(\mathfrak{g})$, then by Chevalley’s theorem there are homogeneous polynomials $I_1, \ldots, I_l$ in $S(\mathfrak{g})^G$ which are algebraically independent and which generate $S(\mathfrak{g})^G$. Thus $S(\mathfrak{g})^G$ can be expressed as $\mathbb{R}[I_1, \ldots, I_l]$.

For $F = \mathbb{C}$ or $F = \mathbb{R}$, we consider the variety $Z_F$ of normalized Jacobi elements of $\mathfrak{g}_F$. Our notation, however, is slightly different from the notation of [16] in the roles of $e_{\alpha_i}$ and $e_{-\alpha_i}$. Thus we let

$$J_F = \left\{ X = x + \sum_{i=1}^l (b_i e_{-\alpha_i} + e_{\alpha_i}) : x \in \mathfrak{h}, b_i \in F \setminus \{0\} \right\},$$

$$Z_F = \left\{ X = x + \sum_{i=1}^l (b_i e_{-\alpha_i} + e_{\alpha_i}) : x \in \mathfrak{h}, b_i \in F \setminus \{0\}, X \in S(F) \right\}.$$

We also allow subsystems which correspond to the cases having some $b_i = 0$:

$$\overset{\circ}{J}_F = \left\{ X = x + \sum_{i=1}^l (b_i e_{-\alpha_i} + e_{\alpha_i}) : x \in \mathfrak{h}, b_i \in F \right\},$$

$$\overset{\circ}{Z}_F = \left\{ X = x + \sum_{i=1}^l (b_i e_{-\alpha_i} + e_{\alpha_i}) : x \in \mathfrak{h}, b_i \in F, X \in S(F) \right\}.$$

Kostant defines in [16] p. 218 a real manifold $Z$ by considering all elements $x + \sum_{i=1}^l (b_i e_{-\alpha_i} + e_{\alpha_i})$ in $Z_\mathbb{R}$ which in addition satisfy $b_i > 0$. We are departing in a crucial way from [16] by allowing the $b_i$ to be negative or even zero when $A \neq \emptyset$. This extension gives the indefinite Toda lattices introduced in [14]. We let for any $\epsilon \in \mathcal{E}$,

$$Z_\epsilon = \left\{ X = x + \sum_{i=1}^l (b_i e_{-\alpha_i} + e_{\alpha_i}) : \epsilon_i b_i > 0, X \in S(\mathbb{R}) \right\}.$$
This is a set of real normalized Jacobi elements, that is, the elements of $Z_{\mathbb{R}}$. Thus the union of all the $Z_\epsilon$ is all of $Z_{\mathbb{R}}$, $Z_{\mathbb{R}} = \bigcup_{\epsilon \in \mathcal{E}} Z_\epsilon$.

The elements in $Z_{\mathbb{R}}$ are thus the signed normalized Jacobi elements (in $S(\mathbb{R})$ and $Z$ simply denotes $Z_{\epsilon_o}$ where $\epsilon_o = (1, \ldots, 1)$).

**Definition 8.2** (Chevalley invariants and isospectral manifold). If $x \in g$ and $g_x \in g'$ is defined by $\langle g_x, y \rangle = (x, y)$ for any $y \in g$ then the map $g \to g'$ sending $x$ to $g_x$ defines an isomorphism. We can then regard $S(g)$ as the algebra of polynomial functions on $g$ itself by setting for $f \in S(g)$ and $x \in g$, $f(x) = f(g_x)$.

The functions $I_1, \ldots, I_l$ now on $g$ and then restricted to $J_{\mathbb{R}}, Z_{\mathbb{R}}$ or to $Z_\mathbb{C}$ are called the Chevalley invariants which are the polynomial functions of $\{a_1, \ldots, a_l, b_1, \ldots, b_l\}$ for $X = \sum_{i=1}^{l} (a_i h_{\alpha_i} + b_i e_{-\alpha_i} + e_{\alpha_i})$. The map $I = (I_1, \ldots, I_l)$ then defines by restriction a map $I = I_F : Z_F \to F^l$.

Fix $\gamma \in F^l$ in the image of the map $I$, and denote $Z(\gamma)_F = I_F^{-1}(\gamma) = \Pi^{-1}(\gamma) \cap Z_F$, which defines the isospectral manifold of Jacobi elements of $g$. Note that in the real (isospectral) manifold $Z(\gamma)$ studied in [16] will just be one connected component of $Z(\gamma)_{\mathbb{R}}$.

**Definition 8.3** (The isotropy subgroup $\tilde{G}^y$ on $\tilde{G}$). Let $G^y_\mathbb{C}$ be the isotropy subgroup of $G_\mathbb{C}$ for an element $y \in g_\mathbb{C}$. The group $G^y_\mathbb{C}$ is an abelian connected algebraic group of complex dimension $l$ (Proposition 2.4 of [16]). If $y \in g$ we denote by $\tilde{G}^y$ the intersection $\tilde{G}^y = G^y_\mathbb{C} \cap \tilde{G}$.

If $x \in g$, the centralizer of $x$ is denoted $g^x$ and dim $g^x \geq l$. We say that $x$ is regular if dim $g^x = l$.

We consider an open subset of $G_\mathbb{C}$ given by the biggest piece in the Bruhat decomposition:

$$(G_\mathbb{C})_* = \mathcal{N}_C H_\mathbb{C} N_\mathbb{C} = \mathcal{N}_C B_\mathbb{C},$$

where $N_\mathbb{C} = \exp(n_\mathbb{C})$, $\mathcal{N}_C = \exp(\mathfrak{n}_\mathbb{C})$ and $B_\mathbb{C} = H_\mathbb{C} N_\mathbb{C}$. We let $\tilde{G}_* = \tilde{G} \cap (G_\mathbb{C})_*$ and, as in Notation 3.9, $B = H_\mathbb{R} \exp(n)$. We then have a map $\mathcal{N}_C \times B_\mathbb{C} \to (G_\mathbb{C})_*$,

given by $(n, b) \mapsto nb$ which is an isomorphism of algebraic varieties. Given $d \in (G_\mathbb{C})_*$, $d$ has a unique decomposition as $d = n_d b_d$ as in (2.4.6) of [16].
We caution the reader that $\tilde{G}^y$ is not an intersection with $\exp(p)H \exp(n)$ but rather with $\exp(p)H_R \exp(n)$. This is the object that appears, for example, in (3.4.10) of [16] and properly contains (3.2.9) in Lemma 3.2 of [16].

The following Proposition gives the relation between $\tilde{G}^y$ and $H_R$:

**Proposition 8.4.** Let $y \in Z_{\alpha_0}$. Then $\tilde{G}^y$ is $\tilde{G}$ conjugate to the Cartan subgroup $H_R$ of $G$.

**Proof.** By Lemma 2.1.1 in [16], $y$ is regular. Then as in Lemma 3.2 of [16], $y$ must be conjugate, under an element in $H$, to an element $x$ in $p$. Using Proposition 2.4 of [16] $G_C^y$ is connected. Since $x$ is also a regular element it follows that $g_C^y$ is a Cartan subalgebra and $G_C^y$ is a Cartan subgroup of $G_C$. Using conjugation by an element in $K$ we may conjugate this Cartan subgroup if necessary to $H_C$ (Proposition 6.61 or Lemma 6.62 of [13]). We can thus assume that $G_C^y = H_C$. We obtain that $\tilde{G}^y = \tilde{G} \cap G_C^y$ is $\tilde{G}$ conjugate to $\tilde{G} \cap H_C = H_R$ (see Notation 3.9).

**8.2. Kostant’s map $\beta_C^y$.** Fix $y \in J(\gamma)_R$. Kostant defines a map

\[
\beta_C^y : (G_C^y)_* \longrightarrow J(\gamma)_C
\]

\[
d \mapsto \text{Ad}(n_d^{-1})(y)
\]

with $d = n_d b_d$, $n_d \in N$ and $b_d \in B$. Note that we have deviated from the convention in [16] by exchanging the roles of $N_C$ and $N_C$. We did not exchange the roles of these two groups in (19) but this is compensated by our use of an inverse in the definition of the Kostant map. Theorem 2.4 of [16] then implies that $\beta_C^y$ is an isomorphism of algebraic varieties. Denote $\beta_C^y$ the restriction of $\beta_C^y$ to the intersection with $\tilde{G}$. Thus we have $\beta^y : G_C^y \rightarrow Z(\gamma)_R$ and $\tilde{G}_s^y = (G_C^y)_* \cap \tilde{G}$.

**Proposition 8.5.** Let $y \in J(\gamma)_R$. The map $\beta_C^y$ is an isomorphism of smooth manifolds $\tilde{G}_s^y \rightarrow J(\gamma)_R$.

**Proof.** The map $\beta_C^y$ is the restriction to the Lie group $\tilde{G}^y$ of the diffeomorphism of complex analytic manifolds $\beta_C^y$. We obtain that $\beta_C^y$ must be an injective map. We show surjectivity. If $z \in J(\gamma)_R$ then by surjectivity of $\beta_C^y$ there is $g_C \in (G_C^y)_*$ such that $\beta_C^y(g_C) = z$ and $g_C = n_C b_C$. Thus $g_C = n_C^c b_C$ with $n_C^c \in N_C$ and $b_C \in C$. Therefore $\beta_C^y(g_C^c) = \text{Ad}(n_C^c)^{-1} y$. Since $y^c = y$ we obtain that $(\text{Ad}(n_C^c)^{-1} y)^c = z^c$. But our assumption is that $z^c = z$. Hence we have obtained that $\beta_C^y(g_C^c) = z$. By the injectivity of $\beta_C^y$ we obtain that $g_C^c = g_C$. Therefore $g_C \in \tilde{G}$ and thus $g_C = g \in \tilde{G}^y$. This proves $\beta^y$ is a bijection.

By Proposition 2.3.1 of [16], $J(\gamma)_R$ is a submanifold of real dimension $l$ of $J(\gamma)_C$. The diffeomorphism $\beta_C^y$ restricts to the smooth nonsingular map $\beta^y$. Since we have shown that $\beta^y$ is a bijection, then it is a diffeomorphism. □
Remark 8.6. If \( y \in Z_{\epsilon_0} \), then \( \tilde{G}^y \) is a Cartan subgroup conjugate to \( H_\mathbb{R} \) (see Proposition 8.4). However, in general it may happen that \( y \in J(\gamma)_{\mathbb{R}} \) is not semisimple or that \( y \) is semisimple but \( \tilde{G}^y \) is a Cartan subgroup not conjugate to \( H_\mathbb{R} \). For example, take \( y = a_1 h_{\alpha_1} + b_1 e_{-\alpha_1} + e_{\alpha_1} \) in the case of \( \mathfrak{sl}(2, \mathbb{R}) \) (\( \tilde{G} = \text{Ad}(SL(2, \mathbb{R})) \)). We get a nilpotent matrix if \( a_1^2 + b_1 = 0 \). When \( a_1^2 + b_1 < 0 \) then \( y \) is semisimple but \( \tilde{G}^y \) is a compact Cartan subgroup and thus it is not conjugate to \( H_\mathbb{R} \). In the cases \( a_1^2 + b > 0 \) one obtains that \( \tilde{G}^y \) is conjugate to \( H_\mathbb{R} \).

Assume that we are in the case when \( \tilde{G}^y \) is conjugate to \( H_\mathbb{R} \) (for example \( y \in Z_{\epsilon_0} \)). Combining Proposition 8.4 and Proposition 8.5, Kostant’s map gives an imbedding of \( Z(\gamma)_\mathbb{R} \), where \( \gamma \in \mathbb{R}^l \) is in the image of \( \mathbb{I} \), into \( \tilde{H}_\mathbb{R} \) as an open dense subset.

8.3. Kostant’s map and toric varieties. We first remark that for a fixed \( y \in Z(\gamma)_\mathbb{R} \) Kostant’s map \( \beta^y \) in (20) is just the map \( d \to n_d^{-1} \) with \( d = n_d b_d \in \tilde{G}^y \) and \( n_d \in \mathbb{N} \), so that it can be described as a map into the flag manifold: \( d \to gB \) in \( \tilde{G}/B \), restricted to \( \tilde{G}^y \).

By Proposition 8.5 the map into the flag manifold is a diffeomorphism onto its image when restricted to \( \tilde{G}^y \). Hence, since the map is given by the action of a Cartan subgroup on the flag manifold; this action has a trivial isotropy group and the map to the flag manifold sends \( \tilde{G}^y \) diffeomorphically to its image.

The Cartan subgroup \( \tilde{G}^y \) is as good as its conjugate \( H_\mathbb{R} \) but, for convenience, we prefer to deal with \( H_\mathbb{R} \) for which we have established notation. We let \( x \) be an element that conjugates \( H_\mathbb{R} \) to \( \tilde{G}^y \), \( x^{-1} H_\mathbb{R} x = \tilde{G}^y \). Then the \( \tilde{G}^y \)-orbit of \( B \) in \( \tilde{G}/B \) is \( x^{-1} H_\mathbb{R} x B \). Since we can translate this set using multiplication by the fixed element \( x \), we can just study the \( H_\mathbb{R} \)-orbit of \( x B \) in \( \tilde{G}/B \).

We denote:

- The map \( q : \tilde{G}^y \to \tilde{G}/B \).
- \( \hat{Z}(\gamma)_{\mathbb{R}} = (q \circ (\beta^y)^{-1} Z(\gamma)_{\mathbb{R}}) = x^{-1}((H_\mathbb{R} x B)) \).

To study \( \hat{Z}(\gamma) \), it is enough to describe in detail the toric variety \( (H_\mathbb{R} x B) \). Thus we focus our attention on objects that have this general form:

Definition 8.7. For any \( n \in \tilde{G} \), such that \( nB \cap H_\mathbb{R} = \{ e \} \), the toric variety \( (H_\mathbb{R} n B) \) is called generic in the sense of [8], if \( n \in \bigcap_{w \in W} w(\mathcal{N} B) w^{-1} \).

We then assume \( n \in \bigcap_{w \in W} w(\mathcal{N} B) w^{-1} \) and \( nB \cap H_\mathbb{R} = \{ e \} \). With these hypotheses we have:

\[
nB = \prod_{\phi_j \in -\Delta_+} \exp(t^e_{\phi_j} B)
\]
for \( t_j^w \in \mathbb{R} \), and
\[
  nB = n(w(\Delta))wB = \prod_{\phi_j \in -\Delta_+} \exp(t_j^w e_{w(\phi_j)})wB
\]
for \( w \in W \) and \( t_j^w \in \mathbb{R} \).

We now define:

**Definition 8.8.** Assume \( n \in \bigcap_{w \in W} w(\overline{NB})w^{-1} \) and \( nB \cap H_{\mathbb{R}} = \{ e \} \). Denote for any \( A \subset \Pi \),
\[
  n(w(\Delta^A))wB = \prod_{\phi_j \in -\Delta_+^A} \exp(t_j^w e_{w(\phi_j)})wB.
\]

We define a map,
\[
  \tilde{B} : \hat{H}_{\mathbb{R}} \rightarrow (H_{\mathbb{R}}nB) \quad \text{ where } g \in H_{\mathbb{R}} \text{ (see Definition 6.1 for } \hat{H}_{\mathbb{R}}).\]

Note here that
\[
  (21) \quad gn(w(\Delta^A))wB = \prod_{\phi_j \in -\Delta_+^A} \exp\left(t_j^w \chi_{w(\phi_j)}(g)e_{w(\phi_j)}\right) wB.
\]

The map \( \tilde{B} \) can be interpreted as a version of Kostant’s map \( \beta^w \) for the subsystem determined by the set \( A \in \Pi \) and its \( w \)-translation. A detailed correspondence with the set \( \tilde{Z}_{\mathbb{R}}(\gamma) \) could be made but it requires additional notation. Note also that we have \( \chi_{w(\phi_j)}(g) = \chi_{w(\phi_j)}^w \circ \text{Ad}^w(\Delta^A)(g) \) for \( \phi_j \in \Delta^A \).

Then we obtain:

**Theorem 8.9.** Assume \( n \in \bigcap_{w \in W} w(\overline{NB})w^{-1} \) and \( nB \cap H_{\mathbb{R}} = \{ e \} \). The function \( \tilde{B} \) is a homeomorphism of topological spaces. The toric variety \( (H_{\mathbb{R}}nB) \) is a smooth manifold and the map \( \tilde{B} \) is a diffeomorphism.

**Proof.** First we point out that we already have a smooth manifold \( \hat{H}_{\mathbb{R}} \) and the assumption \( n \in \bigcap_{w \in W} w(\overline{NB})w^{-1} \) will simply ensure that we can define the map to the flag manifold.

We first show the continuity of \( \tilde{B} \): We use the local coordinates, \( \{ \phi_w : w \in W \} \) for \( \hat{H}_{\mathbb{R}} \) given in Definition 5.4. In these local coordinates, we assume that for each \( i = 1, \ldots, l \), \( \phi_{w_{w_0}}(\text{Ad}^{w_{w_0}}(\Delta^{w_0(A)})(g), [w_{w_0}]^{w_0(A)}) \) (which equals \( \chi_{w(-\alpha_i)}^w(g) \) or zero) converges to a scalar \( \chi_{w(\alpha_i)}^w \). We note that if \( \phi_j = \)
\[ - \sum_{i=1}^{l} c_{ij} \alpha_i \text{ with each } c_{ij} \text{ a nonnegative integer, then} \]

\[ \chi_w(\phi_j) = \prod_{i=1}^{l} \chi_{-w(\alpha_i)} = \prod_{i=1}^{l} \phi_{w_{w_0}(\Delta^w_0(A))}^{c_{ij}}(\text{Ad}_{w_{w_0}(\Delta^w_0(A))}, [w_{w_0}]^{w_0(A)}). \]

We thus let

\[ \chi_w^o(\phi_j) = \prod_{i=1}^{l} \left( \chi_{-w(\alpha_i)}^{o} \right)^{c_{ij}}. \]

Let

\[ A' = \{ \alpha_i \in \Pi : \chi_{-w(\alpha_i)}^{o} = 0 \}. \]

Then note that \( \chi_w^o(\phi_j) = 0 \) if and only if \( c_{ij} \neq 0 \) for some \( \alpha_i \in A' \). Thus the only \( \chi_w^o(\phi_j) \) which are nonzero correspond to roots in \( \Delta^w_{+} \).

By the assumption made, we have

\[ gn(w(\Delta^A))wB = \prod_{\phi_j \in -\Delta_+} \exp \left( t_j^w \chi_w(\phi_j)(g)e_w(\phi_j) \right) wB, \]

which can be written in terms of the coordinate functions \( \phi_{w_{w_0}} \) as

\[ \prod_{\phi_j \in -\Delta_+} \exp \left( t_j^w \prod_{i=1}^{l} \phi_{w_{w_0},i}^{c_{ij}}(\text{Ad}_{w_{w_0}(\Delta^w_0(A))}, [w_{w_0}]^{w_0(A)}e_w(\phi_j)) \right) wB. \]

This then converges (by continuity of \( \phi_{w_{w_0}} \)) to

\[ \prod_{\phi_j \in -\Delta_+} \exp \left( t_j^w \chi_w^o(\phi_j)(g)e_w(\phi_j) \right) wB, \]

which only involves roots in \( \Delta^w_{+} \), and can be written as \( gn(w(\Delta^A))wB \).

Since any \( (\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \in \hat{H}_R \) is completely determined by the coordinates \( \phi_w(\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \) and some \( (\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \) uniquely corresponds to the coordinates \( (\chi_{w(\alpha_1)}^{o}(g), \ldots, \chi_{w(\alpha_l)}^{o}(g)) \), then we can conclude that \( \hat{B}(\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \) converges to \( \hat{B}(\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \) whenever the argument \( (\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \) approaches to \( (\text{Ad}_{w_{w_0}(\Delta^A)}(g), [w]^A) \) in \( \hat{H}_R \). This proves the continuity of the map \( \hat{B} \).

Since \( \hat{H}_R \) is compact (Proposition 7.7) and \( \hat{B}(\hat{H}_R) \) contains the orbit \( H_RnB \), then \( \overline{(H_RnB)} \subset \hat{B}(\hat{H}_R) \). From the construction of the map \( \hat{B} \) it is easy to see that its image is contained in \( \overline{(H_RnB)} \) and thus \( \hat{B}(\hat{H}_R) = \overline{(H_RnB)} \).
The smoothness of the map and that it gives a diffeomorphism follows from the fact that \((s_1, \ldots, s_{|\Delta_s|}) \rightarrow \prod_{j=1}^{|\Delta_s|} \exp(s_j e^{w(\phi_j)})wB\) constitutes a coordinate system in the flag manifold. □

**Remark 8.10.** We caution the reader that if one replaces the Cartan subgroup \(H^1_\mathbb{R}\) of \(G\) instead of \(H_\mathbb{R}\) in the statement of Theorem 8.9 then the closure of the orbit \(H^1_\mathbb{R}nB\) may not be smooth. In fact the structure of \(H^1_\mathbb{R}nB\) can be explicitly described too. Consider only the connected components of each \(H_\mathbb{R}(w(\Delta^A))\) associated to \(\epsilon\) such that \(\text{Ad}(h_\epsilon) \in \text{Ad}(H^1_\mathbb{R})\) in Definition 5.4. This gives a subspace of \(\hat{H}_\mathbb{R}\) that will correspond to \(H^1_\mathbb{R}nB\).

In terms of the coordinate charts \(\phi_w\) one gets locally \(\mathbb{R}^l\) but now some of its \(2^l\) quadrants may be missing (since \(\text{Ad}(H^1_\mathbb{R})\) may have fewer than \(2^l\) connected components). Thus smoothness is obtained exactly when \(\text{Ad}(H^1_\mathbb{R})\) contains \(2^l\) connected components. Examples that lead to nonsmooth closures if one uses the Cartan subgroup \(H^1_\mathbb{R}\) are all the \(G = SL(n, \mathbb{R})\) with \(n\) even. In terms of the Toda lattice this corresponds to considering the indefinite Toda lattice in (1) in the Introduction but leaving out some of the signs \(\epsilon_i\). When \(n = 2\), for example, one obtains a closed interval inside \(G/B\) (which is a circle). The disconnected Lie group \(\tilde{G}\) which leads to the Cartan subgroup \(H_\mathbb{R}\) is then a requirement in all our constructions and main results. We thus have:

**Corollary 8.11.** Assume \(n \in \bigcap_{w \in W} w(\mathbb{N}B)w^{-1} = \emptyset\) and \(nB \cap H^1_\mathbb{R} = \{e\}\). Then \(H^1_\mathbb{R}nB\) is smooth if and only if \(\text{Ad}(H^1_\mathbb{R})\) has \(2^l\) connected components.

We also have:

**Corollary 8.12.** If \(n \in \bigcap_{w \in W} w(\mathbb{N}B)w^{-1} = \emptyset\) and \(nB \cap H_\mathbb{R} = \{e\}\), then the toric variety \(X = (H_\mathbb{R}nB)\) satisfies: \(X^{H_\mathbb{R}} = (G/B)^{H_\mathbb{R}}\), the \(H_\mathbb{R}\)-fixed points.

**Proof.** We use Theorem 8.9. The manifold \(\hat{H}_\mathbb{R}\) has an \(H_\mathbb{R}\)-action and the only fixed points are the \((e, w)\) with \(w \in W\). These get mapped to the fixed points of the \(H_\mathbb{R}\)-action in \(G/B\), the cosets \(wB\), \(w \in W\). Thus \(X^{H_\mathbb{R}} = (G/B)^{H_\mathbb{R}}\). This also follows directly if we consider (21) with \(A = \emptyset\) and we just let each \(\chi_w(\phi_j)\) go to zero and obtain \(wB\). □
Remark 8.13.

1) By Theorem 3.6 in [9] and Remark 3 in p. 257 of [8] we may assume that the map into the flag manifold which appears as a consequence of Kostant’s map in Subsection 8.3 is such that one in fact obtains the $H_R$ orbit of an element $n \in \bigcap_{w \in W} w(\mathcal{N}B)w^{-1}$.

2) In §7 of [2] and in [3] a more restrictive definition of the notion of genericity is given. This discussion is only relevant for the case of toric varieties in $G/P$ where $P$ is a parabolic subgroup rather than just a Borel subgroup as is the case in the present work. In these more general situations sometimes generic varieties as defined in [8] are not normal. In any case, note that if $P = B$ our Corollary 8.12 implies that if $n \in \bigcap_{w \in W} w(\mathcal{N}B)w^{-1}$ and $nB \cap H_R = \{e\}$ then the corresponding toric variety is also generic in the sense discussed in [2] or [3]. Finally we observe that normality is not used in any of our results. Here we rely instead on explicit coordinate charts to obtain the smoothness of our toric varieties and describe their topological structure. (We would like to thank H. Flaschka for sending the papers [2, 3] to us.)

We conclude:

Theorem 8.14. Let $\gamma \in \mathbb{R}^l$, then $\hat{Z}(\gamma)_R$ is a smooth compact manifold diffeomorphic to $\hat{H}_R$.

Proof. This is just Theorem 8.9 and the definition of $\hat{Z}(\gamma)$. The two conditions in Theorem 8.9 are satisfied by Proposition 8.5 (Subsection 8.3) and by Theorem 3.6 in [9] and Remark 3 in p. 257 of [8] as noted above in Remark 8.13.

References


Received January 19, 2000 and revised April 13, 2001. This research was supported by NSF grant DMS-0071523.
E-mail address: casian@math.ohio-state.edu

Department of Mathematics
The Ohio State University
Columbus, OH 43210
E-mail address: kodama@math.ohio-state.edu
A Diophantine monoid $S$ is a monoid which consists of the set of solutions in nonnegative integers to a system of linear Diophantine equations. Given a Diophantine monoid $S$, we explore its algebraic properties in terms of its defining integer matrix $A$. If $d_r(S)$ and $d_c(S)$ denote respectively the minimal number of rows and minimal number of columns of a defining matrix $A$ for $S$, then we prove in Section 3 that $d_r(S) = \text{rank } \text{Cl}(S)$ and $d_c(S) = \text{rank } \text{Cl}(S) + \text{rank } Q(S)$ where $\text{Cl}(S)$ represents the divisor class group of $S$ and $Q(S)$ the quotient group of $S$. The proof relies on the characteristic properties of the so-called essential states of $S$, which are developed in Section 2. We close in Section 4 by offering a characterization of factorial Diophantine monoids and an algorithm which determines if a Diophantine monoid is half-factorial.

1. Introduction.

Because of their applications in commutative algebra, algebraic geometry, combinatorics, number theory, and computational algebra, the study of commutative cancellative monoids has recently increased in popularity (see for example [14]). Let $\mathbb{Z}$ and $\mathbb{N}$ represent the integers and the nonnegative integers respectively. For $1 \leq m, n$ in $\mathbb{N}$ and $A \in \mathbb{Z}^{m \times n}$ set

\begin{equation}
M_A = \mathbb{N}^n \cap \{x \in \mathbb{Z}^n \mid Ax = 0\}.
\end{equation}

We will refer to $M_A$ as a Diophantine monoid and to $A$ as a matrix which determines $M_A$. Admitting the possibility that $m = 0$, we set $M_A = \mathbb{N}^n$ if $m = 0$. The special case of a Diophantine monoid $M_A$ for $A \in \mathbb{Z}^{1 \times n}$ (i.e., one single homogeneous linear Diophantine equation) has been studied in [2], where it was shown that the divisor class group of $M_A$ (denoted $\text{Cl}(M_A)$) in this case must be cyclic [2, Theorem 1.3].

It is natural to consider the question of whether a Diophantine monoid $M_A$ given by a matrix $A \in \mathbb{Z}^{m \times n}$ can be described (up to isomorphy) by another matrix with less rows or less columns. For a Diophantine monoid $S$, let $d_r(S)$ and $d_c(S)$ denote respectively the minimal number of rows and minimal number of columns of a matrix $A$ with $S \simeq M_A$. One has that
\[ d_e(S) = d_r(S) + \text{rank } Q(S) \] where \( Q(S) \) is the quotient group of \( S \). In this paper, we prove for a monoid which is root-closed and finitely generated that \( d_r(S) = \text{rank } \text{Cl}(S) \) and hence \( d_e(S) = \text{rank } \text{Cl}(S) + \text{rank } Q(S) \) (Theorem 3.8). Furthermore, this theorem also shows that for such an \( S \) there exists an \( A \in \mathbb{Z}^{m \times n} \) with \( m = d_r(S), n = d_e(S) \) and \( S \simeq M_A \).

More precisely, we divide our work into three additional sections. After a very brief review of Krull monoids, we focus our considerations in Section 2 on the so-called essential states of these monoids and provide in Lemma 2.3 and Proposition 2.4 their characteristic properties. Using these properties, we obtain the above mentioned result (Theorem 3.8) from our crucial Theorem 3.1 together with Lemma 3.5. Theorem 3.1 demonstrates that any Krull monoid with finitely many essential states must be isomorphic to a Diophantine monoid and describes the structure of the representing matrix as well as the structure of the class group. As a consequence, we obtain in Corollary 3.3 various equivalent characterizations of the above Krull monoids or, equivalently, Krull monoids with finitely generated divisor class group and finitely many prime divisors. For some results related to Corollary 3.3, the interested reader can consult \[7, \text{Theorem 5}], \[8, \text{Lemma 3}], \[10, \text{Proposition 2}], and \[13, \text{Corollary 1}]\. We note that Lemma 3.5 yields a description of the class group of a Diophantine monoid \( S \) purely in terms of linear algebra.

In Section 4, we present some examples to illustrate the results of Sections 2 and 3 and also obtain in Proposition 4.1 a characterization of when \( M_A \) is factorial (recall that \( M_A \) is factorial if any nonzero element has a representation \( \alpha_1 + \cdots + \alpha_k \) by irreducible elements \( \alpha_i \) which is unique up to ordering). We close by presenting an algorithm, which when teamed with known algorithms for computing the set of minimal nonnegative solutions to \( M_A \), will compute the class group of \( M_A \) and determine if \( M_A \) is half-factorial (recall that \( M_A \) is half-factorial if whenever \( \alpha_1, \alpha_2, \ldots, \alpha_k \) and \( \beta_1, \beta_2, \ldots, \beta_l \) are irreducible elements of \( M_A \) with \( \alpha_1 + \cdots + \alpha_k = \beta_1 + \cdots + \beta_l \), then \( k = l \)).

Although the literature concerning the algebraic structure of the monoids \( M_A \) is not extensive, we take interest in this topic partly because of the rich mathematical history behind the study of Diophantine equations. It is easy to determine the solution set of a system of linear Diophantine equations over the integers, but this is not the case for determining the set of solutions over the nonnegative integers. While a modern treatment of the combinatorial aspects of this subject can be found in the works of Stanley (see \[15] and \[16] for example), attempts to determine the set of “irreducible solutions” of \( M_A \) can be traced back almost 100 years to a paper of Elliott \[4], where the author produces generating functions to determine these solutions. Another early attempt at producing this set of minimal solutions can be found in \[5]. The development of modern algorithms connected with these solutions has become a popular topic of research in computational algebra (see \[3] and \[14]).
2. Essential states.

Let $S$ be a commutative cancellative monoid (i.e., a subsemigroup of an abelian group written additively with $0 \in S$). Throughout this paper we assume that $S \neq \{0\}$. If $\{0\}$ is the only subgroup of $S$, then $S$ is called reduced. Let

$$Q(S) := \{x - y \mid x, y \in S\}$$

be the group generated by $S$. A homomorphism $\pi : Q(S) \to \mathbb{Z}$ is called a state of $S$ if $\pi(S) \subseteq \mathbb{N}$. The monoid $S$ is a Krull monoid if there exists, for some set $K$, a monomorphism

$$\varphi : Q(S) \to \mathbb{Z}^{|K|}$$

such that

$$\varphi(S) = \mathbb{N}^{|K|} \cap \varphi(Q(S)) = \mathbb{N}^{|K|} \cap Q(\varphi(S)).$$

As a consequence of our definition, a Krull monoid is always reduced.

For $j \in K$ let $p_j : \mathbb{Z}^{|K|} \to \mathbb{Z}$ be the surjection onto the $j$-th component, $p_j((x_k)_{k \in K}) = x_j$, and

$$\pi_j : Q(S) \to \mathbb{Z}, \quad \pi_j = p_j \circ \varphi.$$

Then

$$(*) \quad S = \{x \in Q(S) \mid \pi_j(x) \geq 0 \text{ for all } j \in K\}$$

where $(\pi_j)_{j \in K}$ is a family of states of $S$ with $\pi_j(x) = 0$ for almost all $j \in K$ and any fixed $x \in Q(S)$. We may assume that $\pi_j \neq 0$ for all $j \in K$.

**Example 2.1.** For every $A \in \mathbb{Z}^{m \times n}$ the Diophantine monoid

$$M_A = \{x \in \mathbb{N}^n \mid Ax = 0\} \subseteq \mathbb{Z}^n$$

is a Krull monoid. With $K = \{1, 2, \ldots, n\}$ and $\varphi : Q(M_A) \to \mathbb{Z}^n$ the canonical embedding, we have

$$M_A = \{(x_1, \ldots, x_n) \in Q(M_A) \mid \pi_j(x) = x_j \geq 0 \text{ for } j \in K\}.$$

When a Krull monoid $S$ is given in the form $(*)$, it is natural to ask for minimal subsets $E \subseteq K$ with the property that

$$S = \{x \in Q(S) \mid \pi_i(x) \geq 0 \text{ for all } i \in E\}.$$

Such subsets exist and can be described by the so-called essential states of $S$.

**Definition 2.2.** A nonzero state $\pi$ of $S$ is called essential, if for every $x, y \in Q(S)$ with $\pi(x) \geq \pi(y)$ there exists $z \in Q(S)$ with

$$\pi(z) = \pi(x), \quad z - x \in S \text{ and } z - y \in S.$$
Lemma 2.3. Let $S$ be a Krull monoid and $(\pi_j)_{j \in J}$ the family of all nonzero states of $S$. For every $i \in J$ the following statements are equivalent:

i) $\pi_i$ is essential.

ii) For every $j \in J$ with $\pi_j \notin \mathbb{Q}\pi_i$ there exists $x \in S$ such that $\pi_i(x) = 0$ and $\pi_j(x) > 0$.

iii) $S \cap \ker \pi_i$ is a maximal element in the set $\{S \cap \ker \pi_j | j \in J\}$ with respect to set inclusion.

iv) If $j \in J$ and $(S \cap \ker \pi_i) \subseteq (S \cap \ker \pi_j)$ then $\pi_j = \alpha \pi_i$ for some $\alpha \in \mathbb{Q}, \alpha > 0$.

In particular, the family of essential states of $S$ is not empty.

Proof. Note that for any finite subset $I \subseteq J$, there exists $x \in S$ with $\pi_i(x) > 0$ for all $i \in I$.

Statements ii) and iv) are obviously equivalent. We begin by showing that i) $\Rightarrow$ iii). Assume that i) holds and that, for some $j \in J$,

$$(S \cap \ker \pi_i) \not\subseteq (S \cap \ker \pi_j).$$

Choose $y \in (S \cap \ker \pi_j) \setminus \ker \pi_i$ and $x \in S$ with $\pi_j(x) > 0$. Then

$$u := \pi_i(y)x - \pi_i(x)y \in \ker \pi_i.$$

Condition i) yields an element $z \in Q(S)$ with

$$\pi_i(z) = \pi_i(\pi_i(y)x) = \pi_i(\pi_i(x)y),$$

$$z - \pi_i(y)x \in S \cap \ker \pi_i \subseteq \ker \pi_j$$

and

$$z - \pi_i(x)y \in S \cap \ker \pi_i \subseteq \ker \pi_j.$$

In particular, $\pi_j(z - \pi_i(y)x) = \pi_j(z - \pi_i(x)y) = 0$. But this gives

$$0 = \pi_j(z - \pi_i(x)y) = \pi_j(z) - \pi_i(x)\pi_j(y) = \pi_j(z),$$

and the contradiction

$$0 = \pi_j(z - \pi_i(y)x) = \pi_j(z) - \pi_i(y)\pi_j(x) = -\pi_i(y) \cdot \pi_j(x) < 0.$$

For the rest of the proof, let $K \subseteq J$ be a subset of $J$ with the property that

$$S = \{x \in Q(S) | \pi_j(x) \geq 0 \text{ for all } j \in K\}$$

where $\pi_j(x) = 0$ for almost all $j \in K$ and any fixed $x \in Q(S)$. The representation (3) shows that such subsets $K$ exist.

For iii) $\Rightarrow$ iv), we fix $i \in J$ and assume that iii) is true for $\pi_i$. Moreover we fix $j \in J$ and assume that $S \cap \ker \pi_i \subseteq S \cap \ker \pi_j$. Define $L := K \cup \{i, j\}$.
Since and π from iii) we know that

\[ L_0 := \{ k \in L \mid (S \cap \text{Ker} \pi_i) \subseteq \text{Ker} \pi_k \} = \{ k \in L \mid S \cap \text{Ker} \pi_i = S \cap \text{Ker} \pi_k \}. \]

We choose some \( v \in S \) with \( \pi_i(v) > 0 \). Then \( \pi_j(v) > 0 \) by our assumption and \( \alpha = \frac{\pi_j(v)}{\pi_i(v)} > 0 \). Let \( w \in S \) be arbitrary. If \( \pi_i(w) = 0 \), then \( \pi_j(w) = 0 \) and \( \pi_i(w) = \alpha \pi_j(w) \). Assume that \( \pi_i(w) > 0 \). Then

\[ i \in L'_0 := \{ k \in L_0 \mid \pi_k(w) \neq 0 \}. \]

Since \( L'_0 \) is finite, we can define \( m = \prod_{k \in L'_0} \pi_k(w) \) and

\[ \lambda = \max \{ r \in \mathbb{N} \mid r \geq 1, \pi_k(mv - rw) \geq 0 \text{ for all } k \in L_0 \}. \]

For some \( k_1 \in L'_0 \) the equality

\[ \pi_{k_1}(mv - \lambda w) = 0 \]

holds. For every \( k \in L_2 := \{ k \in L \setminus L_0 \mid \pi_k(mv - \lambda w) \neq 0 \} \) we choose \( x_k \in S \cap \text{Ker} \pi_i \) with

\[ \pi_k(x_k) > \max \{ 0, \pi_k(\lambda w - mw) \}. \]

Since \( L_2 \) is finite (and possibly empty), the element \( x := \sum_{\mu \in L_2} x_\mu \) is well-defined and \( x \in S \cap \text{Ker} \pi_i \subseteq \text{Ker} \pi_j \).

In the next step, we show that \( u = x + mv - \lambda w \in S \) (i.e., that \( \pi_k(u) \geq 0 \) for all \( k \in L \)). We have already seen that for \( k \in L_0 \) we have

\[ \pi_k(u) = \pi_k(x) + \pi_k(mv - \lambda w) \geq \pi_k(mv - \lambda w) \geq 0. \]

For \( k \in L_2 \) we get

\[ \pi_k(u) = \pi_k(x) + \pi_k(mv - \lambda w) \geq \pi_k(x_k) + \pi_k(mv - \lambda w) \geq 0, \]

and finally for \( k \in L \setminus (L_0 \cup L_2) = (L \setminus L_0) \setminus L_2 \) we have that

\[ \pi_k(u) = \pi_k(x) \geq 0. \]

Therefore \( u \in S \), and it follows for the above chosen \( k_1 \in L'_0 \) with \( \pi_{k_1}(mv - \lambda w) = 0 \) that

\[ \pi_{k_1}(u) = \pi_{k_1}(x) + \pi_{k_1}(mv - \lambda w) = \pi_{k_1}(x) = 0, \]

since \( x \in S \cap \text{Ker} \pi_i = S \cap \text{Ker} \pi_{k_1} \). In particular, \( u \in S \cap \text{Ker} \pi_i \subseteq \text{Ker} \pi_j \).

This gives

\[ \pi_j(u) = 0 = \pi_i(u) = \pi_i(mv - \lambda w) = \pi_j(mv - \lambda w). \]

Hence

\[ \lambda \pi_j(w) = \pi_j(mv), \lambda \pi_i(w) = \pi_i(mv) \]

and

\[ \pi_j(w) = \frac{1}{\lambda} \pi_j(mv) = \frac{\pi_j(mv)}{\pi_i(mv)} \pi_i(w) = \frac{\pi_j(v)}{\pi_i(v)} \pi_i(w) = \alpha \pi_i(w). \]
Since \( w \in S \) was chosen arbitrarily, it follows that \( \pi_j = \alpha \pi_i \).

For iv) \( \Rightarrow \) i), assume that \( x, y \in Q(S) \) are given with \( \pi_i(x) \geq \pi_i(y) \). As above, we define

\[
L := K \cup \{i, j\}, L_0 := \{k \in L \mid (S \cap \text{Ker} \pi_i) \subseteq \text{Ker} \pi_k\}, L_1 := L \setminus L_0,
\]

and \( L_2 := \{k \in L_1 \mid \pi_k(x - y) \neq 0\} \). Note that \( L_2 \) is a finite set. By iv) we know that

\[
\pi_k = \alpha_k \pi_i, \quad \text{for all } k \in L_0 \text{ where } 0 < \alpha_k \in \mathbb{Q}.
\]

For every \( k \in L_1 \) there exists \( x_k \in S \cap \text{Ker} \pi_i \) with

\[
\pi_k(x - y) + \pi_k(x_k) \geq 0.
\]

Defining \( z = x + \sum_{\mu \in L_2} x_\mu \), we get

\[
z - x = \sum_{\mu \in L_2} x_\mu \in S.
\]

We prove that \( z - y \in S \) by showing that \( \pi_k(z - y) \geq 0 \) for all \( k \in L \).

If \( k \in L_2 \), then

\[
\pi_k(z - y) = \pi_k \left( x - y + \sum_{\mu \in L_2} x_\mu \right) = \pi_k(x - y) + \sum_{\mu \in L_2} \pi_k(x_\mu) = \pi_k(x - y) + \pi_k(x_k) + \sum_{\mu \in L_2 \setminus \{k\}} \pi_k(x_\mu) \geq 0.
\]

For \( k \in L_0 \) we get

\[
\pi_k(z - y) = \pi_k(x - y) = \alpha_k \pi_i(x - y) \geq 0.
\]

Finally, for \( k \in L_1 \setminus L_2 \) we have

\[
\pi_k(z - y) = \pi_k \left( \sum_{\mu \in L_2} x_\mu \right) \geq 0.
\]

\[
\square
\]

Again, let the Krull monoid \( S \) be given in the form (\(*\)). Thus,

\[
S = \{x \in Q(S) \mid \pi_j(x) \geq 0 \text{ for all } j \in K\}
\]

where \((\pi_j)_{j \in K}\) is a family of nonzero states of \( S \) such that \( \pi_j(x) = 0 \) for almost all \( j \in K \) and any fixed \( x \in Q(S) \).

**Proposition 2.4.** For every \( \emptyset \neq I \subseteq K \) the following statements are equivalent:

i) \( S = \{x \in Q(S) \mid \pi_i(x) \geq 0 \text{ for all } i \in I\} \).
ii) If $\pi$ is an essential state of $S$, then there exists $i \in I$ and $\alpha \in \mathbb{Q}, \alpha > 0$, such that

$$\pi = \alpha \pi_i.$$

Proof. i) $\implies$ ii) Let $\pi$ be an essential state of $S$. It suffices to show that $\pi \in \mathbb{Q} \pi_i$ for some $i \in I$. Let $i_0 \in I$ and $v, w \in S$ with $\pi(v) > 0, \pi_{i_0}(w) > 0$ and define $u := v + w$. Then $\pi(u) > 0$ and

$$i_0 \in I_1 := \{ i \in I | \pi_i(u) > 0 \}, \quad |I_1| < \infty.$$ We claim that $\pi \in \mathbb{Q} \pi_i$ for some $i \in I_1$. Assume the contrary. Then $\pi_i \notin \mathbb{Q} \pi$ and we apply Lemma 2.3 ii) to get for every $i \in I_1$ an element $x_i \in S$ with $\pi(x_i) = 0$ and $\pi_i(x_i) > 0$. With $x := \sum_{i \in I_1} x_i \in S$ it follows that $\pi(x) = 0$ and for every $i \in I_1$

$$\pi_i(x) = \pi_i(x_i) + \sum_{j \in I \setminus \{ i \}} \pi_i(x_j) > 0.$$ Let $i_1 \in I_1$ with

$$\frac{\pi_{i_1}(x)}{\pi_{i_1}(u)} = \min \left\{ \frac{\pi_i(x)}{\pi_i(u)} \mid i \in I_1 \right\}$$

and $z := \pi_{i_1}(u)x - \pi_{i_1}(x)u$. From i) we get $z \in S$, since

$$\pi_i(z) = \pi_{i_1}(u)\pi_i(x) - \pi_{i_1}(x)\pi_i(u) \geq 0 \quad \text{for } i \in I_1,$$

and

$$\pi_i(z) = \pi_{i_1}(u)\pi_i(x) \geq 0 \quad \text{for } i \in I \setminus I_1.$$ But $\pi(z) = -\pi_{i_1}(x)\pi(u) < 0$, a contradiction to the fact that $\pi(S) \subseteq \mathbb{N}$.

ii) $\implies$ i) Let

$$N := \{ w \in Q(S) \mid \pi_i(w) \geq 0 \quad \text{for all } i \in I \}.$$ For $y \in N$ we consider the finite set $K(y) := \{ j \in K \mid \pi_j(y) < 0 \}$ and prove that $|K(y)| = 0$ (i.e., that $A := \{ y \in N \mid |K(y)| \geq 1 \} = \emptyset$). Assume the contrary and choose $y \in A$ such that $|K(y)|$ is minimal. Let $j_0 \in K$ such that $\pi_{j_0}$ is an essential state of $S$. From ii) we know that $j_0 \notin K(y)$. For every $j \in \{ j_0 \} \cup K(y)$ there exists $x_j \in S$ with $\pi_j(x_j) > 0$. Thus $x := \sum_{j \in \{ j_0 \} \cup K(y)} x_j \in S$ and $\pi_j(x) > 0$ for every $j \in \{ j_0 \} \cup K(y)$. Let $j_1 \in K(y)$ such that

$$0 > \frac{\pi_{j_1}(y)}{\pi_{j_1}(x)} = \max \left\{ \frac{\pi_j(y)}{\pi_j(x)} \mid j \in K(y) \right\}.$$ We define $z := \pi_{j_1}(x)y - \pi_{j_1}(y)x$. Since $\pi_{j_1}(y) < 0$ it follows for every $j \in K \setminus K(y)$ that

$$\pi_j(z) = \pi_{j_1}(x)\pi_j(y) - \pi_{j_1}(y)\pi_j(x) \geq -\pi_{j_1}(y)\pi_j(x) \geq 0.$$
In particular, \( z \in N \) since \( I \subseteq K \setminus K(y) \), and \( K(z) \subseteq K(y) \). Moreover \( |K(z)| < |K(y)| \) because \( j_1 \notin K(z) \). From the minimality of \( |K(y)| \) we conclude \( |K(z)| = 0 \) (i.e., \( z \in S \)). Since
\[
\pi_{j_0}(z) = \pi_{j_1}(x)\pi_{j_0}(y) - \pi_{j_1}(y)\pi_{j_0}(x) \geq -\pi_{j_1}(y)\pi_{j_0}(x) > 0
\]
it follows that
\[
z \in (S \cap \text{Ker } \pi_{j_1}) \subsetneq (S \cap \text{Ker } \pi_{j_0})
\]
But, by Lemma 2.3 iii), there has to be an essential state \( \pi \) of \( S \) with
\[
(S \cap \text{Ker } \pi_{j_1}) \subseteq (S \cap \text{Ker } \pi).
\]
Since \( \pi_{j_0} \) was already arbitrarily chosen, this contradiction finishes the proof.

We call a state \( \pi \) of \( S \) a **normal state**, if \( \pi(Q(S)) = \mathbb{Z} \). For every nonzero state \( \pi : Q(S) \to \mathbb{Z} \), the image \( \pi(Q(S)) \) is a nonzero ideal \( d\mathbb{Z} \) of \( \mathbb{Z} \) where \( d \in \mathbb{N}, d \geq 1 \), and \( \pi_{\text{nor}} := \frac{1}{d} \pi \) is normal. Let \( K_N \) denote the set of all normal states of a given Krull monoid \( S \). From Proposition 2.4 we obtain the following.

**Corollary 2.5.** There is a unique minimal subset \( I(S) \subseteq K_N \) such that
\[
S = \{ x \in Q(S) \mid \pi(x) \geq 0 \quad \text{for all } \pi \in I(S) \}.
\]
The set \( I(S) \) consists exactly of the essential normal states of \( S \). Moreover \( \pi(x) = 0 \) for almost all \( \pi \in I(S) \) and any fixed \( x \in Q(S) \).

**Proof.** Let \( L \) be a set of nonzero states of \( S \) such
\[
S = \{ x \in Q(S) \mid \pi(x) \geq 0 \quad \text{for all } \pi \in L \}
\]
and \( \pi(x) = 0 \) for almost all \( \pi \in L \) and any fixed \( x \in Q(S) \). Define \( \bar{L} := \{ \pi_{\text{nor}} \mid \pi \in L \} \). Obviously \( S = \{ x \in Q(S) \mid \pi_{\text{nor}}(x) \geq 0 \quad \text{for all } \pi \in L \} \), and the set \( I(S) \) of normal essential states of \( S \) is a subset of \( \bar{L} \) by Proposition 2.4. In particular \( \pi(x) = 0 \) for almost all \( \pi \in I(S) \) and any fixed \( x \in Q(S) \). Now we may again apply Proposition 2.4 and conclude that \( S = \{ x \in Q(S) \mid \pi(x) \geq 0 \quad \text{for all } \pi \in I(S) \} \).

We consider the form
\[
S = \{ x \in Q(S) \mid \pi(x) \geq 0 \quad \text{for all } \pi \in I(S) \}
\]
as the (uniquely determined) **normal representation** of the Krull monoid \( S \). It gives rise to a **divisor theory**
\[
\varphi : Q(S) \to \mathbb{Z}^{(I(S))}, \quad \text{(direct sum)}
\]
\[
\varphi(x) := (\pi(x))_{\pi \in I(S)},
\]
with \( \text{Ker } \varphi = S \cap (-S) = \{0\} \), which maps \( Q(S) \) onto a free subgroup of \( \mathbb{Z}^{(I(S))} \) such that

\[
\varphi(S) = \mathbb{N}^{(I(S))} \cap \varphi(Q(S)) = \mathbb{N}^{(I(S))} \cap Q(\varphi(S)).
\]

The quotient group

\[
\text{Cl}(S) := \mathbb{Z}^{(I(S))}/\varphi(Q(S))
\]

is the \textit{divisor class group} of \( S \). It is a direct consequence of the construction that isomorphic Krull monoids have isomorphic divisor class groups.

Now suppose that \( A \in \mathbb{Z}^{m \times n} \) and \( S \) is isomorphic to the Diophantine monoid

\[ M_A = \{ x \in \mathbb{N}^n \mid Ax = 0 \}. \]

For \( 1 \leq i \leq n \) let \( p_i : \mathbb{Z}^n \to \mathbb{Z} \), with \( p_i(x_1, \ldots, x_n) = x_i \), denote the natural surjection. The restriction maps \( \pi_i = p_i|Q(M_A) \) from \( Q(M_A) \) into \( \mathbb{Z} \) are states of \( M_A \). We call them the \textit{canonical projections} of \( Q(M_A) \). Note that a relation \( \pi_i \equiv \pi_j \) does not imply \( i = j \). Moreover, for \( 1 \leq i \leq n \) we define \( c_i = \gcd(x_i \mid x = (x_1, \ldots, x_n) \in M_A) \in \mathbb{N} \) where \( c_i = 0 \) if and only if \( x_i = \pi_i(x) = 0 \) for all \( x \in M_A \). We call the product \( w(M_A) = \prod_{i=1}^{n} c_i \) the \textit{weight} of the monoid \( M_A \). If \( n \geq 2 \) and \( c_i = 0 \), then \( M_A \) is canonically isomorphic to the monoid \( M_{\bar{A}} \subseteq \mathbb{N}^{n-1} \), where \( \bar{A} \in \mathbb{Z}^{m \times (n-1)} \) is given by canceling the \( i \)-th column of \( \bar{A} \). Therefore it suffices to study the situation \( w(M_A) \neq 0 \). All canonical projections \( \pi_i, 1 \leq i \leq n \), are normal if and only if \( w(M_A) = 1 \). In this case, every normal essential state of \( M_A \) is a canonical projection (as follows immediately from Corollary 2.5), hence there are at most \( n \) essential states of \( M_A \). In general, not all normal canonical projections of \( Q(M_A) \) are essential states of \( M_A \). Let

\[ I(M_A) := \{ i \mid 1 \leq i \leq n, \pi_i \text{ is an essential state of } M_A \} \]

and note that the following statements are equivalent:

i) \( w(M_A) \neq 0 \),

ii) there exist elements \( x = (x_1, \ldots, x_n) \in M_A \) with \( x_i > 0 \) for \( 1 \leq i \leq n \).

**Lemma 2.6.** Let \( M_A \) be a Diophantine monoid with \( A \in \mathbb{Z}^{m \times n} \) and \( w(M_A) \neq 0 \). The following statements are true:

i) \( Q(M_A) = \{ x \in \mathbb{N}^n \mid Ax = 0 \} \), and there exists a linearly independent system of vectors \( v_1, v_2, \ldots, v_{n-r} \in M_A \), where \( r = \text{rank } A \).

ii) Let \( c_i \) be defined as above and \( D := \text{diag}(c_1, \ldots, c_n) \in \mathbb{Z}^{n \times n} \). The map \( M_{AD} \to M_A \), defined by \( y \mapsto Dy \), is an isomorphism of monoids and \( w(M_{AD}) = 1 \).

iii) There are at most \( n \) normal essential states of \( M_A \).
Proof. i): The assumption $w(M_A) \neq 0$ yields an element $x = (x_1, \ldots, x_n) \in M_A$ with $x_i > 0$ for $1 \leq i \leq n$. Obviously $Q(M_A) \subseteq G = \{ u \in \mathbb{Z}^n | Au = 0 \}$. If $y \in G$ then $z = kx - y \in \mathbb{N}^n \cap G = M_A$ for $k \in \mathbb{N}$ sufficiently big, hence $y = kx - z \in Q(M_A)$ and $Q(M_A) = G$. Let $r = \text{rank} A$ and $w_1, \ldots, w_{n-r}$ be a basis of the $\mathbb{Q}$-vector space $W = \{ z \in \mathbb{Q}^n | Az = 0 \}$. We may assume that $w_j \in \mathbb{Z}^n$ for $1 \leq j \leq n-r$. Let $m_0 \in \mathbb{N}$ with $m_0x + w_j \in M_A$, $1 \leq j \leq n-r$. Since the family $(kx + w_j | 1 \leq j \leq n-r)$ is linearly independent for all but at most one $k \in \mathbb{N}$, we are done.

We leave the verification of statements ii) and iii) to the reader. \hfill \Box

3. The representation of Krull monoids by matrices.

In this section, we show that any Krull monoid $S$ with a finite number $e$ of essential states must be isomorphic to a Diophantine monoid. We describe the structure of the corresponding matrix and show in principle its computation, as well as the computation of the divisor class group of $S$. For the class group $\text{Cl}(S)$ defined in the previous section, we have the short exact sequence

$$0 \longrightarrow Q(S) \xrightarrow{\varphi} \mathbb{Z}^{(I)} \xrightarrow{\rho} \text{Cl}(S) \longrightarrow 0$$

where $I = I(S)$ indexes the set of all normalized essential states of $S$ and where $\rho$ denotes the canonical epimorphism. To obtain the desired description of $S$ by a matrix, we will use a particular basis for the $\mathbb{Z}$-module $\mathbb{Z}^{(I)}$. For an arbitrary $\mathbb{Z}$-module $M$, let $\text{rank} M$ denote the minimal number of generators of $M$ and $\text{free rank } M$ the maximal length of a free family in $M$. Of course, if $M$ is a free module, then the rank and the free rank of $M$ coincide and are equal to the cardinality of a $\mathbb{Z}$-basis of $M$ (since $M = \{0\}$ is generated by the empty set, one has rank $\{0\} = \text{free rank } \{0\} = 0$).

**Theorem 3.1.** For a (reduced) Krull monoid $S$, the number $e$ of essential states is finite if and only if the rank $r$ of $Q(S)$, the free rank $h$ of $\text{Cl}(S)$, and the rank $k$ of the torsion group of $\text{Cl}(S)$ are all finite. In this case, the following statements apply.

i) There exist natural numbers $\alpha_1, \ldots, \alpha_k \geq 2$ with $\alpha_{i+1} \mid \alpha_i$ for $1 \leq i \leq k-1$ such that

$$\text{Cl}(S) \simeq \mathbb{Z}_{\alpha_1} \oplus \cdots \oplus \mathbb{Z}_{\alpha_k} \oplus \mathbb{Z}^h$$

and $h + r = e$.

ii) $S$ is isomorphic to a Diophantine monoid $M_A$ with $A \in \mathbb{Z}^{m \times n}$ for $m = k + h$ and $n = k + e$ (if $h = k = 0$, then $m = 0$ and $M_A = \mathbb{N}^e$).

iii) The matrix $A$ has a block structure.
A = \[
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & 0
\end{bmatrix}
\]

where $A_{11} \in \mathbb{N}^{k \times e}$ with entries of the $i$-th row in $\{0, 1, \ldots, \alpha_i - 1\}$, $A_{12} = -\text{diag}(\alpha_1, \ldots, \alpha_k) \in \mathbb{N}^{k \times k}$ and $A_{21} \in \mathbb{N}^{h \times e}$ contains in each row at least two strictly positive and two strictly negative entries.

**Proof.** $M = \varphi(Q(S))$ is a submodule of the free module $\mathbb{Z}^{(I)}$ and therefore free. If $e = |I|$ is finite then rank $M$ is finite and, hence, $r$ is finite. By its definition as the quotient of $\mathbb{Z}^{(I)}$ and $M$, the class group is finitely generated and $h$ and $k$ are finite. Conversely, if $r, h$ and $k$ are finite then $k + h + r$ is finite.

**Step I.** In a first step, we analyze certain submodules of $\mathbb{Z}^{(I)}$ for an arbitrary index set $I$. Let $(b_i)_{i \in I}$ be a basis of the free module $\mathbb{Z}^{(I)}$. For a given set $\{\alpha_j \mid j \in J\} \subseteq \mathbb{N}\setminus\{0\}$, $\emptyset \neq J \subseteq I$, let $M$ be the free submodule with basis $(\alpha_j b_j)_{j \in J}$. Consider the epimorphism $\psi: \mathbb{Z}^{(I)} \rightarrow \bigoplus_{i \in J'} \mathbb{Z}_{\alpha_i} \bigoplus \mathbb{Z}^{(I \setminus J)}$ for $J' = \{j \in J \mid \alpha_j \neq 1\}$ defined by

$$
\psi\left(\sum_{i \in I} r_i b_i\right) = \sum_{i \in J'} (r_i \mod \alpha_i) b_i + \sum_{i \in I \setminus J} r_i b_i.
$$

Obviously, $\ker \psi = M$ and, hence,

$$
\mathbb{Z}^{(I)} / M \cong \bigoplus_{i \in J'} \mathbb{Z}_{\alpha_i} \bigoplus \mathbb{Z}^{(I \setminus J)}.
$$

To describe $M$ by a matrix, let $(u_i)_{i \in I}$ be the standard basis of $\mathbb{Z}^{(I)}$ and $u_j = \sum_{i \in I} c_{ij} b_i$, $c_{ij} \in \mathbb{Z}$ for $i, j \in I$. Define a block matrix

$$
A = \[
\begin{bmatrix}
\begin{array}{cc}
I & J' \\
J' & I \setminus J
\end{array}
\end{bmatrix}
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & 0
\end{bmatrix}
\]
$$

where the blocks in the above diagram are given as follows:
Consider now a (reduced) Krull monoid \( S \) for which \( I = I(S) = \{1, \ldots, e\} \) is finite and let \( M = \varphi(Q(S)) \subseteq \mathbb{Z}^I \). Since \( \mathbb{Z}^I \) is finitely generated, by the elementary divisor theorem (see [9]) there exist a basis \((b_i)_{i \in I}\) of \( \mathbb{Z}^I \) and numbers \( \alpha_j \in \mathbb{N} \setminus \{0\} \) for \( 1 \leq j \leq r \leq e \) with \( \alpha_j \neq 1 \) for \( 1 \leq j \leq k \leq r \) and \( \alpha_{j+1} \mid \alpha_j \) for \( 1 \leq j \leq k - 1 \) such that \((\alpha_j b_j)_{1 \leq j \leq r}\) is a basis of \( M \) (\( k = 0 \) admitted). Obviously, \( r = \text{rank} \varphi(Q(S)) = \text{rank} Q(S) \).

Setting \( J = \{1, \ldots, r\} \) and \( J' = \{1, \ldots, k\} \), from (2) in Step I, we obtain for \( M = \varphi(Q(S)) \) that \( \text{Cl}(S) = \mathbb{Z}^I / M \simeq k \bigoplus_{i=1}^{k} \mathbb{Z} \alpha_i \bigoplus \mathbb{Z}^{I \setminus J} \).

It follows that \( k \) is the rank of the torsion group of \( \text{Cl}(S) \) and that \( e - r = |I \setminus J| \) is the free rank of \( \text{Cl}(S) \). This proves Part i) of Theorem 3.1.

Further, for the matrix \( A \) defined by (3) in Step I, we have that \( \pi: N_A \longrightarrow M \) is an isomorphism. Hence the equations in (4) show that \( \pi(x, y) \in \mathbb{N}^I \) implies \((x, y) \in \mathbb{N}^I \times \mathbb{N}^{J'}\). Since \( S \) is a Krull monoid it follows that
φ(S) = \mathbb{N}^I \cap φ(Q(S)) = \mathbb{N}^I \cap M. Combing these facts, we obtain for $M_A := (\mathbb{N}^I \times \mathbb{N}^J) \cap N_A$ that $S \xrightarrow{φ} \mathbb{N}^I \cap M \xrightarrow{π^{-1}} M_A$. Since φ and π⁻¹ are monoid isomorphisms, we have that $S$ is isomorphic to the Diophantine monoid $M_A$.

Also, by (3) of Step I, the matrix $A$ is a block matrix where $A_{11} \in \mathbb{Z}^{k \times e}$ with entries $a_{ij} \in \{0, 1, \ldots, \alpha_i - 1\}$, $A_{12} = -\text{diag}(\alpha_1, \ldots, \alpha_k) \in \mathbb{N}^{k \times k}$ and $A_{21} \in \mathbb{Z}^{h \times e}$. In particular, $A \in \mathbb{Z}^{m \times n}$ with $m = k + h$, $n = k + e$. This proves Part ii) and the first part of iii) of Theorem 3.1.

It remains to show the statement about $A_{21}$ in iii). Pick $i \in I \setminus J$ and let $I_+ = \{j \in I \mid a_{ij} > 0\}$ and $I_- = \{k \in I \mid a_{ik} < 0\}$. Since by Step I the submatrix $A_{21}$ cannot have a zero-row, we must have that $I_+$ or $I_-$ is nonempty. We shall show in fact that both $I_+$ and $I_-$ must be nonempty. Suppose that $j_1 \in I_+$. Since $J \neq \emptyset$ and $i \in I \setminus J$ imply $|I| \geq 2$, there exists $j_2 \in I$, $j_2 \neq j_1$. Let $(π_j)_{j \in I}$ be the set of normalized essential states of $S$. By the isomorphism of $S$ and $M_A$ proved above, we have $x \in \mathbb{N}^I$ with $\sum_{j \in I} a_{ij}x_j = 0$ for $x_j = π_j(s)$ with $j \in I$ and $s \in S$. Since $π_{j_2}$ is essential, there exists by Lemma 2.3 an $s \in S$ with $π_{j_2}(s) = 0$ and $π_{j_1}(s) > 0$. Therefore, there must be some $j \in I$ with $a_{ij} < 0$. Thus, $I_+ \neq \emptyset$ implies $I_- \neq \emptyset$. Similarly, $I_- \neq \emptyset$ implies $I_+ \neq \emptyset$. Therefore, there exist $j_1 \in I_+$ and $k_1 \in I_-$. Suppose now $I_+ = \{j_1\}$. Since $π_{j_1}$ is essential, by Lemma 2.3 there exists $s \in S$ with $π_{j_1}(s) = 0$ and $π_{k_1}(s) > 0$. But then, for $x = (π_j(s))$, we obtain $0 = -a_{ij_1}x_{j_1} = \sum_{j \neq j_1} a_{ij}x_j < 0$, which is a contradiction.

Thus we must have that $|I_+| \geq 2$. Similarly, $|I_-| \geq 2$. This proves the statement about $A_{21}$ in Part iii) of Theorem 3.1.

Remarks 3.2.

1. The Diophantine monoid $M_A$ in Theorem 3.1 ii) consists of two different kinds of equations. The first $k$ equations are of type 1 and the remaining $h$ equations are of type 2 in the sense of [2, Theorem 1.3]. The latter type does not occur if and only if $h = 0$, or, equivalently, $r = e$ which for $e \leq 3$ must happen by Theorem 3.1 iii).

2. Concerning Krull monoids $S$ with infinitely many essential states, Step I in the proof of Theorem 3.1 can be used provided the module $M = φ(Q(S))$ meets the assumptions made there. In general, a submodule $M$ of $\mathbb{Z}^{(I)}$ does not satisfy these assumptions as the following simple example shows. Let $I = \mathbb{N}$, $(u_i)_{i \in \mathbb{N}}$ the standard basis of $\mathbb{Z}^{(N)}$, $Q = \{q_i\}_{i \in \mathbb{N}}$ and $τ: \mathbb{Z}^{(N)} → Q$ be the $\mathbb{Z}$-homomorphism defined by $τ(u_i) = q_i$ for all $i \in \mathbb{N}$. If the submodule $M = \text{Ker} τ$ of $\mathbb{Z}^{(N)}$ would satisfy the assumptions of Step I, then by (2) $Q ∼ \mathbb{Z}^{(N)}/M$ should be isomorphic to a direct sum of copies of $\mathbb{Z}$ and $\mathbb{Z}_n$ (which is not the case). If, however, for $I$ infinite $\text{Cl}(S)$ is finitely generated, then one
can argue that $S$ is isomorphic to a monoid $M_A$ where the number of rows of $A$ is rank $\text{Cl}(S)$ and the “number” of columns of $A$ is given by the cardinality of $I$ augmented by a finite set. Similarly, if $\text{Cl}(S)$ is free, then $S$ is isomorphic to a monoid $M_A$ with $|I\setminus J|$ equations in $|I|$ unknowns.

From Theorem 3.1 we obtain the following characterization of Diophantine monoids.

**Corollary 3.3.** For a reduced monoid $S$ the following statements are equivalent:

i) $S$ is a Krull monoid with finitely many essential states.

ii) $S$ is isomorphic to a Diophantine monoid.

iii) $S$ is isomorphic to a monoid $W \cap \mathbb{N}^n$ for a vector subspace $W$ of $\mathbb{Q}^n$ and some $n \geq 1$.

iv) $S$ is isomorphic to a full and expanded submonoid $T$ of $\mathbb{N}^n$ (i.e., $x, y \in T$, $y - x \in \mathbb{N}^n$ imply $y - x \in T$, and $kd \in T$ for $k \geq 1$, $z \in \mathbb{N}^n$, implies $z \in T$).

v) $S$ is root-closed and finitely generated.

**Proof.** i) $\Rightarrow$ ii) follows directly from Theorem 3.1. Obviously, ii) $\Rightarrow$ iii). Since $T = W \cap \mathbb{N}^n$ is full and expanded, iii) $\Rightarrow$ iv). From iv) it follows by Dickson’s Lemma (see [14, Theorem 5.1]) that $T$, and hence $S$, is finitely generated. Since $T$ is expanded, $S$ must be root-closed (i.e., $kd \in S$ for $k \geq 1$, $z \in Q(S)$ implies $z \in S$). This yields iv) $\Rightarrow$ v). Finally, v) $\Rightarrow$ i) follows from a well-known theorem of Halter-Koch [7, Theorem 5] and the main result in [13], by which a root-closed and finitely generated monoid is a Krull monoid described by finitely many states.

**Remarks 3.4.**

1. In [13, Corollary 1] a description of the class group is given by employing terminology and methods from convex analysis. Using results from [13], in [10, Proposition 2] it is shown that a monoid $S$ satisfying $Q(S) = \mathbb{Z}^n$ is isomorphic to the monoid of nonnegative solutions of a homogeneous system of integral linear equations if and only if $S$ is a Krull monoid which holds if any only if $S$ is finitely generated and root-closed. The method used in the proof of Theorem 3.1 is direct and does not involve convex analysis. Moreover, Theorem 3.1 describes the representing matrix $A \in \mathbb{Z}^{m \times n}$ in terms of the class group and the number of essential states of $S$. This yields in particular that $m = \text{rank Cl}(S)$.

2. Corollary 3.3 applies in particular to additive submonoids $S$ of $\mathbb{N}^n$. For this case it has been shown in [8, Lemma 3] that $S$ is the set of solutions in the nonnegative integers of a system of homogeneous linear equations with rational coefficients if and only if $S = W \cap \mathbb{N}^n$.
for a vector subspace $W$ of $\mathbb{Q}^n$ or, equivalently, $S$ is full and expanded. An additive submonoid $S$ of $\mathbb{N}^n$ is called a full affine semigroup if $S = M \cap \mathbb{N}^n$ for a subgroup $M$ of $\mathbb{Z}^n$ or, equivalently, $S = Q(S) \cap \mathbb{N}^n$ (see [14, Chapter 7] for more about these semigroups). Obviously, such a semigroup is root-closed, but it need not be expanded as the example $S = 2\mathbb{N}$ shows. Furthermore, such a semigroup is a Krull monoid with finitely many essential states. From Corollary 3.3 one concludes that an arbitrary reduced monoid is a Krull monoid with finitely many essential states if and only if it is isomorphic to a full affine semigroup.

By Theorem 3.1 Part ii), we know how to describe an arbitrary Krull monoid $S$ having a finite number of essential states by a particular matrix adapted to $S$. The following Lemma addresses, conversely, a Krull monoid given by an arbitrary matrix. For a matrix $A$, let $\text{im } A$ denote the image of the mapping induced by $A$.

**Lemma 3.5.** Let $S \simeq M_A$ with $A \in \mathbb{Z}^{m \times n}$, $m \geq 1, n \geq 1$ and weight $w(M_A) = 1$. Arrange $A$ such that the first $e$ canonical projections are exactly the normal essential states of $M_A$ and let $A = [A' A'']$ where $A' \in \mathbb{Z}^{m \times e}, A'' \in \mathbb{Z}^{m \times (n-e)}$. The following statements hold:

i) $\text{Cl } (S) \simeq \text{im } A'/\text{im } A' \cap \text{im } A''$.

ii) $n = \text{rank } A + \text{rank } Q(S) = \text{rank } A'' + e$.

iii) $\text{rank } A' - \text{rank } (\text{im } A' \cap \text{im } A'') = \text{free rank } \text{Cl } (S) \leq \text{rank } \text{Cl } (S) \leq \text{rank } A'$.

iv) $\text{rank } \text{Cl } (S) \leq m$ and $\text{rank } \text{Cl } (S) + \text{rank } Q(S) \leq n$.

In particular, if $A' = A$, then one has $\text{Cl } (S) \simeq \mathbb{Z}^k, k = \text{rank } A$ and $\text{rank } \text{Cl } (S) + \text{rank } Q(S) = n$.

**Proof.** It can be assumed that $S = M_A$. By assumption the divisor theory $\varphi : Q(S) \rightarrow \mathbb{Z} = EA$ is given by $\varphi(y_1, \ldots, y_n) = (y_1, \ldots, y_e)$. From Lemma 2.6 it follows that $Q(S) = \text{Ker } A$.

i) Consider $f : \text{im } A' \rightarrow \text{Cl } (S)$ defined by $A'x \mapsto x + \varphi(Q(S))$ for $x \in \mathbb{Z}^e$. The mapping $f$ is well-defined. For, if $A'x = 0$, then $y = \begin{bmatrix} x \\ 0 \end{bmatrix} \in Q(S)$ and, hence, $x = \varphi(y) \in \varphi(Q(S))$. Obviously, $f$ is surjective. Furthermore, if $A'x \in \text{Ker } f$ then $x \in \varphi(Q(S))$ and $A'x + A''y = 0$ for some $y \in \mathbb{Z}^{n-e}$. This shows $\text{Ker } f \subseteq \text{im } A' \cap \text{im } A''$. Conversely, if $A'x \in \text{im } A''$, then $A'x + A''y = 0$ for some $y \in \mathbb{Z}^{n-e}$ and $x = \varphi \begin{bmatrix} x \\ y \end{bmatrix} \in \varphi(Q(S))$. This shows $\text{im } A' \cap \text{im } A'' \subseteq \text{Ker } f$. Thus, $\text{Ker } f = \text{im } A' \cap \text{im } A''$ and we have the short exact sequence

$$0 \rightarrow \text{im } A' \cap \text{im } A'' \xrightarrow{i} \text{im } A' \xrightarrow{f} \text{Cl } (S) \rightarrow 0,$$
where \( i \) denotes the embedding. Therefore, \( \text{Cl}(S) \simeq \text{im} A'/\left(\text{im} A' \cap \text{im} A''\right) \).

ii) Obviously, \( n = \text{rank} \text{im} A + \text{rank} \text{Ker} A \). Since \( \text{Ker} A = Q(S) \) it follows that \( n = \text{rank} A + \text{rank} Q(S) \). Also, \( n-e = \text{rank} \text{im} A'' + \text{rank} \text{Ker} A'' \). We show that \( \text{Ker} A'' = 0 \) which proves ii). If \( z \in \text{Ker} A'' \), then \( y, -y \in \text{Ker} A \) for \( y = \begin{bmatrix} 0 \\ z \end{bmatrix} \). It follows that \( y, -y \in Q(M_A) \) and \( \pi_i(-y) = \pi_i(y) \in \mathbb{N} \) for every \( i \in I = \{1, 2, \ldots, e\} \). Since \( \pi_i \), \( i \in I \) are the essential projections of \( M_A \), we obtain that \( y, -y \in M_A \). Since \( M_A \subseteq \mathbb{N}^n \) we must have that \( y = 0 \) and, hence, \( z = 0 \).

iii) Follows directly from i).

iv) From iii) we have that

\[
\text{rank} \text{Cl}(S) \leq \text{rank} A' \leq \min\{m, \text{rank} A\}
\]

and using ii) we obtain

\[
\text{rank} \text{Cl}(S) + \text{rank} Q(S) \leq \text{rank} A + \text{rank} Q(S) = n.
\]

The result for \( A' = A \) follows immediately since in this case the proof for i) shows that \( \text{Ker} f = \{0\} \).

Concerning the representation of a monoid by a matrix, it is a natural question to ask for a matrix with a minimal number of rows and columns, respectively.

**Definition 3.6.** For a monoid \( S \) the **row degree** of \( S \) is defined by

\[
d_r(S) = \min\{m \in \mathbb{N} \mid S \simeq M_A \text{ for } A \in \mathbb{Z}^{m \times n}\}
\]

and the **column degree** of \( S \) is defined by

\[
d_c(S) = \min\{n \in \mathbb{N} \setminus \{0\} \mid S \simeq M_A \text{ for } A \in \mathbb{Z}^{m \times n}\}.
\]

**Remark 3.7.** Since we defined for \( A \in \mathbb{Z}^{m \times n} \) that \( M_A = \mathbb{N}^n \) if \( m = 0 \), it follows that \( d_r(S) = 0 \) if and only if \( S \simeq \mathbb{N}^n \) for some \( n \). Also, for \( S \simeq \mathbb{N}^n \) it follows from Lemma 3.5 ii) that \( d_c(S) = n \).

From Theorem 3.1 together with Lemma 3.5, we obtain the theorem announced in the Introduction.

**Theorem 3.8.** For a reduced monoid \( S \) which is root-closed and finitely generated the following holds:

i) **Row degree and column degree of** \( S \) **are finite and**

\[
d_r(S) = \text{rank} \text{Cl}(S), \quad d_c(S) = \text{rank} \text{Cl}(S) + \text{rank} Q(S).
\]

ii) **There exists a matrix** \( A \in \mathbb{Z}^{m \times n} \) **with** \( m = d_r(S) \) **and** \( n = d_c(S) \) **such that** \( S \) **is isomorphic to the Diophantine monoid** \( M_A \).

**Proof.** By Corollary 3.3, \( S \) is a Krull monoid with finitely many essential states.
i) By Theorem 3.1, \( d_r(S) \leq \text{rank} \, \text{Cl}(S) < \infty \) and \( d_c(S) \leq \text{rank} \, \text{Cl}(S) + \text{rank} \, Q(S) < \infty \). Suppose \( S \cong M_A \) for \( A \in \mathbb{Z}^{m \times n} \). We may assume that \( w(M_A) \neq 0 \) and, by Lemma 2.6, that \( S \cong M_{AD} \) with \( w(M_{AD}) = 1 \). From Lemma 3.5 iv) we obtain \( \text{rank} \, \text{Cl}(S) \leq m \) and \( \text{rank} \, \text{Cl}(S) + \text{rank} \, Q(S) \leq n \). This proves i).

ii) Follows immediately from i) and Theorem 3.1.

\[ \square \]

4. Some consequences and examples.

The results of the previous section can be used to check if a Diophantine monoid is factorial or half-factorial.

**Proposition 4.1.** Let \( S \cong M_A \) and \( A = [A' \, A''] \) be given as in Lemma 3.5.

i) \( S \) is factorial if and only if each column of \( A' \) is in the column space of \( A'' \).

ii) If any two columns of \( A' \) not in the column space of \( A'' \) have their difference in the column space of \( A'' \), then \( S \) is half-factorial.

**Proof.** It is well-known that any Krull monoid \( S \) is factorial if and only if \( \text{Cl}(S) = \{0\} \) and that \( S \) is half-factorial if \( \text{Cl}(S) \cong \mathbb{Z}_2 \) (cf. [12, Proposition 2]).

i) By Lemma 3.5, \( \text{Cl}(S) = \{0\} \) if and only if \( \text{im} \, A' \subseteq \text{im} \, A'' \). Therefore, \( S \) is factorial if and only if each column of \( A' \) is in the \( \mathbb{Z} \)-span of the columns of \( A'' \).

ii) If all columns of \( A' \) are in \( \text{im} \, A'' \) then by i) \( S \) is factorial and, a fortiori, half-factorial. Suppose there is a column \( A'_{i_0} \) of \( A' \) which is not in \( \text{im} \, A'' \). By assumption, for any column \( A'_i \notin \text{im} \, A'' \) it holds that \( A'_i - A'_{i_0} \in \text{im} \, A'' \). Therefore, for each column \( A'_i \) of \( A' \) either \( A'_i \in \text{im} \, A'' \) or \( A'_i \notin \text{im} \, A'' + \text{im} \, A'' \). Lemma 3.5 implies that \( \text{Cl}(S) \cong \mathbb{Z}_2 \) and, hence, \( S \) is half-factorial.

\[ \square \]

**Remark 4.2.** Neither the condition given in ii) nor the condition \( \text{Cl}(S) \cong \mathbb{Z}_2 \) are necessary for half-factoriality. This is so even for \( m = 1 \); see [2] for various sufficient or necessary conditions of half-factoriality in this particular case.
In [2, Theorem 1.3], the current authors found a formula for computing the class group of a Diophantine monoid given by just one equation. The results of Section 3 yield a simpler proof of this formula, as well as an additional characterization of such monoids.

**Proposition 4.3.**

i) A monoid $S$ is isomorphic to a nonfactorial Diophantine monoid given by just one equation if and only if $S$ is reduced, root-closed, and finitely generated with $\text{rank} \; \text{Cl}(S) = 1$.

ii) Let $S = M_A$ with $A = [a_1a_2\ldots a_n] \in \mathbb{Z}^{1 \times n}$ and, without restriction, $a_i \neq 0$ for all $i$, not all $a_i$ of equal sign and $\gcd(a_1,\ldots,a_n) = 1$. There are exactly two possible cases:

a) $\text{Cl}(S) \simeq \mathbb{Z}_\alpha$ with $\alpha \in \mathbb{N}$. This case occurs if and only if all $a_i$ except one, say $a_n$, are of equal sign. Thus we have $\alpha = \frac{|a_n|}{c}$, where $c = \prod_{i=1}^{n-1} c_i$ and $0 < c_i = \gcd(|a_j| \mid j \neq i)$ for $1 \leq i \leq n$.

b) $\text{Cl}(S) \simeq \mathbb{Z}$. This case occurs if and only if there are at least two $a_i$ with positive sign and at least two $a_i$ with negative sign.

**Proof.** The proof of i) follows directly from Theorem 3.8. For ii), let $\tilde{a}_j = \frac{a_j}{d_j}$ with $d_j = \prod_{i \neq j} c_i$. For $\tilde{A} = [\tilde{a}_1\ldots \tilde{a}_n]$ the mapping $\psi: M_A \rightarrow M_{\tilde{A}}$, $\psi_i(x_1,\ldots,x_n) = \frac{x_i}{c_i}$ is a monoid isomorphism. Hence $\text{Cl}(M_A) \simeq \text{Cl}(M_{\tilde{A}})$. Furthermore, $w(M_{\tilde{A}}) = 1$. Thus, we can assume that $c_i = 1$ for all $i$, $c = 1$ and $w(M_A) = 1$.

a) Let $a_i > 0$ for $1 \leq i \leq n-1$ and $a_n < 0$. Then $A = [A' A'']$ with $A' = [a_1 \ldots a_{n-1}]$ and $A'' = a_n = -\alpha$. Since $c_n = 1$, we have that $1 \in \text{im} \; A'$ and, hence, $\text{im} \; A' = \mathbb{Z}$. Obviously, $\text{im} \; A'' = \alpha \mathbb{Z}$ and from Lemma 3.5 i) we obtain $\text{Cl}(M_A) \simeq \mathbb{Z}_\alpha$.

b) From the assumption in b), all projections must be essential and hence, $A = A'$. From $\gcd(a_1,\ldots,a_n) = 1$ it follows that $\text{im} \; A' = \mathbb{Z}$ and Lemma 3.5 yields $\text{Cl}(M_A) \simeq \mathbb{Z}$.

The following example illustrates the concept of row degree and column degree, respectively and the statements made in Theorem 3.8.

**Example 4.4.** Consider the Diophantine monoid

$$S = \{x \in \mathbb{N}^5 \mid x_1 + x_2 = x_4 + x_5, \; x_2 + x_5 = x_3 + x_4\}.$$
Obviously, \( S = M_A \) for \( A = \begin{bmatrix} 1 & 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 & 1 \end{bmatrix} \in \mathbb{Z}^{2 \times 5} \). All canonical projections are normal and by Lemma 2.3 the projections \( \pi_1 \) to \( \pi_4 \) are essential while \( \pi_5 \) is not. We have \( w(M_A) = 1 \) and \( A = [A'A''] \) with \( A' = \begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & -1 & -1 \end{bmatrix} \) and \( A'' = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \). It follows that \( \text{im} \ A' = \mathbb{Z} \oplus \mathbb{Z} \), \( \text{im} \ A'' = \mathbb{Z} \begin{bmatrix} -1 \\ +1 \end{bmatrix} \), and, by Lemma 3.5 i), \( \text{Cl}(M_A) \simeq \mathbb{Z} \). Since \( \text{rank} \ Q(M_A) = 3 \), it follows from Theorem 3.8 i) that \( d_c(S) = 1 \) and \( d_c(S) = 1 + 3 = 4 \). By Theorem 3.8 ii) there exists a matrix \( B \in \mathbb{Z}^{1 \times 4} \) such that \( M_A \) is isomorphic to the Diophantine monoid \( M_B \) given by the smaller matrix \( B \). Indeed, one can reduce the given system of two equations to just one equation as follows. Eliminating \( x_5 = x_1 + x_2 - x_4 \) yields \( x_2 + (x_1 + x_2 - x_4) = x_3 + x_4 \). That is \( x_1 + 2x_2 - x_3 - 2x_4 = 0 \). One has to make sure, however, that for any solution in \( \mathbb{N} \) of the latter equation, it automatically holds that \( x_5 = x_1 + x_2 - x_4 \geq 0 \). This is obvious in case of \( x_2 \geq x_4 \). If \( x_2 \leq x_4 \), then one obtains \( x_1 + x_2 - x_4 \geq x_1 + 2(x_2 - x_4) = x_3 \geq 0 \).

By Theorem 3.8, one can easily check if a given Diophantine monoid can be described by a smaller matrix without actually carrying out the elimination procedure as in the above example. This latter process might be quite difficult in general.

The following example also illustrates Theorem 3.8, and essentially covers all class group possibilities of rank 2. In contrast to Example 4.4, here there is no possible smaller description of the given monoids.

**Example 4.5.** Let \( G \) be a finitely generated abelian group of rank 2. We show how to construct a Diophantine monoid \( S \) defined by 2 equations such that \( \text{Cl}(S) \simeq G \). There are three cases to consider.

(a) Suppose \( G \simeq \mathbb{Z}_n \oplus \mathbb{Z}_{kn} \) for positive integers \( n > 1 \) and \( k \geq 1 \). Let

\[
S = \{ x \in \mathbb{N}^5 \mid x_1 + x_3 = nx_4, \ x_2 + x_3 = knx_5 \}.
\]

Obviously, \( S = M_A \) for \( A = \begin{bmatrix} 1 & 0 & 1 & -n & 0 \\ 0 & 1 & 1 & 0 & -kn \end{bmatrix} \in \mathbb{Z}^{2 \times 5} \). It is easy to check that all canonical projections are normal and, using Lemma 2.3, that the projections \( \pi_1, \pi_2, \pi_3 \) are essential while \( \pi_4, \pi_5 \) are not. Thus, \( w(M_A) = 1 \) and \( A = [A'A''] \) with \( A' = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \) and \( A'' = \begin{bmatrix} -n & 0 \\ 0 & -kn \end{bmatrix} \). It follows
that \( \text{im} A' = \mathbb{Z} \oplus \mathbb{Z} \) and \( \text{im} A'' = n\mathbb{Z} \oplus kn\mathbb{Z} \). Therefore, by Lemma 3.5 i), \( \text{Cl} (S) \simeq (\mathbb{Z} \oplus \mathbb{Z})/(n\mathbb{Z} \oplus kn\mathbb{Z}) \simeq \mathbb{Z}_n \oplus \mathbb{Z}_{kn} \). Since \( \text{rank} Q(S) = 3 \), by Theorem 3.8 we have that \( d_r(S) = 2 \) and \( d_c(M) = 2 + 3 = 5 \). Therefore, the Diophantine monoid \( S \) can neither be described by one equation alone, nor by less than five variables.

(b) Suppose \( G \simeq \mathbb{Z} \oplus \mathbb{Z}_n \) for some positive integer \( n > 1 \). Let

\[
S = \{ x \in \mathbb{N}^6 \mid x_1 + x_2 - x_3 - x_4 = 0, x_1 + x_3 + x_5 = nx_6 \}.
\]

Clearly, \( S = M_A \) for \( A = \begin{bmatrix} 1 & 1 & -1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & -n \end{bmatrix} \in \mathbb{Z}^{2 \times 6} \). As in (a), \( w(M_A) = 1 \), the projections \( \pi_1, \pi_2, \pi_3, \pi_4 \) and \( \pi_5 \) are essential and \( \pi_6 \) is not essential. Hence, if \( A' = \begin{bmatrix} 1 & 1 & -1 & -1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \) and \( A'' = \begin{bmatrix} 0 \\ -n \end{bmatrix} \), then \( \text{im} A' = \mathbb{Z} \oplus \mathbb{Z} \) and \( \text{im} A'' = n\mathbb{Z} \). Thus, \( \text{Cl} (S) \simeq (\mathbb{Z} \oplus \mathbb{Z})/n\mathbb{Z} \simeq \mathbb{Z} \oplus \mathbb{Z}_n \).

Since \( \text{rank} Q(S) = 4 \), we have that \( d_r(S) = 2 \) and \( d_c(M) = 2 + 4 = 6 \).

(c) Suppose that \( G \simeq \mathbb{Z} \oplus \mathbb{Z} \). Let

\[
S = \{ x \in \mathbb{N}^6 \mid x_1 + x_3 - x_4 - x_6 = 0, x_2 + x_3 - x_5 - x_6 = 0 \}.
\]

Clearly, \( S = M_A \) for \( A = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & -1 \\ 0 & 1 & 1 & 0 & -1 & -1 \end{bmatrix} \in \mathbb{Z}^{2 \times 6} \). As above, \( w(M_A) = 1 \), but here all projections are essential. Thus, \( A' = A \) and Lemma 3.5 yields \( \text{Cl} (S) \simeq \mathbb{Z}^2 \). Since \( \text{rank} Q(S) = 4 \), we have that \( d_r(S) = 2 \) and \( d_c(M) = 2 + 4 = 6 \).

To close Section 4, we present an algorithm which reflects the techniques developed in Sections 2 and 3. The method is based on having the complete set of minimal solutions over the nonnegative integers of the system \( Ax = 0 \). An algorithm for that computation, by García-Sánchez and Rosales, can be found in [14, Section 3, p. 80]. Before proceeding, we will require one additional result. Let \( M \) be a finitely generated submonoid of \((\mathbb{N}^t, +)\) and \( u_1, \ldots, u_t \) the irreducible elements of \( M \). Let \( V \) and \( W \) be the \( \mathbb{Q} \)-vector spaces generated by \( u_1, \ldots, u_t \) and \( u_2 - u_1, \ldots, u_t - u_1 \) respectively. Then \( \text{dim} V - \text{dim} W \leq 1 \) and \( V = W \) if and only if \( u_1 \in W \).

**Lemma 4.6.** Let \( M, V \) and \( W \) be as above. The following statements are equivalent:

i) \( M \) is half factorial
ii) \( \text{dim} V = 1 + \text{dim} W \).

**Proof.** For i) \( \Rightarrow \) ii), assume that \( u_1 = \sum_{j=2}^t \alpha_j (u_j - u_1) \in W \) with each \( \alpha_j \in \mathbb{Q} \). Then \( k_1 u_1 = \sum_{j=2}^t k_j (u_j - u_1) \) for suitable \( k_1, \ldots, k_t \in \mathbb{Z} \), \( k_1 \neq 0 \). Since \( (k_1 + \sum_{j=2}^t k_j) u_1 - \sum_{j=2}^t k_j u_j = 0 \), the monoid \( M \) is not half factorial.
For ii) ⇒ i), let \( r_1, \ldots, r_t \in \mathbb{Z} \) with
\[
\sum_{j=1}^t r_j u_j = 0 = \sum_{j=2}^t r_j (u_j - u_1) + \left( \sum_{j=1}^t r_j \right) u_1.
\]
Since \( u_1 \not\in W \) it follows that \( \sum_{j=1}^t r_j = 0 \). \( \square \)

**Algorithm 4.7.** Assume that \( A \in \mathbb{Z}^{m \times n} \). The following algorithm calculates the class group of \( M_A \) and determines whether or not \( M_A \) is half-factorial. Suppose that \( (u_{\tau}|1 \leq \tau \leq t) \neq \emptyset \) is the family of all irreducible elements of the monoid \( M_A \). Let \( C := (u_{\tau}) \in \mathbb{N}^{n \times t} \) be the matrix with column vectors \( u_{\tau} \) and row vectors \( v_i := (u_{i\tau}) \in \mathbb{N}_{1 \times t}, 1 \leq i \leq n \). Let \( C_1 := (u_{i\tau} - u_{i1}) \in \mathbb{Z}^{n \times t} \). By Lemma 4.6, \( M_A \) is not half factorial if and only if \( \text{im}_Z C = \text{im}_Z C_1 \). The calculation of the class group proceeds as follows.

**I) Reduction of the system.** Let \( v_i := (u_{i\tau}) \in \mathbb{N}^{1 \times t}, 1 \leq i \leq n \), be the \( i \)-th row vector of the matrix \( C \).

Step 1): For all \( i \in \{1, \ldots, n\} \), if \( v_i = 0 \), then cancel the \( i \)-th row of \( C \).

Step 2): For all \( i, j \in \{1, \ldots, n\} \) with \( i < j \), if \( \lambda v_i = v_j \) for some \( \lambda \in \mathbb{Q} \), then cancel the \( j \)-th row of \( C \).

After these canceling steps, we denote the new matrix again with \( C \). Then \( C \in \mathbb{Z}^{\tilde{n} \times t} \) for some \( \tilde{n} \leq n \).

Step 3): For all \( j \in \{1, \ldots, \tilde{n}\} \), calculate \( c_j := \gcd (u_{j\tau}|1 \leq \tau \leq t) \in \mathbb{N} \) and replace the row vector \( v_j \) of \( C \) by \( \frac{1}{c_j} v_j \).

**II) Determining the essential states.**

Step 4): For every \( i \in \{1, \ldots, \tilde{n}\} \), define \( J_i := \{ \tau \in \{1, \ldots, t\}| u_{i\tau} = 0 \} \). If \( J_i \neq \emptyset \), calculate the sums
\[
v_j^{(i)} := \sum_{\tau \in J_i} u_{j\tau}
\]
for all \( j \in \{1, \ldots, \tilde{n}\}, j \neq i \). Define
\[
I := \{ i \in \{1, \ldots, \tilde{n}\}| J_i \neq \emptyset, v_j^{(i)} \neq 0 \text{ for all } j \in \{1, \ldots, \tilde{n}\}, j \neq i \}
\]
and \( e := |I| \). The projections \( \pi_i, i \in I \), are exactly the essential states of \( M \).

**III) The final calculation.**

Step 5): Let \( \tilde{C} \in \mathbb{N}^{e \times t} \) be the matrix with the row vectors \( v_i, i \in I \). Transform \( \tilde{C} \) into its Smith normal form, using e.g., the algorithm given in [14].
From the Smith normal form one gets $k := \text{rank } \tilde{C}$ and $r \geq 0$ elementary divisors $\alpha_i \geq 2$ of $\tilde{C}$ with $\alpha_{i+1} | \alpha_i$ for $1 \leq i \leq r - 1$ if $r \geq 2$. Then

$$Cl(M) = \mathbb{Z}^{e-k} \oplus \mathbb{Z}/\alpha_1 \mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/\alpha_r \mathbb{Z}.$$ 

Acknowledgment. The authors would like to acknowledge many useful suggestions made by an anonymous referee on an earlier draft of the paper.

References


Received September 25, 2000 and revised November 9, 2001. Part of this work was completed while the first author was on an Academic Leave granted by the Trinity University Faculty Development Committee. Part of this work was completed while the second author was visiting Trinity University during the fall of 1999 as the Eva and O.R. Mitchell Visiting Distinguished Professor of Mathematics.

Trinity University  
Department of Mathematics  
715 Stadium Drive  
San Antonio, Texas 78212-7200  
E-mail address: schapman@trinity.edu

Universität Bremen  
Fachbereich Mathematik/Informatik  
DW-2800 Bremen, Germany  
E-mail address: krause@math.uni-bremen.de

Universität Bremen  
Fachbereich Mathematik/Informatik  
DW-2800 Bremen, Germany  
E-mail address: oel@math.uni-bremen.de
THE GROUP OF ISOMETRIES OF A FINSLER SPACE

SHAOQIANG DENG AND ZIXIN HOU

We prove that the group of isometries of a Finsler space is a Lie transformation group on the original manifold. This generalizes the famous result of Myers and Steenrod on a Riemannian manifold and makes it possible to use Lie theory on the study of Finsler spaces.

Introduction.

Let \((M, F)\) be a Finsler space, where \(F\) is positively homogeneous but not necessary absolutely homogeneous. As in the Riemannian case, we have two kinds of definitions of isometry on \((M, F)\). On one hand, we can define an isometry to be a diffeomorphism of \(M\) onto itself which preserves the Finsler function. On the other hand, since on \(M\) we still have the definition of distance function (although generically it is not a real distance), we can define an isometry of \((M, F)\) to be a mapping of \(M\) onto \(M\) which keeps the distance of each pair of points of \(M\).

The equivalence of the two definitions of isometry in the Riemannian case is a famous result of Myers and Steenrod. They used this result to prove that the group of isometries of a Riemannian manifold is a Lie transformation groups on the original manifold \([5]\). This result plays a fundamental role on the theory of homogeneous Riemannian manifolds. Since then, many different proofs were provided, cf., e.g., Palais \([6]\), S. Kobayashi \([4]\).

In this paper we prove that the two definitions of isometry are equivalent for a Finsler space. Then we prove that the group of isometries has a differentiable structure which turns it into a Lie transformation on the manifold. This result makes it possible to use Lie theory on the study of Finsler spaces.

In this paper, Finsler structure \(F\) is only assumed to be positively homogeneous but not necessary absolutely homogeneous. We will not point out this each time. For a mapping \(\phi\) of a manifold \(M\), we use \(d\phi\) to denote its differential. If \(p \in M\), \(d\phi|_p\) will denote the differential of \(\phi\) at \(p\). The notations of forward and backward metric ball in a Finsler spaces comes from the newly published book by D. Bao, S.S. Chern and Z. Shen \([1]\).
1. A result on distance function.

Let \((M, F)\) be a Finsler space, \(d\) be the distance function of \((M, F)\). We first need to prove a result on the distance function.

**Lemma 1.1.** Let \(x \in M\). Then for any \(\epsilon > 0\), there exists a neighborhood \(U\) of the original of \(T_x(M)\) such that \(\exp_x\) is a \(C^1\)-diffeomorphism from \(U\) onto its image and for any \(A, B \in U, A \neq B\), and any \(C^1\) curve \(\sigma_0(s), 0 \leq s \leq 1\), connecting \(A\) and \(B\) which satisfies \(\sigma_0(s) \in U\) and \(\dot{\sigma}_0(s) \neq 0, s \in [0, 1]\), we have

\[
\left| \frac{L(\sigma)}{L(\sigma_0)} - 1 \right| \leq \epsilon,
\]

where \(L(\cdot)\) denotes the arc length of a curve and \(\sigma(s) = \exp_x \sigma_0(s)\).

**Proof.** Let \(B_x(r) = \{ A \in T_x(M) | F(x, A) < r \}\) be a tangent ball in \(T_x(M)\) such that \(\exp_x = \exp\) is a \(C^1\)-diffeomorphism from \(B_x(r)\) onto \(B_x^+(r) = \{ w \in M | d(x, w) \leq r \}\) (cf. [1]). Assume \(A, B \in B_x(r), A \neq B\). Let \(\sigma_0(s), 0 \leq s \leq 1\) be a \(C^1\) curve connecting \(A\) and \(B\) and \(\forall s, \sigma_0(s) \in B_x(r)\) and \(\dot{\sigma}_0(s) \neq 0\).

Then we can write the velocity vector of \(\sigma_0(s)\) as \(\dot{\sigma}_0(s) = t(s)X(s)\), where \(X(s)\) satisfies \(F(x, X(s)) = \frac{\epsilon}{2}, \forall s,\) and \(t(s) \geq 0\) is a continuous function on \([0, 1]\). Therefore the arc length of \(\sigma_0\) is

\[
L(\sigma_0) = \int_0^1 t(s)F(x, X(s))ds.
\]

Denote \(X_1(s) = d(\exp_x)_{\sigma_0(s)}X(s)\). Then the velocity vector of the curve \(\sigma(s) = \exp_x(\sigma_0(s)), 0 \leq s \leq 1\) is

\[
\dot{\sigma}(s) = d(\exp_x)_{\sigma_0(s)}(t(s)X(s)) = t(s)d(\exp_x)_{\sigma_0(s)}(X(s)) = t(s)X_1(s).
\]

Therefore, the arc length of \(\sigma\) is

\[
L(\sigma) = \int_0^1 t(s)F(\sigma(s), X_1(s))ds.
\]

Now we select a neighborhood \(V_1\) of \(x\) in \(M\) with compact closure which is contained in \(B_x^+(r)\) and fix a coordinate system \((x_1, x_2, \ldots, x_n)\) in \(V_1\). Let \(U_1 = \exp^{-1}V_1\). Suppose \(\sigma_0 \subset U_1\). Denote by \(M(s)\) the matrix of \(d(\exp_x)_{\sigma_0(s)}\) under the base \(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\). Given any positive number \(\delta < \frac{\epsilon}{2}\) since \(d(\exp_x)_{\sigma_0} = I_n\) and \(\exp\) is \(C^1\) smooth, there exists a neighborhood \(U_2 \subset U_1\) of the original of \(T_x(M)\) such that for any \(C^1\) curve \(\sigma_0\) satisfying \(\sigma_0(s) \in U_2, \forall s\), we have

\[
\|M(s) - I\| < \frac{\delta}{n}, \quad 0 \leq s \leq 1,
\]
where $\| \cdot \|$ denote the maximum of the absolute value of the entries of a matrix. Write $X(s)$ and $X_1(s)$ as:

$$X(s) = \sum_{j=1}^{n} y_j(s) \frac{\partial}{\partial x_j} x;$$

$$X_1(s) = \sum_{j=1}^{n} y'_j(s) \frac{\partial}{\partial x_j} \sigma(s).$$

Then we have

$$|y'_j(s) - y_j(s)| < \delta, \quad 1 \leq j \leq n.$$

Consider the set

$$C_0 = \left\{ (w, (d(exp_x))_W y) \mid w \in V_1, W = \exp^{-1} w, \right.$$ 

$$\left. y \in T_W(Tx(M)) = Tx(M), F(x, y) = \frac{r}{2} \right\}.$$

Since exp is $C^1$ smooth, the closure of $C_0$ is compact. Hence the Finsler function $F$ is bounded on $C_0$. Suppose $F < r_1$ on $C_0, r_1 > 0$. Now write the Finsler function $F(w, y)$ as $F(w, y_1, y_2, \ldots, y_n)$ for $y = \sum_{j=1}^{n} y_j \frac{\partial}{\partial x_j} w$. Consider the closure $D_1$ of the set $D_0 = \{(w, y) \in TM \mid w \in V_1, F(x, y) \leq \frac{r}{2} + r_1 \}$. Since $F$ is continuous and $D_1$ is compact, $F$ is uniformly continuous on $D_1$. Therefore for the given $\epsilon > 0$, there exists $\delta_1 > 0$ and a neighborhood $V_2 \subset V_1$ of $x$ such that for any $w \in V_2, |y_j - y'_j| < \delta_1, j = 1, 2, \ldots, n, F(x, y_1, y_2, \ldots, y_n) < \frac{r}{2} + r_1, F(w, y'_1, \ldots, y'_n) < \frac{r}{2} + r_1, \text{ we have}$

$$|F(x, y_1, y_2, \ldots, y_n) - F(w, y'_1, y'_2, \ldots, y'_n)| < \frac{r}{2} \epsilon.$$

Therefore if we select the above $\delta$ such that $\delta < \delta_1$. Then for the corresponding $U_2$ and any $C^1$ curve $\sigma_0, \sigma_0 \subset U_2 \cap (\exp)^{-1} V_2$, we have

$$\left| \frac{L(\sigma)}{L(\sigma_0)} - 1 \right| = \frac{\int_{0}^{1} t(s)(F(x, X(s)) - F(\sigma(s), X_1(s))) ds}{\int_{0}^{1} t(s)F(x, X(s)) ds} \leq \frac{\int_{0}^{1} t(s)|F(x, X(s)) - F(\sigma(s), X_1(s))| ds}{r \int_{0}^{1} t(s) ds} \leq \frac{\epsilon}{2} \frac{\int_{0}^{1} t(s) ds}{r \int_{0}^{1} t(s) ds} = \epsilon.$$
Theorem 1.2. Let $x \in M$ and $B_x(r)$ be a tangent ball of $T_x(M)$ such that $\exp_x$ is a $C^1$ diffeomorphism from $B_x(r)$ onto $B_x^+(r)$. For $A, B \in B_x(r)$, $A \neq B$, let $a = \exp_x A$, $b = \exp_x B$. Then we have
\[
\frac{F(x, A - B)}{d(a, b)} \to 1
\]
as $(A, B) \to (0, 0)$.

Proof. Let $B_x^+_x(r) = \{ w \in M | d(w, x) < r \}$. Suppose $r$ is so small that each pair of points in $B_x^+_x(r) \cap B_x^-(x)$ can be joined by a unique minimal geodesic contained in $B_x^+(r)$ (cf. [1]). Let $\Gamma_0(s), 0 \leq s \leq 1$ be the line segment connecting $A$ and $B$, and $\Gamma(s) = \exp_x \Gamma_0(s)$. By Lemma 1.1, we have
\[
\frac{L(\Gamma_0)}{L(\Gamma)} = \frac{F(x, A - B)}{L(\Gamma)} \to 1
\]
as $(A, B) \to (0, 0)$. Now let $a = \exp_x A$, $b = \exp_x B$. Suppose $a, b \in B_x^+(x) \cap B_x^-(x)$. Let $\gamma_{ab}(s), 0 \leq s \leq 1$ be the unique minimal geodesic of constant speed connecting $a$ and $b$. Let $\gamma_0(s), 0 \leq s \leq 1$ be the unique curve in $B_x(r)$ which satisfies $\gamma_{ab}(s) = \exp_x \gamma_0(s)$. Then by Lemma 1.1, we also have
\[
\frac{L(\gamma_0)}{L(\gamma_{ab})} \to 1
\]
as $(A, B) \to (0, 0)$. Since
\[d(a, b) \leq L(\Gamma), L(\gamma_0) \geq F(x, A - B),\]
we have
\[
\frac{F(x, A - B)}{L(\Gamma)} \leq \frac{F(x, A - B)}{d(a, b)} \leq \frac{L(\gamma_0)}{L(\gamma_{ab})}.
\]
Theorem 1.2 follows. \hfill \Box

2. Differentiability of isometries.

First we have:

Proposition 2.1. Let $\| \cdot \|_1$, $\| \cdot \|_2$ be two Minkowski norms on $\mathbb{R}^n$. Let $\phi$ be a mapping of $\mathbb{R}^n$ into itself such that $\| \phi(A) - \phi(B) \|_2 = \| A - B \|_1$, $\forall A, B \in \mathbb{R}^n$. Then $\phi$ is a diffeomorphism.

Proof. Consider $\mathbb{R}^n$ endowed with $\| \cdot \|_j$, $j = 1, 2$ as two Finsler spaces, denoted by $(M_1, F_1)$, $(M_2, F_2)$. Then geodesics in $M_j, j = 1, 2$ are straight lines (cf. [1]). And the distance function of $M_j$ are $d_j(A, B) = \| A - B \|_j$, $j = 1, 2$. Consider $\phi$ as a mapping from the Finsler space $(M_1, F_1)$ to $(M_2, F_2)$. Then $\phi$ preserves the distance function. Since in a Finsler space short geodesics minimize distance between its start and end points (cf. [1]), we can prove (similarly as in the Riemannian case) that $\phi$ transforms geodesics to geodesics. First suppose $\phi(0) = 0$. For $A \in \mathbb{R}^n$, $A \neq 0$, the curve $\phi(tA)$,
$t \geq 0$ is a ray which coincides with the ray $t \phi(A)$ for $t = 0$ and $t = 1$. Therefore they coincide as point sets. Thus $\phi(tA) = \mu(t) \phi(A)$ for some nonnegative function $\mu(t)$. Since

$$\|\phi(tA) - 0\|_2 = \|tA - 0\|_1 = t \|A\|_1$$

$$= \|\mu(t) \phi(A) - 0\|_2 = \mu(t) \|\phi(A)\|_2 = \mu(t) \|A\|_1, t \geq 0,$$

we have $\mu(t) = t$. Thus $\phi(tA) = t \phi(A)$, for $t \geq 0$. Suppose $A \neq B$, a similar argument as the above shows that there exists a nonnegative function $\lambda(t)$ such that $\phi(tA + (1 - t)B) = \lambda(t) \phi(A) + (1 - \lambda(t)) \phi(B)$, $t \geq 0$. And we can similarly show that $\lambda(t) = t$. In particular, for $t = \frac{1}{2}$ we have,

$$\frac{1}{2} \phi(A + B) = \phi \left( \frac{1}{2} (A + B) \right) = \frac{1}{2} \phi(A) + \frac{1}{2} \phi(B).$$

Thus $\phi(A + B) = \phi(A) + \phi(B)$. Taking $A = -B$ in the above equality we have $\phi(-A) = -\phi(A)$. Therefore $\phi$ is a linear transformation. Since $\text{Ker}(\phi) = \{0\}$, it is a diffeomorphism. If $A_1 = \phi(0) \neq 0$, consider the composition mapping $\phi_1 = \pi_{A_1} \circ \phi$, where $\pi_{A_1}(A) = A - A_1$ is the parallel translation, which is a diffeomorphism. Since $\phi_1(0) = 0$ and $\|\phi_1(A) - \phi(B)\|_2 = \|A - B\|_1$, $\phi_1$ is a diffeomorphism. Hence $\phi$ is a diffeomorphism. \qed

**Remark.** The proposition is an interesting application of Finsler geometry to Functional Analysis.

Now we can prove the main result of this paper.

**Theorem 2.2.** Let $(M, F)$ be a Finsler space and $\phi$ be a distance-preserving mapping of $M$ onto itself. Then $\phi$ is a diffeomorphism.

**Proof.** Let $p \in M$ and put $q = \phi(p)$. Let $\rho > 0, \epsilon > 0$ be so small that both $\exp_p$ and $\exp_q$ are $C^1$ diffeomorphisms on the tangent ball $B_p(r+\epsilon), B_q(r+\epsilon)$ of $T_p(M)$ and $T_q(M)$, respectively. For any nonzero $X \in T_p(M)$, consider the radial geodesic $\exp_q(tX), 0 \leq t \leq \frac{r}{2F(q, X)}$. The image $\gamma(t) = \phi(\exp_p(tX))$ is a geodesic since $\phi$ is distance-preserving. Let $X'$ denote the tangent vector of $\gamma$ at the point $q$. We have obtained a mapping $X \to X'$ of $T_p(M)$ into $T_q(M)$. Denoting this mapping by $\phi'$ we have $\phi'(\lambda X) = \lambda \phi'(X)$, for $X \in T_p(M)$ and $\lambda \geq 0$. Let $A, B \in T_p(M), A \neq B$ and $t$ is so small that both $tA$ and $tB$ lie in $B_p(r)$. Let $a_t = \exp_p(tA), b_t = \exp_p(tB)$. Then by Theorem 1.2 we have

$$\lim_{t \to 0^+} \frac{F(p, tA - tB)}{d(a_t, b_t)} = 1.$$

On the other hand, by the definition of $\phi'$ we have

$$\exp_q(\phi'(tX)) = \phi(\exp tX),$$
for any $X$ and $t$ small enough. Thus by Theorem 1.2 we also have

$$\lim_{t \to 0^+} \frac{F(q, \phi'(tA) - \phi'(tB))}{d(\phi(a_t), \phi(b_t))} = 1.$$ 

Since $d(\phi(a_t), \phi(b_t)) = d(a_t, b_t)$, we get

$$1 = \lim_{t \to 0^+} \frac{F(p, tA - tB)}{F(q, \phi'(tA) - \phi'(tB))} = \lim_{t \to 0^+} \frac{tF(p, A - B)}{tF(q, \phi'(A) - \phi'(B))} = \frac{F(p, A - B)}{F(q, \phi'(A) - \phi'(B))}.$$ 

Therefore $F(q, \phi'(A) - \phi'(B)) = F(p, A - B)$. By Proposition 2.1, $\phi'$ is a diffeomorphism of $T_p(M)$ onto $T_q(M)$.

Although on $\mathcal{B}^+_p(r) = \exp_p B_r(p)$ we have $\phi = \exp_q \circ \phi' \circ (\exp_p)^{-1}$, we still cannot conclude that $\phi$ is smooth on $\mathcal{B}^+_p(r)$, since in a Finsler space the exponential mapping is only $C^1$ at the zero section. That is, we can only conclude that $\phi$ is smooth in $\mathcal{B}_p(r) - \{p\}$. To finish the proof, we proceed to take $r$ so small so that every pair of points in $\mathcal{B}^+_p(r) \cap \mathcal{B}^-_q(r)$ can be joint by a unique minimizing geodesic. Select $p_1 \in \mathcal{B}^+_p(\frac{r}{2}) \cap \mathcal{B}^-_q(\frac{r}{2})$, $p_1 \neq p$.

Consider the tangent ball $\mathcal{B}_{p_1}(\frac{r}{2})$ of $T_{p_1}(M)$. The exponential mapping is a $C^1$ diffeomorphism from $\mathcal{B}_{p_1}(\frac{r}{2})$ onto $\mathcal{B}^+_{p_1}(\frac{r}{2})$. The above argument shows that $\phi$ is smooth in $\mathcal{B}^+_{p_1}(\frac{r}{2}) - \{p_1\}$, which is a neighborhood of $p$. This completes the proof. \qed

3. Group of isometries.

Theorem 2.2 justifies the following definition of isometry for a Finsler space.

**Definition 3.1.** Let $(M, F)$ be a Finsler space. A mapping $\phi$ of $M$ onto itself is called an isometry if $\phi$ is a diffeomorphism and for any $x \in M, X \in T_x(M)$, $F(\phi(x), d\phi_x(X)) = F(x, X)$.

In the following we denote the group of isometries of $(M, F)$ by $I(M)$.

Let $N$ be a connected, locally compact metric space and $\mathcal{I}(N)$ be the group of isometries of $N$, for each point $x$ of $N$, let $\mathcal{I}_x(N)$ denote the isotropy subgroup of $\mathcal{I}(N)$ at $x$. Van Danzig and van der Waerden [7] proved that $\mathcal{I}(N)$ is a locally compact topological transformation group on $N$ with respect to the compact-open topology and $\mathcal{I}_x(N)$ is compact.

Now on $M$ we have a distance function $d$ defined by the Finsler function $F$. By Theorem 2.2, the group $I(M)$ coincides with the group of isometries $\mathcal{I}(M)$ of $(M, d)$. Although generically $d$ is not a distance ($d$ is not symmetric unless $F$ is absolutely homogeneous), we still have:

**Theorem 3.2.** Let $(M, F)$ be a connected Finsler space. The compact-open topology turns $I(M)$ into a locally compact transformation group of $M$. Let
$x \in M$ and $I_x(M)$ denote the subgroup of $I(M)$ which leaves $x$ fixed. Then $I_x(M)$ is compact.

Proof. A proof of this result for the Riemannian case was given in Helgason [3] (cf. Helgason [3], pp. 201-204), which is valid in general cases after some minor changes. Just note that on a Finsler manifold the topology generated by the forward metric balls $B^+_p(r) = \{x \in M|d(p,x) < r\}, p \in M, r > 0$ is precisely the underlying manifold topology and this is true for the topology generated by the backward metric balls $B^-_p(r) = \{x \in M|d(x,p) < r\}, p \in M, r > 0$ (cf. [1]).

Bochner-Montgomery [2] proved that a locally compact group of differentiable transformations of a manifold is a Lie transformation group. Therefore we have the following theorem.

**Theorem 3.3.** Let $(M, F)$ be a Finsler space. Then the group of isometries $I(M)$ of $M$ is a Lie transformation group of $M$. Let $x \in M$ and $I_x(M)$ be the isotropy subgroup of $I(M)$ at $x$. Then $I_x(M)$ is compact.

References


Received March 15, 2001. This work is Project 19901015 and 19731004 supported by NSFC.

DEPARTMENT OF MATHEMATICS
NANKAI UNIVERSITY
TIANJIN, P. R. CHINA
E-mail address: matdsq@263.net

DEPARTMENT OF MATHEMATICS
NANKAI UNIVERSITY
TIANJIN, P. R. CHINA
E-mail address: houzx@nankai.edu.cn
ON SPACES OF MATRICES CONTAINING A NONZERO MATRIX OF BOUNDED RANK

Dmitry Falikman, Shmuel Friedland, and Raphael Loewy

Let $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$ be the spaces of $n \times n$ real matrices and real symmetric matrices respectively. We continue to study $d(n, n-2, \mathbb{R})$: The minimal number $\ell$ such that every $\ell$-dimensional subspace of $S_n(\mathbb{R})$ contains a nonzero matrix of rank $n-2$ or less. We show that $d(4, 2, \mathbb{R}) = 5$ and obtain some upper bounds and monotonicity properties of $d(n, n-2, \mathbb{R})$. We give upper bounds for the dimensions of $n-1$ subspaces (subspaces where every nonzero matrix has rank $n-1$) of $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$, which are sharp in many cases. We study the subspaces of $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$ where each nonzero matrix has rank $n$ or $n-1$. For a fixed integer $q > 1$ we find an infinite sequence of $n$ such that any $(q+1)^2$ dimensional subspace of $S_n(\mathbb{R})$ contains a nonzero matrix with an eigenvalue of multiplicity at least $q$.

1. Introduction.

Let $\mathbb{F}$ be a field, $M_{m,n}(\mathbb{F})$ the space of all $m \times n$ matrices over $\mathbb{F}$ and $S_n(\mathbb{F})$ the space of all $n \times n$ symmetric matrices over $\mathbb{F}$. We write $M_n(\mathbb{F})$ for $M_{n,n}(\mathbb{F})$. Let $V$ be either $M_{m,n}(\mathbb{F})$ or $S_n(\mathbb{F})$, and let $k$ be a positive integer. In the last 80 years there has been interest in the following types of subspaces, all related to the rank function:

(a) A subspace of $V$ where each matrix has rank bounded above by $k$.

(b) A subspace of $V$ where each nonzero matrix has rank $k$. A subspace of this type is said to be a $k$-subspace.

(c) A subspace of $V$ where each nonzero matrix has rank bounded below by $k$.

See the works [13], [11], [2], [8], [5], [3], [4], [7] and many others. Roughly speaking these problems are divided into two classes depending on whether $\mathbb{F}$ is algebraically closed or not. The classical case, which goes back to Radon-Hurwitz, discusses the maximal dimension $\rho(n)$ of an $n$-subspace $U$ of $M_n(\mathbb{R})$ where each $0 \neq A \in U$ is an orthogonal matrix times $r \in \mathbb{R}^*$. Write $n = (2a + 1)2^c + 4d$, where $a$ and $d$ are nonnegative integers, and $c \in$
{0, 1, 2, 3}. Then the Radon-Hurwitz number $\rho(n)$ is defined by

\begin{equation}
\rho(n) = 2^c + 8d.
\end{equation}

In his famous work Adams [1] gave a nonlinear version of the Radon-Hurwitz number by showing that $\rho(n) - 1$ is the maximal number of linearly independent vector fields on the $n - 1$ dimensional sphere $S^{n-1}$. ($S^{n-1} \subset \mathbb{R}^n$ denotes the Euclidean sphere of radius one centered at the origin.) From this result Adams deduced that the maximal dimension of an $n$-subspace of $M_n(\mathbb{R})$ is exactly $\rho(n)$. Let $\rho(x) = 0$ if $x$ is not a positive integer. Define now

\begin{equation}
\rho_s(n) = \rho\left(\frac{n}{2}\right) + 1.
\end{equation}

Adams, Lax and Phillips [2] showed that the maximal dimension of an $n$-subspace of $S_n(\mathbb{R})$ is exactly $\rho_s(n)$. Friedland, Robbin and Sylvester [8] and Berger and Friedland [5] gave further nonlinear versions of the above results by considering odd maps $\phi$ from $S^n$ to matrices of rank $n$ in $M_n(\mathbb{R}), S_n(\mathbb{R})$ and $M_{n,n+1}(\mathbb{R})$ respectively. In this paper (§4) we generalize these results to odd maps from $S^n$ to rank $n - 1$ matrices in $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$.

The main motivation of this paper is the following quantity studied in [7]. For an integer $k$, such that $1 \leq k \leq n - 1$, let $d(n, k, \mathbb{F})$ be the smallest integer $\ell$ such that every $\ell$ dimensional subspace of $S_n(\mathbb{F})$ contains a nonzero matrix whose rank is at most $k$. Note that it is clear that the maximal dimension of a subspace of $S_n(\mathbb{F})$ of type (c) above is exactly $d(n, k - 1, \mathbb{F}) - 1$. It is our purpose to continue the study of $d(n, k, \mathbb{R})$ started in [7]. Note that any $d(n, k, \mathbb{R}) - 1$ dimensional subspace of $S_n(\mathbb{R})$ contains a nonzero matrix with an eigenvalue of multiplicity at least $n - k$. (One of the main results of [8] was that any subspace of $S_n(\mathbb{R})$ of dimension $\sigma(n) + 1$, where

\[
\sigma(n) = 2 \text{ if } n \not\equiv 0, \pm 1 \pmod{8}, \\
\sigma(n) = \rho(4b) \text{ if } n = 8b, 8b \pm 1,
\]

contains a nonzero matrix with a multiple eigenvalue.) The equality

\begin{equation}
d(n, k, \mathbb{C}) = \left(\frac{n - k + 1}{2}\right) + 1,
\end{equation}

established in [7] is derived straightforwardly from the following dimension computations. Let $V_{k,n}(\mathbb{C})$ ($V_{k,n}(\mathbb{R})$) be the variety of all matrices in $S_n(\mathbb{C})$ ($S_n(\mathbb{R})$) of rank $k$ or less. Then in the projective space $\mathbb{P}S_n(\mathbb{C})$ the projective variety $\mathbb{P}V_{k,n}(\mathbb{C})$ is an irreducible variety of codimension $d(n, k, \mathbb{C}) - 1$, which yields (1.3). In particular $d(n, n - 1, \mathbb{C}) = 2$. The results of Adams, Lax and Phillips, cited above, yield that $d(n, n - 1, \mathbb{R}) = \rho_s(n) + 1$. This shows that in general the computation of $d(n, k, \mathbb{R})$ is much more difficult than the computation of $d(n, k, \mathbb{C})$. In [7] we gave a simple condition on $n$ when $d(n, k, \mathbb{C}) = d(n, k, \mathbb{R})$ for $k \leq n - 2$, which trivially holds for
It was shown by Harris and Tu \[10\] that the degree of the variety $\mathbb{P}V_{k,n}(\mathbb{C})$ is given by the formula

\[\delta_{k,n} := \deg \mathbb{P}V_{k,n}(\mathbb{C}) = \prod_{j=0}^{n-k-1} \frac{(n+j)}{(2j+1)} \frac{(n-k-j)}{j} \in \mathbb{Z}.\]

Then $d(n, k, \mathbb{C}) = d(n, k, \mathbb{R})$ if $\delta_{k,n}$ is odd. (In this case the result that any $d(n, k, \mathbb{R}) - 1$ dimensional subspace of $S_n(\mathbb{R})$ contains a nonzero matrix with an eigenvalue of multiplicity at least $n - k$ is best possible.) We show that $\delta_{n-q,n}$ is odd if $n > q \geq 1$ and

\[n \equiv \pm q \pmod{2^{\left\lfloor \log_2 q \right\rfloor}}.\]

In particular, under the above conditions,

\[d(n, n - q, \mathbb{C}) = d(n, n - q, \mathbb{R}) = \left(\frac{q+1}{2}\right) + 1.\]

If $n \geq q$ and $n$ satisfies (1.5) then any $\left\lfloor \frac{q+1}{2} \right\rfloor$ subspace of $S_n(\mathbb{R})$ contains a nonzero matrix with an eigenvalue of multiplicity at least $q$. This statement for $q = 2$ yields the original Lax’s result \[12\] that any 3 dimensional subspace of $S_n(\mathbb{R})$ contains a nonzero matrix with a multiple eigenvalue for $n \equiv 2 \pmod{4}$. (This result and its generalization in \[8\] is of importance in the study of singularities of hyperbolic systems.)

An important part of the paper is devoted to the study of the numbers $d(n, n - 2, \mathbb{R})$. Besides the cases given in (1.5-1.6) for $q = 2$ it is easy to see that $d(3, 1, \mathbb{R}) = 6$ \[7\]. (Hence the inequality $d(n, n - 2, \mathbb{R}) \geq \sigma(n) + 2$ established in \[7\] is not sharp for some $n$.) Partial results regarding $d(n, n - 2, \mathbb{R})$ were obtained in \[7\]. In particular, it was shown that for $m \geq 1$

\[d(4m, 4m - 2, \mathbb{R}) \leq 4m + 1.\]

We obtain here additional results on $d(n, n - 2, \mathbb{R})$. In particular, using numerical and symbolic computations we show that $d(4, 2, \mathbb{R}) = 5$, which implies that the upper bound in (1.7) is sharp for $m = 1$.

The paper is organized as follows: In Section 2 we show that $d(4, 2, \mathbb{R}) = 5$. In Section 3 we obtain some upper bounds and some monotonicity properties of $d(n, n - 2, \mathbb{R})$. In Section 4 we consider the existence of continuous and smooth odd maps from the sphere $S^p$ to matrices of rank $n - 1$ in $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$. As a consequence we obtain new upper bounds for the dimension of $(n - 1)$-subspaces of $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$. These upper bounds are shown to be sharp in many cases. In Section 5 we discuss the existence of subspaces $U \subset S_{2n-1}(\mathbb{R})$ of dimension $2n$, which do not contain a nonzero matrix of rank less than $2n - 2$. The last section is devoted to remarks and conjectures.
2. Computation of \(d(4, 2, \mathbb{R})\).

As indicated in the introduction, the first unknown number in the sequence 
\(\{d(n, n - 2, \mathbb{R})\}_{n=3}^\infty\) is \(d(4, 2, \mathbb{R})\). By (1.7) \(d(4, 2, \mathbb{R}) \leq 5\). It is our purpose to prove:

**Theorem 2.1.** \(d(4, 2, \mathbb{R}) = 5\).

**Proof.** It suffices to exhibit a 4-dimensional subspace \(L\) of \(S_4(\mathbb{R})\) with the property that every \(0 \neq A \in L\) has rank 3 or 4.

Let \(L\) be the subspace of \(S_4(\mathbb{R})\) spanned by the matrices

\[
B_1 = \begin{bmatrix}
1 & -1 & -1 & 1 \\
1 & -1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1
\end{bmatrix}, \\
B_2 = \begin{bmatrix}
1 & -1 & -1 & -1 \\
1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 \\
-1 & 1 & -1 & -1
\end{bmatrix}, \\
B_3 = \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 \\
1 & 1 & -1 & 1 \\
1 & 1 & -1 & 1
\end{bmatrix}, \\
B_4 = \begin{bmatrix}
-1 & 1 & 1 & 1 \\
-1 & -1 & -1 & -1 \\
-1 & 1 & 1 & -1 \\
1 & -1 & -1 & 1
\end{bmatrix}.
\]

It is straightforward to check that \(\dim L = 4\). Let \(\alpha_1, \alpha_2, \alpha_3, \alpha_4\) be indeterminates and let \(\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)\), and \(B(\alpha) = \sum_{i=1}^{4} \alpha_i B_i\).

We show that if \(0 \neq \alpha \in \mathbb{R}^4\) then \(\text{rank } B(\alpha) \geq 3\). So let \(d(\alpha) = \det B(\alpha)\), and for \(1 \leq j \leq 4\) let \(d_j(\alpha)\) denote the determinant of the principal submatrix of \(B(\alpha)\) obtained by deleting row and column \(j\). We let Maple compute the common zeros of the 5 polynomials \(d(\alpha), d_1(\alpha), d_2(\alpha), d_3(\alpha), d_4(\alpha)\) in \(\mathbb{C}^4\).

It turns out that there are no common zeros with \(\alpha_1 = 0\) except the trivial solution \(\alpha = 0\). Now we can assume \(\alpha_1 = 1\). We list the ten solutions found by Maple. They are:

(i) The two solutions given by

\[
\alpha = (1, 0, 1 + x, x),
\]

where \(x\) is a zero of the polynomial \(f(z) = z^2 + z + 1\).

(ii) The two solutions given by

\[
\alpha = (1, x, 2x - 1, 3x - 2),
\]

where \(x\) is a zero of the polynomial \(f(z) = 7z^2 - 8z + 3\).

(iii) The two solutions given by

\[
\alpha = (1, x, -x, -5x - 1),
\]

where \(x\) is a zero of the polynomial \(f(z) = 14z^2 + 4z + 1\).
(iv) The four solutions given by
\[ \alpha = \left( 1, x, \frac{1}{2}x^2 + x + \frac{1}{2}, -\frac{1}{2}x^2 - x - \frac{1}{2} \right), \]
where \( x \) is a zero of the polynomial
\[ f(z) = z^4 + 4z^3 + 8z^2 + 4z + 3 = (z + 1)^4 + 2z^2 + 2. \]

It follows that for any \( 0 \neq \alpha \in \mathbb{R}^4 \) either \( d(\alpha) \neq 0 \) or \( \exists 1 \leq j \leq 4 \) such that \( d_j(\alpha) \neq 0 \). Thus, the rank of every nonzero matrix in \( L \) is at least 3.

We explain now why Maple did indeed give us all the common zeros of \( d(\alpha), d_1(\alpha), d_2(\alpha), d_3(\alpha), d_4(\alpha). \) Let \( L_1 \) be the subspace of \( S_4(\mathbb{C}) \) spanned by \( B_1, B_2, B_3, B_4 \). We have by (1.3)
\[ d(4, 2, \mathbb{C}) = 4. \]

Moreover, by (1.4) we get
\[ (2.1) \quad \deg P(V_{2,4}(\mathbb{C})) = \binom{4}{2} \binom{5}{1} = 10. \]

We also use the known fact that a rank \( r \) symmetric matrix with entries in a field has a nonsingular \( r \times r \) principal submatrix. Thus, the ten distinct solutions found by Maple yield ten matrices in \( L_1 \) whose rank is at most 2, and no two of those matrices lie on the same line. In fact, a computation showed that each of those ten matrices has rank 2.

It remains to explain why Maple did not omit other solutions, assuring us it did not miss any real solutions in particular. It follows from (2.1) that if \( L' \) is a four dimensional generic subspace of \( S_4(\mathbb{C}) \) then \( P(V_{2,4}(\mathbb{C})) \) meets \( P(L') \) in at most ten distinct points. Hence, we have to check that \( L_1 \) is indeed generic. It suffices to show that \( P(V_{2,4}(\mathbb{C})) \) and \( P(L_1) \) intersect transversally.

So let \( A \) denote any of the ten matrices in \( L_1 \) of rank 2 found by Maple. Let \( x, y \in \mathbb{C}^4 \) be linearly independent vectors such that \( Ax = Ay = 0 \). A computation showed that the rank of
\[ \begin{bmatrix}
  x^tB_2x & x^tB_3x & x^tB_4x \\
  y^tB_2y & y^tB_3y & y^tB_4y \\
  x^tB_2y & x^tB_3y & x^tB_4y 
\end{bmatrix} \in M_3(\mathbb{C}) \]
is 3. This implies that the linear system of three equations
\[ \begin{align*}
  x^tBx &= 0 \\
  y^tBy &= 0 \\
  x^tBy &= 0,
\end{align*} \]
has no solution \( B \) in \( L_1 \) such that \( A, B \) are linearly independent. This shows that \( L_1 \) is generic.
In addition to the computation using Maple, a numerical procedure (using Matlab) was performed to check that there are no nonzero matrices of rank 1 or 2 in $L$.

3. Upper bounds and monotonicity properties of $d(n, n - 2, \mathbb{R})$.

It follows from (1.5) and (1.6) that if $n \equiv 2 \pmod{4}$ then $d(n, n - 2, \mathbb{R}) = d(n, n - 2, \mathbb{C}) = 4$. It was shown in [7] that

$$d(n, n - 2, \mathbb{R}) \leq 7 \quad \text{for} \ n \equiv 4 \pmod{8}. \quad (3.1)$$

It is our purpose here to get some additional upper bounds and some monotonicity properties of $d(n, n - 2, \mathbb{R})$.

**Proposition 3.1.** Let $n \equiv 3, 5 \pmod{8}$.Then $d(n, n - 2, \mathbb{R}) \leq 7$.

**Proof.** It follows from (1.4) that for every $n \geq 3$

$$\deg P(V_{n-3,n}(\mathbb{C})) = \frac{n(n+1)(n+2)}{1 \cdot 3 \cdot 10} = \frac{(n + 2)(n + 1)n^2(n - 1)(n - 2)}{2^3 \cdot 3^2 \cdot 5},$$

and this is odd for $n \equiv 3, 5 \pmod{8}$. So, as indicated in the introduction, and using (1.3) we have for $n \equiv 3, 5 \pmod{8}$

$$d(n, n - 3, \mathbb{R}) = d(n, n - 3, \mathbb{C}) = \binom{4}{2} + 1 = 7,$$

and since $d(n, n - 2, \mathbb{R}) \leq d(n, n - 3, \mathbb{R})$, the proposition follows. \qed

It follows from [7] that $d(n, n - 2, \mathbb{R}) \geq 4$ for $n \equiv 3, 4, 5 \pmod{8}$. Thus, we have the exact value for $d(n, n - 2, \mathbb{R})$ whenever $n \equiv 2, 6 \pmod{8}$, and good upper and lower bounds whenever $n \equiv 3, 4, 5 \pmod{8}$.

**Proposition 3.2.** Let $k$ be a fixed integer such that $k \in \{3, 4, 5, 11, 12, 13\}$. Then the sequence $\{d(16m + k, 16m + k - 2, \mathbb{R})\}_{m=0}^{\infty}$ is a (weakly) monotone increasing sequence, bounded above by 7.

There exists $M = M(k)$ such that $d(16m + k, 16m + k - 2, \mathbb{R}) = d(16M + k, 16M + k - 2, \mathbb{R})$ for all $m \geq M$.

**Proof.** It suffices to prove the monotonicity of the given sequence, because the required boundedness follows from (3.1) and Proposition 3.1.

Let $L$ be any 8-dimensional 8-subspace of $M_8(\mathbb{R})$. Such a subspace exists because $\rho(8) = 8$ by (1.1). Let

$$L_1 = \left\{ \begin{bmatrix} 0 & A \t \ A \end{bmatrix}, \ A \in L \right\}.$$

Then $L_1$ is an 8-dimensional subspace of $S_{16}(\mathbb{R})$. Note that every nonzero matrix in $L_1$ is nonsingular. Let $\{B_i\}_{i=1}^8$ be a basis of $L_1$.

Let $d_{m,k} = d(16m + k, 16m + k - 2, \mathbb{R})$ and $d_0 = d_{m,k} - 1$. Let $W_{m,k}$ be any $d_0$-dimensional subspace of $S_{16m+k}(\mathbb{R})$ such that the rank of any
nonzero matrix in $W_{m,k}$ is at least $16m + k - 1$. Note that $d_0 \leq 7 - 1 = 6$. Let $\{C_i\}_{i=1}^{d_0}$ be a basis of $W_{m,k}$. Let $\tilde{W}_{m,k} = \text{span}\{C_i \oplus B_i : 1 \leq i \leq d_0\}$. Then $\tilde{W}_{m,k}$ is a $d_0$-dimensional subspace of $S_{16(m+1)+k}(\mathbb{R})$, and clearly every nonzero matrix in it has rank $\geq 16 + 16m + k - 1 = 16(m+1) + k - 1$. This shows that

$$d(16(m+1)+k,16(m+1)+k-2,\mathbb{R}) \geq d_0 + 1 = d_{m,k} = d(16m+k,16m+k-2,\mathbb{R}).$$

\[\square\]

The following proposition justifies some claims we made in the introduction.

**Proposition 3.3.** Let $1 \leq k \leq n - 2$. Then any $d(n,k,\mathbb{R}) - 1$ dimensional subspace of $S_n(\mathbb{R})$ contains a nonzero matrix which has an eigenvalue of multiplicity at least $n - k$. If $d(n,k,\mathbb{R}) = d(n,k,\mathbb{C})$ then this result is sharp.

**Proof.** Let $V$ be a $d(n,k,\mathbb{R}) - 1$ dimensional subspace of $S_n(\mathbb{R})$. If $I_n \in V$ then 1 is the eigenvalue of $I_n$ of multiplicity $n \geq n - k$. Suppose that $I_n \notin V$. Let $\hat{V} = \text{span}\{V,I_n\}$. Then $\hat{V}$ contains $0 \neq \hat{A}$ such that rank $\hat{A} \leq k$. $\hat{A} = A + aI, 0 \neq A \in V$, so $-a$ is an eigenvalue of $A$ of multiplicity at least $n - k$.

Let $\Sigma_{n-k} \subset S_n(\mathbb{R})$ be the variety of all $A \in S_n(\mathbb{R})$ such that $A$ has an eigenvalue of multiplicity at least $n - k$. Clearly

$$\Sigma_{n-k} = \{B \in S_n(\mathbb{R}) : B = A + aI_n \text{ for some } A \in V_{k,n}(\mathbb{R}) \text{ and } a \in \mathbb{R}\}.$$ 

Since $V_{k,n}(\mathbb{R})$ is an irreducible variety of codimension $d(n,k,\mathbb{C}) - 1$ it follows that $\Sigma_{n-k}$ is an irreducible variety of codimension $d(n,k,\mathbb{C}) - 2$. Hence there exists an $d(n,k,\mathbb{C}) - 2$ dimensional subspace $W \subset S_n(\mathbb{R})$ such that $\Sigma_{n-k} \cap W = \{0\}$.

4. **Upper bounds for the dimension of $(n - 1)$-subspaces of $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$.**

In this section we obtain upper bounds for the dimension of $(n - 1)$-subspaces of $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$. This is done by considering certain odd, smooth maps from $S^p$ into $M_n(\mathbb{R})$ and $S_n(\mathbb{R})$. We also need several known results.

Let $M_n(k)(\mathbb{R})(S_n(k)(\mathbb{R}))$ denote the set of all rank $k$ matrices in $M_n(\mathbb{R})(S_n(\mathbb{R}))$. Given any field $\mathbb{F}$ and positive integers $m, n$, let $M_{m,n}^0(\mathbb{F})$ denote the set of all matrices in $M_{m,n}(\mathbb{F})$ whose rank is equal to $\min\{m,n\}$.

**Lemma 4.1.** Let $C \in M_{m,n}^0(\mathbb{F})$. For $j = 1, 2, \ldots, n + 1$, let $C^{(j)}$ denote the $n \times n$ matrix obtained from $C$ by deleting its $j$-th column. Then the
solution of the homogeneous linear system \( Cy = 0 \) is a line spanned by

\[
y = \left(-\det C^{(1)}, \det C^{(2)}, -\det C^{(3)}, \ldots, (-1)^{n+1}\det C^{(n+1)}\right)^t.
\]

Proof. This is an easy consequence of well-known properties of the determinant. \(\square\)

The next theorem appears as a part of Theorems A and F in [8]. See there how it is related to classical results due to Radon, Hurwitz, Adams and Adams-Lax-Phillips.

**Theorem 4.1.** Let \( \varphi : S^p \to M_n(\mathbb{R}) \) be an odd continuous map, i.e., \( \varphi(-\alpha) = -\varphi(\alpha) \quad \forall \alpha \in S^p. \)

\(\text{(i)}\) Suppose that \( \varphi(S^p) \subset M_n^{(2)}(\mathbb{R}) = GL(n, \mathbb{R}). \) Then, \( p \leq \rho(n) - 1. \)

\(\text{(ii)}\) Suppose that \( \varphi(S^p) \subset S_n^{(n)}(\mathbb{R}). \) Then, we have \( p = 0 \) if \( n \) is odd, and \( p \leq \rho(\frac{n}{2}) \) if \( n \) is even.

All the inequalities in (i) and (ii) are sharp.

The next theorem appears in [5].

**Theorem 4.2.** Let \( \varphi : S^p \to M_{n,n+1}^0(\mathbb{R}) \) be an odd continuous map. Then \( p \leq \max\{\rho(n) - 1, \rho(n + 1) - 1\} \), and this inequality is sharp.

**Theorem 4.3.** Let \( n \geq 2. \) Let \( \varphi : S^p \to M_n^{(n-1)}(\mathbb{R}) \) be \( n \) odd smooth map. Then, if \( n \neq 3, 5, 9 \)

\[
p \leq \max\{\rho(n-1) - 1, \rho(n) - 1, \rho(n + 1) - 1, 2\}.
\]

Furthermore, if \( \varphi(S^p) \subset S_n^{(n-1)}(\mathbb{R}) \) then \( p = 0 \) if \( n \) is even, and

\[
p \leq \max\left\{\rho\left(\frac{n-1}{2}\right), \rho(n + 1) - 1\right\} \quad \text{if} \quad n \text{ is odd}.
\]

Proof. We prove first (4.2). If \( p \leq 1 \) (4.2) certainly holds, so we may assume \( p \geq 2. \) Hence \( S^p \) is simply connected. Therefore, we can choose \( x(\alpha) \in S^{n-1} \) in a well-defined and continuous way such that

\[
x^t(\alpha)\varphi(\alpha) = 0 \quad \forall \alpha \in S^p.
\]

For \( \alpha \in S^p \) let \( B(\alpha) \) be the matrix in \( M_{n,n+1}(\mathbb{R}) \) obtained from \( \varphi(\alpha) \) by augmenting it with the column vector \( x(\alpha) \), i.e.,

\[
B(\alpha) = (\varphi(\alpha), x(\alpha)) \in M_{n,n+1}(\mathbb{R}).
\]

Then \( \operatorname{rank} \varphi(\alpha) = n - 1 \) and (4.4) imply that \( B(\alpha) \in M_{n,n+1}^0(\mathbb{R}) \) for all \( \alpha \in S^p. \)

Note that the smoothness of \( \varphi \) implies that \( x(\alpha) \) is also smooth. It is also clear that for each \( \alpha \in S^p \) \( x(-\alpha) = \pm x(\alpha) \). Hence, we must have \( x(-\alpha) = -x(\alpha) \) for all \( \alpha \in S^p \), or \( x(-\alpha) = x(\alpha) \) for all \( \alpha \in S^p. \)
Case 1. Suppose that \( x(-\alpha) = -x(\alpha) \) for all \( \alpha \in S^p \). Then the map from \( S^p \) to \( M_{n,n+1}(\mathbb{R}) \) defined by \( \alpha \to B(\alpha) \) is an odd continuous map, so by Theorem 4.2 we have

\[
(4.5) \quad p \leq \max\{\rho(n) - 1, \rho(n + 1) - 1\},
\]
and (4.2) holds.

Case 2. We now assume \( x(-\alpha) = x(\alpha) \) for all \( \alpha \in S^p \). We apply Lemma 4.1 to \( B(\alpha) \in M_{n,n+1}(\mathbb{R}) \), for each \( \alpha \in S^p \), and denote the normalized solution given by (4.1) (that is, \( \frac{1}{|y|} y \)) by \( \psi(\alpha) \). Let

\[
\psi(\alpha) = (\psi_1(\alpha), \psi_2(\alpha), \ldots, \psi_n(\alpha), \psi_{n+1}(\alpha)).
\]

There are two subcases now.

Case 2a. Suppose that \( n \) is even. It follows from (4.1) that

\[
\psi(-\alpha) = (-\psi_1(\alpha), -\psi_2(\alpha), \ldots, -\psi_n(\alpha), \psi_{n+1}(\alpha)).
\]

We define

\[
C(\alpha) = \begin{bmatrix} B(\alpha) \\ \psi(\alpha) \end{bmatrix} \in M_{n+1}(\mathbb{R}).
\]

This matrix is nonsingular because \( \psi(\alpha) \) is orthogonal to all the rows of \( B(\alpha) \). Let \( \eta(\alpha) = (\psi_1(\alpha), \psi_2(\alpha), \ldots, \psi_n(\alpha)) \) and let

\[
D(\alpha) = \begin{bmatrix} \varphi(\alpha) \\ \eta(\alpha) \end{bmatrix} \in M_{n+1,n}(\mathbb{R}),
\]

that is, \( D(\alpha) \) is the matrix obtained from \( C(\alpha) \) by deleting its last column. Its rank is \( n \) for each \( \alpha \in S^p \). Since \( D(-\alpha) = -D(\alpha) \) for all \( \alpha \in S^p \) we apply Theorem 4.2 again and conclude that (4.5) holds, so (4.2) holds.

Case 2b. Suppose that \( n \) is odd. So \( n \geq 3 \). Suppose first that \( p < n - 1 \). Since \( x(\alpha) \) is smooth, it follows from Sard’s theorem that \( \{x(\alpha) : \alpha \in S^p\} \) has measure zero in \( S^{n-1} \). In particular, \( \exists \xi \in S^{n-1} \) such that \( x(\alpha) \neq \pm \xi \) for all \( \alpha \in S^p \).

For every \( \alpha \in S^p \) let \( L_\alpha = \text{span}\{x(\alpha), \xi\} \). Let \( Q(\alpha) \) be the unique \( n \times n \) orthogonal matrix satisfying: \( Q(\alpha) \) is the identity on \( L_\alpha \) and its restriction to \( L_\alpha \) is the rotation in that plane by an angle \( < \pi \) that sends \( x(\alpha) \) to \( \xi \). Since \( x(-\alpha) = x(\alpha) \) we have \( Q(-\alpha) = Q(\alpha) \). Also, \( Q(\alpha) \) is continuous in \( \alpha \) (cf. the Proposition in [8]). It follows from (4.4) that

\[
(4.6) \quad \xi^t(Q^t(\alpha)\varphi(\alpha)Q(\alpha)) = 0, \quad \text{for all } \alpha \in S^p.
\]

Let \( \varphi_1(\alpha) = Q^t(\alpha)\varphi(\alpha)Q(\alpha) \). Then \( \varphi_1(\alpha) \) is an odd continuous function. Without loss of generality we may assume \( \xi^t = (0, 0, \ldots, 0, 1) \). It follows from (4.6) that for each \( \alpha \in S^p \) the last row of \( \varphi_1(\alpha) \) is 0, so deleting it
results in an \((n - 1) \times n\) matrix of rank \(n - 1\). Applying Theorem 4.2 again we conclude that

\[
p \leq \max\{\rho(n - 1) - 1, \rho(n) - 1\}
\]

so (4.2) holds.

Now suppose that \(p \geq n - 1\). We may consider \(S^{n-2}\) as contained in \(S^p\), and then consider the restriction of the given \(\varphi(\alpha)\) and \(x(\alpha)\) to \(S^{n-2}\). We repeat the proof given for \(p < n - 1\) and conclude in the same way that

\[
n - 2 \leq \max\{\rho(n - 1) - 1, \rho(n) - 1\} = \rho(n - 1) - 1.
\]

Since this cannot happen by (1.1) when \(n \neq 3, 5, 9\) we have completed the proof of (4.2).

Suppose now that \(\varphi(S^p) \subset S^{(n-1)}(\mathbb{R})\). Recall that the inertia of \(A \in S_n(\mathbb{R})\) is the triple \((\pi, \nu, \delta)\), where \(\pi\) is the number of positive eigenvalues of \(A\), \(\nu\) is the number of negative eigenvalues of \(A\) and \(\delta\) is the number of zero eigenvalues of \(A\).

Let \(n\) be even and suppose that \(p > 0\). Let \(\alpha \in S^p\), and consider a path in \(S^p\) from \(\alpha\) to \(-\alpha\). Let \(J\) denote the image of this path under \(\varphi\). Since every matrix in \(J\) has rank \(n - 1\), it follows that the inertia of each matrix on this path is equal to inertia \((\varphi(\alpha))\). In particular, inertia \((\varphi(-\alpha)) = \text{inertia}(\varphi(\alpha))\), but this is impossible since \(\varphi(-\alpha) = -\varphi(\alpha)\). Hence \(p = 0\).

Let \(n\) be odd. Suppose first that \(n = 3\). We show that \(p \leq 1\) in this case, so (4.3) holds. Suppose to the contrary that \(p \geq 2\). It is clear that \(\varphi(\alpha)\) has the same inertia for each \(\alpha \in S^p\). In particular, given any \(\alpha \in S^p\), \(\varphi(\alpha)\) and \(\varphi(-\alpha) = -\varphi(\alpha)\) have the same inertia. Hence each \(\varphi(\alpha)\) has one positive, one negative and one zero eigenvalue. So for each \(\alpha \in S^p\), the eigenvalues of \(\varphi(\alpha)\) are pairwise distinct, contradicting Theorem B of [8].

Note that while we have shown that there is no odd continuous map from \(S^2\) to \(S^{(2)}_3(\mathbb{R})\), there is an odd continuous map from \(S^2\) to \(M^{(2)}_3(\mathbb{R})\). For example, consider the map that sends \((x, y, z) \in S^2\) to

\[
\begin{bmatrix}
0 & x & y \\
-x & 0 & z \\
-y & -z & 0
\end{bmatrix}.
\]

We assume now \(n \geq 5\). Since \(\max\{\rho(n - 1), \rho(n + 1) - 1\} \geq 2\), (4.3) holds if \(p \leq 1\). Hence we may assume that \(p \geq 2\). We go along the proof of the nonsymmetric part of the theorem. If \(x(\alpha)\) is odd, then (4.5) holds, and since \(\rho(n) = 1\) for odd \(n\), we get

\[
p \leq \rho(n + 1) - 1 \leq \max\left\{\rho\left(\frac{n - 1}{2}\right), \rho(n + 1) - 1\right\},
\]

and (4.3) holds.
So we may assume that \( x(\alpha) \) is even. Suppose also that \( p < n - 1 \). Then we get (4.6) again, where \( \varphi_1(\alpha) = Q^t(\alpha)\varphi(\alpha)Q(\alpha) \) is an odd continuous function, and we may assume \( \xi' = (0, 0, \ldots, 0, 1) \). It follows that for each \( \alpha \in S^p \) the matrix obtained from \( \varphi_1(\alpha) \) by deleting its last row and column is in \( S^{(n-1)}(\mathbb{R}) \). Hence, by part (ii) of Theorem 4.1 we get

\[
p \leq \rho \left( \frac{n - 1}{2} \right),
\]

so (4.3) holds.

If \( p \geq n - 1 \) we consider \( S^{n-2} \) as contained in \( S^p \), and then consider the restriction of \( \varphi(\alpha), x(\alpha) \) to \( S^{n-2} \). A repetition of the argument for \( p < n - 1 \) yields \( n - 2 \leq \rho \left( \frac{n - 1}{2} \right) \), which is impossible for \( n \) odd, \( n \geq 5 \). This completes the proof. \( \square \)

**Corollary 4.1.** Let \( n \geq 2 \) and let \( L \) be an \((n - 1)\)-subspace of \( M_n(\mathbb{R}) \).

(a) If \( n \neq 3, 5, 9 \), then

\[
\dim L \leq \max\{\rho(n - 1), \rho(n), \rho(n + 1), 3\}.
\]

(b) If \( L \subset S_n(\mathbb{R}) \) (that is, \( L \) is a subspace of \( S_n(\mathbb{R}) \)), then

\[
\dim L \leq \begin{cases} 1 & \text{if } n \text{ is even}, \\ \max\left\{\rho \left( \frac{n-1}{2} \right) + 1, \rho(n+1)\right\} & \text{if } n \text{ is odd}. \end{cases}
\]

**Proof.** (a) Let \( d = \dim L \) and let \( A_1, A_2, \ldots, A_d \) be a basis of \( V \). For \( A \in V \) we write \( A = \sum_{i=1}^{d} \alpha_i A_i \) and let \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \). Then we may view the map \( \alpha \to A \), restricted to \( \alpha \) such that \( \|\alpha\| = 1 \), as an odd smooth map from \( S^{d-1} \) to \( M_n^{(n-1)}(\mathbb{R}) \). Then \( d - 1 \) is bounded by the right-hand side of (4.2), so (4.7) follows.

(b) The proof is similar to the previous case, using now (4.3) instead of (4.2). \( \square \)

Consider next the sharpness of the inequalities (4.2), (4.3), (4.7) and (4.8). For that purpose we have the following lemma.

**Lemma 4.2.**

(i) There exists a \( \rho(n - 1) \) dimensional \((n - 1)\)-subspace \( V \) of \( M_n(\mathbb{R}) \).

(ii) There exists an \((n - 1)\)-subspace \( V \) of \( S_n(\mathbb{R}) \) such that

\[
\dim V = \begin{cases} 1 & \text{if } n \text{ is even}, \\ \rho \left( \frac{n-1}{2} \right) + 1, & \text{if } n \text{ is odd}. \end{cases}
\]

**Proof.** (i) As indicated in the introduction, there exists an \( \rho(n - 1) \) dimensional \((n - 1)\)-subspace \( V_1 \) of \( M_{n-1}(\mathbb{R}) \). Then we let \( V \) be the subspace of \( M_n(\mathbb{R}) \) obtained from \( V_1 \) by appending a row and column of 0’s to each matrix in \( V_1 \).
(ii) The claim is trivial if \( n \) is even, so suppose \( n \) is odd. As indicated in the introduction, there exists an \((n-1)\)-subspace \( V_1 \) of \( S_{n-1}(\mathbb{R}) \) of dimension \( \rho\left(\frac{n-1}{2}\right) + 1 \). Now \( V \) is obtained from \( V_1 \) as in Part (i). \( \square \)

**Corollary 4.2.** Let \( n \equiv 1 \) (mod 4). Then the inequalities (4.3) and (4.8) are sharp, and the inequalities (4.2) and (4.7) are sharp provided that \( n \neq 5, 9 \).

**Proof.** For \( n \equiv 1 \) (mod 4) the maximum of the right-hand sides of (4.2) and (4.7) are \( \rho(n-1)-1 \) and \( \rho(n-1) \), respectively. Let \( V \) be the \( \rho(n-1) \)-dimensional \((n-1)\)-subspace defined in the Proof of Lemma 4.2, Part (i). This shows the sharpness of (4.7) provided that \( n \neq 5, 9 \). The sharpness of (4.2) is obtained by considering an odd smooth map from \( S^{\rho(n-1)-1} \) to \( M_{n(n-1)}(\mathbb{R}) \) as in the proof of Corollary 4.1 Part (a).

The sharpness of (4.3) and (4.8) is proved similarly, using the \( \rho\left(\frac{n-1}{2}\right) + 1 \) dimensional \((n-1)\)-subspace \( V \) of \( S_n(\mathbb{R}) \) in Part (ii) of Lemma 4.2. \( \square \)

**Remark.** In Theorem 4.3 it is possible to replace the assumption that \( \varphi \) is an odd smooth map by the assumption that \( \varphi \) is an odd continuous map.

The proof of the remark is achieved as follows: First approximate \( \varphi(\alpha) \) arbitrarily by a smooth odd \( \tilde{\varphi}(\alpha) \) (which is in \( S_n(\mathbb{R}) \) if \( \varphi(S^p) \subset S_n^{(n-1)}(\mathbb{R}) \)). We can assume that for each \( \alpha \in S^p \), \( \tilde{\varphi}(\alpha) \) has a simple eigenvalue \( \lambda(\alpha) \) which is the smallest in absolute value among all eigenvalues of \( \tilde{\varphi}(\alpha) \) (and is also real if \( \varphi(S^p) \subset S_n^{(n-1)}(\mathbb{R}) \)). Clearly \( \lambda(-\alpha) = -\lambda(\alpha) \forall \alpha \in S^p \). Let \( \psi(\alpha) \) be the projection of \( \tilde{\varphi}(\alpha) \) corresponding to \( \lambda(\alpha) \). Then \( \psi(\alpha) \) is a smooth odd function of rank at most 1, and it follows that \( \tilde{\varphi}(\alpha) - \psi(\alpha) \) satisfies the conditions of Theorem 4.3, so (4.2) and (4.3) hold respectively.

**5. Existence of certain \( 2n \) dimensional subspaces of \( S_{2n-1}(\mathbb{R}) \).**

Let \( n \geq 4 \) be an even integer. It follows from [7] and this paper that \( d(n,n-2,\mathbb{R}) \leq n+1 \), with equality for \( n = 4 \). Now suppose that \( n \geq 5 \) is an odd integer. We do not know if a similar result holds, although it seems plausible. In this section we show that if there exists an \( n+1 \) dimensional subspace \( L \) of \( S_n(\mathbb{R}) \) such that each nonzero matrix in \( L \) has rank \( n-1 \) or \( n \), then there is a severe restriction on the possible inertias attained in \( L \). We also derive here additional results on \( n+1 \) dimensional subspaces of \( n \times n \) matrices where each nonzero matrix has rank \( n-1 \) or \( n \).

Let \( \mathbb{F} \) be any field and let \( V \) be an \( n+1 \) dimensional subspace of \( M_n(\mathbb{F}) \) spanned by \( A_1, A_2, \ldots, A_{n+1} \). Sometimes we find it convenient to identify \( \mathbb{F}^{n+1} \) with \( \mathbb{F}^{n+1} \) and \( \mathbb{P} V \) with \( \mathbb{P}^{n} \). That is, we identify \( \tilde{\alpha} \in \mathbb{F}^{n+1} \) and \( \tilde{A} \in \tilde{V} := \mathbb{P} V \) with lines spanned by \( \alpha \neq (\alpha_1, \alpha_2, \ldots, \alpha_{n+1})^t \) and \( 0 \neq A \in \mathbb{P}^{n+1} \) and \( V \) respectively. Let \( A(\alpha) = \sum_{i=1}^{n+1} \alpha_i A_i \). We identify \( \tilde{\alpha} \) with \( A(\alpha) \).
Given any $0 \neq y \in \mathbb{F}^n$, let
\begin{equation}
L_y = \{ A \in V : Ay = 0 \}.
\end{equation}
Since $\dim V \geq n + 1$ it is clear that $\dim L_y \geq 1$ for every $y \neq 0$. Let
\begin{equation}
M(y) = [A_1y, A_2y, \ldots, A_{n+1}y] \in M_{n,n+1}(\mathbb{F}).
\end{equation}
For $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{n+1})^t$ it is clear that $(\sum_{i=1}^{n+1} \alpha_i A_i)y = 0$ if and only if $M(y)\alpha = 0$. It follows that
\[ \dim L_y = 1 \text{ if and only if } \operatorname{rank} M(y) = n. \]
Observe that if $\operatorname{rank} M(y) = n$ then, by Lemma 4.1, $\ker M(y)$ is spanned
\begin{equation}
\begin{pmatrix}
\alpha_1(y), \\
\alpha_2(y), \\
\vdots \\
\alpha_{n+1}(y)
\end{pmatrix}.
\end{equation}
Let
\begin{equation}
\tilde{T} = \{ \tilde{y} \in \mathbb{P}^{n-1} : \alpha(y) = 0 \}.
\end{equation}
Hence, $T$ consists of those $\tilde{y}$ for which $\operatorname{rank} M(y) < n$.
Given $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{n+1})$ we consider the homogeneous polynomial
\begin{equation}
\psi(\alpha) = \det \left( \sum_{i=1}^{n+1} \alpha_i A_i \right),
\end{equation}
and let
\begin{equation}
\tilde{Z}(\psi) = \text{ the zero set of } \psi(\alpha), \text{ considered as a subset of } \mathbb{P}^{n-1}.
\end{equation}

**Lemma 5.1.** Let $V$ be an $n+1$ dimensional subspace of $M_n(\mathbb{C})$, spanned by $A_1, A_2, \ldots, A_{n+1}$, and such that:
\begin{enumerate}
\item $\exists A \in V$ such that $\det A \neq 0$.
\item $\exists B \in V$ such that $\operatorname{rank} B = n-1$ and $\ker B$ is spanned by a vector $0 \neq u \in \mathbb{C}^n$ with $\dim L_u = 1$.
\end{enumerate}
Then the following holds:
Let $H \subset \mathbb{P} \mathbb{C}^n$ be the irreducible component of the hypersurface $\tilde{Z}(\psi)$ passing through $\tilde{B}$. Then there exists a birational map
\[ \varphi : H \dashrightarrow \mathbb{P} \mathbb{C}^{n-1}, \quad \theta = \varphi^{-1} : \mathbb{P} \mathbb{C}^{n-1} \dashrightarrow H \]
defined as follows:
\begin{enumerate}
\item For each $\tilde{\alpha} \in H$ with $\operatorname{rank} C = n-1$, where $C = \sum_{i=1}^{n+1} \alpha_i A_i$, $\varphi(\tilde{\alpha}) = \tilde{y}$, where $y \in \mathbb{C}^n$ is the basis of $\ker C$.
\item For each $\tilde{y} \in \mathbb{P} \mathbb{C}^{n-1}$ such that $\dim L_y = 1$, $\theta(\tilde{y}) = \tilde{\alpha}$, where $\alpha = \alpha(y)$.
\end{enumerate}
Proof. The maps $\varphi$ and $\theta$ are rational. For $y$ in the neighborhood of $u$, $\dim L_y = 1$. Hence $\theta$ is holomorphic in the neighborhood of $u$. Let $B = \sum_{i=1}^{n+1} \beta_i A_i$. The assumptions (ii) imply that $\varphi$ is holomorphic in the neighborhood of $\beta$ and $\varphi \cdot \theta$ is the identity map (as a rational map). Hence $\theta : \mathbb{P}^{n-1} \rightarrow H$.

□

Remark. As $\theta : \mathbb{P}^{n-1} \rightarrow H$ is a rational map, it is not holomorphic on a variety of codimension at least 2 (cf. [9, Chapter 4, Section 2]). Hence $\text{codim } T \geq 2$.

Lemma 5.2. Let $V$ be an $n + 1$ dimensional subspace of $M_n(\mathbb{R})$, such that each nonzero matrix in $V$ has rank $n - 1$ or $n$. Then:

(i) $\exists A \in V$ such that $\det A \neq 0$.

(ii) $\exists B \in V$ such that $\text{rank } B = n - 1$ and $\text{ker } B$ is spanned by a vector $0 \neq y \in \mathbb{R}^n$ with $\dim L_y = 1$.

Proof. By Theorem 2 of [3], $V$ cannot be an $(n - 1)$-subspace or an $n$-subspace. Hence $V$ contains nonsingular matrices and singular matrices. Let

$$\tilde{V}_s = \{ \tilde{A} \in \tilde{V} : \det \tilde{A} = 0 \}.$$  

Let

$$\eta : \tilde{V}_s \rightarrow \mathbb{P}^{n-1}$$

be defined as follows:

Suppose that $B \in V$ is any singular matrix, and suppose $0 \neq y \in \mathbb{R}^n$ satisfies $By = 0$. Then let $\eta(B) = \tilde{y}$.

It is known that the set of singular points of $V$ is a proper subvariety of $V$. As $\eta(V) = \mathbb{P}^{n-1}$, it follows that there exists a regular point $\tilde{B}$ of $\tilde{V}$ such that $D\eta(\tilde{B})$ has the rank $n - 1$. Let $y \in S^{n-1}$ be such that $By = 0$. Suppose that $\dim L_y \geq 2$. Then it is clear that $\dim \text{ker } D\eta(\tilde{B}) \geq 1$.

It follows that the dimension of the tangent space to $\tilde{B}$ in $\tilde{V}_s$ has dimension $\geq n$, implying that $\dim \tilde{V}_s \geq n$, a contradiction. □

Lemma 5.3. Let $n$ be odd, $n \geq 3$, and let $V$ be an $n + 1$ dimensional subspace of $S_n(\mathbb{R})$ such that each nonzero matrix in $V$ has rank $n - 1$ or $n$. Suppose that $V^\perp = \{ B \in S_n(\mathbb{R}) : \text{tr } (AB) = 0 \ \forall A \in V \}$ does not contain rank one matrices. Let $\tilde{V}_s$ be as in (5.6). Then, $\tilde{V}_s$ is a smooth connected variety in $\mathbb{P}^{n-1}$.

Proof. Let $A_k = (a^{(k)}_{ij})_{i,j=1}^{n+1}$, $k = 1, 2, \ldots, n + 1$ be a basis of $V$. Let $\psi(\alpha) = \det \left( \sum_{i=1}^{n+1} \alpha_i A_i \right)$. For $\alpha \in \tilde{V}_s$ we have $\psi(\alpha) = 0$, so $A(\alpha) = \sum_{i=1}^{n+1} \alpha_i A_i$ has
rank $n - 1$. Hence the $(n-1)$th compound of $A(\alpha), C_{n-1}(A(\alpha))$, has rank 1. So there exists $0 \neq x(\alpha) = (x_1(\alpha), x_2(\alpha), \ldots, x_n(\alpha))^t$ and $\epsilon = \pm 1$ such that $C_{n-1}(A(\alpha)) = \epsilon x(\alpha)x^t(\alpha)$. Let $C_{n-1}(\alpha) = (c_{ij}(\alpha))_{i,j=1}^n$. Observe that

$$\frac{\partial \psi(\alpha)}{\partial \alpha_k} = \sum_{i,j=1}^n a_{ij}^{(k)}c_{ij}(\alpha) = \epsilon x^t(\alpha)A_kx(\alpha), \ k = 1, 2, \ldots, n + 1.$$

Since $V^\perp$ does not contain a rank one matrix, we deduce that $\nabla \psi(\alpha) \neq 0$. Thus each $\bar{\alpha} \in \bar{V}_s$ is a smooth point and the local dimension of $\bar{V}_s$ at $\bar{\alpha}$ is $n - 1$.

Let $V_C$ be the complexification of $V$. Let $T_C, T_R$ be defined by (5.5) for $F = C, R$, respectively. By the remark following Lemma 5.1, codim$_C T_C \geq 2$. Since $T_R = T_C \cap \mathbb{P}R^{n-1}$, we have codim$_R T_R \geq 2$.

Let $Y_r = \mathbb{P}R^{n-1} \setminus T_R$. Then $Y_r$ is connected. Using (ii) of Lemma 5.2 we can define $H$ and $\theta$ as in Lemma 5.1. Let $H_R = H \cap \mathbb{P}R^n$. Then $\theta(Y_r)$ is connected in $H_R$, and this implies that $\theta(Y_r) = H_R$ is connected.

To finish the proof it suffices to show that $\bar{V}_s = H_R$. Since $\bar{Y}_r = \mathbb{P}R^{n-1}$ it follows that for any $\bar{y} \in \mathbb{P}R^{n-1}$ there exists $\bar{\beta} \in H_R$ such that $A(\beta)y = 0$. Let $\gamma$ be an arbitrary point in $\bar{V}_s$. Then $\exists$ $0 \neq u \in \mathbb{R}^n$ such that $A(\gamma)u = 0$. If dim $L_u = 1$, then $\gamma \in H_R$. So suppose dim $L_u \geq 2$. Then $\exists$ $\beta \in H_R$ such that $A(\beta) \in L_u$. Hence $\gamma \in L_u$, and clearly $\beta$ and $\gamma$ are connected in $L_u \subset \bar{V}_s$. As $\bar{V}_s$ is a smooth variety, it follows that $\gamma \in H_R$. Hence $\bar{V}_s = H_R$.

Recall that the standard inner product in $S_n(\mathbb{R})$ is given by $\langle A, B \rangle = \text{tr}(AB)$, so with respect to this inner product a matrix $A$ is normalized if and only if $\text{tr}(A^2) = 1$.

**Corollary 5.1.** Let $n$ and $V$ satisfy the assumptions of Lemma 5.3. Let

$$V_\nu = \{ A \in V : \text{tr}(A^2) = 1 \}$$

and

$$V_{\nu s} = \{ A \in V_\nu : \det A = 0 \}.$$

Then $V_{\nu s}$ is a smooth connected hypersurface in $V_\nu$.

**Proof.** Assume that the basis $A_1, A_2, \ldots, A_{n+1}$ is orthonormal with respect to the standard inner product in $S_n(\mathbb{R})$. Let $\bar{V}_s, T_R$ and $Y_r$ be as in Lemma 5.3. Identify $\bar{V}_s$ with $V_{\nu s}/\{ \pm I \}$. It follows from Lemma 5.3 that $V_{\nu s}$ is smooth. To show that $V_{\nu s}$ is connected, it suffices to show we can connect some $A \in V_{\nu s}$ to $-A$ in $V_{\nu s}$.

Take $y \in S^{n-1}$ such that $\bar{y} \in Y_r$. Connect $y$ to $-y$ by a path $\mathcal{J}$ (in $S^{n-1}$) such that $\mathcal{J} \subset Y_r$. For $z \in S^{n-1}$ such that $\bar{z} \notin T_R$, let

$$\hat{\alpha}(z) = \frac{1}{||\alpha(z)||} \alpha(z),$$

where $\hat{\alpha}(z)$ is the normalized vector at $z$. Then $\hat{\alpha}(z) \in V_{\nu s}$. This completes the proof.

**Corollary 5.2.** Let $n$ and $V$ satisfy the assumptions of Lemma 5.3. Let

$$V_\nu = \{ A \in V : \text{det}(A^2) = 1 \}$$

and

$$V_{\nu s} = \{ A \in V_\nu : \det A = 0 \}.$$

Then $V_{\nu s}$ is a smooth connected hypersurface in $V_\nu$.
where $\alpha(z)$ is given by (5.3). Since $\alpha$ is odd, $A(\hat{\alpha}(z)), z \in \mathcal{J}$, connects $A(\hat{\alpha}(y))$ to $A(\hat{\alpha}(-y)) = -A(\hat{\alpha}(y))$. \hfill \Box

**Theorem 5.1.** Let $n$ be odd, $n \geq 3$, and let $V$ be an $n + 1$ dimensional subspace of $S_n(\mathbb{R})$ such that each nonzero matrix in $V$ has rank $n - 1$ or $n$. Then each nonzero matrix in $V$ has at least $\frac{n-1}{2}$ positive and negative eigenvalues.

**Proof.** Let $A_1, A_2, \ldots, A_{n+1}$ be a basis of $V$. We distinguish two cases.

**Case 1.** Suppose that $V_{\perp}$ does not contain a matrix of rank one. Let $V_{\nu s}$ be as in Corollary 5.1. Then, each matrix in $V_{\nu s}$ has rank $n - 1$, and by the connectedness of $V_{\nu s}$ it follows that $A$ and $-A$ have the same inertia for each $A \in V_{\nu s}$. It follows that every singular matrix in $V\backslash\{0\}$ has $\frac{n-1}{2}$ positive eigenvalues, $\frac{n-1}{2}$ negative eigenvalues, and 1 zero eigenvalue. Continuity yields immediately that each $0 \neq A \in V$ has at least $\frac{n-1}{2}$ positive and negative eigenvalues.

**Case 2.** Suppose now that $V_{\perp}$ does contain a rank one matrix. Note that we have $\dim \mathbb{P}(V_{\perp}) = \frac{n(n+1)}{2} - (n + 1) - 1$, while $\dim \mathbb{P}V_{1,n}(\mathbb{R}) = n - 1$. Hence, for any $\epsilon > 0$ there exist $A_{1,\epsilon}, A_{2,\epsilon}, \ldots, A_{n+1,\epsilon} \in S_n(\mathbb{R})$ which satisfy the following three conditions:

(i) $\|A_{i,\epsilon} - A_i\| < \epsilon, i = 1, 2, \ldots, n + 1$.

(ii) $V_\epsilon = \text{span}\{A_{1,\epsilon}, A_{2,\epsilon}, \ldots, A_{n+1,\epsilon}\}$ is $n + 1$ dimensional.

(iii) $V_{\epsilon_{\perp}}$ does not contain a rank one matrix.

For $\epsilon > 0$ small enough, we can apply Case 1 and conclude that any nonzero matrix in $V_\epsilon$ has at least $\frac{n-1}{2}$ positive and negative eigenvalues. Let $\epsilon \to 0$ and use continuity to finish the proof. \hfill \Box

6. Concluding remarks and conjectures.

**Theorem 6.1.** Let $1 \leq q \in \mathbb{Z}$. Then $\deg \mathbb{P}V_{n-q,n}(\mathbb{C})$ is odd if $n > q$ and

$$n \equiv \pm q \pmod{2^{\lceil \log_2 q \rceil}}. \tag{6.1}$$

**Proof.** Let $0 \neq q \in \mathbb{Z}$. Let $\mu(q) = \lceil \log_2 2q \rceil$ and $0 \leq \nu(q) \in \mathbb{Z}$ be the largest integer such that $2^{\nu(q)}$ divides $q$. Then $\nu(q) + 1 \leq \mu(q)$. Clearly if $\ell \equiv m \not\equiv 0 \pmod{2^\mu}$ for some $1 \leq \mu \in \mathbb{Z}$ then $\nu(\ell) = \nu(m)$. Let $\delta_{k,n} := \deg \mathbb{P}V_{k,n}(\mathbb{C})$. We claim that for $n$ satisfying (6.1)

$$\nu(\delta_{n-q,n}) = \sum_{j=0}^{q-1} \nu\left(\binom{\hat{q} + j}{q - j}\right) - \nu\left(\binom{2j + 1}{j}\right), \quad \hat{q} = \pm q. \tag{6.2}$$
Here for any real \( x \) and \( 0 \leq k \in \mathbb{Z} \) we let
\[
\binom{x}{k} = \frac{x(x-1) \ldots (x-k+1)}{k!}.
\]
In particular
\[
(6.3) \quad \binom{-x}{k} = (-1)^k \binom{x+k-1}{k}, \quad k = 1, \ldots.
\]
Observe that for \( j \leq q - 1 \)
\[
\binom{n+j}{q-j} = \frac{(n+j)(n+j-1) \ldots (n-q+2j+1)}{1 \cdot 2 \ldots (q-j)}.
\]
For \( n \) satisfying (6.1) it follows that \( n + \ell \equiv \pm q + \ell \not\equiv 0 \pmod{2^\mu(q)} \) if \(-q + 1 \leq \ell \leq q - 1\). Hence (6.2) holds.

Our theorem is equivalent to the statement that \( \nu(\delta_{n-q,n}) = 0 \) if (6.1) holds. We first consider the case \( n \equiv -q \pmod{2^\mu(q)} \). Then
\[
\nu(\delta_{n-q,n}) = \sum_{j=0}^{q-1} \nu\left( \binom{-q+j}{q-j} - \nu\binom{2j+1}{j} \right)
= \sum_{j=0}^{q-1} \nu\left( (-1)^{q-j} \binom{2(q-j)-1}{q-j} \right) - \sum_{j=0}^{q-1} \nu\left( \binom{2j+1}{j} \right)
= \sum_{j=0}^{q-1} \nu\left( \binom{2(q-j)-1}{q-j-1} \right) - \sum_{j=0}^{q-1} \nu\left( \binom{2j+1}{j} \right) = 0.
\]
We now consider the case \( n \equiv q \pmod{2^\mu(q)} \). In this case it is enough to show the identity
\[
(6.4) \quad \prod_{j=0}^{q-1} \frac{(q+j)}{(2j+1)} = 1, \quad q = 1, \ldots.
\]
We prove the above identity by induction on \( q \). For \( q = 1 \) (6.4) trivially holds. Assume that (6.4) holds for \( q = m \geq 1 \). Let \( q = m + 1 \). Use the identity
\[
\binom{m+1+j}{m+1-j} = \frac{m+1+j}{m+1-j} \binom{m+j}{m-j}, \quad j = 0, 1, \ldots, m,
\]
to deduce (6.4) for \( q = m + 1 \) from \( q = m \).

\[\square\]

**Corollary 6.1.** Let \( n > q \) and let (6.1) hold. Then
\[
d(n, n-q, \mathbb{R}) = d(n, n-q, \mathbb{C}) = \binom{q+1}{2} + 1.
\]
Corollary 6.2. Let $n \geq q \geq 2$ and assume that (6.1) holds. Then any \((\frac{q+1}{2})\) dimensional subspace $U$ of $S_n(\mathbb{R})$ contains a nonzero matrix with an eigenvalue of multiplicity $q$ at least. Furthermore, there exists an \((\frac{q+1}{2}) - 1\) dimensional subspace $V \subset S_n(\mathbb{R})$ such that each eigenvalue of $0 \neq A \in V$ has a multiplicity less than $q$.

Proof. For $n = q$, $U = S_n(\mathbb{R})$ and the eigenvalue 1 of $I_q$ has multiplicity $q$. Assume that $V \subset S_q(\mathbb{R})$ does not contain $I_q$. Then each eigenvalue of $0 \neq A \in V$ has multiplicity less than $q$. Assume that $q < n$. Then the claim of the corollary follows from Corollary 6.1 and Proposition 3.3.

Clearly $\delta_{n-1,n} = n$ is odd if and only if $n$ is odd. As $d(n,n-1,\mathbb{R}) = \rho_s(n)+1$ it follows that $d(n,n-1,\mathbb{R}) > d(n,n-1,\mathbb{C}) = 2$ if $n$ is even.

Conjecture 6.1. Let $n > q \geq 2$ and assume that (6.1) does not hold. Then $\delta_{n-q,n}$ is even. Furthermore $d(n,n-q,\mathbb{R}) > d(n,n-q,\mathbb{C})$.

For $q = 2,3,4,5$ it is straightforward to check that $\delta_{n-q,n}$ is odd if and only if (6.1) holds.

Finally, we return to the sequence \(\{d(n,n-2,\mathbb{R})\}_{n=3}^\infty\) and state two questions. The first unknown number in this sequence is $d(5,3,\mathbb{R})$. As we have seen in Section 3, $d(5,2,\mathbb{R}) = 7$, so $d(5,3,\mathbb{R}) \leq 7$.

Question 1. Is $d(5,3,\mathbb{R}) = 6$?

We think that the answer is yes. In [6] it is shown that $d(5,3,\mathbb{R}) \leq 6$. Let $L$ be the 5 dimensional subspace of $S_5(\mathbb{R})$ spanned by the matrices

\[
B_1 = \begin{bmatrix}
1 & 1 & 1 & 1 & -1 \\
1 & 1 & -1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 \\
-1 & 1 & -1 & 1 & 1
\end{bmatrix}, \quad B_2 = \begin{bmatrix}
1 & 1 & 1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 \\
-1 & 1 & 1 & 1 & 1 \\
-1 & -1 & 1 & 1 & 1
\end{bmatrix},
\]

\[
B_3 = \begin{bmatrix}
-1 & 1 & 1 & -1 & -1 \\
1 & -1 & 1 & 1 & 1 \\
1 & -1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 & 1 \\
-1 & 1 & -1 & 1 & 1
\end{bmatrix}, \quad B_4 = \begin{bmatrix}
1 & 1 & -1 & 1 & -1 \\
1 & -1 & -1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 \\
-1 & 1 & 1 & 1 & 1 \\
1 & -1 & -1 & -1 & 1
\end{bmatrix},
\]

\[
B_5 = \begin{bmatrix}
1 & 1 & 1 & 1 & -1 \\
1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1
\end{bmatrix}.
\]

A local minimization procedure using Matlab seems to indicate that every nonzero matrix in $L$ has rank 4 or 5. This has yet to be confirmed by other means, but if it is correct, then $d(5,3,\mathbb{R}) = 6$. 

Note that for every field \( F \) we have
\[
\begin{align*}
    d(n, 1, F) & \geq d(n, 2, F) \geq \cdots \geq d(n, n-2, F) \geq d(n, n-1, F).
\end{align*}
\] (6.5)

**Question 2.** Is there strict inequality everywhere in (6.5) if \( F = \mathbb{R} \)?

**Acknowledgement.** We would like to thank H. Tracy Hall for his help in the computations using Maple associated with the example that is discussed in Theorem 2.1.

**References**


Received March 5, 2001 and revised August 6, 2001

DEPARTMENT OF INDUSTRIAL ENGINEERING
TECHNION – ISRAEL INSTITUTE OF TECHNOLOGY
E-mail address: dfalikm@ie.technion.ac.il

DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE
UNIVERSITY OF ILLINOIS AT CHICAGO
CHICAGO, IL 60607-7045
The second author was a Lady Davis Visiting Professor at the Technion in the Fall of 2000. The research of the third author was supported by the Fund for the Promotion of Research at the Technion.
ON THE COMMUTATOR FORMULA OF A SPLIT BN-PAIR

GWENAËLLE GENET

The purpose of this note is to prove in an elementary way and with geometric considerations that the Levi decomposition of a finite group with a split BN-pair of characteristic $p$ (a prime integer) implies the commutator formula.

1. Introduction.

Split BN-pairs are defined by a set of axioms devised to study finite reductive groups (see [CR, 69.1], [R, 3.1]). Such finite groups $G$ are assumed to contain subgroups $B$, $N$ and $S \subseteq W := N/T$, with $T := N \cap B$, such that among other things, $T \trianglelefteq N$, $(W,S)$ is a Coxeter system, $B$ can be written as a semi-direct $B = UT$ where $U$ is a Sylow $p$-subgroup of $G$ and $B \cap B^{w_0} = T$ ($w_0$ the longest element of $W$). Recalling the geometric representation of $W$ in an euclidean space $E$, we denote by $\Phi$ the associated root system (a subset of the unit sphere), by $\Delta \subseteq \Phi^+$ the fundamental and positive systems of $\Phi$, so that the fundamental reflections $\{s_\delta, \delta \in \Delta\}$ correspond with the elements of $S$ (see [CR, 64.28] or [B]). For $\delta \in \Delta$, one defines $X_\delta := U \cap U^{w_0 s_\delta}$. An elementary consequence of the axioms of BN-pairs ([CR, 69.2]) is:

(R0) For arbitrary $\gamma \in \Phi$, written as $\gamma = w(\delta)$ with $\delta \in \Delta$, $w \in W$, one may define $X_\gamma := w X_\delta$. This only depends on $\gamma$ and $X_\gamma \neq \{1\}$.

A nice extra property, satisfied in all finite reductive groups, is the commutator formula:

(1) $\forall \alpha, \beta \in \Phi$, $\alpha \neq \pm \beta$, $[X_\alpha, X_\beta] \subseteq \langle X_{i\alpha + j\beta}, i > 0, j > 0, i\alpha + j\beta \in \Phi \rangle$.

This is useful to check the crucial property of Levi decompositions, which is that for all subsets $I \subseteq \Delta$, $U \cap U^{w_I}$ is normal in $U$ ($w_I$ the longest element of $W_I$). In particular,

(2) $\forall \delta \in \Delta$, $U \cap U^{s_\delta} \trianglelefteq U$.

N. Tinberg ([T]) has shown that (2) implies in an elementary way the Levi decomposition.

Here, we show a little more, namely that (1) follows from (2), without using the classification of BN-pairs or the case of rank 2 ([FS]). Note that (2)
is always satisfied when the root subgroups $X_\delta$ have order $p$. Our arguments are easy considerations in the reflection representation space. We recall two elementary properties of BN-pairs. Denote by $l_S$ the length in $W$ relative to $S$.

(R1) (See [CR, 69.2] or [R, 3.3] by iteration.) There is at least one sequence such that $\Phi^+ = \{\gamma_1, \ldots, \gamma_N\}$, $N = |\Phi^+|$ and $U = X_{\gamma_1} \cdots X_{\gamma_N}$. If $\gamma \in \Phi^+$, then $X_{-\gamma} \cap U = \{1\} = U \cap U^{u_0}$.

(R2) ([R, 2.3]) If $w \in W$ and $s \in S$ is such that $l_S(ws) = l_S(w) + 1$, then $U \cap U^{ws} \subseteq U \cap U^s$.

The author thanks M. Cabanes for his suggestions.

Notation. If $E'$ is a subset of $E$, we denote $X(E') := \langle X_\gamma, \gamma \in E' \cap \Phi \rangle$ and if $A$ is a nonempty subset of $W$, we denote $V_A := \cap_{w \in A} U^w$.

2. First results.

The following proposition shows how to express $V_A$, $A \subseteq W$, $A \neq \emptyset$, with the root groups $X_\gamma$, $\gamma \in \Phi$. For this, we need:

Lemma 1. Let $m \geq 1$, let $\gamma_1, \ldots, \gamma_m$ be pairwise distinct in $\Phi^+$ and for every $i$, $1 \leq i \leq m$, let $x_i \in X_{\gamma_i} \setminus \{1\}$. Let $w \in W$, then $x_1 x_2 \cdots x_m \in U^w$ if and only if for every $i$, $w(\gamma_i) \in \Phi^+$.

Proof. If for all $i$, $w(\gamma_i) \in \Phi^+$, then for all $i$, $X_{w(\gamma_i)} = X_{\gamma_i} \subseteq U$ because of (R1). We prove the converse by induction on $l_S(w)$. If $w = 1$, this is obvious.

If $w = s_\delta$, $\delta \in \Delta$, one must check $\delta$ is none of the $\gamma_i$'s. Suppose on the contrary that $\delta = \gamma_{i_0}$. Then $s_\delta(\gamma_i) \in \Phi^+$ for any $i \neq i_0$. So $x_1 \ldots x_{i_0} \in U^{s_\delta}$ and $x_{i_0+1} \ldots x_m \in U^{s_\delta}$ by the ‘if’. Then $x_{i_0} \in U^{s_\delta} \cap X_{\delta} \subseteq U^{s_\delta} \cap U^{u_{i_0}s_\delta} = \{1\}$, a contradiction.

For an arbitrary $w \in W$ with $l_S(w) \geq 1$, we can write $w = w's_\delta$ for $w' \in W$, $\delta \in \Delta$ and $l_S(w') = l_S(w) - 1$. Then $U \cap U^{w'} \subseteq U \cap U^{s_\delta}$ by (R2). We have just seen that, for all $i$, $s_\delta(\gamma_i) \in \Phi^+$. Set $\gamma'_i := s_\delta(\gamma_i)$, $x'_i := x_i^{s_\delta} \in X_{\gamma'_i}$ where $s_\delta \in N$ is a representative of $s_\delta$. The result comes by applying the induction hypothesis to $x'_1 \cdots x'_m \in U^{w'}$.

Proposition 1. For all nonempty subset $A \subseteq W$, we have $V_A = X(\Psi_A)$ where $\Psi_A := \{\gamma \in \Phi \mid \forall w \in A, w(\gamma) \in \Phi^+\}$. Moreover, $\Psi_A = \{\gamma \in \Phi, X_\gamma \subseteq V_A\}$.

Proof. Suppose $1 \in A$, so that $\Psi_A \subseteq \Phi^+$. Take a sequence $\gamma_1, \ldots, \gamma_N \in \Phi^+$ as in (R1), so that $\Psi_A = \{\gamma_{i_1}, \ldots, \gamma_{i_m}\}$ for $1 \leq i_1 < \cdots < i_m \leq N$. Then any $x \in V_A \subseteq U$, can be written as $x = x_1 \cdots x_N$ with $x_i \in X_{\gamma_i}$. The above lemma implies that $x_i = 1$ whenever $\gamma_i \notin \Psi_A$. Then $V_A = X(\Psi_A)$. So it makes clear that $\Psi_A \subseteq \{\gamma \in \Phi, X_\gamma \subseteq V_A\}$. But if $\gamma \in \Phi$ is such that
Proof. One inclusion is clear. For the other, let $x$ find a linear form $f$ such that $f(C) \subseteq \mathbb{R}_+^*$ and $0 \notin f(\Psi)$. □

The following is a slight adaptation to our needs of the standard theorem about separation of convex sets.

**Lemma 2.** Let $C$ be a closed convex cone of $E$ such that $C \cap -C = \{0\}$. Let $\Psi$ be a finite set of $E \setminus \{0\}$ and let $F$ be the subset of the dual $E^\vee$ of $E$, $F = \{f \in E^\vee : f(C) \subseteq \mathbb{R}_+^* \text{ and } 0 \notin f(\Psi)\}$. Then $C = \bigcap_{f \in F} f^{-1}(\mathbb{R}_+^*) \cup \{0\}$.

**Proof.** One inclusion is clear. For the other, let $x \notin C$, so $x \neq 0$. We will find a linear form $f$ such that $f(C) \subseteq \mathbb{R}_+^*$, $f(x) < 0$ and $0 \notin f(\Psi)$. One may assume that $x$ has norm 1. If $r > 0$, let $C_r$ (resp. $D_r$) be the open cone generated by the elements of the unit sphere at distance $< r$ from the elements of $C$ (resp. from $x$). For $r$ sufficiently small, we clearly have $C_r \cap D_r = C_r \cap -C_r = \emptyset$. Taking a hyperplane separating the open convex sets $C_r$ and $D_r$, it is clear that such a hyperplane must contain 0. This tells us that the set $G := \{f \in E^\vee, f(C) \subseteq \mathbb{R}_+^*, f(x) < 0\}$ is not empty. Now, it is easy to see that $G$ is an open set of the dual $E^\vee$ and that the set $\{f \in G, 0 \notin f(\Psi)\} \neq G$ since $\Psi$ is finite. So the outcome is still nonempty. Thus our claim.

The consequence of the above on “convex” subsets of $\Phi$ and corresponding subgroups of $G$ is as follows.

**Proposition 2.** Let $C$ be a closed convex cone of $E$ such that $C \cap -C = \{0\}$. Denote $A(C) := \{w \in W \mid C \cap \Phi \subseteq w^{-1}(\Phi^+)\}$. Then

(i) $C \cap \Phi = \Psi_{A(C)}$ with the notation of Proposition 1.

(ii) $X(C) = V_{A(C)}$.

**Proof.** (i) The inclusion $C \cap \Phi \subseteq \Psi_{A(C)} = \cap_{w \in A(C)} w^{-1}(\Phi^+)$ is trivial. On the other hand, we have $C \cap \Phi = \cap_{f \in F} f^{-1}(\mathbb{R}_+^*) \cap \Phi$ where the set $F$ consists of the $f \in E^\vee$ such that $f(C) \subseteq \mathbb{R}_+^*$ and $0 \notin f(\Phi)$ (Lemma 2). But for $f \in F$, $f^{-1}(\mathbb{R}_+^*) \cap \Phi$ defines an order relation on $\Phi$ ([B]) and therefore some unique fundamental system. By transitivity of $W$ on fundamental systems, there is $w_f \in W$ such that $f^{-1}(\mathbb{R}_+^*) \cap \Phi = w_f^{-1}(\Phi^+)$. Obviously, $w_f \in A(C)$. So we get the reverse inclusion we seek.

(ii) By (i) above, we have $X(C) = X(\Psi_{A(C)})$. But Proposition 1 tells us that the latter is indeed $V_{A(C)}$.

**Corollary 1.** Let $C$ and $D$ be two closed convex cones of $E$ such that $C \cap -C = D \cap -D = \{0\}$. Then, $X(C) \cap X(D) = X(C \cap D)$.

**Proof.** With Proposition 2 (ii), we clearly have $X(C) \cap X(D) = V_{A(C)} \cap V_{A(D)} = V_{A(C) \cup A(D)}$, and the latter is $X(\Psi_{A(C) \cup A(D)})$ by Proposition 1. But
Theorem 1. If $G$ is a finite group with a split BN-pair satisfying the hypothesis (2), then it satisfies the commutator formula (1).

Proof. For all finite subset $Y \subseteq E$, we denote by $C(Y)$ the closed convex cone generated by $Y$. We use the abbreviation $C(y, z) := C(\{y, z\})$.

Suppose $G$ satisfies (2). Denote $U_\delta = U \cap U^{s_\delta}$ when $\delta \in \Delta$. Let $\alpha \in \Delta$ and $\beta \in \Phi^+$, $\alpha \neq \beta$. We have $U_\alpha = V_{\{1, s_\alpha\}} = X(\Psi_{\{1, s_\alpha\}})$ and $X_\beta \subseteq U_\alpha$ by Proposition 1 because $\Psi_{\{1, s_\alpha\}} = \Phi^+ \setminus \{\alpha\}$. But $\Phi^+ \setminus \{\alpha\} = \Phi \cap C(\Phi^+ \setminus \{\alpha\})$ because any positive linear combination of positive roots that is again a root is a positive one and $\alpha$ is a fundamental root, so is a minimal positive linear combination of positive roots. Therefore, $U_\alpha = X(C(\Phi^+ \setminus \{\alpha\}))$. Besides, since $X_\alpha \subseteq U$, $[X_\alpha, X_\beta] \subseteq [X_\alpha, U_\alpha] \subseteq U_\alpha$ by hypothesis (2).

On the other hand, both $X_\alpha$ and $X_\beta$ are subgroups of $X(C(\alpha, \beta))$, so $[X_\alpha, X_\beta] \subseteq X(C(\alpha, \beta))$. The closed convex cones $X(C(\Phi^+ \setminus \{\alpha\}))$ and $X(C(\alpha, \beta))$ satisfy $C \cap -C = \{0\}$ by property of positive roots ([B]) then Corollary 1 implies $[X_\alpha, X_\beta] \subseteq U_\alpha \cap X(C(\alpha, \beta)) = X(C(\Phi^+ \setminus \{\alpha\}) \cap C(\alpha, \beta)) = X(C(\alpha, \beta) \setminus \{\alpha\})$. So we get, for all $\alpha \in \Delta$, $\beta \in \Phi^+ \setminus \{\alpha\}$,

$$
[X_\alpha, X_\beta] \subseteq X(C(\alpha, \beta) \setminus \{\alpha\}).
$$

If $\alpha \in \Delta$, $\beta \in \Phi^+ \setminus \{-\alpha\}$, we also have (3) by applying the above to $w_0 s_\alpha(\alpha) \in \Delta$, $w_0 s_\alpha(\beta) \in \Phi^+$ and conjugating the corresponding subgroups of $G$ by $s_\alpha w_0 \in W$. So we have (3) for any $\alpha \in \Delta$, $\beta \in \Phi \setminus \{\pm \alpha\}$.

If $\alpha$, $\beta$ are any non-proportional arbitrary roots, there is $w \in W$ such that $w(\alpha) \in \Delta$, then again we have (3). Now, exchanging the rôles of $\alpha$ and $\beta$, (3) becomes $[X_\alpha, X_\beta] \subseteq X(C(\alpha, \beta) \setminus \{\beta\})$. It is clear that each set $\Phi \cap (C(\alpha, \beta) \setminus \{\alpha\})$, $\Phi \cap (C(\alpha, \beta) \setminus \{\beta\})$, and $\Phi \cap (C(\alpha, \beta) \setminus \{\alpha, \beta\})$ is of the type $\Phi \cap C$ where $C$ is a closed convex cone such that $C \cap -C = \{0\}$ (draw a picture in the plane generated by $\alpha$ and $\beta$). Then Corollary 1 gives $X(C(\alpha, \beta) \setminus \{\alpha\}) \cap X(C(\alpha, \beta) \setminus \{\beta\}) = X(C(\alpha, \beta) \setminus \{\alpha, \beta\})$. We get our claim.

3. The commutator formula.

References


Received January 8, 2001 and revised May 4, 2001.

UFR de Mathématiques - Case 7012
Université Denis Diderot - Paris 7
2, Place Jussieu
75251 Paris Cedex 05
France
E-mail address: genet@math.jussieu.fr
PHRAGMÈN–LINDELÖF THEOREM FOR MINIMAL SURFACE EQUATIONS IN HIGHER DIMENSIONS

Chun-Chung Hsieh, Jenn-Fang Hwang, and Fei-Tsen Liang

Here we prove that if \( u \) satisfies the minimal surface equation in an unbounded domain which is properly contained in a half space of \( \mathbb{R}^n \), with \( n \geq 2 \), then the growth rate of \( u \) is of the same order as that of the shape of \( \Omega \) and the boundary value of \( u \).

1. Introduction.

Consider the minimal surface equation
\[
\text{div } T u = 0,
\]
where
\[
T u = \frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \quad \text{and} \quad \nabla u = (u_x, \ldots, u_{x_n}).
\]
In 1965, Nitsche [7] announced the following result: “Let \( \Omega_\alpha \subset \mathbb{R}^2 \) be a sector with angle \( 0 < \alpha < \pi \). If \( u \) satisfies the minimal surface equation with vanishing boundary value in \( \Omega_\alpha \), then \( u \equiv 0 \)”. Hwang extends this result in [4], [5], [6] and proves that, in an unbounded domain \( \Omega \) properly contained in the half plane in \( \mathbb{R}^2 \), if \( u \) satisfies the minimal surface equation, then, the growth property of \( u \) is determined completely by the shape of \( \Omega \) and the boundary value of \( u \). In this respect, the Phragmèn-Lindelöf theorem for the minimal surface equation is better than that for the Laplace equation. (Indeed, if \( u \) satisfies the Laplace equation in an unbounded domain \( \Omega \), the growth property of \( u \) cannot be determined completely by the shape of \( \Omega \) and the boundary data of \( u \) alone (cf. [10]).)

The purpose of this paper is to generalize the two-dimensional Phragmèn-Lindelöf theorems in [4], [5] and [6], to higher dimensions. In §2, we review the statements of the Phragmèn-Lindelöf theorem of [4], [5] and [6]. The higher-dimensional version is similar in content, but proof is different. In §3, based on an argument of [2], we established the suitable comparison principle. In §4, we compute the mean curvature of our comparison function, and use it to finish the proof of our main theorems in §5.
2. Preliminary.

The main purpose of this paper is to generalize the two-dimensional Phragmèn-Lindelöf theorem in [4], [5], [6] to higher dimensions. We may, first of all, recall some results in these papers and consider functions

\[ f : [0, \infty) \to [0, \infty), \quad f \in C^2([0, \infty)), \quad f' \equiv \frac{df(y)}{dy} > 0, \]

from which we define

\[ p(f) = 1 - \frac{f f''}{(f')^2}. \]

In particular, for \( f(y) = y^m \), \( m \) being a positive constant, we have

\[ p(f) = \frac{1}{m}, \]

which is precisely the reciprocal of the order of \( f \), while for \( f(y) = e^y \), we have

\[ p(f) = 0; \]

moreover, in case \( f \) grows faster than the exponential function, we can assume \( p(f) \geq -\epsilon \) for some small positive constant \( \epsilon \), essentially (cf. [5, Remark 2.7]). Accordingly, we may proceed to solve the ordinary differential equation in \([-1, 1]\)

\[
(1 - p(f))(h - th')(1 + h'^2) + h''(h^2 + t^2) = 0
\]

with initial values

\[
h(-1) = 0 \quad \text{and} \quad h'(-1) = \tan \left( \left(1 - p(f)\right)\frac{\pi}{2} \right),
\]

and then denote its solution, if exists, by \( h_m \) if \( f(y) = y^m \) (and hence \( p(f) = \frac{1}{m} \)), and by \( h_\infty \) if \( f(y) = e^y \) (and hence \( p(f) = 0 \)). In general, (*) and (**) cannot be solved explicitly; but, for some specific \( m \), its solution can be written out explicitly. For example, we have

\[ h_2 = \frac{1 - t^2}{2}, \]

and also

\[ h_\infty = \sqrt{1 - t^2}. \]

It is useful to know some interesting properties of \( h_m, 0 < m \leq \infty \), in the following:

**Lemma 1** ([6]). For \( 1 < m, m' \leq \infty \) and \( t \in (-1, 1) \), then we have

(i) \( h_m(t) > h_{m'}(t) \), whenever \( m > m' \),

and

(ii) \( h_m(t) < h_{m'}(t') \), whenever \( |t| > |t'| \).
The Phragmèn-Lindelöf theorems in [5], [6] can now be formulated as follows.

**Theorem 2.** Let \( \Omega \subseteq \{(x, y) \in \mathbb{R}^2 | -ay^m < x < ay^m, y > 0\} \subseteq \mathbb{R}^2 \) be an unbounded domain, where \( a \) and \( m \) are positive constants, \( m \geq 1 \). Let \( u \in C^2(\Omega) \cap C^0(\overline{\Omega}) \) and suppose that

\[
\begin{align*}
\text{div } Tu & \geq 0 \quad \text{in } \Omega \\
u & \leq ay^m h_m \left( \frac{x}{ay^m} \right) \quad \text{on } \partial \Omega.
\end{align*}
\]

Then we have

\[
u \leq ay^m h_m \left( \frac{x}{ay^m} \right) \leq ay^m h_\infty \left( \frac{x}{ay^m} \right) = \sqrt{a^2 y^{2m} - x^2} \text{ in } \Omega.
\]

**Theorem 2*.** Let \( \Omega \subseteq \{(x, y) \in \mathbb{R}^2 | -ae^b y < x < ae^b y, y > 0\} \) be an unbounded domain where \( a, b \) are positive constants. Let \( u \in C^2(\Omega) \cap C^0(\overline{\Omega}) \) and suppose that

\[
\begin{align*}
\text{div } Tu & \geq 0 \quad \text{in } \Omega \\
u & \leq \sqrt{a^2 e^{2b} y - x^2} \quad \text{on } \partial \Omega.
\end{align*}
\]

Then we have \( u \leq \sqrt{a^2 e^{2b} y - x^2} \) in \( \Omega \).

**Theorem 3.** Let \( f \in C^2([0, \infty)), f > 0, f' > 0 \) in \( (0, \infty) \) and \( p(f) \geq p_0 \), where \( p_0 \) is a negative constant, and let \( f_1 \in C^0([0, \infty)) \) and \( f_1 > 0 \) in \( (0, \infty) \). For a given unbounded open domain

\[
\Omega \subset \{(x, y) \in \mathbb{R}^2 | -f_1(y) < x < f_1(y), y > 0\},
\]

and \( u \in C^2(\Omega) \cap C^0(\overline{\Omega}) \) with

\[
\begin{align*}
\text{div } Tu & \geq 0 \quad \text{in } \Omega \\
u & \leq a \sqrt{f^2 - x^2} \quad \text{on } \partial \Omega,
\end{align*}
\]

where \( f^2 \geq \frac{(a^2-1)(2-p_0)}{(a^2-(1-p_0))} f_1^2 \) and \( a \) is a positive constant satisfying

\[
a^2 - 1 + p_0 > 0.
\]

Then, we have

\[
u \leq a \sqrt{f^2 - x^2} \quad \text{in } \Omega.
\]

**Remark.** In Theorem 3, since \( p_0 < 0 \) and \( a > 0 \), we have

\[
\frac{(a^2-1)(2-p_0)}{(a^2-(1-p_0))} = \left( \frac{a^2-1}{a^2-(1-p_0)} \right)(2-p_0) > 2.
\]

Thus, in case \( u \leq 0 \) on \( \partial \Omega \), our estimates are not good enough since we use worse boundary conditions, whereas the best estimates remain unknown.

These theorems will be generalized to higher dimensions in §5.
3. A comparison principle.

To establish the higher-dimensional Phragmèn-Lindelöf theorem, we shall need the following comparison principle.

**Lemma 4.** Let $\Omega$ be an unbounded domain in $\mathbb{R}^n$, and let $u, v \in C^2(\Omega) \cap C^0(\overline{\Omega})$. Suppose that

\[
\begin{aligned}
\text{div } Tu - \text{div } Tv &\geq C & \text{in } \Omega, \\
\text{ } u &\leq v & \text{on } \partial \Omega,
\end{aligned}
\]

where $C$ is a positive constant. Then we have $u \leq v$ in $\Omega$.

**Proof.** The idea of proof is analogous to that of [2].

Suppose that this lemma fails to hold. There then exists a positive constant $\epsilon$ such that

\[
\Omega' = \{x \in \Omega \mid u(x) > v(x) + \epsilon\}
\]

is not empty; by Sard’s theorem, we may further assume that $\partial \Omega' \cap \Omega$ is smooth. For every $R > 0$, set

\[
\begin{aligned}
B_R &= \{x \in \mathbb{R}^n \mid |x| < R\}, \\
\Omega_R &= B_R \cap \Omega', \\
\Gamma_R &= \partial B_R \cap \partial \Omega_R,
\end{aligned}
\]

and

\[
|\Gamma_R| = \text{the Hausdorff } (n-1) - \text{dimensional measure of } \Gamma_R.
\]

Also, let

\[
g(R) = \int_{\partial \Omega_R} \tan^{-1}(u - v - \epsilon)(Tu - Tv) \cdot \nu
\]

\[
= \int_{\Gamma_R} \tan^{-1}(u - v - \epsilon)(Tu - Tv) \cdot \nu
\]

where $\nu$ is the unit outward normal of $\partial \Omega_R$.

Then we have

\[
g(R) = \int_{\Omega_R} \frac{(\nabla u - \nabla v) \cdot (Tu - Tv)}{1 + (u - v - \epsilon)^2} + \int_{\Omega_R} \tan^{-1}(u - v - \epsilon)(\text{div } Tu - \text{div } Tv).
\]

Since the integrand of the right-hand side of (1) is nonnegative, Fubini’s theorem tells us that $g'(R)$ exists for almost all $R > 0$, and whenever it
exists, we have, by (2),

\[
\begin{align*}
g'(R) &= \int_{\Gamma_R} \frac{(\nabla u - \nabla v) \cdot (Tu - Tv)}{1 + (u - v - \epsilon)^2} \\
&\quad + \int_{\Gamma_R} \tan^{-1}(u - v - \epsilon)(\text{div } Tu - \text{div } Tv) \\
&\geq C \int_{\Gamma_R} \tan^{-1}(u - v - \epsilon), \quad \text{(by assumption)} \\
&\geq \frac{C}{2} \int_{\Gamma_R} \tan^{-1}(u - v - \epsilon) |Tu - Tv|, \\
&\quad \text{(since } |Tu| < 1 \text{ and } |Tv| < 1) \\
&\geq \frac{C}{2} g.
\end{align*}
\]

Since $g$ is an increasing function of $R$ and $g \geq 0$, it is easy to see that Lemma 4 holds in the case that $g \equiv 0$. If, on the other hand, $g \not\equiv 0$, there would exist a positive constant $R_0$ such that $g(R) > 0$ for all $R \geq R_0$, and hence, for every $R > R_0$, in virtue of (3)

\[
\int_{R_0}^R \frac{g'(r)}{g(r)} \, dr \geq \frac{C}{2} (R - R_0),
\]

i.e.,

\[
\left. \log g(r) \right|_{R_0}^R \geq \frac{C}{2} (R - R_0),
\]

and therefore,

\[
g(R) \geq g(R_0) e^{\frac{C}{2} (R - R_0)}.
\]

However, we have, by (1)

\[
g(R) \leq \int_{\Gamma_R} \frac{\pi}{2} \cdot 2 \leq \pi |\Gamma_R|;
\]

since $\Gamma_R \subset \partial B_R$, this yields a positive constant $C_1$ completely determined by $n$ such that

\[
g(R) \leq C_1 R^{n-1},
\]

which contradicts (4) and yields the truth of Lemma 4. $\square$

**Remark.** The above proof works well and so the lemma is valid if $v = +\infty$ on some parts of $\partial \Omega$. 


Henceforth, we will denote \( \Omega \) as an unbounded domain in \( \mathbb{R}^n, n \geq 2 \), such that, for some \( f \in C^2 ([0, \infty)) \), \( f > 0 \), \( f' > 0 \) and \( f'' > 0 \) in \( (0, \infty) \), we have

\[
\Omega \subset \{(x, y) \in \mathbb{R}^2 \mid -f(y) < x < f(y), \ y > 0\} \times \mathbb{R}^{n-2} \subset \mathbb{R}^n.
\]

We shall extend the results in §2 to such a domain \( \Omega \).

First, for every positive constant \( y_0 \), since \( f > 0, f' > 0 \) and \( f'' > 0 \) in \( (0, \infty) \), it is easy to see that there exists a positive constant \( \delta_1 \), depending on \( y_0 \), such that \( \{(f(y), y) \in \mathbb{R}^2 \mid y > 0\} \cap \{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 = 0\} \) has exactly one point. And also, \( \{(f(y), y) \in \mathbb{R}^2 \mid y > 0\} \cap \{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 = 0\} \) has exactly two points, say \((f(y_1), y_1)\) and \((f(y_2), y_2)\) with \( 0 < y_1 < y_2 \), for all \( \delta \) with \( 0 < \delta < \delta_1 \). In general, we have \( y_1 = y_1(y_0, \delta), \ y_2 = y_2(y_0, \delta) \) and also \( \lim_{y \to 0} y_1(y_0, \delta) = y_0 \). From now on, we always assume that the positive constant \( \delta \) is less than the above \( \delta_1 \).

To apply Lemma 4 to estimate the speed of growth of solutions in \( \Omega \), we may consider comparison functions of the following form

\[
F_{y_0, \delta} = \frac{A(f^2(y) - x^2)^{\frac{1}{2}}}{y_0 - y + \frac{\delta}{2} x^2},
\]

which is defined on

\[
\Omega_{y_0, \delta} = \Omega \cap \left( \{ (x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 > 0, \ 0 < y < y_1 \} \times \mathbb{R}^{n-2} \right),
\]

where \( \delta, y_0, \) and \( A \) are positive constants. We first proceed to calculate the mean curvature of \( F_{y_0, \delta} \). For convenience of computation, we may set

\[
F = A \cdot P^2 Q^{-1},
\]

where \( P = f^2(y) - x^2 \) and \( Q = y_0 - y + \frac{\delta}{2} x^2 \). We observe that

\[
\text{div} \ TF = \frac{(1 + F_x^2) F_{yy} - 2 F_x F_y F_{xy} + (1 + F_y^2) F_{xx}}{(1 + F_x^2 + F_y^2)^{\frac{3}{2}}}
\]

\[
= \frac{1}{P^2} \left( \frac{F^2}{P^2} F_{yy} - 2 \frac{F_x}{P^2} F_x F_{xy} + \frac{1}{P^2} \left( \frac{F^2}{P^2} F_{xx} \right) \right)
\]

\[
= \left( \frac{1}{P^2} + \frac{F_y^2}{P^2} \right) F_{yy} - 2 \frac{F_x}{P} F_y F_{xy} + \left( \frac{1}{P^2} + \frac{F_x^2}{P^2} \right) F_{xx}.
\]

Denoting

\[
I = \frac{F_x^2}{F^2} \frac{F_{yy}}{F} + \frac{F_y^2}{F^2} \frac{F_{xx}}{F} - 2 \frac{F_x}{F} \frac{F_y}{F} F_{xy}
\]

and

\[
II = \frac{F_{xx}}{F^3} + \frac{F_{yy}}{F^3},
\]
we note that the numerator in (5) is the sum of these two expressions and we shall treat them separately. For the first expression, we have

\[ I = \frac{F^2_x}{F^2} \left( \frac{\partial_y}{F} \left( \frac{F_y}{F} \right) + \left( \frac{F_y}{F} \right)^2 \right) + \frac{F^2_y}{F^2} \left( \frac{\partial_x}{F} \left( \frac{F_x}{F} \right) + \left( \frac{F_x}{F} \right)^2 \right) \]

\[ - 2 \frac{F_x F_y}{F^2} \left[ \frac{\partial_x}{F} \left( \frac{F_y}{F} \right) + \frac{F_x F_y}{F^2} \right] \]

\[ = \frac{F^2_x}{F^2} \left( \frac{\partial_y}{F} \left( \frac{F_y}{F} \right) \right) + \frac{F^2_y}{F^2} \left( \frac{\partial_x}{F} \left( \frac{F_x}{F} \right) \right) - 2 \frac{F_x F_y}{F^2} \left( \frac{\partial_x}{F} \left( \frac{F_y}{F} \right) \right) \]

\[ = I^* + I^{**} \]

where

\[ I^* = \frac{F^2_x}{F^2} \left( - \frac{1}{2} \frac{P^2_y}{P^2} + \frac{Q^2_y}{Q^2} \right) + \frac{F^2_y}{F^2} \left( - \frac{1}{2} \frac{P^2_x}{P^2} + \frac{Q^2_x}{Q^2} \right) \]

\[ - 2 \frac{F_x F_y}{F^2} \left( - \frac{1}{2} \frac{P_x P_y}{P^2} + \frac{Q_x Q_y}{Q^2} \right), \]

and

\[ I^{**} = \frac{F^2_x}{F^2} \left( \frac{1}{2} \frac{P_y}{P} - \frac{Q_y}{Q} \right) + \frac{F^2_y}{F^2} \left( \frac{1}{2} \frac{P_x}{P} - \frac{Q_x}{Q} \right) \]

\[ - 2 \frac{F_x F_y}{F^2} \left( \frac{1}{2} \frac{P_{xy}}{P} - \frac{Q_{xy}}{Q} \right). \]

By a direct computation,

\[ I^* = -\frac{1}{4} \frac{1}{P^2 Q^2} (P_y Q_x - P_x Q_y)^2, \]

while

\[ I^{**} = \left( \frac{1}{2} \frac{P_x}{P} - \frac{Q_x}{Q} \right)^2 \left( \frac{1}{2} \frac{P_y}{P} - \frac{Q_y}{Q} \right) \]

\[ + \left( \frac{1}{2} \frac{P_y}{P} - \frac{Q_y}{Q} \right)^2 \left( \frac{1}{2} \frac{P_x}{P} - \frac{Q_x}{Q} \right) \]

\[ - 2 \left( \frac{1}{2} \frac{P_x}{P} - \frac{Q_x}{Q} \right) \left( \frac{1}{2} \frac{P_y}{P} - \frac{Q_y}{Q} \right) \left( \frac{1}{2} \frac{P_{xy}}{P} - \frac{Q_{xy}}{Q} \right). \]

Thus, in particular, we have

(6) \quad I^* \leq 0.

As for \( I^{**} \) and \( II \), we recall that

\[ P = f^2(y) - x^2 \quad \text{and} \quad Q = y_0 - y + \frac{\delta}{2} x^2, \]
and hence
\[ P_x = -2x, \quad P_y = 2f'f, \quad Q_x = \delta x, \quad Q_y = -1; \]
moreover
\[ P_{xx} = -2, \quad P_{xy} = 0, \quad P_{yy} = 2(f''f + f'f') \]
and
\[ Q_{xx} = \delta, \quad Q_{xy} = 0, \quad Q_{yy} = 0. \]
Thus, we have
\[ I^{**} = \frac{1}{P^3} \left[ x^2(f''f + f'^2) - f^2f'^2 \right] + \frac{(-2)}{P^2Q} \left[ -\delta x^2(f''f + f'^2) + f'f + \frac{\delta}{2}f^2f'^2 \right] + \frac{1}{PQ^2} \left[ \delta^2x^2(f''f + f'^2) - 1 - 2\delta ff' \right] - \delta Q^{-3}, \]
and also
\[ II = \frac{Q^2}{A^2P} \left( \frac{\partial_x \left( \frac{F_x}{F} \right)}{F} + \frac{\partial_y \left( \frac{F_y}{F} \right)}{F^2} + \frac{F_y}{F} \right) \]
\[ = \frac{Q^2}{A^2P} \left( \partial_x \left( \frac{1}{2} \frac{P_x}{P} - \frac{Q_x}{Q} \right) + \partial_y \left( \frac{1}{2} \frac{P_y}{P} - \frac{Q_y}{Q} \right) \right) \]
\[ = \frac{Q^2}{A^2} \left( \frac{1}{P^3} \left[ f^2(f''f - 1) - x^2(f f'' + f'^2) \right] + \frac{1}{P^2Q} \left( 2ff' - \delta f^2 + 3\delta x^2 \right) + \frac{2}{PQ^2}(\delta^2x^2 + 1) \right). \]
Thus the numerator of \( \text{div} \, TF \) is
\[ (7) \quad I + II = I^* + I^{**} + II \]
\[ = \frac{1}{P^3} \left[ x^2(f f'' + f'^2) - f^2f'^2 + \frac{Q^2}{A^2}((f^2(f f'' - 1) - x^2(f f'' + f'^2))) \right] \]
\[ + \frac{1}{P^2Q} \left[ \delta x^2(f f'' + (f')^2) - 2ff' - \delta^2f^2f'^2 \right. \]
\[ + \frac{Q^2}{A^2}(-\delta f^2 + 2ff' + 3\delta x^2) \]
\[ + \frac{1}{PQ^2} \left[ \delta^2x^2(f''f + (f')^2) - 1 - 2\delta ff' + \frac{2}{A^2}(\delta^2x^2 + 1) \right] \]
\[ - \delta Q^{-3} - \frac{1}{4}P^{-2}Q^{-2}(P_xQ_y - Q_xP_y)^2. \]
We want to choose $\delta$ and $A$ to make the third bracket of the right-hand side of (7) negative. For this, substituting the expression for $Q$ in the bracket and rewriting it as

\begin{equation}
III = -1 + \left( \frac{2}{A^2} \right) \left( y_0 - y + \frac{\delta x^2}{2} \right)^2 (1 + \delta^2 x^2) + \delta^2 x^2 (f'' f + (f')^2) - 2\delta f f'.
\end{equation}

For any given $\lambda$, $0 < \lambda < \frac{1}{4}$, if we take $\delta$ such that

\begin{equation}
0 < \delta < \inf_{y \in (0, y_1)} \min \left\{ \frac{\lambda y_0}{f^2}, \frac{\lambda}{f}, \frac{\lambda}{f (f'' f + (f')^2)^{\frac{\lambda}{2}}} \right\}
\end{equation}

and $A = 4\sqrt{y_0}$, then we have

\begin{equation}
III \leq -1 + \frac{1}{4(2y_0)^2} \left( y_0 + \frac{\lambda y_0}{2} \right)^2 (1 + \lambda^2) + \lambda^2 \leq -1 + \frac{1 + \lambda^2}{4} + \lambda^2 < 0.
\end{equation}

As of the second bracket of the right-hand side of (7), to make it negative, it clearly suffices to make the following expression negative, namely

\begin{equation}
IV = \left( \frac{Q^2}{A^2} \right) \left( f f' + \frac{3}{2} \delta x^2 \right) - f f' + \delta x^2 (f f'' + (f')^2).
\end{equation}

For this, we observe that, as $x^2 < f^2$ in $\Omega$ and $f > 0$, $f' > 0$ and $f'' > 0$ in $(0, \infty)$,

\begin{equation}
\frac{3}{2} \delta x^2 + f f' \leq \frac{3}{2} \delta f^2 + f f' = f f' \left( 1 + \frac{3}{2} \delta \frac{f^2}{f f'} \right),
\end{equation}

while

\begin{equation}
-f f' + \delta x^2 (f f'' + (f')^2) \leq -f f' + \delta f^2 (f f'' + (f')^2)
\end{equation}

\begin{equation}
= -f f' \left( 1 - \frac{f^2 (f f'' + (f')^2)}{f f'} \right)
\end{equation}

and furthermore, if we require that

\begin{equation}
(9^*) \delta < \inf_{y \in (0, y_1)} \min \left\{ \frac{\lambda f f'}{f^2 (f f'' + (f')^2)^{\frac{\lambda}{2}}}, \frac{\lambda f f'}{2f^2} \right\}
\end{equation}

it follows from (9) that

\begin{equation}
\frac{Q^2}{A^2} \leq \frac{1}{4}
\end{equation}

And also, we have

\begin{equation}
IV \leq f f' \left( \frac{Q^2}{A^2} (1 + \lambda) + \lambda - 1 \right) \leq f f' \left( \frac{1}{4} (1 + \lambda) + \lambda - 1 \right) \leq -\frac{1}{4} f f'.
\end{equation}

Thus, the condition that $f > 0$, $f' > 0$ in $(0, \infty)$ ensures us of the negativity of (10). It remains to consider the first bracket of the right-hand side of (7).
To make it negative, it suffices to make negative the following expression
\[ V = x^2(f f'' + (f')^2) - f^2 f'^2 + \frac{Q^2}{A^2} f^2 (f f'' - 1), \]
or, in view of (11),
\[ V \leq x^2(f f'' + (f')^2) - f^2 f'^2 + \frac{1}{4} f^2 (f f''). \]
Recall that for given function \( f \) as above, we define
\[ p(f) = 1 - \frac{f f''}{(f')^2}. \]
For \( \S 5 \), and from now on, we assume that \(-1 \leq p(f) \leq 1 \), following a remark concerning \( p(f) \) for our interesting functions, \([5, \text{Remark 2.7}] \). And so, in particular for \( f(y) = (y + z)^m \), \( p(f) = \frac{1}{m} \) and for \( f(y) = ae^{by} \), \( p(f) = 0 \) where \( z, m > 1, a, \) and \( b \) are positive constants, and also it is easy to see that for \( f(y) = e^{\alpha y} \), with \( \alpha > 1 \), then \( p(f) \to 0^- \) as \( y \to +\infty \).

Rewriting (12) in terms of \( p(f) \), and noticing that \( \left( \frac{3}{4} + \frac{p}{4} \right) \times \left( \frac{1}{2-p} \right) \geq \frac{1}{6} \), we have
\[ V \leq (2-p)(f')^2 \left( x^2 - \frac{3}{2} + \frac{p}{2} f^2 \right) \leq (2-p)(f')^2 \left( x^2 - \frac{1}{6} f^2 \right), \]
and so if we assume furthermore that
\[ \Omega \subseteq \left\{ (x, y) \in \mathbb{R}^2 \mid - \frac{1}{\sqrt{6}} f(y) < x < \frac{1}{\sqrt{6}} f(y), y > 0 \right\} \times \mathbb{R}^{n-2} \subseteq \mathbb{R}^n, \]
then \( V \leq 0 \) and get the following conclusion about the estimation of our comparison function: If \( F = \sqrt{f^2(y) - x^2} \left( \frac{4\sqrt{2}y_0}{(y_0 - y + \frac{\sqrt{2}}{2} x^2)} \right) \) with \( \delta \) as in our assumptions, \((9), (9^*) \), then \( \text{div} TF \leq 0 \) in \( \Omega_{y_0, \delta} \), where \( \Omega \) is assumed as in (14). Now we state what we achieved as follows:

**Proposition 5.** Let \( f_1 : [0, \infty) \to [0, \infty) \), and \( f_1 \in C^2([0, \infty)) \) with \( f_1 > 0, f_1' > 0, f_1'' > 0 \) on \( [0, \infty), \) and \(-1 \leq p(f_1) \leq 1 \). Suppose that \( \Omega \subseteq \{(x, y) \in \mathbb{R}^2 \mid - f_1(y) < x < f_1(y), y > 0 \} \times \mathbb{R}^{n-2} \subseteq \mathbb{R}^n \) and that \( u \in C^2(\Omega) \cap C^0(\overline{\Omega}) \) and for some constant \( \beta \) with \( 0 < \beta < 1 \) satisfying
\[
\begin{align*}
\text{div} \, Tu &\geq 0 & \text{in } \Omega \\
u &\leq 4\sqrt{2} \beta \sqrt{6 f_1^2(y) - x^2} & \text{on } \partial \Omega.
\end{align*}
\]
Then \( u \leq 4\sqrt{2} \sqrt{6 f_1^2(y) - x^2} \) in \( \Omega \).

**Proof.** Set \( f(y) = \sqrt{6} f_1(y) \) and define \( F(x, y) = 4\sqrt{2} y_0 \left( \frac{f^2(y) - x^2}{(y_0 - y + \frac{\sqrt{2}}{2} x^2)} \right) \) as above, where \( y_0 > 0 \) and \( \delta > 0 \), small as in \((9) \) and \((9^*) \) and we also require
that \( \delta \leq \frac{(2 - 2\beta)y_0}{\beta(f(y_1))^2} \). Then following the computation as above, in particular that of (7), and also noticing that the first three brackets of the right-hand side of (7) are negative in \( \Omega_{y_0,\delta} \) as shown above, it is easy to see that

\[
\text{div } TF = \left( \frac{1}{F^2} + \frac{F_y^2}{F^2} \right) F_{yy} - \frac{2}{F^2} \frac{F_y F_x F_{xy}}{F} + \left( \frac{1}{F^2} + \frac{F_x^2}{F^2} \right) F_{xx} < -\delta \left( y_0 - y + \frac{\delta}{2} x^2 \right)^{-3},
\]

and

\[
\left( \frac{1}{F^2} + \frac{F^2}{F_x^2} \right) \frac{F_{yy}}{F} - \frac{2}{F^2} \frac{F_x F_y F_{xy}}{F} + \left( \frac{1}{F^2} + \frac{F_y^2}{F^2} \right) \frac{F_{xx}}{F} < -\delta \left( y_0 - y + \frac{\delta}{2} x^2 \right)^{-3},
\]

when \((x, y)\) is close to \(\{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 = 0\}\).

And so, noticing that \(P = f^2(y) - x^2\), \(Q = y_0 - y + \frac{\delta}{2} x^2\) and \(A = 4\sqrt{2}y_0\), when \((x, y)\) is close to \(\{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 = 0\}\), we have

\[
\text{div } TF \leq -\delta Q^{-3} \left( \frac{1}{F^2} + \left| \nabla F \right|^2 \right)^{-\frac{3}{2}}
\leq -\delta Q^{-3} \left( \frac{Q^2}{A^2 P} + \left| \frac{1}{2} \nabla P - \nabla Q \right|^2 \right)^{-\frac{3}{2}}
\leq -\delta \left( \frac{Q^4}{A^2 P} + \left| \frac{1}{4} \nabla P \cdot \nabla Q \right|^2 \right)^{-\frac{3}{2}}
\leq -\delta \left( \frac{Q^4}{A^2 P} + \left| \nabla P \right|^2 + \left| \nabla Q \right|^2 - \frac{Q}{P} \nabla P \cdot \nabla Q \right)^{-\frac{3}{2}}
\leq -\frac{\delta}{2} (1 + \delta^2 x^2)^{-\frac{3}{2}},
\]

since \(-\frac{Q}{P} \nabla P \cdot \nabla Q \geq 0\) and \(\left| \nabla Q \right|^2 = 1 + \delta^2 x^2\).

But the bounded connected component of the closure of \(\{(x, y) \in \mathbb{R}^2 \mid -f_1(y) < x < f_1(y), y > 0\} \cap \{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 > 0\}\), which is denoted as \(\Omega^*\), is compact. And we have \(\Omega_{y_0,\delta} \subseteq \Omega^* \times \mathbb{R}^{n-2}\), and so, there exists a positive constant \(c\), such that

\[
\begin{cases}
\text{div } TF \leq -c & \text{in } \Omega_{y_0,\delta}, \\
F \geq u & \text{on } \partial \Omega_{y_0,\delta} \cap \{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 > 0\} \times \mathbb{R}^{n-2}, \\
F = +\infty & \text{on } \partial \Omega_{y_0,\delta} \cap \{(x, y) \in \mathbb{R}^2 \mid y_0 - y + \frac{\delta}{2} x^2 = 0\} \times \mathbb{R}^{n-2}.
\end{cases}
\]

Now, by Lemma 4, we have \(u \leq F\) in \(\Omega_{y_0,\delta}\), which is

\[
u(x, y, z_1, \ldots, z_{n-2}) \leq 4\sqrt{2}y_0 \left( \frac{6f^2_1(y) - x^2}{2} \right)^{\frac{1}{2}} \quad \text{in } \Omega_{y_0,\delta}.
\]
Now, let $\delta \to 0$ and then let $y_0 \to +\infty$, we get the conclusion of the proof. $\square$

5. Phragmèn-Lindelöf theorem in higher dimensions.

First, let’s generalize Theorem 2 as follows:

**Theorem 6.** Let $\Omega \subseteq \{(x, y) \in \mathbb{R}^2 \mid -ay^m < x < ay^m, y > 0\} \times \mathbb{R}^{n-2} \subseteq \mathbb{R}^n$ be an unbounded domain, where $m \geq 1$ and $a$ are positive constants. Let $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ and suppose that 

$$
\begin{cases}
\text{div} \, Tu \geq 0 & \text{in } \Omega \\
u \leq ay^m h_m(\frac{x}{ay^m}) & \text{on } \partial \Omega.
\end{cases}
$$

Then we have $u \leq ay^m h_m(\frac{x}{ay^m}) \leq ay^m h_\infty(\frac{x}{ay^m}) = \sqrt{a^2 y^{2m} - x^2}$ in $\Omega$.

**Proof.** For every given positive constant $\epsilon > 0$, we now set $f_\epsilon(x, y) = a(y + \epsilon)^m + c$, $F_\epsilon(x, y, z_1, z_2, \ldots, z_{n-2}) = a(y + \epsilon)^m h_m(\frac{x}{a(y + \epsilon)^m})$, where $(x, y, z_1, z_2, \ldots, z_{n-2}) \in \Omega$.

Since $-ay^m < x < ay^m, y > 0$, we have

$$
\left| \frac{x}{a(y + \epsilon)^m} \right| \leq \frac{y^m}{(y + \epsilon)^{m+\epsilon}} \to 0 \quad \text{as} \quad y \to +\infty.
$$

By Lemma 1, $h_m(\frac{x}{a(y + \epsilon)^m}) \to h_m(0)$ uniformly as $y \to +\infty$, and so it is easy to see that there exists a large constant $y_3$ such that $F_\epsilon(x, y) \geq \sqrt{200a^2 y^{2m} - x^2}$ for $y \geq y_3$.

Next by [6, Theorem 2], setting $f_\epsilon(y) = a(y + \epsilon)^{m+\epsilon}, t = \frac{x}{f_\epsilon(y)}$ and recalling that $p(f_\epsilon) = \frac{1}{m+\epsilon},$ we have

$$
\text{div} \, TF_\epsilon = (1 + | \nabla F_\epsilon |^2)^{-\frac{1}{2}} \frac{(f_\epsilon')^2}{f_\epsilon} \cdot \left( (1 - p(f_\epsilon))(h_{m+\epsilon} - th_{m+\epsilon})(h_{m+\epsilon}'(h_{m+\epsilon}')(2 + 1) + h_m h_m + t^2 + h_m'(f_\epsilon')^2) \right).
$$

Since $h_m(\epsilon)$ is the solution of $(\ast)$ and $(\ast\ast)$ with $p(f_\epsilon) = \frac{1}{m+\epsilon}$, we have

$$
\text{div} \, TF_\epsilon = (1 + | \nabla F_\epsilon |^2)^{-\frac{1}{2}} \frac{(f_\epsilon')^2}{f_\epsilon} \cdot \frac{h_m'(f_\epsilon')^2}{(f_\epsilon')^2}
$$

and so obviously that $\text{div} \, TF_\epsilon < 0$ on $\Omega'$ where $\Omega' = \Omega \cap \{ (x, y, z_1, \ldots, z_{n-2}) \in \mathbb{R}^n | 0 < y < y_3 \}$.

And so, there exists a positive constant $C_1 > 0$ such that

$$
\text{div} \, TF_\epsilon \leq -C_1 \quad \text{on } \Omega'.
$$
But, noticing that
\[ u \leq ay^{m}h_{m}\left(\frac{x}{ay^{m}}\right) \leq \sqrt{a^{2}y^{2m} - x^{2}} \leq 4\sqrt{2}\beta\sqrt{6a^{2}y^{2m} - x^{2}}, \]
for some constant \( \beta < 1 \) on \( \partial\Omega \), by Proposition 5, we also have
\[ u \leq 4\sqrt{2}\sqrt{6a^{2}y^{2m} - x^{2}} \leq \sqrt{200} \cdot a^{2}y^{2m} - x^{2} \]
in \( \Omega \setminus \Omega' \).

By Lemma 4, we have
\[ u \leq F_{\epsilon} \quad \text{in} \quad \Omega'. \]
In conclusion, we have
\[ u \leq F_{\epsilon} \quad \text{in} \quad \Omega, \]
and let \( \epsilon \to 0 \), the proof is done. \( \square \)

As a corollary of Theorem 6, we state a generalization of Nitsche’s theorem [7] as follows.

**Corollary.** Let \( \Omega = \{(x,y) \in \mathbb{R}^2 \mid -ay < x < ay, \ y > 0\} \times \mathbb{R}^{n-2} \) be a wedge domain, where \( a \) is a positive constant. Let \( u \in C^2(\Omega) \cap C^{0}(\overline{\Omega}) \) and suppose that
\[
\begin{cases}
\text{div} \ T u = 0 & \text{in} \ \Omega, \\
u = 0 & \text{on} \ \partial\Omega.
\end{cases}
\]
Then \( u \equiv 0 \) in \( \Omega \).

**Proof.** Apply Theorem 6 to functions \( u \) and \(-u\), we have \( u \leq 0 \) in \( \Omega \) and \(-u \leq 0 \) in \( \Omega \), and so \( u \equiv 0 \) as claimed. \( \square \)

Next, let’s generalize Theorem 2* as follows:

**Theorem 6*.** Let \( \Omega \subseteq \{(x,y) \in \mathbb{R}^2 \mid -ae^{by} < x < ae^{by}, \ y > 0\} \times \mathbb{R}^{n-2} \subseteq \mathbb{R}^n \), where \( a, b \) are positive constants. Let \( u \in C^2(\Omega) \cap C^{0}(\overline{\Omega}) \) and suppose that
\[
\begin{cases}
\text{div} \ T u \geq 0 & \text{in} \ \Omega \\
u \leq \sqrt{a^2e^{2by} - x^2} & \text{on} \ \partial\Omega.
\end{cases}
\]
Then we have \( u \leq \sqrt{a^2e^{2by} - x^2} \) in \( \Omega \).

**Proof.** The proof is similar to that of Theorem 6.
For every \( \epsilon > 0 \), we consider the following function
\[ F_{\epsilon}(x, y, z_{1}, z_{2}, \ldots, z_{n-2}) = ae^{(b+\epsilon)y}h_{\infty}\left(\frac{x}{ae^{(b+\epsilon)y}}\right) = \sqrt{a^2e^{2(b+\epsilon)y} - x^2} \]
with \( (x, y, z_{1}, \ldots, z_{n-2}) \in \Omega \).
Since \(-ae^{by} < x < ae^{by}, \ y > 0\), we have
\[
\left|\frac{x}{ae^{(b+\epsilon)y}}\right| \leq \frac{ae^{by}}{ae^{(b+\epsilon)y}} \to 0 \quad \text{as} \quad y \to +\infty
\]
and notice that \( F_{\epsilon} = ae^{(b+\epsilon)y}(1 - \frac{x^2}{a^2e^{2(b+\epsilon)y}})^{\frac{1}{2}} \).
Hence, there exists a positive constant $y_3 > 0$ such that
\[
F_{\epsilon} \geq \sqrt{200a^2e^{2by} - x^2} \quad \text{for } y \geq y_3.
\]

Next, by [6, Theorem 2], and setting $f_{\epsilon}(y) = a e^{(b+\epsilon)y}$, $t = \frac{x}{f(y)}$, and noticing that $p(f_{\epsilon}) = 0$, we have
\[
\text{div } TF_{\epsilon} = (1 + |\nabla F_{\epsilon}|^2)^{-\frac{3}{2}} \frac{(f'_{\epsilon})^2}{f_{\epsilon}} \cdot \left( (h_{\infty} - th_{\infty}'(2h_{\infty}^2 + 1) + h_{\infty}'(h_{\infty}^2 + t^2) + \frac{h_{\infty}''}{f_{\epsilon}^2}) \right)
= (1 + |\nabla F_{\epsilon}|^2)^{-\frac{3}{2}} \frac{(f'_{\epsilon})^2}{f_{\epsilon}} \frac{h_{\infty}''}{(f'_{\epsilon})^2}.
\]
So, we have
\[
\text{div } TF_{\epsilon} < 0 \quad \text{on } \Omega',
\]
where
\[
\Omega' = \Omega \cap \{(x, y, z_1, \ldots, z_{n-2}) \in \mathbb{R}^n \mid 0 < y < y_3\},
\]
and so there exists a positive constant $C_1 > 0$ such that
\[
\text{div } TF_{\epsilon} \leq -C_1 \quad \text{in } \Omega'.
\]
Finally, by Proposition 5, notice that
\[
u \leq \sqrt{a^2e^{2by} - x^2} \leq 4\sqrt{2}\beta \sqrt{6a^2e^{2by} - x^2},
\]
for some constant $\beta < 1$ on $\partial \Omega$, we also have
\[
u \leq 4\sqrt{2}\sqrt{6a^2e^{2by} - x^2} \leq F_{\epsilon} \quad \text{in } \Omega \setminus \Omega'.
\]
So, by Lemma 4, we have
\[
u \leq F_{\epsilon} \quad \text{on } \Omega',
\]
and so obviously, we get
\[
u \leq F_{\epsilon} \quad \text{in } \Omega,
\]
and let $\epsilon \to 0$, the proof is finished.

Finally, let's generalize Theorem 3 as follows:

**Theorem 7.** Let $f_1 \in C^2([0, \infty))$ with $f_1 \geq 0$, $f_1' > 0$, and $f_1'' \geq 0$ in $(0, \infty)$ such that $p(f_1) \geq p_0$, where $p_0$ is a constant with $-1 \leq p_0 \leq 0$. Suppose that $\Omega \subseteq \{(x, y) \in \mathbb{R}^2 \mid -f_1(y) < x < f_1(y), y > 0\} \times \mathbb{R}^{n-2} \subseteq \mathbb{R}^n$ and $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$ satisfying
\[
\begin{cases}
\text{div } Tu \geq 0 & \text{in } \Omega \\
u \leq a\sqrt{f^2(y) - x^2} & \text{on } \partial \Omega,
\end{cases}
\]
where $f = \left(\frac{(a^2-1)(2-p_0)}{(a^2-1+p_0)}\right)^{\frac{1}{2}} f_1$ and $a$ is a positive constant with $a^2 - 1 + p_0 > 0$.

Then we have $u \leq a\sqrt{f^2(y) - x^2}$ in $\Omega$.  

Proof. For any given $\epsilon > 0$, we define $f_\epsilon(y) = e^{\epsilon y}f(y+\epsilon)$ and $F_\epsilon(x,y,z_1,\ldots,z_{n-2}) = a\sqrt{f_\epsilon^2(y) - x^2}$, then there exists $y_3 > 0$ such that

$$F_\epsilon \geq a(\epsilon^2 y f_\epsilon^2(y) - x^2)\frac{1}{2} \geq (200 \cdot a^2 f_\epsilon^2(y) - x^2)\frac{1}{2} \quad \text{for} \quad y > y_3.$$ 

Computing the mean curvature of $F_\epsilon$ and using the definition,

$$p(f_\epsilon) = 1 - \frac{f_\epsilon f_\epsilon''(f_\epsilon')}{2},$$

we have

$$\text{div} \, TF_\epsilon = (1 + |\nabla F_\epsilon|^2)^{-\frac{3}{2}}(f_\epsilon^2 - x^2)^{\frac{3}{2}} \cdot \left( a(f_\epsilon')^2[(a^2 - 1)(2 - p(f_\epsilon))x^2 - f_\epsilon^2(a^2 - 1 + p(f_\epsilon)) - a f^2_\epsilon) \right).$$

Obviously, we have

$$f_\epsilon^2(y) \geq f^2(y) \geq \frac{(a^2 - 1)(2 - p_0)}{(a^2 - 1 + p_0)} f^2_1(y) \geq \frac{(a^2 - 1)(2 - p(f_\epsilon))}{(a^2 - 1 + p_0)} f^2_1(y),$$

and so, we have

$$\text{div} \, TF_\epsilon \leq -a(1 + |\nabla F_\epsilon|^2)^{-\frac{3}{2}} f^2_\epsilon < 0 \quad \text{in} \quad \Omega,$$

and by compactness, there exists a positive constant $C_1 > 0$ such that

$$\text{div} \, TF_\epsilon \leq -C_1 \quad \text{in} \quad \Omega_1 = \{(x,y,z_1,z_2,\ldots,z_{n-2}) \in \Omega \mid y < y_3\}.$$

But by Proposition 5, notice that

$$u \leq a\sqrt{f^2(y) - x^2} \leq 4\sqrt{2}\beta \sqrt{6a^2 f^2 - x^2},$$

for some constant $\beta < 1$ on $\partial \Omega$, we also have

$$u \leq \sqrt{200a^2 f^2(y) - x^2} \leq F_\epsilon \quad \text{in} \quad \Omega \setminus \Omega_1.$$

By Lemma 4, we have

$$u \leq F_\epsilon \quad \text{in} \quad \Omega_1.$$

In conclusion, we have $u \leq F_\epsilon$ in $\Omega$, and then let $\epsilon \to \infty$.

We thus finish the proof. \hfill \Box

References


Received February 27, 2001.

Mathematics Department
Academia Sinica
Nan-kang
Taipei, 115, Taiwan
E-mail address: macchsieh@ccvax.sinica.edu.tw

Mathematics Department
Academia Sinica
Nan-kang
Taipei, 115, Taiwan
E-mail address: majfh@ccvax.sinica.edu.tw

Mathematics Department
Academia Sinica
Nan-kang
Taipei, 115, Taiwan
E-mail address: liang@math.sinica.edu.tw
REMARK ON THE RATE OF DECAY OF SOLUTIONS TO LINEARIZED COMPRESSIBLE NAVIER–STOKES EQUATIONS

Takayuki Kobayashi and Yoshihiro Shibata

We consider the $L_p-L_q$ estimates of solutions to the Cauchy problem of linearized compressible Navier–Stokes equation. Especially, we investigate the diffusion wave property of the compressible Navier–Stokes flows, which was studied by D. Hoff and K. Zumbrum and Tai-P. Liu and W. Wang.

1. Introduction.

In this paper, we consider the Cauchy problem of the following linearized compressible Navier-Stokes equations:

\[
\begin{align*}
\rho_t + \gamma \text{div} \, v &= 0 \quad \text{in } (0, \infty) \times \mathbb{R}^n, \\
v_t - \alpha \Delta v - \beta \text{div} \, v + \gamma \nabla \rho &= 0 \quad \text{in } (0, \infty) \times \mathbb{R}^n, \\
\rho|_{t=0} &= \rho_0, \quad v|_{t=0} = v_0 \quad \text{in } \mathbb{R}^n,
\end{align*}
\]

where $v(t, x) = T(v_1(t, x), \ldots, v_n(t, x))$ a vector valued unknown function, $\rho = \rho(t, x)$ is a scalar valued unknown function; $t$ is time variable; we denote the spatial point of $n$-dimensional Euclidian Space $\mathbb{R}^n$ by $x = (x_1, \ldots, x_n)$ ($n \geq 2$);

\[
\begin{align*}
\rho_t &= \frac{\partial \rho}{\partial t}, \\
v_t &= \frac{\partial v}{\partial t}, \\
\Delta v &= \sum_{j=1}^{n} \frac{\partial^2 v}{\partial x_j^2}, \\
\text{div} \, v &= \sum_{j=1}^{n} \frac{\partial v_j}{\partial x_j}, \\
\nabla \rho &= \left( \frac{\partial \rho}{\partial x_1}, \ldots, \frac{\partial \rho}{\partial x_n} \right);
\end{align*}
\]

$\rho_0$ and $v_0$ are given initial data; $\alpha$ and $\gamma$ are positive constants and $\beta$ a non-negative constant. Concerning the decay property, asymptotically, the solution decomposed into sum of two parts under the influence of a hyperbolic aspect and a parabolic aspect. One of which dominates in $L_p$ for $2 \leq p \leq \infty$, the other $1 \leq p < 2$. For $p \geq 2$, the time asymptotic behavior of solutions is similar to the solution of pure diffusion problem. Namely, the decay at the rate of the solution is similar to the solution of a linear, second order, strictly parabolic system with $L_1$ initial data. Moreover, the decay order of the term that is given by the convolution of Green functions of diffusion equation and
wave equation is better than the solution to pure diffusion system. On the other hand, for $p < 2$, the asymptotically dominant term reflects the spreading effect of the solution operator for the standard multi-dimensional wave equation. As a result, the solution may grow without bound in $L^p$ for $p < 2$.

This result was investigated by D. Hoff and K. Zumbrun \cite{2, 3} in the case of the Navier-Stokes system describing the compressible fluid flow, and Y. Shibata \cite{6} in the case of the linear viscoelastic equation. D. Hoff and K. Zumbrun \cite{2, 3} considered the linear effective artificial viscosity system as the first approximation of the compressible Navier-Stokes equation in several space dimension. The Green function of this system is written exactly by the convolution of the Green function of diffusion equation and wave equation. In view of this, they gave the pointwise estimate and $L^p$ estimate of the Green function in \cite{2, 3}, and $L^p$ estimate for the solutions to the nonlinear problem in \cite{2}. But, the Green function of the system (1.1) and the linear viscoelastic equation is not written exactly. Tai-P. Liu and W. Wang \cite{4} gave the pointwise estimate for the solutions to the system (1.1) and the nonlinear problem in odd multi-dimension case, and Y. Shibata \cite{6} gave the $L^p$ estimate for the solution to the linear viscoelastic equations by directly using Fourier transform method. The main difference of the structure to the solutions between (1.1) or effective artificial viscosity system and linear viscoelastic equation is the Riesz kernel $R_j(x) = F^{-1}\left[\xi_j/|\xi|\right](x)$, where $F^{-1}$ denotes the Fourier inverse transform. The Green matrix of the system (1.1) and effective artificial viscosity system includes the Riesz kernel. Since the convolution operator $u \to R_j * u$ is not bounded from $L^1$ to $L^1$ and from $L^\infty$ to $L^\infty$, if we consider $L^1$ or $L^\infty$ estimate, then these features will lead to a great deal of cancellation in the convolution operator of the Green function. D. Hoff and K. Zumbrun \cite{2} overcame this difficulty by applying the weak version of the Paley-Wiener theorem to the general, symmetrizable, hyperbolic-strictly parabolic systems. In this paper, we shall estimate directly using Fourier transform method in \cite{6}. In particular, we shall detect the cancellation in the Green function.

2. Main results.

First of all, we shall introduce the solution operator of (1.1). Applying the Fourier transform with respect to $x = (x_1, \ldots, x_n)$, (1.1) is reduced to the following ordinary differential equation with parameter $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$:

\[
\begin{aligned}
\frac{d\hat{\rho}}{dt}(t, \xi) + i\gamma \xi \cdot \hat{v}(t, \xi) &= 0, \\
\frac{d\hat{v}}{dt}(t, \xi) + \alpha |\xi|^2 \hat{v}(t, \xi) + \beta \xi (\xi \cdot \hat{v}(t, \xi)) + i\gamma \xi \hat{\rho}(t, \xi) &= 0, \\
\hat{\rho}(0, \xi) &= \hat{\rho}_0(\xi), \quad \hat{v}(0, \xi) = \hat{v}_0(\xi),
\end{aligned}
\]  

(2.1)
The roots \( \lambda \) of (2.3) are given by the formula
\[
\lambda_{\pm}(\xi) = -A \left( |\xi|^2 \pm \sqrt{|\xi|^4 - B^2|\xi|^2} \right),
\]
where \( A = (\alpha + \beta)/2, B = 2\gamma/(\alpha + \beta) \). When \( |\xi| \neq 0, B \), the solution of (2.2) is given by the formula
\[
\hat{\rho}(t, \xi) = \frac{\lambda_+(\xi)e^{\lambda_-(\xi)t} - \lambda_-(\xi)e^{\lambda_+(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \hat{\rho}_0(\xi) - i\gamma \frac{e^{\lambda_+(\xi)t} - e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \xi \cdot \hat{\nu}_0(\xi).
\]
Since \( \lambda_+(\xi) = \lambda_-(\xi) \) when \( |\xi| = B \), as the solution of (2.2), when \( B/2 < |\xi| < 2B \), we use the following formula
\[
\hat{\rho}(t, \xi) = \frac{1}{2\pi i} \int_{\Gamma} \frac{(z + |\xi|^2)e^{zt}}{z^2 + (\alpha + \beta)|\xi|^2z + \gamma^2|\xi|^2} \, dz \hat{\rho}_0(\xi)
+ \frac{i\gamma}{2\pi i} \int_{\Gamma} \frac{e^{zt}}{z^2 + (\alpha + \beta)|\xi|^2z + \gamma^2|\xi|^2} \, dz \xi \cdot \hat{\nu}_0(\xi),
\]
where \( \Gamma \) is a closed path containing \( \lambda_{\pm}(\xi) \) and contained in \( \{ z \in \mathbb{C} \mid \text{Re} z \leq -c_0 \} \) and \( c_0 \) is a positive number such that
\[
\max_{\frac{B}{2} \leq |\xi| \leq 2B} \text{Re} \lambda_{\pm}(\xi) \leq -2c_0.
\]
Also, by (2.1) we have
\[
\begin{aligned}
\left\{ \frac{d\hat{\rho}}{dt}(t, \xi) + \alpha|\xi|^2\hat{\nu}(t, \xi) &= \hat{f}(t, \xi), \\
\hat{\nu}(0, \xi) &= \hat{\nu}_0(\xi),
\end{aligned}
\]
where
\[
\hat{f}(t, \xi) = \frac{\xi}{i\gamma} \left\{ \beta \frac{d\hat{\rho}}{dt}(t, \xi) + \gamma^2\hat{\rho}(t, \xi) \right\}.
\]
Therefore, by (2.5) and (2.8), the solution of (2.6) given by the formula:
when $|\xi| \neq 0, B,$

$$
(2.9) \quad \hat{v}(t, \xi) = e^{-\alpha|\xi|^2t}\hat{\varphi}_0(\xi) + \int_0^t e^{-\alpha|\xi|^2(t-s)}\hat{f}(s, \xi)\,ds
$$

\begin{align*}
&= e^{-\alpha|\xi|^2t}\hat{\varphi}_0(\xi) - i\gamma\xi\left(\frac{e^{\lambda_+(\xi)t} - e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)}\right)\hat{\rho}_0(\xi) \\
&\quad + \left(\frac{\lambda_+(\xi)e^{\lambda_+(\xi)t} - \lambda_-(\xi)e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} - e^{-\alpha|\xi|^2t}\right)\frac{\xi \cdot \hat{\varphi}_0(\xi)}{|\xi|^2},
\end{align*}

and when $B/2 < |\xi| < 2B,$

$$
\hat{v}(t, \xi) = e^{-\alpha|\xi|^2t}\hat{\varphi}_0(\xi) - \frac{i\gamma\xi}{2\pi i} \int_{\Gamma} \frac{e^{zt}}{z^2 + (\alpha + \beta)|\xi|^2z + \gamma^2|\xi|^2} \,dz \hat{\rho}_0(\xi) \\
+ \left(\frac{1}{2\pi i} \int_{\Gamma} \frac{(z + |\xi|^2)e^{zt}}{z^2 + (\alpha + \beta)|\xi|^2z + \gamma^2|\xi|^2} \,dz - e^{-\alpha|\xi|^2t}\right)\frac{\xi \cdot \hat{\varphi}_0(\xi)}{|\xi|^2}.
$$

Let $\varphi_0(\xi), \varphi_M(\xi)$ and $\varphi_\infty(\xi)$ be functions in $C^\infty(\mathbb{R}^n)$ such that

$$
(2.10) \quad \varphi_0(\xi) = \begin{cases} 1 & |\xi| \leq B/2, \\ 0 & |\xi| \geq B/\sqrt{2}, \end{cases} \quad \varphi_\infty(\xi) = \begin{cases} 1 & |\xi| \geq 2B, \\ 0 & |\xi| \leq \sqrt{2B}, \end{cases} \\
\varphi_M(\xi) = 1 - \varphi_0(\xi) - \varphi_\infty(\xi).
$$

Put

$$
(2.11) \quad E_0(t) = (E_{0,\rho}(t), E_{0,v}(t)), \\
E_\infty(t) = (E_{\infty,\rho}(t), E_{\infty,v}(t)), \\
E_{0,\rho}(t)(\rho_0, v_0)(x) = \mathcal{F}^{-1}[\varphi_0(\xi)\hat{\rho}(t, \xi)](x), \\
E_{0,v}(t)(\rho_0, v_0)(x) = \mathcal{F}^{-1}[\varphi_0(\xi)\hat{v}(t, \xi)](x), \\
E_{\infty,\rho}(t)(\rho_0, v_0)(x) = \mathcal{F}^{-1}[(\varphi_M(\xi) + \varphi_\infty(\xi))\hat{\rho}(t, \xi)](x), \\
E_{\infty,v}(t)(\rho_0, v_0)(x) = \mathcal{F}^{-1}[(\varphi_M(\xi) + \varphi_\infty(\xi))\hat{v}(t, \xi)](x).
$$

Noting that $\varphi_M(\xi) = 1$ for $B/\sqrt{2} \leq |\xi| \leq \sqrt{2}B$ and $\varphi_M(\xi) = 0$ for $|\xi| \geq 2B$ or $|\xi| \leq B/2,$ by (2.10) and (2.11) we see that $(\rho(t, x), v(t, x)) = E_0(t)(\rho_0, v_0)(x)$ is a solution of (1.1). The main purpose of the paper is to show the following two theorems.
Theorem 2.1 \((L_1 - L_\infty \text{ and } L_1 - L_1 \text{ estimate of } E_0(t))\).

1. For any \(t > 0\), we have
   \[
   \|\partial_t^j \partial_x^\alpha E_{0,\rho}(t)(\rho_0, v_0)\|_{L_\infty(\mathbb{R}^n)} \\
   \leq C_{j,\alpha,n}(1 + t)^{-\left(\frac{3n-1}{4} + \frac{j + |\alpha|}{2}\right)} \left[\|\rho_0\|_{L_1(\mathbb{R}^n)} + \|v_0\|_{L_1(\mathbb{R}^n)}\right] \\
   + C_{j,\alpha,n}(1 + t)^{-\left(\frac{n}{2} + \frac{j + |\alpha|}{2}\right)} \|v_0\|_{L_1(\mathbb{R}^n)}.
   \]

Here and hereafter, we write
\[
\partial_t^j = \frac{\partial^j}{\partial t^j}, \quad \partial_x^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}},
\]
\(\alpha = (\alpha_1, \ldots, \alpha_n), \quad |\alpha| = \alpha_1 + \cdots + \alpha_n\),
\(C_{A,B,...}\) means the constant depending on \(A, B, \ldots\).

2. For any \(t > 0\), we have
   \[
   \|\partial_t^j \partial_x^\alpha E_{0}(t)(\rho_0, v_0)\|_{L_1(\mathbb{R}^n)} \\
   \leq C_{j,\alpha,n}(1 + t)^{q(n) - \left(\frac{3n-1}{4} + \frac{j + |\alpha|}{2}\right)} \left[\|\rho_0\|_{L_1(\mathbb{R}^n)} + \|v_0\|_{L_1(\mathbb{R}^n)}\right],
   \]
where
\[
q(n) = \begin{cases} 
\frac{n - 1}{4} & \text{if } n \geq 3 \text{ and } n \text{ is an odd number,} \\
\frac{n}{4} & \text{if } n \geq 2 \text{ and } n \text{ is an even number.}
\end{cases}
\]

Remark. The estimate (1) is better than [3, Theorem 1.2] when \(n = 2, j = 0\) and \(|\alpha| = 0\).

Theorem 2.2 \((L_1 - L_1 \text{ and } L_\infty - L_\infty \text{ estimate of } E_\infty(t))\). Let \(p = 1 \text{ or } \infty\). For any \(t > 0\), we have
\[
\|\partial_t^j \partial_x^\alpha E_{\infty,\rho}(t)(\rho_0, v_0)\|_{L_p(\mathbb{R}^n)} \\
\leq C_{j,\alpha,n} e^{-ct} \left[ C_{k} t^{-(j-k)}\|\rho_0\|_{W_p^{2k+|\alpha|-1}(\mathbb{R}^n)} + \|\rho_0\|_{W_p^{|\alpha|}(\mathbb{R}^n)} \right] \\
+ C_{j,\alpha,n} e^{-ct} \left[ C_{k} t^{-(j-k)}\|v_0\|_{W_p^{2k+|\alpha|}(\mathbb{R}^n)} + \|v_0\|_{W_p^{|\alpha|}(\mathbb{R}^n)} \right];
\]
\[
\|\partial_t^j \partial_x^\alpha E_{\infty}(t)(\rho_0, v_0)\|_{L_p(\mathbb{R}^n)} \\
\leq C_{j,\alpha,n} e^{-ct} \left[ C_{k} t^{-(j-k)}\|\rho_0\|_{W_p^{2k+|\alpha|}(\mathbb{R}^n)} + \|\rho_0\|_{W_p^{|\alpha|}(\mathbb{R}^n)} \right] \\
+ C_{j,\alpha,n} e^{-ct} \left[ C_{k} t^{-(j-k)}\|v_0\|_{W_p^{2k+|\alpha|}(\mathbb{R}^n)} + \|v_0\|_{W_p^{|\alpha|}(\mathbb{R}^n)} \right].
\]
Here and hereafter, we put $K^+ = \max(K,0)$ and
\[ W_p^k(\mathbb{R}^n) = \left\{ u \in L_p(\mathbb{R}^n) \mid \|u\|_{W_p^k(\mathbb{R}^n)} = \sum_{|\alpha| \leq k} \|\partial_\alpha u\|_{L_p(\mathbb{R}^n)} < \infty \right\}. \]

Y. Shibata [6] gave the $L_p - L_q$ type estimates for the solution to the linear viscoelastic equation:

\begin{equation}
\begin{aligned}
&v_{tt} - \Delta v - \Delta v_t = 0 \quad \text{in } [0, \infty) \times \mathbb{R}^n, \\
v(0) = v_0, \quad v_t(0) = v_1 \quad \text{in } \mathbb{R}^n.
\end{aligned}
\end{equation}

The solution of (2.12) are represented by the Fourier transform as follows: When $|\xi| \neq 0, 2$

\[ \hat{v}(t, \xi) = \frac{\lambda_+ (\xi) e^{\lambda_- (\xi) t} - \lambda_- (\xi) e^{\lambda_+ (\xi) t}}{\lambda_+ (\xi) - \lambda_- (\xi)} \hat{v}_0 (\xi) + \frac{e^{\lambda_+ (\xi) t} - e^{\lambda_- (\xi) t}}{\lambda_+ (\xi) - \lambda_- (\xi)} \hat{v}_1 (\xi), \]

where

\[ \lambda_\pm (\xi) = \frac{-|\xi|^2 \pm \sqrt{|\xi|^4 - |\xi|^2}}{2} \]

and when $1 < |\xi| < 4$,

\[ \hat{v}(t, \xi) = \frac{1}{2\pi i} \int_{\gamma} \frac{(z + |\xi|^2)^e^{zt}}{z^2 + |\xi|^2 z + |\xi|^2} d\hat{v}_0 (\xi) + \frac{1}{2\pi i} \int_{\gamma} \frac{e^{zt}}{z^2 + |\xi|^2 z + |\xi|^2} d\hat{v}_1 (\xi), \]

where $\gamma$ is a closed path containing $\lambda_\pm (\xi)$ and contained in $\{z \in \mathbb{C} \mid \text{Re} z \leq -c_0 \}$ and $c_0$ is a positive number such that

\[ \max_{1 \leq |\xi| \leq 4} \text{Re} \lambda_\pm (\xi) \leq -2c_0. \]

The difference of the structure to the solutions between (1.1) and (2.12) is the Riesz kernel $R_j (x) = F^{-1} (\xi_j / |\xi|)(x)$. The Green matrix of the solution of (1.1) includes the Riesz kernel (cf. (2.9)). Since the convolution operator $u \rightarrow R_j * u$ is bounded from $L_p$ to $L_p$ for $1 < p < \infty$, the following theorems directly follow from [6, Theorems 2.1 and 2.2].

**Theorem 2.3** ($L_p - L_q$ estimate of $E_0(t)$).

1. Let $M$ be the positive number $\geq 1$ and let $1 \leq p \leq q \leq \infty$, $(p, q) \neq (\infty, \infty), (1, 1)$. Then, for any $t \in [0, M]$, we have

\[ \|\partial_t^j \partial_x^\alpha E_0(t)(\rho_0, v_0)\|_{L_q(\mathbb{R}^n)} \leq C_{n, p, q, j, \alpha, M} \left[ \|\rho_0\|_{L_p(\mathbb{R}^n)} + \|v_0\|_{L_p(\mathbb{R}^n)} \right]. \]

2. Let $1 \leq p \leq 2 \leq q \leq \infty$. For any $t > 0$, we have

\[ \|\partial_t^j \partial_x^\alpha E_0(t)(\rho_0, v_0)\|_{L_q(\mathbb{R}^n)} \leq C_{n, p, q, j, \alpha} \left( 1 + t \right)^{\left( \frac{\gamma}{2} \left( 1 - \frac{\gamma}{p} \right) + \frac{\gamma |\alpha|}{2} \right)} \left[ \|\rho_0\|_{L_p(\mathbb{R}^n)} + \|v_0\|_{L_p(\mathbb{R}^n)} \right]. \]
Remark.
(1) The estimate (1) in Theorem 2.1 is better than the estimate (2) in Theorem 2.3 with \((p, q) = (1, \infty)\).
(2) By Theorem 2.1 and Theorem 2.2, the estimate (1) in Theorem 2.3 also holds when \((p, q) = (1, 1)\) or \((\infty, \infty)\).

**Theorem 2.4** \((L_p - L_p\) estimate of \(E_\infty(t))\). Let \(1 < p < \infty\). For any \(t > 0\), we have

\[
\|\partial_t^2 \partial_x^2 E_\infty(t)(\rho_0, v_0)\|_{L_p(\mathbb{R}^n)} \leq C_{j, a, p, N} e^{-ct} \left[ t^{-\frac{N}{2}} \|\rho_0\|_{W_p^{(2j + |a| - N - 2)^+}(\mathbb{R}^n)} + \|\rho_0\|_{W_p^{(a)_+}(\mathbb{R}^n)} \right] + C_{j, a, p, N} e^{-ct} \left[ t^{-\frac{N}{2}} \|v_0\|_{W_p^{(2j + |a| - N - 1)^+}(\mathbb{R}^n)} + \|v_0\|_{W_p^{(a)_+}(\mathbb{R}^n)} \right];
\]

\[
\|\partial_t^2 \partial_x^2 E_\infty(t)(\rho_0, v_0)\|_{L_p(\mathbb{R}^n)} \leq C_{j, a, n, N} e^{-ct} \left[ t^{-\frac{N}{2}} \|\rho_0\|_{W_p^{(2j + |a| - N - 1)^+}(\mathbb{R}^n)} + \|\rho_0\|_{W_p^{(a)_+}(\mathbb{R}^n)} \right] + C_{j, a, n, N} e^{-ct} \left[ t^{-\frac{N}{2}} \|v_0\|_{W_p^{(2j + |a| - N)^+}(\mathbb{R}^n)} + \|v_0\|_{W_p^{(a)_+}(\mathbb{R}^n)} \right].
\]

3. Proof of Theorem 2.1 (1).

To prove Theorem 2.1 (1), we put

\[
L_{11}(t, x) = \mathcal{F}^{-1} \left[ \frac{\lambda_+(\xi) e^{\lambda_-(\xi)t} - \lambda_-(\xi) e^{\lambda_+(\xi)t}}{\lambda_+ - \lambda_-} \varphi_0(\xi) \right](x),
\]

\[
L_{12}(t, x) = -i\gamma \mathcal{F}^{-1} \left[ t^\xi \frac{e^{\lambda_+(\xi)t} - e^{\lambda_-\xi t}}{\lambda_+ - \lambda_-} \varphi_0(\xi) \right](x),
\]

\[
L_{21}(t, x) = t L_{12}(t, x),
\]

\[
L_{22}(t, x) = K_1(t, x) + K_2(t, x) - K_3(t, x),
\]

\[
K_1(t, x) = \mathcal{F}^{-1} \left[ e^{-\alpha|\xi|^2 t} \varphi_0(\xi) \right](x), \quad I \text{ is unit matrix},
\]

\[
K_2(t, x) = \mathcal{F}^{-1} \left[ \frac{\lambda_+(\xi) e^{\lambda_+(\xi)t} - \lambda_-(\xi) e^{\lambda_-(\xi)t}}{\lambda_+ - \lambda_-} \frac{\xi_k}{|\xi|^2} \varphi_0(\xi) \right](x),
\]

\[
K_3(t, x) = \mathcal{F}^{-1} \left[ e^{-\alpha|\xi|^2 t} \frac{\xi_k}{|\xi|^2} \varphi_0(\xi) \right](x),
\]

and then, from (2.5) and (2.9) it follows that

\[
E_0(t)(\rho_0, v_0) = \left( \begin{array}{cc} L_{11}(t, \cdot) & L_{12}(t, \cdot) \\ L_{21}(t, \cdot) & L_{22}(t, \cdot) \end{array} \right) \ast \left( \begin{array}{c} \rho_0 \\ v_0 \end{array} \right),
\]
where \( * \) denotes the spatial convolution. In view of the Young inequality, in order to get Theorem 2.1 (1) it suffices to show that for \( t > 0 \)

\[
\| \partial_t^j \partial_x^\alpha L_{11}(t, \cdot) \|_{L_\infty(\mathbb{R}^n)} \leq C_{j,\alpha,n} t^{-\left(\frac{3n-1}{4} + \frac{1}{2}\|\alpha\|_\infty\right)},
\]

\[
\| \partial_t^j \partial_x^\alpha L_{12}(t, \cdot) \|_{L_\infty(\mathbb{R}^n)} \leq C_{j,\alpha,n} t^{-\left(\frac{3n-1}{4} + \frac{1}{2}\|\alpha\|_\infty\right)},
\]

\[
\| \partial_t^j \partial_x^\alpha K_1(t, \cdot) \|_{L_\infty(\mathbb{R}^n)} \leq C_{j,\alpha,n} t^{-\left(\frac{2}{3} + \frac{1}{2}\|\alpha\|_\infty\right)},
\]

\[
\| \partial_t^j \partial_x^\alpha K_2(t, \cdot) \|_{L_\infty(\mathbb{R}^n)} \leq C_{j,\alpha,n} t^{-\left(\frac{3n-1}{4} + \frac{1}{2}\|\alpha\|_\infty\right)},
\]

\[
\| \partial_t^j \partial_x^\alpha K_3(t, \cdot) \|_{L_\infty(\mathbb{R}^n)} \leq C_{j,\alpha,n} t^{-\left(\frac{2}{3} + \frac{1}{2}\|\alpha\|_\infty\right)}.
\]

It is obvious that

\[
\left| \partial_t^j \partial_x^\beta F^{-1} \left[ e^{-\alpha|\xi|^2 t} \left( \delta_{ik} - \frac{\xi_i \xi_k}{|\xi|^2} \right) \varphi_0(\xi) \right] (x) \right| \leq C_{j,\beta,n} \int_{\mathbb{R}^n} e^{-\alpha|\xi|^2 t} |\xi|^{2j+|\beta|} \, d\xi
\]

\[
\leq C_{j,\beta,n} t^{-\left(\frac{2}{3} + \frac{|\beta|}{2}\right)},
\]

which show (3.5) and (3.7). In view of (3.1), we put

\[
K_{\psi,0}(t, x) = F^{-1} \left[ \frac{e^{\lambda_+(\xi)t} - e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \psi(\xi) \varphi_0(\xi) \right](x),
\]

\[
K_{\psi,1}(t, x) = F^{-1} \left[ \frac{\lambda_+(\xi)e^{\lambda_+(\xi)t} - \lambda_-(\xi)e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \psi(\xi) \varphi_0(\xi) \right](x),
\]

\[
K_{\psi,2}(t, x) = F^{-1} \left[ \frac{\lambda_-(\xi)e^{\lambda_+(\xi)t} - \lambda_+(\xi)e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \psi(\xi) \varphi_0(\xi) \right](x),
\]

where \( \psi = \psi(\omega) \in C_\infty(S^{n-1}), S^{n-1} = \{ \xi \in \mathbb{R}^n \mid |\xi| = 1 \} \) and \( \psi(\xi) = \psi(\xi/|\xi|) \). By (2.3) and (2.4), we know that

\[
\lambda_+(\xi)\lambda_-(\xi) = A^2B^2|\xi|^2,
\]

\[
\lambda_+(\xi) + \lambda_-(\xi) = -2A|\xi|^2,
\]

\[
\lambda_+(\xi)^2 + 2A\lambda_+(\xi)|\xi|^2 + \gamma|\xi|^2 = 0,
\]

and then

\[
\begin{aligned}
K_{\psi,1}(t, x) &= \partial_t K_{\psi,0}(t, x), \\
K_{\psi,2}(t, x) &= -\partial_t K_{\psi,0}(t, x) + 2A\Delta K_{\psi,0}(t, x).
\end{aligned}
\]

Therefore, in order to show (3.3), (3.4) and (3.6) it suffices to show the following theorem:

**Theorem 3.1.** Let \( n \geq 2 \). For any \( t \geq 0 \), we have

\[
\| \partial_t^j \partial_x^\alpha K_{\psi,0}(t, \cdot) \|_{L_\infty(\mathbb{R}^n)} \leq C_{j,\alpha,n} (1 + t)^{-\left(\frac{3n-1}{4} + \frac{1}{2}\|\alpha\|_\infty\right)}.
\]
To prove this theorem, first of all, we shall estimate $K_{\psi,0}(t, x)$ near the light cone. Namely, we shall show that for $t \geq \max(1, (R/R_0)^4)$ and $|x| \geq R_0 t$

\begin{equation}
|\partial_t^j \partial_x^\alpha K_{\psi,0}(t, x)| \leq C_{j, \alpha, n}(1 + t)^{-\left(\frac{3n-3}{4} + \frac{j + |\alpha|}{2}\right)},
\end{equation}

where $R$ is the number appearing in Lemma 3.2, below and $R_0$ is the fixed number such that $R_0 \leq \gamma/4$. To obtain (3.13), we shall use the following lemma concerning the stationary phase method (cf. Vainberg [9, pp. 29-35]):

**Lemma 3.2.** Let $g(\omega) \in C^\infty(S^{n-1})$, $S^{n-1} = \{\xi \in \mathbb{R}^n \mid |\xi| = 1\}$. Then, there exist a $R > 1$ and a $C_g$ such that

\[
\left| \int_{S^{n-1}} e^{ir(\hat{x}, \omega)} g(\omega) dS_\omega \right| \leq C_g r^{-\frac{n-1}{2}}, \quad \hat{x} \in S^{n-1}, \quad r \geq R.
\]

If we put $|\xi| = r$, we have

\[
\lambda_{\pm}(r) = -A \left( r^2 \pm i r \sqrt{B^2 - r^2} \right) = \lambda_{\pm}(r).
\]

Since we may assume that $\varphi_0(\xi) = \varphi_0(|\xi|) = \varphi_0(r)$, by using the polar coordinate we have

\[
\partial_t^j \partial_x^\alpha K_{\psi,0}(t, x) = \left(\frac{1}{2\pi}\right)^n \int_0^\infty \frac{\lambda_+(r)e^{\lambda_+(r)t} - \lambda_-(r)e^{\lambda_-(r)t}}{\lambda_+(r) - \lambda_-(r)} r^{\alpha+n-1} \varphi_0(r) dr
\]

\[
\cdot \int_{S^{n-1}} e^{i(\hat{x}, \omega)r|x|} (i \omega)^\alpha \psi(\omega) dS_\omega,
\]

where $\hat{x} = x/|x|$. Let $\epsilon > 0$ be a number determined later on. Let us consider the case where $|x| \geq R$, below. Since $r|x| \geq \epsilon |x| \geq R$ when $r \geq \epsilon$, by Lemma 3.2

\[
\left| \int_{S^{n-1}} e^{i(\hat{x}, \omega)r|x|} (i \omega)^\alpha \psi(\omega) dS_\omega \right| \leq C_\alpha |r|x|^{-\frac{n-1}{2}}.
\]

Noting that $\varphi_0(r) = 0$ when $r \geq B/\sqrt{2}$ (cf. (2.10)), we have

\[
\left| \partial_t^j \partial_x^\alpha K_{\psi,0}(t, x) \right| \leq C \left\{ \int_0^\epsilon r^{n+j+|\alpha|-2} dr + \int_\epsilon^\infty e^{-A r^2} r^{-n-j+|\alpha|} (r|x|)^{-\frac{n-1}{2}} dr \right\}.
\]

If we make the change of variable; $r\sqrt{t} = s$ in the last integration and if we use the assumption: $|x| \geq R_0 t$, then we have

\[
\left| \partial_t^j \partial_x^\alpha K_{\psi,0}(t, x) \right| \leq C_{j, \alpha, n} \left\{ \epsilon^{n+j+|\alpha|-1} + t^{-\left(\frac{3n-3}{4} + \frac{j + |\alpha|}{2}\right)} \right\}.
\]

Choose $\epsilon > 0$ in such a way that

\[
\epsilon^{n+j+|\alpha|-1} = t^{-\left(\frac{3n-3}{4} + \frac{j + |\alpha|}{2}\right)}.
\]
When $|x| \geq R_0 t$ and $t \geq \max(1, (R/R_0)^4)$, we see that
\[
|x|e \geq R_0 t \cdot t^{-\left(\frac{3n-3+|\alpha|}{2} + \frac{1+|\alpha|}{2} (n+j+|\alpha| - 1)\right)} = R_0 \frac{t^{\frac{1}{2}}}{n+j+|\alpha| - 1} \geq R_0 t^{\frac{1}{2}} \geq R.
\]
Therefore, we have (3.13).

Now, we shall show that for $t \geq 1$ and $|x| \leq R_0 t$
\[
(3.14) \quad \left| \partial_t^j \partial_x^\alpha K_{\psi,0}(t,x) \right| \leq C_{j,\alpha,n} t^{-\left(\frac{3n-3+|\alpha|}{4} + \frac{j+|\alpha|}{4}\right)}.
\]

If we put
\[
f(\xi) = \sqrt{1 - |\xi|^2 B^{-2}} = 1 + |\xi|^2 g(|\xi|^2), \quad g(s) = -\frac{1}{2B^2} \int_0^1 \frac{1}{\sqrt{1 - \theta s B^{-2}}} d\theta,
\]
then by Taylor’s formula we have
\[
(3.15) \quad \frac{e^{\lambda_+\xi t} - e^{-\lambda_-\xi t}}{\lambda_+\xi - \lambda_-\xi} = \sum_{\ell=0}^N \frac{1}{\ell!} \left( \frac{\partial^\ell}{\partial^\ell \xi} \frac{\sin \gamma \xi}{|\xi|} \right) \frac{e^{-A|\xi|^2 t}}{\gamma f(|\xi|)} \left( |\xi|^2 g(|\xi|^2) t \right)^\ell + e^{-A|\xi|^2 t} R_N(t,|\xi|),
\]
where
\[
R_N(t,|\xi|) = \frac{1}{2i\gamma |\xi| f(|\xi|) N!} \int_0^1 (1 - \theta)^N \left[ e^{i\gamma|\xi|^2 t + i\gamma |\xi|^3 g(|\xi|^2) t^\theta} \left( i\gamma |\xi|^3 g(|\xi|^2) t \right)^N + e^{-i\gamma|\xi|^2 t - i\gamma |\xi|^3 g(|\xi|^2) t^\theta} \left( -i\gamma |\xi|^3 g(|\xi|^2) t \right)^N \right] d\theta.
\]
In fact,
\[
e^{\lambda_{\pm}\xi t} = e^{-A|\xi|^2 t} e^{\mp i\gamma |\xi|^3 g(|\xi|^2) t} = e^{-A|\xi|^2 t} e^{\mp i\gamma |\xi|^3 g(|\xi|^2) t}.
\]
Put
\[
h(\theta) = e^{\pm i\gamma |\xi|^3 g(|\xi|^2) t^\theta}, \quad h^{(k)}(\theta) = \frac{d^k h}{d\theta^k}(\theta).
\]
Since
\[
h(1) = h(0) + h'(0) + \cdots + \frac{1}{N!} h^{(N)}(0) + \frac{1}{N!} \int_0^1 (1 - \theta)^N h^{(N+1)}(\theta) d\theta,
\]
we have
\[
e^{\pm i\gamma |\xi|^3 g(|\xi|^2) t} = 1 + (\pm i\gamma |\xi|^3 g(|\xi|^2) t) + \cdots + \frac{1}{N!} \left( \pm i\gamma |\xi|^3 g(|\xi|^2) t \right)^N + \frac{1}{N!} \int_0^1 (1 - \theta)^N e^{\pm i\gamma |\xi|^3 g(|\xi|^2) t^\theta} \left( \pm i\gamma |\xi|^3 g(|\xi|^2) t \right)^{N+1}.
\]
Since
\[
(i\gamma|\xi|^3g(|\xi|^2)t)^N e^{i\gamma|\xi|t} - (i\gamma|\xi|^3g(|\xi|^2)t)^N e^{-i\gamma|\xi|t}
= \left\{ \partial_t^N \left( e^{i\gamma|\xi|t} - e^{-i\gamma|\xi|t} \right) \right\} (|\xi|^2g(|\xi|^2)t)^N
= 2i (\partial_t^N \sin \gamma|\xi|t) (|\xi|^2g(|\xi|^2)t)^N,
\]
noting that \(\lambda_+(\xi) - \lambda_-(\xi) = -2i\gamma|\xi|f(|\xi|)\), we have (3.15).

We shall use the following lemma.

**Lemma 3.3** (cf. Mizohata [5], Evans [1]). Put
\[
w(t,x) = F^{-1} \left[ \sin \frac{|\xi| t}{|\xi|} \hat{h}(\xi) \right](x).
\]
Then, for suitable constants \(a_\alpha\) we have
\[
w(t,x) = \sum_{0 \leq |\alpha| \leq \frac{n-3}{2}} a_\alpha t^{|\alpha|+1} \int_{|z|=1} z^\alpha (\partial^\alpha_x h) (x + tz) dS
\]
for odd \(n \geq 3\); and
\[
w(t,x) = \sum_{0 \leq |\alpha| \leq \frac{n-2}{2}} a_\alpha t^{|\alpha|+1} \int_{|z|=1} \frac{z^\alpha (\partial^\alpha_x h) (x + tz)}{\sqrt{1-|z|^2}} dz
\]
for even \(n \geq 2\).

Regarding (3.15), we put
\[
G_\ell(t,x) = F^{-1} \left[ e^{-A|\xi|^2t} \left( |\xi|^2g(|\xi|^2)t \right)^\ell \psi(\xi) \varphi_0(\xi) \right](x),
\]
\[
\omega_\ell(t,x) = \frac{1}{\gamma} F^{-1} \left[ \left( \partial_t^\ell \sin \gamma|\xi|t \right) \frac{|\xi|}{|\xi|} G_\ell(t,\xi) \right](x).
\]
Since
\[
F^{-1} \left[ \partial_t^\ell \left( \frac{\sin \gamma|\xi|t}{|\xi|} \right) \hat{h}(\xi) \right](x) = \partial_t^\ell F^{-1} \left[ \sin \gamma|\xi|t \frac{|\xi|}{|\xi|} \hat{h}(\xi) \right](x),
\]
by Lemma 3.3 we have
\[
\partial_t^\ell \partial_x^\beta K_{\psi,0}(t,x) = \sum_{\ell=0}^N \partial_t^\ell \partial_x^\beta \omega_\ell(t,x)
+ \partial_t^\ell \partial_x^\beta F^{-1} \left[ e^{-A|\xi|^2t} R_N(t,|\xi|) \right](x),
\]
where
\[(3.17) \quad \partial^j_t \partial^\ell_x \omega (t, x) = \begin{cases} \frac{1}{\gamma} \sum_{k=0}^j \sum_{|\alpha| \leq \frac{n-1}{2}} a_{\alpha} \sum_{m=0}^{\min(\ell+k,|\alpha|+1)} \left( \begin{array}{c} \ell + k \\ m \end{array} \right) \partial^m_t (\gamma t)^{|\alpha|+1} \\
\quad \cdot \sum_{|\delta| = \ell + k - m} \int_{|z| = 1} z^\alpha \delta \left( \partial_x^{\alpha + \beta + \delta} \partial_t^{j-k} G_\ell \right)(t, x + \gamma tz) \, dS \\
\text{for odd } n \geq 3; \\
\frac{1}{\gamma} \sum_{k=0}^j \sum_{|\alpha| \leq \frac{n-2}{2}} a_{\alpha} \sum_{m=0}^{\min(\ell+k,|\alpha|+1)} \left( \begin{array}{c} \ell + k \\ m \end{array} \right) \partial^m_t (\gamma t)^{|\alpha|+1} \\
\quad \cdot \sum_{|\delta| = \ell + k - m} \int_{|z| \leq 1} \frac{z^\alpha \delta \left( \partial_x^{\alpha + \beta + \delta} \partial_t^{j-k} G_\ell \right)(t, x + \gamma tz)}{\sqrt{1 - |z|^2}} \, dz \\
\text{for even } n \geq 2. \end{cases}\]

The following proposition and lemma play an essential role to prove Theorem 3.1.

**Proposition 3.4** (Shibata-Shimizu [7]). Let \( \alpha \) be a number \( > -n \) and put \( \alpha = N + \sigma - n \) where \( N \geq 0 \) is an integer and \( 0 < \sigma \leq 1 \). Let \( f(\xi) \) be a function in \( C^\infty (\mathbb{R}^n - \{0\}) \) such that
\[ \partial_\gamma^\ell f(\xi) \in L_1(\mathbb{R}^n), \quad |\gamma| \leq N; \]
\[ \left| \partial_\gamma^\ell f(\xi) \right| \leq C_\gamma |\xi|^{\alpha - |\gamma|}, \quad \xi \neq 0, \quad \forall \gamma. \]
Then, we have
\[ |F^{-1}[f(\xi)](x)| \leq C_{a, n} \left( \max_{|\gamma| \leq N+2} C_\gamma \right) |x|^{-(n+|\alpha|)}, \quad x \neq 0, \]
where \( C_{a, n} \) is a constant depending essentially only on \( n \) and \( \alpha \).

**Lemma 3.5.** Let \( \alpha \) be a nonnegative number and \( \psi(t, \xi) \) be a function such that
\[ \psi(t, \cdot) \in C^\infty (\mathbb{R}^n - \{0\}), \quad \forall t \geq 0; \]
\[ \left| \partial_\gamma^\ell \psi(t, \xi) \right| \leq C_\gamma |\xi|^{\alpha - |\gamma|}, \quad \xi \neq 0, \quad \forall \gamma, \quad \forall t \geq 0. \]
Put
\[ g(t, x) = F^{-1} \left[ e^{-|\xi|^2 t} \psi(t, \xi) \right](x), \]
where \( \beta > 0 \). Then, we have
\[(3.18) \quad |g(t, x)| \leq C_{a, \beta, n} |x|^{-(\alpha+n)}, \quad x \neq 0, \]
\[ |g(t, x)| \leq C_{a, \beta, n} t^{-\frac{\alpha+n}{2}}, \quad t > 0. \]
Moreover,
\[
\|g(t, \cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{\alpha, \beta, n} t^{-\frac{\alpha}{2}}, \quad \alpha > 0, \ t > 0,
\]
\[
\int_{|x| \leq At} |g(t, x)| \, dx \leq C_{\beta, n, A} (1 + \log(1 + t)), \quad \alpha = 0, \ t > 0.
\]

Proof. By the formula of derivative of composed function (cf. Simader [8, p. 202]):
\[
(3.19) \quad \partial^\gamma_{\xi} h(g(\xi)) = \sum_{\nu=1}^{\gamma} \nu! (\partial^{\alpha_1}_\xi g(\xi)) \cdots (\partial^{\alpha_\nu}_\xi g(\xi)) \left[ \sum_{\alpha_1 + \cdots + \alpha_\nu = \gamma} \frac{\partial^{\alpha_1}_\xi |\xi|^2 \cdots \partial^{\alpha_\nu}_\xi |\xi|^2}{|\alpha_1| \geq 1} \right],
\]
we have
\[
\partial^\gamma_{\xi} e^{-\beta|\xi|^2 t} = \sum_{\nu=1}^{\gamma} (\beta t)^\nu e^{-\beta|\xi|^2 t} \left[ \sum_{\alpha_1 + \cdots + \alpha_\nu = \gamma} \frac{\partial^{\alpha_1}_\xi |\xi|^2 \cdots \partial^{\alpha_\nu}_\xi |\xi|^2}{|\alpha_1| \geq 1} \right].
\]
Since
\[
(3.20) \quad \left| \partial^{\alpha_1}_\xi |\xi|^M \right| \leq C_{\alpha_1} |\xi|^{M-|\alpha_1|}, \quad \xi \neq 0,
\]
\[
(t|\xi|^2)^M e^{-\beta|\xi|^2 t} \leq C_{\alpha, \beta} e^{-\frac{\beta}{2}|\xi|^2 t},
\]
we have
\[
(3.21) \quad \left| \partial^\gamma_{\xi} e^{-\beta|\xi|^2 t} \right| \leq C_{\gamma} \sum_{\nu=1}^{\gamma} (\beta t)^\nu e^{-\beta|\xi|^2 t} \left[ \sum_{\alpha_1 + \cdots + \alpha_\nu = \gamma} \frac{|\xi|^{2\nu-(|\alpha_1|+\cdots+|\alpha_\nu|)}}{|\alpha_1| \geq 1} \right]
\]
\[
\leq C_{\gamma} \sum_{\nu=1}^{\gamma} (\beta |\xi|^2 t)^\nu e^{-\beta|\xi|^2 t} |\xi|^{-|\gamma|}
\]
\[
\leq C_{\beta, \gamma} |\xi|^{-|\gamma|} e^{-\frac{\beta}{2}|\xi|^2 t}, \quad \xi \neq 0,
\]
and the Leibniz’s rule we have
\[
(3.22) \quad \left| \partial^\gamma_{\xi} \left( e^{-\beta|\xi|^2 t} \psi(t, \xi) \right) \right| \leq C_{\alpha, \beta, \gamma} e^{-\frac{\beta}{2}|\xi|^2 t} |\xi|^{\alpha-|\gamma|}, \quad \xi \neq 0, \ \forall \gamma.
\]
Therefore, by Proposition 3.4 we have (3.18).

By the assumptions, we have
\[
(3.23) \quad |g(t, x)| \leq C \int_{\mathbb{R}^n} e^{-\beta|\xi|^2 t} |\xi|^\alpha \, d\xi = C t^{-\frac{\alpha+n}{2}} \int_{\mathbb{R}^n} e^{-\beta|\eta|^2} |\eta|^\alpha \, d\eta
\]
\[
\leq C_{\alpha, \beta, n} t^{-\frac{\alpha+n}{2}}.
\]
By (3.18) and (3.23), we have
\[ \|g(t, \cdot)\|_{L^1(R^n)} \leq C_{\alpha, \beta, n} t^{-\frac{\alpha + n}{2}} \int_{|x| \leq \sqrt{t}} |x|^{-(\alpha + n)} \, dx \]
and also
\[ \int_{|x| \leq A\sqrt{t}} |g(t, x)| \, dx \leq C_{\beta, n} t^{-\frac{n}{2}} \int_{|x| \leq \sqrt{t}} |x|^{-n} \, dx \leq C_{\beta, n} (1 + \log t), \quad \alpha = 0, \quad t \geq 1 \]
which completes the Proof of Lemma 3.5.

Concerning the estimate \( G_\ell(t, x) \), we have
\[
\left| \partial^j_{t} (\partial^{\alpha + \beta + \delta}_{x} G_\ell(t, x + \gamma tz)) \right| \leq C_{j, k, \alpha, \beta, \delta, \ell} |x + \gamma tz|^{-(2(j - k) + |\alpha| + |\beta| + |\delta| + n)}.
\]
In fact,
\[
\partial^j_{t} (\partial^{\alpha + \beta + \delta}_{x} G_\ell(t, x)) = \mathcal{F}^{-1} \left[ \sum_{m=0}^{\min(j - k, \ell)} \binom{j - k}{m} e^{-\frac{A|\xi|^2}{t}} \frac{f(|\xi|)}{f(|\xi|)} (-A|\xi|^2)^{j-k-m} (i\xi)^{\alpha + \beta + \delta} \partial^m_{\xi} \left( |\xi|^2 g(|\xi|^2) t^\ell \psi(\xi) \varphi_0(\xi) \right) \right](x).
\]
By (3.20), (3.21) and (3.22) we have
\[
\left| \partial^j_{\xi} \left( \sum_{m=0}^{\min(j - k, \ell)} \binom{j - k}{m} e^{-\frac{A|\xi|^2}{t}} \frac{f(|\xi|)}{f(|\xi|)} (-A|\xi|^2)^{j-k-m} (i\xi)^{\alpha + \beta + \delta} \right) \partial^m_{\xi} \left( |\xi|^2 g(|\xi|^2) t^\ell \psi(\xi) \varphi_0(\xi) \right) \right| \leq C_{j, k, \alpha, \beta, \delta, \ell, \mu} |\xi|^{2(j - k) + |\alpha| + |\beta| + |\delta| - |\mu|}.
\]
Therefore, by Lemma 3.5, we have (3.24).

First we consider the case when \( n \) is an odd \( \geq 3 \). When \( |z| = 1, |x| \leq R_0 t \), and \( R_0 \leq \gamma/4 \), we have
\[
|x + \gamma tz| \geq \gamma t - |x| \geq (\gamma - R_0)t \geq \frac{\gamma}{2} t.
\]
Therefore, applying (3.24) to (3.17), we have for $|x| \leq R_0$ and $t \geq 1$

(3.25)

\[
\left| \frac{\partial^j_t}{\partial x^\beta} \omega(t, x) \right| \leq \frac{1}{\gamma} \sum_{k=0}^{j} \sum_{|\alpha| \leq \frac{n-2}{2}} |a_\alpha| \sum_{m=0}^{\min(\ell+k,|\alpha|+1)} \left( \frac{\ell + k}{m} \right) \frac{(|\alpha|+1)!}{m!} (\gamma t)^{\alpha+1} t^{-m} \\
\cdot \sum_{|\delta|=\ell+k-m} C_{j,k,\alpha,\beta,\delta,\ell} \int_{|z|=1} |x + \gamma tz|^{-(2(j-k)+|\alpha|+|\beta|+|\delta|+n)} dS
\]

\[
\leq \frac{1}{\gamma} \sum_{k=0}^{j} \sum_{|\alpha| \leq \frac{n-2}{2}} |a_\alpha| \sum_{m=0}^{\min(\ell+k,|\alpha|+1)} \left( \frac{\ell + k}{m} \right) \frac{(|\alpha|+1)!}{m!} (\gamma t)^{\alpha+1} t^{-m} \\
\cdot \sum_{|\delta|=\ell+k-m} C_{j,k,\alpha,\beta,\delta,\ell} t^{-(2(j-k)+|\alpha|+|\beta|+|\delta|+n)} \int_{|z|=1} dS
\]

\[
\leq C_{j,\beta,\ell,n} t^{-(j+|\beta|+\ell+n-1)}.
\]

Next, we consider the case when $n$ is even $\geq 2$. By (3.17) we have

\[
\left| \frac{\partial^j_t}{\partial x^\beta} \omega(t, x) \right| \leq \frac{1}{\gamma} \sum_{k=0}^{j} \sum_{|\alpha| \leq \frac{n-2}{2}} |a_\alpha| \sum_{m=0}^{\min(\ell+k,|\alpha|+1)} \left( \frac{\ell + k}{m} \right) \frac{(|\alpha|+1)!}{m!} (\gamma t)^{\alpha+1} t^{-m} \\
\cdot \sum_{|\delta|=\ell+k-m} C_{j,k,\alpha,\beta,\delta,\ell} \int_{|z| \leq 1} \left| \frac{\partial^j_t}{\partial x^{\alpha+\beta+\delta}} G_{\ell_t}(t, x + \gamma tz) \right| \sqrt{1-|z|^2} dS.
\]

Put

\[
\int_{|z| \leq 1} \left| \frac{\partial^j_t}{\partial x^{\alpha+\beta+\delta}} G_{\ell_t}(t, x + \gamma tz) \right| \sqrt{1-|z|^2} dS = I + II,
\]

where

\[
I = \int_{\frac{1}{2} \leq |z| \leq 1} \left| \frac{\partial^j_t}{\partial x^{\alpha+\beta+\delta}} G_{\ell_t}(t, x + \gamma tz) \right| \sqrt{1-|z|^2} dS,
\]

\[
II = \int_{|z| \leq \frac{1}{2}} \left| \frac{\partial^j_t}{\partial x^{\alpha+\beta+\delta}} G_{\ell_t}(t, x + \gamma tz) \right| \sqrt{1-|z|^2} dS.
\]
When $1/2 \leq |z| \leq 1$, $|x| \leq R_0 t$, $t \geq 1$ and $R_0 \leq \gamma/4$, we have

$$|x + \gamma tz| \geq \frac{\gamma t}{2} - |x| \geq \left(\frac{\gamma}{2} - R_0\right) t \geq \frac{\gamma}{4} t.$$  

Then, by (3.24) we have

$$p \leq C_{j,k,a,\beta,\delta,\ell} t^{-\frac{1}{2}(2(j-k)+|\alpha|+|\beta|+|\delta|+n)} \int_{\frac{1}{2}\leq |z| \leq 1} \frac{dz}{\sqrt{1-|z|^2}}.$$

When $|z| \leq 1/2$, $|x| \leq R_0 t$, $t \geq 1$ and $R_0 \leq \gamma/4$, we have

$$|x + \gamma tz| \leq \frac{\gamma t}{2} + R_0 t \leq \frac{3}{4} \gamma t, \quad \sqrt{1-|z|^2} \geq \frac{\sqrt{3}}{2}.$$

Therefore, putting $p = x + \gamma tz$, by Lemma 3.5 we have

$$II \leq \frac{2}{\sqrt{3}} \int_{|p| \leq \frac{1}{4} \gamma t} \left|\left(\partial_t^{j-k} \partial_x^{n+\beta+\delta} G_\xi\right)(t, p)\right| dp$$

$$\leq C_{j,k,a,\beta,\delta,\ell,n} \begin{cases} t^{-n} t^{-\frac{1}{2}(2(j-k)+|\alpha|+|\beta|+|\delta|)} & \text{when } 2(j-k) + |\alpha| + |\beta| + |\delta| \geq 1, \\ t^{-n} (1 + \log t) & \text{when } 2(j-k) + |\alpha| + |\beta| + |\delta| = 0 \\ t^{-n} & \text{and } \ell = 0, \\ t^{-n} & \text{and } \ell \geq 1. \end{cases}$$

Combining these estimations, we have for $|x| \leq R_0 t$ and $t \geq 1$

(3.26)
\[ + \sum_{m=0}^{\min(j+\ell,1)} \left( \frac{j+\ell}{m} \right) \frac{|a_\alpha|}{m!} t^{1-m} \cdot C_{j,\beta,\ell,n} \left\{ t^{-(\ell+j-m+|\beta|+n)} + t^{-n-\frac{1}{2}(\ell+j-m+|\beta|)}(1 + \log(1 + t)) \right\} \leq C_{j,\beta,\ell,n} \left\{ (1 + \log(1 + t))t^{-\left(\frac{3n-3+|\beta|}{2}\right)} - \frac{n-\ell}{2} + t^{-\left(\frac{3n-3+|\beta|}{2}\right)} - \frac{n-\ell}{2} \right\}. \]

Next, we shall estimate the remainder term. By (3.15) and (3.16), we have
\[ \partial_\ell^j \partial_\beta^\ell \mathcal{F}^{-1} \left[ e^{-A|\xi|^2} R_N(t, \xi) \right](x) = \sum_{k=0}^{j} \binom{j}{k} \mathcal{F}^{-1} \left[ e^{-A|\xi|^2} (-A|\xi|^2)^k (i|\xi|)^\beta \partial_\ell^{j-k} R_N(t, \xi) \right](x), \]
and
\[ \left| \partial_\ell^{j-k} R_N(t, \xi) \right| \leq \frac{1}{2\gamma |\xi| f(|\xi|) N!} \int_0^1 (1 - \theta)^N \sum_{\ell_1=0}^{j-k} \binom{j-k}{\ell_1} \partial_\ell^{j-k-\ell_1} e^{i\gamma |\xi|^3 g(|\xi|^2)} t^\theta \cdot \sum_{\ell_2=0}^{\ell_1} \binom{\ell_1}{\ell_2} \partial_\ell^{\ell_1-\ell_2} e^{i\gamma |\xi|^3 g(|\xi|^2)} t^{N+1} \, d\theta \leq C_{j,\beta,N} \sum_{\ell_1=0}^{j-k} \sum_{\ell_2=0}^{\ell_1} |\xi|^{3(N+j-k)+2(1-\ell_1) - \ell_2} t^{N+1-\ell_2}. \]

Combining these estimations, we have
\[ (3.27) \]
\[ \left| \partial_\ell^j \partial_\beta^\ell \mathcal{F}^{-1} \left[ e^{-A|\xi|^2} R_N(t, \xi) \right](x) \right| \leq C_{j,\beta,N} \sum_{\ell_1=0}^{j-k} \sum_{\ell_2=0}^{\ell_1} t^{N+1-\ell_2} \int_{\mathbb{R}^n} |\xi|^{3(N+j)+2(1-\ell_1) - \ell_2 - k+|\beta|} e^{-A|\xi|^2} \, d\xi \leq C_{j,\beta,N} t^{-\frac{N+j+|\beta|}{2}}. \]
In fact, putting (4.2) where (4.1), we have

\[ (3.28) \quad \left| \partial_t^i \partial_x^j K_{\psi,0}(t, x) \right| \leq C_{j,\alpha,n} t^{-\left(\frac{2n-3 + n|\beta|}{4}\right)}, \quad t \geq \max(1, (R/R_0)^4). \]

To complete the Proof of Theorem 3.1, we have to estimate the case when 0 ≤ t ≤ max(1, (R/R_0)^4). But, it is obvious that

\[ \left| \partial_t^i \partial_x^j K_{\psi,0}(t, x) \right| \]

\[ \leq \left( \frac{1}{2\pi} \right)^n \int_{\mathbb{R}^n} |i\xi|^\beta \lambda_\pm(\xi) e^{\lambda_\pm(\xi)t} - \lambda_\pm(\xi) e^{\lambda_\pm(\xi)t} \psi(\xi) \varphi_0(\xi) \]

\[ \leq C_{j,\beta} \int_0^\sqrt{2} r^{j + |\beta| + n - 2} dr \]

\[ \leq C_{j,\beta,n}. \]

Therefore, the Proof of Theorem 3.1 is completed.

4. Proof of Theorem 2.1 (2).

In this section, we shall show Theorem 2.1 (2). In view of (3.1), (3.2) and Young inequality, it suffices to show that

\[ (4.1) \quad \| \partial_t^i \partial_x^j L_{ij}(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j,\alpha,n} (1 + t)^{q(n) - \frac{j + |\beta|}{4}}, \]

where q(n) = (n - 1)/4 for odd n ≥ 3 and = n/4 for even n ≥ 2. Since the kernel of L_{11}(t, x), L_{12}(t, x) = tL_{21}(t, x) are the same as those of (2.12), (4.1) directly follows from the results of [6, Theorem 2.1] when (i, j) = (1, 1), (1, 2) and (2, 1). Therefore, our task is to show (4.1) when (i, j) = (2, 2). In view of (3.1), we put

\[ L_0(t, x) = K_2(t, x) - K_3(t, x) \]

\[ = \mathcal{F}^{-1} \left[ \left( \frac{\lambda_+(\xi) e^{\lambda_+(\xi)t} - \lambda_-(\xi) e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} - e^{-\alpha |\xi|^2 t} \right) \frac{\xi_j \xi_k}{|\xi|^2} \right](x). \]

Then, we have

\[ L_{22}(t, x) = K_1(t, x) + L_0(t, x), \quad K_1(t, x) = \mathcal{F}^{-1} \left[ e^{-\alpha |\xi|^2 t} \varphi_0(\xi) \right](x)I. \]

Noting that \( \varphi_0(\xi) = 0 \) when |\xi| ≥ B/\sqrt{2}, we have

\[ (4.2) \quad \| \partial_t^i \partial_x^j K_1(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j,\beta,n} (1 + t)^{-j - \frac{|\beta|}{4}}. \]

In fact, putting

\[ \chi(x) = \mathcal{F}^{-1} [\varphi_0(\xi)](x) \in \mathcal{S}(\mathbb{R}^n), \]

we have

\[ K_1(t, x) = \frac{1}{(4\pi t)^\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{2ot}} \chi(y) dy. \]
By Young inequality, we see that
\[ \|K_1(t,x)\|_{L^1(\mathbb{R}^n)} \leq C_n, \quad t > 0. \]
When \( j + |\beta| \geq 1 \), we have
\[ \partial_t^j \partial_x^\beta \mathcal{F}^{-1} \left[ e^{-\alpha|\xi|^2t} \varphi_0(\xi) \right] = \mathcal{F}^{-1} \left[ (i\xi)^\beta (-\alpha|\xi|^2)^j e^{-\alpha|\xi|^2t} \varphi_0(\xi) \right], \]
and
\[ \left| \partial_t^j \left( (i\xi)^\beta (-\alpha|\xi|^2)^j \varphi_0(\xi) \right) \right| \leq C_{j,\beta,n} |\xi|^{2j+|\beta|-|\mu|}, \quad \xi \neq 0. \]
Therefore, by Lemma 3.5 we have
\[ \|\partial_t^j \partial_x^\beta K_1(t,x)\|_{L^1(\mathbb{R}^n)} \leq C_{j,\beta,n} t^{-j-\frac{|\beta|}{2}}, \quad t > 0. \]
When \( 0 < t \leq 1 \), since
\[ \partial_t^j \partial_x^\beta \mathcal{F}^{-1} \left[ e^{-\alpha|\xi|^2t} \varphi_0(\xi) \right] (x) = (\alpha \Delta)^j \partial_x^\beta \left\{ \frac{1}{(4\pi \alpha t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4\alpha t}} \chi(y) dy \right\} = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-|\xi|^2} \left( (\alpha \Delta)^j \partial_x^\beta \chi \right) (x - \sqrt{2\alpha t}z) dz, \]
we have
\[ \|\partial_t^j \partial_x^\beta K_1(t,\cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{j,\beta,n}, \quad 0 < t \leq 1. \]
Combining these estimations, we have (4.2).
Now, we shall show that
\[ \|\partial_t^j \partial_x^\beta L_0(t,\cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{j,\beta,n} t^{q(n)-j-\frac{|\beta|}{2}}, \quad t \geq 1. \]
By (3.15), we have
\[ \frac{\lambda_+(\xi)e^{\lambda_+(\xi)t} - \lambda_-(\xi)e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} = \partial_t \left( \frac{e^{\lambda_+(\xi)t} - e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \right) = \partial_t \left\{ \sum_{\ell=0}^{N} \frac{1}{\ell!} \partial_t^\ell \left( \frac{\sin \gamma_{\xi}|t|}{|\xi|} \right) e^{-A|\xi|^2t} \frac{1}{\gamma f(|\xi|)} (|\xi|^2 g(|\xi|^2) t)^\ell + e^{-A|\xi|^2t} R_N(t, |\xi|) \right\} = \sum_{\ell=0}^{N} \frac{1}{\ell!} \partial_t^{\ell+1} \left( \frac{\sin \gamma_{\xi}|t|}{|\xi|} \right) e^{-A|\xi|^2t} \frac{1}{\gamma f(|\xi|)} (|\xi|^2 g(|\xi|^2) t)^\ell + \sum_{\ell=0}^{N} \frac{1}{\ell!} \partial_t^\ell \left( \frac{\sin \gamma_{\xi}|t|}{|\xi|} \right) \partial_t \left( e^{-A|\xi|^2t} \frac{1}{\gamma f(|\xi|)} (|\xi|^2 g(|\xi|^2) t)^\ell \right) + \partial_t \left( e^{-A|\xi|^2t} R_N(t, |\xi|) \right). \]
Combining these two estimations, we have

\[
\frac{\partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) e^{-A|\xi|^2 t}}{\gamma f(|\xi|)} = e^{-A|\xi|^2 t} \left\{ \frac{\partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right)}{\gamma f(|\xi|)} - \partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) \right\}.
\]

Since \( f(|\xi|) = 1 + |\xi|^2 g(|\xi|^2) \), we have

\[
\partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) e^{-A|\xi|^2 t} = e^{-A|\xi|^2 t} \left\{ \partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) - \partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) \right\}.
\]

Combining these two estimations, we have

\[
L_0(t, x) = L_1(t, x) + L_2(t, x) - M_0^1(t, x) + \sum_{\ell=1}^N \frac{1}{\ell !} M_\ell^1(t, x) + R_N(t, x),
\]

where

\[
L_1(t, x) = \mathcal{F}^{-1} \left[ \partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) - 1 \right] e^{-A|\xi|^2 t} \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \big( x \big),
\]

\[
L_2(t, x) = \mathcal{F}^{-1} \left[ e^{-A|\xi|^2 t} - e^{-\alpha |\xi|^2 t} \right] \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \big( x \big),
\]

\[
M_0^1(t, x) = \mathcal{F}^{-1} \left[ \partial_t \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) e^{-A|\xi|^2 t} \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right] \big( x \big),
\]

\[
M_\ell^1(t, x) = \mathcal{F}^{-1} \left[ \partial_t^{\ell+1} \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) e^{-A|\xi|^2 t} \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right] \big( x \big),
\]

\[
M_\ell^2(t, x) = \mathcal{F}^{-1} \left[ \partial_t^\ell \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) e^{-A|\xi|^2 t} \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right] \big( x \big),
\]

\[
M_\ell^3(t, x) = \mathcal{F}^{-1} \left[ \partial_t^\ell \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) e^{-A|\xi|^2 t} \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right] \big( x \big),
\]

and

\[
R_N(t, x) = \mathcal{F}^{-1} \left[ \partial_t \left( e^{-A|\xi|^2 t} R_N(t, |\xi|) \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right) \right] \big( x \big).
\]

First, we shall show that

\[
(4.4) \quad \| L_1(\cdot) \|_{L_1(\mathbb{R}^n)} \leq C_n t^{q(n)}, \quad t \geq 1.
\]

Put

\[
g_0(t, x) = \mathcal{F}^{-1} \left[ e^{-A|\xi|^2 t} \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right] \big( x \big).
\]
Since
\[ L_1(t, x) + g_0(t, x) = \partial_t \mathcal{F}^{-1} \left[ \left( \frac{\sin \gamma |\xi| t}{\gamma |\xi|} \right) \tilde{g}_0(t, \xi) \right](x) - \mathcal{F}^{-1} \left[ \frac{\sin \gamma |\xi| t}{\gamma |\xi|} \tilde{\partial_t} \tilde{g}_0(t, \xi) \right](x), \]

by Lemma 3.3 we have
\[ (4.5) \quad L_1(t, x) + g_0(t, x) = \partial_t \left\{ \frac{1}{\gamma} \sum_{|\alpha| \leq \frac{n-3}{2}} a_\alpha (\gamma t)^{|\alpha|+1} \int_{|z|=1} z^\alpha (\partial_x^\alpha g_0)(t, x + \gamma tz) dS \right\} \]
\[ - \mathcal{F}^{-1} \left[ \frac{\sin \gamma |\xi| t}{\gamma |\xi|^2} \tilde{\partial_t} \tilde{g}_0(t, \xi) \right](x) \]
\[ = \sum_{|\alpha| \leq \frac{n-3}{2}} a_\alpha (|\alpha| + 1)(\gamma t)^{|\alpha|+1} \int_{|z|=1} z^\alpha (\partial_x^\alpha g_0)(t, x + \gamma tz) dS \]
\[ + \sum_{|\alpha| \leq \frac{n-3}{2}} a_\alpha (\gamma t)^{|\alpha|+1} \sum_{|\delta|=1} \int_{|z|=1} z^{\alpha+\delta} (\partial_x^{\alpha+\delta} g_0)(t, x + \gamma tz) dS \]

when \( n \) is an odd \( \geq 3 \). In view of Lemma 3.3, we see that
\[ (4.6) \quad a_0 \int_{|z|=1} dS = 1. \]

In fact, putting
\[ w(t, x) = \mathcal{F}^{-1} \left[ \frac{\sin |\xi| t}{|\xi|} \tilde{h}(\xi) \right](x), \]

by Lemma 3.3, we have
\[ h(x) = w_t(0, x) \]
\[ = \sum_{|\alpha| \leq \frac{n-3}{2}} a_\alpha (1 + |\alpha|) t^{|\alpha|} \int_{|z|=1} z^\alpha (\partial_x^\alpha h)(x + tz) dS \bigg|_{t=0} \]
\[ + \sum_{|\alpha| \leq \frac{n-3}{2}} a_\alpha t^{|\alpha|+1} \sum_{|\delta|=1} \int_{|z|=1} z^{\alpha+\delta} (\partial_x^{\alpha+\delta} h)(x + tz) dS \bigg|_{t=0} \]
\[ = \left( a_0 \int_{|z|=1} dS \right) h(x), \]
which implies (4.6). Combining (4.5) and (4.6), we have

(4.7) \[ L_1(t, x) = a_0 \int_{|z|=1} \{ g_0(t, x + \gamma t z) - g_0(t, x) \} \ dS \]

\[ + \sum_{1 \leq |\alpha| \leq \frac{n-3}{2}} a_{\alpha}(1 + |\alpha|)(\gamma t)^{|\alpha|} \int_{|z|=1} z^\alpha (\partial_x^\alpha g_0)(t, x + \gamma t z) \ dS \]

\[ + \sum_{|\alpha| \leq \frac{n-3}{2}} a_{\alpha}(\gamma t)^{|\alpha|+1} \sum_{|\delta|=1} \int_{|z|=1} z^{\alpha+\delta} (\partial_x^{\alpha+\delta} g_0)(t, x + \gamma t z) \ dS. \]

Similarly, we see that

(4.8) \[ L_1(t, x) = a_0 \int_{|z| \leq 1} \frac{g_0(t, x + \gamma t z) - g_0(t, x)}{\sqrt{1 - |z|^2}} dz \]

\[ + \sum_{1 \leq |\alpha| \leq \frac{n-2}{2}} a_{\alpha}(1 + |\alpha|)(\gamma t)^{|\alpha|} \int_{|z| \leq 1} \frac{z^\alpha (\partial_x^\alpha g_0)(t, x + \gamma t z)}{\sqrt{1 - |z|^2}} dz \]

\[ + \sum_{|\alpha| \leq \frac{n-2}{2}} a_{\alpha}(\gamma t)^{|\alpha|+1} \sum_{|\delta|=1} \int_{|z| \leq 1} \frac{z^{\alpha+\delta} (\partial_x^{\alpha+\delta} g_0)(t, x + \gamma t z)}{\sqrt{1 - |z|^2}} dz. \]

when \( n \) is an even \( \geq 2 \). Concerning the estimate \( g_0(t, x) \), we have

(4.9) \[ \| \partial_t^j \partial_x^\alpha g_0(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j, \alpha, n} t^{-(j + |\alpha|)/2}, \quad j + |\alpha| \geq 1. \]

In fact, we have

\[ \partial_t^j \partial_x^\alpha g_0(t, x) = \mathcal{F}^{-1} \left[ e^{-A|\xi|^2} (-A|\xi|^2)^j (i\xi)^\alpha \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right](x) \]

and

\[ \left| \partial_\xi^\mu \left( (-A|\xi|^2)^j (i\xi)^\alpha \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi) \right) \right| \leq C_{j, k, \alpha, \mu} |\xi|^{2\ell + |\alpha| - |\mu|}, \quad \xi \neq 0. \]

Therefore, (4.9) follows from Lemma 3.5. Since

\[ g_0(t, x + \gamma t z) - g_0(t, x) = \int_0^1 \frac{d}{d\theta} \{ g_0(t, x + \gamma t z \theta) \} \ d\theta \]

\[ = \int_0^1 (\nabla_x g_0)(t, x + \gamma t z \theta) \ d\theta \cdot \gamma t z \]
by (4.9) we have

\[ (4.10) \quad \left\| a_0 \int_{|z|=1} \{ g_0(t, \cdot + \gamma tz) - g_0(t, \cdot) \} \, dS \right\|_{L_1(\mathbb{R}^n)} \leq a_0 \gamma t \int_{\mathbb{R}^n} \int_{|z|=1} \int_0^1 |z| \left| (\nabla_x g_0)(t, x + \gamma tz) \right| \, d\theta dS dx \leq C_n t^{\frac{1}{2}} \]

when \( n \) is an odd \( \geq 3 \); and

\[ (4.11) \quad \left\| a_0 \int_{|z| \leq 1} \frac{g_0(t, \cdot + \gamma tz) - g_0(t, \cdot)}{\sqrt{1 - |z|^2}} \, dz \right\|_{L_1(\mathbb{R}^n)} \leq a_0 \gamma t \int_{\mathbb{R}^n} \int_{|z| \leq 1} \int_0^1 |z| \left| (\nabla_x g_0)(t, x + \gamma tz) \right| \, d\theta dz dx \leq C_n t^{\frac{1}{2}} \]

when \( n \) is an even \( \geq 2 \). Therefore, by (4.7), (4.9) and (4.10) we have

\[ \| L_1(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_n \left\{ t^{\frac{1}{2}} + \sum_{1 \leq |\alpha| \leq \frac{n-3}{2}} t^{\frac{|\alpha|}{2}} + \sum_{|\alpha| \leq \frac{n-3}{2}} t^{\frac{|\alpha|+1}{2}} \right\} \leq C_n t^{\frac{n-1}{4}} \]

when \( n \) is an odd \( \geq 3 \); and by (4.7), (4.8) and (4.11) we have

\[ \| L_1(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_n \left\{ t^{\frac{1}{2}} + \sum_{1 \leq |\alpha| \leq \frac{n-2}{2}} t^{\frac{|\alpha|}{2}} + \sum_{|\alpha| \leq \frac{n-2}{2}} t^{\frac{|\alpha|+1}{2}} \right\} \leq C_n t^{\frac{n}{2}} \]

when \( n \) is an even \( \geq 2 \), which implies (4.4).

Next, we shall show that

\[ (4.12) \quad \| \partial_t^j \partial_x^\beta L_1(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j, \beta, n} t^{q(n)-j+|\beta|}, \quad t \geq 1, \ j + |\beta| \geq 1. \]
By (4.7) we have

\[
\begin{align*}
(4.13) & \quad \partial_t^j \partial_x^\beta L_1(t,x) \\
& = a_0 \int_{|z|=1} \partial_t^j \partial_x^\beta \left\{ g_0(t, x + \gamma tz) - g_0(t, x) \right\} \, dS \\
& + \sum_{k=0}^j \binom{j}{k} \sum_{0 \leq |\alpha| \leq \frac{n-3}{2}} a_\alpha (1 + |\alpha|) \sum_{m=0}^{\min(k+1,|\alpha|)} \frac{|\alpha|!}{m!} (\gamma t)^{|\alpha|} t^{-m} \\
& \cdot \sum_{|\delta| = k-m} \int_{|z|=1} z^{\alpha+\delta} (\partial_x^{\alpha+\beta+\delta} \partial_t^{j-k} g_0)(t, x + \gamma tz) \, dS \\
& + \sum_{k=0}^j \binom{j}{k} \sum_{0 \leq |\alpha| \leq \frac{n-5}{2}} a_\alpha \sum_{m=0}^{\min(k,|\alpha|+1)} \frac{(|\alpha|+1)!}{m!} (\gamma t)^{|\alpha|} t^{1-m} \\
& \cdot \sum_{|\delta| = k-m+1} \int_{|z|=1} z^{\alpha+\delta} (\partial_x^{\alpha+\beta+\delta} \partial_t^{j-k} g_0)(t, x + \gamma tz) \, dS
\end{align*}
\]

when \( n \) is an odd \( \geq 3 \); and by (4.8) we have

\[
\begin{align*}
(4.14) & \quad \partial_t^j \partial_x^\beta L_1(t,x) \\
& = a_0 \int_{|z|\leq1} \partial_t^j \partial_x^\beta \left\{ g_0(t, x + \gamma tz) - g_0(t, x) \right\} \frac{dz}{\sqrt{1 - |z|^2}} \\
& + \sum_{k=0}^j \binom{j}{k} \sum_{1 \leq |\alpha| \leq \frac{n-3}{2}} a_\alpha (1 + |\alpha|) \sum_{m=0}^{\min(k+1,|\alpha|)} \frac{|\alpha|!}{m!} (\gamma t)^{|\alpha|} t^{-m} \\
& \cdot \sum_{|\delta| = k-m} \int_{|z|\leq1} \frac{z^{\alpha+\delta} (\partial_x^{\alpha+\beta+\delta} \partial_t^{j-k} g_0)(t, x + \gamma tz)}{\sqrt{1 - |z|^2}} \, dz \\
& + \sum_{k=0}^j \binom{j}{k} \sum_{0 \leq |\alpha| \leq \frac{n-5}{2}} a_\alpha \sum_{m=0}^{\min(k,|\alpha|+1)} \frac{(|\alpha|+1)!}{m!} (\gamma t)^{|\alpha|} t^{1-m} \\
& \cdot \sum_{|\delta| = k-m+1} \int_{|z|\leq1} \frac{z^{\alpha+\delta} (\partial_x^{\alpha+\beta+\delta} \partial_t^{j-k} g_0)(t, x + \gamma tz)}{\sqrt{1 - |z|^2}} \, dz
\end{align*}
\]
when \( n \) is an even \( n \geq 2 \). By (4.9) we have

\[
(4.15) \quad \left\| a_0 \int_{|z|=1} \partial_t^j \partial_x^\beta \{ g_0(t, \cdot + \gamma tz) - g_0(t, \cdot) \} \, dS \right\|_{L_1(\mathbb{R}^n)} \\
\leq a_0 \int_{|z|=1} \left\{ \| \partial_t^j \partial_x^\beta g_0(t, \cdot + \gamma tz) \|_{L_1(\mathbb{R}^n)} + \| \partial_x^\beta \partial_t^j g_0(t, \cdot) \|_{L_1(\mathbb{R}^n)} \right\} \, dS \\
\leq C_{j,\beta,n} t^{-\left(j+\frac{|eta|}{2}\right)},
\]

and

\[
(4.16) \quad \left\| a_0 \int_{|z|\leq1} \frac{\partial_t^j \partial_x^\beta \{ g_0(t, \cdot + \gamma tz) - g_0(t, \cdot) \}}{\sqrt{1-|z|^2}} \, dz \right\|_{L_1(\mathbb{R}^n)} \\
\leq a_0 \int_{|z|=1} \left\{ \| \partial_t^j \partial_x^\beta g_0(t, \cdot + \gamma tz) \|_{L_1(\mathbb{R}^n)} + \| \partial_x^\beta \partial_t^j g_0(t, \cdot) \|_{L_1(\mathbb{R}^n)} \right\} \, dz \\
\leq C_{j,\beta,n} t^{-\left(j+\frac{|eta|}{2}\right)}.
\]

Putting

\[
p(n) = \begin{cases} 
\frac{n-3}{2}, & \text{when } n \text{ is an odd } \geq 3, \\
\frac{n-2}{2}, & \text{when } n \text{ is an even } \geq 2,
\end{cases}
\]

by (4.9), (4.13), (4.14), (4.15) and (4.16), we have

\[
\| \partial_t^j \partial_x^\beta L_1(t, \cdot) \|_{L_1(\mathbb{R}^n)} \\
\leq C_{j,\beta,n} \left\{ t^{-\left(j+\frac{|eta|}{2}\right)} + \sum_{k=0}^j \sum_{0 \leq |\alpha| \leq p(n)} \min(k+1,|\alpha|) \sum_{m=0}^{\min(k+1,|\alpha|)} t^{-\left(j+\frac{|eta|-|\alpha|+m-k}{2}\right)} \\
+ \sum_{k=0}^j \sum_{0 \leq |\alpha| \leq p(n)} \min(k,|\alpha|+1) \sum_{m=0}^{\min(k,|\alpha|+1)} t^{-\left(j+\frac{|eta|-|\alpha|+m-k-1}{2}\right)} \right\} \\
\leq C_{j,\beta,n} t^{q(n)-\left(j+\frac{|eta|}{2}\right)},
\]

which implies (4.12).

Now we shall estimate

\[
\partial_t^j \partial_x^\beta L_2(t, x) = (\alpha - A)^j \mathcal{F}^{-1} \left[ \left( e^{-|\xi|^2t} - e^{-\alpha |\xi|^2t} \right) |\xi|^{2\ell(i\xi)^\beta \xi_k \xi_k} \varphi_0(\xi) \right](x).
\]

Since

\[
e^{-A|\xi|^2t} - e^{-\alpha |\xi|^2t} = (\alpha - A)|\xi|^2 t \int_0^1 e^{-\theta A|\xi|^2t - (1-\theta)\alpha |\xi|^2t} \, d\theta,
\]

we have

\[
\| \partial_t^j \partial_x^\beta L_2(t, \cdot) \|_{L_1(\mathbb{R}^n)} \\
\leq C_{j,\beta,n} t^{\frac{n-3}{2}} \left\{ t^{-\left(j+\frac{|eta|}{2}\right)} \sum_{k=0}^j \sum_{0 \leq |\alpha| \leq p(n)} \min(k+1,|\alpha|) \sum_{m=0}^{\min(k+1,|\alpha|)} t^{-\left(j+\frac{|eta|-|\alpha|+m-k}{2}\right)} \\
+ \sum_{k=0}^j \sum_{0 \leq |\alpha| \leq p(n)} \min(k,|\alpha|+1) \sum_{m=0}^{\min(k,|\alpha|+1)} t^{-\left(j+\frac{|eta|-|\alpha|+m-k-1}{2}\right)} \right\} \\
\leq C_{j,\beta,n} t^{q(n)-\left(j+\frac{|eta|}{2}\right)},
\]
we have
\[
\left| \partial_\xi^\ell \left\{ \left( e^{-A|\xi|^2 t} - e^{-\alpha|\xi|^2 t} \right) |\xi|^{2\ell} (i\xi)^{\beta} \xi^k \xi^k_0(\xi) \right\} \right| \\
\leq C_{\ell,\beta,n} |\xi|^{2\ell + |\beta| + 2 - |n|} t, \quad \xi \neq 0.
\]
Therefore, by Lemma 3.5 we have
\[
\| \partial_\xi^\ell \partial_\xi^\beta L_2(t, \cdot) \|_{L^1(\mathbb{R}^n)} \leq C_{\ell,\beta,n} t^{-\left( j + \frac{|\beta|}{2} \right)}, \quad t > 0.
\]
Next, we shall show that for \( t \geq 1 \)
\[
\begin{align*}
\| \partial_\xi^\ell \partial_\xi^\beta M_0^0(t, \cdot) \|_{L^1(\mathbb{R}^n)} &\leq C_{j,\beta,n} t^{q(n) - \frac{j + |\beta| + 1}{2}}, \quad \ell \geq 1, \\
\| \partial_\xi^\ell \partial_\xi^\beta M_1^0(t, \cdot) \|_{L^1(\mathbb{R}^n)} &\leq C_{j,\beta,n} t^{q(n) - \frac{j + |\beta| + \ell - 1}{2}}, \quad \ell \geq 1, \\
\| \partial_\xi^\ell \partial_\xi^\beta M_2^0(t, \cdot) \|_{L^1(\mathbb{R}^n)} &\leq C_{j,\beta,n} t^{q(n) - \frac{j + |\beta| + \ell}{2}}, \quad \ell \geq 0.
\end{align*}
\]
To do this, we put
\[
\begin{align*}
\psi_0^1(t, x) &= e^{-\frac{4}{3} |\xi|^2 t} \frac{g(|\xi|^2)}{\gamma f(|\xi|)} \xi_j \xi_k \varphi_0(\xi), \\
\psi_1^1(t, x) &= e^{-\frac{4}{3} |\xi|^2 t} \frac{1}{\gamma f(|\xi|)} (|\xi|^2 g(|\xi|^2 t) \ell \xi_j \xi_k \varphi_0(\xi), \\
\psi_2^1(t, x) &= e^{-\frac{4}{3} |\xi|^2 t} \frac{-A}{\gamma f(|\xi|)} (|\xi|^2 g(|\xi|^2 t) \ell \xi_j \xi_k \varphi_0(\xi), \\
\psi_3^1(t, x) &= e^{-\frac{4}{3} |\xi|^2 t} \frac{g(|\xi|^2)}{\gamma f(|\xi|)} (|\xi|^2 g(|\xi|^2 t) \ell - 1 \xi_j \xi_k \varphi_0(\xi).
\end{align*}
\]
Then, we have
\[
\begin{align*}
M_0^0(t, x) &= F^{-1} \left[ \partial_\xi \left( \frac{\sin \gamma |\xi|^t}{|\xi|} \right) e^{-\frac{4}{3} |\xi|^2 t} \psi_0^1(t, \xi) \right] (x), \\
M_1^0(t, x) &= -F^{-1} \left[ \partial_\xi \left( \frac{\sin \gamma |\xi|^t}{|\xi|} \right) e^{-\frac{4}{3} |\xi|^2 t} \psi_1^1(t, \xi) \right] (x), \\
M_2^0(t, x) &= F^{-1} \left[ \partial_\xi \left( \frac{\sin \gamma |\xi|^t}{|\xi|} \right) e^{-\frac{4}{3} |\xi|^2 t} \psi_2^1(t, \xi) \right] (x), \\
M_3^0(t, x) &= F^{-1} \left[ \partial_\xi \left( \frac{\sin \gamma |\xi|^t}{|\xi|} \right) e^{-\frac{4}{3} |\xi|^2 t} \psi_3^1(t, \xi) \right] (x).
\end{align*}
\]
Concerning the estimate \( \psi_1^\ell(t, \xi) \), we have
\[
\left| \partial_\xi^\mu \psi_1^\ell(t, \xi) \right| \leq C_{\mu,\ell} |\xi|^{2 - |\mu|}, \quad \forall \mu.
\]
Therefore, if we put
\[
g_1^\ell(t, x) = F^{-1} \left[ \psi_1^\ell(t, \xi) \right] (x),
\]
then, by Lemma 3.5 we have
\[
\| \partial_\xi^\ell \partial_\xi^\beta g_1^\ell(t, \cdot) \|_{L^1(\mathbb{R}^n)} \leq C_{j,\beta,k,\ell,n} t^{-\left( j + \frac{1+|\beta|}{2} \right)}.
\]
In view of (4.19) and (4.20), we consider the function:

$$N(t, x) = \mathcal{F}^{-1} \left[ \partial_t^\ell \left( \frac{\sin \gamma |\xi| t}{|\xi|} \right) \hat{G}(t, \xi) \right](x),$$

where $G(t, x)$ satisfies the following conditions:

$$\| \partial_t^j \partial_x^\beta G(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j, \beta, n} t^{-\frac{(1+j+|\beta|)}{2}}. \quad (4.21)$$

In order to prove (4.18), it suffices to show that

$$\| \partial_t^j \partial_x^\beta N(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j, \beta, \ell, n} t^{q(n)-\frac{j+|\beta|+\ell}{2}}. \quad (4.22)$$

By (3.17), we have

$$\partial_t^j \partial_x^\beta N(t, x)$$

when $n$ is an odd $\geq 3$; and

$$\partial_t^j \partial_x^\beta N(t, x)$$

when $n$ is an even $\geq 2$. Therefore, by (4.21) we have

$$\| \partial_t^j \partial_x^\beta N(t, \cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j, \beta, \ell, n} t^{q(n)-\frac{j+|\beta|+\ell}{2}},$$

which implies (4.22).

In order to estimate the remainder term $\mathcal{R}_N(t, x)$, we consider the function:

$$R_{\mathcal{R}_N}(t, x) = \mathcal{F}^{-1} \left[ \tau_{\mathcal{R}_N}(t, \xi) \right](x),$$
where
\[
r_{N,\psi}^\pm(t, \xi) = \frac{e^{-A|\xi|^2 t}}{2i\gamma|\xi|f(\xi)} \int_0^1 (1 - \theta)^N e^{\pm i\gamma|\xi|(1 + \theta|\xi|^2)g(|\xi|^2) t} d\theta
\]
\[
\cdot \left(\pm i\gamma|\xi|^3 g(|\xi|^2)\right)^N \psi(\xi) \varphi_0(\xi),
\]
and \(\psi \in C^\infty(\mathbb{R}^n - \{0\})\) satisfies the condition:
\[
\left| \partial_\xi^\gamma \psi(\xi) \right| \leq C_{\gamma} |\xi|^{-|\gamma|}, \quad \forall \xi \neq 0.
\]
If \(\psi(\xi) = \xi_j \xi_k |\xi|^2\), then we have
\[
(4.23) \quad R_{N}(t, x) = \partial_t \left\{ \left( R_{N,\psi}^+(t, x) - R_{N,\psi}^-(t, x) \right) t^N + 1 \right\}.
\]
First, we observe that
\[
(4.24) \quad \left| \partial_\xi^\delta e^{\pm i\gamma|\xi|(1 + \theta|\xi|^2)g(|\xi|^2) t} \right| \leq C_{\delta} |\xi|^{|\delta|} \left( t |\xi|^2 \right), \quad 0 \leq \theta \leq 1, \ \xi \neq 0.
\]
In fact, by the formula of derivative of composed function (cf. (3.19)), we have
\[
(4.25) \quad \partial_\xi^\delta e^{\pm i\gamma|\xi|(1 + \theta|\xi|^2)g(|\xi|^2) t}
\]
\[
= \sum_{\ell=1}^{\delta} (\pm i\gamma t)^\ell e^{\pm i\gamma|\xi|(1 + \theta|\xi|^2)g(|\xi|^2) t} \cdot \sum_{|\alpha_1| + \cdots + |\alpha_\ell| = |\delta|} \partial_\xi^{\alpha_1} \left\{ |\xi|(1 + \theta|\xi|^2)g(|\xi|^2) \right\} \cdots \partial_\xi^{\alpha_\ell} \left\{ |\xi|(1 + \theta|\xi|^2)g(|\xi|^2) \right\}.
\]
Since
\[
\left| \partial_\xi^{\alpha} \left\{ |\xi|(1 + \theta|\xi|^2)g(|\xi|^2) \right\} \right| \leq C_{\alpha,\nu} |\xi|^{1-|\alpha|}, \quad \xi \in \text{supp } \varphi_0,
\]
by (4.25) we have
\[
\left| \partial_\xi^\delta e^{\pm i\gamma|\xi|(1 + \theta|\xi|^2)g(|\xi|^2) t} \right| \leq C_{\delta} \sum_{\ell=1}^{\delta} (t |\xi|^2)^{\ell} |\xi|^{-|\delta|}, \quad \xi \in \text{supp } \varphi_0.
\]
Therefore, we have for \( \xi \in \text{supp} \varphi_0 \) and \( \xi \neq 0 \)
\[
\left| \partial^\delta_x \left\{ (-A|\xi|^2 + i\gamma|\xi|(1 + \theta)|\xi|^2)g(|\xi|^2)j(i\xi)^3 r_{N,\psi}^\pm(t, \xi) \right\} \right|
\leq C_\delta \sum_{0 \leq \nu \leq \delta} |\xi|^{3N+2+j+|\beta|-|\nu|} e^{-\frac{3}{4}|\xi|^2t} \sum_{\ell=1}^\infty (t|\xi|)^\ell |\xi|^{-\delta-\nu}
\leq C_\delta \sum_{0 \leq \nu \leq \delta} |\xi|^{3N+2+j+|\beta|-|\nu|} e^{-\frac{3}{4}|\xi|^2t} \sum_{\ell=1}^\infty (t|\xi|)^\ell |\xi|^{-\ell}.
\]
Since \( |\xi|^{-\ell} \leq C_\delta |\xi|^{-|\delta|} \) \((0 \leq \ell \leq |\delta|)\) when \( \xi \in \text{supp} \varphi_0 \) and \( \xi \neq 0 \), by the above inequality we have (4.24). Since
\[
(4.26) \quad e^{ix\cdot \xi} = \sum_{j=1}^n \frac{x_j}{i|x|^2} \partial_{x_j} e^{ix\cdot \xi},
\]
in view of (4.24) by \( n + 1 \)-times integration by parts, we have
\[
\partial_t^j \partial_x^\beta R_{N,\psi}^\pm(t, x)
= \sum_{|\nu|=n+1} \left( \frac{ix}{|x|^2} \right)^{\delta-n} \left( \frac{1}{2\pi} \right)^n \int_0^1 \int_{\mathbb{R}^n} e^{-x\cdot \xi} \partial^\delta \xi
\cdot \left\{ (-A|\xi|^2 + i\gamma|\xi|(1 + \theta)|\xi|^2)g(|\xi|^2)j(i\xi)^3 r_{N,\psi}^\pm(t, \xi) \right\} d\xi d\theta.
\]
when \( N > n/3 \). Therefore, by (4.24) we have
\[
\left| \partial_t^j \partial_x^\beta R_{N,\psi}^\pm(t, x) \right| \leq C_{j,\beta,N,n} |x|^{-(1+n)} \int_{\mathbb{R}^n} |\xi|^{3N+2+j+|\beta|-2(n+1)} e^{-\frac{3}{4}|\xi|^2t} d\xi
\leq C_{j,\beta,N,n} |x|^{-(n+1)} t^{-\frac{3N+2+j-|\beta|-n}{2}}.
\]
On the other hand, by (4.24) with \( \delta = 0 \) we have
\[
\left| \partial_t^j \partial_x^\beta R_{N,\psi}^\pm(t, x) \right| \leq C_{j,\beta,N} \int_{\mathbb{R}^n} |\xi|^{3N+2+j+|\alpha|} e^{-\frac{3}{4}|\xi|^2t} d\xi
\leq C_{j,\beta,N} t^{-\frac{3N+2+j+|\beta|+n}{2}}.
\]
Combining these two estimations, we have
\[
(4.27) \quad \| \partial_t^j \partial_x^\beta R_{N,\psi}^\pm(t, \cdot) \|_{L_1(\mathbb{R}^n)}
\leq C_{j,\beta,N,n} \left\{ \int_{|x|\leq \sqrt{t}} t^{-\frac{3N+2+j+|\beta|+n}{2}} dx + \int_{|x|\geq \sqrt{t}} t^{-\frac{3N+2+j+|\alpha|-n}{2}} |x|^{-n+1} dx \right\}
\leq C_{j,\beta,N,n} t^{-\frac{3N+|\alpha|+2+j+|\beta|+1-n}{2}}, \quad t \geq 1.
\]
By (4.31) and Leibniz’ rule, we have
\[
\|\partial_t^j \partial_x^3 R_N(t, \cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{j, \beta, N} t^{-\frac{N-n+3|\beta|-1}{2}}, \quad t \geq 1, \; N \geq \frac{n}{3}.
\]
Combining (4.4), (4.12), (4.17), (4.18) and (4.28), we have (4.3). To complete the Proof of Theorem 2.1 (2), we have to show that
\[
\|\partial_t^j \partial_x^3 L_0(t, \cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{j, \alpha, n}, \quad 0 \leq t \leq 1.
\]
Regarding the relations:
\[
\lambda_{\pm}(\xi) = -A|\xi|^2 \mp i\gamma|\xi|f(|\xi|),
\]
we put
\[
L_0(t, x) = \sum_{j=1}^3 \mathcal{F}^{-1} [\psi_j(t, \xi)](x),
\]
where
\[
\psi_1(t, \xi) = \frac{A|\xi|\epsilon^{-A|\xi|^2t}}{2i\gamma f(|\xi|)} \left( e^{-i\gamma|\xi|f(|\xi|)t} - e^{i\gamma|\xi|f(|\xi|)t} \right) \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi),
\]
\[
\psi_2(t, \xi) = e^{-A|\xi|^2t} \left( e^{-i\gamma|\xi|f(|\xi|)t} - e^{i\gamma|\xi|f(|\xi|)t} \right) \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi),
\]
\[
\psi_3(t, \xi) = \left( e^{-A|\xi|^2t} - e^{-\alpha|\xi|^2t} \right) \frac{\xi_j \xi_k}{|\xi|^2} \varphi_0(\xi).
\]
First, we shall estimate \( \mathcal{F}^{-1} [\psi_1(t, \xi)](x) \). By the formula of derivative of composed function (cf. (3.19)), we have
\[
\partial^\delta_{\xi} e^{\pm i\gamma|\xi|f(|\xi|)t} = \sum_{\ell=1}^{[\delta]} (\pm i\gamma t)^\ell e^{\pm i\gamma|\xi|f(|\xi|)t} \sum_{|\alpha_1|+\cdots+|\alpha_{\ell}|=|\delta| \atop |\alpha_i| \geq 1} \partial^{\alpha_1}_{\xi} \{ |\xi|f(|\xi|) \} \cdots \partial^{\alpha_{\ell}}_{\xi} \{ |\xi|f(|\xi|) \}.
\]
Since
\[
\left| \partial^\nu_{\xi} |\xi|f(|\xi|) \right| \leq C_{\alpha, \nu} |\xi|^{1-|\alpha|}, \quad |\xi| \leq \frac{B}{\sqrt{2}}, \; \xi \neq 0,
\]
by (4.30) we have
\[
\left| \partial^\delta_{\xi} e^{\pm i\gamma|\xi|f(|\xi|)t} \right| \leq C_{\delta} \sum_{\ell=1}^{[\delta]} |\xi|^{\ell-|\delta|} \leq C_{\delta} |\xi|^{-|\delta|}, \quad |\xi| \leq \frac{B}{\sqrt{2}}, \; \xi \neq 0.
\]
By (4.31) and Leibniz’ rule, we have
\[
\left| \partial^\delta_{\xi} \left\{ \partial_t^j \psi_1(t, \xi) (i\xi)^\alpha \right\} \right| \leq C_{j, \delta, \alpha} |\xi|^{1-|\delta|}, \quad |\xi| \leq \frac{B}{\sqrt{2}}, \; \xi \neq 0.
\]
Since $\text{supp} \psi_1(t, \cdot) \subset \{ \xi \in \mathbb{R}^n \mid |\xi| \leq B/\sqrt{2} \}$, by Proposition 3.4 with $(\alpha, N, \sigma) = (1, n, 1)$ we have

$$\left| \partial_t^\alpha \partial_x^\gamma F^{-1} [\psi_1(t, \xi)] (x) \right| \leq C_{j,\alpha,n} |x|^{-n-1}, \quad \forall x \neq 0.$$ 

On the other hand, we have

$$\left| \partial_t^\alpha \partial_x^\gamma F^{-1} [\psi_1(t, \xi)] (x) \right| \leq C_{j,\alpha,n} \int_{|\xi| \leq B/\sqrt{2}} |\xi| \, d|\xi| \leq C_{j,\alpha,n}.$$ 

Therefore, combining these two estimations, we have

(4.32) $$\| \partial_t^\alpha \partial_x^\gamma F^{-1} [\psi_1(t, \xi)] (\cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j,\alpha,n} \left\{ \int_{|x| \leq 1} dx + \int_{|x| \geq 1} |x|^{-(n+1)} \, dx \right\} \leq C_{j,\alpha,n}.$$ 

Next, we shall estimate $F^{-1} [\psi_2(t, \xi)] (x)$. By Taylor's formula, we have

$$\psi_2(t, \xi) = -i \gamma f(|\xi|) te^{-A|\xi|^2 t} \int_0^1 \sin(\theta \gamma |\xi| f(|\xi|)) t \, d\theta \xi_j \xi_k \phi_0(\xi).$$

Therefore, we have

$$\left| \partial_{\xi}^\delta \left\{ (i\xi)^\alpha \partial_t^\gamma \psi_2(t, \xi) \right\} \right| \leq C_{j,\alpha,\delta} |\xi|^{1-|\delta|}, \quad \xi \neq 0.$$ 

Employing the same argument as in $F^{-1} [\psi_1(t, \xi)] (x)$, we have

(4.33) $$\| \partial_t^\alpha \partial_x^\gamma F^{-1} [\psi_2(t, \xi)] (\cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j,\alpha,n}.$$ 

Finally, we shall estimate $F^{-1} [\psi_3(t, \xi)] (x)$. Since

$$\psi_3(t, x) = (A - \alpha) t \int_0^1 e^{-(1-\theta)(A+\theta \alpha)|\xi|^2 t} d\theta \xi_j \xi_k \phi_0(\xi),$$

we have

$$\left| \partial_{\xi}^\delta \left\{ (i\xi)^\alpha \partial_t^\gamma \psi_3(t, \xi) \right\} \right| \leq C_{j,\alpha,\delta} |\xi|^{2-|\delta|}.$$ 

Therefore, employing the same argument as in $F^{-1} [\psi_1(t, \xi)] (x)$, we have

(4.34) $$\| \partial_t^\alpha \partial_x^\gamma F^{-1} [\psi_3(t, \xi)] (\cdot) \|_{L_1(\mathbb{R}^n)} \leq C_{j,\alpha,n}.$$ 

Combining (4.32), (4.33) and (4.34), we have (4.29), which completes the proof.
5. Proof of Theorem 2.2.

In this section, we shall prove Theorem 2.2. First, we consider the part where $|\xi| \geq \sqrt{2}B$. Since $\lambda_{\pm}(\xi) = -A \left( |\xi|^2 \pm \sqrt{|\xi|^4 - B^2|\xi|^2} \right)$ when $|\xi| \geq \sqrt{2}B$, we write:

\begin{equation}
\lambda_{\pm}(\xi) = -A|\xi|^2 + 1 + \mu(\xi), \quad \lambda_{\mp}(\xi) = -1 - \mu(\xi),
\end{equation}

where

\[
\mu(\xi) = \frac{AB^4}{4|\xi|^2} g \left( \frac{B^2}{|\xi|^2} \right), \quad g(s) = \int_0^1 (1 - \theta s)^{-\frac{3}{2}} (1 - \theta) d\theta.
\]

Note that $g(B^2/|\xi|^2) \in C^\infty$ when $|\xi| \geq 2$. In view of (2.5) and (2.9), we put

\[
L_{\pm}(t)u(x) = \mathcal{F}^{-1} \left[ \frac{\lambda_{\mp}(\xi)e^{\lambda_{\pm}(\xi)t}}{\lambda_{\pm}(\xi) - \lambda_{\mp}(\xi)} \varphi_{\infty}(\xi) \hat{u}(\xi) \right](x),
\]

\[
M_{\pm,\beta}(t)u(x) = \mathcal{F}^{-1} \left[ \frac{\xi^{\beta}e^{\lambda_{\pm}(\xi)t}}{\lambda_{\pm}(\xi) - \lambda_{\mp}(\xi)} \varphi_{\infty}(\xi) \hat{u}(\xi) \right](x), \quad |\beta| = 1,
\]

\[
K_{\pm,\infty}(t)v_0(x) = \mathcal{F}^{-1} \left[ \frac{\lambda_{\pm}(\xi)e^{\lambda_{\pm}(\xi)t}}{\lambda_{\pm}(\xi) - \lambda_{\pm}(\xi)} \frac{\xi \xi_k}{|\xi|^2} \varphi_{\infty}(\xi) \hat{v}_0(\xi) \right](x),
\]

\[
K_{1,\infty}(t)v_0(x) = \mathcal{F}^{-1} \left[ e^{-\alpha|\xi|^2t} \left( \delta_{jk} - \frac{\xi \xi_k}{|\xi|^2} \right) \varphi_{\infty}(\xi) \hat{v}_0(\xi) \right](x).
\]

By [6, Theorem 4.2.1], we have for $p = 1$ or $\infty$

\begin{equation}
\|\partial_t^j \partial_x^\alpha L_{\pm}(t)u\|_{L_p(\mathbb{R}^n)} \leq C_{j,k,\alpha} t^{-(j-k)} e^{-ct} \|u\|_{W_p^{2k+|\alpha|+1}(\mathbb{R}^n)},
\end{equation}

\[
\|\partial_t^j \partial_x^\alpha M_{\pm,\beta}(t)u\|_{L_p(\mathbb{R}^n)} \leq C_{j,k,\alpha} t^{-(j-k)} e^{-ct} \|u\|_{W_p^{2k+|\alpha|}(\mathbb{R}^n)}, \quad |\beta| = 1,
\]

\[
\|\partial_t^j \partial_x^\alpha (L_{-}(t)u - e^{-t}u)\|_{L_p(\mathbb{R}^n)} \leq C_{j,\alpha} t^{-\frac{j}{2}} e^{-ct} \|u\|_{W_p^{1+|\alpha|}(\mathbb{R}^n)},
\]

\[
\|\partial_t^j \partial_x^\alpha M_{-}(\beta)u\|_{L_p(\mathbb{R}^n)} \leq C_{j,\alpha} t^{-\frac{j}{2}} e^{-ct} \|u\|_{W_p^{1+|\alpha|}(\mathbb{R}^n)}, \quad |\beta| = 1.
\]

Now we shall show that for $p = 1$ or $\infty$

\begin{equation}
\|\partial_t^j \partial_x^\alpha K_{+,\infty}(t)v_0\|_{L_p(\mathbb{R}^n)} \leq C_{j,\alpha,n} \left( 1 + t^{-\frac{j}{2}} \right) t^{-(j-k)} e^{-ct} \|v_0\|_{W_p^{2k+|\alpha|}(\mathbb{R}^n)},
\end{equation}

\[
\|\partial_t^j \partial_x^\alpha K_{1,\infty}(t)v_0\|_{L_p(\mathbb{R}^n)} \leq C_{j,\alpha,n} \left( 1 + t^{-\frac{j}{2}} \right) t^{-(j-k)} e^{-ct} \|v_0\|_{W_p^{2k+|\alpha|}(\mathbb{R}^n)},
\]
Put

\[ K_{+j,k}(t, x) = \mathcal{F}^{-1} \left[ \frac{\lambda_+^j e^{\lambda_+^j + \lambda_-^j t}(1 + |\xi|^2)^{j-k}}{1 + |\xi|^2} \frac{\xi_m \xi_\ell}{|\xi|^2} \varphi_{\infty}(\xi) \right](x), \]

\[ K_{1j,k}(t, x) = \mathcal{F}^{-1} \left[ \frac{(-\alpha |\xi|^2)^j e^{-\alpha |\xi|^2 t}(1 + |\xi|^2)^{j-k}}{1 + |\xi|^2} \right. \]

\[ \cdot \left( \delta_{m\ell} - \frac{\xi_m \xi_\ell}{|\xi|^2} \right) \varphi_{\infty}(\xi)(x), \]

for \( j \leq k \leq 0 \), and then

\[ \partial_t^j \partial_\nu^r K_{+\omega}(t) v_0 = K_{+j,k}(t, \cdot) * \partial_\nu^r (1 - \Delta)^k v_0; \]

\[ \partial_t^j \partial_\nu^r K_{1\omega}(t) v_0 = K_{1j,k}(t, \cdot) * \partial_\nu^r (1 - \Delta)^k v_0. \]

By (5.1) we have for \(|\xi| \geq \sqrt{2B}\)

\[ \partial_\xi^r \left( \lambda_+^j e^{\lambda_+^j + \lambda_-^j} \frac{|\xi_m \xi_\ell}{|\xi|^2} \varphi_{\infty}(\xi) \right) \]

\[ \leq C_{j,k,\nu} t^{-(j-k)} e^{-c_1 t} |\xi|^{-|\nu|} e^{-c_2 |\xi|^2 t}, \quad \forall \nu, \]

and also we have

\[ \partial_\xi^r \left( (-\alpha |\xi|^2)^j e^{-\alpha |\xi|^2 t} \frac{|\xi_m \xi_\ell}{|\xi|^2} \varphi_{\infty}(\xi) \right) \]

\[ \leq C_{j,k,\nu} t^{-(j-k)} e^{-c_1 t} |\xi|^{-|\nu|} e^{-c_2 |\xi|^2 t}, \quad \forall \nu. \]

Therefore, using (4.26) and the integration by parts \( n + 1 \) times, by (5.6) we have

\[ |L_{+j,k}(t, x)| \leq C_{j,k,n} \frac{t^{-(j-k)} e^{-c_1 t}}{|x|^{n+1}} \int_{|\xi| \geq \sqrt{2B}} |\xi|^{-n-1} e^{-c_2 |\xi|^2 t} d\xi \]

\[ \leq C_{j,k,n} \frac{t^{-(j-k)} e^{-c_1 t}}{|x|^{n+1}}, \]

and by (5.7) we have

\[ |L_{1j,k}(t, x)| \leq C_{j,k,n} t^{-(j-k)} e^{-c_1 t}. \]

On the other hand, by (5.6) we have

\[ |L_{+j,k}(t, x)| \leq C_{j,k,n} t^{-(j-k)} e^{-c_1 t} \int_{|\xi| \geq \sqrt{2B}} e^{-c_2 |\xi|^2 t} d\xi \]

\[ \leq C_{j,k,n} t^{-(j-k)} \frac{n}{2} e^{-c_1 t}, \]
and by (5.7) we have
\[ |L_{1,j,k}(t, x)| \leq C_{j,k,n} t^{-(j-k)} e^{-ct}. \]

Therefore, by (5.8) and (5.10) we have
\[ \|L_{+,j,k}(t, \cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{j,k,n} t^{-(j-k)} e^{-ct}. \]

By (5.12), (5.13) and the Young inequality, we have (5.3).

Next, we shall show that for \( p = 1 \) or \( \infty \)
\[ \|\partial_j t \partial_\alpha x K_{-,\infty}(t)v_0\|_{L^p(\mathbb{R}^n)} \leq C_{j,k,\alpha,n} t^{-(j-k)} e^{-ct} \|v_0\|_{W^{2k+|\alpha|}(\mathbb{R}^n)}. \]

Put
\[ \ell_-(t, x) = F^{-1} \left[ \frac{\lambda_-(\xi)e^{\lambda_-(\xi)t}}{\lambda_+(\xi) - \lambda_-(\xi)} \xi_j \xi_k \varphi_\infty(\xi) \right] (x), \]
and then
\[ K_{-,\infty}(t)v_0(x) = \ell_-(t, \cdot) * v_0. \]
Now, we shall prove that
\[ \|\partial_j^a \ell_-(t, \cdot)\|_{L^1(\mathbb{R}^n)} \leq C_j e^{-ct}. \]

By (5.17) we have for \( |\xi| \geq \sqrt{2}B \)
\[ \|\partial_j^a \ell_-(t, \cdot)\|_{L^1(\mathbb{R}^n)} \leq C_{j,\beta}(1+t)^{|\beta|} e^{-ct} |\xi|^{-|\beta|}. \]

Therefore, using (4.26) and the integration by parts \( n - 1 \) times, by (5.17) we have
\[ \left| \partial_j^a \ell_-(t, x) \right| \leq C_{j,n} e^{-ct} \begin{cases} |x|^{-(n-1)}, & 0 < |x| \\ |x|^{-(n+1)}, & |x| \geq 1. \end{cases} \]
which implies (5.16). By (5.16) and the Young inequality, we have (5.14).

In order to complete the Proof of Theorem 2.2, we have to estimate the part where \( B/2 \leq |\xi| \leq 2B \) (cf. (2.10)). In view of (2.6) and (2.9), below, if we put
(5.18)\[ N_0,\psi(t, x) = \frac{1}{2\pi i} \mathcal{F}^{-1} \left[ \int_{\Gamma} \frac{(z + |\xi|^2)e^{zt}}{z^2 + (\alpha + \beta)|\xi|^2z + \gamma^2|\xi|^2} \, dz \psi(\xi) \varphi_M(\xi) \right] (x); \]
\[ N_1,\psi(t, x) = \frac{1}{2\pi i} \mathcal{F}^{-1} \left[ i\gamma \xi \int_{\Gamma} \frac{e^{zt}}{z^2 + (\alpha + \beta)|\xi|^2z + \gamma^2|\xi|^2} \, dz \psi(\xi) \varphi_M(\xi) \right] (x); \]
\[ N_2,\psi(t, x) = \mathcal{F}^{-1} \left[ e^{-\alpha|\xi|^2t}\psi(\xi) \varphi_M(\xi) \right] (x), \]
where \( \psi \in C^\infty(S^{n-1}) \) and \( \psi = \psi(\xi/|\xi|) \), then we have
\[ \mathcal{F}^{-1} [\varphi_M(\xi) \hat{\rho}(t, \xi)] (x) = N_0,\psi(t, \cdot) * \rho_0 + N_1,\psi(t, \cdot) * v_0; \]
\[ \mathcal{F}^{-1} [\varphi_M(\xi) \hat{v}(t, \xi)] (x) = N_1,\psi(t, \cdot) * \rho_0 + N_0,\psi(t, \cdot) * v_0 + N_2,\psi(t, \cdot) * v_0. \]

If we use (4.26) and (2.7), then we see easily that
\[ \left| \partial_t^j \partial_x^\alpha N_{t,\psi}(t, x) \right| \leq C_{j,\alpha,N} e^{-ct|x|^{-N}}, \quad \forall \, N \geq 0, \text{ integer}. \]

Therefore, applying the Young inequality to (5.18) we have
(5.19)\[ \| \mathcal{F}^{-1} [\varphi_M(\xi) \hat{\rho}(t, \xi)] \|_{L^p(\mathbb{R}^n)} \leq C_{j,\alpha,p} e^{-ct} \| (\rho_0, v_0) \|_{L^p(\mathbb{R}^n)}, \quad 1 \leq p \leq \infty. \]

Combining (5.2), (5.3), (5.14) and (5.19), we have Theorem 2.2.

References

Received February 27, 2001.

DEPARTMENT OF MATHEMATICS
KYUSHU INSTITUTE OF TECHNOLOGY
1-1 SENSUI-CHO, TOBATA-KU
KITAKYUSUYU 804-8550
JAPAN
E-mail address: kobayasi@math.tsukuba.ac.jp

DEPARTMENT OF MATHEMATICS
SCHOOL OF SCIENCE AND ENGINEERING
WASEDA UNIVERSITY
3-4-1 OHKUBO SHINJUKU-KU
TOKYO 169-8555
JAPAN
E-mail address: yshibata@waseda.jp
PURE-PERIODIC MODULES AND A STRUCTURE OF PURE-PROJECTIVE RESOLUTIONS

Daniel Simson

Dedicated to Stanislaw Balcerzyk on the occasion of his seventieth birthday

We investigate the structure of pure-syzygy modules in a pure-projective resolution of any right $R$-module over an associative ring $R$ with an identity element. We show that a right $R$-module $M$ is pure-projective if and only if there exists an integer $n \geq 0$ and a pure-exact sequence $0 \rightarrow M \rightarrow P_n \rightarrow \cdots \rightarrow P_0 \rightarrow M \rightarrow 0$ with pure-projective modules $P_n, \ldots, P_0$.

As a consequence we get the following version of a result in Benson and Goodearl, 2000: A flat module $M$ is projective if $M$ admits an exact sequence $0 \rightarrow M \rightarrow F_n \rightarrow \cdots \rightarrow F_0 \rightarrow M \rightarrow 0$ with projective modules $F_n, \ldots, F_0$.

1. Introduction.

Throughout this paper $R$ is an associative ring with an identity element. We denote by Mod($R$) the category of all right $R$-modules. We recall (see [12]) that an exact sequence $\cdots \rightarrow X_{n-1} \rightarrow X_n \rightarrow X_{n+1} \rightarrow \cdots$ in Mod($R$) is said to be pure (in the sense of Cohn [6]) if the induced sequence $\cdots \rightarrow X_{n-1} \otimes_R L \rightarrow X_n \otimes_R L \rightarrow X_{n+1} \otimes_R L \rightarrow \cdots$ of abelian groups is exact for any left $R$-module $L$. An epimorphism $f : Y \rightarrow Z$ in Mod($R$) is said to be pure if the exact sequence $0 \rightarrow \text{Ker } f \rightarrow Y \rightarrow Z \rightarrow 0$ is pure. A submodule $X$ of a right $R$-module $Y$ is said to be pure if the exact sequence $0 \rightarrow X \rightarrow Y \rightarrow Y/X \rightarrow 0$ is pure. A module $P$ in Mod($R$) is said to be pure-projective if for any pure epimorphism $f : Y \rightarrow Z$ in Mod($R$) the induced group homomorphism $\text{Hom}_R(P, f) : \text{Hom}_R(P, Y) \rightarrow \text{Hom}_R(P, Z)$ is surjective. The following facts are well-known (see [14], [15], [29], [31]):

(i) A module $P$ in Mod($R$) is pure-projective if and only if $P$ is a direct summand of a direct sum of finitely presented modules.

(ii) Every module $M$ in Mod($R$) admits a pure-projective pure resolution $P_\ast$ in Mod($R$), that is, there is a pure-exact sequence

$$\cdots \rightarrow P_n \xrightarrow{d_n} P_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \rightarrow 0$$

where the modules $P_0, \ldots, P_n, \ldots$ are pure-projective.

The main results of the paper are the following two theorems:
Theorem 1.2. Let $M$ be a right $R$-module and (1.1) a pure-exact sequence in $\text{Mod}(R)$ such that the modules $P_0, \ldots, P_n, \ldots$ are pure-projective. Then, for each $n \geq 0$, the $n$-th pure-syzygy module $\text{Ker} d_n$ of $M$ is an $\aleph_n$-directed union of $\aleph_n$-generated pure-projective $R$-modules, which are pure submodules of $P_n$ and of $\text{Ker} d_n$.

Theorem 1.3. Let $M$ be a right $R$-module. If there exists a pure-exact sequence

$$0 \to M \to P_n \to P_{n-1} \to \ldots \to P_1 \to P_0 \to M \to 0$$

in $\text{Mod}(R)$ such that the modules $P_0, \ldots, P_n$ are pure-projective, then $M$ is pure-projective.

In other words, every pure-periodic $R$-module $M$ is pure-projective. As a consequence of Theorem 1.3 we get the following version of a recent result by Benson and Goodearl in [5].

Corollary 1.4. Let $M$ be a right flat $R$-module. If there exists an exact flat sequence

$$0 \to M \to F_n \to F_{n-1} \to \ldots \to F_1 \to F_0 \to M \to 0$$

in $\text{Mod}(R)$ such that the modules $F_0, \ldots, F_n$ are projective, then $M$ is pro-jective.

Our Theorem 1.2 extends the main projective resolution structure theorem [23, Theorem 1.5] (see also [17, Theorem 3.3] for $n = 0$) from flat modules to arbitrary modules. The Corollary 1.4 with $n = 0$ coincides with [5, Theorem 2.5].

The main results of the paper are proved in Sections 3 and 4. In Section 2 we collect preliminary facts and notation we need throughout the paper. In Section 5 we show that Theorems 1.2 and 1.3 remain valid in any locally finitely presented Grothendieck category $\mathcal{A}$.

Throughout this paper we use freely the module theory terminology and notation introduced in [1] and [12]. The reader is referred to [8], [12], [14], [16], [26], [29], [31] and to the expository papers [9] and [28] for a basic background and historical comments on purity and pure homological dimensions.

2. Preliminaries on the pure-projective dimension.

We start this section by collecting basic definitions, notation and elementary facts we need throughout this paper.

Given right $R$-modules $M$ and $N$ the $n$-th pure extension group $\text{Pext}_R^n(M, N)$ is defined to be the $n$-th cohomology group of the complex $\text{Hom}_R(P_s, N)$, where $P_s$ is a pure-projective resolution of $M$ in $\text{Mod}(R)$.
The pure-projective dimension $P_{pd} M$ of $M$ is defined to be the minimal integer $m \geq 0$ (or infinity) such that $\text{Pext}_R^m(M, -) = 0$. The right pure global dimension $r.P_{gl}.\dim R$ of $R$ is defined to be the minimal integer $n \geq 0$ (or infinity) such that $\text{Pext}_R^n = 0$. Following [27] we call the ring $R$ right pure semisimple if $r.P_{gl}.\dim R = 0$.

The left pure global dimension $l.P_{gl}.\dim R$ of a ring $R$ was introduced in 1967 by R. Kielpinski [14] and in 1970 by P. Griffith [7]. It was shown in [26, Theorem 2.12] that the right pure global dimension $r.P_{gl}.\dim R$ of the ring $R$ is the supremum of $P_{pd} M$, where $M$ runs through all right $R$-modules $M$ such that $P_{pd} M$ is finite. This means that the right finitistic pure global dimension of $R$ and the right pure global dimension of $R$ coincide.

Throughout this paper we denote by $\aleph$ an infinite cardinal number and by $\aleph_0$ the cardinality of a countable set. A right $R$-module $M$ is said to be $\aleph$-generated if it is generated by a set of cardinality $\aleph$, and $M$ is $\aleph$-presented if $M$ is $\aleph$-generated and for any epimorphism $f : L \to M$ with $\aleph$-generated module $L$ the kernel $\text{Ker} f$ is $\aleph$-generated, or equivalently, $M$ is a limit of a direct system $\{M_j, h_{ij}\}$ of cardinality $\aleph$ consisting of finitely presented modules $M_j$ (see [18], [20], [22], [26]). We say that $M$ is an $\aleph$-directed union of submodules $M_j$, $j \in J$, if for each subset $J_0$ of $J$ of cardinality $\aleph$ there exist $j_0 \in J$ such that $M_t \subseteq M_{j_0}$ for all $t \in J_0$.

A union $\bigcup_{\xi < \gamma} M_\xi$ of submodules $M_\xi$ of $M$ is well-ordered and continuous if $\gamma$ is an ordinal number, $M_0 = (0)$, $M_\xi \subseteq M_\eta$ for $\xi < \eta < \gamma$, and $M_\tau = \bigcup_{\xi < \tau} M_\xi$ for any limit ordinal number $\tau \leq \gamma$ (see [16]).

The following pure version of the well-known Auslander result [2, Proposition 3] is of importance.

**Proposition 2.1.** Assume that the right $R$-module $M$ is a continuous well-ordered union of submodules $M_\xi$, with $\xi < \gamma$, where $\gamma$ is an ordinal number. If $P_{pd} M_{\xi+1}/M_\xi \leq m$ for all $\xi < \gamma$, then $P_{pd} M \leq m$.

**Proof.** The arguments of Auslander in the proof of [2, Proposition 3] generalise to our situation (see [16, Proposition 1.2]).

We also need the following pure version of the well-known Osofsky result [19] (see also [3]) proved in [16, Corollary 1.4], [8], and in [26, Theorem 2.12] in a general context of Grothendieck categories.

**Proposition 2.2.** Assume that $m \geq 0$ is an integer and $M$ is an arbitrary $\aleph_m$-presented right $R$-module. Then $P_{pd} M \leq m + 1$.

By applying the definition of a pure submodule one proves the following useful criterion:

**Lemma 2.3.** Assume that $P$ is a pure-projective right $R$-module and let $K$ be a submodule of $P$. The following conditions are equivalent:
(a) $K$ is a pure submodule of $P$.
(b) For any finitely generated submodule $X$ of $K$ there exists an $R$-homomorphism $\varphi : P \to K$ such that $\text{Im} \varphi$ is contained in a finitely generated $R$-submodule of $K$ and $\varphi|_X = \text{id}_X$.
(c) For any finitely generated submodule $X$ of $K$ there exists an $R$-homomorphism $\varphi : P \to K$ such that $\varphi|_X = \text{id}_X$.

Proof. Since the module $P$ is pure projective, there exists a module $P'$ such that $P \oplus P'$ is a direct sum of finitely presented modules. Assume that $K$ is a submodule of $P$ and let $u : K \to P$ be the embedding.

(a) $\Rightarrow$ (b) Assume that $u : K \to P$ is a pure monomorphism and $X$ is a finitely generated submodule of $K$. Then the monomorphism $(u, 0) : K \to P \oplus P'$ is pure and there exists a finitely presented direct summand $L$ of $P \oplus P'$ such that $(u, 0)(X) \subseteq L$. Consider the commutative diagram

\[
\begin{array}{c}
0 \to K \xrightarrow{(u,0)} P \oplus P' \xrightarrow{\pi} \text{Coker}(u,0) \to 0 \\
0 \xrightarrow{h'} \to X \xrightarrow{u'} \to L \xrightarrow{p} \to \text{Coker}(u,0) \to 0
\end{array}
\]

with exact rows, where $h'$ is the embedding of $X$ into $K$, $h$ is a direct summand embedding, $\pi$ is a pure epimorphism and the module $L$ is finitely presented. It follows that there exists $v'' \in \text{Hom}_R(L, P \oplus P')$ such that $\pi v'' = h''$, and consequently there exists $v' \in \text{Hom}_R(L, K)$ such that $v'u' = h'$. Let $\varphi' : P \oplus P' \to K$ be an extension of $v'$ to $P \oplus P'$ such that $v' = \varphi' h'$ and $\text{Im} \varphi'$ is finitely generated. Let $\varphi : P \to K$ be the restriction of $\varphi'$ to $P$. It follows that $\text{Im} \varphi$ is contained in the finitely generated $R$-submodule $\text{Im} \varphi'$ of $K$ and, for any $x \in X$, we have $x = h'(x) = v'u'(x) = \varphi'hu'(x) = \varphi'(h'(x), 0) = \varphi'(x, 0) = \varphi(x)$. This shows that $\varphi|_X = \text{id}_X$ and (b) follows.

The implication (b) $\Rightarrow$ (c) is obvious.

(c) $\Rightarrow$ (a) Assume that, for any finitely generated submodule $X$ of $K$, there exists an $R$-homomorphism $\varphi : P \to K$ such that $\varphi|_X = \text{id}_X$. We shall prove that $K$ is a pure submodule of $P$ by showing that the canonical epimorphism $\pi : P \to P/K$ is pure. Let $f : L \to P/K$ be a homomorphism from a finitely presented module $L$ to $P/K$. Then $L \cong F/N$, where $F$ is a finitely generated free module and $N$ is a finitely generated submodule of $F$. It is clear that there exists a commutative diagram

\[
\begin{array}{c}
0 \to K \xrightarrow{u} P \xrightarrow{\pi} P/K \to 0 \\
\uparrow{f''} \quad \quad \quad \quad \quad \quad \uparrow{f'} \quad \quad \quad \quad \quad \quad \uparrow{f} \\
0 \to N \xrightarrow{u'} F \xrightarrow{p} L \to 0
\end{array}
\]

with exact rows, where $p$ is the canonical epimorphism and $u'$ is the canonical embedding. Then $X = f''(N)$ is a finitely generated submodule of $K$ and,
according to our assumption, there exits an $R$-homomorphism $\varphi : P \to K$ such that $\varphi|_X = \text{id}_X$. Note that the homomorphism $v' = \varphi f' : F \to K$ satisfies the equality $f'' = v'u'$. It follows that there exists $v'' \in \text{Hom}_R(L, P)$ such that $\pi v'' = f$. This shows that $\pi$ is a pure epimorphism and finishes the proof of the lemma.

Let $P$ be a pure-projective right $R$-module and let $K$ be a pure submodule of $P$. Following [23, Proposition 1.4] and [24] we define a pure-closure $L^\Diamond$ of any $R$-submodule $L$ of $K$ as follows. Set $L_0 = L$ and fix a set $L'$ of generators of $L$. By Lemma 2.3, for any finite subset $\lambda$ of $L'$ we find an $R$-homomorphism $\varphi_\lambda : P \to K$ such that $\text{Im} \varphi_\lambda$ is contained in a finitely generated $R$-submodule $K_\lambda$ of $K$, and $\varphi_\lambda|_X = \text{id}_X$. Let $L_1$ be the $R$-submodule of $L$ generated by the set $L'' = \bigcup_{\lambda \subseteq L'} K_\lambda$, where $\lambda$ runs over all finite subsets of $L'$. It is clear that $L = L_0 \subseteq L_1$ and, for any finitely generated submodule $X$ of $L_0 = L$, there exists an $R$-homomorphism $\varphi : P \to L_1$ such that $\text{Im} \varphi$ is contained in a finitely generated $R$-submodule of $L_1$, and $\varphi|_X = \text{id}_X$. By choosing a set $L'_1$ of generators of $L_1$ and applying the procedure above with $L'$ and $L'_1$ interchanged, we construct a submodule $L_2$ containing $L_1$ such that for any finitely generated submodule $X$ of $L_1$ there exists an $R$-homomorphism $\varphi : P \to L_2$ such that $\text{Im} \varphi$ is contained in a finitely generated $R$-submodule of $L_2$, and $\varphi|_X = \text{id}_X$. Continuing this way we define an ascending sequence

$$L = L_0 \subseteq L_1 \subseteq L_2 \subseteq \cdots \subseteq L_m \subseteq L_{m+1} \subseteq \cdots$$

of $R$-submodules of $K$, and sets $L'_0, L'_1, L'_2, \ldots L'_m, L'_{m+1}, \ldots$ of their generators in such a way that, for each $m \geq 0$ and for any finitely generated submodule $X$ of $L_m$, there exists an $R$-homomorphism $\varphi : P \to L_{m+1}$ such that $\text{Im} \varphi$ is contained in a finitely generated $R$-submodule of $L_{m+1}$, and $\varphi|_X = \text{id}_X$. By Lemma 2.3, the submodule

$$L^\Diamond = \bigcup_{m=0}^{\infty} L_m$$

of $K$ is a pure submodule of $P$ (and of $K$), and we call it a pure-closure of the $R$-submodule $L$ of $K$. It is clear that $L^\Diamond$ is not determined uniquely by $L$ and depends on the choice of the modules $K_\lambda$, the sets $L'_0, L'_1, L'_2, \ldots L'_m, L'_{m+1}, \ldots$ and the $R$-homomorphisms $\varphi_\lambda : P \to K$. However, if $\aleph$ is an infinite cardinal number and the module $L$ is $\aleph$-generated then the sets $L'_0, L'_1, L'_2, \ldots L'_m, L'_{m+1}, \ldots$ can be chosen of cardinality $\aleph$ and we get the following result.

**Lemma 2.5.** Assume that $P$ is a pure-projective right $R$-module, $K$ a pure submodule of $P$ and $L$ an $\aleph$-generated submodule of $K$, where $\aleph$ is an infinite
cardinal number. Then there exists an $\aleph$-generated submodule $L^\Diamond$ of $K$ such that $L \subseteq L^\Diamond$ and $L^\Diamond$ is a pure submodule of $P$ (and of $K$).

We also need the following technical result.

**Lemma 2.6.** Assume that $\aleph$ is an infinite cardinal number, $h : P \rightarrow K$ is a pure epimorphism in $\text{Mod}(R)$, $P$ is an $\aleph$-generated pure-projective module and $K$ is a pure submodule of a pure-projective module.

(a) The module $K$ has a directed union form $K = \bigcup_{\lambda \in \Omega} K_\lambda$, where $\Omega$ is a set of cardinality $\leq \aleph$ and $K_\lambda$ is a countably generated pure-projective pure submodule of $K$, for each $\lambda \in \Omega$.

(b) The module $\text{Ker } h$ is $\aleph$-generated.

**Proof.** Let $h : P \rightarrow K$ be a pure-epimorphism. We set $L = \text{Ker } h$ and assume that the module $P$ is $\aleph$-generated. Then there exist a set $\Omega$ of cardinality $\leq \aleph$ and a family of finitely generated submodules $P_\lambda$ of $P$, with $\lambda \in \Omega$, such that $P = \bigcup_{\lambda \in \Omega} P_\lambda$ is a directed union. By our assumption, $K$ is a pure submodule of a pure-projective module $P_0$. Let $P'_0$ be a right $R$-module such that $P_0 \oplus P'_0$ is a direct sum of finitely presented modules.

For each $\lambda \in \Omega$, we consider the commutative diagram

$$
\begin{array}{ccc}
0 & \rightarrow & L \cap P_\lambda \\
\downarrow u'_\lambda & & \downarrow u'_\lambda \\
0 & \rightarrow & L
\end{array}
\quad
\begin{array}{ccc}
P_\lambda & \xrightarrow{g_\lambda} & P_\lambda \\
\downarrow r_\lambda & & \downarrow r_\lambda \\
0 & \rightarrow & P
\end{array}
\quad
\begin{array}{ccc}
P \xrightarrow{h} & & K \\
\downarrow & & \downarrow \\
0 & \rightarrow & 0
\end{array}
$$

with exact rows, where $\overline{P}_\lambda = P_\lambda/L \cap P_\lambda$, $u_\lambda$, $u'_\lambda$, $u''_\lambda$, $v$ are the embeddings and $r_\lambda$ is the natural $R$-module homomorphism induced by $u''_\lambda$. Since $V_\lambda = \text{Im } r_\lambda = h(P_\lambda)$ is a finitely generated submodule of $K$ and $K$ is a pure submodule of $P_0 \oplus P'_0$ then $V_\lambda$ is $\aleph_0$-generated and, according to Lemma 2.5, there exists an $\aleph_0$-generated pure submodule $V^\Diamond_\lambda$ of $P_0 \oplus P'_0$ contained in $K$ and containing $V_\lambda$. It follows that $V^\Diamond_\lambda$ is a pure submodule of an $\aleph_0$-generated direct summand $P'$ of $P_0 \oplus P'_0$. Then the module $P'/V^\Diamond_\lambda$ is $\aleph_0$-presented and Proposition 2.2 yields (see also [10])

$$
P \cdot \text{pd } P'/V^\Diamond_\lambda \leq 1.
$$

It follows that the submodule $V^\Diamond_\lambda$ of $K$ is pure-projective. If we set $K_\lambda = V^\Diamond_\lambda$, then obviously $K = \bigcup_{\lambda \in \Omega} K_\lambda$ is a directed union and $K_\lambda$ is a countably generated pure-projective pure submodule of $K$ for each $\lambda \in \Omega$. This proves Statement (a).

Since the epimorphism $h : P \rightarrow K$ is pure, the embedding $w_\lambda : V^\Diamond_\lambda \hookrightarrow K$ extends to an $R$-module homomorphism $f_\lambda : V^\Diamond_\lambda \rightarrow P$ such that $hf_\lambda =$
the proof. Because \((\ast)\) there exists an \(R\)-module homomorphism \(\varphi_\lambda : P_\lambda \to L\) such that \(\varphi_\lambda u_\lambda = u'_\lambda\). Hence we easily conclude that \(L = \sum_{\lambda \in \Omega} \text{Im} \varphi_\lambda\) and therefore \(L\) is \(\aleph\)-generated, because \(|\Omega| \leq \aleph\) and \(\text{Im} \varphi_\lambda\) is finitely generated for any \(\lambda \in \Omega\). This finishes the proof. \(\square\)


The aim of this section is to prove Theorem 1.2 on the pure-projective structure of the \(n\)-th pure-syzygy module of any right \(R\)-module \(M\), that is, the pure submodule \(\text{Ker} \, d_n\) of \(P_n\) in a pure-projective resolution \((\ast\ast)\) of \(M\).

We start with the following key proposition.

Proposition 3.1. Assume that \(R\) is a ring, \(\aleph\) is an infinite cardinal number, \(M\) is a right \(R\)-module, \(n \geq 0\) an integer and

\[(\ast) \quad 0 \to K_n \to P_n \xrightarrow{d_n} P_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \to 0\]

is a pure-exact sequence, where \(K_n = \text{Ker} \, d_n\) and the modules \(P_0, \ldots, P_n\) are pure-projective.

(a) For any \(\aleph\)-generated submodule \(N\) of \(K_n\) and any \(\aleph\)-generated submodule \(L\) of \(K_0 = \text{Ker} \, d_0\) there exist an \(\aleph\)-generated pure submodule \(N^{\bullet_n}\) of \(P_n\), and \(\aleph\)-generated pure submodule \(L^{\bullet_0}\) of \(P_0\), an \(\aleph\)-generated direct summands \(P'_1, \ldots, P'_n\) of \(P_1, \ldots, P_n\), respectively, such that \(d_j(P'_j) \subseteq P'_{j-1}\) for \(j = 1, \ldots, n\), \(N \subseteq N^{\bullet_n} \subseteq K_n = \text{Ker} \, d_n\), \(L \subseteq L^{\bullet_0} \subseteq K_0 = \text{Ker} \, d_0\), and, for each \(n \geq 1\), the sequence

\[(\ast\ast) \quad 0 \to N^{\bullet_n} \to P'_n \xrightarrow{d'_n} P'_{n-1} \xrightarrow{d'_{n-1}} \cdots \xrightarrow{d'_2} P'_1 \xrightarrow{d'_1} P' \xrightarrow{d'_0} L^{\bullet_0} \to 0\]

is pure-exact, where \(d'_j\) is the restriction of \(d_j\) to \(P'_j\). In case \(n = 0\) we have \(N^{\bullet_n} = L^{\bullet_0}\).

(b) If \(n \geq 1\) and there is an \(R\)-module isomorphism \(K_n \cong K_0\), then there exists a pure-exact sequence \((\ast\ast)\) satisfying the conditions listed in (a) and such that \(N^{\bullet_n} \cong L^{\bullet_0}\).

Proof. (a) Since any pure-projective module is a direct summand of a direct sum of finitely presented modules then, according to the well-known Kaplansky theorem [13], there are pairwise disjoint sets \(I_0, I_1, \ldots, I_n\) and countably generated pure-projective modules \(Q_t\), with \(t \in I_0 \cup I_1 \cup \cdots \cup I_n\), such that, for each for \(j \in \{0, 1, \ldots, n\}\), the pure-projective module \(P_j\) in \((\ast)\) has the form

\[P(I_j) = \bigoplus_{t \in I_j} Q_t,\]

up to isomorphism. Without loss of generality we can suppose that \(P_j = P(I_j)\) for \(j = 0, 1, \ldots, n\).
Instead of the statement (a) we prove inductively on \( n \geq 0 \), like [23, Proposition 1.4], the following stronger form of (a).

Assume that, for each \( j \in \{0, 1, \ldots, n\} \), the pure-projective module \( P_j \) in (*) has the form \( P(I_j) \) as above. Then the following two statements hold.

(a1) For any \( \aleph \)-generated submodule \( N \) of \( K_n \) and any \( \aleph \)-generated submodule \( L \) of \( K_0 = \text{Ker} \ d_0 \) there exist an \( \aleph \)-generated pure submodule \( N' \) of \( P_n = P(I_n) \), \( \aleph \)-generated pure submodule \( L' \) of \( P_0 = P(I_0) \), and subsets \( I'_0, \ldots, I'_n \) of \( I_0, \ldots, I_n \), respectively, of cardinality \( \leq \aleph \) such that \( d_j(P(I'_j)) \subseteq P(I'_{j-1}) \) for \( j = 1, \ldots, n \), \( N' \subseteq N' \subseteq K_n = \text{Ker} \ d_n \), \( L \subseteq L' \subseteq K_0 = \text{Ker} \ d_0 \), and, for each \( n \geq 1 \), the sequence (**) is pure-exact, where \( P'_j = P(I'_j) \) and \( d'_j \) is the restriction of \( d_j \) to \( P'_j \). In case \( n = 0 \) we have \( N' = L' \).

(a2) Assume that \( N, L, N', L', I'_0, \ldots, I'_n \) are such that the statement (a1) holds, and let \( N' \) and \( L' \) be \( \aleph \)-generated submodules of \( K_n \) and \( K_0 \) containing \( N \) and \( L \), respectively. Then there exist an \( \aleph \)-generated pure submodule \( N'' \) of \( P_n = P(I_n) \), \( \aleph \)-generated pure submodule \( L'' \) of \( P_0 = P(I_0) \) and subsets \( I''_0, \ldots, I''_n \) of \( I_0, \ldots, I_n \), respectively, of cardinality \( \leq \aleph \) such that \( d_j(P(I''_j)) \subseteq P(I''_{j-1}) \) for \( j = 1, \ldots, n \), \( N'' \subseteq N'' \subseteq K_n \), \( L' \subseteq L'' \subseteq K_0 \), \( N'' \subseteq N'' \), \( L'' \subseteq L'' \), the diagram

\[
\begin{array}{ccc}
0 \rightarrow N'' & \xrightarrow{d'_n} & P(I''_{n-1}) \rightarrow \cdots \rightarrow d'_1 \rightarrow L'' \rightarrow 0 \\
\downarrow & & \downarrow \\
0 \rightarrow N'' & \xrightarrow{d''_n} & P(I''_{n-1}) \rightarrow \cdots \rightarrow d''_1 \rightarrow L'' \rightarrow 0 \\
\end{array}
\]

is commutative and has pure-exact rows, where the vertical arrows are natural embeddings induced by the inclusions \( I'_0 \subseteq I''_0, \ldots, I'_n \subseteq I''_n \) and \( d'_j \) is the restriction of \( d_j \) to \( P(I''_j) \) for \( j = 1, \ldots, n \). In case \( n = 0 \) we have \( N'' = L'' \).

Assume that \( n = 0 \). Since the submodules \( N \) and \( L \) of \( K_0 \) are \( \aleph \)-generated then applying Lemma 2.5 to the \( \aleph \)-generated submodule \( N + L \) of \( K_0 \) we get an \( \aleph \)-generated pure submodule \( (N + L) \) of \( P(I_0) \) and of \( K_0 \). It follows that there is a subset \( I'_0 \) of \( I_0 \) of cardinality \( \leq \aleph \) such that \( (N + L) \) is a pure submodule of \( P(I'_0) \) of \( P(I_0) \). If we set \( N' = L' = (N + L) \) we get (a1). Statement (a2) follows in a similar way.

Assume that \( n \geq 1 \). First we prove the following:

**Claim.** For any \( \aleph \)-generated submodule \( Y \) of \( K_{n-1} \) and any subset \( \mathcal{Y} \) of \( I_n \) of cardinality \( \leq \aleph \) there exists an \( \aleph \)-generated pure submodule \( Y' \) of \( K_{n-1} \)
containing $Y$ and a subset $\mathcal{Y}'$ of $I_n$ of cardinality $\leq \aleph$ containing $\mathcal{Y}$ such that:

(c1) $d_n(P(\mathcal{Y}')) = Y'$,
(c2) the restriction $d_n' : P(\mathcal{Y}') \rightarrow Y'$ of $d_n$ to $P(\mathcal{Y}')$ is a pure epimorphism, and
(c3) the submodule $\ker d_n'$ of $P(\mathcal{Y}')$ is $\aleph$-generated.

Let $Y$ be an $\aleph$-generated submodule of $K_{n-1}$ and $\mathcal{Y}$ a subset of $I_n$ of cardinality $\leq \aleph$. We construct the subset $\mathcal{Y}'$ of $I_n$ as the union $\mathcal{Y}' = \bigcup_{j=1}^{\infty} \mathcal{Y}_j$ of subsets

($+$) $\mathcal{Y} \subseteq \mathcal{Y}_1 \subseteq \mathcal{Y}_2 \subseteq \cdots \subseteq \mathcal{Y}_j \subseteq \mathcal{Y}_{j+1} \subseteq \cdots$

of $I_n$ of cardinality $\leq \aleph$, and the module $Y'$ as the union $Y' = \bigcup_{j=1}^{\infty} Y^{(j)}$ of $\aleph$-generated pure submodules

($++$) $Y \subseteq Y^{(1)} \subseteq Y^{(2)} \subseteq \cdots \subseteq Y^{(j)} \subseteq Y^{(j+1)} \subseteq \cdots$

of $K_{n-1}$ such that the image of the restriction $d^{(j)} : P(\mathcal{Y}^{(j)}) \rightarrow K_{n-1}$ of $d_n$ to $P(\mathcal{Y}^{(j)})$ contains $Y^{(j)}$ and it is contained in $Y^{(j+1)}$ for $j \geq 1$, and for any finitely generated $R$ module $Z$ and any $R$-homomorphism $f : Z \rightarrow Y^{(j)}$ there exists an $R$-homomorphism $f' : Z \rightarrow P(\mathcal{Y}^{(j)})$ such that $f = d^{(j)} f'$.

It is clear that the above properties imply Conditions (c1) and (c2) of claim. In view of Lemma 2.6, Condition (c3) is a consequence of (c2), because $Y'$ is an $\aleph$-generated pure submodule of $K_{n-1}$.

We construct the sequences ($+$) and ($++$) inductively as follows. By applying Lemma 2.5 to the pure submodule $K = K_{n-1}$ of the pure-projective module $P = P(I_{n-1})$ and $L = Y$ we get an $\aleph$-generated pure submodule $Y^{\varnothing}$ of $K_{n-1}$ containing $Y$. We set $Y^{(1)} = Y^{\varnothing}$. By Lemma 2.6, the module $Y^{(1)}$ has a directed union form $Y^{(1)} = \bigcup_{\lambda \in \Omega_1} Y^{(1)}_{\lambda}$, where $\Omega_1$ is a set of cardinality $\leq \aleph$ and $Y^{(1)}_{\lambda}$ is a countably generated pure-projective pure submodule of $K_{n-1}$ for each $\lambda \in \Omega_1$. Since the epimorphism $d_n : P(I_n) \rightarrow K_{n-1}$ is pure and $Y^{(1)}_{\lambda}$ is pure-projective, then for each $\lambda \in \Omega_1$ the embedding $v_{\lambda} : Y^{(1)}_{\lambda} \rightarrow Y^{(1)}$ has a factorisation $v_{\lambda} = d_n f_{\lambda}$, where $f_{\lambda} \in \operatorname{Hom}_R(Y^{(1)}_{\lambda}, P(I_n))$. Since $f_{\lambda}(Y^{(1)}_{\lambda})$ is a countably generated submodule of $P(I_n)$, $|\Omega_1| \leq \aleph$ and $\aleph \geq \aleph_0$, then there exists a subset $\mathcal{Y}^{(1)}$ of $I_n$ of cardinality $\leq \aleph$ containing $Y$ such that $\sum_{\lambda \in \Omega_1} f_{\lambda}(Y^{(1)}_{\lambda}) \subseteq P(\mathcal{Y}^{(1)})$. It follows that the image of the restriction $d^{(1)} : P(\mathcal{Y}^{(1)}) \rightarrow K_{n-1}$ of $d_n$ to $P(\mathcal{Y}^{(1)})$ contains $Y^{(1)} \supseteq Y$. Moreover, for any finitely generated $R$ module $Z$ and any $R$-homomorphism $f : Z \rightarrow Y^{(1)}$ there exists an $R$-homomorphism $f' : Z \rightarrow P(\mathcal{Y}^{(1)})$ such
that \( f = d^{(1)} f' \). Indeed, \( \text{Im} \, f \) is a finitely generated submodule of \( Y^{(1)} \) and therefore there exists \( \lambda \in \Omega_1 \) such that \( \text{Im} \, f \subseteq Y^{(1)}_{\lambda} \). If we set \( f' = f_\lambda f \), we get the required equality \( f = d^{(1)} f' \). Hence we conclude \( Y^{(1)} \subseteq \text{Im} \, d^{(1)} \).

Since \( |Y^{(1)}| \leq \aleph \), the submodule \( \text{Im} \, d^{(1)} \) of \( K_{n-1} \) is \( \aleph \)-generated, and according to Lemma 2.5 there exists an \( \aleph \)-generated pure submodule \((\text{Im} \, d^{(1)})^\diamond\) of \( K_{n-1} \) containing \( \text{Im} \, d^{(1)} \). We set \( Y^{(2)} = (\text{Im} \, d^{(1)})^\diamond \).

If \( j \geq 1 \) and \( Y^{(j)} \), \( Y^{(j)} \) are constructed, we construct \( Y^{(j+1)} \) by applying the above construction of \( Y^{(1)} \), \( Y^{(1)} \) and \( Y^{(2)} \) to \( Y^{(j)} \) and the set \( Y^{(j)} \). The details are left to the reader. This finishes the proof of claim.

Now we prove the inductive step. Assume that \( n \geq 1 \) and that Statements (a1) and (a2) hold for \( n - 1 \). In order to prove (a1) and (a2) for \( n \), we assume that \( N \) is an \( \aleph \)-generated submodule of \( K_n \) and \( L \) is an \( \aleph \)-generated submodule of \( K_0 \). We set \( L_0 = L \). By Lemma 2.5, there exists an \( \aleph \)-generated pure submodule \( N_0^{\diamond n} \) of \( P(I_n) \) such that \( N \subseteq N_0^{\diamond n} \subseteq K_n \).

Let \( J_{n,0} \) be a subset of \( I_n \) of cardinality \( \leq \aleph \) such that \( N_0^{\diamond n} \subseteq P(J_{n,0}) \subseteq P(I_n) \). Then the submodule \( T_0 = d_n(P(J_{n,0}')) \) of \( K_{n-1} = \text{Ker} \, d_{n-1} \subseteq P(I_{n-1}) \) is \( \aleph \)-generated. By applying the induction hypothesis to \( T_0 \subseteq K_{n-1} \) and \( L_0 = L \subseteq K_0 \) one gets subsets \( J_{n-1,0} \subseteq I_{n-1} \), \( J_{0,0} \subseteq I_0 \) of cardinality \( \leq \aleph \), an \( \aleph \)-generated pure submodule \( T_0^{\diamond n-1} \subseteq K_{n-1} \) containing \( T_0^{\diamond n-1} \) and \( P(J_{n-1,0}) \) containing \( T_0 \), an \( \aleph \)-generated pure submodule \( L_0^{\diamond 0} \subseteq K_0 \) of \( P(J_{0,0}) \) containing \( L_0 \) such that the sequence

\[
0 \to T_0^{\diamond n-1} \to P(J_{n-1,0}) \xrightarrow{d_{n-1,0}} P(J_{n-2,0}) \to \cdots \to P(J_{1,0}) \xrightarrow{d_{1,0}} L_0^{\diamond 0} \to 0
\]

is pure-exact, where \( d_{j,0} \) is the restriction of \( d_j \) to \( P(J_{j,0}) \) for \( j = 1, \ldots, n-1 \).

By our claim applied to \( Y = T_0^{\diamond n-1} \) and \( Y = J_{n,0} \), there exist a subset \( J_{n,0} \) of \( I_n \) of cardinality \( \leq \aleph \) containing \( J_{n,0} \) and an \( \aleph \)-generated pure submodule \( T_1 = (T_0^{\diamond n-1})' \) of \( K_{n-1} \) containing \( T_0^{\diamond n-1} \) such that \( J_{n,0} \subseteq J_{n,0} \), the restriction of \( d_n \) to \( P(J_{n,0}) \) yields a pure epimorphism

\[
d_{n,0} : P(J_{n,0}) \to T_1
\]

and the pure submodule \( \text{Ker} \, d_{n,0} \) of \( P(J_{n,0}) \) is \( \aleph \)-generated. It is clear that \( N \subseteq N_0^{\diamond n} \subseteq \text{Ker} \, d_{n,0} \).

By applying the induction hypothesis to \( T_1 \subseteq K_{n-1} \) and \( L_1 = L_0^{\diamond 0} \subseteq K_0 \), one gets subsets \( J_{n-1,1} \subseteq I_{n-1} \), \( J_{0,1} \subseteq I_0 \) of cardinality \( \leq \aleph \), an \( \aleph \)-generated pure submodule \( T_1^{\diamond n-1} \subseteq K_{n-1} \) containing \( T_1 \), an \( \aleph \)-generated pure submodule \( L_1^{\diamond 0} \subseteq K_0 \) of \( P(J_{0,1}) \) containing \( L_1 \) such that
the sequence
\[ 0 \to T_1^\otimes n^{-1} \to P(J_{n-1,1}) \to \cdots \]
\[ \quad \to P(J_{1,1}) \to L_1^\otimes 0 \to 0 \]

is pure-exact, where \( d_j \) is the restriction of \( d_j \) to \( P(J_{j,1}) \) and \( J_{j,0} \subseteq J_{j,1} \subseteq I_j \) for \( j = 1, \ldots, n-1 \).

By our claim applied to \( Y = T_1^\otimes n^{-1} \) and \( Y = J_{n,0} \), there exist a subset \( J_{n,1} \) of \( I_n \) of cardinality \( \leq \aleph \) containing \( J_{n,0} \) and an \( \aleph \)-generated pure submodule \( T_2 = (T_1^\otimes n^{-1})' \) of \( K_{n-1} \) containing \( T_1^\otimes n^{-1} \) such that the restriction of \( d_n \) to \( P(J_{n,1}) \) yields a pure epimorphism
\[ d_{n,1} : P(J_{n,1}) \to T_2, \]

the submodule \( \text{Ker } d_{n,1} \) of \( P(J_{n,1}) \) is \( \aleph \)-generated and \( N \subseteq N_0^\otimes n \subseteq \text{Ker } d_{n,0} \subseteq \text{Ker } d_{n,1} \).

Continuing this way, we construct two sequences

- \( T_0 \subseteq T_n^\otimes n^{-1} \subseteq T_1^\otimes n^{-1} \subseteq \cdots \subseteq T_s^\otimes n^{-1} \subseteq \cdots, \)
- \( L = L_0 \subseteq L_1 = L_0^\otimes 0 \subseteq L_2 = L_1^\otimes 0 \subseteq \cdots \subseteq L_s = L_{s-1}^\otimes 0 \subseteq \cdots \)

of \( \aleph \)-generated submodules of \( K_{n-1} \subseteq P(I_{n-1}) \) and \( K_0 \subseteq P(I_0) \), respectively, and, for each \( j \in \{1, \ldots, n\} \), a chain
\[ J_{j,0} \subseteq J_{j,1} \subseteq J_{j,2} \subseteq \cdots \subseteq J_{j,s} \subseteq J_{j,s+1} \subseteq \cdots \]

of subsets \( J_{j,s} \) of \( I_j \) such that \( |J_{j,s}| \leq \aleph \), \( T_s^\otimes n^{-1} \subseteq P(J_{n-1,s}) \) and \( L_s \subseteq P(J_{0,s-1}) \) are pure embeddings and the restriction of \( d_n \) to \( P(J_{n,s}) \) yields a pure epimorphism
\[ d_{n,s} : P(J_{n,s}) \to T_{s+1}. \]

It follows that, for each \( j \in \{1, \ldots, n\} \), there is a chain
\[ P(J_{j,0}) \subseteq P(J_{j,1}) \subseteq P(J_{j,2}) \subseteq \cdots \subseteq P(J_{j,s}) \subseteq P(J_{j,s+1}) \subseteq \cdots \]

of submodules \( P(J_{j,s}) \) of \( P(I_j) \), and we get an infinite commutative diagram
with pure-exact rows, where the vertical homomorphisms are the $R$-module embeddings constructed above. Let

$$0 \to N^\bullet = \bigcup_{s=1}^{\infty} \text{Ker } d_{n,s}, \quad L^\bullet = \bigcup_{s=1}^{\infty} L_s \quad \text{and} \quad I'_j = \bigcup_{s=0}^{\infty} J_{j,s}$$

for $j = 1, \ldots, n$. It follows that the limit sequence is pure-exact, consists of $\aleph$-generated modules, $N^\bullet = \text{Ker } d'_n$ is a pure submodule of $P(I'_n)$ (and of $K_n$) containing $N$, the module

$$\text{Im } d'_n = \bigcup_{s=1}^{\infty} T_s = \bigcup_{s=1}^{\infty} T_s^{\diamondsuit s-1} = \text{Ker } d'_{n-1}$$

is a pure submodule of $K_{n-1}$ and $L^\bullet = \bigcup_{s=1}^{\infty} L_s$ is a pure submodule of $P(I'_0)$ as well as of $K_0$. By Lemma 2.6, the module $N^\bullet = \text{Ker } d'_n$ is $\aleph$-generated.
This finishes the proof of (a1). The method used in the inductive step of (a1) above also proves the inductive step of (a2). We leave it to the reader. This finishes the proof of (a).

(b) Assume that \( n \geq 1 \) and \( K_n \cong K_0 \). Let \( N \) be an \( R \)-generated submodule of \( K_n \) and \( L \) an \( R \)-generated submodule of \( K_0 \). Fix an \( R \)-module isomorphism \( f : K_n \to K_0 \). Keeping the notation above and by applying (a), we construct inductively an infinite commutative diagram

\[
\begin{array}{c}
0 \to N_1^\bullet \to P(I_{n,1}') \xrightarrow{d_{n,1}} P(I_{n-1,1}') \xrightarrow{d_{n-1,1}} \ldots \to P(I_{1,1}') \xrightarrow{d_{1,1}} L_1^\bullet \to 0 \\
\downarrow \quad \quad \downarrow \quad \quad \downarrow \quad \quad \downarrow \quad \quad \downarrow \\
0 \to N_2^\bullet \to P(I_{n,2}') \xrightarrow{d_{n,2}} P(I_{n-1,2}') \xrightarrow{d_{n-1,2}} \ldots \to P(I_{1,2}') \xrightarrow{d_{1,2}} L_2^\bullet \to 0 \\
\downarrow \quad \quad \downarrow \quad \quad \downarrow \quad \quad \downarrow \quad \quad \downarrow \\
0 \to N_3^\bullet \to P(I_{n,3}') \xrightarrow{d_{n,3}} P(I_{n-1,3}') \xrightarrow{d_{n-1,3}} \ldots \to P(I_{1,3}') \xrightarrow{d_{1,3}} L_3^\bullet \to 0 \\
\vdots \quad \quad \vdots \quad \quad \vdots \quad \quad \vdots \quad \quad \vdots
\end{array}
\]

with pure-exact rows, where the vertical homomorphisms are \( R \)-module embeddings induced by the inclusions \( I_{j,1}' \subseteq I_{j,2}' \subseteq I_{j,3}' \subseteq \ldots \), for \( j = 1, \ldots, n \). We set \( N_1 = N + \cdot^{-1}(L) \) and \( L_1 = f(N) + L \). If the modules \( N_j, L_j \) and \( N_j^\bullet, L_j^\bullet \) are defined we set

\[
N_{j+1} = N_j^\bullet + f^{-1}(L_j^\bullet) \quad \text{and} \quad L_{j+1} = f(N_j^\bullet) + L_j^\bullet.
\]

It is clear that \( N_j \subseteq N_j^\bullet \subseteq N_{j+1} \), \( L_j \subseteq L_j^\bullet \subseteq L_{j+1} \), \( f(N_1) = L_1 \) and, for each \( j \geq 1 \), we get \( f(N_{j+1}) = L_{j+1} \). Let

\[
0 \to N^\bullet \to P(I_{n}') \xrightarrow{d_n} P(I_{n-1}') \xrightarrow{d_{n-1}} \ldots \xrightarrow{d_2} P(I_1') \xrightarrow{d_1} L^\bullet \to 0
\]

be the direct limit of the above system of pure-exact sequences, where

\[
N^\bullet = \bigcup_{s=1}^{\infty} N_s^\bullet = \bigcup_{s=1}^{\infty} N_s, \quad L^\bullet = \bigcup_{s=1}^{\infty} L_s^\bullet = \bigcup_{s=1}^{\infty} L_s \quad \text{and} \quad I_j' = \bigcup_{s=0}^{\infty} I_{j,s}'
\]

for \( j = 1, \ldots, n \). It is easy to see that \( f(N^\bullet) = L^\bullet \). Thus the modules \( N^\bullet, L^\bullet \) are isomorphic and the statement (b) follows. This finishes the proof. \( \square \)
The claim proved in the above proof yields the following useful result.

**Corollary 3.2.** Assume that $\aleph$ is an infinite cardinal number, $h : P \to K$ is a pure epimorphism in $\text{Mod}(R)$, $P$ is an $\aleph$-generated pure-projective module and $K$ is a pure submodule of a pure-projective module. For any $\aleph$-generated submodule $Y$ of $K$ and any subset $X$ of $P$ of cardinality $\leq \aleph$ there exist an $\aleph$-generated direct summand $P'$ of $P$ containing $X$ and an $\aleph$-generated pure submodule $Y'$ of $K$ containing $Y$ such that $h(P') = Y'$, the restriction $h' : P' \to Y'$ of $h$ to $P'$ is a pure epimorphism and the module $\text{Ker} h'$ is $\aleph$-generated.

**Proof.** Let $h : P \to K$ be a pure-epimorphism. By our assumption, $K$ is a pure submodule of a pure-projective module $P_0$. We apply the claim in the proof above to $n = 1$, $P_1 = P$, $d_1 = h$, $M = P_0/K$ and $d_0 : P_0 \to M$ the canonical epimorphism. By Kaplansky theorem [13], the module $P$ has the form $P_1 = P(I_1) = \bigoplus_{t \in I_1} Q_t$, where $Q_t$ is a countably generated pure-projective module for $t \in I_1$, in the notation introduced above. Since $X$ is a subset of $P$ of cardinality $\leq \aleph$ there exist a subset $Y$ of $I_1$ of cardinality $\leq \aleph$ containing $X$. Then the corollary is an immediate consequence of Statements (c1), (c2) and (c3) of the claim. \hfill \Box

One of the main results of this paper is the following theorem describing a pure-projective structure of pure-syzygy modules of any $R$-module (compare with [23, Theorem 1.5]).

**Theorem 3.3.** Let $R$ be a ring, $M$ a right $R$-module and

$$
\cdots \to P_n \xrightarrow{d_n} P_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \to 0
$$

a pure-exact sequence in $\text{Mod}(R)$ such that the modules $P_0, \ldots, P_n, \ldots$ are pure-projective. Then, for each $n \geq 0$, the $n$th pure-syzygy module $\text{Ker} d_n$ of $M$ is an $\aleph_n$-directed union of $\aleph_n$-generated pure-projective pure $R$-submodules of $\text{Ker} d_n$, which are pure submodules of $P_n$.

**Proof.** Fix $n \geq 0$ and consider the $n$th pure-syzygy submodule $K_n = \text{Ker} d_n$ of $P_n$. Set $K_0 = \text{Ker} d_0$. It is sufficient to show that any $\aleph_n$-generated submodule $N$ of $K_n$ is a submodule of an $\aleph_n$-generated pure-projective submodule of $K_n$, which is pure in $K_n$. Let $N$ be an $\aleph_n$-generated submodule of $K_n$. By applying Proposition 3.1 to $N \subseteq K_n$ and to the submodule $L = (0)$ of $K_0$, one gets a pure-exact sequence

$$
0 \to N^{\bullet_n} \longrightarrow P' \xrightarrow{d'_n} P'_{n-1} \xrightarrow{d'_{n-1}} \cdots \xrightarrow{d'_2} P'_1 \xrightarrow{d'_1} L^{\bullet_0} \to 0
$$

consisting of $\aleph_n$-generated modules, where $P'_n, \ldots, P'_1$ are pure-projective modules, $N^{\bullet_0}$ is a pure submodule of $K_n$, $L^{\bullet_0}$ is a pure submodule of $P_0$ and $N \subseteq N^{\bullet_n}$. In case $n = 0$ we get just $N^{\bullet_0} = L^{\bullet_0}$.
Let $P''_0$ be a right $R$-module such that the module $P' = P_0 \oplus P''_0$ is a direct sum of finitely presented modules. Since $L^\oplus_0$ is a pure submodule of $P_0$, it is a pure submodule of $P'$ and therefore $L^\oplus_0$ is a pure submodule of an $\aleph_n$-generated pure-projective direct summand $P$ of $P_0$. Then the module $P/L^\oplus_0$ is $\aleph_n$-presented and Proposition 2.2 yields

$$p.d. \frac{P}{P/L^\oplus_0} \leq n + 1.$$ 

It follows that the submodule $N^\bullet_n$ of $P'_n$ in the pure-exact sequence (3.4) is pure-projective, because $P'_n, \ldots, P'_1$ are pure-projective modules. This finishes the proof. □

As a consequence of Theorem 3.3 we get the following structure theorem on syzygy modules of flat modules proved by the author in [23, Theorem 1.5].

**Corollary 3.5.** Let $R$ be a ring, $M$ a right flat $R$-module and $n \geq 0$ an integer. If

(3.6) \hspace{1cm} 0 \rightarrow K_n \rightarrow F_n \rightarrow F_{n-1} \rightarrow \ldots \rightarrow F_1 \rightarrow F_0 \rightarrow M \rightarrow 0

is an exact sequence in $\text{Mod}(R)$ such that the modules $F_0, \ldots, F_n$ are projective, then the $n$-th syzygy module $K_n$ of $M$ is an $\aleph_n$-directed union of $\aleph_n$-generated projective pure $R$-submodules of $K_n$ (which are pure submodules of $F_n$).

**Proof.** Since $M$ is flat then the sequence (3.6) is pure-exact and the projective modules $F_0, \ldots, F_n$ are obviously pure-projective. It follows from Theorem 3.3 that $K_n$ is an $\aleph_n$-directed union of $\aleph_n$-generated pure-projective $R$-submodules $U_\mu$ of $K_n$, with $\mu \in \Omega_n$, which are pure submodules of the projective module $F_n$. Hence each of the modules $U_\mu$ is flat and therefore any epimorphism $h_\mu : P_\mu \rightarrow U_\mu$ from a projective module $P_\mu$ to $U_\mu$ is a pure epimorphism. Since the module $U_\mu$ is pure-projective, then the epimorphism $h_\mu$ splits and, consequently, the module $U_\mu$ is projective for any $\mu \in \Omega_n$. This completes the proof. □

### 4. Pure-periodic modules.

**Definition 4.1.** Let $R$ be a ring. A right $R$-module $T$ is defined to be pure-periodic of period $p \geq 1$ if there exists a pure-exact sequence

$$0 \rightarrow T \rightarrow P_p \rightarrow P_{p-1} \rightarrow \ldots \rightarrow P_1 \rightarrow T \rightarrow 0$$

in $\text{Mod}(R)$ such that the modules $P_1, \ldots, P_p$ are pure-projective.

We start with two useful lemmata on pure-periodic modules.

**Lemma 4.2.** Assume that $T$ is a pure-periodic right $R$-module of period $p \geq 1$ and $\aleph$ is an infinite cardinal number. Then, for any $\aleph$-generated
submodule \( U \) of \( T \), there exists an \( \aleph \)-generated pure submodule \( U^\bullet \) of \( T \) such that \( U \subseteq U^\bullet \) and \( U^\bullet \) is pure-periodic of the same period \( p \).

**Proof.** Since \( T \) is pure-periodic of period \( p \geq 1 \), then \( T \) is a pure submodule of a pure-projective module \( P_0 \) and there exists a pure-exact sequence

\[
0 \to K_p \to P_0 \xrightarrow{d_p} P_{p-1} \xrightarrow{d_{p-1}} \ldots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \to 0
\]

where \( M = P_0/T, \) \( d_0 \) is the canonical projection, \( \text{Ker} \ d_0 = T \cong K_p = \text{Ker} \ d_p \) and the modules \( P_1, \ldots, P_p \) are pure-projective. Let \( f : T \to K_p \) be an \( R \)-module isomorphism and let \( N = f(U) \). Then Proposition 3.1 (b) applies to the above sequence with \( K_0 = T \) and \( N \cong L = U \). Consequently, there exists an \( \aleph \)-generated pure submodule \( U^\bullet \) of \( T \) satisfying the required conditions. \( \square \)

**Lemma 4.3.** Assume that \( \aleph \geq \aleph_1 \) is an infinite cardinal number and \( T \) an \( \aleph \)-generated pure submodule of a pure-projective right \( R \)-module.

(a) The module \( T \) is a continuous well-ordered union of pure submodules \( T_\xi \), with \( \xi < \gamma \), such that, for each \( \xi \), the module \( T_\xi \) is generated by a set of cardinality \( < \aleph \).

(b) If, in addition, \( T \) is pure-periodic of period \( p \geq 1 \) then, for each \( \xi < \gamma \), the pure submodule \( T_\xi \) of \( T \) can be chosen pure-periodic of the same period \( p \).

**Proof.** Let \( \aleph' < \aleph \) be an infinite cardinal. It follows from Lemma 2.5 that any \( \aleph' \)-generated submodule \( X \) of \( T \) can be embedded in an \( \aleph' \)-generated pure submodule \( X^\diamond \) of \( T \). If, in addition, \( T \) is pure-periodic of period \( p \geq 1 \) then, according to Lemma 4.2, \( X \) can be embedded in an \( \aleph' \)-generated pure-periodic pure submodule \( X^\bullet \) of \( T \) of the same period \( p \).

It is well-known that \( T \) can be represented as a continuous well-ordered union of submodules \( U_\xi \), with \( \xi < \gamma \), such that, for each \( \xi \), the module \( U_\xi \) is generated by a set of cardinality \( < \aleph \) (see [11, Lemma 1.4] and [19, Lemma 2.2]). Let us define a transfinite increasing chain of pure-submodules \( T_\xi \) of \( T \), with \( \xi < \gamma \), having \( < \aleph \) generators as follows (compare with [19, Theorem 1.5] and [16, Lemma 1.7]). We set \( T_0 = (0) \) and \( T_1 = U_1^\diamond \). If \( T_\xi \) is defined and \( \xi + 1 < \gamma \), we set \( T_{\xi+1} = (T_\xi + U_{\xi+1})^\diamond \). Finally, we set \( T_\tau = \bigcup_{\xi < \tau} T_\xi \) if \( \tau \) is a limit ordinal number. This proves Statement (a).

In order to prove (b), we assume that \( T \) is pure-periodic of period \( p \geq 1 \). It follows that \( T \) is a pure submodule of a pure-projective module \( P_0 \) and there exists a pure-exact sequence

\[
0 \to K_p \to P_0 \xrightarrow{d_p} P_{p-1} \xrightarrow{d_{p-1}} \ldots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \to 0
\]

where \( M = P_0/T, \) \( d_0 \) is the canonical projection, \( \text{Ker} \ d_0 = T \cong K_p = \text{Ker} \ d_p \) and the modules \( P_1, \ldots, P_p \) are pure-projective. Let \( f : T \to K_p \) be an \( R \)-module isomorphism and let \( N = f(U) \).
In the notation introduced in (a1) and (a2) of the proof of Proposition 3.1 with \( n = p \), we set \( K_p = \text{Ker} \cdot d_p \cong K_0 = T \), \( P_j = P(I_j) \) for \( j = 0, 1, \ldots, p \). We define a transfinite increasing chain of pure-periodic pure-submodules \( T_\xi \) of \( T \) generated by sets of cardinality \( \mathfrak{c} \), a transfinite chain of subsets \( I_{j, \xi} \) of \( I_j \) of cardinality \( |I_{j, \xi}| < \mathfrak{c} \) for \( \xi < \gamma \) and \( j = 0, 1, \ldots, p \), such that \( I_{j, \xi} \subseteq I_{j, \xi + 1} \) and there is a commutative diagram

\[
\begin{array}{c}
0 \rightarrow f(T_\xi) \rightarrow P(I_{p, \xi}) \xrightarrow{d_{p, \xi}} P(I_{p-1, \xi}) \xrightarrow{\cdots} P(I_1, \xi) \xrightarrow{d_{1, \xi}} T_\xi \rightarrow 0 \\
\end{array}
\]

with pure-exact rows, where the vertical arrows are natural embeddings induced by the inclusions \( I_{j, \xi} \subseteq I_{j, \xi + 1} \) for \( j = 1, \ldots, p \) and the module \( T_\xi \) is a pure submodule of \( P(I_{0, \xi}) \).

We set \( T_0 = (0) \) and \( I_{j, 0} = \emptyset \) for \( j = 0, 1, \ldots, p \). If \( T_\xi \), \( I_{j, \xi} \) and a corresponding pure-exact sequence are defined, we set \( T_{\xi + 1} = (T_\xi + U_{\xi + 1}) \) and we define the sets \( I_{p, \xi + 1}, I_{p-1, \xi + 1}, \ldots, I_{0, \xi + 1} \) by applying Proposition 3.1 (b) and its proof in such a way that the above diagram \((*)_\xi \) is commutative, the rows are pure-exact and the module \( T_\xi \) in \((*)_\xi \) is a pure submodule of \( P(I_{0, \xi}) \) generated by a set of cardinality \( < \mathfrak{c} \) and the module \( f(T_\xi) \) in \((*)_\xi \) is a pure submodule of \( P(I_{p, \xi}) \).

If \( \tau \) is a limit ordinal number we set \( T_\tau = \bigcup_{\xi<\tau} T_\xi \) and \( I_{j, \tau} = \bigcup_{\xi<\tau} I_{j, \xi} \). Then the direct limit of the sequences \((*)_\xi \) is a pure-exact sequence

\[
0 \rightarrow f(T_\tau) \rightarrow P(I_{p, \tau}) \xrightarrow{d_{p, \tau}} P(I_{p-1, \tau}) \xrightarrow{d_{p-1, \tau}} \cdots \xrightarrow{d_{2, \tau}} P(I_{1, \tau}) \xrightarrow{d_{1, \tau}} T_\tau \rightarrow 0
\]

and the required conditions are satisfied. This finishes the proof. \( \square \)

Now we are able to prove main results of this section. The following theorem implies Theorem 1.3:

**Theorem 4.4.** Let \( R \) be a ring. Every pure-periodic \( R \)-module is pure-projective.

**Proof.** Assume that \( T \) is a pure-periodic right \( R \)-module of period \( p \geq 1 \), and \( T \) is generated by a set of cardinality \( \leq \mathfrak{c} \). We prove the theorem by transfinite induction on \( \mathfrak{c} \).

First, we suppose that \( \mathfrak{c} = \aleph_n \), where \( n \geq 0 \). Since \( T \) is pure-periodic, it follows that \( T \) is a pure submodule of a pure-projective module \( P_0 \) and
there exists a pure-exact sequence

\[ 0 \to \text{Ker } d_p \to P_p \xrightarrow{d_p} P_{p-1} \to \ldots \to P_1 \xrightarrow{d_1} P_0 \to M \to 0 \]

such that the modules \(P_0, \ldots, P_p\) are pure-projective, \(M = P_0/T\) and \(T = \text{Im } d_1\). Without loss of generality we can suppose that \(p \geq n\). In that case \(T\) is also \(\aleph_p\)-generated, and therefore we can assume \(p = n\).

By Theorem 3.3, the \(p\)th pure-syzygy module \(\text{Ker } d_p \cong T\) of \(P_p\) is an \(\aleph_p\)-directed union of \(\aleph_p\)-generated pure-projective pure submodules. It follows that \(T\) is pure-projective, because it is \(\aleph_p\)-generated.

Next, we assume that \(T\) is an arbitrary \(\aleph\)-generated pure-periodic \(R\)-module, \(\aleph > \aleph_n\) for all integers \(n \geq 1\), and the theorem holds for all cardinals smaller than \(\aleph\). By Lemma 4.3, \(T\) is a continuous well-ordered union of pure-periodic pure submodules \(T_\xi\), \(\xi < \gamma\), such that for each \(\xi < \gamma\) the module \(T_\xi\) is generated by a set of cardinality \(< \aleph\). By the inductive hypothesis, each of the modules \(T_\xi\) is pure-projective and therefore \(\text{P.pd } T_\xi \leq 1\). It follows from Proposition 2.1 that \(\text{P.pd } T \leq 1\). Hence the submodule \(\text{Ker } d_2\) of \(P_2\) in the pure-exact sequence above is pure-projective and consequently the monomorphism \(T \cong \text{Ker } d_p \hookrightarrow P_p\) splits. This shows that \(T\) is pure-projective and finishes the proof.

\textbf{Corollary 4.5.} Let \(M\) be a right \(R\)-module and

\[ (\ast) \quad \ldots \to P_n \xrightarrow{d_n} P_{n-1} \xrightarrow{d_{n-1}} \ldots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \to 0 \]

a pure-exact sequence, where the modules \(P_0, \ldots, P_n, \ldots\) are pure-projective. Assume that there exist two integers \(n \geq 0\) and \(p \geq 1\) such that \(\text{Ker } d_{n+p} \cong \text{Ker } d_n\). Then \(\text{P.pd } M \leq n + 1\).

\textit{Proof.} By our assumption, the \(n\)th pure syzygy module \(\text{Ker } d_n\) is pure-periodic of period \(p \geq 1\). It then follows from Theorem 4.4 that the module \(\text{Ker } d_n\) is pure-projective and, consequently, \(\text{P.pd } M \leq n + 1\).

As a consequence of Theorem 4.4 we get the following form of [5, Theorem 2.5]:

\textbf{Corollary 4.6.} Let \(R\) be a ring, \(M\) a right flat \(R\)-module and \(n \geq 1\) an integer. If there exists an exact sequence

\[ (4.7) \quad 0 \to M \to F_n \to F_{n-1} \to \ldots \to F_1 \to F_0 \to M \to 0 \]

in \(\text{Mod}(R)\) such that the modules \(F_0, \ldots, F_n\) are projective, then \(M\) is projective.

\textit{Proof.} Since \(M\) is flat, the sequence (4.7) is pure-exact and the projective modules \(F_0, \ldots, F_n\) are obviously pure-projective. It follows that \(M\) is pure-periodic and, according to Theorem 4.4, the sequence (4.7) splits. Consequently, \(M\) is projective.
We finish this section with the following interesting question suggested by referee and related with the problems studied in [4] and [5].

**Problem 4.8.** Assume that $R$ is an associative ring with an identity element and $H$ is a subgroup of finite index in a group $G$. Let $M$ be an arbitrary right module over the group ring $RG$. Is the $RG$-module $M$ pure-projective, if $M$ is pure-projective, when viewed as an $RH$-module?

5. A structure of pure-projective resolutions in Grothendieck categories.

We show in this section that the main results of Sections 3 and 4 on pure-syzygies and pure-periodic modules generalize from the module category $\text{Mod}(R)$ to an arbitrary locally finitely presented Grothendieck category $\mathcal{A}$ (see [21]).

Throughout we denote by $\mathcal{A}$ a locally finitely presented Grothendieck category. We recall that an object $L$ of $\mathcal{A}$ is said to be finitely presented if the additive functor $\text{Hom}_\mathcal{A}(L, -) : \mathcal{A} \to Ab$ commutes with filtered direct limits (see [21], [26], [30]). A long exact sequence $\cdots \to X_{n-1} \to X_n \to X_{n+1} \to \cdots$ in $\mathcal{A}$ is said to be **pure** if the induced sequence $\cdots \to \text{Hom}_\mathcal{A}(L, X_{n-1}) \to \text{Hom}_\mathcal{A}(L, X_n) \to \text{Hom}_\mathcal{A}(L, X_{n+1}) \to \cdots$ of abelian groups is exact for any finitely presented object $L$. An epimorphism $f : Y \to Z$ in $\mathcal{A}$ is said to be **pure** if the exact sequence $0 \to \text{Ker}f \to Y \xrightarrow{f} Z \to 0$ is pure. A subobject $X$ of $Y$ in $\mathcal{A}$ is said to be **pure** if the exact sequence $0 \to X \to Y \to Y/X \to 0$ is pure. An object $P$ in $\mathcal{A}$ is said to be **pure-projective** if for any pure-epimorphism $f : Y \to Z$ in $\mathcal{A}$ the induced group homomorphism $\text{Hom}_\mathcal{A}(P, f) : \text{Hom}_\mathcal{A}(P, Y) \to \text{Hom}_\mathcal{A}(P, Z)$ is surjective.

It is well-known that:

- An object $P$ in locally finitely presented Grothendieck category $\mathcal{A}$ is pure-projective if and only if $P$ is a direct summand of a coproduct of finitely presented objects.
- Every object $M$ in $\mathcal{A}$ admits a pure-projective pure resolution in $\mathcal{A}$, that is, there is a pure-exact sequence

\[
\cdots \to P_n \xrightarrow{d_n} P_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_2} P_1 \xrightarrow{d_1} P_0 \xrightarrow{d_0} M \to 0,
\]

where $P_0, \ldots, P_n, \ldots$ are pure-projective objects (see [26], [30]).

Then, for each $n \geq 0$, the functor $\text{Pext}_\mathcal{A}^n : \mathcal{A}^{op} \times \mathcal{A} \to Ab$ is naturally defined and, for any object $M$ in $\mathcal{A}$, the pure-projective dimension $\text{P.pd} M$ of $M$ is defined in a natural way (see [30] and [26]).

The main results of Sections 3 and 4 extend to the context of locally finitely presented Grothendieck categories as follows:

**Theorem 5.2.** Let $M$ be an arbitrary object of a locally finitely presented Grothendieck category $\mathcal{A}$. If (5.1) is a pure-exact sequence in $\mathcal{A}$ such that
the objects $P_0,\ldots, P_n,\ldots$ are pure-projective, then for each $n \geq 0$ the $n$-th pure-syzygy object $\text{Ker} \ d_n$ of $M$ is an $\aleph_n$-directed union of $\aleph_n$-generated pure-projective objects, which are pure subobjects of $P_n$.

**Theorem 5.3.** Let $M$ be an object of a locally finitely presented Grothendieck category $\mathcal{A}$. If there exists a pure-exact sequence

$$0 \to M \to P_n \to P_{n-1} \to \ldots \to P_1 \to P_0 \to M \to 0$$

in $\mathcal{A}$ such that the objects $P_0,\ldots, P_n$ are pure-projective, then $M$ is pure-projective.

In other words, every pure-periodic object $M$ of $\mathcal{A}$ is pure-projective.

**Outline of the proof.**

First we note that Proposition 2.1 (a pure version of a theorem of Auslander [2]) extends to an arbitrary Grothendieck category $\mathcal{A}$ (see [25] and [26, Proposition 2.6]). Further, by [26, Theorem 2.12], if $M$ is an $\aleph_n$-presented object of a locally finitely presented Grothendieck category $\mathcal{A}$ and $n \geq 0$, then $\text{P.pd} \ M \leq n+1$. Finally, we note that the proof of Lemma 2.5 uses only categorical arguments and therefore extends to our situation. It then follows that also our Lemma 2.6 remains valid with $\text{Mod}(R)$ and $\mathcal{A}$ interchanged. Since the Kaplansky theorem [13] also remains valid for objects of a Grothendieck category $\mathcal{A}$ (see [21]) then the proof of Proposition 3.1 works with $\text{Mod}(R)$ and $\mathcal{A}$ interchanged. Thus, applying the arguments in the proof of Theorems 3.3 and 4.4, we easily get Theorems 5.2 and 5.3. The details are left to the reader. \quad \Box

**Acknowledgment.** The author is indebted to the referee for useful remarks, comments and improvements.

**References**


[28] ______, A class of potential counter-examples to the pure semisimplicity conjecture, 
  Proc. Conf. ‘Model Theory and Modules’, Essen, 1994; Algebra, Logic and Applica-
  tions, Vol. 9, Gordon & Breach Science Publishers, Amsterdam, 1997, 345-373, 

[29] B. Stenström, Pure submodules, Arkiv för Mat., 7 (1967), 159-171, MR 36 #6473, 
  Zbl 0153.34302.

  Zbl 0157.04204.


Received February 12, 2001 and revised February 26, 2002. This author was partially 
  supported by Polish KBN Grant 5 P0 3A 015 21.

Faculty of Mathematics and Informatics 
Nicholas Copernicus University 
ul. Chopina 12/18 
87-100 Toruń, Poland 
E-mail address: simson@mat.uni.torun.pl
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.
Fundamental solutions of invariant differential operators on a semisimple Lie group II
GUILLERMO AMES 1

Free stochastic measures via noncrossing partitions II
MICHAEL ANSHELEVICH 13

Groups acting on Cantor sets and the end structure of graphs
BRIAN H. BOWDITCH 31

On the Diophantine equation \(\frac{x^n - 1}{x - 1} = \frac{y^n - 1}{y - 1}\)
YANN BUGEAUD AND T.N. SHOREY 61

Toda lattice and toric varieties for real split semisimple Lie algebras
LUIS G. CASIAN AND YUI KODAMA 77

On diophantine monoids and their class groups
SCOTT T. CHAPMAN, ULRICH KRAUSE AND EBERHARD OELJEKLAUS 125

The group of isometries of a Finsler space
SHAOQIANG DENG AND ZIXIN HOU 149

On spaces of matrices containing a nonzero matrix of bounded rank
DMITRY FALIKMAN, SHMUEL FRIEDLAND AND RAPHAEL LOEWY 157

On the commutator formula of a split BN-pair
GWENAELLE GENET 177

Phragmén–Lindelof theorem for minimal surface equations in higher dimensions
CHUN–CHUNG HSEIH, JENN–FANG HWANG AND FEI–TSEN LIANG 183

Remark on the rate of decay of solutions to linearized compressible Navier–Stokes equations
TAKAYUKI KOBAYASHI AND YOSHIHITO SHIBATA 199

Pure-periodic modules and a structure of pure-projective resolutions
DANIEL SIMSON 235